**ETL Conventions for use with PEDSnet CDM V1.0**

Revision Date: October 31, 2014

Version Tracking

|  |  |
| --- | --- |
| 10/20/2014 | Version 1: Includes observation period, uses CHOP’s codes for height/weight |
| 10/20/2014 | Version 1a: Readjusted logic for SBP/DBP and position. Incorporated changes from TO |
| 10/22/2014 | Incorporated (most of) changes from CB, GS, and KM |
| 10/24/2014 | Post ETL call changes; added CF’s future elements list |
| 10/27/2014 | Incorporated edits from JB |
| 10/28/2014 | Incorporated comments from St. Louis  Removed BMI from observations -- all sites can provide source height and weight. DCC will calculate and insert BMI observations using one common algorithm  Replaced vocabulary 99 with identical and additional concept\_ids from new vocabulary 60, which includes new concept\_id codes for "unknown" and "no information."  Untangled confusion regarding which fields can be/should be obfuscated (should be obfuscate key source values rather than ID fields |
| 10/29/2014 | Added suggestion that \*\_id field could be used as obfuscated value of source value as long as mapping between ID field and real source value is maintained by site.  Added text that biobank flag and chart availability can be entered into observations table more than once if patient changes consent status over time.  Clarified that discharge disposition and discharge status observations occur once per visit\_occurrence. |
| 10/31/2014 | Made care\_site\_source\_value and provider\_source\_value required fields and reverted observation\_period\_stop\_date to the standard observation\_period\_end\_date. |
|  |  |
|  |  |

The PEDSnet Common Data Model is an evolving specification, based in structure on the OMOP Common Data Model, but expanded to accommodate requirements of both the PCORnet Common Data Model and the primary research cohorts established in PEDSnet.

Version 1 of the PEDSnet CDM reflects the ETL processes developed during the first six months of network development. As such, it closely follows version 1 of the PCORnet CDM. We anticipate that the PEDSnet CDM will expand to include additional data domains such as medication usage, laboratory testing, and other types of clinical observations. However, in order to minimize discordance with the PCORnet CDM, specification of these domains has been deferred until the cognate portion of the PCORnet CDM is developed, or active research in one of the PEDSnet cohorts requires those data types.

This document provides the ETL processing assumptions and conventions developed by the PEDSnet data partners that should be used by a data partner for ensuring common ETL business rules. This document will be modified as new situations are identified, incorrect business rules are identified and replaced, as new analytic use cases impose new/different ETL rules, and as the PEDSnet CDM continues to evolve.

Comments on this specification and ETL rules are welcome. Please send email to [pedsnetdcc@email.chop.edu](mailto:pedsnetdcc@email.chop.edu), or contact the PEDSnet project management office (details available via <http://www.pedsnet.info>).

