**Sentinel Common Data Model**

Version 7.1.0

Prepared by the Sentinel Operations Center (SOC)

Department of Population Medicine,

Harvard Pilgrim Health Care Institute and Harvard Medical School, Boston, MA

December 28, 2020

**Sentinel Common Data Model**

Version 7.1.0

Table of Contents

[1 Overview and Description of the Common Data Model v7.1.0. 1](#_Toc60996771)

[1.1 Sentinel Common Data Model 1](#_Toc60996772)

[2 SCDM: Enrollment Table Structure 2](#_Toc60996773)

[3 SCDM: Demographic Table Structure 4](#_Toc60996774)

[4 SCDM: Dispensing Table Structure 6](#_Toc60996775)

[5 SCDM: Encounter Table Structure 8](#_Toc60996776)

[6 SCDM: Diagnosis Table Structure 13](#_Toc60996777)

[7 SCDM: Procedure Table Structure 17](#_Toc60996778)

[8 SCDM: Laboratory Result Table Structure 20](#_Toc60996779)

[9 SCDM: Vital Signs Table Structure 36](#_Toc60996780)

[10 SCDM: Death Table Structure 40](#_Toc60996781)

[11 SCDM: Cause of Death Table Structure 42](#_Toc60996782)

[12 SCDM: Inpatient Pharmacy Table Structure 44](#_Toc60996783)

[13 SCDM: Inpatient Transfusion Table Structure 46](#_Toc60996784)

[14 SCDM: Mother-Infant Linkage Table Structure 49](#_Toc60996785)

**History of Modifications**

|  |  |  |  |
| --- | --- | --- | --- |
| Version | Date | Modification | Author |
| v1.1 | 12/08/10 | Minor text modifications. | SOC |
| v2.0 | 12/29/11 | Added new tables: Laboratory, Vitals, Summary Tables.  Revised definitions of Encounter table variables: Discharge Disposition, Discharge Status. | SOC |
| v2.1\_edits | 08/03/12 | Updated 6.1 Laboratory table with new lab structure; Added 6.1.1 Lab Details, 6.1.2 Examples & Stand. Units, 6.1.3 Additional Info, 6.1.4 LOINC Info, and 7. State Vaccine tables. | SOC |
| v2.2 | 08/21/12 | Updated Laboratory Information for Influenza, including test types, subtypes, LOINCs, and specimen sources. Minor text modifications. | SOC |
| v2.3 | 09/17/12 | Removed NTBAND as a value for MS\_Test\_Name on the 6.1 Laboratory table. Removed NTBAND as a test type on the 6.1.1 Lab Details table. | SOC |
| v2.4\_EDITS | 10/10/12 | Updated Lab tables, MS\_Test\_Sub\_Category values updated, edited LOINCs information. | SOC |
| v3.0 | 09/16/13 | Added updated Lab tables, added Age Groups summary table, updated Summary Tables based on Query Tool documentation (v1.4, Jan 2013), added Incident Summary Tables, changed DX length to support SNOMED codes and PX length to support LOINCs, added allowable value "IN" for the IIS variable in the State Vaccine table, added DX\_CodeType and PX\_CodeType allowable values, replaced all descriptions in the List of Tables and added a Source column, aligned page headers and footers with page margins, adjusted fonts to be consistent, left-aligned all table column headers, changed "Variable Format" column headers to "Variable Type and Length", changed all references of "file" to "table", changed all references of "field" to "variable", added version number to all footers, reformatted text and made other updates. | SOC |
| v4.0 | 12/03/13 | Added Chart variable to Enrollment table and updated table description and notes; Added ZIP and ZIP\_Date variables to Demographic table and updated table description; removed ALP LOINCs "16182-8" and "33063-9" from Laboratory LOINC Info; added ANC LOINCs "26499-4" and "30451-9", added D\_DIMER\_QN LOINC "55449-3", added HGBA1C LOINCs "62388-4" and "718875-9", and added PG\_QN LOINC "2217-0" to Laboratory LOINC Info; removed Specimen\_Source value "NS" from Laboratory Result table; added Specimen\_Source values "NPWASH", NWASH" and "OTHER" to Laboratory Result table for INF\_A, INF\_AB, INF\_B, and INF\_NS tests; changed Specimen\_Source value "NS" to "UNK" in the Laboratory Result table for INF\_A, INF\_AB, INF\_B, and INF\_NS tests; added Specimen\_Source value "UNK" in the Laboratory Result table for PG\_QN and PG\_QL tests. | SOC |
| v5.0 | 07/31/15 | Updated Laboratory Result Table as follows. Added a new variable Fast\_Ind, to indicate whether the test was performed under fasting conditions. Added a new variable Result\_Type to indicate whether lab result is expected to be numeric or character. Removed “QN” and “QL” from allowable MS\_Test\_Name values. New tests (and preliminary guidance) added: CHOL\_HDL = cholesterol high density lipoprotein, CHOL\_LDL = cholesterol low density lipoprotein, CHOL\_TOT = cholesterol total, SODIUM= sodium, TRIG = triglycerides, TSH = thyroid stimulating hormone. Acceptable MS\_TEST\_SUB\_CATEGORY changed; Added: CLC = Calculated, DIRECT = Direct; Removed: FST = Fasting, RAN = Random. Increased Orig\_Result character length from 8 to 50. Increased Orig\_Result\_Unit character length from 11 to 20. Increased MS\_Result\_C character length from 12 to 50. Added acceptable LOINC and Specimen\_Source values for both new and existing tests. Removed acceptable LOINC values for some existing tests. Added new CPT codes for new tests. Corrected typos and examples. Created Sentinel Common Data Model Laboratory Result Table Documentation (v1.0) for comprehensive guidance to populate the Laboratory Result Table. | SOC |
| V5.0.1 | 02/02/16 | Minor text modifications. | SOC |
| V6.0 | 09/13/16 | Added new tables: Inpatient Pharmacy, Inpatient Transfusion. Added Padmit variable to the Diagnosis table (variable only populated by one Data Partner). Minor text modifications. | SOC |
| V6.0.1 | 04/20/17 | Added two additional allowable values for variable Padmit in the Diagnosis table. Changed EncounterID description to "Arbitrary encounter-level identifier. Used to link across the Encounter, Diagnosis, Procedure, Vital Signs, Inpatient Pharmacy, & Inpatient Transfusion tables", for all tables that contain the EncounterID variable. Removed State Vaccine Table tab. The State Vaccine Table is not regularly maintained and updated so is not a part of the SCDM itself. Removed Summary Table (ST) tabs. Summary Tables are derived from SCDM tables and are not considered to be part of the SCDM itself. Modified several examples that were inconsistent with the variable description. Other minor text and formatting modifications. | SOC |
| V6.0.2 | 10/02/17 | Updated Inpatient Pharmacy table description. | SOC |
| V7.0.0 | 10/31/18 | Added new table: Mother-Infant Linkage. | SOC |
| V7.0.1 | 12/31/19 | Updated Table Description for Table 8. Inpatient Transfusion. | SOC |
| V7.1.0 | 7/18/20 | Updated Laboratory Result Table as follows. Added new allowable values to several existing variables in order to include new Covid-19 diagnostic test results. New MS\_Test\_Name value of SARS\_COV\_2 added. New MS\_Test\_Sub\_Category value of IA\_RAP = immunoassay, rapid and SEQ=sequencing added. New Specimen\_Source value of SALIVA=saliva added. | SOC |

# Overview and Description of the Common Data Model v7.1.0.

The primary goal of Sentinel is to build and operate a national public health surveillance system to monitor the safety of FDA-regulated medical products, including drugs, biologics, and devices. Sentinel is part of the Sentinel Initiative, the FDA’s response to a congressional mandate to create an active surveillance system using electronic health data.

