

**Common Data Model (CDM) Specification, Version 5.1**

|  |
| --- |
| **1. Table of Contents** |

# [Table of Contents 1](#_bookmark0)

# [Overview of the PCORnet Common Data Model (CDM) 4](#_bookmark1)

* 1. [**Introduction 4**](#_bookmark2)
  2. [**History of Releases and Modifications 5**](#_bookmark3)
  3. [**Overview Diagram 9**](#_bookmark4)
  4. [**Implementation Expectations 10**](#_bookmark5)

# [Design of the CDM 10](#_bookmark6)

* 1. [**Special Topics for CDM Modeling 10**](#_bookmark7)
  2. [**Development Notes 13**](#_bookmark8)
  3. [**Comments on Protected Health Information (PHI) 15**](#_bookmark9)
  4. [**The Continuum of Medication-related Data Domains 16**](#_bookmark10)

# [Implementation Guidance 17](#_bookmark11)

# [Core CDM Table Specifications 21](#_bookmark12)

* 1. [**Table: DEMOGRAPHIC 21**](#_bookmark13)
  2. [**Table: ENROLLMENT 28**](#_bookmark14)
  3. [**Table: ENCOUNTER 32**](#_bookmark15)
  4. [**Table: DIAGNOSIS 46**](#_bookmark16)
  5. [**Table: PROCEDURES 54**](#_bookmark17)
  6. [**Table: VITAL 60**](#_bookmark18)
  7. [**Table: DISPENSING 69**](#_bookmark19)
  8. [**Table: LAB\_RESULT\_CM 75**](#_bookmark20)

[Implementation Guidance Reference Table 1: Laboratory Results & LOINC Codes 88](#_bookmark21)

[Implementation Guidance Reference Table 2: Laboratory Results and CPT Codes 89](#_bookmark22)

[Implementation Guidance Reference Table 3: Laboratory Standard Abbreviations 89](#_bookmark23)

* 1. [**Table: CONDITION 90**](#_bookmark24)
  2. [**Table: PRO\_CM 95**](#_bookmark25)

[CDM Reference Table: PRO Common Measures 108](#_bookmark26)

* 1. [**Table: PRESCRIBING 109**](#_bookmark27)

[Implementation Guidance Reference Table 4: Ordering of RxNorm Term Types 117](#_bookmark28)

[Implementation Guidance Reference Table 4a: RxNorm Term Types with examples 118](#_bookmark29)

* 1. [**Table: PCORNET\_TRIAL 120**](#_bookmark30)
  2. [**Table: DEATH 125**](#_bookmark31)
  3. [**Table: DEATH\_CAUSE 127**](#_bookmark32)
  4. [**Table: MED\_ADMIN 130**](#_bookmark33)
  5. [**Table: PROVIDER 136**](#_bookmark34)
  6. [**Table: OBS\_CLIN 139**](#_bookmark35)
  7. [**Table: OBS\_GEN 145**](#_bookmark36)
  8. [**Table: HASH\_TOKEN 153**](#_bookmark37)
  9. [**Table: LDS\_ADDRESS\_HISTORY 156**](#_bookmark38)
  10. [**Table: IMMUNIZATION 162**](#_bookmark39)
  11. [**Table: HARVEST 169**](#_bookmark40)

# [Supplemental Table Specifications 189](#_bookmark41)

* 1. [**Supplemental Table: PRIVATE\_DEMOGRAPHIC 189**](#_bookmark42)
  2. [**Supplemental Table: PRIVATE\_ADDRESS\_HISTORY 199**](#_bookmark43)
  3. [**Supplemental Table: PRIVATE\_ADDRESS\_GEOCODE 206**](#_bookmark44)

**Important Links and References**

The PCORnet CDM documentation can be accessed online at: <https://pcornet.org/data-driven-common-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: <https://pcornet.org/data-driven-common-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/~~](http://www.pcornet.org/pcornet-common-data-model/)

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

The CDM specifications for version 3.1 and above incorporate the Implementation Guidance that has been developed for PCORnet. The Implementation Guidance is intended to help reduce the variability in how network partners populate their CDM datamarts. It provides recommendations and preferred approaches when there are multiple interpretations of the CDM specification or if there is unexpected complexity in a partner’s source data. The Implementation Guidance is intended to be a living document, and as such, will be updated more frequently than the CDM specification itself.

To accommodate the addition of this material, the CDM page size has been increased from US Letter to US Legal (8.50”x14”). For best results when printing, use Legal-size paper.

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

|  |
| --- |
| **2.1. License and Use** |

This work is licensed under the Creative Commons Attribution 4.0 International License, and is provided on an "as is" basis without warranties or conditions of any kind. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

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 the governance established by PCORnet’s stakeholders.

The PCORnet CDM was originally based on the Mini-Sentinel Common Data Model v4.0 (MSCDM v4.0; [www.mini-sentinel.org](http://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.

This material contains content from LOINC® (http://loinc.org). The LOINC Table, LOINC Table Core, LOINC Panels and Forms File, LOINC Answer File, LOINC Part File, LOINC Group File, LOINC Document Ontology File, LOINC Hierarchies, LOINC Linguistic Variants File, LOINC/RSNA Radiology Playbook, and LOINC/IEEE Medical Device Code Mapping Table are copyright © 1995-2017, Regenstrief Institute, Inc. and the Logical Observation Identifiers Names and Codes (LOINC) Committee and is available at no cost under the license at [http://loinc.org/license.](http://loinc.org/license) (Updated in v4.0.)

|  |
| --- |
| **2.2. Introduction** |

***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 to be a “…functional distributed research network that facilitates multi-site patient–centered research across the Clinical Research Networks (CRNs) 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>

|  |
| --- |
| **2.3. History of Releases and Modifications** |

Note on version conventions: Major releases are denoted in whole number increments (e.g., v1.0, v2.0, v3.0). Minor releases are denoted with decimal increments (e.g., v1.1, v1.2) and will be used for bug fixes and minor adjustments. Updates to the HARVEST table are not listed in the release notes as this metadata table is expected to change to reflect every CDM expansion.

|  |  |  |
| --- | --- | --- |
| **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. 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. |
| v3.1 | 2016-11-15 | Please note: New and modified fields are 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 |

|  |  |  |
| --- | --- | --- |
| **Reference Table: History of Releases** | | |
| *Version* | *Date of Release* | *Description of Release* |
| V4.0 | 2018-01-02 | The v4.0 release includes:   * Four new tables (PROVIDER, OBS\_CLIN, OBS\_GEN, MED\_ADMIN) * Thirty-two new fields in existing tables (PAT\_PREF\_LANGUAGE\_SPOKEN, PAYER\_TYPE\_PRIMARY, PAYER\_TYPE\_SECONDARY, FACILITY\_TYPE, DX\_POA, PPX, DISPENSE\_DOSE\_DISP, DISPENSE\_DOSE\_DISP\_UNIT, DISPENSE\_ROUTE, RESULT\_SNOMED, PRO\_TYPE, PRO\_ITEM\_LOINC, PRO\_ITEM\_NAME, PRO\_RESPONSE\_TEXT, PRO\_ITEM\_VERSION, PRO\_MEASURE\_NAME, PRO\_MEASURE\_SEQ, PRO\_MEASURE\_SCORE, PRO\_MEASURE\_THETA, PRO\_MEASURE\_SCALED\_TSCORE, PRO\_MEASURE\_STANDARD\_ERROR, PRO\_MEASURE\_COUNT\_SCORED, PRO\_MEASURE\_VERSION, PRO\_ITEM\_FULLNAME, PRO\_ITEM\_TEXT, PRO\_MEASURE\_FULLNAME, RX\_DOSE\_ORDERED, RX\_DOSE\_ORDERED\_UNIT, RX\_PRN\_FLAG, RX\_ROUTE, RX\_SOURCE, RX\_DISPENSE\_AS\_WRITTEN) * Renamed PRO\_RESPONSE field to PRO\_RESPONSE\_NUM. * PRO\_ITEM field deprecated. * Renamed PRO\_LOINC field to PRO\_ITEM\_LOINC. * ADMITTING\_SOURCE value set expanded to include intra-hospital admitting source. * DX\_ORIGIN and PX\_SOURCE value sets expanded to include diagnoses/procedures derived or imputed through analytical procedures (e.g., natural language processing). * SPECIMEN\_SOURCE value set expanded to include all values from the LOINC SYSTEM part. * RESULT\_QUAL value set expanded. * RESULT\_UNIT value set expanded to include common UCUM units. * RX\_QUANTITY\_UNIT value set modified to align with RxNorm terminology. * RX\_FREQUENCY value set expanded to include “every evening” and “once” concepts. * Renamed RX\_QUANTITY\_UNIT field to RX\_DOSE\_FORM for clarity and consistency. * LAB\_NAME field deprecated. * Date management and refresh date fields for new tables added to HARVEST table. * Required/not null ENCOUNTERID constraint removed from DIAGNOSIS and PROCEDURES tables. * Required/not null PRO\_RESPONSE constraint removed from PRO\_CM table. * Domain descriptions updated for DISPENSING, LAB\_RESULT\_CM and PRO\_CM tables. * Deprecation of Implementation Guidance Reference Table 3 and CDM Reference Table (PRO Common Measures). * Modifications to the foreign key descriptions for several tables. * Concept of PRIVATE and Supplemental tables introduced. * Various updates to the Implementation Guidance |

|  |  |  |
| --- | --- | --- |
| **Reference Table: History of Releases** | | |
| *Version* | *Date of Release* | *Description of Release* |
| V4.1 | 2018-05-15 | The v4.1 release includes:   * Field length updates for PAYER\_TYPE\_PRIMARY and PAYER\_TYPE\_SECONDARY. * Foreign key updates for DISPENSING.PATID, PCORNET\_TRIAL.PATID and DEATH.PATID. * Typo correction for MEDADMIN.MEDADMIN\_PROVIDERID. * Renamed MEDADMIN\_END\_DATE\_MGMT to MEDADMIN\_STOP\_DATE\_MGMT. * Update to description of PRO\_MEASURE\_FULLNAME. * Various updates to the Implementation Guidance. * Updates to entries of the FACILITY\_TYPE, ROUTE and PAYER\_TYPE value sets included in the Value Set Appendix. * Removal of the DISPENSE\_FORM value set from the Value Set Appendix. |
| V5.0 | 2019-07-16 | The v5.0 release includes:   * Three new tables: HASH\_TOKEN, LDS\_ADDRESS\_HISTORY, IMMUNIZATION. * Modifications to field length and definition of ENCOUNTER.FACILITY\_LOCATION. * Modifications to value sets for six fields (CONDITION\_TYPE, CONDITION\_SOURCE, OBSGEN\_TYPE, VITAL\_SOURCE, DEATH\_SOURCE, DEATH\_CAUSE\_SOURCE) * Seven new fields in existing tables (DX\_DATE, DISPENSE\_SOURCE, LAB\_RESULT\_SOURCE, LAB\_LOINC\_SOURCE, PRO\_SOURCE, OBSCLIN\_SOURCE, OBSGEN\_SOURCE) * Three new PRIVATE tables: PRIVATE\_DEMOGRAPHIC, PRIVATE\_ADDRESS\_HISTORY, PRIVATE\_ADDRESS\_GEOCODE * Various updates to the Implementation Guidance |

|  |  |  |
| --- | --- | --- |
| **Reference Table: History of Releases** | | |
| *Version* | *Date of Release* | *Description of Release* |
| V5.1 | 2019-09-10 | Please note: v5.0 updates are highlighted in green and updates introduced in v5.1 are highlighted in yellow to assist with visually scanning the document (in addition to the descriptive comments).  The v5.1 release includes:   * Three new fields in HASH\_TOKEN: TOKEN\_03, TOKEN\_04, TOKEN\_16. * Deprecation of 5 fields in HASH\_TOKEN: TOKEN\_12, TOKEN\_17, TOKEN\_21, TOKEN\_22, TOKEN\_23. * Updates to the field descriptions in HASH\_TOKEN: TOKEN\_01, TOKEN\_02, TOKEN\_05. * Corrections to the constraints of PRIVATE\_DEMOGRAPHIC. * Corrections to the field length and definition of PRIVATE\_GEOCODE.GEOCODE\_COUNTY. * Updates to the Implementation Guidance in PRESCRIBING, HASH\_TOKEN, PRIVATE\_DEMOGRAPHIC and PRIVATE\_ADDRESS\_HISTORY. * Correction to PRO\_SOURCE value set to align with the parseable file for this field. * HARVEST table: REFRESH\_LDS\_ADDRESS\_HISTORY\_DATE is shortened to REFRESH\_LDS\_ADDRESS\_HX\_DATE. * Clarified implementation guidance for future dates and PATID consistency. * Clarified field implementation guidance for ADDRESS\_PERIOD\_END in the ADDRESS\_HISTORY table. |

|  |
| --- |
| **2.1. Overview Diagram** |

**PCORnet Common Data Model v5.1**

**New to v5.0**

**IMMUNIZATIONID**

**PATID VX\_CODE VX\_CODE\_TYPE VX\_STATUS** *ETC*

**IMMUNIZATION**

**DISPENSINGID PATID DISPENSE\_DATE NDC** DISPENSE\_SOURCE *ETC...*

DISPENSE\_DOSE\_DISP\_UNIT

DISPENSE\_ROUTE

**DISPENSING**

**PATID**

**DEATH\_CAUSE DEATH\_CAUSE\_CODE DEATH\_CAUSE\_TYPE DEATH\_CAUSE\_SOURCE** *ETC...*

**DEATH\_CAUSE**

**DIAGNOSISID**

**PATID DX DX\_TYPE**

**DX\_SOURCE** DX\_DATE *ETC...*

DX\_POA

**DIAGNOSIS**

**PATID DEATH\_SOURCE** *ETC...*

**DEATH**

**PRESCRIBINGID**

**PATID**

*ETC...*

RX\_DOSE\_ORDERED RX\_DOSE\_ORDERED\_UNIT RX\_ROUTE

RX\_SOURCE RX\_DISPENSE\_AS\_WRITTEN RX\_PRN\_FLAG

**PRESCRIBING**

**PATID ENR\_START\_DATE ENR\_BASIS**

*ETC...*

**ENROLLMENT**

**PRO\_CM\_ID PATID** ENCOUNTERID **PRO\_DATE** PRO\_TIME PRO\_TYPE PRO\_ITEM\_NAME PRO\_ITEM\_LOINC

PRO\_RESPONSE\_TEXT PRO\_RESPONSE\_NUM PRO\_METHOD PRO\_MODE

PRO\_CAT PRO\_SOURCE *ETC*

**PRO\_CM**

**ENCOUNTERID PATID ADMIT\_DATE ENC\_TYPE** *ETC...*

PAYER\_TYPE\_PRIMARY PAYER\_TYPE\_SECONDARY FACILITY\_TYPE

**ENCOUNTER**

**PATID TRIALID**

**PARTICIPANTID**

*ETC...*

**PCORNET\_TRIAL**

**VITALID PATID**

**MEASURE\_DATE VITAL\_SOURCE** *ETC...*

**VITAL**

**LAB\_RESULT\_CM\_ID PATID RESULT\_DATE**

LAB\_RESULT\_SOURCE LAB\_LOINC\_SOURCE *ETC* RESULT\_SNOMED

**LAB\_RESULT\_CM**

**PATID**

*ETC...*

PAT\_PREF\_LANGUAGE\_SPOKEN

**DEMOGRAPHIC**

**ADDRESSID PATID ADDRESS\_USE ADDRESS\_TYPE**

**ADDRESS\_PREFERRED**

*ETC*

**LDS\_ADDRESS\_HISTORY**

**OBSGENID PATID** ENCOUNTERID

OBSGEN\_PROVIDERID **OBSGEN\_DATE** OBSGEN\_TIME

*ETC* OBSGEN\_SOURCE RAW\_OBSGEN\_UNIT

**OBS\_GEN**

**OBSCLINID PATID** ENCOUNTERID

OBSCLIN\_PROVIDERID **OBSCLIN\_DATE** OBSCLIN\_TIME OBSCLIN\_TYPE OBSCLIN\_CODE

*ETC* OBSCLIN\_SOURCE *ETC* RAW\_OBSCLIN\_UNIT

**OBS\_CLIN**

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

**CONDITIONID PATID CONDITION CONDITION\_TYPE**

**CONDITION\_SOURCE**

*ETC...*

**CONDITION**

**NETWORKID DATAMARTID** *ETC...*

**HARVEST**

**PATID** TOKEN\_01 *ETC* TOKEN\_16

**HASH\_TOKEN**

**MEDADMINID**

**PATID MEDADMIN\_START\_DATE** ENCOUNTERID MEDADMIN\_START\_TIME MEDADMIN\_STOP\_DATE MEDADMIN\_STOP\_TIME PRESCRIBINGID

*ETC..*.

MEDADMIN\_SOURCE

**MED\_ADMIN**

**PROVIDERID**

PROVIDER\_SEX PROVIDER\_SPECIALTY\_PRIMARY PROVIDER\_NPI PROVIDER\_NPI\_FLAG

**PROVIDER**

**PROCEDURESID**

**PATID PX PX\_TYPE** *ETC...*

PPX

**PROCEDURES**

|  |
| --- |
| **2.2. Implementation Expectations** |

Partners should populate all core CDM tables if data are available in their source system(s). **All core CDM 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. The fields that are required to be populated for all records in a given table are listed in the “constraints” section of each table description. Any table and/or field in the CDM may be required for a partner’s participation in a given study or other PCORnet activity. In assessing foundational data quality, the PCORnet Data Curation query packages may query any CDM table or field.

|  |
| --- |
| **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 stack1. 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

SAS Number fields have a *byte length* of 8 [SAS Numeric(8)]. This corresponds to an 8-byte floating-point number of approximately 16 significant digits. When deciding on the precision/scale for their RDBMS Number fields, partners should ensure that they do not store numbers in a way that would overflow the SAS numeric data type, which would result a loss of data when generating a SAS dataset from the RDBMS. 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.

When deciding on the parameters to choose for their RDBMS number fields, network partners should choose a combination that does not result in additional, artificial decimal precision. For example:

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

Instead of specifying a precision (total number of digits) and scale (digits to the right of a decimal point) for RDBMS Number data types, the CDM spec has been revised to just read “RDBMS Number.” Partners should specify the parameters that are most appropriate for their RDBMS that that does not cause a loss of data when generating SAS datasets from the RDBMS or nor result in additional, artificial decimal precision.

# 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:

**A data field is not present** in the source system. (populate with null)

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)

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) 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.

1 <https://msdn.microsoft.com/en-us/data/ef.aspx>

# 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 are case-sensitive. Schema implementations for Oracle, Microsoft SQL Server, and PostgreSQL should be in uppercase (table name, column names, etc.). Value set codes should reflect the case formatting in the CDM specification and/or Value Set Appendix.

# 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,

“CRNs 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.

# Expanded value sets

Version 4.0 of the PCORnet CDM introduced the concept of expanded value sets for fields with dozens or hundreds of allowable options (e.g., LAB\_RESULT\_CM.RESULT\_UNIT, PRESCRIBING.RX\_DOSE\_FORM). To reduce the size of the CDM specification document, these value sets are provided in a supplementary Value Set Appendix (ValueSet\_ReferenceFile), which is co-located with the CDM specification on the PCORnet web site (<http://pcornet.org/pcornet-common-data-model/>). It is expected that these value sets will only be updated as part of a CDM version update, but there may be extenuating circumstances where an out-of-sequence update is required. For each value set, we list the raw/expected source value and the corresponding analytics-friendly string to be used when populating the CDM.

# Supplemental tables

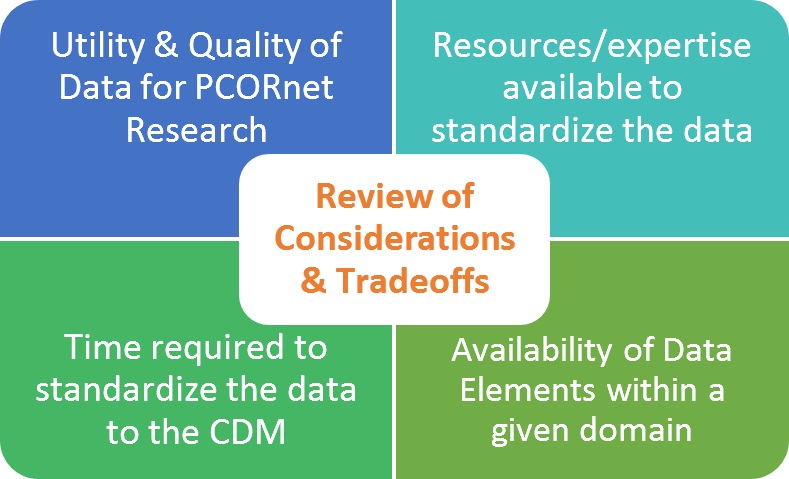
Version 4.0 of the PCORnet CDM also formalized the concept of “supplemental” CDM tables. These are tables outside of the core CDM that used to support study-specific activities that involve the participation of many network partners. These table definitions may be managed in separate document(s) outside of the core CDM specification and may be refined out of cycle with the rest of the CDM. Over time, some of these tables may be promoted to the core CDM, at which point they will be governed by the versioning processes of the core. Network partners are not expected to populate these tables unless they are participating in a study that utilizes them.

|  |
| --- |
| **3.2.** **Development Notes** |

PCORnet CDM Guiding Principle #2 states,

“**It is expected that not all CRNs will have data needed to populate all parts of the PCORnet CDM.** It is the responsibility of the CRNs 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 (e.g., 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, “**CRNs will control how their data are used as allowed by internal governance policies**. Data resources developed for PCORnet will stay within the CRNs 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):

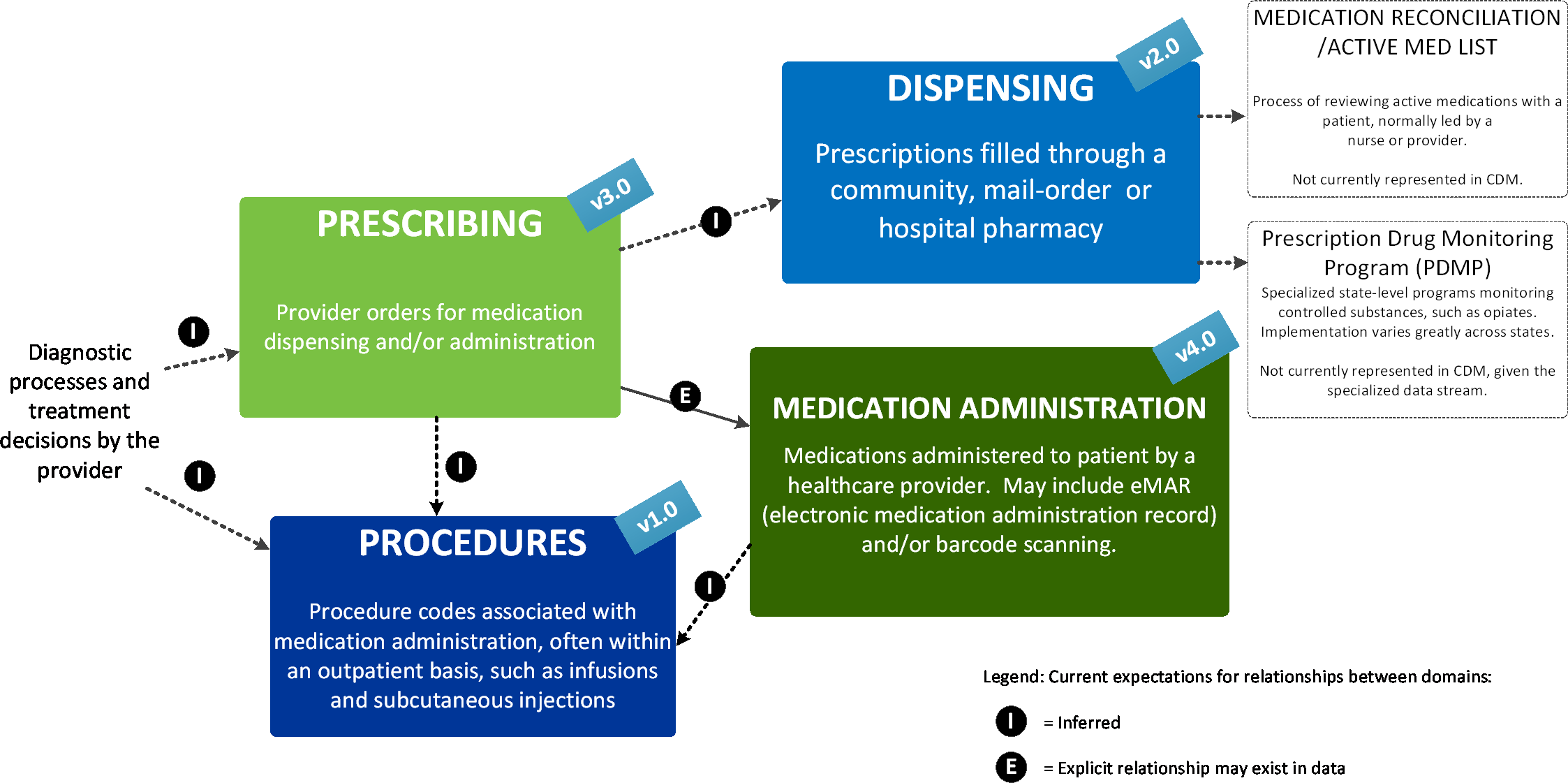
* PATID crosswalk
* PROVIDER crosswalk

# PRIVATE tables

Version 4.0 of the PCORnet CDM introduced the concept of “PRIVATE” tables, which are intended to provide standardized representations for the commonly-used PHI elements that are necessary for certain analytic activities (e.g., patient linkage, geocoding). These tables will not be directly queried by the DRN OCand can remain physically and logically separate from the rest of the CDM. These tables will initially be developed as Supplemental tables through the process described in Section 3.1.

|  |
| --- |
| **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. Implementation Guidance**

Implementation Guidance is intended to help reduce the variability in how partners populate their CDM datamarts. It provides recommendations and describes preferred approaches when there are multiple interpretations of the CDM specifications or if there is unexpected complexity in a partner’s source data. The Implementation Guidance is divided into three sections, plus a set of reference tables:

1. *General guidance*: The guidance applies to more than 1 CDM table/domain. These items are included in this section of the CDM Specification.
2. *Table-level guidance*: The guidance applies to the table/domain in general or applies to more than one field in the table. Table-level guidance is provided in each table’s description before the table specification.
3. *Field-level guidance*: The guidance applies to implementation decisions that are specific to a given field in a table. Field-level guidance is provided as an extra column in the CDM table specification.
4. *Reference tables*: When applicable, reference tables have been created to provide additional guidance to network partners.