**PEDSnet Data Standards and Interoperability Policies:**

1. The PEDSnet data network will store data using structures compatible with the PEDSnet Common Data Model (PCDM).
2. The PCDM v1 is based on the Observational Medical Outcomes Partnership (OMOP) data model, version 4. OMOP will be expanded to include the PCORnet and pediatric-specific data standards, as developed by PEDSnet. The next release of PCDM will be based on OMOP CDM Version 5, which incorporates additional requirements for realizing PCORnet CDM V2.
3. A subset of data elements in the PCDM will be identified as principal data elements (PDEs). The PDEs will be used for population-level queries. Data elements which are NOT PDEs will be marked as “Optional” (ETL at site discretion) or “Non-PDE” (ETL required, but data need not be transmitted to DCC), and will not be used in queries without prior approval of site.
4. It is anticipated that PEDSnet institutions will make a good faith attempt to obtain as many of the data elements not marked as “Optional” as possible.
5. The data elements classified as PDEs and those included in the PCDM will be approved by the PEDSnet Executive Committee (comprised of each PEDSnet institution’s site principal investigator).
6. Concept IDs are taken from OMOP v4.5 vocabularies for PEDSnet CDM v1, using the complete (“restricted”) version that includes licensed terminologies such as CPT and others.
7. PCORnet CDM V1.0 requires data elements that are not currently considered "standard concepts". Vocabulary version 4.5 has a new vocabulary (vocabulary\_id = 60) that was added by OMOP to capture all of the PCORnet concepts that are not in the standard terminologies. We use concept\_ids from vocabulary\_id 60 where there are no existing standard concepts. We highlight where we are pulling concept\_ids from vocabulary\_id 60 in the tables. While terms from vocabulary\_id = 60 violates the OMOP rule to use only concept\_ids from **standard vocabularies** (vocabulary 60 is a **non-standard** vocabulary), this convention enables a clean extraction from PEDSnet CDM to PCORnet CDM.
8. Some source fields may be considered “sensitive” by data sites. Potential examples include patient\_source\_value, provider\_source\_value, care\_site\_source\_value. Many of these fields are used to generate an “ID” field, such as PATIENT.patient\_source\_value 🡪 PATIENT.patient\_id, that is used as a primary key in PATIENT and a foreign key in many other tables. Sites are free to obfuscate or not provide source values that are used to create “ID” variables. Sites must maintain a mapping from the ID variable back to the original site-specific value for local re-identification tasks.
   1. Source fields that contain clinical data, such as source condition occurrence, should be included
   2. The PEDSnet DCC will never release source values to external data partners.
   3. Source value obfuscation techniques may include replacing the real source value with a random number, an encrypted derivative value/string, or some other site-specific algorithm.
9. Regarding the nullability of all source value (string) fields only, the PEDSnet CDM will accommodate the following values, taken from the PCORnet CDM:

|  |  |
| --- | --- |
| **Null Name** | **Definition of each field** |
| NULL | A data field is not present in the source system |
| NI = No Information | A data field is present in the source system, but the source value is null or blank |
| UN = Unknown | A data field is present in the source system, but the source value explicitly denotes an unknown value |
| OT = Other | A data field is present in the source system, but the source value cannot be mapped to the CDM |

**ETL Recommendation:** Due to PK/FK constraints, the most efficient order for ETL table is location, organization, care\_site, provider, person, visit\_occurrence, condition\_occurrence, observation, procedure\_occurrence, and observation\_period

## PERSON

The person domain contains records that uniquely identify each patient in the source data who is time at-risk to have clinical observations recorded within the source systems. Each person record has associated demographic attributes, which are assumed to be constant for the patient throughout the course of their periods of observation. All other patient-related data domains have a foreign-key reference to the person domain.

PEDSnet uses a specific definition of an active PEDSnet patient. Only patients who meet the PEDSnet definition of an active patient should be included in this table. The criteria for identifying an active patient are:

* Has a unique identifier AND
* At least 1 “in person” clinical encounter on or after January 1, 2009 AND
* At least 1 coded diagnoses recorded on or after January 1, 2009 AND
* Is not a test patient or a research-only patient

The definition of an “in person” clinical encounter remains heuristic –any encounter type that involves a meaningful \*\*physical\*\* interaction with a clinician that involved clinical content. An encounter for a suture removal or a telephone encounter or a lab blood draw does not meet this definition.

**NOTE: While the 1/1/2009 date and “in person” clinical encounter restrictions apply to defining an active PEDSnet patient, once a patient has met this criteria, PEDSnet will extract \*\*\*ALL\*\*\* available clinical encounters/clinical data of any type across all available dates. That is, “1/1/2009 and 1 ‘in person’ clinical encounter” applies only to defining the active patient cohort. It does NOT apply to data extraction on active patients.**