The Sentinel program will undertake three major types of activities: (1) prospective evaluation of accumulating experience about specific medical products and specific suspected safety problems; (2) evaluation of the impact of FDA actions (e.g., labeling changes) on medical practice and health outcomes; and (3) rapid assessment of past experience in response to FDA questions about specific exposures and outcomes.

Sentinel Collaborating Institutions maintain data resources and provide technical, scientific, and methodological expertise as needed to meet the public health surveillance requirements of Sentinel. The Collaborating Institutions also participate as members of the Planning Board, the Safety Science Committee, the Sentinel Operations Center (SOC), and various Sentinel workgroups.

## Sentinel Common Data Model

The SOC coordinates the network of Sentinel Data Partners and leads development of the Sentinel Common Data Model (SCDM), a standard data structure that allows Data Partners to quickly execute distributed programs against local data. The SOC manages creation of the Sentinel Distributed Database (SDD) using the SCDM, and maintains complete documentation of the implementation and characteristics of the SDD. The SCDM was developed in accordance with the Mini-Sentinel Common Data Model Guiding Principles and was modeled after the Health Care Systems Research Network (formerly known as HMO Research Network) Virtual Data Warehouse.

The SCDM currently includes 13 tables that represent information for the data elements needed for Sentinel activities. Records are linked across tables by a unique person identifier, PatID. Details of the 13 tables are provided in this document.

This data model is freely available to all.

* For more information about Sentinel visit the website at: [www.sentinelinitiative.org](https://www.sentinelinitiative.org/)
* For comments and suggestions, please email: [info@sentinelsystem.org](file:///C:\repos\DEV-11439\info@sentinelsystem.org)

# SCDM: Enrollment Table Structure

Description: The SCDM Enrollment Table has a start/stop structure that contains one record per continuous enrollment period. Members with medical coverage, drug coverage, or both should be included. A unique combination of PatID, Enr\_Start, Enr\_End, MedCov, DrugCov, and Chart identifies a unique record. A break in enrollment (of at least one day) or a change in either the medical or drug coverage variables should generate a new record.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable Name | Variable Type and Length (Bytes) | Values | Definition / Comments / Guideline | Example |
| PatID1 | Char (Site specific length) | Unique member identifier | Arbitrary person-level identifier. Used to link across tables. A new enrollment period generates a new record, but the same person should have the same PatID on subsequent records. | 123456789012345 |
| Enr\_Start2 | Numeric (4) | SAS date | Date of the beginning of the enrollment period. If the exact date is unknown, use the first day of the month. Enr\_Start should not be before January 1, 2000. | 1/1/2005 |
| Enr\_End2,3 | Numeric (4) | SAS date | Date of the end of the enrollment period. If the exact date is unknown, use the last day of the month. | 12/31/2005 |
| MedCov | Char (1) | Y = Yes  N = No  U = Unknown | Mark as “Y” if 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). | Y |
| DrugCov | Char (1) | Y = Yes  N = No  U = Unknown | Mark as “Y” if 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). | Y |
| Chart4 | Char (1) | Y = Yes  N = No | Chart abstraction flag to answer the question, “Are you able to request charts for this member?” This flag does not address chart availability. Mark as “Y” if there are no contractual restrictions between you and the member (or sponsor) that would prohibit you from requesting any chart for this member. | Y |

NOTES:

1. 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.
2. Adjacent and overlapping enrollment periods with the same PatID, Enr\_Start, Enr\_End, MedCov, DrugCov, and Chart values should be collapsed. Enrollment periods separated by more than one day should not be bridged. For example, an Enr\_End date of 1/31/2005 should be bridged with an Enr\_Start date of 2/1/2005, but should not be bridged with an Enr\_Start date of 2/2/2005.
3. Enr\_End should not be imputed using the date of death found in the Death table.
4. Chart variable aims to identify enrollment periods for which medical charts cannot be requested. Potential scenarios include:
   1. Charts cannot be requested for Medicare members (all enrollment periods for Medicare members should be assigned Chart='N')
   2. Charts cannot be requested for administrative services only (ASO) populations (all ASO enrollment periods should be assigned Chart='N')

* If there is no definitive information indicating that medical charts cannot be requested for member enrollment period(s), records should be assigned Chart = 'Y'.

# SCDM: Demographic Table Structure

Description: The SCDM Demographic Table contains one record per PatID with the most recent information on Birth\_Date, Sex, Race/Ethnicity, and Zip Code.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable Name | Variable Type and Length (Bytes) | Values | Definition / Comments / Guideline | Example |
| PatID1 | Char (Site specific length) | Unique member identifier | Arbitrary person-level identifier. Used to link across tables. | 123456789012345 |
| Birth\_Date | Numeric (4) | SAS date | Date of birth. | 12/5/1971 |
| Sex | Char (1) | A = Ambiguous (e.g., transgender/hermaphrodite)  F = FemaleM = Male  U = Unknown | Sex. | F |
| Hispanic | Char (1) | N = No  Y = Yes  U = Unknown | A person of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin, regardless of race. | N |
| Race | Char (1) | 0 = Unknown | Please use only one race value per member. | 2 |
|  |  | 1 = 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. |  |
|  |  | 2 = 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. |  |
|  |  | 3 = Black or African American | A person having origins in any of the black racial groups of Africa. |  |
|  |  | 4 = Native Hawaiian or Other Pacific Islander | A person having origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands. |  |
|  |  | 5 = White | A person having origins in any of the original peoples of Europe, the Middle East, or North Africa. |  |
| Zip | Char (5) | Zip code | First 5 digits of the ZIP code of the member’s most recent primary residence. | 04090 |
| Zip\_Date | Numeric (4) | SAS date | Earliest date that the ZIP code is believed to be valid. Date will be updated/overwritten as ZIP code changes over time. | 12/12/2009 |

NOTES:

1. 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.

# SCDM: Dispensing Table Structure

Description: The SCDM Outpatient Pharmacy Dispensing Table contains one record per unique combination of PatID, NDC, and RxDate. Each record represents an outpatient pharmacy dispensing. Rollback transactions and other adjustments should be processed before populating this table.1,2

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable Name | Variable Type and Length (Bytes) | Values | Definition / Comments / Guideline | Example |
| PatID3 | Char (Site specific length) | Unique member identifier | Arbitrary person-level identifier. Used to link across tables. | 123456789012345 |
| RxDate | Numeric (4) | SAS date | Dispensing date (as close as possible to date the person received the dispensing). | 11/29/2005 |
| NDC | Char (11) | National Drug Code | Please expunge any place holders (e.g., ‘-’ or extra digit). | 00006007431 |
| RxSup2 | Numeric (4) | 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 dispensings record. It should not be necessary to calculate this variable for use in the SCDM. Positive integer values are expected. | 30 |
| RxAmt2 | Numeric (4) | Amount dispensed | Number of units (pills, tablets, vials) dispensed. Net amount per NDC per dispensing. This amount is typically found on the dispensings record. It should not be necessary to calculate this variable for use in the SCDM. Positive values are expected. | 60 |

NOTES:

1. Medications distributed in other settings such as infusions given in medical practices or inpatient hospitals are captured in the utilization tables. Medication prescriptions (as opposed to dispensings) are not currently captured in the SCDM.
2. Rollback transactions and other adjustments that are indicative of a dispensing being canceled or not picked up by the member should be processed before populating this table. This may be handled differently by Data Partners and may be affected by billing cycles.
3. 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.