# Guidance updated as part of a CDM revision is highlighted in green. Guidance updated between CDM versions is highlighted in yellow.

|  |  |
| --- | --- |
| **GENERAL Implementation Guidance (spans more than 1 table)** | |
| *Topic* | *Guidance* |
| 1 – Population of RAW fields | If a given record/observation can be directly mapped into the PCORnet CDM, there is no need to populate the RAW values. RAW values may be used when partners need to employ a mapping in order to populate a given table (e.g., converting local diagnosis codes to ICD-9 or ICD-10, mapping internal procedure codes to one of the specified PCORnet procedure terminologies or converting several dozen race values to the values specified in the PCORnet CDM). Populating the RAW fields is optional, but if a partner chooses to do so, they should use the following strategy:   * If there is a 1:1 relationship between the source value and the CDM, the RAW value can be populated on the same record.   In cases where there is a 1:many relationship for multiple records that are part of encounter or have the same timestamp (e.g., 3 local diagnosis codes recorded in the same encounter that map to a single ICD9 code), all of the local values should be concatenated into a single RAW observation, with the values separated by a pipe delimiter (“|”). This prevents the creation of duplicate records in the PCORnet CDM. |
| 2 – Date Obfuscation and/or Truncation | ~~The preferred approach for PCORnet is that partners retain the actual dates reported within their source systems. If it is necessary to use~~ ~~date obfuscation (also known as date shifting), partners should ensure that the obfuscation is internally consistent across all dates for a given~~ ~~patient (i.e., all dates for that patient are shifted by the same amount). This would include dates in the PCORNET\_TRIAL table, if~~ ~~populated. Failure to adhere to this guidance will make it all but impossible to use that partner’s data in any analytic query. Do not shift~~ ~~dates into the future. Only apply a backwards shift. If partners are also truncating dates (i.e., defining a “valid” time window for their~~ ~~dataset and removing patients/data that correspond to encounters whose dates have been shifted out of that window) partners should~~ ~~document their strategy in their ETL ADD. Truncation is discouraged, and partners will need a way to flag those patients who have~~ ~~incomplete data as a result of the truncation process in order to retain the ability to use the datamart analytically.~~ (Guidance deprecated as of July 2019 to reflect terms of the PCRF Clinical Research Network Statement of Work.)  Do not include shifted or truncated dates within the CDM, with the exception of birth date for patients >89, if required by local regulations. Aggregate or de-identified queries that return age for patients >89 will appropriately bucket the results before they are returned. |

|  |  |
| --- | --- |
| **GENERAL Implementation Guidance (spans more than 1 table)** | |
| *Topic* | *Guidance* |
| 3 – Patients with age > 89 | PCORnet queries issued by the DRN OC will bucket all patients age 90+ into one age group to limit the risk of re-identification. Some partners’ local regulatory restrictions may require additional protection of date-related data on patients age > 89. For example, some partners might be asked to **remove** the birth date while others might require that it be **shifted** to mask the patient’s true date of birth. If the birth date must be **shifted**, these regulatory restrictions would subsequently prevent the same shift from being applied to the rest of the patient’s dates of service, resulting in inconsistencies in the data (see General Guidance #2). In this case, partners should consider shifting the value in the DEMOGRAPHIC.BIRTH\_DATE field for these patients to a dummy date of January 1, 1900. If this approach is taken, partners should structure their ETL to shift birth dates from patients currently over 89, as well as those that will “age out” and turn 89 before the next  expected datamart refresh. |
| 4 – Mapping source data to standard terminologies when multiple options exist | Some partners with Epic EHRs are able to access diagnoses/problem list entries coded via vocabulary mapping software/middleware (e.g., IMO). Such software/middleware provides mappings to multiple terminologies, so partners have several options from which to choose when populating their datamart.  Partners should utilize the vocabulary or terminology that most closely reflects the standard typically used to encode data in that domain at the time the observation was recorded. For instance, diagnosis codes would use ICD-9 before October 2015, and ICD-10 afterwards; procedure codes would be encoded in CPT/HCPCS or in ICD-9 before October 2015 and in CPT/HCPCS or ICD-10 afterwards; problem list values would be encoded in SNOMED. This will allow partners to avoid the potential duplication of records (e.g., if combining facility billing diagnoses coded in ICD with professional diagnoses coded in SNOMED or mapping the same EHR record to ICD-9 and ICD-10) and more readily allow the application of existing validated algorithms. In these situations, the original codes can be retained within the RAW fields to allow for additional studies that seek to execute exploratory analyses using alternative mappings. This guidance may be revised in the future to incorporate the published findings of subsequent validation studies that compare the downstream analytical results that arise from various mapping strategies. |
| 5 – Approach when there are known errors or missingness in TYPE fields (e.g., PROCEDURES.PX\_TYPE, LAB\_RESULT\_CM.LAB\_RES ULT\_PX\_TYPE, ENCOUNTER.DRG\_TYPE, or DIAGNOSIS.DX\_TYPE) | The PCORnet CDM Guiding Principles ask that data from source systems be populated in the CDM “as is.” However, some fields in the CDM that are used to identify attributes about the data, such as the TYPE fields, are critical to the operation of the PCORnet analytic queries. If the TYPE field is incorrect, the query will end up using a mismatched code set and return invalid/empty results. If the errors in the source for the TYPE fields are systemic (e.g., an interface reports all procedure billing codes as ICD-10, when they are really a mix of ICD-10 and CPT), and the partner is able to isolate and rectify the issue, then the preference is that these data be cleaned before populating the CDM. The same guidance applies if TYPE fields are not populated or are not transmitted from the source system. The heuristic or algorithm used by the partner should be documented in their ETL ADD. *If there are known systemic issues in other fields or domains, these data should not be cleaned, but partners should make every attempt to resolve or rectify the issue within their source system(s)*. Any known  systemic issues should also be documented in the ETL ADD. |
| 6 – Corrected/updated source values | If a measure/laboratory result has been corrected in the source system, and both the original and corrected values are present in the source system, partners should only include the corrected (e.g., most recent) value in their CDM datamart. If a measure/laboratory result is included in their CDM datamart that has later been found to be canceled, partners should remove that value from their CDM datamart in  subsequent refreshes. |

|  |  |
| --- | --- |
| **GENERAL Implementation Guidance (spans more than 1 table)** | |
| *Topic* | *Guidance* |
| 7 – Precision/scale for RDBMS  NUMBER fields | This guidance has been deprecated as of CDM v3.1. |
| 8 – Study-specific datamart restrictions | ~~Some networks that use a centralized model may have datamarts that contain data from sites that are not participating in a given study. In~~ ~~these cases, datamart teams are responsible for ensuring that only the data from participating sites are queried. Possible options include~~ ~~creating site flags or views/subsets of the approved SAS datamart tables and then running the study queries against that subset of the~~  ~~data. Partners have latitude to choose the approach that is most appropriate for their network.~~ (Guidance deprecated as of July 2019 to reflect terms of the PCRF Clinical Research Network Statement of Work.)  The DRN OC will work with networks that operate using a centralized model to ensure that site-specific results are available for all queries in the same fashion as distributed data marts. |
| 9 – PATIDs & other IDs | ~~To the extent possible,~~ Partners should maintain PATIDs and all other ID fields (ENCOUNTERID, DIAGNOSISID, PROVIDERID, etc.)  across refreshes. In addition, the lengths of these fields should be harmonized across all tables to facilitate cross-table querying. |
| 10 – Units of Measure | Starting with PCORnet CDM v4.0, the value set for the units of measure fields (e.g., RESULT\_UNIT in LAB\_RESULT\_CM, OBSGEN and CLINOBS, and \*\_STRENGTH\_\*\_UNIT in PRESCRIBING, DISPENSING and MED\_ADMIN) has been expanded to include values from the Unified Code for Units of Measure (UCUM). The value set in the Value Set Appendix reflects the list of values curated by LOINC that are most often used in clinical reporting (<https://loinc.org/usage/units/)>. This list likely includes more values than a partner would have in their source system, but we are providing a more comprehensive list to err on the side of completeness.  When mapping from source data to the UCUM values, partners should choose the code that is the closest match to their source data. It is expected that most units will have a 1:1 match with the UCUM list, though in some cases, there will be different spelling variations or representations that need to be harmonized (e.g., mapping PERCENT to %). For laboratory test results and clinical observations, partners should ensure that there is no discordance between the selected unit of measure and the corresponding test code (e.g., choosing % as a unit of measure when the LOINC code indicates the result is reported as mg/dL).  It is possible that a partner may have unit values that are not part of the LOINC-curated list. This is because the universe of UCUM codes is essentially unbounded (<http://unitsofmeasure.org/ucum.html#datyp2apdxatblxmp>). We have standardized on the LOINC-curated list to allow for testing of model conformance. Units that fall outside of this value set should be mapped to “OT,” with the source value stored in the appropriate RAW field, even if they are technically “valid” UCUM codes.  Please note that there are several instances where a single unit code maps to multiple entries (e.g., nmol/mg{prot} maps to both “nanomole of ½ cystine per milligram of protein” and “nanomole per milligram of protein”). This is a known issue and reflective of how the curated list of codes were created by LOINC. The values were distinct at one point but collapsed due to field length restrictions. In addition, there are instances where units differ solely by letter case (e.g., “A” – ampere, “a” – year; “Ms” – megasecond, “ms” – millisecond). We will retain case sensitivity when making comparisons on this field. |

|  |  |
| --- | --- |
| **GENERAL Implementation Guidance (spans more than 1 table)** | |
| *Topic* | *Guidance* |
| 11 – Use of “OT” when mapping to expanded value sets | As partners map their source data to the expanded value sets, the value of “OT” (Other) may be used in cases where a source value is present, but does not match any of the responses in the value set. Additionally, “OT” may be used in those cases when a data partner has not completely mapped all of their source data to the CDM value set for a given variable (e.g., \_UNIT, \_ROUTE, FACILITY\_TYPE, etc.). In this case, partners can include the code from the Value Set Appendix for those entries that have been mapped and “OT” for the remainder  until they have been assigned a code from the value set. Unmapped values should be specified in the corresponding RAW field. |
| 12 – Required vs. Optional fields | The fields that are required to be populated for all records in a given table are listed in the “constraints” section of each table  description. Any table and/or field in the CDM may be required for a partner’s participation in a given study or other PCORnet activity. In assessing foundational data quality, the PCORnet Data Curation query packages may query any CDM table or field (Language replicated from Section 2.5 – Implementation Expectations). |
| 13 – Storing PROs or other observations with date-type responses | When loading PROs with a date-type response, store the character/string version of the date in PRO\_RESPONSE\_TEXT field, and the SAS Date format (Numeric) of the response in PRO\_RESPONSE NUM. Analytically, the PRO\_RESPONSE\_NUM value will be used in queries. The same logic applies for observations with date-type responses that might be stored within OBS\_CLIN or OBS\_GEN. |
| 14 – Future Dates | If the source system assigns all unspecified end dates with a far-off future date instead of null (e.g., year 4700), these values should be set to null when loading the CDM. Only do this for true “dummy” dates that should truly be blank/null, and only for dates that are systematically assigned at the source. |
| 15 – Correspondence between LOINC classes and CDM tables | The LOINC Class Type variable has 4 values (as of May 2018): 1 = Lab Class, 2 = Clinical Class, 3 = Claims Attachment, 4 = Survey. This information can be found on search.loinc.org or in the LOINC Core files. In general, observations with a class type of “1” will be stored in  LAB\_RESULT\_CM. Observations with a class type of “2” will be loaded into OBS\_CLIN, and observations with a class type of “4” will generally be stored in PRO\_CM, but not definitively. |
| 16 – Immunizations | Vaccines captured as part of an immunization workflow should be stored in the IMMUNIZATION table. If immunizations are present in other data streams (e.g., medication administrations, procedures, etc.), however, they should not be removed from those tables. |
| 17 – Lab LOINC mappings | If partners discover that laboratory results have been assigned an incorrect LOINC code within their source system(s), they should make every effort to have it corrected by the personnel who maintain those mappings. In the meantime, it is acceptable to assign the proper  LOINC code before loading those results into the CDM. |

|  |
| --- |
| **5.** **Core CDM Table Specifications** |

|  |
| --- |
| **5.1. 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 Implementation Guidance** |
| *Guidance* |
| * The most recently available information should be populated for BIRTH\_DATE, SEX, and other characteristics. If these attributes have been updated in the patient record, use the most recent value. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-Level Implementation Guidance* |
| 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 |  |
| 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 |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-Level Implementation Guidance* |
| SEX | RDBMS  Text(2) | SAS Char(2) | A=Ambiguous F=Female M=Male NI=No  information UN=Unknown OT=Other | Sex assigned at birth. | MSCDM v4.0 with modified field size and value set  Source: Administrative Sex (HL7)  [http://phinvads.cdc.g](http://phinvads.cdc.gov/vads/ViewValueSet.action?id=06D34BBC-617F-DD11-B38D-00188B398520) [ov/vads/ViewValueS](http://phinvads.cdc.gov/vads/ViewValueSet.action?id=06D34BBC-617F-DD11-B38D-00188B398520) [et.action?id=06D34B](http://phinvads.cdc.gov/vads/ViewValueSet.action?id=06D34BBC-617F-DD11-B38D-00188B398520) [BC-617F-DD11-](http://phinvads.cdc.gov/vads/ViewValueSet.action?id=06D34BBC-617F-DD11-B38D-00188B398520) [B38D-](http://phinvads.cdc.gov/vads/ViewValueSet.action?id=06D34BBC-617F-DD11-B38D-00188B398520) [00188B398520](http://phinvads.cdc.gov/vads/ViewValueSet.action?id=06D34BBC-617F-DD11-B38D-00188B398520) | The “Ambiguous” category may be used for individuals who are physically undifferentiated from birth. The “Other” category may be used for individuals who are undergoing gender re-assignment. |
| SEXUAL\_ORIENTATI ON | 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. | Source: Health IT Certification Criteria, 2015 Base Edition, modified with expert advisory within PCORnet  [https://www.federalre](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base) [gister.gov/documents](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base)  [/2015/10/16/2015-](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base)  [25597/2015-edition-](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base) [health-information-](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base) [technology-health-it-](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base) [certification-criteria-](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base) [2015-edition-base](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* | *Field-Level Implementation Guidance* |
| 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/ Non-binary SE=Something else  MU=Multiple gender categories DC=Decline to answer  NI=No  information UN=Unknown OT=Other | Current gender identity. | Source: Health IT Certification Criteria, 2015 Base Edition, modified with expert advisory within PCORnet  [https://www.federalre](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base) [gister.gov/documents](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base)  [/2015/10/16/2015-](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base)  [25597/2015-edition-](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base) [health-information-](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base) [technology-health-it-](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base) [certification-criteria-](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base) [2015-edition-base](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base) | Use Genderqueer (“GQ”) for patients who report a non-binary gender identify. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-Level Implementation Guidance* |
| 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) | **Populating RACE and HISPANIC if race and ethnicity are not captured separately within the source system** (e.g., “Hispanic or Latino” is included as a selection under Race) - for patients with a known race (e.g., Race is something other than “Hispanic or Latino”, partners should set HISPANIC to "OT" and RACE to the appropriate race code. For patients who are listed as having a race of “Hispanic,” partners should set HISPANIC to "Y" and RACE to "OT". In this situation, the combined race/ethnicity field is treated as known field capturing values for both race and ethnicity, which is why the preference is to use  “OT” instead of “NI”. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-Level Implementation Guidance* |
| 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 | This field is a derived attribute and is not expected to be an explicit data field within a source system |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-Level Implementation Guidance* |
| PAT\_PREF\_LANGUAG E\_SPOKEN | RDBMS  Text(3) | SAS Char(3) | See Value Set Appendix for a list of acceptable values. | Preferred spoken language of communication as expressed by the patient. | ISO639-2 | * This information may be documented in the EHR or an external registry. * Do not impute or derive if not expressly defined in the source system. * Analytically, will assume that “NI” corresponds to a preferred language of English. * Use the value of “ZHO” (Chinese) for both Mandarin and Cantonese, and place the specific value in the RAW field. Within the ISO639-2 value set, there is no distinction between the two. [https://www.loc.gov/standards/iso6 39-2/faq.html#24](https://www.loc.gov/standards/iso639-2/faq.html#24) * Use “OT” for American Sign Language |
| RAW\_SEX | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value of field, prior to  mapping into the PCORnet CDM value set. | PCORnet |  |
| RAW\_  SEXUAL\_ORIENTATI ON | RDBMS  Text(x) | SAS Char(x) | . | 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) | . | Field for originating value, prior to mapping into  the PCORnet CDM value set. | PCORnet |  |
| RAW\_HISPANIC | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, prior to mapping into  the PCORnet CDM value set. | PCORnet |  |
| RAW\_RACE | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, prior to mapping into  the PCORnet CDM value set. | PCORnet |  |
| RAW\_PAT\_PREF\_LAN  GUAGE\_SPOKEN | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, prior to mapping into  the PCORnet CDM value set. | PCORnet |  |

|  |
| --- |
| **5.2. 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.**

# 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.

**Composite Primary Key:** PATID, ENR\_START\_DATE, ENR\_BASIS

# Foreign Key:

ENROLLMENT.PATID is a foreign key to DEMOGRAPHIC.PATID (many-to-one 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 Implementation Guidance** |
| *Guidance* |
| * For partners that do not have insurance-based enrollment information for some of their patients, other approaches can be used to identify periods during which complete medical capture is expected. * Members with medical insurance coverage, with or without drug coverage, or should be included. If a patient has both medical and drug coverage, create the appropriate enrollment records for each. * A break in insurance coverage of at least one day or a change in the chart abstraction flag should generate a new record. * The ENROLLMENT table provides an important analytic basis for identifying periods during which medical care should be observed, for calculating person-time, and for inferring the meaning of unobserved care (i.e., if care is not observed, it likely did not happen). The most recently available information should be populated for BIRTH\_DATE, SEX, and other characteristics. If these attributes have been updated in the patient record, please use the most recent value. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-Level Implementation Guidance* |
| 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 | * For implementation of the CDM, a long span of longitudinal data is desirable. However, an enrollment record is intended to represent the dates for which there is complete capture of all medically-attended events, so the data partner’s knowledge of the validity and usability of the data should inform their choice of enrollment start date, especially for historical data more than a decade old. * If ENR\_BASIS is encounter- based (“E”), the enrollment start date should not be earlier than the earliest encounter date in the datamart for that patient. If ENR\_BASIS is based on insurance coverage (“I”), then the enrollment start date *may* occur before the earliest encounter date in the datamart for that patient. |
| 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 |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-Level Implementation Guidance* |
| 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 | * Mark as "Yes" if there are no contractual or other restrictions between you and the individual (or sponsor) that would prohibit you from requesting any chart for this patient. * This field is a derived attribute and is not expected to be an explicit data field within a source system |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-Level Implementation Guidance* |
| 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.) | PCORnet  Based upon the HMORN VDW and  Sentinel CDM v6.0 | * When an algorithmic method is used to determine the ENR\_BASIS, the exact details should be described in the ETL ADD. * This field is a derived attribute and is not expected to be an explicit data field within a source system |
|  |  |  |  | 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. |  |

|  |
| --- |
| **5.3.** **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 (many-to-one relationship) ENCOUNTER.PROVIDERID is a foreign key to PROVIDER.PROVIDERID (many-to-one relationship)

# Constraints:

ENCOUNTERID (unique; required, not null) PATID (required, not null)

ADMIT\_DATE (required, not null) ENC\_TYPE (required, not null)

|  |
| --- |
| **ENCOUNTER Table Implementation Guidance** |
| *Guidance* |
| * Each ENCOUNTERID will generally reflect a unique combination of PATID, ADMIT\_DATE, PROVIDERID and ENC\_TYPE. * Every diagnosis and procedure recorded during the encounter should have a separate record in the DIAGNOSIS or PROCEDURES Tables. * Multiple visits to the **same** provider on the same day may be considered one encounter, especially if defined by a reimbursement basis; if so, the ENCOUNTER record should be associated with all diagnoses and procedures that were recorded during those visits. * Visits to **different** providers for different encounter types on the same day, however, such as a physician appointment that leads to a hospitalization, would generally correspond to multiple encounters within the ENCOUNTER table. * Rollback or voided transactions and other adjustments should be processed before populating this table. * Although “Expired” is represented in both DISCHARGE\_DISPOSITION and DISCHARGE\_STATUS, this overlap represents the reality that both fields are captured in hospital data systems but with variation in how each field is populated. * Do not include scheduled encounters. * Partners should ensure that “administrative” encounters (e.g., e-mail, phone, documentation-only), are coded to the appropriate encounter type, which is typically “OA” for outpatient visits. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| 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 | All PATIDs in this table must be present in the DEMOGRAPHIC  table. |
| 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* | *Field-level Implementation Guidance* |
| 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 | Should be populated for all Inpatient Hospital Stay (IP), Non- Acute Institutional Stay (IS) encounter types and ED-to- Inpatient (EI) encounter types.  May be populated for Emergency Department (ED) encounter types. Should be missing for ambulatory visit (AV or OA) encounter types, though it may be present for  Observation Stays. |
| 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 |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| PROVIDERID | RDBMS  Text(x) | SAS Char(x) | . | Code for the provider who is most responsible for this encounter. As with the PATID, the provider code is a pseudoidentifier with a consistent crosswalk to the real identifier. | MSCDM v4.0 | * PROVIDERID generally refers to the person most responsible for providing care during the encounter, though it may also correspond to a device (e.g., MRI) for certain procedure-only encounters. * If populated, PROVIDERID must be present in the PROVIDER table. |
| FACILITY\_LOCATION | RDBMS  Text(5) | SAS Char(5) | . | Geographic location (5-digit zip code). | MSCDM v4.0 with  modified field type | Updated in CDM v5.0 to  support 5-digit zip codes. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 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.  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.  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.](https://www.medicare.gov/Pubs/pdf/11435.pdf)  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. This includes physician consults for patients during inpatient encounters that are not directly related to the cause of the admission (e.g. a ophthalmologist consult for a patient with diabetic ketoacidosis) (guidance updated in v4.0).  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 | * **Observation stays**– If partners are able to identify observation stays within their data, these encounters should be labeled “OS.” Typical observation stays last 24-48 hours. If partners find that they have observation stays that last significantly longer (e.g., weeks), this should also be documented in the ETL ADD. * Same-day surgery, OT/PT, and provider office visits for treatment/testing should be labeled as “AV.” * For the situation where an Emergency Department (ED) encounter leads to a hospital admission   + The optimal, preferred state is to have one record for the ED (ENC\_TYPE=ED), and a separate record for the hospital admission (ENC\_TYPE=IP)   + However, this separation does not always exist in source data records. If the source system combines the ED and IP basis into one concept, a permissible substitution is to use ENC\_TYPE=EI   + Never merge separate ED and IP records together. * **Services rendered in an inpatient setting that cannot be combined with the facility encounter** – Inpatient (IP) and ED to Inpatient (EI) encounter types should be limited to encounters which include the facility component of the admission since these data are required to fully populate the expected fields (e.g. Discharge Date, Admitting Source, Discharge Disposition, Discharge Status). If a partner has data for professional services that occur in an inpatient care setting that cannot be combined with the associated facility   encounter, the partner should map |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
|  |  |  |  |  |  | these services to "OT" and document this in their ETL ADD.   * Generally, a reimbursement basis will determine the source system classification, instead of physical location. For example, a patient may occupy a hospital bed during an observation that is not classified as an inpatient hospital stay. * Please note that stand-alone urgent care facilities are usually not established as Emergency Departments. |
| 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 | * If populating FACILITY\_TYPE, ensure that multiple FACILITY\_TYPE values are not used to describe the same FACILITYID. * If a facility exists that operates both inpatient and outpatient units and is described in the source system by the same facility id, a potential solution is to append the source id with a suffix to create “sub facilities” that can be used to distinguish locations with different levels of service. |
| 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 | Should be populated for Inpatient Hospital Stay (IP), Non-Acute Institutional Stay (IS) and ED-to- Inpatient (EI) encounter types.  May be populated for Emergency Department (ED) encounter types. Should be missing for ambulatory visit (AV or OA) encounter types,  though it may be present for Observation Stays. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| 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=Rehabilitatio n 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 | Should be populated for Inpatient Hospital Stay (IP), Non-Acute Institutional Stay (IS) and ED-to-Inpatient (EI) encounter types. May be populated for Emergency Department (ED) encounter types. Should be missing for ambulatory visit (AV or OA) encounter types, though it may be present for Observation Stays. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| 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 | * Should be populated for Inpatient Hospital Stay (IP), Non-Acute Institutional Stay (IS) and ED-to-Inpatient (EI) encounter types. May be populated for Emergency Department (ED) encounter types. Should be missing for ambulatory visit (AV or OA) encounter types, though it may be present for Observation Stays. * Use leading zeroes for codes less than 100. * For records with multiple DRGs assigned, choose the most appropriate one based on the source data. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| 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 | * MS-DRG (current system) began on 10/1/2007. * Should be populated for Inpatient Hospital Stay (IP), Non-Acute Institutional Stay (IS) and ED-to-Inpatient (EI) encounter types. May be populated for Emergency Department (ED) encounter types. Should be missing for ambulatory visit (AV or OA) encounter types, though it may be present for Observation Stays. * This field is a derived attribute and is not expected to be an explicit data field within a source system |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| 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=Rehabilitatio n Facility RS=Residential Facility SN=Skilled Nursing Facility IH=Intra-hospital NI=No  information UN=Unknown  OT=Other | Admitting source. | MSCDM v4.0 with modified value set | * Should be populated for Inpatient Hospital Stay (IP), Non-Acute Institutional Stay (IS) and ED-to-Inpatient (EI) encounter types. May be populated for Emergency Department (ED) encounter types. Should be missing for ambulatory visit (AV or OA) encounter types, though it may be present for Observation Stays. * When a patient is discharged from one distinct entity in a hospital and admitted to another, *resulting in a separate claim*, use “IH” |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| PAYER\_TYPE\_ PRIMARY | RDBMS  Text(5) | SAS Char(5) | See Value Set Appendix for a list of acceptable values. | Categorization of payer type for primary payer associated with the encounter | PHDSC Source of Payment Typology | * Do not derive if not already assigned through a validated process (e.g., by hospital billing department) * Map to the most granular value the source data will support. * Additional information can be found in the Payer Typology User Guide, located here: [http://www.phdsc.org/standards/pay er-typology.asp](http://www.phdsc.org/standards/payer-typology.asp) * Even if payer type is not specified, populating RAW\_PAYER\_TYPE\_PRIM ARY and RAW\_PAYER\_NAME\_PRI MARY will allow a value to be determined through PCORnet-wide or study-   specific process |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| PAYER\_TYPE\_ SECONDARY | RDBMS  Text(5) | SAS Char(5) | See Value Set Appendix for a list of acceptable values. | Categorization of payer type for secondary payer associated with the encounter | PHDSC Source of Payment Typology | * Do not derive if not already assigned through a validated process (e.g., by hospital billing department) * Map to the most granular value the source data will support. * Additional information can be found in the Payer Typology User Guide, located here: [http://www.phdsc.org/standards/pay er-typology.asp](http://www.phdsc.org/standards/payer-typology.asp) * Even if payer type is not specified, populating RAW\_PAYER\_TYPE\_SECO NDARY and RAW\_PAYER\_NAME\_SEC ONDARY will allow a value to be determined through PCORnet-wide or study- specific process |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| FACILITY\_TYPE | RDBMS  Text(x) | SAS Char(x) | See Value Set Appendix for a list of acceptable values. | Description of the facility where the encounter occurred. | PCORnet | * Do not assign more than one FACILITY\_TYPE to a single FACILITYID * If a facility exists that operates both inpatient and outpatient units and is described in the source system by the same facility id, a potential solution is to append the source id with a suffix to create “sub facilities” that can be used to distinguish locations with different levels of service. * For office visits at an academic medical center, select the most appropriate hospital outpatient clinic value. * Pediatric specialty clinics should map the relevant specialty clinic, if available. The PEDIATRIC facility types would best be applied to a General Pediatrics clinic. * A draft mapping between the CMS Place of Service value set and FACILITY\_TYPE can be found here: https://github.com/CDMFORUM/C DM-GUIDANCE/issues/67 |
| RAW\_SITEID | RDBMS  Text(x) | SAS Char(x) | . | 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 |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| RAW\_ENC\_TYPE | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, prior to mapping  into the PCORnet CDM value set. | PCORnet |  |
| RAW\_DISCHARGE\_DISPOSIT  ION | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, prior to mapping  into the PCORnet CDM value set. | PCORnet |  |
| RAW\_DISCHARGE\_STATUS | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, prior to mapping  into the PCORnet CDM value set. | PCORnet |  |
| RAW\_DRG\_TYPE | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, prior to mapping  into the PCORnet CDM value set. | PCORnet |  |
| RAW\_ADMITTING\_SOURCE | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, prior to mapping  into the PCORnet CDM value set. | PCORnet |  |
| RAW\_FACILITY\_TYPE | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, prior to mapping  into the PCORnet CDM value set. | PCORnet |  |
| RAW\_PAYER\_TYPE\_PRIMAR  Y | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, prior to mapping  into the PCORnet CDM value set. | PCORnet |  |
| RAW\_PAYER\_NAME\_PRIMA RY | RDBMS  Text(x) | SAS Char(x) | . | Primary payer name as denoted in the source system. Used to derive PAYER\_TYPE\_PRIMARY if validated  process does not exist. | PCORnet | * Name of secondary payer associated with the encounter. * Partners should only populate if local governance allows it. |
| RAW\_PAYER\_ID\_PRIMARY | RDBMS  Text(x) | SAS Char(x) | . | Primary PAYER identifier as denoted in the source system. Used to derive PAYER\_TYPE\_PRIMARY if validated process does not exist. | PCORnet | * Identifier associated with the primary payer. * Partners should only populate if local governance allows it. |
| RAW\_PAYER\_TYPE\_SECOND  ARY | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, prior to mapping  into the PCORnet CDM value set. | PCORnet |  |
| RAW\_PAYER\_NAME\_SECON DARY | RDBMS  Text(x) | SAS Char(x) | . | Secondary payer name as denoted in the source system. Used to derive PAYER\_TYPE\_SECONDARY if validated  process does not exist. | PCORnet | * Name of secondary payer associated with the encounter. * Partners should only populate if local governance allows it. |
| RAW\_PAYER\_ID\_SECONDAR Y | RDBMS  Text(x) | SAS Char(x) | . | Secondary PAYER identifier as denoted in the source system. Used to derive PAYER\_TYPE\_SECONDARY if validated  process does not exist. | PCORnet | * Identifier associated with the secondary payer. * Partners should only populate if local governance allows it. |

|  |
| --- |
| **5.4. 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 (many-to-one relationship) DIAGNOSIS.ENCOUNTERID is a foreign key to ENCOUNTER.ENCOUNTERID (many-to-one relationship) DIAGNOSIS.PROVIDERID is a foreign key to PROVIDER.PROVIDERID (many-to-one relationship)