| **Field** | **Required** | **Description** | **PEDSnet Conventions** |
| --- | --- | --- | --- |
| person\_id | Yes | A unique identifier for each person; this is created by each contributing site. | This is not a value found in the EHR.  PERSON\_ID must be unique for all patients within a single data set.  Sites may choose to use a sequential value for this field |
| gender\_concept\_id | Yes | A foreign key that refers to a standard concept identifier in the Vocabulary for the gender of the person. | Please include valid concept ids (consistent with OMOP CDMv4). Predefined value set (valid concept\_ids found in CONCEPT table where vocabulary\_id = 12 and vocabulary\_id = 60 where noted):   * Ambiguous: concept\_id = 8570 * Female: concept\_id = 8532 * Male: concept\_id = 8507 * No Information: concept\_id = 44814667 (Vocabulary 60) * Unknown: concept\_id = 8551 * Other: concept\_id = 8521 |
| year\_of\_birth | Yes | The year of birth of the person. | For data sources with date of birth, the year is extracted. For data sources where the year of birth is not available, the approximate year of birth is derived based on any age group categorization available.  Please keep all accurate/real dates (No date shifting) |
| month\_of\_birth | No | The month of birth of the person. | For data sources that provide the precise date of birth, the month is extracted and stored in this field.  Please keep all accurate/real dates (No date shifting) |
| day\_of\_birth | No | The day of the month of birth of the person. | For data sources that provide the precise date of birth, the day is extracted and stored in this field.  Please keep all accurate/real dates (No date shifting) |
| pn\_time\_of\_birth | No | The time of birth at the birth day | Include whatever is the relevant time zone when insert this value  Please keep all accurate/real dates (No date shifting) |
| race\_concept\_id | No | A foreign key that refers to a standard concept identifier in the Vocabulary for the race of the person. | 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.  For patients with multiple races (i.e. biracial), race is considered a single concept, meaning there is only one race slot. If there are multiple races in the source system, concatenate all races into one race\_source\_value (see below) and use concept\_id code as ‘Multiple Race’.  Predefined values (valid concept\_ids found in CONCEPT table where vocabulary\_id = 13 and vocabulary\_id = 60):   * American Indian/Alaska Native: concept\_id = 8657 * Asian: concept\_id = 8515 * Black or African American: concept\_id = 8516 * Native Hawaiian or Other Pacific Islander: concept\_id = 8557 * White: concept\_id = 8527 * Multiple Race: concept\_id = 44814659 (vocabulary 60) * Refuse to answer: concept\_id = 44814660 (vocabulary 60) * No Information: concept\_id = 44814661 (vocabulary 60) * Unknown: concept\_id = 8552 * Other: concept\_id = 8522 |
| ethnicity\_concept\_id | No | A foreign key that refers to the standard concept identifier in the Vocabulary for the ethnicity of the person. | For PEDSnet, a person with Hispanic ethnicity is defined as “A person of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin, regardless of race.”  Please include valid concept ids (consistent with OMOP CDMv4). Predefined value set (valid concept\_ids found in CONCEPT table where vocabulary\_id = 44 or vocabulary 60 where noted):   * Hispanic: concept\_id = 38003563 * Not Hispanic: concept\_id = 38003564 * No Information: concept\_id = 44814650 (vocabulary 60) * Unknown: concept\_id = 44814653 (vocabulary 60) * Other: concept\_id = 44814649 (vocabulary 60) |
| location\_id | No | A foreign key to the place of residency (ZIP code) for the person in the location table, where the detailed address information is stored. |  |
| provider\_id | No | Foreign key to the primary care provider – the person is seeing in the provider table. | For PEDSnet CDM V1.0: Sites will use site-specific logic to determine the “best” primary care provider and document how that decision was made (e.g., billing provider).  PEDSnet CDM 2.0/OMOP V5, multiple providers may be asserted based on specific use cases that require multiple providers in all provider\_id fields. |
| care\_site\_id | Yes | A foreign key to the site of primary care in the care\_site table, where the details of the care site are stored | For patients who receive care at multiple care sites, use site-specific logic to select a care site that best represents where the patient obtains the majority of their recent care. If a specific site within the institution cannot be identified, use a care\_site\_id representing the institution as a whole. |
| pn\_gestational\_age | No | The post-menstrual age in weeks of the person at birth, if known | Use granularity of age in weeks as is recorded in local EHR. |
| person\_source\_value | Yes | An encrypted key derived from the person identifier in the source data. | Insert a pseudo-identifier (random number, encrypted identifier) into the field. Do not insert the actual MRN or PAT\_ID from your site. A mapping from the pseudo-identifier for person\_source\_value in this field to a real patient ID or MRN from the source EHR must be kept at the local site. This mapping is not shared with the data coordinating center. It is used only by the site for re-identification for study recruitment or for data quality review.  Sites may consider using the person\_id field value in this table as the pseudo-identifier as long as a local mapping from person\_id to the real site identifier is maintained. |
| gender\_source\_value | No | The source code for the gender of the person as it appears in the source data. | The person’s gender is mapped to a standard gender concept in the Vocabulary; the original value is stored here for reference. See gender\_concept\_id |
| race\_source\_value | No | The source code for the race of the person as it appears in the source data. | The person race is mapped to a standard race concept in the Vocabulary and the original value is stored here for reference.  For patients with multiple races (i.e. biracial), race is considered a single concept, meaning there is only one race slot. If there are multiple races in the source system, concatenate all races into one source value, and use the conceptt\_id for ‘Multiple Race’. |
| ethnicity\_source\_value | No | The source code for the ethnicity of the person as it appears in the source data. | The person ethnicity is mapped to a standard ethnicity concept in the Vocabulary and the original code is, stored here for reference. |