# SCDM: Encounter Table Structure

Description: The SCDM Encounter Table contains one record per PatID and EncounterID. Each encounter should have a single record in the SCDM Encounter Table. Each diagnosis and procedure recorded during the encounter should have a separate record in the Diagnosis or Procedure Tables. Multiple visits to the same provider on the same day should be considered one encounter and should include all diagnoses and procedures that were recorded during those visits. Visits to different providers on the same day, such as a physician appointment that leads to a hospitalization, should be considered multiple encounters. Rollback transactions and other adjustments should be processed before populating this table.1

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable Name | Variable Type and Length (Bytes) | Values | Definition / Comments / Guideline | Example |
| PatID2 | Char (Site specific length) | Unique member identifier | Arbitrary person-level identifier. Used to link across tables. | 123456789012345 |
| EncounterID3 | Char (Site specific length) | Unique encounter identifier | Arbitrary encounter-level identifier. Used to link across the Encounter, Diagnosis, Procedure, Vital Signs, Inpatient Pharmacy, & Inpatient Transfusion tables. | 123456789012345\_12242005\_99218766\_IP |
| ADate | Numeric (4) | SAS date | Encounter or admission date. | 12/24/2005 |
| DDate | Numeric (4) | SAS date | Discharge date. Should be populated for all Inpatient Hospital Stay (IP) and Non-Acute Institutional Stay (IS) encounter types. May be populated for Emergency Department (ED) encounter types. Should be missing for ambulatory visit (AV or OA) encounter types. | 12/31/2005 |
| Provider4 | Char (Site specific length) | Unique provider identifier | Provider code for the provider who is most responsible for this encounter. For encounters with multiple providers choose **one** so the encounter can be linked to the diagnosis and procedure tables. As with the PatID, the provider code is a pseudoidentifier with a consistent crosswalk to the real identifier. | 99218766 |
| Facility\_Location | Char (3) | Geographic location (3 digit zip code) | Should be left blank if missing. | 902 |
| EncType | Char (2) | AV = Ambulatory Visit | Includes visits at outpatient clinics, same day surgeries, urgent care visits, and other same-day ambulatory hospital encounters, but excludes emergency department encounters. | IP |
|  |  | ED = Emergency Department | Includes Emergency Department encounters that become inpatient stays (in which case inpatient stays would be a separate encounter). Excludes urgent care visits. Emergency Department claims should be pulled before hospitalization claims to ensure that ED with subsequent admission won’t be rolled up in the hospital event. |  |
|  |  | IP = Inpatient Hospital Stay | Includes all inpatient stays, same-day hospital discharges, hospital transfers, and acute hospital care where the discharge is after the admission date. |  |
|  |  | IS = Non-Acute Institutional Stay | Includes hospice, skilled nursing facility (SNF), rehab center, nursing home, residential, overnight non-hospital dialysis and other non-hospital stays. |  |
|  |  | OA = Other Ambulatory Visit | Includes other non overnight AV encounters such as hospice visits, home health visits, skilled nursing facility visits, other non-hospital visits, as well as telemedicine, telephone and email consultations. |  |
| Facility\_Code | Char (Site specific length) | Servicing provider identifier | Local facility code that identifies hospital or clinic. Taken from facility claims. Used for chart abstraction and validation. | FC12345678 |
| Discharge\_ Disposition | Char (1) | A = Discharged aliveE = ExpiredU = Unknown | Should be populated for Inpatient Hospital Stay (IP) and Non-Acute Institutional Stay (IS) encounter types. May be populated for Emergency Department (ED) encounter types. Should be missing for ambulatory visit (AV or OA) encounter types. | A |
| Discharge\_Status | 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)  OT = Other  RH = Rehabilitation Facility  RS = Residential Facility  SH = Still In Hospital  SN = Skilled Nursing Facility  UN = Unknown | Should be populated for Inpatient Hospital Stay (IP) and Non-Acute Institutional Stay (IS) encounter types. May be populated for Emergency Department (ED) encounter types. Should be missing for ambulatory visit (AV or OA) encounter types. | SN |
| DRG | Char (3) | 3-digit Diagnosis Related Group | Diagnosis Related Group. Should be populated for IP and IS encounter types. May be populated for ED encounter types. Should be missing for AV or OA encounters. Use leading zeroes for codes less than 100. | 372 |
| DRG\_Type | Char (1) | 1 = CMS-DRG (old system)  2 = MS-DRG (current system) | DRG code version. MS-DRG (current system) began on 10/1/2007. Should be populated for IP and IS encounter types. May be populated for ED encounter types. Should be missing for AV or OA encounters. | 1 |
| Admitting\_Source | 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)  OT = Other  RH = Rehabilitation Facility  RS = Residential Facility  SN = Skilled Nursing Facility  UN = Unknown | Should be populated for Inpatient Hospital Stay (IP) and Non-Acute Institutional Stay (IS) encounter types. May be populated for Emergency Department (ED) encounter types. Should be missing for ambulatory visit (AV or OA) encounter types. | HH |

NOTES:

1. Rollback transactions and other adjustments should be processed before populating this table. This may be handled differently by Data Partners and may be affected by billing cycles.
2. 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.
3. Medical utilization data is captured in 3 tables:
   * Encounter: the encounter record that characterizes the outpatient visit or hospital stay
   * Diagnosis: the diagnosis or other clinical code(s) associated with the encounter record
   * Procedure: the procedure code(s) associated with the encounter record.

* These 3 tables and the Inpatient Pharmacy, Inpatient Transfusion, Vital Signs, and Mother-Infant Linkage tables are linked by EncounterID. All diagnoses and procedures for an encounter should have the same EncounterID. It is allowable to have “orphan” diagnosis or procedure records with EncounterIDs that do not have a match in the Encounter table.

1. The provider variable must be consistent within a health plan. An inpatient stay must only have one Provider, even if multiple providers performed procedures.

# SCDM: Diagnosis Table Structure

Description: The SCDM Diagnosis Table contains one record per unique combination of PatID, EncounterID, DX, and DX\_CodeType. This table should capture all uniquely recorded diagnoses for all encounters.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable Name | Variable Type and Length (Bytes) | Values | Definition / Comments / Guideline | Example |
| PatID1 | Char (Site specific length) | Unique member identifier | Arbitrary person-level identifier. Used to link across tables. | 123456789012345 |
| EncounterID2 | Char (Site specific length) | Unique encounter identifier | Arbitrary encounter-level identifier. Used to link across the Encounter, Diagnosis, Procedure, Vital Signs, Inpatient Pharmacy, & Inpatient Transfusion tables. | 123456789012345\_12242005\_99218766\_IP |
| ADate | Numeric (4) | SAS date | Encounter or admission date. | 12/24/2005 |
| Provider | Char (Site specific length) | Unique provider identifier | Provider code for the provider who is most responsible for this encounter. For encounters with multiple providers choose one so the encounter can be linked to the diagnosis and procedure tables. As with the PatID, the provider code is a pseudoidentifier with a consistent crosswalk to the real identifier. | 99218766 |
| EncType | Char (2) | AV = Ambulatory Visit | Includes visits at outpatient clinics, same day surgeries, urgent care visits, and other same-day ambulatory hospital encounters, but excludes emergency department encounters. | IP |
|  |  | ED = Emergency Department | Includes ED encounters that become inpatient stays (in which case inpatient stays would be a separate encounter). Excludes urgent care 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. |  |
|  |  | IP = Inpatient Hospital Stay | Includes all inpatient stays, same-day hospital discharges, hospital transfers, and acute hospital care where the discharge is after the admission date. |  |
|  |  | IS = Non-Acute Institutional Stay | Includes hospice, skilled nursing facility (SNF), rehab center, nursing home, residential, overnight non-hospital dialysis and other non-hospital stays. |  |
|  |  | OA = Other Ambulatory Visit | Includes other non overnight AV encounters such as hospice visits, home health visits, skilled nursing facility visits, other non-hospital visits, as well as telemedicine, telephone and email consultations. |  |
| DX3 | Char (18) | Diagnosis code | For ICD codes this variable can include decimal points or not. Remove site specific suffixes and prefixes. Other codes should be listed as recorded in the source data. | 761.5 |
| Dx\_Codetype4 | Char (2) | 09 = ICD-9-CM10 = ICD-10-CM11 = ICD-11-CMSM = SNOMED CTOT = Other | Diagnosis code type. This field combined with the DX field should be used to capture any type of diagnosis or clinical concept available in the source data. We provide values for ICD and SNOMED code types. Other code types will be added as new terminologies are used. | 09 |
| OrigDX | Char (Site specific length) | Original diagnosis from source table, if different | Used if Data Partner has to map internal codes to standard codes. |  |
| PDX | Char (1) | P = PrincipalS = SecondaryX = Unable to Classify | Principal discharge diagnosis flag. Relevant only on IP and IS encounters. For ED, AV, and OA encounter types, mark as missing. One principal diagnosis is expected, although in some instances more than one diagnosis may be flagged as principal | P |
| PAdmit5 | Char (1) | N = No Y = Yes U = Unknown or unable to determineX = Unreported/not used | Indicates whether the diagnosis code is indicative of a condition present at admission. | Y |