# Constraints:

DIAGNOSISID (unique; required, not null) PATID (required, not null)

DX (required, not null) DX\_TYPE (required, not null) DX\_SOURCE (required, not null)

|  |
| --- |
| **DIAGNOSIS Table Implementation Guidance** |
| *Guidance* |
| * This table should capture all uniquely recorded diagnoses for all encounters, with the exception of problem list entries. If partners have access to multiple versions of each diagnosis within a given encounter (e.g., admitting, interim, final), the preference is to prioritize final or discharge diagnoses. A value should be specified in DX\_SOURCE to indicate the classification of the diagnosis. * Diagnoses from problem lists will be captured in the CONDITION table. * If a patient has multiple diagnoses associated with one encounter, then there would be one record in this table for each diagnosis. * ENCOUNTERID should be populated for DIAGNOSIS and PROCEDURES. The definitions of the DIAGNOSIS and PROCEDURES tables are dependent upon a healthcare context; therefore, the encounter basis is necessary and the ENCOUNTERID, PROVIDERID, ENCOUNTER\_TYPE, and ADMIT\_DATE from the associated ENCOUNTER record should be included. While not desirable, a low percentage of orphan records is permissible to accommodate instances in which the associated ENCOUNTER details are missing from the source data. * Data in this table are expected to be from healthcare-mediated processes and reimbursement drivers, including technical/facility billing, professional billing and other data streams. **Do not omit** billing data unless it is unavailable from the source system or the partner is certain that the diagnoses loaded from the non-billing system (e.g., the EHR) represents completely the diagnosis data available from the billing system. Data from these different streams have different analytical utility so there is a benefit to including both if available. * Diagnoses are often only related to the **treatment** of the patient during the specific encounter. Chronic conditions that are not be pertinent to the treatment of a specific encounter, for example, would not be expected to be present. * If a local vocabulary is used, but *canno*t be mapped to a standard vocabulary such as ICD-9-CM, DX\_TYPE should be populated as “Other” and the local value stored in DX. If the local value can be mapped to a standard vocabulary, follow the guidance around the population of Raw fields (General Guidance #1). * Partners should continue to populate ADMIT\_DATE, even if they are populating DX\_DATE. Analyses may leverage either date, or both. DX\_DATE can be particularly useful for identifying diagnoses or conditions that might have developed over the course of a long inpatient encounter. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| DIAGNOSISID | RDBMS Text(x) | SAS Char(x) | . | Arbitrary identifier for each  unique record. | PCORnet |  |
| PATID | RDBMS Text(x) | SAS Char(x) | . | Arbitrary person-level identifier. Used to link across tables. | MSCDM v4.0 | All PATIDs in this table must be present in the DEMOGRAPHIC  table. |
| ENCOUNTERID | RDBMS Text(x) | SAS Char(x) | . | Arbitrary encounter-level identifier. Used to link across  tables. | MSCDM v4.0 | All ENCOUNTERIDs in this table must be present in the ENCOUNTER table. |
| 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 | Should be non-null for all records replicated from ENCOUNTER table (guidance added in v4.0). |
| 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 |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| PROVIDERID | RDBMS Text(x) | SAS Char(x) | . | Identifier associated with the provider most responsible for the diagnosis. | MSCDM v4.0 | * The PROVIDERID from the ENCOUNTER can be used if provider assigning the diagnosis is unknown. * Use the ID of the attending provider if the diagnosis is assigned by a resident/intern. * All PROVIDERIDs must be   present in the PROVIDER table. |
| 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. | MSCDM v4.0 | * Please populate the exact textual value of this diagnosis code, but remove source- specific suffixes and prefixes. Other codes should be listed as recorded in the source data. * Do not include leading zeros beyond those that are part of the code itself (i.e., represent ICD-9 diagnosis 001.9 as “001.9”, not “000001.9” or some other variation). * Decimal points may or may not be present for ICD-9/ICD-10 codes. If the decimal point is missing from source data, do not add. If it is present in source data, do not remove. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| 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 | This field is a derived attribute and is not expected to be an explicit data field within a source system |
| DX\_DATE | RDBMS Date | SAS Date (Numeric) | . | Date diagnosis was recorded, if known. | PCORnet | * This field is meant to represent when a diagnosis was first recorded in the source system during a given encounter, not necessarily when a patient was first diagnosed with a given condition. * If the source system records a date for each diagnosis for every day of the encounter, only populate this field if it actually represents the date the diagnosis was first recorded. Otherwise,   leave blank. |
| 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 | * It is not necessary to populate interim diagnoses unless readily available. * Ambulatory encounters would generally be expected to have a source of “Final.” |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| DX\_ORIGIN | RDBMS Text(2) | SAS Char(2) | OD=Order/EHR BI=Billing CL=Claim DR=Derived 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. | PCORnet | * Use “OD” for diagnoses entered into the EHR that are associated with an Order. * Use “OD” for any diagnosis associated with an encounter that is entered into the EHR by a provider. * Use “BI” for all diagnoses that are generated through the physician and hospital billing process. * Use “DR” for all diagnoses that are derived or imputed through analytical procedures (e.g., natural language processing). This does not apply to diagnoses that have been mapped from vocabulary mapping software/middleware (e.g., IMO) See General Guidance #4. In those instances, use “OD” or “BI”, depending on the provenance of the diagnosis. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| 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 | * Value expected for IP, IS, EI, and OS encounters. May also be present for other encounter types. * One principal diagnosis per encounter is expected, although in some instances more than one diagnosis may be flagged as principal. * **Professional vs. Facility** – If a partner has access to both professional and facility diagnosis data for a given encounter, the facility diagnoses should be used to populate this field. Partners should document their approach in   their ETL ADD. |
| DX\_POA | RDBMS Text(2) | SAS Char(2) | Y = Diagnosis present  N = Diagnosis not present U = Insufficient documentation  W = Clinically undetermined  1 = Unreported / not used NI=No information UN=Unknown OT=Other | Flag to denote whether diagnosis was present on inpatient admission.  Y = Diagnosis present at time of inpatient admission  N = Diagnosis not present at time of inpatient admission  U = Documentation insufficient to determine if the condition was present at the time of inpatient admission  W = Clinically undetermined. Provider unable to clinically determine whether the condition was present at the time of inpatient admission  1 = Unreported / not used. Exempt from present-on-admission reporting. | CMS | * Include for EI, IP visits only * If data are populated for some inpatient diagnoses, but not all, use “UN” for diagnoses with blank/null value * Only assign a value of “1” to a diagnosis if it is reported that way in the source system. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| RAW\_DX | RDBMS Text(x) | SAS Char(x) | . | Field for originating value, prior to mapping into the PCORnet  CDM value set. | PCORnet |  |
| RAW\_DX\_TYPE | RDBMS Text(x) | SAS Char(x) | . | Field for originating value, prior  to mapping into the PCORnet CDM value set. | PCORnet |  |
| RAW\_DX\_SOURCE | RDBMS Text(x) | SAS Char(x) | . | Field for originating value, prior to mapping into the PCORnet  CDM value set. | PCORnet |  |
| RAW\_PDX | RDBMS Text(x) | SAS Char(x) | . | Field for originating value, prior  to mapping into the PCORnet CDM value set. | PCORnet |  |
| RAW\_DX\_POA | RDBMS Text(x) | SAS Char(x) | . | Field for originating value, prior  to mapping into the PCORnet CDM value set. | PCORnet |  |

|  |
| --- |
| **5.5.** **Table: PROCEDURES** |

**PROCEDURES Domain Description:**

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

# Relational Integrity:

The PROCEDURES table contains one record per PROCEDURESID.

**Primary Key:** PROCEDURESID

# Foreign Keys:

PROCEDURES.PATID is a foreign key to DEMOGRAPHIC.PATID (many-to-one relationship) PROCEDURES.ENCOUNTERID is a foreign key to ENCOUNTER.ENCOUNTERID (many-to-one relationship) PROCEDURES.PROVIDERID is a foreign key to PROVIDER.PROVIDERID (many-to-one relationship)

# Constraints:

PROCEDURESID (unique; required, not null) PATID (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 Implementation Guidance** |
| *Guidance* |
| * This table should capture all uniquely recorded procedures for all encounters, including office or evaluation and management visits, diagnostic testing, laboratory test orders, medication administrations, or other services rendered by a clinician. * If a patient has multiple procedures ordered during one encounter, then there would be one record in this table for each procedure. * ENCOUNTERID should be populated for DIAGNOSIS and PROCEDURES. The definitions of the DIAGNOSIS and PROCEDURES tables are dependent upon a healthcare context; therefore, the encounter basis is necessary and the ENCOUNTERID, PROVIDERID, ENCOUNTER\_TYPE, and ADMIT\_DATE from the associated ENCOUNTER record should be included. While not desirable, a low percentage of orphan records is permissible to accommodate instances in which the associated ENCOUNTER details are missing from the source data. * Data in this table are expected to be from healthcare-mediated processes and reimbursement drivers, including technical/facility billing, professional billing and other data streams. **Do not omit** billing data unless it is unavailable from the source system or the partner is certain that the procedures loaded from the non-billing system (e.g., the EHR) represents completely the procedure data available from the billing system * If a local vocabulary is used, but cannot be mapped to a standard vocabulary such as ICD-9-CM, PX\_TYPE should be populated as “Other” and the local value stored in PX. If the local value can be mapped to a standard vocabulary, follow the guidance around the population of Raw fields (General Guidance #1). * Evidence of medications administered in outpatient settings should be present in the PROCEDURES table if that information is included with other billed/ordered PROCEDURES. * Evidence of inpatient administrations should be present in the PROCEDURES table if that information is included with other billed/ordered PROCEDURES. * **DO NOT** include records from medication administration sources (e.g., electronic medication administration records) in this table. * If possible to determine from the source data, only include procedures that have actually occurred. * **Inclusion of laboratory orders –** If possible, partners should include laboratory orders within the PROCEDURES table to support potential studies of appropriate laboratory monitoring. This includes those orders without a corresponding result in the LAB\_RESULT\_CM table. Do not include canceled orders. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| PROCEDURESID | RDBMS Text(x) | SAS Char(x) | . | Arbitrary identifier for each unique record. | PCORnet |  |
| PATID | RDBMS Text(x) | SAS Char(x) | . | Arbitrary person-level identifier. Used to link across tables. | MSCDM v4.0 | All PATIDs in this table must be present in the DEMOGRAPHIC  table. |
| ENCOUNTERID | RDBMS Text(x) | SAS Char(x) | . | Arbitrary encounter-level identifier. Used to link across tables. | MSCDM v4.0 | All ENCOUNTERIDs in this table  must be present in the ENCOUNTER table. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| 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 | Should be non-null for all records replicated from ENCOUNTER table (guidance added in v4.0). |
| 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) | . | Identifier of the PROVIDER most  associated with the procedure order. | MSCDM v4.0 | All PROVIDERIDs must be present  in the PROVIDER table. |
| PX\_DATE | RDBMS Date | SAS Date (Numeric) | . | Date the procedure was performed. | PCORnet | PX\_DATE should reflect the date the procedure was performed, if known. Leave blank if you cannot determine if or when the procedure was performed. |
| PX | RDBMS  Text(11) | SAS Char(11) | . | Procedure code. | MSCDM v4.0 | Decimal points may or may not be present for ICD-9/ICD-10 procedure codes. If the decimal point is missing, do not add. If it is present,  do not remove. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| PX\_TYPE | RDBMS Text(2) | SAS Char(2) | 09=ICD-9-CM  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. | MSCDM v4.0 with modified field name and value set | * **Expected/known length of codes for some terminologies in PX\_TYPE**:   o ICD-9-CM (09): 3-4  numbers, including a decimal  o ICD-10-PCS (10): 7  alphanumeric characters  o CPT/HCPCS: 5 alphanumeric characters; may be longer if modifiers are included   * CPT and HCPCS codes should be assigned a value of “CH.” * This field may be a derived attribute. In these cases, it would not be expected to be an explicit data field within a source system |
|  |  |  |  | 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. |  |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| PX\_SOURCE | RDBMS Text(2) | SAS Char(2) | OD=Order/EHR BI=Billing CL=Claim DR=Derived 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. | PCORnet | * If evaluation and management (E/M) or level of service (LOS) codes are available, they should be included * This field is a derived attribute and is not expected to be an explicit data field within a source system * Use “OD” for procedures entered into the EHR that are associated with an Order. * Use “OD” for any procedures associated with an encounter that is entered into the EHR by a provider. * Use “BI” for all procedures that are generated through the physician and hospital billing process. * Use “DR” for all procedure records that are derived or imputed through analytical procedures (e.g., natural language processing). This does not apply to procedures mapped from a superset terminology   (General Guidance #4). |
| PPX | RDBMS Text(2) | SAS Char(2) | P=Principal S=Secondary NI=No information UN=Unknown OT=Other | Principal procedure flag. | PCORnet | * Value may be present for IP, IS, EI, and OS encounters. * One principal procedure per encounter is expected, although in some instances more than one procedure may be flagged as   principal. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| RAW\_PX | RDBMS Text(x) | SAS Char(x) | . | Field for originating value, prior to mapping into the PCORnet CDM  value set. | PCORnet |  |
| RAW\_PX\_TYPE | RDBMS Text(x) | SAS Char(x) | . | Field for originating value, prior to mapping into the PCORnet CDM  value set. | PCORnet |  |
| RAW\_PPX | RDBMS Text(x) | SAS Char(x) | . | Field for originating value, prior to mapping into the PCORnet CDM  value set. | PCORnet |  |

|  |
| --- |
| **5.6.** **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 (many-to-one relationship) VITAL.ENCOUNTERID is a foreign key to ENCOUNTER.ENCOUNTERID (zero/many-to-one relationship)

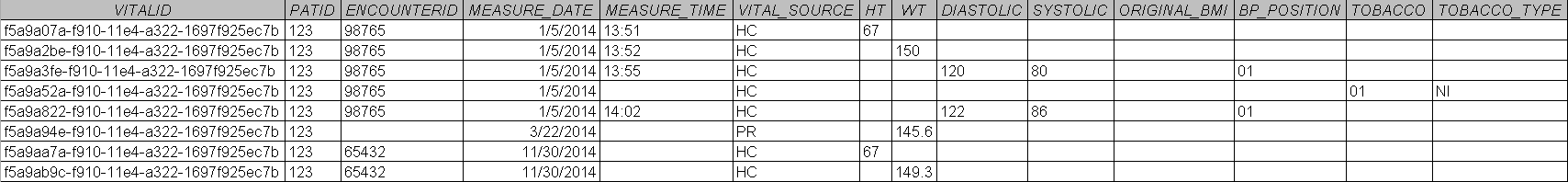
# Constraints:

VITALID (unique; required, not null) PATID (required, not null) MEASURE\_DATE (required, not null) VITAL\_SOURCE (required, not null)

|  |  |
| --- | --- |
| **VITAL Table Implementation Guidance** | |
|  | *Guidance* |
| * This table includes measurements recorded in both healthcare and non-healthcare settings. * The VITAL table contains one record per result/entry. Multiple measurements may exist in source data (for example, 3 blood pressure readings on the same day); in this case, each measurement would be a separate record. If multiple vitals are collected at the same time (e.g., height, weight and blood pressure recorded at the start of an encounter), it is permissible to store these values in a single record. This table should be populated with all available measures, with the possible exception(s) noted below. * If a partner has access to vital signs that are sourced from a device feed, they should make an assessment about data volume before including these measures, particularly if multiple readings per day are present for a large percentage of their population. Measures should not be averaged or aggregated.   o For healthcare device data sources: If multiple readings are available and the volume of data is judged by the data partner to be too burdensome for inclusion, using the set of values that were recorded directly in the medical record preferred over any algorithmic selection process.   * For personal device data sources: If multiple readings are available and the volume of data is judged by the data partner to be too high for inclusion, the project/study leadership should define a method for selecting individual measurements and this logic should be documented in the ETL ADD. | |

Figure 1. Example of populated VITAL table.

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* | *Field-level Implementation Guidance* |
| VITALID | RDBMS Text(x) | SAS Char(x) | . | Arbitrary identifier for each unique  VITAL record. | PCORnet |  |
| PATID | RDBMS Text(x) | SAS Char(x) | . | Arbitrary person-level identifier. Used  to link across tables. | MSCDM v4.0 | All PATIDs must be present in the DEMOGRAPHIC table. |
| ENCOUNTERID | RDBMS Text(x) | SAS Char(x) | . | Arbitrary encounter-level identifier.  Not all vital sign measures will be associated with a healthcare encounter. | PCORnet | * ENCOUNTERID should generally be present if the vitals were measured as part of the healthcare delivery captured by this datamart. * All ENCOUNTERIDs in this table must be present in the   ENCOUNTER table. |
| 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* | *Field-level Implementation Guidance* |
| 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 Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and*  *Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level Implementation Guidance* |
| 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  DR=Derived NI=No information UN=Unknown OT=Other | Please note: The “Patient-reported” category can include reporting by patient’s family or guardian. | PCORnet | * This field is a derived attribute and is not expected to be an explicit data field within a source system. * If the source of a given vital sign is not explicitly present in the source system, partners should infer a value for VITAL\_SOURCE based on the data and workflow used to collect them. If there is uncertainty as to whether the values come directly from a device, partners should use the more general value/context (patient- reported or healthcare delivery setting). If it is not possible to infer whether the value is patient-reported or was collected in a healthcare delivery setting, partners should choose NI (“no information”). * Use “DR” for all vital signs that are derived or imputed through analytical procedures (e.g., natural language processing). |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| HT | RDBMS  Number(x) | 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. | MSCDM v4.0 |  |
| WT | RDBMS  Number(x) | 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. | MSCDM v4.0 |  |
| DIASTOLIC | RDBMS  Number(x) | 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. | MSCDM v4.0 |  |
| SYSTOLIC | RDBMS  Number(x) | 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. | MSCDM v4.0 |  |
| ORIGINAL\_BMI | RDBMS  Number(x) | SAS  Numeric(length 8) | . | BMI if calculated in the source system. Decimal precision is permissible.  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. | 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 |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 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-](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) [certification/2014-edition-draft-test-procedures/170-314-a-](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) [11-smoking-status-2014-test-procedure-draft-v1.0.pdf](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/Regul](http://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/downloads/9_Record_Smoking_Status.pdf) [ations-and-](http://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/downloads/9_Record_Smoking_Status.pdf) [Guidance/Legislation/EHR](http://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/downloads/9_Record_Smoking_Status.pdf) [IncentivePrograms/downlo](http://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/downloads/9_Record_Smoking_Status.pdf) [ads/9\_Record\_Smoking\_St](http://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/downloads/9_Record_Smoking_Status.pdf) [atus.pdf](http://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/downloads/9_Record_Smoking_Status.pdf) [retrieved January 11, 2015] |  |
| TOBACCO | RDBMS Text(2) | SAS Char(2) | 01=Current user  02=Never 03=Quit/former user | Indicator for any form of tobacco. | MSCDM v4.0  with modified |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
|  |  |  | 04=Passive or environmental exposure 06=Not asked  NI=No information UN=Unknown OT=Other |  | 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) | . | Field for originating value, prior to mapping into the PCORnet CDM value  set. | PCORnet |  |
| RAW\_SYSTOLIC | RDBMS Text(x) | SAS Char(x) | . | Field for originating value, prior to mapping into the PCORnet CDM value set. | PCORnet |  |
| RAW\_BP\_POSITION | RDBMS Text(x) | SAS Char(x) | . | Field for originating value, 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* | *Field-level Implementation Guidance* |
| RAW\_SMOKING | RDBMS Text(x) | SAS Char(x) | . | Field for originating value, prior to mapping into the PCORnet CDM value  set. | PCORnet |  |
| RAW\_TOBACCO | RDBMS Text(x) | SAS Char(x) | . | Field for originating value, prior to  mapping into the PCORnet CDM value set. | PCORnet |  |
| RAW\_TOBACCO\_TYPE | RDBMS Text(x) | SAS Char(x) | . | Field for originating value, prior to mapping into the PCORnet CDM value  set. | PCORnet |  |

|  |
| --- |
| **5.7. Table: DISPENSING** |

**DISPENSING Domain Description:**

Prescriptions filled through a community, mail-order or hospital pharmacy. Outpatient dispensing may not be directly captured within healthcare systems.

(Domain description updated in v4.0)

# Relational Integrity:

The DISPENSING table contains one record per DISPENSINGID.

**Primary Key:** DISPENSINGID

# Foreign Keys:

DISPENSING.PATID is a foreign key to DEMOGRAPHIC.PATID (many-to-one 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 Implementation Guidance** |
| *Guidance* |
| * Each record represents an outpatient pharmacy dispensing. * This domain is commonly available in claims data, but may not be available in many EHR data sources. * Dispensing records are different from medication orders or prescribing records, data from medication administration activities, as well as the medication reconciliation of the active medication list. * Administered medications should NOT be stored in this table. They should be stored in the MED\_ADMIN table. Evidence of medications administered in outpatient settings, such as infusions given in medical practices, or those administered in an inpatient setting may be present in the PROCEDURES table if that level of detail is available in the source procedure data. * Rollback transactions and other adjustments that are indicative of a dispensing being canceled or not picked up by the member should be processed (removed) before populating this table. This may be handled differently by Data Partners and may be affected by billing cycles. * In the uncommon situation where one NDC is dispensed more than once for a given patient on a given day, it is acceptable to combine the values from the multiple dispensings for days supply and   number of units. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| DISPENSINGID | RDBMS  Text(x) | SAS Char(x) | . | Arbitrary identifier for each unique record. | PCORnet |  |
| PATID | RDBMS  Text(x) | SAS Char(x) | . | Arbitrary person-level identifier. Used to link across  tables. | MSCDM v4.0 | All PATIDs must be present in the DEMOGRAPHIC table. |
| 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 | NDC must be in HIPAA format. Guidance on normalization for other forms of NDC can be found here: [http://www.nlm.nih.gov/researc](http://www.nlm.nih.gov/research/umls/rxnorm/docs/2012/rxnorm_doco_full_2012-1.html) [h/umls/rxnorm/docs/2012/rxnor](http://www.nlm.nih.gov/research/umls/rxnorm/docs/2012/rxnorm_doco_full_2012-1.html)  [m\_doco\_full\_2012-1.html](http://www.nlm.nih.gov/research/umls/rxnorm/docs/2012/rxnorm_doco_full_2012-1.html) (see section 6) |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| DISPENSE\_SOURCE | RDBMS  Text(2) | SAS Char(2) | OD=Order/EHR BI=Billing CL=Claim PM=Pharmacy Benefit Manager DR=Derived NI=No  information UN=Unknown OT=Other | Source of the dispensing information. | PCORnet | * This field is a derived attribute and is not expected to be an explicit data field within a source system * Use “OD” for dispensing records that are sourced from the EHR or internal pharmacy systems. * Use “BI” for all dispensing records that are generated through the physician and hospital billing process. * Use “CL” for dispensing records that are sourced from pharmacy or medical claims. * Use “DR” for all dispensing records that are derived or imputed through analytical procedures (e.g., natural language processing). * Use “PM” for dispensing records sourced from a pharmacy benefit manager (e.g., Surescripts, Express   Scripts). |
| DISPENSE\_SUP | RDBMS  Number(x) | SAS  Numeric(len gth 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. | 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* | *Field-level Implementation Guidance* |
| DISPENSE\_AMT | RDBMS  Number(x) | SAS  Numeric(len gth 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. | MSCDM v4.0 with modified field name |  |
| DISPENSE\_DOSE\_DI SP | RDBMS  Number(x) | SAS  Numeric(len gth 8) | . | Dose of a given mediation, as dispensed | PCORnet | Do not impute or derive. Populate only if captured in the source system as a discrete  value. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| DISPENSE\_DOSE\_DI SP\_UNIT | RDBMS  Text(x) | SAS Char(x) | See Value Set Appendix for a list of acceptable values. | Units of measure associated with the dose of the medication as dispensed | UCUM | * Do not impute or derive. Populate only if captured in the source system as a discrete value. * Choose the standardized unit of measure that is most reflective of the source data. * ~~The Value Set Appendix contains a list of the units most commonly associated with medication records. Partners can use this table to aid in their mapping efforts, but they should refer back to the full value set if they have a medication record with a unit of measure that is not present in this curated list~~. (Guidance deprecated as of CDM v5.0) * This is a mixed case value set and entries should be handled accordingly. |
| DISPENSE\_ROUTE | RDBMS  Text(x) | SAS Char(x) | See Value Set Appendix for a  list of acceptable values. | Route of delivery |  | Do not impute or derive. Populate only if captured in the source system as a discrete value. |
| RAW\_NDC | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, prior to mapping into the  PCORnet CDM value set. |  |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| RAW\_DISPENSE\_DO  SE\_DISP | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, prior to mapping into the  PCORnet CDM value set. |  |  |
| RAW\_DISPENSE\_DO  SE\_DISP\_UNIT | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, prior to mapping into the  PCORnet CDM value set. |  |  |
| RAW\_DISPENSE\_RO  UTE | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, prior to mapping into the  PCORnet CDM value set. |  |  |

|  |
| --- |
| **5.8.** **Table: LAB\_RESULT\_CM** |

**LAB\_RESULT\_CM Domain Description:**

This table is used to store quantitative and qualitative measurements from blood and other body specimens.