## DEATH

The death domain contains the clinical event for how and when a person dies. Living patients should not contain any information in the death table.

|  |  |  |  |
| --- | --- | --- | --- |
| Field | Required | Description | PEDSnet Conventions |
| person\_id | Yes | A foreign key identifier to the deceased person. The demographic details of that person are stored in the person table. | See PERSON.person\_id (primary key) |
| death\_date | Yes | The date the person was deceased. | If the precise date including day or month is not known or not allowed, December is used as the default month, and the last day of the month the default day. If no date available, use date recorded as deceased.  When the date of death is not present in the source data, use the date the source record was created. |
| death\_type\_concept\_id | Yes | A foreign key referring to the predefined concept identifier in the Vocabulary reflecting how the death was represented in the source data. | Please include valid concept ids (consistent with OMOP CDMv4). Predefined value set (valid concept\_ids found in CONCEPT table where vocabulary\_id = 45)  select \* from concept where vocabulary\_id = 45 yields 8 valid concept\_ids.  If none are correct, use concept\_id = 0  Note: Most current ETLs are extracting data from EHR so most likely concept\_id to insert here is 38003569 ("Recorded from EHR")  Note: These terms only describe the source from which the death was reported. It does not describe our certainty/source of the date of death, which may have been created by one of the heuristics described in death\_date. |
| cause\_of\_death\_concept\_id | No | A foreign referring to a standard concept identifier in the Vocabulary for conditions. |  |
| cause\_of\_death\_source\_value | No | The source code for the cause of death as it appears in the source. This code is mapped to a standard concept in the Vocabulary and the original code is stored here for reference. |  |

### Additional Notes

* Each Person may have more than one record of death in the source data. It is OK to insert multiple death records for an individual.
* If the Death Date cannot be precisely determined from the data, the best approximation should be used.

## LOCATION

The Location table represents a generic way to capture physical location or address information. Locations are used to define the addresses for Persons and Care Sites. NOTE: In OMOP CDM V5, this table is eliminated so don’t spend a lot of time on this table. The most important field is ZIP for location-based queries.