NOTES:

1. 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.
2. For efficiency medical utilization data is captured in 3 tables:
   * Encounter: the encounter record that characterizes the outpatient visit or hospital stay
   * Diagnosis: the diagnosis code(s) associated with the encounter record
   * Procedure: the procedure code(s) associated with the encounter record

* These 3 tables and the Inpatient Pharmacy, Inpatient Transfusion, Vital Signs, and Mother-Infant Linkage tables are linked by EncounterID. All diagnoses and procedures for an encounter should have the same EncounterID. It is allowable to have “orphan” diagnosis or procedure records with EncounterIDs that do not have a match in the Encounter table.

1. For ICD codes, some Data Partners will have a decimal point in the DX variable and others will not. We recommend that users of the data strip the decimal point during data analyses.
2. For those who collect SNOMED CT codes as part of routine care, those codes can be stored in this table, using the “SM” DX\_CodeType
3. Sentinel Data Partners receive individual guidance for populating the PAdmit field.

# SCDM: Procedure Table Structure

Description: The SCDM Procedure Table contains one record per unique combination of PatID, EncounterID, PX, and PX\_CodeType. This table should capture all uniquely recorded procedures for all encounters.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable Name | Variable Type and Length (Bytes) | Values | Definition / Comments / Guideline | Example |
| PatID1 | Char (Site specific length) | Unique member identifier | Arbitrary person-level identifier. Used to link across tables. | 123456789012345 |
| EncounterID2 | Char (Site specific length) | Unique encounter identifier | Arbitrary encounter-level identifier. Used to link across the Encounter, Diagnosis, Procedure, Vital Signs, Inpatient Pharmacy, & Inpatient Transfusion tables. | 123456789012345\_12262005\_99218766\_IP |
| ADate | Numeric (4) | SAS date | Encounter or admission date. | 12/26/2005 |
| Provider | Char (Site specific length) | Unique provider identifier | Provider code for the provider who is most responsible for this encounter. For encounters with multiple providers choose one so the encounter can be linked to the diagnosis and procedure tables. As with the PatID, the provider code is a pseudoidentifier with a consistent crosswalk to the real identifier. | 99218766 |
| EncType | Char (2) | AV = Ambulatory Visit | Includes visits at outpatient clinics, same day surgeries, urgent care visits, and other same-day ambulatory hospital encounters, but excludes emergency department encounters. | IP |
|  |  | ED = Emergency Department | Includes ED encounters that become inpatient stays (in which case inpatient stays would be a separate encounter). Excludes urgent care 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. |  |
|  |  | IP = Inpatient Hospital Stay | Includes all inpatient stays, same-day hospital discharges, hospital transfers, and acute hospital care where the discharge is after the admission date. |  |
|  |  | IS = Non-Acute Institutional Stay | Includes hospice, skilled nursing facility (SNF), rehab center, nursing home, residential, overnight non-hospital dialysis and other non-hospital stays. |  |
|  |  | OA = Other Ambulatory Visit | Includes other non overnight AV encounters such as hospice visits, home health visits, skilled nursing facility visits, other non-hospital visits, as well as telemedicine, telephone and email consultations. |  |
| PX | Char (11) | Procedure code | Convert local codes to standard codes. | 76815 |
| PX\_CodeType | Char (2) | 09 = ICD-9-CM  10 = ICD-10-CM  11 = ICD-11-CM  C2 = CPT Category II  C3 = CPT Category III  C4 = CPT-4 (i.e., HCPCS Level I)  H3 = HCPCS Level III  HC = HCPCS (i.e., HCPCS Level II)  LC = LOINC  LO = Local homegrown  ND = NDC  OT = Other  RE = Revenue | Procedure code type. | C4 |
| OrigPX | Char (Site specific length) | Original procedure code from source table, if different. | Used if Data Partner has to map internal codes to standard codes. |  |

NOTES:

1. 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.
2. For efficiency medical utilization data is captured in 3 tables:
   * Encounter: the encounter record that characterizes the outpatient visit or hospital stay
   * Diagnosis: the diagnosis code(s) associated with the encounter record
   * Procedure: the procedure code(s) associated with the encounter record

* These 3 tables and the Inpatient Pharmacy, Inpatient Transfusion, Vital Signs, and Mother-Infant Linkage tables are linked by EncounterID. All diagnoses and procedures for an encounter should have the same EncounterID. It is allowable to have “orphan” diagnosis or procedure records with EncounterIDs that do not have a match in the Encounter table.