(Domain description updated in v4.0)

# 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 (many-to-one relationship) LAB\_RESULT\_CM.ENCOUNTERID is a foreign key to ENCOUNTER.ENCOUNTERID (zero/many-to-one relationship)

# Constraints:

LAB\_RESULT\_CM\_ID (unique; required, not null) PATID (required, not null)

RESULT\_DATE (required, not null)

|  |
| --- |
| **LAB\_RESULT\_CM Table Implementation Guidance** |
| *Guidance* |
| * Only records with actual lab results should be included in this table. If the result suggests that the test was run (e.g., result is "borderline" or "inconclusive") include it. But if the test is not resulted for any reason (specimen not sufficient, patient did not show), then do not include it. * If lab results are stored using local or custom codes, partners should ensure that the assigned LOINC code has been validated by a subject matter expert or similar process * If a LOINC code is available for a given result, the LAB\_LOINC field should be populated. If a LOINC code is available for the *order*, that value can be used to populate the LAB\_PX field. Note that one order can correspond to many different results. Each result should have its own record in the LAB\_RESULT\_CM table. If the same LOINC code is used to populate both the order and the result, partners should ensure that the LAB\_LOINC field is populated. * **Inclusion of additional lab results -** Partners should include *all* available laboratory results within their LAB\_RESULT\_CM table. If the result has a *validated* LOINC code, the LAB\_LOINC field should be populated. Otherwise, the LAB\_LOINC field should be blank. The RAW\_LAB\_NAME field can be used to keep track of the various lab results until the appropriate LOINC code is assigned. Lab results beyond the 11 originally included in the PCORnet CDM are being requested in order to establish a denominator of potentially available lab results. Over time, the number of unmapped results is expected to decrease. Results for labs performed as a service for outside institutions do not need to be included. Results from external vendors (e.g., LabCorp, Quest) should be included when available. * **Clinical LOINC Concepts** – Only include Laboratory LOINC concepts in this table. Do not include clinical LOINC concepts (e.g., EKG results). These records may be stored in the OBS\_CLIN table. * **Standing orders** - Partners should populate the date fields to the best of their ability. For results that are tied to standing laboratory orders, even if LAB\_ORDER\_DATE reflects the date of the original standing order, SPECIMEN\_DATE and/or RESULT\_DATE would be expected to correspond to the time when the sample was collected/resulted. Analyses will take both dates into consideration. * **Units of measure** – A given LOINC code may have many acceptable units of measure. If the RESULT\_UNIT field is not populated, it may not be possible to use a result analytically.    |

|  |  |  |
| --- | --- | --- |
| **Verifying LOINC mappings** – At most health systems, laboratory results are typically not associated with a LOINC code at the time they are generated but are assigned a code after the fact. In order | | |
| to verify that the LOINC code has been appropriately assigned, the PCORnet Coordinating Center will verify that the metadata associated with the result, such as SPECIMEN\_SOURCE and | |  |
| RESULT\_UNIT, are valid options. Partners should ensure that these fields are populated. **Do not derive these values based on metadata associated with the selected LOINC code.** |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| 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 | All PATIDs are expected to  be in the DEMOGRAPHIC table. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| ENCOUNTERID | RDBMS Text(x) | SAS Char(x) | . | Arbitrary encounter-level identifier.  Not all lab results will be associated with a healthcare encounter. | PCORnet (modeled upon VITAL table) | * Populate with the ENCOUNTERID where the lab specimen was collected (i.e., encounter when the lab test was administered). * All ENCOUNTERIDs in this table must also be present in the ENCOUNTER 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=Creatinin~~ ~~e~~  ~~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~~ | This field is deprecated effective v4.0. Partners should prioritize mapping their labs to LOINC. If the LOINC code for a given result is unknown, partners should populate the name of the lab in RAW\_LAB\_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* | *Field-level Implementation Guidance* |
| SPECIMEN\_SOURCE | RDBMS Text(x) | SAS Char(x) | See Value Set Appendix for a list of acceptable values. | Specimen source. All records will have a specimen source; some tests have several possible values for SPECIMEN\_SOURCE. | LOINC | Do not derive a specimen source based on the LOINC code. Please map the specimen source present in the source system to the appropriate code in the Value  Set Appendix. |
| 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 | * Use this field to store the LOINC code of the laboratory *result*. * Expected format of LOINC codes: Length of 3-7, hyphen in the penultimate position, no alphabetical characters. * Do not populate the LOINC field with dummy codes. If the LOINC code for a result is unknown, leave blank. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| LAB\_RESULT\_SOURCE | RDBMS Text(2) | SAS Char(2) | OD=Order/EHR | Source of the information for the lab | PCORnet | * This field is a derived attribute and is not expected to be an explicit data field within a source system * Use “OD” for lab results that are sourced from the EHR or laboratory information management system (LIMS). * Use “BI” for all lab results that are generated through the physician and hospital billing process (it is unlikely that this value will be used). * Use “CL” for laboratory results that are sourced from pharmacy or medical claims. * Use “DR” for all lab results that are derived or imputed through analytical procedures   (e.g., natural language processing). |
|  |  |  | BI=Billing | result. |  |
|  |  |  | CL=Claim |  |  |
|  |  |  | DR=Derived |  |  |
|  |  |  | NI=No information |  |  |
|  |  |  | UN=Unknown |  |  |
|  |  |  | OT=Other |  |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| LAB\_LOINC\_SOURCE | RDBMS Text(2) | SAS Char(2) | IN=Instrument LM=LIMS (Standalone or EHR)  HL=HL7 feed or other interface  DW=Data warehouse PC=PCORnet ETL  DM=Other CDM NI=No information UN=Unknown OT=Other | Source/provenance of the LOINC code for this result.  Details of categorical definitions: Instrument: Assigned by the instrument used to process the sample and generate the result.  Laboratory Information Management System (LIMS): Code is assigned by the system used to manage laboratory results, either stand-alone or as part of the EHR.  HL7 Feed: Code is assigned from an external HL7 feed or other interface (e.g., by a Health Information Exchange). | PCORnet | * This field is a derived attribute and is not expected to be an explicit data field within a source system. * If a LOINC code is present in laboratory results transmitted from a LIMS to the EHR via an HL7 feed, use “LM” or “IN” as appropriate, unless the LOINC code is assigned via the HL7 interface itself. Then use “HL” |
|  |  |  |  | Data warehouse: LOINC codes are generated during the extract-transform-load (ETL) procedures used to populate a standalone data warehouse or reporting database. |  |
|  |  |  |  | PCORnet ETL: LOINC codes are assigned as part of the ETL process used to populate the PCORnet CDM. |  |
|  |  |  |  | Other CDM: Codes are assigned through the ETL used to populate a CDM that is upstream from the PCORnet ETL. |  |
| PRIORITY | RDBMS Text(2) | SAS Char(2) | E=Expedite | Immediacy of test. The intent of this | MSCDM v4.0 with |  |
|  |  |  | R=Routine | variable is to determine whether the | modified value set |
|  |  |  | S=Stat | test was obtained as part of routine | and modified field |
|  |  |  | NI=No information | care or as an emergent/urgent | name |
|  |  |  | UN=Unknown | diagnostic test (designated as Stat or |  |
|  |  |  | OT=Other | Expedite). |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| 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) | . | Variable for local and standard procedure codes, used to identify the originating order for the lab test. | MSCDM v4.0 with modified field name | Can be used to store the procedure code of the laboratory *order*. If the same LOINC procedure code is used to identify both the order and the result, make sure  LAB\_LOINC is populated. |
| LAB\_PX\_TYPE | RDBMS Text(2) | SAS Char(2) | 09=ICD-9-CM  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. | MSCDM v4.0 with modified field name and value set | * CPT and HCPCS codes should be assigned a value of “CH.” * This field may be a derived attribute. In these situations, it is not expected to be an explicit data field within a source system |
| 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 |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| 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 |  |
| RESULT\_DATE | RDBMS Date | SAS Date (Numeric) |  | Result date. | MSCDM v4.0 with modified field name | If RESULT\_DATE is  unavailable, partners should use the date that is the closest match in their source data.  Partners are permitted to populate RESULT\_DATE with the value from SPECIMEN\_DATE if  necessary and note this in their ETL ADD. |
| 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(x) | SAS Char(x) | See Value Set Appendix for a list of acceptable values. | Standardized result for qualitative results. This variable should be NI for quantitative results. | LOINC | If qualitative result cannot be harmonized to a value in RESULT\_QUAL value set, please ensure that  RAW\_RESULT is populated with result value. |
| RESULT\_SNOMED | RDBMS Text(x) | SAS Char(x) | . | If the qualitative result has been mapped to SNOMED CT, the corresponding SNOMED code can  be placed here. | PCORnet | Do not impute or derive. Populate only if captured in the source system as a discrete value. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| RESULT\_NUM | RDBMS  Number(x) | SAS  Numeric(length 8) | . | Standardized/converted result for quantitative results. | MSCDM v4.0 with modified field name | Used to store quantitative results, including the numeric component of numeric results that contain operators (e.g., “<200”, “>= 0.5”). See  guidance for RESULT\_MODIFIER for  further details. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| 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 | * For quantitative results, a non-null RESULT\_MODIFER must be present if they data are to be used analytically. * Any symbols in the RAW\_RESULT value should be reflected in the   RESULT\_MODIFIER |
|  |  |  |  |  |  | variable. |
|  |  |  |  |  |  | For example, if the |
|  |  |  |  |  |  | original source data value |
|  |  |  |  |  |  | is "<=200" then |
|  |  |  |  |  |  | RAW\_RESULT is |
|  |  |  |  |  |  | “<=200” and |
|  |  |  |  |  |  | RESULT\_MODIFIER is |
|  |  |  |  |  |  | LE. RESULT\_NUM |
|  |  |  |  |  |  | would also be set to “200”. |
|  |  |  |  |  |  | If the original source data |
|  |  |  |  |  |  | value is text, then |
|  |  |  |  |  |  | RESULT\_MODIFIER=T |
|  |  |  |  |  |  | X If the original source |
|  |  |  |  |  |  | data value is a numeric |
|  |  |  |  |  |  | value, then |
|  |  |  |  |  |  | RESULT\_MODIFIER=E |
|  |  |  |  |  |  | Q |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| RESULT\_UNIT | RDBMS Text(x) | SAS Char(x) | See Value Set Appendix for a list of acceptable values. | Converted/standardized units for the quantitative result. | UCUM | * Chose the standardized unit of measure that is most reflective of the source data. Do not derive based on the LOINC code. * A given LOINC code may have many acceptable units of measure (e.g., tests reporting results as mass concentration can utilize mg/dL, g/dL, or similar). It is important that this field be populated in order to use the data analytically.   ~~ The Value Set Appendix~~ ~~contains a list of the units~~ ~~associated with the most~~ ~~common laboratory~~ ~~results (based on the Top~~ ~~2000 LOINC codes).~~  ~~Partners can use this table~~ ~~to aid in their mapping~~ ~~efforts, but they should~~ ~~refer back to the full value set if they have a~~ ~~laboratory result with a~~ ~~unit of measure that is not~~ ~~present in this curated list.~~ (Guidance deprecated as  of CDM v5.0) |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
|  |  |  |  |  |  | * This is a mixed case value set and entries should be handled accordingly. |
| 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)   1. 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* | *Field-level Implementation Guidance* |
| 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)   1. 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* | *Field-level Implementation Guidance* |
| 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. If field is blank in source data, map to the appropriate flavor of null (guidance added in v4.0). | 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 |  |
| 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). | 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 |  |

|  |
| --- |
| **Implementation Guidance Reference Table 1: Laboratory Results & LOINC Codes** |

This table has been deprecated as of version 2.0 of the PCORnet Implementation Guidance.

|  |
| --- |
| **Implementation Guidance Reference Table 2: Laboratory Results and CPT Codes** |

This table has been deprecated as of version 2.0 of the PCORnet Implementation Guidance

|  |
| --- |
| **Implementation Guidance Reference Table 3: Laboratory Standard Abbreviations** |

This table has been deprecated as of version 4.0 of the PCORnet CDM.

|  |
| --- |
| **5.9. 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 (many-to-one relationship) CONDITION.ENCOUNTERID is a foreign key to ENCOUNTER.ENCOUNTERID (zero/many-to-one 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 Implementation Guidance** |
| *Guidance* |
| * This table includes both healthcare and non-healthcare settings. * Rollback or voided transactions and other adjustments should be processed (removed) before populating this 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* | *Field-level Implementation Guidance* |
| CONDITIONID | RDBMS  Text(x) | SAS Char(x) | . | Arbitrary identifier for each unique  record. | PCORnet |  |
| PATID | RDBMS  Text(x) | SAS Char(x) | . | Arbitrary person-level identifier. Used  to link across tables. | MSCDM v4.0 | All PATIDs in this table must be present in the DEMOGRAPHIC table. |
| ENCOUNTERID | RDBMS  Text(x) | SAS Char(x) | . | Arbitrary encounter-level identifier used to link across tables. This 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) | * 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. * All ENCOUNTERIDs in this table must also be present in the ENCOUNTER 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) | 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. |
| 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* | *Field-level Implementation Guidance* |
| 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. | PCORnet | A value should only be provided where it exists in the source data. It is not calculated. |
| 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) | The value of IN=Inactive may be used in situations where a condition is not resolved, but is not currently active (for example, psoriasis). |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| 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. | PCORnet (modeled upon DIAGNOSIS table) |  |
| CONDITION\_TYPE | RDBMS  Text(2) | SAS Char(2) | 09=ICD-9-CM  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. | PCORnet (modeled upon DIAGNOSIS table) | * This field is a derived attribute and is not expected to be an explicit data field within a source system. * Do not include ICD procedure codes. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| 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  DR=Derived 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) | * “Registry cohort” generally refers to cohorts of patients flagged with a certain set of characteristics for management within a health system. * “Patient-reported” can include self-reported medical history and/or current medical conditions, not captured via healthcare problem lists or registry cohorts. * Use “DR” for all conditions that are derived or imputed through analytical procedures (e.g., natural language processing). * This field is a derived attribute and is not expected to be an explicit data field within a source system |
| RAW\_CONDITION\_STATUS | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, prior to mapping into the PCORnet CDM value set. | PCORnet |  |
| RAW\_CONDITION | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, prior to mapping into the PCORnet CDM value  set. | PCORnet |  |
| RAW\_CONDITION\_TYPE | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, prior to mapping into the PCORnet CDM value  set. | PCORnet |  |
| RAW\_CONDITION\_SOURCE | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, prior to  mapping into the PCORnet CDM value set. | PCORnet |  |

|  |
| --- |
| **5.10.** **Table: PRO\_CM** |

**PRO\_CM Domain Description:**

This table is used to store responses to patient-reported outcome measures (PROs) or questionnaires. This table can be used to store item-level responses as well as the overall score for each measure.

(Domain description updated in v4.0)

# Relational Integrity:

The PRO\_CM 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 (many-to-one relationship) PRO\_CM.ENCOUNTERID is a foreign key to ENCOUNTER.ENCOUNTERID (zero/many-to-one relationship)

# Constraints:

PRO\_CM\_ID (unique; required, not null) PATID (required, not null)

PRO\_DATE (required, not null)

|  |
| --- |
| **PRO\_CM Table Implementation Guidance** |
| *Guidance* |
| * This version of the PRO\_CM table is not optimized for representational efficiency. Certain values will be duplicated across records, and many fields will be blank for certain records. Over time, the structure of this table is expected to evolve as PCORnet better defines the analytical use of PROs across the network. Until then, this table has been defined to support the broadest range of possible use cases at the expense of representational efficiency. * The PRO\_CM table can be used to store both individual item-level responses, as well as the overall score for the measure/instrument. Each item response will be stored in an individual record. Measure/instrument scores can be stored along with the item-level responses that are associated with that measure (where available). See the figure below for an example of how to populate this table. * If partners are populating PRO item responses or measure scores and are unsure of the PRO\_ITEM\_NAME, PRO\_ITEM\_LOINC, PRO\_MEASURE\_NAME and/or PRO\_MEASURE\_LOINC, they should populate PRO\_ITEM\_FULLNAME and PRO\_MEASURE\_FULLNAME instead. These fields can be considered analogous to RAW fields. * For the PRO\_CM fields with variable field lengths, partners should choose an appropriate field length based on the characteristics of the data are loading into the table. As we use these tables analytically as part of PCORnet studies, we will determine whether it is more efficient to define specific field lengths. * If a patient completes a survey, but skips a question, create a record in the PRO\_CM table as you would for other items in the survey (i.e., include the appropriate date/time fields and other relevant metadata). Then leave PRO\_RESPONSE\_TEXT and PRO\_RESPONSE\_NUM blank, as these fields are not required. Do not create empty records if the patient did not actually see the question. * The PRO\_CM table can be used to store the results from questionnaires where the provider or caregiver are providing their interpretation or assessment of the patient’s status. Despite the name, it is not restricted solely to patient-reported outcomes. The table is designed to represent survey-type responses. General observations about patients, however, like pain scores recorded in an inpatient or surgical setting, should be stored in the OBS\_CLIN table. See General Guidance #15 for additional details or contact the DRN OC with questions. * For responses that can be represented as either a number or a text (e.g., a numeric value of 2 corresponds to “rarely” in a value set), the expectation is to store whatever is recorded in the source system. Both should be populated if present, but partners are not expected to derive the other if not. If values are represented based on a sequence number, store the actual value, not the sequence number. |