| **Field** | **Required** | **Description** | **PEDSnet Conventions** |
| --- | --- | --- | --- |
| location\_id | Yes | A unique identifier for each geographic location. | This is not a value found in the EHR. Sites may choose to use a sequential value for this field |
| state | No | The state field as it appears in the source data. |  |
| zip | No | The zip code. For US addresses, valid zip codes can be 3, 5 or 9 digits long, depending on the source data. | While optional, this is the most important field in this table to support location-based queries. |
| location\_source\_value | No | Optional - Do not transmit to DCC  The verbatim information that is used to uniquely identify the location as it appears in the source data. | If location source values are deemed sensitive by your organization, insert a pseudo-identifier (random number, encrypted identifier) into the field. Sites electing to obfuscate location\_source\_values will keep the mapping between the value in this field and the original clear text location source value. This value is only used for site-level re-identification for study recruitment and for data quality review.  Sites may consider using the location\_id field value in this table as the pseudo-identifier as long as a local mapping from location\_id to the real site identifier is maintained. |
| address\_1 | No | Optional - Do not transmit to DCC |  |
| address\_2 | No | Optional - Do not transmit to DCC |  |
| city | No | Optional - Do not transmit to DCC |  |
| county | No | Optional - Do not transmit to DCC |  |

### Additional Notes

* Each address or Location is unique and is present only once in the table

## CARE\_SITE

The Care Site table contains a list of uniquely identified physical or organizational units where healthcare delivery is practiced (offices, wards, hospitals, clinics, etc.). Future definitions of PEDSnet CDM will more precisely define the fields in this table. The most important field in this table is organization\_id, which is the tie back to the contributing PEDSnet data partner (CHOP versus Colorado versus St. Louis).

|  |  |  |  |
| --- | --- | --- | --- |
| Field | Required | Description | PEDSnet Conventions |
| care\_site\_id | Yes | A unique identifier for each defined location of care within an organization. Here, an organization is defined as a collection of one or more care sites that share a single EHR database. | This is not a value found in the EHR.  Sites may choose to use a sequential value for this field |
| place\_of\_service\_concept\_id | No | A foreign key that refers to a place of service concept identifier in the Vocabulary | Please include valid concept ids (consistent with OMOP CDMv4). Predefined value set (valid concept\_ids found in CONCEPT table where vocabulary\_id = 24 and vocabulary 60 where noted)  select \* from concept where vocabulary\_id = 24 yields 4 valid concept\_ids.  If none are correct, use concept\_id = 0   * Inpatient Hospital Stay: concept\_id = 9201 * Ambulatory Visit: concept\_id = 9202 * Emergency Department: concept\_id = 9203 * Non-Acute Institutional Stay: concept\_id = 42898160 (vocabulary 60) * Unknown: concept\_id = 44814713 (vocabulary 60) * Other: concept\_id = 44814711 (vocabulary 60) * No information: concept\_id = 44814712 (vocabulary 60) |
| location\_id | No | A foreign key to the geographic location of the administrative offices of the organization in the location table, where the detailed address information is stored. |  |
| care\_site\_source\_value | Yes | The identifier for the organization in the source data, stored here for reference. | If care site source values are deemed sensitive by your organization, insert a pseudo-identifier (random number, encrypted identifier) into the field. Sites electing to obfuscate care site\_source\_values will keep the mapping between the value in this field and the original clear text location source value. This value is only used for site-level re-identification for study recruitment and for data quality review.  For EPIC EHRs, map care\_site\_id to Clarity Department.  Sites may consider using the care\_site\_id field value in this table as the pseudo-identifier as long as a local mapping from care\_site\_id to the real site identifier is maintained. |
| place\_of\_service\_source\_value | No | The source code for the place of service as it appears in the source data, stored here for reference. |  |
| organization\_id | Yes | A foreign key to the organization record. | See ORGANIZATION.organization\_id (primary key) |

### Additional Notes

* Care sites are primarily identified based on the specialty or type of care provided, and secondarily on physical location, if available (e.g. North Satellite Endocrinology Clinic)
* The Care Site Source Value typically contains the name of the Care Site.
* The Place of Service Concepts are based on a catalog maintained by the CMS (see vocabulary for values)