# SCDM: Laboratory Result Table Structure

Description: The SCDM Laboratory Result Table contains one record per result/entry. Only include resulted lab tests.1 Data Partners are strongly encouraged to review the comprehensive Sentinel Common Data Model Laboratory Result Table Documentation for details on how to populate each variable.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable Name | Variable Type and Length (Bytes) | Values | Definition / Comments / Guideline | Example |
| PatID2 | Char (Site specific length) | Unique member identifier | Arbitrary person-level identifier. Populated for all records. Used to link across tables. | 123456789012345 |
| MS\_Test\_Name | Char (10) | ALP = alkaline phosphatase  ALT = alanine aminotransferase  ANC = absolute neutrophil count  BILI\_TOT = total bilirubin  CHOL\_HDL = cholesterol high density lipoprotein  CHOL\_LDL = cholesterol low density lipoprotein  CHOL\_TOT = cholesterol total  CK = creatine kinase total  CK\_MB = creatine kinase MB  CK\_MBI = creatine kinase MB / creatine kinase total  CREATININE = creatinine  D\_DIMER = d-dimer  GLUCOSE = glucose  HGB = hemoglobin  HGBA1C = glycosylated hemoglobin  INF\_A = influenza virus A  INF\_AB = influenza virus A + B  INF\_B = influenza virus B  INF\_NS = influenza virus not specified  INR = international normalized ratio  LIPASE = lipase  PG = pregnancy test  PLATELETS = platelet count  SARS\_COV\_2= sars\_cov-2 / covid-19  SODIUM = sodium  TRIG = triglycerides  TROP\_I = troponin I cardiac  TROP\_T = troponin T cardiac  TSH = thyroid stimulating hormone | Abbreviated test name. Populated for all records. Several LOINC codes (LOINC) and/or local codes (LOCAL\_CD) can point to one MS\_Test\_Name. | ALP |
| Result\_Type | Char (1) | N = numericC = character | Indicates whether the laboratory test result is numeric (e.g., 100 ug/mL, <100 ug/mL, >100 ug/mL) and MS\_Result\_N will be populated, or character (e.g., for character results such as: +, POS, POSITIVE, and ranges, such as 50-100 mg/mL) and MS\_Result\_C will be populated, based on Data Partner source data. | N |
| MS\_Test\_Sub\_ Category | Char (6) | BHCG = beta human choriogonadotropin  CLC = calculated  DDU = d-dimer units  DIRECT= direct  EIA = enzyme immunoassay  FEU = fibrinogen equivalent units  IA\_RAP = immunoassay, rapid  HCG = human choriogonadotropin  IF = immunofluorescence  NS = not specified  PCR = probe and target amplification  SEQ = sequencing  VTC = organism-specific culture | Sub-category for MS\_Test\_Name. Sub-categories apply to only select laboratory tests. DIRECT and CALCULATED is only populated for MS\_Test\_Name = CHOL\_LDL. DDU and FEU is only populated for MS\_Test\_Name = D\_DIMER, Result\_Type = N. BHCG AND HCG is only populated for MS\_Test\_Name = PG. | PCR |
| Fast\_Ind | Char (1) | F= fasting  R= random  X= not applicable | Fasting Indicator. Not all laboratory tests are differentiated by fasting status. For laboratory tests where fasting status is not considered to interpret results, Fast\_Ind is marked X (not applicable). For laboratory test results where fasting status may be considered to interpreting results (e.g., glucose), Fast\_Ind is marked F (fasting) only if the source data included with the laboratory test result indicates fasting test. All other laboratory test results where fasting status may be considered that is not explicitly marked as fasting, Fast\_Ind is marked R (random). | F |
| Specimen\_Source | Char (6) | BAL = bronchoalveolar lavage  BALBX = bronchoalveolar biopsy  BLOOD = blood  CSF = cerebrospinal fluid  NPH = nasopharyngeal swab  NPWASH = nasopharyngeal wash  NSWAB = nasal swab or nose specimen  NWASH = nasal wash  OTHER = other  PLASMA = plasma  PPP = platelet poor plasma  SALIVA = saliva  SERUM = serum  SPUTUM = sputum  SR\_PLS = serum/plasma  THRT = throat swab, oropharyngeal swab  UNK = unknown or null  URINE = urine | Specimen source. Populated for all records. Some laboratory tests have several possible valid values for Specimen\_Source. | SERUM |
| LOINC | Char (10) | LOINC code | Logical Observation Identifiers, Names, and Codes (LOINC) is a universal coding system for tests, measurements, and observations developed by the Regenstrief Institute. LOINC provides very granular information about the tests such as the long name of the test, test definition/description, component, system, property, timing, scale, method, and examples of units. All parts of the LOINC code, including the hyphen, are included. The last digit of the LOINC code is a check digit and is always preceded by a hyphen. If a record utilizes a local LOINC (e.g., LOINC candidate codes), the result of that record is included, however, the LOINC variable is set to null. Values in the LOINC variable do not contain leading zeros. LOINC is populated when available, thus not all records have LOINC. | 16182-8 |
| Stat | Char (1) | E = Expedite  R = Routine  S = Stat  U = Unknown or null | Immediacy of test. The intent of this variable is to determine whether the test was obtained as part of routine care or as an emergent/urgent diagnostic test (designated as Stat or Expedite). | E |
| Pt\_Loc | Char (1) | E = Emergency department  H = Home  I = Inpatient  O = Outpatient  U = Unknown or null | Patient location where the lab specimen was obtained. | O |
| Result\_Loc | Char (1) | L = Lab  P = Point of Care | 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 are not any null values. | L |
| LOCAL\_CD | Char (Site specific length) | Unique to each Data Partner | Local code (non-LOINC) related to an individual lab test. Values for LOCAL\_CD are not required. LOCAL\_CD is only populated if available in source data. This variable will not be used in queries, but may be used by local programmers to identify and extract the required CDM tests. |  |
| BATTERY\_CD | Char (Site specific length) | Unique to each Data Partner | Local code (non-LOINC) related to a battery or panel of lab tests. Values for BATTERY\_CD are not required. BATTERY\_CD is only populated if available in source data. This variable will not be used in queries, but may be used by local programmers to identify and extract the required CDM tests. |  |
| PX | Char (Site specific length) | Procedure code | Procedure Code associated with the laboratory result record. Values for PX are not required. PX is only populated if available in source data. PX is populated if PX\_CodeType is populated. | 76815 |
| PX\_CodeType | Char (2) | 09 = ICD-9-CM  10 = ICD-10-CM  11 = ICD-11-CM  C2 = CPT Category II  C3 = CPT Category III  C4 = CPT-4 (i.e., HCPCS Level I)  H3 = HCPCS Level III  HC = HCPCS (i.e., HCPCS Level II)  LO = Local homegrown  OT = Other  RE = Revenue | Procedure code type associated with the laboratory result record. Values for PX\_CodeType are not required. PX\_CodeType is only populated if available in source data. PX\_CodeType is populated if PX is populated. | C4 |
| Order\_dt3 | Numeric (4) | SAS date | Date that the test was ordered, represented as a SAS date value. SAS format is MMDDYY10. This date could fall anywhere from the same day the specimen was collected to months before the specimen was collected. Please see footnote 3. | 11/29/2009 |
| Lab\_dt3 | Numeric (4) | SAS date | Date that the specimen was collected, represented as a SAS date value. SAS format is MMDDYY10. For most Sentinel activities, this is the most relevant date. Please see footnote 3. | 11/29/2009 |
| Lab\_tm | Numeric (4) | SAS time | Time of day that the specimen was collected, represented as a SAS time value. SAS format is HHMM. Valid values are between 00:00 to 23:59. Please note that Lab\_tm is associated with Lab\_dt. | 18:00 |
| Result\_dt3 | Numeric (4) | SAS date | Date that the laboratory test was resulted, represented as a SAS date value. SAS format is MMDDYY10. Dependent on the time of the test, this date could be the same day the specimen was collected or any date up to weeks later. Please see footnote 3. | 12/1/2009 |
| Result\_tm | Numeric (4) | SAS time | Time that the laboratory test was resulted, represented as a SAS time value. SAS format is HHMM. Valid values are between 00:00 to 23:59. | 11:30 |
| Orig\_Result | Char (50) | Text | Orig\_Result is populated for all records. If Result\_Type equals N, then the Orig\_Result value represents the numeric portion of the test result, stripped of any Modifiers (e.g., >, LE, GT) and/or result units (e.g., ng/ml, cells/mm3, %). If Result\_Type equals C, Orig\_Result value reflects the test string based on source data. Additionally, for records where Result\_Type = C, values may include a decimal point (.), a sign (-, +) or text (e.g., POSITIVE, NEGATIVE, DETECTED). For records where Result\_Type = N, the symbols >, <, >=, <= are removed from the value and stored in the Modifier variable and result units are removed and stored as Orig\_Result\_unit variable. | + |