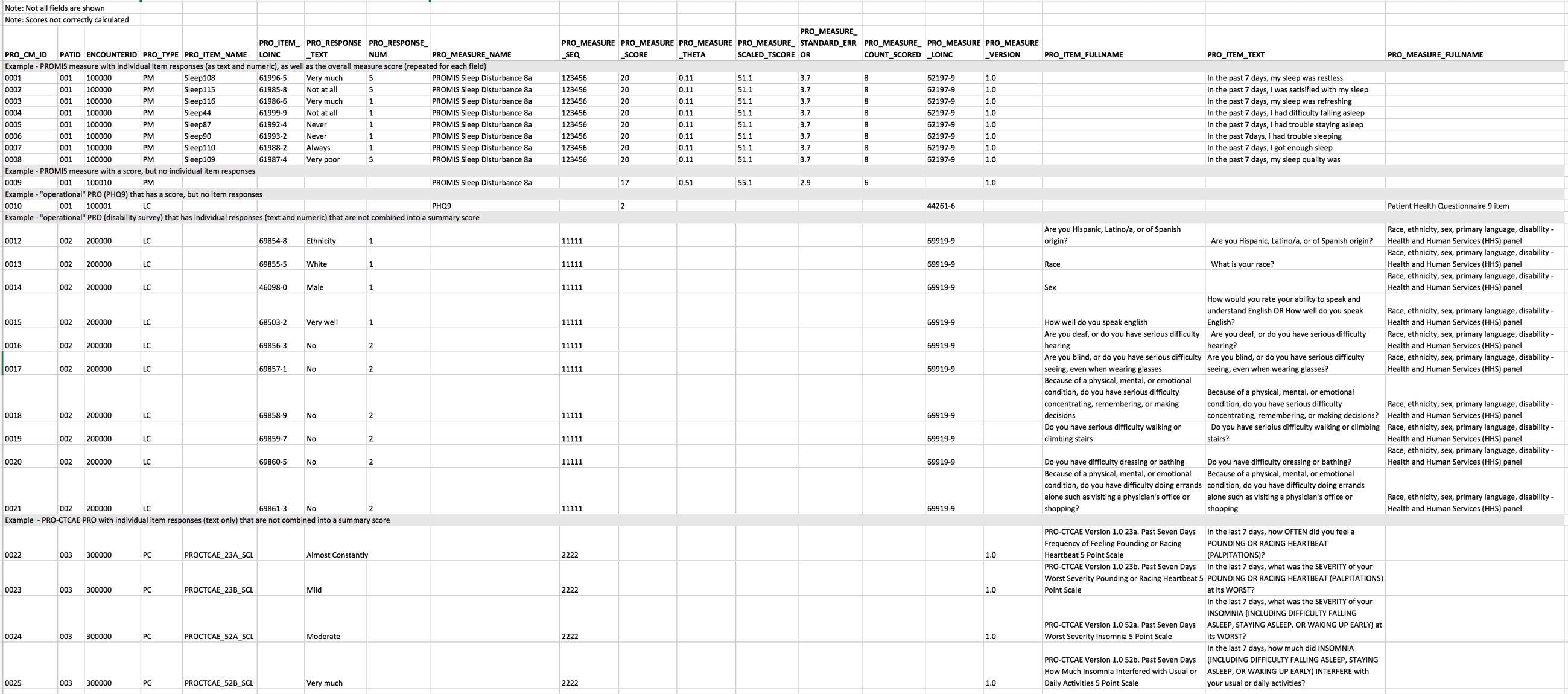


Figure: Example of a populated PRO\_CM table (note: not all required fields are shown).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| PRO\_CM\_ID | RDBMS Text(x) | SAS Char(x) | . | Arbitrary identifier for each unique  record. | 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* | *Field-level Implementation Guidance* |
| 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 | All PATIDs in this table must be present in the DEMOGRAPHIC table. |
| ENCOUNTERID | RDBMS Text(x) | SAS Char(x) | . | Arbitrary encounter-level identifier used to link across tables. This should only be populated if the item was collected as  part of a healthcare encounter. | PCORnet (modeled upon VITAL table) | All ENCOUNTERIDs in this table must be present in the ENCOUNTER table. |
| PRO\_DATE | RDBMS Date | SAS Date  (Numeric) | . | The date of the response submission. | 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 submission. | PCORnet  Source of time format: ISO 8601 |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| ~~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\_C~~ ~~APS PN\_0009=PAININ9~~ ~~PN\_0010=3793R1~~ ~~PN\_0011=28676R1~~ ~~PN\_0012=EOS\_P\_011~~ ~~PN\_0013=PEDSGLOBA~~ ~~L2~~ ~~PN\_0014=PEDSGLOBA~~ ~~L5~~ ~~PN\_0015=PEDSGLOBA~~ ~~L6~~ ~~PN\_0016=GLOBAL03~~ ~~PN\_0017=GLOBAL04~~ ~~PN\_0018=EDANX53~~ ~~PN\_0019=SAMHSA~~ ~~PN\_0020=CAHPS 4.0~~ ~~PN\_0021=PA070~~  ~~NI=No information~~  ~~UN=Unknown~~ ~~OT=Other~~ | ~~PCORnet identifier for the specific Common~~ ~~Measure item. Please see the Common Measures~~ ~~Reference Table for more details.~~ | ~~PCORnet~~ | ~~Non-PCORnet Common~~ ~~Measure PROs may also be~~ ~~stored in this table. These~~ ~~measures should be labeled~~ ~~with a value of “OT”.~~  This field has been deprecated as of CDM v4.0. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| PRO\_TYPE | RDBMS Text(2) | SAS Char(2) | PM=PROMIS  NQ=Neuro-QoL AM=ASQC-Me  NT=NIH Toolbox PC=PRO\_CTCAE LC=LOINC HC=HCAHPS  NI=No information UN=Unknown OT=Other | Terminology / vocabulary used to describe the PRO item.  More information on PROMIS, Neuro- QoL and ASQC-Me and the NIH Toolbox can be found on the HealthMeasures website. (www.healthmeasures.net)  The Patient-Reported Outcome version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE™) is maintained by the National Cancer Institute. (<https://healthcaredelivery.cancer.gov/pro-ctcae/)> | PCORnet | * For items/measures that belong to one of the listed terminologies and can also be found in LOINC (e.g., PROMIS), list the native terminology for PRO\_TYPE. * A value of “LOINC” should be used for those items/measures that do not belong to one of the other specified terminologies but can be found in LOINC (e.g., PHQ-9, WHO-5). * Information on PRO- CTCAE can also be found in the NCI Common Data Element browser (https://cdebrowser.nci.nih.gov) |
|  |  |  |  | Information on the Hospital Consumer Assessment of Healthcare Providers and Systems (HPCAHPS) is located here: [http://www.hcahpsonline.org](http://www.hcahpsonline.org/) |  |
| PRO\_ITEM\_NAME | RDBMS Text(x) | SAS Char(x) | . | Short name or code of the PRO item in | PCORnet | If a short name or code for |
|  |  |  |  | the vocabulary/terminology specified in |  | the PRO item does not exist |
|  |  |  |  | PRO\_TYPE. |  | within the specified |
|  |  |  |  |  |  | terminology, do not create |
|  |  |  |  |  |  | one. Populate |
|  |  |  |  |  |  | PRO\_ITEM\_FULLNAME |
|  |  |  |  |  |  | instead. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| PRO\_ITEM\_LOINC | RDBMS  Text(10) | SAS Char(10) | . | LOINC® code for the PRO item, if available.  Logical Observation Identifiers, Names, and Codes (LOINC) from the Regenstrief Institute. 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\_RESPONSE\_TEXT | RDBMS Text(x) | SAS Char(x) | . | Text version of the response recorded for  the item, if available/applicable. | PCORnet |  |
| PRO\_RESPONSE\_NUM | RDBMS  Number(x) | SAS  Numeric(length 8) | . | The numeric response recorded for the item, if available/applicable. | PCORnet |  |
| 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\_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* | *Field-level Implementation Guidance* |
| PRO\_MODE | RDBMS Text(2) | SAS Char(2) | SF=Self without | The person who responded on behalf of | PCORnet |  |
|  |  |  | assistance | the patient for whom the response was |  |
|  |  |  | SA= Self with | captured. A proxy report is a |  |
|  |  |  | assistance | measurement based on a report by |  |
|  |  |  | PR=Proxy without | someone other than the patient reporting |  |
|  |  |  | assistance | as if he or she is the patient, such as a |  |
|  |  |  | PA=Proxy with | parent responding for a child, or a |  |
|  |  |  | assistance | caregiver responding for an individual |  |
|  |  |  | NI=No information | unable to report for themselves. |  |
|  |  |  | UN=Unknown | Assistance excludes providing |  |
|  |  |  | OT=Other | interpretation of the patient’s response. |  |
| PRO\_CAT | RDBMS Text(2) | SAS Char(2) | Y=Yes | Indicates whether Computer Adaptive | PCORnet |  |
|  |  |  | N=No | Testing (CAT) was used to administer |  |
|  |  |  | NI=No information | the survey or instrument that the item |  |
|  |  |  | UN=Unknown | was part of. May apply to electronic (EC) |  |
|  |  |  | OT=Other | and telephonic (PH or IV) modes. |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| PRO\_SOURCE | RDBMS Text(2) | SAS Char(2) | OD=Order/EHR | Source of the information for the PRO | PCORnet | * This field is a derived attribute and is not expected to be an explicit data field within a source system * Use “OD” for PRO records that are sourced from the EHR, its patient portal, or any other system used to capture PROs as part of the care process. * Use “BI” for all PRO records that are generated through the physician and hospital billing process (it is unlikely that this value will be used). * Use “SR” for PRO records that are generated from external survey systems or mobile apps as part of a non-care process. * Use “CL” for PRO records that are sourced from pharmacy or medical claims. * Use “DR” for all PRO records that are derived or imputed through analytical procedures (e.g., natural   language processing). |
|  |  |  | BI=Billing | result. |  |
|  |  |  | CL=Claim |  |  |
|  |  |  | SR=Survey |  |  |
|  |  |  | system/mobile app |  |  |
|  |  |  | DR=Derived |  |  |
|  |  |  | NI=No information |  |  |
|  |  |  | UN=Unknown |  |  |
|  |  |  | OT=Other |  |  |
| PRO\_ITEM\_VERSION | RDBMS Text(x) | SAS Char(x) | . | Version of the item/question. | 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* | *Field-level Implementation Guidance* |
| PRO\_MEASURE\_NAME | RBDMS Text(x) | SAS Char(x) | . | Short name or code of the PRO measure/form that item belongs to, if item is being administered as part of a measure | PCORnet | * Will be blank if item is not being administered as part of a measure/form * If measure does not have a short name or code within the specified PRO terminology, do not create one. Populate PRO\_MEASURE\_FULLN AME instead. * If item is part of a PRO measure, the value for this field will be replicated for all items in the measure. |
| PRO\_MEASURE\_SEQ | RDBMS Text(x) | SAS Char(x) | . | Arbitrary ID/sequence number used to link PRO item responses that are associated with the same measure/form. | PCORnet | * All PRO item responses associated with the same form/measure should have the same value for PRO\_MEASURE\_SEQ. * Will be blank if item is not part of a PRO measure/form. |
| PRO\_MEASURE\_SCORE | RDBMS  Number(x) | SAS  Numeric(length 8) | . | Overall raw score for the PRO measure. | PCORnet | * Will be blank if item is not part of a PRO measure/form. * If item is part of a PRO measure, the value for this field will be replicated for all items in the measure. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| PRO\_MEASURE\_THETA | RDBMS  Number(x) | SAS  Numeric(length 8) | . | The value of theta reported from the CAT PROMIS results. Only applies to items that are administered as part of a measure. | PCORnet | * Expected when scoring any measure/instrument that uses computer-adaptive testing (CAT) that has been calibrated with item- response theory (IRT), but may not be present for PRO measures/instruments that are not using CAT. * Will be blank if item is not part of a PRO measure/form. * If item is part of a PRO measure, the value for this field will be replicated for all items in the measure. |
| PRO\_MEASURE\_SCALED  \_TSCORE | RDBMS  Number(x) | SAS  Numeric(length 8) | . | Standardized score based on the total raw score for the instrument. Only applies to items that are administered as part of a measure. | PCORnet | * Expected when scoring any measure/instrument that uses CAT that has been calibrated with IRT, but may not be present for PRO measures/instruments that are not using CAT. * Will be blank if item is not part of a PRO measure/form. * If item is part of a PRO measure, the value for this field will be replicated for all items in the measure. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| PRO\_MEASURE\_STAND ARD\_ERROR | RDBMS  Number(x) | SAS  Numeric(length 8) | . | Possible range of the actual final score based on the scaled T-score. Only applies to items that are administered as part of a measure. | PCORnet | * Expected when scoring any measure/instrument that uses CAT that has been calibrated with IRT, but may not be present for PRO measures/instruments that are not using CAT. * Will be blank if item is not part of a PRO measure/form. * If item is part of a PRO measure, the value for this field will be replicated for all items in the measure. |
| PRO\_MEASURE\_COUNT  \_SCORED | RDBMS  Number(x) | SAS Numeric (length 8) | . | Number of PRO item responses that were involved in the scoring of the measure. | PCORnet | * Expected when scoring any measure/instrument that uses CAT that has been calibrated with IRT, but may not be present for PRO measures/instruments that are not using CAT. * Will be blank if item is not part of a PRO measure/form. * If item is part of a PRO measure, the value for this field will be replicated for all items in the measure. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| PRO\_MEASURE\_LOINC | RDBMS  Text(10) | SAS Char(10) | . | LOINC® code for the PRO item, if available.  Logical Observation Identifiers, Names, and Codes (LOINC) from the Regenstrief Institute. 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 | * Will be blank if item is not part of a PRO measure/form. * If item is part of a PRO measure, the value for this field will be replicated for all items in the measure. |
| PRO\_MEASURE\_VERSIO N | RDBMS Text(x) | SAS Char(x) | . | Version of the measure. | PCORnet | * Will be blank if item is not part of a PRO measure/form. * If item is part of a PRO measure, the value for this field will be replicated for all items in the measure. |
| PRO\_ITEM\_FULLNAME | RDBMS Text(x) | SAS Char(x) | . | Full name of the PRO item. | PCORnet |  |
| PRO\_ITEM\_TEXT | RDBMS Text(x) | SAS Char(x) | . | Text of the PRO item question. | PCORnet |  |
| PRO\_MEASURE\_FULLN AME | RDBMS Text(x) | SAS Char(x) | . | Full name of the PRO measure. | PCORnet | * Will be blank if item is not part of a PRO measure/form. * If item is part of a PRO measure, the value for this field will be replicated for all items in the measure. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| ~~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~~ | This field is deprecated as of CDM v4.0 |
| ~~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~~ | This field is deprecated as of CDM v4.0 |

|  |
| --- |
| **CDM Reference Table: PRO Common Measures** |

This table has been deprecated as of version 4.0 of the PCORnet CDM.

|  |
| --- |
| **5.11.** **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.

# Relational Integrity:

The PRESCRIBING table contains one record per PRESCRIBINGID.

**Primary Key:** PRESCRIBINGID

# Foreign Keys:

PRESCRIBING.PATID is a foreign key to DEMOGRAPHIC.PATID (many-to-one relationship) PRESCRIBING.ENCOUNTERID is a foreign key to ENCOUNTER.ENCOUNTERID (zero/many-to-one relationship) PRESCRIBING.RX\_PROVIDERID is a foreign key to PROVIDER.PROVIDERID (many-to-one relationship)

# Constraints:

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

|  |
| --- |
| **PRESCRIBING Table Implementation Guidance** |
| *Guidance* |
| * If a medication cannot be mapped to RxNorm, it should still be present and RAW\_RX\_MED\_NAME should be populated. * This table can be used to store all medication orders, regardless of encounter type (e.g., inpatient, outpatient, ED) and can include orders for medications that are to be dispensed as well as for those that are to be administered. * If including orders derived through natural language processing (NLP), make sure that RX\_SOURCE has been populated for *all* records. * See Reference Table 4 for the ordering strategy for RxNorm Term Types. * Do not populate CDM fields with information derived from the RXCUI (e.g., RX\_DOSE\_ORDERED, RX\_DOSE\_FORM). Populate fields only if data are captured in the source system as a discrete value. * Populate records with the RXCUI as it existed at the time the order was entered, even if the RXCUI is no longer active. Do not attempt to update inactive RXCUIs with a more recent value. * ~~If a medication mixture contains multiple RXCUIs (e.g., inpatient mixture),~~ Medications with approved formulations should have an RXCUI that can adequately represent all ingredients with a single code (e.g., SBD, SCD, MIN). For medication mixtures that lack RXCUIs that can represent all of the component ingredients (e.g., IV mixtures prepared at an inpatient or compounding pharmacy), each individual medication from the order set should be included as a separate record with a unique PRESCRIBINGID. If partners wish to preserve the fact that the records belong to the same order, they do so by creating and populating a new *optional* ORDERID field. Medications with a 1:1 correspondence between the order and RXCUI could have the PRESCRIBINGID stored in the ORDERID. Orders with a 1:many RXCUI relationship would have different PRESCRIBINGIDs but the same ORDERID. Future versions of the CDM may formalize this guidance. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| PRESCRIBINGID | RDBMS  Text(x) | SAS Char(x) | . | Arbitrary identifier for each unique  PRESCRIBING record. | PCORnet |  |
| PATID | RDBMS  Text(x) | SAS Char(x) | . | Arbitrary person-level identifier used to link across tables. | MSCDM v4.0 | All PATIDs must be present in the DEMOGRAPHIC table. |
| ENCOUNTERID | RDBMS  Text(x) | SAS Char(x) | . | Arbitrary encounter-level identifier. This should be present if the prescribing activity is directly associated with an  encounter. | MSCDM v4.0 | All ENCOUNTERIDs in this table must be present in the ENCOUNTER table. |
| 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 | All PROVIDERIDs in this table must be present in the PROVIDER table. |
| RX\_ORDER\_DATE | RDBMS Date | SAS Date  (Numeric) | . | Order date of the prescription by the  provider. | MSCDM v4.0 | If RX\_ORDER\_DATE is not known, a permissible substitution is |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
|  |  |  |  |  |  | the RX\_START\_DATE (if known)  or the date the order was signed by the provider. |
| 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\_DOSE\_ORDERED | RDBMS  Number(x) | SAS  Numeric(lengt h 8) | . | Dose of a given mediation, as ordered by the provider | PCORnet | Do not impute or derive. Populate only if captured in the source system as a discrete value. |
| RX\_DOSE\_ORDERED\_ UNIT | RDBMS  Text(x) | SAS Char(x) | See Value Set Appendix for a list of acceptable values. | Units of measure associated with the dose of the medication as ordered by the provider | UCUM | * Do not impute or derive. Populate only if captured in the source system as a discrete value. * Choose the standardized unit of measure that is most reflective of the source data. * ~~The Value Set Appendix contains a list of the units most commonly associated with medication records. Partners can~~   ~~use this table to aid in their~~ |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
|  |  |  |  |  |  | ~~mapping efforts, but they should~~ ~~refer back to the full value set if~~ ~~they have a medication record~~ ~~with a unit of measure that is not~~ ~~present in this curated list~~. (Guidance deprecated as of CDM v5.0)   * This is a mixed case value set   and entries should be handled accordingly. |
| RX\_QUANTITY | RDBMS  Number(x) | SAS  Numeric(lengt h 8) | . | Quantity ordered. | Based on OMOP and ESP | Do not impute or derive. Populate only if captured in the source system as a discrete value. |
| RX\_DOSE\_FORM | RDBMS  Text(x) | SAS Char(x) | See Value Set Appendix for a list of acceptable values. | The unit associated with the quantity prescribed. This is equivalent to RxNorm Dose Form. | PCORnet, based on RxNorm attributes | Do not impute or derive. Populate only if captured in the source system as a discrete value. |
| RX\_REFILLS | RDBMS  Number(x) | SAS  Numeric(lengt h 8) | . | Number of refills ordered (not including the original prescription). If no refills are ordered, the value should be zero. | Based on OMOP and ESP | * If value is non-numeric, leave field blank and populate RAW\_RX\_REFILLS with originating source value. * Do not impute or derive. Populate only if captured in the source system as a discrete value. |
| RX\_DAYS\_SUPPLY | RDBMS  Number(x) | SAS  Numeric(lengt h 8) | . | Number of days supply ordered, as specified by the prescription. | Based on OMOP | Do not impute or derive. Populate only if captured in the source system as a discrete value. |
| RX\_FREQUENCY | RDBMS  Text(2) | SAS Char(2) | 01=Every day 02=Two times a day  (BID) | Specified frequency of medication. | PCORnet | Do not impute or derive. Populate only if captured in the source system as a discrete value. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
|  |  |  | 03=Three times a day (TID)  04=Four times a day (QID)  05=Every morning 06=Every afternoon 07=Before meals 08=After meals 10=Every evening 11=Once  NI=No information UN=Unknown OT=Other |  |  |  |
| RX\_PRN\_FLAG | RDBMS  Text(1) | SAS Char(1) | Y=Yes N=No | Flag to indicate that all or part of medication frequency instructions includes “as needed.” | PCORnet | * Select Y if medication order includes instructions to take medication “as needed” or with any other frequency “as needed” (e.g., Two times a day, as needed). * This field is a derived attribute and is not expected to be an explicit data field within a source system |
| RX\_ROUTE | RDBMS  Text(x) | SAS Char(x) | See Value Set Appendix for a list of acceptable values. | Route of medication delivery. | RxNorm (SNOMED) | * The value set for Route is derived from SNOMED and may include values that are more granular than what is present in the source system. If a direct mapping is available, use the appropriate SNOMED code. If there is any possible ambiguity, use “OT” and   then store the source value in |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
|  |  |  |  |  |  | RAW\_RX\_ROUTE. For  example, an Injection could map to Subcutaneous or Intramuscular or Intraocular, depending on the drug, so that would best be mapped to “OT.”   * Do not impute or derive. Populate only if captured in the source system as a discrete value. |
| RX\_BASIS | RDBMS | SAS Char(2) | 01=Order to | Basis of the medication order. The | PCORnet | This field is a derived attribute and is |
| Text(2) |  | Dispense 02=Order to | PRESCRIBING table can contain orders for many different activities, and this |  | not expected to be an explicit data  field within a source system |
|  |  |  | administer | field is intended to connect the |  |  |
|  |  |  | NI=No information | provider’s prescribing order with how |  |  |
|  |  |  | UN=Unknown | the order was fulfilled (such as |  |  |
|  |  |  | OT=Other | outpatient dispensing or administration |  |  |
|  |  |  |  | by a healthcare professional). (Value set |  |  |
|  |  |  |  | items updated and field definition |  |  |
|  |  |  |  | expanded in v3.1.) |  |  |
| 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. | PCORnet and NLM UMLS | * **Ordering Strategy**– The ordering strategy for RxNorm Term Types has been updated to indicate a preference of brand name CUIs over generics, when available. It has also been expanded to include addition RxNorm Term Types. Please see Reference Table 4 for more information. * Do not assign a RxNorm Term Type that is not supported by the   source data (i.e., if the |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
|  |  |  |  |  |  | medication order in source system is missing strength information, do not use any of the RxNorm Term Types that incorporate the strength component).   * This field may be a derived attribute and is not necessarily expected to be an explicit data field within a source system * Expected format of RXNORM\_CUI codes: Length of 2-7, no alphabetical characters. * Do not assign more than one RxCUI per order. Ensure that a single prescribing record is not assigned multiple RxCUIs. |
| RX\_SOURCE | RDBMS  Text(2) | SAS Char(2) | OD=Order/EHR DR=Derived NI=No information UN=Unknown OT=Other | Source of the prescribing information. | PCORnet | * This field is a derived attribute and is not expected to be an explicit data field within a source system * Use “OD” for medication orders entered into the EHR or for electronic prescriptions. * Use “DR” for all medication orders that are derived or imputed through analytical procedures (e.g., natural language processing). This does not apply to medication orders   mapped from a superset |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
|  |  |  |  |  |  | terminology or drug database (e.g., MediSpan, FDB). For those records, use “OD” (General Guidance #4). |
| RX\_DISPENSE\_AS\_W RITTEN | RDBMS  Text(2) | SAS Char(2) | Y=Yes N=No  NI=No information UN=Unknown OT=Other | Flag to indicate whether the provider indicated that the medication order was to be dispensed as written. | PCORnet | * This information is typically captured within EHRs or e- prescribing as part of the ordering process. * Do not impute or derive. Populate only if captured in the source system as a discrete value. |
| RAW\_RX\_MED\_NAM  E | RDBMS  Text(x) | SAS Char(x) | . | Field for originating, full textual  medication name from the source. | PCORnet |  |
| RAW\_RX\_FREQUENC Y | RDBMS  Text(x) | SAS Char(x) |  | Field for originating value, prior to  mapping into the PCORnet CDM value set. | PCORnet |  |
| RAW\_RXNORM\_CUI | RDBMS  Text(x) | SAS Char(x) |  | Field for originating value, prior to mapping into the PCORnet CDM value  set. | PCORnet |  |
| RAW\_RX\_QUANTITY | RDBMS  Text(x) | SAS Char(x) |  | Field for originating value, prior to mapping into the PCORnet CDM value set. | PCORnet |  |
| RAW\_RX\_NDC | RDBMS  Text(x) | SAS Char(x) |  | Field for originating value, prior to mapping into the PCORnet CDM value  set. | PCORnet |  |
| RAW\_RX\_DOSE\_ORD ERED | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, prior to  mapping into the PCORnet CDM value set. | PCORnet |  |
| RAW\_RX\_DOSE\_ORD ERED\_UNIT | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, 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* | *Field-level Implementation Guidance* |
| RAW\_RX\_ROUTE | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, prior to mapping into the PCORnet CDM value  set. | PCORnet |  |
| RAW\_RX\_REFILLS | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, prior to  mapping into the PCORnet CDM value set. |  |  |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Implementation Guidance Reference Table 4: Ordering of RxNorm Term Types** | | | | | | | |
| **(Content from the UMLS [**[**https://www.nlm.nih.gov/research/umls/rxnorm/docs/2015/appendix5.html**](https://www.nlm.nih.gov/research/umls/rxnorm/docs/2015/appendix5.html)**] – Accessed October 2016)** | | | | | | | |
|  | **RxNorm Term Type** | | **Information incorporated** | | | |  |
|  | **Code** | **Description** | **Ingredient(s)** | **Strength** | **Dose Form** | **Brand Name** | **Notes** |
| *Most Preferred* | SBD | Semantic Branded Drug | X | X | X | X |  |
|  | SCD | Semantic Clinical Drug | X | X | X |  |  |
|  | BPCK | Brand Name Pack | X | X | X | X |  |
|  | GPCK | Generic Pack | X | X | X |  |  |
|  | SBDF | Semantic Branded Drug Form | X |  | X | X |  |
|  | SCDF | Semantic Clinical Drug Form | X |  | X |  |  |
|  | SBDG | Semantic Branded Dose Form Group |  |  | X | X |  |
|  | SCDG | Semantic Clinical Dose Form Group | X |  | X |  |  |
|  | SBDC | Semantic Branded Drug Component | X | X |  | X |  |
|  | BN | Brand Name |  |  |  | X |  |
|  | MIN | Multiple Ingredients | X |  |  |  |  |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | SCDC | Semantic Clinical Drug Component | X | X |  |  | May not be enough to distinguish medication for analysis purposes. If medication contains multiple ingredients, include a record in the PRESCRIBING table for each one. |
|  | PIN | Precise Ingredient | X |  |  |  |  |
| *Least Preferred* | IN | Ingredient | X |  |  |  | May not be enough to distinguish medication for analysis purposes. If medication contains multiple ingredients, include a record in the PRESCRIBING  table for each one. |
| *Do not use* | DF | Dose Form |  |  | X |  | Non-specific |
| *Do not use* | DFG | Dose Form Group |  |  | X |  | Non-specific |
| *Do not use* | PSN | Prescribable Name |  |  |  |  | Synonym of another TTY; Use original TTY |
| *Do not use* | SY | Synonym |  |  |  |  | Synonym of another TTY; Use original TTY |
| *Do not use* | TMSY | Tall Man Lettering Synonym |  |  |  |  | Synonym of another TTY; Use original TTY |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Implementation Guidance Reference Table 4a: RxNorm Term Types with examples** | | | | | | |
| **(obtained from RxNav [**[**https://mor.nlm.nih.gov/RxNav/**](https://mor.nlm.nih.gov/RxNav/)**] – Accessed October 2016)** | | | | | | |
|  | **RxNorm Term Type** | |  |  |  |  |
|  | **Code** | **Description** | **Example (Augmentin XR 12 HR 1000 MG Extended Oral Release Tablet)** | **RxCUI(s)** | **Example (Z-PAK)** | **RxCUI(s)** |
| *Most*  *Preferred* | SBD | Semantic Branded Drug | Augmentin XR 12 HR 1000 MG Extended  Release Oral Tablet | 861689 | Zithromax 250 MG Oral Tablet | 212446 |
|  | SCD | Semantic Clinical Drug | 12 HR Amoxicillin 1000 MG / Clavulanate 62.5  MG Extended Release Oral Tablet | 617995 | Axithromycin 250 MG Oral  Tablet | 308460 |
|  | BPCK | Brand Name Pack | N/A | N/A | Z-PAK | 750149 |
|  | GPCK | Generic Pack | N/A | N/A | {6 (Azithromycin 250 MG Oral  Tablet) } Pack | 749783 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | SBDF | Semantic Branded Drug  Form | Amoxicillin / Clavulanate Extended Release  Oral Tablet [Augmentin] | 618038 | Azithromycin Oral Tablet  [Zithromax] | 367697 |
|  | SCDF | Semantic Clinical Drug  Form | Amoxicillin / Clavulanate Extended Release  Oral Tablet | 617988 | Azithromycin Oral Tablet | 370976 |
|  | SBDG | Semantic Branded Dose  Form Group | Augmentin Oral Product;  Augmentin Pill | 1174397; 1174308 | Zithromax Oral Product;  Zithromax Pill | 1187674; 1187675 |
|  | SCDG | Semantic Clinical Dose Form Group | Amoxicillin / Clavulanate Oral Product; Amoxicillin / Clavulanate Pill | 1152874; 1152875 | Azithromycin Oral Product; Azithromycin Pill | 1155011; 1155012 |
|  | SBDC | Semantic Branded Drug  Component | Amoxicillin 1000 MG / Clavulanate 62.5 MG  [Augmentin] | 618037 | Azithromycin 250 MG  [Zithromax] | 564001 |
|  | BN | Brand Name | Augmentin | 151392 | Zithromax | 169474 |
|  | MIN | Multiple Ingredients | Amoxicillin / Clavulanate | 19711 | N/A | N/A |
|  | SCDC | Semantic Clinical Drug Component | Amoxicillin 1000 MG; Clavulanate 62.5 MG | 331055; 617303 | Azithromycin 250 MG | 315449 |
|  | PIN | Precise Ingredient | N/A | N/A | N/A | N/A |
| *Least Preferred* | IN | Ingredient | Amoxicillin; Clavulanate | 723; 48203 | Azithromycin | 18631 |
| *Do not use* | DF | Dose Form | Extended Release Oral Tablet | 316945 | Oral Capsule; Oral Tablet | 316965; 317541 |
| *Do not use* | DFG | Dose Form Group | Oral Product; Pill | 1151131; 1151133 | Oral Product; Pill | 1151131; 1151133 |
| *Do not use* | PSN | Prescribable Name |  |  |  |  |
| *Do not use* | SY | Synonym |  |  |  |  |
| *Do not use* | TMSY | Tall Man Lettering  Synonym |  |  |  |  |

|  |
| --- |
| **5.12.** **Table: PCORNET\_TRIAL** |

**PCORNET\_TRIAL Domain Description:**

Patients who are enrolled in PCORnet clinical trials and PCORnet studies.

# 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 (many-to-one relationship)

# 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:**



Associate the

CDM domains specified in the trial protocol

Associate the

study records

Work with CDM

data in the correct timeframe

TRIAL\_ENROLL\_DATE TRIAL\_END\_DATE TRIAL\_WITHDRAW\_DATE

TRIAL\_INVITE\_CODE

If used by trial

Which

person?

PARTICIPANTID

Which trial?

TRIALID

Which patient?