## ORGANIZATION

Note: This table will be incorporated into CARE SITE in OMOP CDM V5/PEDSnet CDM V2

|  |  |  |  |
| --- | --- | --- | --- |
| Field | Required | Description | PEDSnet Conventions |
| organization\_id | Yes | A unique identifier for each defined location of care within an organization. Here, an organization is defined as a collection of one or more care sites that share a single EHR database. | This is not a value found in the EHR. Sites may choose to use a sequential value for this field |
| place\_of\_service\_concept\_id | No | A foreign key that refers to a place of service concept identifier in the Vocabulary | Please include valid concept ids (consistent with OMOP CDMv4). Predefined value set (valid concept\_ids found in CONCEPT table where vocabulary\_id = 14)  select \* from concept where vocabulary\_id = 14 yields 49 valid concept\_ids.  If none are correct, use concept\_id = 0  Make a best-guess mapping. |
| location\_id | No | A foreign key to the geographic location of the administrative offices of the organization in the location table, where the detailed address information is stored. | See LOCATION.location\_id (primary key) |
| place\_of\_service\_source\_value | No | The source code for the place of service as it appears in the source data, stored here for reference. |  |
| organization\_source\_value | Yes | A foreign key to the organization record, stored here for reference. | In PEDSnet, expected organizations are Boston, CCHMC, CHOP, Colorado, Nationwide, Nemours, Seattle, and St. Louis. |

## PROVIDER

The Provider table contains a list of uniquely identified health care providers. These are typically physicians, nurses, etc.

| **Field** | **Required** | **Description** | **PEDSnet Conventions** |
| --- | --- | --- | --- |
| provider\_id | Yes | A unique identifier for each provider.  Each site must maintain a map from this value to the identifier used for the provider in the source data. | This is not a value found in the EHR. Sites may choose to use a sequential value for this field  See “Additional Comments” below. Sites should document who they have included as a provider. |
| specialty\_concept\_id | No | A foreign key to a standard provider's specialty concept identifier in the Vocabulary. | Please include valid concept ids (consistent with OMOP CDMv4). Predefined value set (valid concept\_ids found in CONCEPT table where vocabulary\_id = 48)  select \* from concept where vocabulary\_id = 48 yields 111 valid concept\_ids.  If none are correct, use concept\_id = 0  For providers with more than one specialty, use site-specific logic to select one specialty and document the logic used. For example, sites may decide to always assert the \*\*first\*\* specialty listed in their data source. |
| care\_site\_id | Yes | A foreign key to the main care site where the provider is practicing.. | See CARE\_SITE.care\_site\_id (primary key) |
| NPI | No | Optional - Do not transmit to DCC  The National Provider Identifier (NPI) of the provider. |  |
| DEA | No | Optional - Do not transmit to DCC  The Drug Enforcement Administration (DEA) number of the provider. |  |
| provider\_source\_value | Yes | The identifier used for the provider in the source data, stored here for reference. | Insert a pseudo-identifier (random number, encrypted identifier) into the field. Do not insert the actual PROVIDER\_ID from your site. A mapping from the pseudo-identifier for provider\_source\_value in this field to a real provider ID from the source EHR must be kept at the local site. This mapping is not shared with the data coordinating center. It is used only by the site for re-identification for study recruitment or for data quality review.  Sites may consider using the provider\_id field value in this table as the pseudo-identifier as long as a local mapping from provider\_id to the real site identifier is maintained. |
| specialty\_source\_value | No | The source code for the provider specialty as it appears in the source data, stored here for reference. | Optional. May be obfuscated if deemed sensitive by local site. |

### additional notes

* For PEDSnet, a provider is any individual (MD, DO, NP, PA, RN, etc) who is authorized to document care.
* Providers are not duplicated in the table.

## VISIT\_OCCURRENCE

The visit domain contains the spans of time a person continuously receives medical services from one or more providers at a care site in a given setting within the health care system.