| MS\_Result\_C | Char (50) | BORDERLINE  NEGATIVE  POSITIVE  UNDETERMINED  RANGE: start\|end[unit] | This standardized result value is only populated for text or character results (Result\_Type = C) and is null for numeric results (Result\_Type = N). If Result\_Type = C and the source result is a range (e.g., 50-100 mg/mL), then populate MS\_Result\_C using Orig\_Result with the start and end values of the range delimited by a vertical bar (e.g., “50-100 mg/mL” becomes “50|100 mg/mL”). | “POSITIVE”"50 |
| MS\_Result\_N | Numeric (8) | Numeric digits with or withouta decimal | Standardized/converted numeric result for records where Result\_Type=N. Acceptable values are numeric digits with or without a decimal (“.”). If the standardized result unit (Std\_Result\_unit) differs from an acceptable MS\_Result\_unit for a numeric test (Result\_Type=N), then Orig\_Result is converted prior to populating the MS\_Result\_N value. Additionally, MS\_Result\_unit reflects this conversion. This variable is only populated for numeric results (Result\_Type = N) and does not contain negative values. This variable is null for text/character results (Result\_Type = C). | 100 |
| Modifier | Char (2) | EQ = equal  GE = greater than or equal to  GT = greater than  LE = less than or equal to  LT = less than  TX = text | Modifier for result values. Any relational operators in the original source data value (e.g., <, >, or = ) are reflected in the Modifier variable. For example, if the original source data value is <=200, then Orig\_Result = 200 and Modifier = LE. If the original source data result value is text, then Modifier = TX. If the original source data result value is numeric (digits with or without decimal) and does not contain an operator, then Modifier = EQ. | LE |
| Orig\_Result\_unit | Char (20) | Text | Original units for the test result, as reported in source data. This variable is directly related to Orig\_Result and Modifier. This variable does not include the test name, or any special characters, unless that character is part of the unit value. For example, special characters are included in “10^9/L”. However, special characters are not included in “U/L”, as these carats are not part of the unit value. Some laboratory tests may not have a result unit. | 10^9/L |
| Std\_Result\_unit | Char (11) | Text | Standardized units for the result. The purpose of Std\_Result\_unit is to modify Orig\_Result\_unit from free text to a standardized unit, as an intermediary step in converting from Orig\_Result\_unit to MS\_Result\_unit. This variable is only populated for records where Result\_Type = N. Common rules and guidelines for populating Std\_Result\_unit, include: converting all text values for Orig\_Result\_unit to uppercase, and using standard abbreviations as provided by SOC. This variable does not include special characters, unless that character is part of the unit. For example, special characters are included in “10^9/L”. However, special characters are not included in “U/L”, as these carats are not part of the unit value. This value is not usually null, though there are exceptions, such as when the test result is a ratio (e.g., International Normalized Ratio *INR*). This value is null for character tests (Result\_Type=C) until that test has been reviewed and characterized by the Clinical Data Elements Workgroup. | CELL/MM3 |
| MS\_Result\_unit | Char (11) | Text | Converted/standardized result units for the value populated in MS\_Result\_N. This value is null for records where Result\_Type = C, and null for Laboratory Tests that have not been characterized and reviewed by the Clinical Data Elements Workgroup. This value may be null for some numeric tests (e.g., as International Normalized Ratio is a ratio, it does not have a result unit). For tests that require a result unit, MS\_Result\_unit is set to UNKNOWN for all records where the original result unit is missing or blank, “NULL”, “N/A”, “NA”, or “UNK”. Guidance is provided for only Laboratory Tests that have been reviewed by the Clinical Data Elements Workgroup. | G/L |
| Norm\_Range\_low | Char (8) | Text | Lower bound of the normal reference range, as assigned by the laboratory. The normal range associated with a test, as assigned by the laboratory is parsed out into the following variables: Norm\_Range\_low, Modifier\_low, Norm\_Range\_high, Modifier\_high, and reflects what is seen in source data. Value only contains the value of the lower bound of the normal reference range. This value is not converted and unit of measure is not included. It is assumed that the associated unit is the same as the original result unit from the source data. The symbols >, <, >=, <= are removed. For example, if the normal range for a test is >100 and <300, then 100 is entered. Additionally, this value is null for records where Result\_Type = C. | 100 |
| Modifier\_low | Char (2) | EQ = equal  GE = greater than or equal to  GT = greater than | Modifier for Norm\_Range\_low values. The normal range associated with a test, as assigned by the laboratory is parsed out into the following variables: Norm\_Range\_low, Modifier\_low, Norm\_Range\_high, Modifier\_high and reflects what is seen in source data. This value is null for records where Result\_Type=C. For numeric results one of the following needs to be true:  Both Modifier\_low and Modifier\_high contain EQ (e.g., normal values fall in the range 3-10).  Modifier\_low contains GT or GE and Modifier\_high is null (e.g., normal values are >3 with no upper boundary).  Modifier\_high contains LT or LE and Modifier\_low is null (e.g., normal values are <=10 with no lower boundary). | EQ |
| Norm\_Range\_high | Char (8) | Text | Upper bound of the normal reference range, as assigned by the laboratory. The normal range associated with a test, as assigned by the laboratory is parsed out into the following variables: Norm\_Range\_low, Modifier\_low, Norm\_Range\_high, Modifier\_high and reflects what is seen in source data. Value only contains the value of the upper bound of the normal reference range. This value is not converted and unit of measure is not included. It is assumed that the associated unit is the same as the original result unit from source data. The symbols >, <, >=, <= are removed. For example, if the normal range for a test is >100 and <300, then 100 is entered. Additionally, this value is null for records where Result\_Type = C. | 300 |
| Modifier\_high | Char (2) | EQ = equal  LE = less than or equal to  LT = less than | Modifier for Norm\_Range\_high values. The normal range associated with a test, as assigned by the laboratory is parsed out into the following variables: Norm\_Range\_low, Modifier\_low, Norm\_Range\_high, Modifier\_high and reflects what is seen in source data. The value is null for records where Result\_Type=C. For numeric results one of the following needs to be true:  Both Modifier\_low and Modifier\_high contain EQ (e.g., normal values fall in the range 3-10).  Modifier\_low contains GT or GE and Modifier\_high is null (e.g., normal values are >3 with no upper boundary).  Modifier\_high contains LT or LE and Modifier\_low is null (e.g., normal values are <=10 with no lower boundary). | LT |
| Abn\_ind | Char (2) | AB = abnormal  AH = abnormally high  AL = abnormally low  CH = critically high  CL = critically low  CR = critical  IN = inconclusive  NL = normal  UN = unknown | Abnormal result indicator. This value comes directly from the source data; this value is not created programmatically based on variables such as MS\_Result\_N, Norm\_Range\_high, or Norm\_Range\_low. | CR |
| Order\_dept | Char (Site Specific length) | Unique to each Data Partner | Local code for ordering provider department. Values for Order\_dept are not required. Populate only if ordering provider department is available in source data. |  |
| Facility\_Code | Char (Site Specific length) | Servicing provider identifier | Local facility code that identifies the hospital or clinic. Taken from facility claims. Values for Facility\_Code are not required. Populate only if facility code is available in source data. This information is useful when locating medical charts for review. | FC12345678 |