PATID

**PCORNET\_TRIAL**

**PCORnet Trial Database (eg, ADAPTABLE)**

May contain:

* Consent module
* Randomization assignment
* Study-specific data collection
* Study-specific schedule of assessments

**CDM Tables**

**(within a specific CDRN Datamart)**

|  |
| --- |
| **PCORNET\_TRIAL Table Implementation Guidance** |
| *Guidance* |
| * Partners may use the PCORNET\_TRIAL table to maintain mappings between PCORnet CDM PATIDs and external trial or study IDs. * Partners wishing to use this table will need to register their TRIALID with the DRN OC. Please contact the DRN OC if you plan to utilize this table. * TRIALIDs that start with “PT\_” and “PS\_” are reserved for PCORnet Trials and PCORnet Studies. Partners should refrain from using TRIALIDs that start with these characters. The TRIALID “ADPT” is also reserved. * One patient participating in multiple trials or studies will have multiple records * Each PCORnet trial or study will define its parameters for enrollment * Patients who decline to participate in a trial or study or do not meet eligibility criteria should not be included in this table * Patients who enroll in a trial or study but later withdraw should be included in this table so that their withdrawal status and date are recognized and used to appropriately manage reporting of CDM data back to the coordinating center * In most cases, trials will be expected to have a separate trial database that is separate from the CDM * Randomization assignment is not included in this table due to the potential for unblinding. * PATID is not generally appropriate for use as a PARTICIPANTID because it is not disambiguated across networks. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| PATID | RDBMS  Text(x) | SAS Char(x) | . | Arbitrary person-level identifier  used to link across tables. | MSCDM v4.0 | All PATIDs must be present in the DEMOGRAPHIC table. |
| TRIALID | RDBMS  Text(20) | SAS  Char(20) | . | Each TRIALID is assigned by the PCORnet trial or study’s coordinating center. | PCORnet |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| PARTICIPANTID | RDBMS  Text(x) | SAS Char(x) | . | Arbitrary person-level identifier used to uniquely identify a participant in a PCORnet trial or study.  PARTICIPANTID is never repeated or reused for a specific clinical trial or study, and is generally assigned by trial or study-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 or study coordinating center. | PCORnet | * This field is a derived attribute and is not expected to be an explicit data field within a source system * This field will typically represent the patient’s recruiting site for clinical trials. A given patient may have data in more than one PCORnet DataMart. |
| TRIAL\_ENROLL\_DATE | RDBMS  Date | SAS Date (Numeric) | . | Date on which the participant enrolled in the trial or study (generally coincides with trial or  study consent process). | PCORnet |  |
| TRIAL\_END\_DATE | RDBMS  Date | SAS Date (Numeric) |  | Date on which the participant completes participation in the trial  or study. |  |  |
| TRIAL\_WITHDRAW\_DATE | RDBMS  Date | SAS Date (Numeric) | . | If applicable, date on which the participant withdraws consent from  the trial or study. | PCORnet |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| 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 |  |

|  |
| --- |
| **5.13.** **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 (many-to-one relationship)

# Constraints (modified in v3.1)

PATID + DEATH\_SOURCE (unique)

PATID (required, not null) DEATH\_SOURCE (required, not null)

|  |
| --- |
| **DEATH Table Implementation Guidance** |
| *Guidance* |
| * One patient may potentially have multiple records in this table because their death may be reported by different sources. * Deaths represented in the ENCOUNTER.DISCHARGE\_DISPOSITION and ENCOUNTER.DISCHARGE\_STATUS would generally be expected to be present in this table (see field-level guidance for DEATH.DEATH\_SOURCE). |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| PATID | RDBMS  Text(x) | SAS Char(x) | . | Arbitrary person-level identifier used to link across  tables. | MSCDM v4.0 | All PATIDs in this table must be present in the DEMOGRAPHIC table. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| DEATH\_DATE | RDBMS  Date | SAS Date (Numeric) |  | Date of death. | MSCDM v4.0 with modified field name and  data type | If the death date is completely unknown (e.g., fully imputed), partners should leave it blank. |
| 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 | This field is a derived attribute and is not expected to be an explicit data field within a source system |
| 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  DR=Derived NI=No information UN=Unknown  OT=Other |  | MSCDM v4.0 with modified field name and additional guidance | * “Other, locally defined” may be used to indicate presence of deaths reported from EHR systems, such as in-patient hospital deaths or dead on arrival. * This field is a derived attribute and is not expected to be an explicit data field within a source system * Use “DR” for all death records that are derived or imputed through analytical procedures (e.g., natural language processing). |
| DEATH\_MATCH\_CONF IDENCE | 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 | * Should not be present where DEATH\_SOURCE is L (locally- defined). May not be applicable for DEATH\_SOURCE=T (tumor registry data). * This field is a derived attribute and is not expected to be an explicit data field within a source system |

|  |
| --- |
| **5.14.** **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 (many-to-one 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 Implementation Guidance** |
| *Guidance* |
| * When legacy data have conflicting reports, please make a local determination as to which to use. There is typically a 1-2 year lag in death registry data. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| PATID | RDBMS  Text(x) | SAS Char(x) | . | Arbitrary person-level identifier used to link across  tables. | MSCDM v4.0 | All PATIDs in this table must be present in the DEMOGRAPHIC table. |
| 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 | This field is a derived attribute and is not expected to be an explicit data field within a source system |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation Guidance* |
| 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  DR=Derived NI=No information UN=Unknown  OT=Other | Source of cause of death information. | MSCDM v4.0 with modified field name | * “Other, locally defined” may be used to indicate presence of deaths reported from EHR systems, such as in-patient hospital deaths or dead on arrival. * This field is a derived attribute and is not expected to be an explicit data field within a source system * Use “DR” for all death cause records that are derived or imputed through analytical procedures (e.g., natural language processing). |
| DEATH\_CAUSE\_CONFIDEN | RDBMS | SAS Char(2) | E=Excellent | Confidence in the accuracy of | MSCDM v4.0 with | This field is a derived attribute and is |
| CE | Text(2) |  | F=Fair P=Poor | the cause of death based on source, match, number of | modified field name | not expected to be an explicit data  field within a source system |
|  |  |  | NI=No information | reporting sources, |  |  |
|  |  |  | UN=Unknown | discrepancies, etc. |  |  |
|  |  |  | OT=Other |  |  |  |

|  |
| --- |
| **5.15. Table: MED\_ADMIN** |

**MED\_ADMIN Domain Description:**

Records of medications administered to patients by healthcare providers. These administrations may take place in any setting, including inpatient, outpatient or home health encounters.

# Relational Integrity:

The MED\_ADMIN table contains one record per MEDADMINID.

**Primary Key:** MEDADMINID

# Foreign Keys:

MEDADMIN.PATID is a foreign key to DEMOGRAPHIC.PATID (many-to-one relationship) MEDADMIN.ENCOUNTERID is a foreign key to ENCOUNTER.ENCOUNTERID (many-to-one relationship) MEDADMIN.MEDADMIN\_PROVIDERID is a foreign key to PROVIDER.PROVIDERID (many-to-one relationship)

# Constraints:

MEDADMINID (unique; required, not null) PATID (required, not null) MEDADMIN\_START\_DATE (required, not null)

|  |
| --- |
| **MED\_ADMIN Table Implementation Guidance** |
| *Guidance* |
| * If a medication cannot be mapped to RxNorm or NDC, it should still be present and RAW\_MEDADMIN\_NAME should be populated. * Only include administrations that were actually delivered to the patient, if that level of specificity is available in the source system. * Patient-reported medication administrations are not within the scope of this table. * See Reference Table 4 for the ordering strategy for RxNorm Term Types. * Do not populate CDM fields with information derived from the RXCUI (e.g., MEDADMIN\_DOSE\_ADMIN). Populate fields only if data are captured in the source system as a discrete value. * Populate records with the RXCUI as it existed at the time the order was entered, even if the RXCUI is no longer active. Do not attempt to update inactive RXCUIs with a more recent value. * If a medication mixture contains multiple RXCUIs (e.g., inpatient mixture), each individual medication from the order set should be included as an individual record with a unique MEDADMINID. Each individual medication is expected to have a unique dose. * ENCOUNTERID is expected to be present for records in the MED\_ADMIN table. * For administrations where the amount ordered is listed as a rate (e.g., infusions), if the DOSE/DOSE\_UNIT values are specified as a rate, and those fields are stored discretely in your EHR, populate the relevant CDM fields. Assuming that START\_DATE/START\_TIME and STOP\_DATE/STOP\_TIME are also populated, it will be possible to compute the rate analytically. Otherwise, leave blank. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **MED\_ADMIN Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and Descriptive Text for*  *Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level Implementation Guidance* |
| MEDADMINID | RDBMS  Text(x) | SAS Char(x) | . | Arbitrary identifier for each unique  MED\_ADMIN record. | PCORnet |  |
| PATID | RDBMS  Text(x) | SAS Char(x) | . | Arbitrary person-level identifier used to  link across tables. | MSCDM v4.0 | All PATIDs must be present in the DEMOGRAPHIC table. |
| ENCOUNTERID | RDBMS  Text(x) | SAS Char(x) | . | Arbitrary encounter-level identifier. The  ENCOUNTERID should be present. | MSCDM v4.0 | All ENCOUNTERIDs must be present in the ENCOUNTER table. |
| 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  administration records. | PCORnet |  |
| MEDADMIN\_PROVID ERID | 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 | All PROVIDERIDs must be present in the PROVIDER table. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **MED\_ADMIN Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and*  *Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level Implementation Guidance* |
| MEDADMIN\_ START\_DATE | RDBMS Date | SAS Date (Numeric) | . | Date medication administration started/occurred | PCORnet | Populate for single point-in-time administrations, as well as continuous time administrations,  such as infusions. |
| MEDADMIN\_ START\_TIME | RDBMS  Text(5): Format as HH:MI using 24-hour clock and zero- padding for  hour and minute | SAS Time (Numeric) | . | Time medication administration started/occurred | PCORnet | Populate for single point-in-time administrations, as well as continuous time administrations, such as infusions. |
| MEDADMIN\_STOP\_D ATE | RDBMS Date | SAS Date (Numeric) | . | Date medication administration ended | PCORnet | Populate for continuous time administrations, such as infusions. Leave blank if the administration is a  single point-in-time event. |
| MEDADMIN\_ STOP\_TIME | RDBMS  Text(5): Format as HH:MI using 24-hour clock and zero- padding for hour and minute | SAS Time (Numeric) | . | Time medication administration ended | PCORnet | Populate for continuous time administrations, such as infusions. Leave blank if the administration is a single point-in-time event. |
| MEDADMIN\_TYPE | RDBMS  Text(2) | SAS Char(2) | ND=NDC RX=RXNORM  NI=No information UN=Unknown OT=Other | Medication code type. | PCORnet | * This field is a derived attribute and is not expected to be an explicit data field within a source system * If mapping from medication database (e.g., MediSpan, FDB), and it is possible to map to   RxNorm and NDC, RxNorm is |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **MED\_ADMIN Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and*  *Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level Implementation Guidance* |
|  |  |  |  |  |  | the preferred term type. If medication administration records are stored natively as NDC, do not convert to RxNorm. |
| MEDADMIN\_CODE | RDBMS  Text(x) | SAS Char(x) | . | Medication code | PCORnet |  |
| MEDADMIN\_DOSE\_A DMIN | RDBMS  Number(x) | SAS  Numeric(lengt h 8) | . | Dose of a given mediation, as administered by the provider | PCORnet | Do not impute or derive. Populate only if captured in the source system as a discrete value. |
| MEDADMIN\_DOSE\_A DMIN\_UNIT | RDBMS  Text(x) | SAS Char(x) | See Value Set Appendix for a list of acceptable values. | Units of measure associated with the dose of the medication as administered by the provider | UCUM | * Do not impute or derive. Populate only if captured in the source system as a discrete value. * Choose the standardized unit of measure that is most reflective of the source data.   ~~ The Value Set Appendix~~ ~~contains a list of the units most~~ ~~commonly associated with~~ ~~medication records. Partners can~~ ~~use this table to aid in their~~ ~~mapping efforts, but they should~~ ~~refer back to the full value set if~~ ~~they have a medication record~~ ~~with a unit of measure that is not~~ ~~present in this curated list~~. (Guidance deprecated as of  CDM v5.0) |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **MED\_ADMIN Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and*  *Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level Implementation Guidance* |
|  |  |  |  |  |  | * This is a mixed case value set and entries should be handled accordingly. |
| MEDADMIN\_ROUTE | RDBMS  Text(x) | SAS Char(x) | See Value Set Appendix for a list of acceptable values. | Route of medication delivery. | RxNorm (SNOMED) | * The value set for Route is derived from SNOMED and may include values that are more granular than what is present in the source system. If a direct mapping is available, use the appropriate SNOMED code. If there is any possible ambiguity, use “OT” and then store the source value in RAW\_MEDAMIN\_ROUTE. For example, an Injection could map to Subcutaneous or Intramuscular or Intraocular, depending on the drug, so that would best be mapped to “OT.” * Do not impute or derive. Populate only if defined in the source system as a discrete value. |
| MEDADMIN\_SOURCE | RDBMS  Text(2) | SAS Char(2) | OD=Order/EHR DR=Derived NI=No information UN=Unknown OT=Other | Source of the medication administration record. | PCORnet | * This field is a derived attribute and is not expected to be an explicit data field within a source system * Use “OD” for medication orders entered into the EHR. * Use “DR” for all medication   orders that are derived or |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **MED\_ADMIN Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and*  *Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level Implementation Guidance* |
|  |  |  |  |  |  | imputed through analytical procedures (e.g., natural language processing). This does not apply to administrations mapped from a superset terminology or drug database (e.g., MediSpan, FDB). For those records, use “OD” (General Guidance #4). |
| RAW\_MEDADMIN\_M  ED\_NAME | RDBMS  Text(x) | SAS Char(x) | . | Field for originating, full textual  medication name from the source. | PCORnet |  |
| RAW\_MEDADMIN\_C ODE | RDBMS  Text(x) | SAS Char(x) |  | Field for originating value, prior to mapping into the PCORnet CDM value  set. | PCORnet |  |
| RAW\_MEDADMIN\_D OSE\_ADMIN | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, prior to mapping into the PCORnet CDM value  set. | PCORnet |  |
| RAW\_MEDADMIN\_D OSE\_ADMIN \_UNIT | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, prior to  mapping into the PCORnet CDM value set. | PCORnet |  |
| RAW\_MEDADMIN\_R OUTE | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, prior to mapping into the PCORnet CDM value  set. | PCORnet |  |

|  |
| --- |
| **5.16. Table: PROVIDER** |

**PROVIDER Domain Description:**

Data about the providers who are involved in the care processes documented in the CDM.

# Relational Integrity:

The PROVIDER table contains one record per PROVIDERID.

**Primary Key:** PROVIDERID

# Foreign Keys:

PROVIDER.PROVIDERID is a foreign key to ENCOUNTER.PROVIDERID (one-to-many relationship) PROVIDER.PROVIDERID is a foreign key to DIAGNOSIS.PROVIDERID (one-to-many relationship) PROVIDER.PROVIDERID is a foreign key to PROCEDURES.PROVIDERID (one-to-many relationship) PROVIDER.PROVIDERID is a foreign key to PRESCRIBING.RX\_PROVIDERID (one-to-many relationship) PROVIDER.PROVIDERID is a foreign key to MEDADMIN.MEDADMIN\_PROVIDERID (one-to-many relationship)

# Constraints:

PROVIDERID (unique; required, not null)

|  |
| --- |
| **PROVIDER Table Implementation Guidance** |
| *Guidance* |
| * Include one record per provider. * When populating provider specialty, if multiple values are available, use the specialty believed to be primary. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **PROVIDER Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and*  *Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level Implementation Guidance* |
| PROVIDERID | RDBMS  Text(x) | SAS Char(x) | . | Arbitrary identifier for each unique PROVIDER record. Does not need to be persistent across refreshes, and may be  created by methods such as sequence or GUID. | PCORnet |  |
| PROVIDER\_SEX | RDBMS  Text(2) | SAS Char(2) | A=Ambiguous F=Female M=Male  NI=No information UN=Unknown OT=Other | Sex assigned at birth. | MSCDM v4.0 with modified field size and value set  Source: Administrative Sex (HL7)  [http://phinvads.cdc.gov/v](http://phinvads.cdc.gov/vads/ViewValueSet.action?id=06D34BBC-617F-DD11-B38D-00188B398520) [ads/ViewValueSet.action](http://phinvads.cdc.gov/vads/ViewValueSet.action?id=06D34BBC-617F-DD11-B38D-00188B398520)  [?id=06D34BBC-617F-](http://phinvads.cdc.gov/vads/ViewValueSet.action?id=06D34BBC-617F-DD11-B38D-00188B398520) [DD11-B38D-](http://phinvads.cdc.gov/vads/ViewValueSet.action?id=06D34BBC-617F-DD11-B38D-00188B398520) [00188B398520](http://phinvads.cdc.gov/vads/ViewValueSet.action?id=06D34BBC-617F-DD11-B38D-00188B398520) | The “Ambiguous” category may be used for individuals who are physically undifferentiated from birth. The “Other” category may be used for individuals who are undergoing gender re-assignment. |
| PROVIDER\_SPECIALT Y\_PRIMARY | RDBMS  Text(x) | SAS Char(x) | See Value Set  Appendix for a list of acceptable values. | Primary specialty of the provider | Healthcare Provider Taxonomy Code Set |  |
| PROVIDER\_NPI | RDBMS  Number(x) | SAS  Numeric(lengt h 8) | . | National Provider Identifier (NPI) of the provider. | PCORnet | * Partners should only consider populating this field if their local governance allows it. * The expectation is that this field will primarily be used to support study-specific activities, though partners may also populate it to support their internal work. |
| PROVIDER\_NPI\_FLAG | RDBMS  Text(1) | SAS Char(1) | Y=Yes N=No | Flag to indicate whether partner has access to the National Provider  Identifier (NPI) of the provider. | PCORnet | This field is a derived attribute and is not expected to be an explicit data field within a source system |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **PROVIDER Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and*  *Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level Implementation Guidance* |
| RAW\_PROVIDER\_SPE CIALTY\_PRIMARY | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value of field, prior to mapping into the PCORnet CDM  value set. | PCORnet |  |

|  |
| --- |
| **5.17. Table: OBS\_CLIN** |

**OBS\_CLIN Domain Description:**

Standardized qualitative and quantitative clinical observations about a patient.

# Relational Integrity:

The OBS\_CLIN table contains one record OBSCLINID

**Primary Key:** OBSCLINID

# Foreign Keys:

OBSCLIN.PATID is a foreign key to DEMOGRAPHIC.PATID (many-to-one relationship) OBSCLIN.ENCOUNTERID is a foreign key to ENCOUNTER.ENCOUNTERID (many-to-one relationship) OBSCLIN.PROVIDERID is a foreign key to PROVIDER.PROVIDERID (many-to-one relationship)

# Constraints:

OBSCLINID (unique; required, not null) PATID (required, not null) OBSCLIN\_DATE (required, not null)

|  |
| --- |
| **OBS\_CLIN Table Implementation Guidance** |
| *Guidance* |
| * The OBS\_CLIN table is intended to store standardized clinical observations that have been recorded about a patient. * Examples of the types of observations that can be stored in this table include pulmonary function test results (e.g., FEV1, FVC, FEV1/FVC), echocardiogram results (e.g., left ventricle ejection fraction), vital signs not included in the VITAL table (e.g., temperature), etc. * Decisions on what to include in this table and how to prioritize the population of those records are expected to be driven primarily by potential funding opportunities. * This table provides a generalized structure for storing observations and is not optimized for analytical efficiency. As elements from this table are used in studies and/or distributed queries, additional representations of those data elements (i.e., new table structures) may be required to better support those activities. * If partners are populating pain scores (not pain-related PRO surveys) captured in an inpatient or surgical setting, these values would be expected to be present in this table, not PRO\_CM. * If an observation has a value set that includes an option of “not documented or assessed” (or similar), these values should be included in the CDM if they are present in the source system. Do not derive if an observation is missing. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **OBS\_CLIN Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level Implementation Guidance* |
| OBSCLINID | RDBMS Text(x) | SAS Char(x) | . | Arbitrary identifier for each unique  OBS\_CLIN record. | PCORnet |  |
| PATID | RDBMS Text(x) | SAS Char(x) | . | Arbitrary person-level identifier. Used to link across tables. | MSCDM v4.0 | All PATIDs must be present  in the DEMOGRAPHIC table. |
| ENCOUNTERID | RDBMS Text(x) | SAS Char(x) | . | Arbitrary encounter-level identifier used to link across tables. | MSCDM v4.0 | * Populate with the ENCOUNTERID where the observation was obtained. * All ENCOUNTERIDs must be present in the ENCOUNTER table. |
| OBSCLIN\_PROVIDERID | RDBMS Text(x) | SAS Char(x) | . | Provider code for the provider who ordered the observation. The provider code is a pseudoidentifier with a consistent crosswalk to the  real identifier. | PCORnet | All PROVIDERIDs must be present in the PROVIDER table. |
| OBSCLIN\_DATE | RDBMS Date | SAS Date  (Numeric) |  | Date of observation/measurement | MSCDM v4.0 with  modified field name |  |
| OBSCLIN\_TIME | RDBMS Text(5):  Format as HH:MI using 24-hour clock and zero- padding for hour and minute | SAS Time (Numeric) |  | Time of observation/measurement. | MSCDM v4.0 with modified field name |  |
| OBSCLIN\_TYPE | RDBMS Text(2) | SAS Char(2) | LC=LOINC SM=SNOMED CT  (observable entity) NI=No information UN=Unknown OT=Other | Terminology / vocabulary used to describe the clinical observation. | PCORnet |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **OBS\_CLIN Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level Implementation Guidance* |
| OBSCLIN\_CODE | RDBMS Text(x) | SAS Char(x) | . | Code of the clinical observation in the vocabulary/terminology specified in OBSCLIN\_TYPE. | PCORnet | * Results with local versions of LOINC codes (e.g., LOINC candidate codes) should be included in the RAW\_ table field. * 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 codes with leading zeros. |
| OBSCLIN\_RESULT\_QUA L | RDBMS Text(x) | SAS Char(x) | See Value Set Appendix for a list of acceptable values. | Standardized result for qualitative results. This variable should be NI for quantitative results. | LOINC | If qualitative result cannot be harmonized to a value in OBSCLIN\_RESULT\_QUAL  value set, please ensure that RAW\_OBSCLIN\_RESULT  is populated with result value. |
| OBSCLIN\_RESULT\_TEXT | RDBMS Text(x) | SAS Char(x) | . | Narrative/textual clinical  observations | PCORnet |  |
| OBSCLIN\_RESULT\_SNO MED | RDBMS Text(x) | SAS Char(x) | . | If the qualitative result has been mapped to SNOMED CT, the corresponding SNOMED code can  be placed here. | PCORnet | Partners are not expected to derive or impute if not present in the source system. |
| OBSCLIN\_RESULT\_NUM | RDBMS  Number(x) | SAS  Numeric(length 8) | . | Standardized/converted result for quantitative results. | MSCDM v4.0 with modified field name | Used to store quantitative results, including the numeric component of numeric results that contain operators (e.g., “<200”, “>= 0.5”). See  guidance for RESULT\_MODIFIER for  further details. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **OBS\_CLIN Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level Implementation Guidance* |
| OBSCLIN\_RESULT\_MODI FIER | 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 | Any symbols in the RAW\_RESULT value should be reflected in the RESULT\_MODIFIER  variable.  For example, if the original source data value is "<=200" then RAW\_RESULT=200  and RESULT\_MODIFIER=LE.  RESULT\_NUM would also be set to “200”. If the original source data value is text, then RESULT\_MODIFIER=TXIf  the original source data value  is a numeric value, then RESULT\_MODIFIER=EQ |
| OBSCLIN\_RESULT\_UNIT | RDBMS Text(x) | SAS Char(x) | See Value Set Appendix for a list of acceptable values. | Converted/standardized units for the result. | UCUM | * Chose the standardized unit of measure that is most reflective of the source data (if applicable). * This is a mixed case value set and entries should be handled accordingly. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **OBS\_CLIN Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level Implementation Guidance* |
| OBSCLIN\_SOURCE | RDBMS Text(2) | SAS Char(2) | OD=Order/EHR BI=Billing CL=Claim  RG=Registry / ancillary system  DR=Derived NI=No information UN=Unknown OT=Other | Source of the information for the lab result. | PCORnet | * This field is a derived attribute and is not expected to be an explicit data field within a source system * Use “OD” for clinical observations that are sourced from the EHR. * Use “BI” for all clinical observations that are generated through the physician and hospital billing process (it is unlikely that this value will be used). * Use “CL” for all clinical observations that are sourced from pharmacy or medical claims. * Use “RG” for clinical observations that are sourced from a separate registry or ancillary clinical system. * Use “DR” for all clinical observations that are derived or imputed through analytical procedures (e.g., natural   language processing). |
| RAW\_OBSCLIN\_NAME | RDBMS Text(x) | SAS Char(x) | . | Local name related to an individual  clinical observation/measurement. | PCORnet |  |
| RAW\_OBSCLIN\_CODE | RDBMS Text(x) | SAS Char(x) | . | Local code related to an individual clinical observation/measurement. | PCORnet |  |
| RAW\_OBSCLIN\_TYPE | RDBMS Text(x) | SAS Char(x) | . | Terminology related to the code in  RAW\_OBSGEN\_CODE. | PCORnet |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **OBS\_CLIN Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level Implementation Guidance* |
| RAW\_OBSCLIN\_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\_OBSCLIN\_MODIFIE  R | RDBMS Text(x) | SAS Char(x) | . | The original modifier text as  represented in your source data. | PCORnet |  |
| RAW\_OBSCLIN\_UNIT | RDBMS Text(x) | SAS Char(x) | . | Original units for the result in your  source data. | PCORnet |  |

|  |
| --- |
| **5.18. Table: OBS\_GEN** |

**OBS\_GEN Domain Description:**

Table to store everything else.

# Relational Integrity:

The OBS\_GEN table contains one record OBSGENID

**Primary Key:** OBSGENID

# Foreign Keys:

OBSGEN.PATID is a foreign key to DEMOGRAPHIC.PATID (many-to-one relationship) OBSGEN.ENCOUNTERID is a foreign key to ENCOUNTER.ENCOUNTERID (zero/many-to-one relationship) OBSGEN.PROVIDERID is a foreign key to PROVIDER.PROVIDERID (many-to-one relationship)

# Constraints:

OBSGENID (unique; required, not null) PATID (required, not null) OBSGEN\_DATE (required, not null)

|  |
| --- |
| **OBS\_GEN Table Implementation Guidance** |
| *Guidance* |
| * Partners may use this table to store network- or study-specific data elements. * Records in this table are not expected to be used in queries distributed by the DRN OC. * This table provides a generalized structure for storing observations and is not optimized for analytical efficiency. As elements from this table are used in studies and/or distributed queries, additional representations of those data elements (i.e., new table structures) may be required to better support those activities. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **OBS\_GEN Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level Implementation Guidance* |
| OBSGENID | RDBMS Text(x) | SAS Char(x) | . | Arbitrary identifier for each unique  OBS\_GEN record. | PCORnet |  |
| PATID | RDBMS Text(x) | SAS Char(x) | . | Arbitrary person-level identifier. Used to link across tables. | MSCDM v4.0 | All PATIDs must be present  in the DEMOGRAPHIC table. |
| ENCOUNTERID | RDBMS Text(x) | SAS Char(x) | . | Arbitrary encounter-level identifier used to link across tables. This field should be populated if the observation was recorded as part of a healthcare encounter. | PCORnet | * Populate with the ENCOUNTERID where the observation was recorded. * All ENCOUNTERIDs in this table must also be present in the ENCOUNTER table. |
| OBSGEN\_PROVIDERID | RDBMS Text(x) | SAS Char(x) | . | Provider code for the provider who recorded the observation. The provider code is a pseudoidentifier with a consistent crosswalk to the  real identifier. | PCORnet | All PROVIDERIDs must be present in the PROVIDER table. |
| OBSGEN\_DATE | RDBMS Date | SAS Date  (Numeric) |  | Date of observation/measurement | MSCDM v4.0 with  modified field name |  |
| OBSGEN\_TIME | RDBMS Text(5):  Format as HH:MI using 24-hour clock and zero- padding for hour and minute | SAS Time (Numeric) |  | Time of observation/measurement. | MSCDM v4.0 with modified field name |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **OBS\_GEN Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level Implementation Guidance* |
| OBSGEN\_TYPE | RDBMS Text(30) | SAS Char(30) | 09DX=ICD-9-CM  09PX=ICD-10-PCS  10DX=ICD-10-CM~~/PCS~~  10PX=ICD-10-PCS  11DX=ICD-11-CM~~/PCS~~  11PX=ICD-11-PCS  ON=ICD-O (Oncology) SM=SNOMED  HP=Human Phenotype Ontology  HG=Human Genome Organization LC=LOINC RX=RXNORM ND=NDC  CH=CPT or HCPCS  GM=Global Medical Device Nomenclature ~~CVX=Vaccine~~ ~~administered~~ UD\_\*=User-defined PC\_\*=PCORnet reserved NI=No information  UN=Unknown OT=Other | Terminology/vocabulary used to describe the observation.  Networks/partners can define their own terminologies with strings starting with “UD\_”.  Strings that start with “PC\_” are reserved for network-wide activities and will be assigned by the Coordinating Center. | PCORnet | * For user-defined values, use the listed convention. * If a study involves this table and spans multiple networks, participants should ensure that they chose a UD string that is not in use in those locations. * The ICD CM and PCS terminologies are listed as separate options, since codes may be present without decimals. This can lead to collisions between values if they are co-mingled. |
| OBSGEN\_CODE | RDBMS Text(x) | SAS Char(x) | . | Standardized code denoting the observations based on the terminology/vocabulary specified in  OBSGEN\_TYPE | PCORnet |  |
| OBSGEN\_RESULT\_QUAL | RDBMS Text(x) | SAS Char(x) | See Value Set Appendix for a list of acceptable values. | Standardized result for qualitative results. This variable should be NI for quantitative results. | LOINC | Use RAW\_OBSGEN\_RESULT to  store qualitative results that cannot be harmonized to the  defined value set. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **OBS\_GEN Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level Implementation Guidance* |
| OBSGEN\_RESULT\_TEXT | RDBMS Text(x) | SAS Char(x) | . | Narrative/textual observations. | PCORnet |  |
| OBSGEN\_RESULT\_NUM | RDBMS  Number(x) | SAS  Numeric(length 8) | . | Standardized/converted result for quantitative results. | MSCDM v4.0 with modified field name | Used to store quantitative results, including the numeric component of numeric results that contain operators (e.g., “<200”, “>= 0.5”). See  guidance for RESULT\_MODIFIER for  further details. |
| OBSGEN\_RESULT\_MODI FIER | 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 | Any symbols in the RAW\_RESULT value should be reflected in the RESULT\_MODIFIER  variable.  For example, if the original source data value is "<=200" then RAW\_RESULT=200  and RESULT\_MODIFIER=LE.  RESULT\_NUM would also be set to “200”. If the original source data value is text, then RESULT\_MODIFIER=TX If  the original source data value is a numeric value, then  RESULT\_MODIFIER=EQ |
| OBSGEN\_RESULT\_UNIT | RDBMS Text(x) | SAS Char(x) | See Value Set Appendix for a list of acceptable values. | Converted/standardized units for the result. | UCUM | * Chose the standardized unit of measure that is most reflective of the source data (if applicable). * This is a mixed case value set and entries should be handled accordingly. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **OBS\_GEN Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level Implementation Guidance* |
| OBSGEN\_TABLE\_MODIFI ED | RDBMS Text(3) | SAS Char(3) | ENR=ENROLLMENT ENC=ENCOUNTER DX=DIAGNOSIS PX=PROCEDURES VT=VITAL DSP=DISPENSING LAB=LAB\_RESULT\_C M  CON=CONDITION PRO=PRO\_CM RX=PRESCRIBING PT=PCORNET\_TRIAL DTH=DEATH DC=DEATH\_CAUSE MA=MED\_ADMIN OC=OBS\_CLIN OB=OBS\_GEN  OT=Other | Table name when observation describes attributes of an existing record in the CDM. | PCORnet | * If observation record modifies something other than the patient (i.e., attribute about an encounter), a link to that table can be included here. * If a value is listed in OBSGEN\_TABLE\_MODIF IED, then a corresponding ID should be listed in OBSGEN\_ID\_MODIFIED. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **OBS\_GEN Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level Implementation Guidance* |
| OBSGEN\_ID\_MODIFIED | RDBMS Text(x) | SAS Char(x) | . | Identifier when observation describes attributes of an existing record in the CDM. | PCORnet | * If observation record modifies something other than the patient (i.e., attribute about an encounter), a link to that record can be included here. * If a value is listed in OBSGEN\_TABLE\_MODIF IED, then a corresponding ID should be listed in OBSGEN\_ID\_MODIFIED. * If modifying a record in OBS\_GEN, the value of OBSGEN\_ID\_MODIFIED must be different than the value of OBSGENID for that record. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **OBS\_GEN Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level Implementation Guidance* |
| OBSGEN\_SOURCE | RDBMS Text(2) | SAS Char(2) | OD=Order/EHR BI=Billing CL=Claim  RG=Registry / ancillary system  SR=Survey system / mobile app DR=Derived NI=No information UN=Unknown OT=Other | Source of the information for the lab result. | PCORnet | * This field is a derived attribute and is not expected to be an explicit data field within a source system * Use “OD” for observations that are sourced from the EHR. * Use “BI” for observations that are generated through the physician and hospital billing process (it is unlikely that this value will be used). * Use “CL” for observations that are sourced from pharmacy or medical claims. * Use “RG” for observations that are sourced from a separate registry or ancillary clinical system. * Use “SR” for observations that are sourced from an external survey system or mobile app. * Use “DR” for |
|  |  |  |  |  |  | observations that are |
|  |  |  |  |  |  | derived or imputed |
|  |  |  |  |  |  | through analytical |
|  |  |  |  |  |  | procedures (e.g., natural |
|  |  |  |  |  |  | language processing). |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **OBS\_GEN Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level Implementation Guidance* |
| RAW\_OBSGEN\_NAME | RDBMS Text(x) | SAS Char(x) | . | Local name related to an individual  clinical observation/measurement. | PCORnet |  |
| RAW\_OBSGEN\_CODE | RDBMS Text(x) | SAS Char(x) | . | Local code related to an individual  clinical observation/measurement. | PCORnet |  |
| RAW\_OBSGEN\_TYPE | RDBMS Text(x) | SAS Char(x) | . | Terminology related to the code in  RAW\_OBSGEN\_CODE. | PCORnet |  |
| RAW\_OBSGEN\_RESULT | RDBMS Text(x) | SAS Char(x) | . | The original test result value as seen  in your source data. | PCORnet |  |
| RAW\_OBSGEN\_UNIT | RDBMS Text(x) | SAS Char(x) | . | Original units for the result in your source data. | PCORnet |  |

|  |
| --- |
| **5.19. Table: HASH\_TOKEN** |

**HASH\_TOKEN Domain Description:**

Encrypted, keyed secure hash tokens that are used to match patient records across DataMarts using privacy-preserving record linkage methods.

**Relational Integrity:**

The HASH\_TOKEN table contains one record per patient.

**Primary Key:** PATID

**Foreign Key**:

HASH\_TOKEN.PATID is a foreign key to DEMOGRAPHIC.PATID (one-to-one relationship)

**Constraints:**

PATID (unique; required, not null)

|  |
| --- |
| **HASH\_TOKEN Table Implementation Guidance** |
| *Guidance* |
| * Every patient in the DEMOGRAPHIC table is expected to have one record in the HASH\_TOKEN table. * Tokens are generated from personally-identifiable information (PII) that is stored in each partner’s PRIVATE\_DEMOGRAPHIC table and PRIVATE\_ADDRESS\_HISTORY table. The PII is used as input to the Datavant DeId module. Tokens should not be placed into the CDM until they have been transformed into Site-PCORnet transit tokens using the Datavant Link module. * Tokens are generated based on data availability. If input data is not present for a given token strategy (e.g., combination of PII elements), no token will be generated and an error code will be produced instead. These token error codes should be loaded into the HASH\_TOKEN table (i.e., there should not be any null values). Do not suppress the error codes in the output of the Datavant software. * Each successfully generated token has a fixed length of 44 characters. Do not enforce a 44-character constraint, however, to accommodate the error codes generated in the case of tokenization failure. * Tokens should be generated as part of every refresh. Partners can choose to generate tokens for all patients, or only for those patients who were added between refreshes or had updates to their PII. * Select tokens generated using the Datavant DeID module are certified as de-identified data via the HIPAA Expert Determination method in accordance with the HIPAA Privacy Rule (45 CFR parts 160 and 164). All tokens in the HASH\_TOKEN table satisfy this criteria and are controlled via the Datavant DeID PCORnet configuration settings. * Additional token strategies are available and can be implemented as needed on a per-study basis based on the study-specific data dictionary. * See the Supplemental Guide on Privacy-Preserving Record Linkage for additional implementation details and guidance (separate document). |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **HASH\_TOKEN Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level*  *Implementation Guidance* |
| PATID | RDBMS  Text(x) | SAS Char(x) | . | Arbitrary person-level identifier. Used to link across tables. PATID is passed through the Datavant DeID module in order to be associated with the generated encrypted keyed hashes. | MSCDM v4.0 | All PATIDs should be present in the DEMOGRAPHIC  table. |
| TOKEN\_01 | RDBMS  Text(x) | SAS Char(x) | . | Encrypted keyed hash generated from PII using token strategy 01 in Datavant DeID. Enforced through PCORnet configuration  setting. | PCORnet |  |
| TOKEN\_02 | RDBMS  Text(x) | SAS Char(x) | . | Encrypted keyed hash generated from PII using token strategy 02 in Datavant DeID. Enforced through PCORnet configuration  setting. | PCORnet |  |
| TOKEN\_03 | RDBMS  Text(x) | SAS Char(x) | . | Encrypted keyed hash generated from PII using token strategy 03 in Datavant DeID. Enforced through PCORnet configuration  setting. | PCORnet |  |
| TOKEN\_04 | RDBMS  Text(x) | SAS Char(x) | . | Encrypted keyed hash generated from PII using token strategy 04 in Datavant DeID.  Enforced through PCORnet configuration setting. | PCORnet |  |
| TOKEN\_05 | RDBMS  Text(x) | SAS Char(x) | . | Encrypted keyed hash generated from PII using token strategy 05 in Datavant DeID. Enforced through PCORnet configuration  setting. | PCORnet |  |
| ~~TOKEN\_12~~ | ~~RDBMS~~  ~~Text(x)~~ | ~~SAS Char(x)~~ | . | ~~Encrypted hash generated from PII using token~~  ~~strategy 12.~~ | ~~PCORnet~~ | This field is deprecated  as of CDM v5.1 |
| TOKEN\_16 | RDBMS  Text(x) | SAS Char(x) | . | Encrypted keyed hash generated from PII using token strategy 16 in Datavant DeID.  Enforced through PCORnet configuration setting. | PCORnet |  |
| ~~TOKEN\_17~~ | ~~RDBMS~~  ~~Text(x)~~ | ~~SAS Char(x)~~ | . | ~~Encrypted hash generated from PII using token~~  ~~strategy 17.~~ | ~~PCORnet~~ | This field is deprecated  as of CDM v5.1 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **HASH\_TOKEN Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level*  *Implementation Guidance* |
| ~~TOKEN\_21~~ | ~~RDBMS~~  ~~Text(x)~~ | ~~SAS Char(x)~~ | . | ~~Encrypted hash generated from PII using token~~ ~~strategy 21.~~ | ~~PCORnet~~ | This field is deprecated as of CDM v5.1 |
| ~~TOKEN\_22~~ | ~~RDBMS~~  ~~Text(x)~~ | ~~SAS Char(x)~~ | . | ~~Encrypted hash generated from PII using token~~  ~~strategy 22.~~ | ~~PCORnet~~ | This field is deprecated  as of CDM v5.1 |
| ~~TOKEN\_23~~ | ~~RDBMS~~  ~~Text(x)~~ | ~~SAS Char(x)~~ | . | ~~Encrypted hash generated from PII using token~~  ~~strategy 23.~~ | ~~PCORnet~~ | This field is deprecated  as of CDM v5.1 |

|  |
| --- |
| **5.20. Table: LDS\_ADDRESS\_HISTORY** |

**LDS\_ADDRESS\_HISTORY Domain Description:** Longitudinal record of a patient’s address that adheres to the requirements of a Limited Data Set.

**Relational Integrity:**

The LDS\_ADDRESS\_HISTORY table can contain many records per patient.

**Primary Key:** ADDRESSID

**Foreign Key**:

ADDRESS\_HISTORY.PATID is a foreign key to DEMOGRAPHIC.PATID (many-to-one relationship)

**Constraints:**

ADDRESSID (unique; required, not null) PATID (required, not null) ADDRESS\_USE (required, not null) ADDRESS\_TYPE (required, not null)

ADDRESS\_PREFERRED (required, not null)

|  |
| --- |
| **LDS\_ADDRESS\_HISTORY Table Implementation Guidance** |
| *Guidance* |
| * Expect multiple records per individual * This table is currently limited to addresses in the United States. * Partners can limit records in this table to validated addresses if known. * If partners have difficulty constructing a longitudinal address history for patients within their DataMart, they should prioritize populating the current address for each patient. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **LDS\_ADDRESS\_HISTORY Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level Implementation*  *Guidance* |
| ADDRESSID | RDBMS  Text(x) | SAS Char(x) | . | Arbitrary identifier for each unique address  record. | PCORnet |  |
| PATID | RDBMS  Text(x) | SAS Char(x) | . | Arbitrary person-level identifier. Used to link across tables. | MSCDM v4.0 | All PATIDs should be present in the DEMOGRAPHIC  table. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **LDS\_ADDRESS\_HISTORY Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level*  *Implementation Guidance* |
| ADDRESS\_USE | RDBMS  Text(2) | SAS Char(2) | HO=Home WO=Work TP=Temp OL=Old/Incorrect NI=No  information UN=Unknown OT=Other | Purpose of the address.  Details of categorical definitions:  Home: A communication address at home.  Work: An office address. First choice for business- related contacts during business hours.  Temp: A temporary address.  Old/Incorrect: This address is no longer in use (or was never correct but retained for records). | FHIR - ADDRESS | * This field may be a derived attribute that is not an explicit data field within a source system. * Use the period start/end fields to indicate if an address is no longer valid. **Do not change values of HO/WO/TP to OL if a new address is available**. * The old/incorrect value is included in case partners are doing a bulk load and it is present in their source system. It is acceptable to exclude these records, however. * If addresses within the source system are reasonably expected to represent the patient’s home address, it is acceptable to mark these as “HO.” |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **LDS\_ADDRESS\_HISTORY Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level*  *Implementation Guidance* |
| ADDRESS\_TYPE | RDBMS  Text(2) | SAS Char(2) | PO=Postal PH=Physical BO=Both NI=No  information UN=Unknown OT=Other | Type of address.  Details of categorical definitions:  Postal: mailing address – PO Boxes and care-of addresses.  Physical: A physical address that can be visited. Both: An address that is both physical and postal. | FHIR - ADDRESS | * This field may be a derived attribute that is not an explicit data field within a source system. * Addresses that are clearly not postal-only addresses can be marked as “Both” |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ADDRESS\_PREFERRED | RDBMS  Text(2) | SAS Char(2) | Y=Yes N=No | Indicates whether this address is the preferred one for a given patient, address use and address type within a given address period. | PCORnet | * This field may be a derived attribute that is not an explicit data field within a source system. * It is possible to have an address marked as preferred for every address type/use within each defined address period. * If a new address period is defined, it is not necessary to set the flag(s) back to “N” for any of the existing address periods. * If only one address is present for a given period, that address should be marked as preferred. * If multiple addresses are present for a period, one should be denoted as preferred. The rest should be labeled as “N”. * If multiple addresses are present for a period without clear institutional guidance as to which is preferred, partners can   use a heuristic to make |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **LDS\_ADDRESS\_HISTORY Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level*  *Implementation Guidance* |
|  |  |  |  |  |  | a determination (i.e., address associated with most recent encounter, address used for billing, etc.) |
| ADDRESS\_CITY | RDBMS  Text(x) | SAS Char(x) | . | The name of the city, town, village or other  community. | FHIR -  ADDRESS |  |
| ADDRESS\_STATE | RDBMS  Text(2) | SAS Char(2) | See Value Set Appendix for a list of acceptable  values. | State, as represented by 2-digit postal abbreviation. | PCORnet |  |
| ADDRESS\_ZIP5 | RDBMS  Text(5) | SAS Char(5) | . | 5-digit postal code for the address. | PCORnet |  |
| ADDRESS\_ZIP9 | RDBMS  Text(9) | SAS Char(9) | . | 9-digit postal code for the address. | PCORnet | Do not include hyphens. |
| ADDRESS\_PERIOD\_START | RDBMS Date | SAS Date (Numeric) | . | Initial date when the address is known to be in use. | FHIR - ADDRESS | * If the date the address is known to be in use is unknown, it is acceptable to use the date the record was created in the source system or the date the record was first created in the CDM. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **LDS\_ADDRESS\_HISTORY Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level*  *Implementation Guidance* |
| ADDRESS\_PERIOD\_END | RDBMS Date | SAS Date (Numeric) | . | Date when address was no longer in use. | FHIR - ADDRESS | * Only the current address period is expected to have a null value. All previous periods are expected to have a defined end date. |

|  |
| --- |
| **5.21.** **Table: IMMUNIZATION** |

**IMMUNIZATION Domain Description:**

Records of immunizations that have been delivered within the health system as well as reports of those administered elsewhere.

**Relational Integrity:**

The IMMUNIZATION table contains one record per IMMUNIZATIONID.

**Primary Key:** IMMUNIZATIONID

**Foreign Keys:**

IMMUNIZATION.PATID is a foreign key to DEMOGRAPHIC.PATID (many-to-one relationship) IMMUNIZATION.ENCOUNTERID is a foreign key to ENCOUNTER.ENCOUNTERID (zero/many-to-one relationship) IMMUNIZATION.VX\_PROVIDERID is a foreign key to PROVIDER.PROVIDERID (zero/many-to-one relationship) IMMUNIZATION.PROCEDURESID is a foreign key to PROCEDURES.PROCEDURESID (zero/many-to-one relationship)

**Constraints:**

IMMUNIZATIONID (unique; required, not null) PATID (required, not null)

VX\_CODE (required, not null) VX\_CODE\_TYPE (required, not null) VX\_STATUS (required, not null)

|  |
| --- |
| **IMMUNIZATION Table Implementation Guidance** |
| *Guidance* |
| * Do not include study vaccines. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **IMMUNIZATION Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and*  *Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level Implementation Guidance* |
| IMMUNIZATIONID | RDBMS  Text(x) | SAS Char(x) | . | Arbitrary identifier for each unique  IMMUNIZATION record. | PCORnet |  |
| PATID | RDBMS  Text(x) | SAS Char(x) | . | Arbitrary person-level identifier used to  link across tables. | MSCDM v4.0 | All PATIDs must be present in  the DEMOGRAPHIC table. |
| ENCOUNTERID | RDBMS  Text(x) | SAS Char(x) | . | Arbitrary encounter-level identifier. This should be present if the immunization activity is directly  associated with an encounter. | MSCDM v4.0 | All ENCOUNTERIDs in this table must be present in the ENCOUNTER table. |
| PROCEDURESID | RDBMS  Text(x) | SAS Char(x) | . | This is an optional relationship to the PROCEDURES table and is not expected to be available for all  immunizations. One procedure may generate multiple immunization records. | PCORnet | All PROCEDURESIDs in this table must be present in the PROCEDURES table. |
| VX\_PROVIDERID | RDBMS  Text(x) | SAS Char(x) | . | Provider code for the provider who delivered the immunization. The provider code is a pseudoidentifier with a consistent crosswalk to the real  identifier. | PCORnet | All PROVIDERIDs in this table must be present in the PROVIDER table. |
| VX\_RECORD\_DATE | RDBMS Date | SAS Date  (Numeric) | . | Date immunization was recorded (i.e.,  date record was created). | FHIR – Immunization |  |
| VX\_ADMIN\_DATE | RDBMS Date | SAS Date (Numeric) | . | Date immunization was administered, if known. | FHIR - Immunization | * Leave blank if unknown. * Partial dates are allowed.   Follow the CDM date imputation guidance and ensure the HARVEST table is properly updated. |
| VX\_CODE\_TYPE | RDBMS  Text(2) | SAS Char(2) | CX=CVX ND=NDC CH = CPT or HCPCS  RX=RXNORM  NI=No information | Immunization code type. | PCORnet | * This field is a derived attribute and is not expected to be an explicit data field within a source system |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **IMMUNIZATION Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and*  *Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level Implementation Guidance* |
|  |  |  | UN=Unknown OT=Other |  |  | * If immunizations are represented within the source system by multiple terminologies, choose the one with the highest level of granularity. |
| VX\_CODE | RDBMS  Text(x) | SAS Char(x) | . | Immunization code | PCORnet |  |
| VX\_STATUS | RDBMS  Text(2) | SAS Char(2) | CP=Completed ER=Entered in error ND=Not Done IC=Incomplete NI=No information UN=Unknown  OT=Other | Status of the immunization. | FHIR – Immunization; PCORnet | It is not necessary to load immunizations that were entered in error if partners can easily distinguish these data in the source system(s). |
| VX\_STATUS\_REASON | RDBMS  Text(2) | SAS Char(2) | IM=Immunity MP=Medical precaution OS=Out of stock PO=Patient objection  NI=No information UN=Unknown OT=Other | Reason immunization is incomplete or not done. | FHIR – Immunization | This field is not expected to be populated for immunizations with a status of “CP” or “ER” |
| VX\_SOURCE | RDBMS  Text(2) | SAS Char(2) | OD=Internal administration EF=External feed IS=Immunization Information Systems PR=Patient-reported DR=Derived  NI=No information UN=Unknown OT=Other | Source of the prescribing information. | PCORnet | * This field is a derived attribute, and is not expected to be an explicit data field within a source system * Use “DR” for all immunizations that are derived or imputed through analytical procedures (e.g., natural language processing). |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **IMMUNIZATION Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and*  *Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level Implementation Guidance* |
|  |  |  |  |  |  | * Use “IS” for immunizations sourced from the CDC Immunization Information Systems |
| VX\_DOSE | RDBMS  Number(x) | SAS  Numeric(lengt h 8) | . | Dose of a given immunization | PCORnet | Do not impute or derive. Populate  only if captured in the source system as a discrete value. |
| VX\_DOSE\_UNIT | RDBMS  Text(x) | SAS Char(x) | See Value Set Appendix for a list of acceptable values. | Units of measure associated with the dose of the immunization as delivered by the provider | UCUM | * Do not impute or derive. Populate only if captured in the source system as a discrete value. * Choose the standardized unit of measure that is most reflective of the source data. * This is a mixed case value set and entries should be handled   accordingly. |
| VX\_ROUTE | RDBMS  Text(x) | SAS Char(x) | See Value Set Appendix for a list of acceptable values. | Route of immunization delivery. | SNOMED | * The value set for Route is derived from SNOMED and may include values that are more granular than what is present in the source system. If a direct mapping is available, use the appropriate SNOMED code. If there is any possible ambiguity, use “OT” and then store the source value in RAW\_VX\_ROUTE. For example, an Injection could map to Subcutaneous or Intramuscular or Intraocular, depending on the drug, so that would best be   mapped to “OT.” |

Page 166 of 208

<https://pcornet.org/data-driven-common-model/>

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **IMMUNIZATION Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and*  *Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level Implementation Guidance* |
|  |  |  |  |  |  | * Do not impute or derive. Populate only if captured in the source system as a discrete value. |
| VX\_BODY\_SITE | RDBMS Text (x) | SAS Char(x) | See Value Set Appendix for a list of acceptable values. | Body site where the immunization was delivered. | FHIR – Immunization (SNOMED) | * Most immunizations are expected in the right arm (“RA”) or left arm (“LA”) |
| VX\_MANUFACTURER | RDBMS  Text(x) | SAS Char(x) | See Value Set  Appendix for a list of acceptable values. | Manufacturer of the immunization, coded using MVX terminology. | PCORnet |  |
| VX\_LOT\_NUM | RDBMS  Text(x) | SAS Char(x) | . | Lot number of the immunization. | FHIR - Immunization | * This information is typically captured within EHRs or e- prescribing as part of the ordering process. * Do not impute or derive. Populate only if captured in the source system as a discrete value. |
| VX\_EXP\_DATE | RDBMS Date | SAS Date  (Numeric) | . | Expiration date of the immunization. | FHIR – Immunization |  |
| RAW\_VX\_NAME | RDBMS  Text(x) | SAS Char(x) | . | Field for originating, full textual  immunization name from the source. | PCORnet |  |
| RAW\_VX\_CODE | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, prior to mapping into the PCORnet CDM value  set. | PCORnet |  |
| RAW\_VX\_CODE\_TYP E | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, prior to  mapping into the PCORnet CDM value set. | PCORnet |  |
| RAW\_VX\_DOSE | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, prior to  mapping into the PCORnet CDM value set. | PCORnet |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **IMMUNIZATION Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and*  *Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level Implementation Guidance* |
| RAW\_VX\_DOSE\_UNI T | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, prior to mapping into the PCORnet CDM value  set. | PCORnet |  |
| RAW\_VX\_ROUTE | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, prior to  mapping into the PCORnet CDM value set. | PCORnet |  |
| RAW\_VX\_BODY\_SITE | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, prior to mapping into the PCORnet CDM value  set. | PCORnet |  |
| RAW\_VX\_STATUS | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, prior to  mapping into the PCORnet CDM value set. | PCORnet |  |
| RAW\_VX\_STATUS\_R EASON | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, prior to  mapping into the PCORnet CDM value set. | PCORnet |  |
| RAW\_VX\_MANUFAC TURER | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, prior to mapping into the PCORnet CDM value  set. | PCORnet |  |