NOTES:

1. 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. Additionally, negative numeric values (e.g., -5.0 mg/ml) are not included.
2. 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.
3. There are three dates that can be associated to the laboratory results: Order date (ORDER\_DT), Laboratory date (Lab\_dt), and Result date (RESULT\_DT). The typical sequence of these date variables, from earliest to latest occurrence, is: ORDER\_DT ≤ Lab\_dt ≤ RESULT\_DT. One or more of these dates is populated for each record in the Laboratory Result Table. These dates are all populated if available in source data.

# SCDM: Vital Signs Table Structure

Description: The SCDM Vital Signs Table contains one record per result/entry.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable Name | Variable Type and Length (Bytes) | Values | Definition / Comments / Guideline | Example |
| PatID1 | Char (Site specific length) | Unique member identifier | Arbitrary person-level identifier. Used to link across tables. | 123456789012345 |
| EncounterID2 | Char (Site specific length) | Unique encounter identifier | Arbitrary encounter-level identifier. Used to link across the Encounter, Diagnosis, Procedure, Vital Signs, Inpatient Pharmacy, & Inpatient Transfusion tables. | 123456789012345\_12242005\_99218766\_IP |
| Measure\_Date2 | Numeric (4) | SAS date | Date the vital signs were measured. | 12/1/2005 |
| Measure\_Time | Numeric (4) | SAS time | Time associated with the vital signs record. This may be the time an actual blood pressure measurement was taken or it may be a check-in time from encounter. |  |
| HT | Numeric (8) | Height (in inches) | ####.## = If HT can be represented in inches. Only populated if height was taken on this date. If missing, leave blank. | 60.50 |
| WT | Numeric (8) | Weight (in lbs) | ####.## = If WT can be represented in pounds. Only populated if weight was taken on this date. If missing, leave blank. | 170.25 |
| Diastolic | Numeric (4) | Diastolic blood pressure | ### = If Diastolic can be represented in mmHg. Only populated if diastolic blood pressure was taken on this date. If missing, leave blank. | 70 |
| Systolic | Numeric (4) | Systolic blood pressure | ### = If Systolic can be represented in mmHg. Only populated if systolic blood pressure was taken on this date. If missing, leave blank. | 120 |
| BP\_Type | Char (1) | E = Extended  M = Multiple  O = Orthostatic  R = Rooming | Type of blood pressure taken. | E |
| Position | Char (1) | 1 = Sitting  2 = Standing  3 = Supine | Position for orthostatic blood pressure. If unknown, leave blank. | 1 |
| Tobacco | Numeric (3) | 1 = Current user  2 = Never  3 = Quit/former user  4 = Passive  5 = Environmental exposure  6 = Not asked  7 = Conflicting | Tobacco status as of the visit date. Unknown values should be left blank. The “Not asked” value should be used only when it is a valid response from your system (e.g. this is a valid value for EPIC). The “Conflicting” value should be used when you receive tobacco information from multiple sources that disagree. | 3 |
| Tobacco\_Type | Numeric (3) | 1 = Cigarettes only  2 = Other tobacco only  3 = Cigarettes and other tobacco4 = None | Type of tobacco used. Unknown values should be left blank. | 4 |

NOTES:

1. 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.
2. The Encounter, Diagnosis, Procedure, Inpatient Transfusion, Inpatient Pharmacy, Vital Signs, and Mother-Infant Linkage tables are linked by EncounterID. All diagnoses & procedures for an encounter should have the same EncounterID.

# SCDM: Death Table Structure

Description: The SCDM Death Table contains one record per PatID1. 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.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable Name | Variable Type and Length (Bytes) | Values | Definition / Comments / Guideline | Example |
| PatID2 | Char (Site specific length) | Unique member identifier | Arbitrary person-level identifier. Used to link across tables. | 123456789012345 |
| DeathDt | Numeric (4) | SAS date | Date of death. | 1/1/2006 |
| DtImpute | Char (1) | B = Both month and day imputedD = Day imputedM = Month imputedN = Not imputed | When DeathDt is imputed, this variable indicates which parts of the date were imputed. | N |
| Source | Char (1) | L = Other, locally definedN = National Death IndexS = State Death filesT = Tumor data | Source of death information. | S |
| Confidence | Char (1) | E = ExcellentF = FairP = Poor | Confidence that the patient drawn from the Source data represents the actual patient (contrasts with Confidence in the Cause of Death table). | E |

NOTES:

1. For efficiency death data is captured in 2 tables:
   * Death: the record that characterizes the death date and source of that information
   * Cause of Death: the cause(s) of death associated with the death record

* These 2 tables are linked by PatID. All Cause of Death records have a matching PatID in the Death Table.

1. 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.

# SCDM: Cause of Death Table Structure

Description: The SCDM Cause of Death Table contains one record per unique combination of PatID and COD.1 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.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable Name | Variable Type and Length (Bytes) | Values | Definition / Comments / Guideline | Example |
| PatID2 | Char (Site specific length) | Unique member identifier | Arbitrary person-level identifier. Used to link across tables. | 123456789012345 |
| COD | Char (8) | Diagnosis code | Cause of death code. Please include the decimal point in ICD codes (if any). | J18.0 |
| CodeType | Char (2) | 09 = ICD-9  10 = ICD-10 | Cause of death code type. | 09 |
| CauseType | Char (1) | C = Contributory  I = Immediate / Primary  O = Other  U = Underlying | Cause of death type. There should be only one underlying cause of death. | C |
| Source | Char (1) | L = Other, locally defined  N = National Death Index  S = State Death files  T = Tumor data | Source of cause of death information. | S |
| Confidence | Char (1) | E = Excellent  F = Fair  P = Poor | Confidence in the accuracy of the cause of death based on source, match, number of reporting sources, discrepancies, etc. | E |

NOTES:

1. For efficiency death data is captured in 2 tables:
   * Death: the record that characterizes the death date and source of that information
   * Cause of Death: the cause(s) of death associated with the death record

* These 2 tables are linked by PatID. All Cause of Death records have a matching PatID in the Death Table.

1. 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.

# SCDM: Inpatient Pharmacy Table Structure

Description: The SCDM Inpatient Pharmacy table contains data on inpatient drug administrations. It contains one record per unique combination of PatID, NDC, RxADate, RxATime, and RxID. Each record represents a unique inpatient pharmacy dispensing administration.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable Name | Variable Type and Length (Bytes) | Values | Definition / Comments / Guideline | Example |
| PatID1 | Char (Site specific length) | Unique member identifier | Arbitrary person-level identifier. Used to link across tables. | 123456789012345 |
| EncounterID2 | Char (Site specific length) | Unique encounter identifier | Arbitrary encounter-level identifier. Used to link across the Encounter, Diagnosis, Procedure, Vital Signs, Inpatient Pharmacy, & Inpatient Transfusion tables. | 123456789012345\_12242005\_99218766\_IP |
| NDC | Char (11) | National Drug Code | Please expunge any place holders (e.g., ‘-’ or extra digit) | 12345678910 |
| RxID | Char (15) | Unique Rx administration identifier | Useful to map back to source data |  |
| RxADate | Numeric (4) | SAS date value | Rx Administration date | 12/1/2005 |
| RxATime | Numeric (4) | SAS time value HH:MM | Rx Administration time | 20:56 |
| RxRoute | Char (10) | Text | Actual / administered route. Standard list of allowable values under development. | IV |
| RxDose | Num (8) | Numeric digits with or without a decimal | Actual / administered dose. Intended to be analyzed in conjunction with the RxUOM (unit of measure) field and product strength data associated with the NDC (available from drug databases). Format captures maximum # of whole and decimal digits allowed by software technology for numeric data. | 100 |
| RxUOM | Char (10) | Text | Actual / administered unit of measure. Intended to be analyzed in conjunction with the RxDose field and product strength data associated with the NDC (available from drug databases). Standard list of allowable values under development. | ML |

NOTES:

1. 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.
2. The Encounter, Diagnosis, Procedure, Inpatient Transfusion, Inpatient Pharmacy, Vital Signs, and Mother-Infant Linkage tables are linked by EncounterID. All diagnoses and procedures for an encounter should have the same EncounterID.