|  |
| --- |
| **5.22.** **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 Implementation Guidance** |
| *Guidance* |
| * If partners need to impute date values, whether for a portion of the date (e.g., month) or the entire string, a value of “02” should be chosen for the relevant DATE\_MGMT field(s). * If partners must impute the entire date for a field, this should only be done for those dates that are required. Optional fields should be left blank in these situations. * All date obfuscation and imputation strategies must be thoroughly documented in the Extract, Transform and Load (ETL) Annotated Data Dictionary (ADD). * Partners should refrain from obfuscating dates within the CDM (see General Guidance #2). Future versions of the CDM may remove options “03” and “04” from the value set of the DATE\_MGMT fields. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level Implementation*  *Guidance* |
| NETWORKID | RDBMS  Text(10) | SAS Char(10) | . | This identifier is assigned by the PCORnet  Distributed Research Network Operations Center (DRN OC) | 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) | PCORnet |  |
| DATAMARTID | RDBMS  Text(10) | SAS Char(10) | . | This identifier is assigned by the PCORnet Distributed Research Network Operations  Center (DRN OC) | 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) | 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* | *Field-level*  *Implementation Guidance* |
| 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(x) | SAS  Numeric(length 8) | . | Version currently implemented within this datamart (for example, 1.0, 2.0, 3.0). | PCORnet |  |
| 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 |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level*  *Implementation Guidance* |
| 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 employed for the BIRTH\_DATE field in the DEMOGRAPHIC table. | 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 employed for the ENR\_START\_DATE field in the ENROLLMENT table. | 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* | *Field-level*  *Implementation Guidance* |
| 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 employed for the ENR\_END\_DATE field in the ENROLLMENT table. | 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 employed for the ADMIT\_DATE field in the ENCOUNTER table. | 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* | *Field-level*  *Implementation Guidance* |
| 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 employed for the DISCHARGE\_DATE field in the ENCOUNTER table. | PCORnet |  |
| DX\_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 employed for the DX\_DATE field in the ENCOUNTER table. | 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* | *Field-level*  *Implementation Guidance* |
| 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 employed for the PX\_DATE field in the PROCEDURES table. | 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 employed for the RX\_ORDER\_DATE field in the PRESCRIBING table. | 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* | *Field-level*  *Implementation Guidance* |
| 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 employed for the RX\_START\_DATE field in the PRESCRIBING table. | PCORnet |  |
| 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 employed for the RX\_END\_DATE field in the PRESCRIBING table. | 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* | *Field-level*  *Implementation Guidance* |
| 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 employed for the DISPENSE\_DATE field in the DISPENSING table. | 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 employed for the LAB\_ORDER\_DATE field in the LAB\_RESULT\_CM table. | 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* | *Field-level*  *Implementation Guidance* |
| 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 employed for the SPECIMEN\_DATE field in the LAB\_RESULT\_CM table. | 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 employed for the RESULT\_DATE field in the LAB\_RESULT\_CM table. | 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* | *Field-level*  *Implementation Guidance* |
| 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 employed for the MEASURE\_DATE field in the VITAL table. | PCORnet |  |
| 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 employed for the ONSET\_DATE field in the CONDITION table. | 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* | *Field-level*  *Implementation Guidance* |
| 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 employed for the REPORT\_DATE field in the CONDITION table. | 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 employed for the RESOLVE\_DATE field in the CONDITION table. | 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* | *Field-level*  *Implementation Guidance* |
| 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 employed for the PRO\_DATE field in the PRO\_CM table. | PCORnet |  |
| DEATH\_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 employed for the DEATH\_DATE field in the DEATH table. | 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* | *Field-level*  *Implementation Guidance* |
| MEDADMIN\_START\_DATE\_MGM T | 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 employed for the MEDADMIN\_START\_DATE field in the MED\_ADMIN table. | PCORnet |  |
| MEDADMIN\_STOP\_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 employed for the MEDADMIN\_STOP\_DATE field in the MED\_ADMIN table. | 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* | *Field-level*  *Implementation Guidance* |
| OBSCLIN\_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 employed for the OBSCLIN\_DATE field in the OBS\_CLIN table. | PCORnet |  |
| OBSGEN\_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 employed for the OBSGEN\_DATE field in the OBS\_GEN table. | 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* | *Field-level*  *Implementation Guidance* |
| ADDRESS\_PERIOD\_START\_MGM T | 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 employed for the ADDRESS\_PERIOD\_START field in the LDS\_ADDRESS\_HISTORY table. | PCORnet |  |
| ADDRESS\_PERIOD\_END\_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 employed for the ADDRESS\_PERIOD\_END field in the LDS\_ADDRESS\_HISTORY table. | 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* | *Field-level*  *Implementation Guidance* |
| VX\_RECORD\_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 employed for the VX\_RECORD\_DATE field in the IMMUNIZATION table. | PCORnet |  |
| VX\_ADMIN\_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 employed for the VX\_ADMIN\_DATE field in the IMMUNIZATION table. | 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* | *Field-level*  *Implementation Guidance* |
| VX\_EXP\_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 employed for the VX\_EXP\_DATE field in the IMMUNIZATION table. | 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 |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level*  *Implementation Guidance* |
| 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 |  |
| 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\_DAT E | 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\_DAT E | 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 |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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* | *Field-level*  *Implementation Guidance* |
| 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 |  |
| REFRESH\_MED\_ADMIN\_DATE | RDBMS Date | SAS Date (Numeric) | . | Most recent date on which the present data were loaded into the MED\_ADMIN table. This date should be null if the table does not have  records. | PCORnet |  |
| REFRESH\_OBS\_CLIN\_DATE | RDBMS Date | SAS Date (Numeric) | . | Most recent date on which the present data were loaded into the OBS\_CLIN table. This  date should be null if the table does not have records. | PCORnet |  |
| REFRESH\_PROVIDER\_DATE | RDBMS Date | SAS Date (Numeric) | . | Most recent date on which the present data were loaded into the PROVIDER table. This date should be null if the table does not have  records. | PCORnet |  |
| REFRESH\_OBS\_GEN\_DATE | RDBMS Date | SAS Date (Numeric) | . | Most recent date on which the present data were loaded into the OBS\_GEN table. This  date should be null if the table does not have records. | PCORnet |  |
| REFRESH\_HASH\_TOKEN\_DATE | RDBMS Date | SAS Date (Numeric) | . | Most recent date on which the present data were loaded into the HASH\_TOKEN table. This date should be null if the table does not  have records. | PCORnet |  |
| REFRESH\_LDS\_ADDRESS\_HX\_DA TE | RDBMS Date | SAS Date (Numeric) | . | Most recent date on which the present data were loaded into the LDS\_ADDRESS\_HISTORY table. This date should be null if the table does not have  records. | PCORnet |  |
| REFRESH\_IMMUNIZATION\_DATE | RDBMS Date | SAS Date (Numeric) | . | Most recent date on which the present data were loaded into the IMMUNIZATION table. This date should be null if the table does not  have records. | PCORnet |  |

|  |
| --- |
| **6. Supplemental Table Specifications** |

|  |
| --- |
| **6.1. Supplemental Table: PRIVATE\_DEMOGRAPHIC** |

**PRIVATE\_DEMOGRAPHIC Domain Description:**

Protected table that is intended to provide a standardized representation of the personally-identifiable information (PII) that is needed to support local activities related to record linkage.

**Relational Integrity:**

The PRIVATE\_DEMOGRAPHIC table contains one record per PATID.

**Primary Key:** PATID

**Foreign Key**:

PRIVATE\_DEMOGRAPHIC.PATID is a foreign key to DEMOGRAPHIC.PATID (one-to-one relationship)

**Constraints:**

PATID (unique, required, not null) PAT\_FIRSTNAME (required, not null) PAT\_LASTNAME (required, not null)

|  |
| --- |
| **PRIVATE\_DEMOGRAPHIC Table Implementation Guidance** |
| *Guidance* |
| * ~~Tokens will be generated from the personally-identifiable information (PII) that is stored in this table and PRIVATE\_ADDRESS\_HISTORY.~~ This table is intended to provide a standardized representation for PII used in PCORnet privacy-preserving record linkage. It is not necessary to create this table if existing processes can generate the relevant PII from local source system(s). * For DataMarts that incorporate data from multiple organizations/institutions, PRIVATE\_DEMOGRAPHIC can be used to assist de-duplication efforts. See field-level guidance for more information. * Include most recent values for each field. * This table can be kept logically and physically separate from the rest of the CDM. * This table replicates the fields of the DEMOGRAPHIC table. They are included in case partners wish to create a master table that can be used to populate DEMOGRAPHIC within the CDM. SAS data types are included for the replicated fields to help with ETL programming, but partners are not expected to create a SAS-based version of this PRIVATE table It is not necessary to populate the replicated fields in this table if DEMOGRAPHIC is loaded through a separate ETL process. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **PRIVATE\_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* | *Field-level Implementation*  *Guidance* |
| PATID | RDBMS  Text(x) | SAS Char(x) | . | Arbitrary person-level identifier. Used to link across tables. Corresponds to PATID in the  DEMOGRAPHIC table. | MSCDM v4.0 | All PATIDs should be present in the  DEMOGRAPHC table. |
| ORG\_PATID | RDBMS  Text(x) |  | . | Corresponds to PATID in the originating CDM if receiving data from multiple managing organizations. | PCORnet | Can leave null if not combining records for multiple organizations. |
| MANAGING\_ORG | RDBMS  Text(x) |  | . | Organization where record originated | PCORnet |  |
| MRN | RDBMS  Text(x) |  |  | Local Medical Record Number for the  patient | PCORnet |  |
| PAT\_FIRSTNAME | RDBMS  Text(x) |  | . | Given name of the patient. | PCORnet |  |
| PAT\_MIDDLENAME | RDBMS  Text(x) |  |  | Middle name of the patient. | PCORnet |  |
| PAT\_LASTNAME | RDBMS  Text(x) |  | . | Surname or family name. | PCORnet |  |
| PAT\_MAIDENNAME | RDBMS  Text(x) |  | . | Surname or family name prior to marriage. | PCORnet |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **PRIVATE\_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* | *Field-level*  *Implementation Guidance* |
| PAT\_SSN | RDBMS  Text(x) |  | . | Social Security Number (SSN; whole or part). If only part of the SSN is available, replace digits with the letter "X" (e.g. XXX- XX-8453). | FHIR-managing Organization | * Hyphens can be included. |
| BIRTH\_DATE | RDBMS Date | SAS Date (Numeric) | . | Date of birth. Corresponds to BIRTH\_DATE in the DEMOGRAPHIC  table. | MSCDM v4.0 |  |
| 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. Corresponds to BIRTH\_TIME in the DEMOGRAPHIC  table. | PCORnet  Source of time format: ISO 8601 |  |
| PRIMARY\_EMAIL | RDBMS  Text(x) |  | . | Primary e-mail address for the patient. | PCORnet |  |
| PRIMARY\_PHONE | RDBMS  Text(10) |  | . | Primary phone number for the patient (if known). 10-digit US phone number. | PCORnet. | * Remove punctuation (parentheses, hyphens). * Cell phone preferred over home phone. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **PRIVATE\_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* | *Field-level*  *Implementation Guidance* |
| SEX | RDBMS  Text(2) | SAS Char(2) | A=Ambiguous F=Female M=Male NI=No  information UN=Unknown OT=Other | Sex assigned at birth. Corresponds to SEX in the DEMOGRAPHIC table. | MSCDM v4.0  with modified field size and value set  Source: Administrative Sex (HL7)  [http://phinvads.c](http://phinvads.cdc.gov/vads/ViewValueSet.action?id=06D34BBC-617F-DD11-B38D-00188B398520) [dc.gov/vads/Vie](http://phinvads.cdc.gov/vads/ViewValueSet.action?id=06D34BBC-617F-DD11-B38D-00188B398520) [wValueSet.actio](http://phinvads.cdc.gov/vads/ViewValueSet.action?id=06D34BBC-617F-DD11-B38D-00188B398520) [n?id=06D34BB](http://phinvads.cdc.gov/vads/ViewValueSet.action?id=06D34BBC-617F-DD11-B38D-00188B398520) [C-617F-DD11-](http://phinvads.cdc.gov/vads/ViewValueSet.action?id=06D34BBC-617F-DD11-B38D-00188B398520) [B38D-](http://phinvads.cdc.gov/vads/ViewValueSet.action?id=06D34BBC-617F-DD11-B38D-00188B398520)  [00188B398520](http://phinvads.cdc.gov/vads/ViewValueSet.action?id=06D34BBC-617F-DD11-B38D-00188B398520) | The “Ambiguous” category may be used for individuals who are physically undifferentiated from birth. The “Other” category may be used for individuals who are undergoing gender re- assignment. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **PRIVATE\_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* | *Field-level*  *Implementation Guidance* |
| 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. Corresponds to SEXUAL\_ORIENTATION in the DEMOGRAPHIC table. | Source: Health IT Certification Criteria, 2015 Base Edition, modified with expert advisory within PCORnet  [https://www.fede](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base) [ralregister.gov/d](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base) [ocuments/2015/1](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base) [0/16/2015-](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base)  [25597/2015-](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base)  [edition-health-](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base) [information-](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base) [technology-](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base) [health-it-](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base) [certification-](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base) [criteria-2015-](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base)  [edition-base](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base) |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **PRIVATE\_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* | *Field-level*  *Implementation Guidance* |
| 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/ Non-binary SE=Something else  MU=Multiple gender categories DC=Decline to answer  NI=No  information UN=Unknown OT=Other | Current gender identity. Corresponds to GENDER\_IDENTITY in the DEMOGRAPHIC table. | Source: Health IT Certification Criteria, 2015 Base Edition, modified with expert advisory within PCORnet  [https://www.fede](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base) [ralregister.gov/d](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base) [ocuments/2015/1](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base) [0/16/2015-](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base)  [25597/2015-](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base)  [edition-health-](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base) [information-](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base) [technology-](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base) [health-it-](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base) [certification-](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base) [criteria-2015-](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base) [edition-base](https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base) | Use Genderqueer (“GQ”) for patients who report a non- binary gender identify. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **PRIVATE\_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* | *Field-level*  *Implementation Guidance* |
| 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.  Corresponds to HISPANIC in the DEMOGRAPHIC table. | MSCDM v4.0  with modified field size and value set  Compatible with “OMB Hispanic Ethnicity” (Hispanic or Latino, Not Hispanic or Latino) | **Populating RACE and HISPANIC if race and ethnicity are not captured separately within the source system** (e.g., “Hispanic or Latino” is included as a selection under Race) - for patients with a known race (e.g., Race is something other than “Hispanic or Latino”, partners should set HISPANIC to "OT" and  RACE to the appropriate race code. For patients who are listed as having a race of “Hispanic,” partners should set HISPANIC to "Y" and RACE to "OT". In this situation, the combined race/ethnicity field is treated as known field capturing values for both race and ethnicity, which is why the preference is to  use “OT” instead of “NI”. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **PRIVATE\_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* | *Field-level*  *Implementation Guidance* |
| 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. Corresponds to RACE in the DEMOGRAPHIC table.  **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. | 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 |  |
|  |  |  |  | White: A person having origins in any of the original peoples of Europe, the Middle East, or North Africa. |  |
| 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). | PCORnet | This field is a derived attribute and is not expected to be an explicit data field within a source system |
|  |  |  |  | If no known biobanked specimens are available, this field should be marked “No”. |  |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **PRIVATE\_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* | *Field-level*  *Implementation Guidance* |
| PAT\_PREF\_LANGUAGE\_SPOKEN | RDBMS  Text(3) | SAS Char(3) | See Value Set Appendix for a list of acceptable values. | Preferred spoken language of communication as expressed by the patient. Corresponds to PAT\_PREF\_LANGUAGE\_SPOKEN in the DEMOGRAPHIC table. | ISO639-2 | * This information may be documented in the EHR or an external registry. * Do not impute or derive if not expressly defined in the source system. * Analytically, will assume that “NI” corresponds to a preferred language of English. * Use the value of “ZHO” (Chinese) for both Mandarin and Cantonese, and place the specific value in the RAW field. Within the ISO639-2 value set, there is no   distinction between the two. [https://www.loc.gov/st](https://www.loc.gov/standards/iso639-2/faq.html#24) [andards/iso639-](https://www.loc.gov/standards/iso639-2/faq.html#24) [2/faq.html#24](https://www.loc.gov/standards/iso639-2/faq.html#24)   * Use “OT” for American Sign Language |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **PRIVATE\_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* | *Field-level*  *Implementation Guidance* |
| RAW\_PAT\_NAME | RDBMS  Text(x) |  | . | Full name of the patient, as represented in the source. | PCORnet | * If components of patient name are not stored as separate fields, can be used to store full name before parsing into first/middle/last. * If components of name are stored separately, they can be combined into a single string in this field (first middle last). |
| RAW\_SEX | RDBMS  Text(x) | SAS Char(x) | . | 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) | . | 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) | . | Field for originating value, prior to mapping  into the PCORnet CDM value set. | PCORnet |  |
| RAW\_HISPANIC | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, prior to mapping  into the PCORnet CDM value set. | PCORnet |  |
| RAW\_RACE | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, prior to mapping  into the PCORnet CDM value set. | PCORnet |  |
| RAW\_PAT\_PREF\_LANGUAGE\_SP  OKEN | RDBMS  Text(x) | SAS Char(x) | . | Field for originating value, prior to mapping  into the PCORnet CDM value set. | PCORnet |  |

|  |
| --- |
| **6.2. Supplemental Table: PRIVATE\_ADDRESS\_HISTORY** |

**PRIVATE\_ADDRESS\_HISTORY Domain Description:** Protected table that can be used to store elements of a patient’s address that are considered personal health information (PHI). These data can be used for geocoding or other linkage projects.

**Relational Integrity:**

The PRIVATE\_ADDRESS\_HISTORY table can contain many records per patient.

**Primary Key:** ADDRESSID

**Foreign Key**:

PRIVATE\_ADDRESS\_HISTORY.PATID is a foreign key to DEMOGRAPHIC.PATID (many-to-one relationship)

**Constraints:**

ADDRESSID (unique; required, not null) PATID (required, not null) ADDRESS\_USE (required, not null) ADDRESS\_TYPE (required, not null)

ADDRESS\_PREFERRED (required, not null)

|  |
| --- |
| **PRIVATE\_ADDRESS\_HISTORY Table Implementation Guidance** |
| *Guidance* |
| * ~~Tokens will be generated in part from the patient’s current home address.~~ The only address-related information used by the current configuration of the PCORnet privacy-preserving record linkage solution is a patient’s current zip code. It is not necessary to create and populate this table solely for that field. However, partners who are participating in the PCORnet geographic surveillance pilots or those who wish to have their patient populations included in any future geographic characterizations of PCORnet will need to populate LDS\_ADDRESS\_HISTORY. In those situations, there may be a benefit in using PRIVATE\_ADDRESS\_HISTORY as a local staging table. * Expect multiple records per individual. * This table is currently limited to addresses in the United States. * Partners can limit records in this table to validated addresses if known. * If partners have difficulty constructing a longitudinal address history for patients within their DataMart, they should prioritize populating the current address for each patient. * Geocoded information about the address should be stored in PRIVATE\_ADDRESS\_GEOCODE. * This table can be kept logically and physically separate from the rest of the CDM. The LDS\_ADDRESS\_HISTORY table represents a subset of this table. It removes the fields that are considered personal health information under HIPAA (ADDRESS\_STREET, ADDRESS\_DETAIL, RAW\_ADDRESS\_TEXT). * This table replicates the fields of the LDS\_ADDRESS\_HISTORY table. They are included in case partners wish to create a master table that can be used to populate LDS\_ADDRESS\_HISTORY within the CDM. SAS data types are included for the replicated fields to help with ETL programming, but partners are not expected to create a SAS-based version of this PRIVATE table. It is not necessary to populate the replicated fields in this table if LDS\_ADDRESS\_HISTORY is loaded through a separate ETL process. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **PRIVATE\_ADDRESS\_HISTORY Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level Implementation*  *Guidance* |
| ADDRESSID | RDBMS  Text(x) | SAS Char(x) | . | Arbitrary identifier for each unique address  record. | PCORnet |  |
| PATID | RDBMS  Text(x) | SAS Char(x) | . | Arbitrary person-level identifier. Used to link across tables. | MSCDM v4.0 | All PATIDs should be present in the DEMOGRAPHIC table. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **PRIVATE\_ADDRESS\_HISTORY Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level*  *Implementation Guidance* |
| ADDRESS\_USE | RDBMS  Text(2) | SAS Char(2) | HO=Home WO=Work TP=Temp OL=Old/Incorrect NI=No  information UN=Unknown OT=Other | Purpose of the address.  Details of categorical definitions:  Home: A communication address at home. Work: An office address. First choice for business-related contacts during business hours.  Temp: A temporary address.  Old/Incorrect: This address is no longer in use (or was never correct but retained for records). | FHIR - ADDRESS | * This field may be a derived attribute that is not an explicit data field within a source system. * Use the period start/end fields to indicate if an address is no longer valid. **Do not change values of HO/WO/TP to OL if a new address is available**. * The old/incorrect value is included in case partners are doing a bulk load and it is present in their source system. It is acceptable to exclude these records, however. * If addresses within the source system are reasonably expected to represent the patient’s home address, it is acceptable to mark these as “HO.” |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **PRIVATE\_ADDRESS\_HISTORY Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level*  *Implementation Guidance* |
| ADDRESS\_TYPE | RDBMS  Text(2) | SAS Char(2) | PO=Postal PH=Physical BO=Both NI=No  information UN=Unknown OT=Other | Type of address.  Details of categorical definitions:  Postal: mailing address – PO Boxes and care- of addresses.  Physical: A physical address that can be visited.  Both: An address that is both physical and postal. | FHIR - ADDRESS | * This field may be a derived attribute that is not an explicit data field within a source system. * Addresses that are clearly not postal-only addresses can be marked as “Both” |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ADDRESS\_PREFERRED | RDBMS  Text(2) | SAS Char(2) | Y=Yes N=No | Indicates whether this address is the preferred one for a given patient, address use and address type within a given address period. | PCORnet | * This field is a derived attribute and is not expected to be an explicit data field within a source system. * It is possible to have an address marked as preferred for every address type/use within each defined address period. Partners do not need to set the flag back to “N” every time the address is updated. * If only one address is present for a given period, that address can be marked as preferred. * If multiple addresses are present for a period, one should be denoted as preferred. * If multiple addresses are present for a period without clear institutional guidance as to which is preferred, partners can use a heuristic to make a determination (i.e., address associated   with most recent |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **PRIVATE\_ADDRESS\_HISTORY Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level*  *Implementation Guidance* |
|  |  |  |  |  |  | encounter, address used for billing, etc.) |
| ADDRESS\_STREET | RDBMS  Text(x) |  | . | Primary address line (e.g., street name and number) | PCORnet | If possible, standardize according to US Post  Office conventions. |
| ADDRESS\_DETAIL | RDBMS  Text(x) |  | . | Remaining address details (e.g., suite, post office box, other details) |  | * Information in this field may be used for deduplication/linkage of people who live in the same building. * Contents are unlikely to add additional information when geocoding. |
| ADDRESS\_COUNTY | RDBMS  Text(x) |  | . | The name of the county associated with the  address. | PCORnet |  |
| ADDRESS\_CITY | RDBMS  Text(x) | SAS Char(x) | . | The name of the city, town, village or other  community. | FHIR -  ADDRESS |  |
| ADDRESS\_STATE | RDBMS  Text(2) | SAS Char(2) | See Value Set Appendix for a list of acceptable  values. | State, as represented by 2-digit postal abbreviation. | PCORnet |  |
| ADDRESS\_ZIP5 | RDBMS  Text(5) | SAS Char(5) | . | 5-digit postal code for the address. | FHIR -  ADDRESS |  |
| ADDRESS\_ZIP9 | RDBMS  Text(9) | SAS Char(9) | . | 9-digit postal code for the address. |  | Do not include  hyphens. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **PRIVATE\_ADDRESS\_HISTORY Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level*  *Implementation Guidance* |
| ADDRESS\_PERIOD\_START | RDBMS Date | SAS Date (Numeric) | . | Initial date when the address known to be in use. | FHIR - ADDRESS | If the date the address is known to be in use is unknown, it is acceptable to use the date the record was created in the source system or the date the record was first created in  the CDM. |
| ADDRESS\_PERIOD\_END | RDBMS Date | SAS Date (Numeric) | . | Date when address was no longer in use. | FHIR - ADDRESS | Only the current address period should have a null value. All previous periods are expected to  have a defined end date. |
| RAW\_ADDRESS\_TEXT | RDBMS  Text(x) |  | . | Text representation of the address | FHIR - ADDRESS | The full address can be represented in this field and may contain information beyond the  street and the PO Box. |

|  |
| --- |
| **6.3. Supplemental Table: PRIVATE\_ADDRESS\_GEOCODE** |

**PRIVATE\_ADDRESS\_GEOCODE Domain Description:** Protected table that can be used to store the results of geocoding algorithms.

**Relational Integrity:**

The PRIVATE\_ADDRESS\_GEOCODE table can contain many records per ADDDRESSID.

**Primary Key:** GEOCODEID

**Foreign Key**:

PRIVATE\_ADDRESS\_GEOCODE.ADDRESSID is a foreign key to PRIVATE\_ADDRESS.ADDRESSID (many-to-one relationship)

**Constraints:**

GEOCODEID (unique; required, not null) ADDRESSID (required, not null)

|  |
| --- |
| **PRIVATE\_ADDRESS\_GEOCODE Table Implementation Guidance** |
| *Guidance* |
| * This table can contain more than one record per address. * This table is designed to accommodate documentation of algorithmic linkage. * This table can be kept logically and physically separate from the rest of the CDM. * There are no implementation expectations related to this table at this time. Future PCORnet projects may reference it in some fashion, but in the meantime, partners may utilize it if they require geocoding for any of their local analyses. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **PRIVATE\_ADDRESS\_GEOCODE Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level Implementation*  *Guidance* |
| GEOCODEID | RDBMS  Text(x) |  | . | Arbitrary identifier for each unique geocode  record. | PCORnet |  |
| ADDRESSID | RDBMS  Text(x) |  | . | Arbitrary identifier for each unique address record. | PCORnet | All ADDRESSIDs should be in the PRIVATE\_ADDRESS\_  HISTORY table. |
| GEOCODE\_STATE | RDBMS  Text(2) |  | See Value Set Appendix for a list of acceptable  values. | State associated with the address. | PCORnet |  |
| GEOCODE\_COUNTY | RDBMS  Text(x) |  | . | County associated with the address. | PCORnet |  |
| GEOCODE\_LONGITUDE | RDBMS  Text(x) |  | . | Longitude assigned to the address by the geocoding algorithm. WGS84 format. | FHIR –  Location.positi on.longitude |  |
| GEOCODE\_LATITUDE | RDBMS  Text(x) |  | . | Latitude assigned to the address by the geocoding algorithm. WGS84 format. | FHIR –  Location.positi on.latitude |  |
| GEOCODE\_BLOCK | RDBMS  Text(x) |  | . | Census Block associated with the address. | PCORnet |  |
| GEOCODE\_TRACT | RDBMS  Text(x) |  | . | Census Tract associated with the address. | PCORnet |  |
| GEOCODE\_GROUP | RDBMS  Text(x) |  | . | Census Block Group associated with the  address. | PCORnet |  |
| GEOCODE\_ZCTA | RDBMS  Text(x) |  | . | Zip code tabulation area associated with the  address. | PCORnet |  |
| GEOCODE\_CUSTOM | RDBMS  Text(x) |  | . | Field that can be used to store a custom/local  geocode for a given address. | PCORnet |  |
| GEOCODE\_CUSTOM\_TEXT | RDBMS  Text(x) |  | . | Text description of the terminology/approach  used to assign the custom geocode. | PCORnet |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **PRIVATE\_ADDRESS\_GEOCODE Table Specification** | | | | | | |
| *Field Name* | *RDBMS Data Type* | *SAS Data Type* | *Predefined Value Sets and Descriptive Text for Categorical Fields* | *Definition / Comments* | *Data Element Provenance* | *Field-level*  *Implementation Guidance* |
| SHAPEFILE | RDBMS  Text(x) |  | . | Name of shapefile used in the geocoding process. | PCORnet | * Include year along with file name. |
| GEO\_ACCURACY | RDBMS  Text(2) |  | Z9=ZIP9 Z5=ZIP5  CN=County CY=City ST=State TR=Tract BL=Block SR=Street NI=No  information UN=Unknown OT=Other | Level of accuracy of the geocoded address based on the method used.  ZIP9: 9-digit zip code ZIP5: 5-digit zip code County: County level  City: City, village or community State: State  Tract: Census tract Block: Census block Street: Street | PCORnet | * The geocoding software should provide this information after addresses are geocoded. * Important when merging data collected/aggregated at different levels. |
| GEO\_PROV\_REF | RDBMS  Text(x) |  | . | Reference to the methodology/software and  parameters used to assign the geocode. | PCORnet |  |
| ASSIGNMENT\_DATE | RDBMS Date |  | . | Date that the geocoding was completed. | PCORnet |  |