# SCDM: Inpatient Transfusion Table Structure

Description: The SCDM Inpatient Transfusion table contains data on inpatient blood transfusion administrations. It contains one record per unique combination of PatID and TransID. Each record represents a unique inpatient pharmacy transfusion administration, as defined by unique value combinations of PatID / TransCode / TransCode\_Type / TDate\_Start / Ttime\_Start.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable Name | Variable Type and Length (Bytes) | Values | Definition / Comments / Guideline | Example |
| PatID1 | Char (Site specific length) | Unique member identifier | Arbitrary person-level identifier. Used to link across tables. | 123456789012345 |
| EncounterID2 | Char (Site specific length) | Unique encounter identifier | Arbitrary encounter-level identifier. Used to link across the Encounter, Diagnosis, Procedure, Vital Signs, Inpatient Pharmacy, & Inpatient Transfusion tables. | 123456789012345\_12242005\_99218766\_IP |
| TransID | Char (15) | Unique transfusion administration identifier | Retain because useful to map back to source data | 123456789012345 |
| TransCode | Char (15) | Code value for an infusion product | Must be paired with the correct TransCode\_Type | 123451224200599 |
| TransCode\_Type | Char(2) | CD = CODABAR  IS = ISBT | Transfusion product code type. This variable combined with the TransCode variable should be used to capture any type of Inpatient Infusion product in the source data. Other code types will be added as new terminologies are used. | CD |
| Orig\_TransProd | Char (Site specific length) | Original product name/mneumonic | Name of product within Data Partner | Thawed POOLED PLATELETS |
| BloodType | Char(3) | A, B, O, AB (upper case) with RH factor (+, -, or null only) | Blood type and Rh factors, left-justified. Convert any text Rh factor to symbols (e.g., “pos” to “+”, “negative” to “-”). Rh factor can be blank. | AB+ |
| TDate\_Start | Numeric (4) | SAS date value | Administration start date. | 12/1/2005 |
| TTime\_Start | Numeric (4) | SAS time value HH:MM | Administration start time. | 14:27 |
| TDate\_End | Numeric (4) | SAS date value | Administration end date. | 12/1/2005 |
| TTime\_End | Numeric (4) | SAS time value HH:MM | Administration end time. | 20:56 |
| EncType | Char (2) | ED = Emergency Department | Includes ED encounters that become inpatient stays (in which case inpatient stays would be a separate encounter). Excludes urgent care 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. | IP |
|  |  | IP = Inpatient Hospital Stay | Includes all inpatient stays, same-day hospital discharges, hospital transfers, and acute hospital care where the discharge is after the admission date. |  |
|  |  | IS = Non-Acute Institutional Stay | Includes hospice, skilled nursing facility (SNF), rehab center, nursing home, residential, overnight non-hospital dialysis and other non-hospital stays. |  |
|  |  | OA = Other Ambulatory Visit | Includes other non overnight AV encounters such as hospice visits, home health visits, skilled nursing facility visits, other non-hospital visits, as well as telemedicine, telephone and email consultations. |  |

NOTES:

1. 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.
2. The Encounter, Diagnosis, Procedure, Inpatient Transfusion, Inpatient Pharmacy, Vital Signs, and Mother-Infant Linkage tables are linked by EncounterID. All diagnoses and procedures for an encounter should have the same EncounterID.

# SCDM: Mother-Infant Linkage Table Structure

Description: The SCDM Mother-Infant Linkage Table contains one record per MPatID, CPatID, and ADate. This table is created following identification of mothers (via evidence of live delivery by women aged 10-54 inclusive) and infants (via date of birth) in the Sentinel Distributed Database (SDD). The file may include:

1. Live birth deliveries (with MPatID and ADate) that were linked to a child (CPatID);
2. Live birth deliveries (with MPatID and ADate) that were not linked to a child (CPatID, CBirth\_Date, Sex, and CEnr\_Start will have missing values); and
3. Children (with CPatID) who were not linked to a mother (MPatID, MBirth\_Date, Age, EncounterID, EncType, ADate, DDate, Birth\_Type, and Birth\_Type\_Primes will have missing values).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable Name | Variable Type and Length (Bytes) | Values | Definition / Comments / Guideline | Example |
| MPatID1 | Char(Site specific length) | Unique member identifier. Text string; left justified. | Arbitrary person-level identifier. Used to link across tables. Length is DP specific. Must match mom “PatID” value in all other SCDM tables.Blank for child-only records. | 123456789012345 |
| MBirth\_Date | Numeric(4) | SAS Date | Mother Birth\_Date value from SCDM Demographic table. Blank for child-only records. | 12/5/1971 |
| Age | Numeric(3) | 10-54 inclusive | Mother’s age as of ADate.Blank for child-only records. | 32 |
| EncounterID2 | Char(Site specific length) | Unique encounter identifier. | EncounterID value from SCDM Encounter table, for mother’s delivery encounter.Blank for child-only records. | 123456789012345\_12242005\_99218766\_IP |
| EncType2 | Char(2) | AV = Ambulatory Visit  ED = Emergency Department  IP = Inpatient Hospital Stay  IS = Non-Acute Institutional Stay  OA = Other Ambulatory Visit | EncType value from SCDM Encounter table, for mother’s delivery encounter. Blank for child-only records. | IP |
| ADate | Numeric(4) | SAS Date | ADate value from SCDM Encounter table, for mother’s delivery encounter. Blank for child-only records. | 12/24/2005 |
| DDate | Numeric(4) | SAS Date; may be null | DDate value from SCDM Encounter table, for mother’s delivery encounter. Blank for child-only records. | 12/31/2005 |
| CPatID | Char(Site specific length) | Unique member identifier. Text string; left justified. | Arbitrary person-level identifier. Used to link across tables. Length is DP specific. Must match child “PatID” value in all other SCDM tables. Blank for mother/delivery-only records. | 12341234 |
| CBirth\_Date | Numeric(4) | SAS Date | Child Birth\_Date value from SCDM Demographic table. Blank for mother/delivery-only records. | 1/2/2015 |
| Sex | Char(1) | A = Ambiguous (e.g., transgender/hermaphrodite)  F = Female  M = Male  U = Unknown | Child Sex value from SCDM Demographic table. Blank for mother/delivery-only records. | F |
| CEnr\_Start | Numeric(4) | SAS Date | Earliest Enr\_Start from Enrollment table.Blank for mother/delivery-only records. | 1/1/2005 |
| Birth\_Type | Numeric(3) | 0 = Unspecified # of live births  1 = 1 live birth  2 = 2 live births  3 = 3 live births  4 = 4 live births  5 = 5 live births  8 = Multiple live births, unspecified number  9 = Conflicting code(s) for number of live births | Based upon ICD-9-CM/ICD-10-CM codes in the health plan data for the delivery admission.Blank for child-only records. | 3 |
| Birth\_Type\_Primes | Numeric(8) | 2 = Birth\_Type of 0  3 = Birth\_Type of 1  5 = Birth\_Type of 2  7 = Birth\_Type of 3  11 = Birth\_Type of 4  13 = Birth\_Type of 5  17 = Birth\_Type of 8 | Multiplication of all prime numbers assigned to all Birth\_Types found in delivery codes within the selected encounter. This provides a record of all values of Birth\_Type for the selected encounter.Missing/null for child-only records. | 5 |
| MatchMethod | Char(2) | BC = Birth Certificate  RE = DP maintained birth registry  SI = health plan subscriber or family number  LA = exact or probabilistic last name and address match based upon health plan administrative data  OT = other | Prioritized method of linkage for mom-baby match, or reason for unlinked record.For linked records, prioritize so that only one method is listed: RE > SI > LA > BC > OT | RE |
|  |  | N1 = No subscriber/family IDs available for linkage  N2 = No name/address available for linkage  N3 = Neither subscriber/family IDs nor name/address available for linkage  NA = no linkage made; any other reason | Prioritized method of linkage for mom-baby match, or reason for unlinked record. For cases where a mother/delivery is not linked to a child OR a child is not linked to a mother/delivery, the value of this variable should be one of N1, N2, N3, or NA only | N2 |

NOTES

1. 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.
2. The Encounter, Diagnosis, Procedure, Inpatient Transfusion, Inpatient Pharmacy, Vital Signs, and Mother-Infant Linkage tables are linked by EncounterID. All diagnoses and procedures for an encounter should have the same EncounterID. If more than 1 delivery encounter occur on the same ADate, then the values are based on the encounter selected from the following hierarchy: IP > ED > AV > IS > OA