| **Field** | **Required** | **Description** | **PEDSnet Conventions** |
| --- | --- | --- | --- |
| visit\_occurrence\_id | Yes | A unique identifier for each person’s visits or encounter at a healthcare provider. | This is not a value found in the EHR. Sites may choose to use a sequential value for this field. Do not use institutional encounter ID. |
| person\_id | Yes | A foreign key identifier to the person for whom the visit is recorded. The demographic details of that person are stored in the person table. |  |
| visit\_start\_date | Yes | The start date of the visit. | No date shifting |
| visit\_end\_date | No | The end date of the visit. | No date shifting  If this is a one-day visit the end date should match the start date.  If the encounter is on-going at the time of ETL, this should be null. |
| provider\_id | No | A foreign key to the provider in the provider table who was associated with the visit. | Use attending or billing provider for this field if available, even if multiple providers were involved in the visit. Otherwise, make site-specific decision on which provider to associate with visits and document.  **NOTE: this is NOT in OMOP CDM v4, but appears in OMOP CDMv5.** |
| care\_site\_id | No | A foreign key to the care site in the care site table that was visited. | See CARE\_SITE.care\_site\_id (primary key) |
| place\_of\_service\_concept\_id | Yes | A foreign key that refers to a place of service concept identifier in the vocabulary. | Please include valid concept ids (consistent with OMOP CDMv4). Predefined value set (valid concept\_ids found in CONCEPT table where vocabulary\_id = 24)  select \* from concept where vocabulary\_id = 24 yields 4 valid concept\_ids.  If none are correct, use concept\_id = 0   * Inpatient Hospital Stay: concept\_id = 9201 * Ambulatory Visit: concept\_id = 9202 * Emergency Department: concept\_id = 9203 * Non-Acute Institutional Stay: concept\_id = 42898160 * Unknown: concept\_id = 44814713 * Other: concept\_id = 44814711 (vocabulary 60) * No information: concept\_id = 44814712 (vocabulary 60) |
| place\_of\_service\_source\_value | No | The source code used to reflect the type or source of the visit in the source data. Valid entries include office visits, hospital admissions, etc. These source codes can also be type-of service codes and activity type codes. |  |

### Additional Notes

* The 1/1/2009 date limitation that is used to define a PEDSnet active patient is \*\*NOT\*\* applied to visit\_occurrence. All visits, of all types (physical and virtual) are included for an active patient.
* A Visit Occurrence is recorded for each visit to a healthcare facility.
* If a visit includes moving between different “place\_of\_service\_concepts” (ED -> inpatient) this should be split into separate visit\_occurrences to meet PCORnet’s definitions
* Each Visit is standardized by assigning a corresponding Concept Identifier based on the type of facility visited and the type of services rendered.
* At any one day, there could be more than one visit.
* One visit may involve multiple attending or billing providers (e.g. billing, attending, etc), in which case the ETL must specify how a single provider id is selected or leave the provider\_id field null.
* One visit may involve multiple care sites, in which case the ETL must specify how a single care\_site id is selected or leave the care\_site\_id field null.

## CONDITION\_OCCURRENCE

The condition occurrence domain captures records of a disease or a medical condition based on diagnoses, signs and/or symptoms observed by a provider or reported by a patient.

Conditions are recorded in different sources and levels of standardization. For example:

* Medical claims data include ICD-9-CM diagnosis codes that are submitted as part of a claim for health services and procedures.
* EHRs may capture a person’s conditions in the form of diagnosis codes and symptoms as ICD-9-CM codes, but may not have a way to capture out-of-system conditions.

|  |  |  |  |
| --- | --- | --- | --- |
| Field | Required | Description | PEDSnet Conventions |
| condition\_occurrence\_id | Yes | A unique identifier for each condition occurrence event. | This is not a value found in the EHR. Sites may choose to use a sequential value for this field |
| person\_id | Yes | A foreign key identifier to the person who is experiencing the condition. The demographic details of that person are stored in the person table. |  |
| condition\_concept\_id | Yes | A foreign key that refers to a standard condition concept identifier in the Vocabulary. | Please include valid concept ids (consistent with OMOP CDMv4). Predefined value set (valid concept\_ids found in CONCEPT table where vocabulary\_id = 1)  select \* from concept where vocabulary\_id = 1 yields ~400,000 valid concept\_ids.  If none are correct, use concept\_id = 0 |
| condition\_start\_date | Yes | The date when the instance of the condition is recorded. | No date shifting |
| condition\_end\_date | No | The date when the instance of the condition is considered to have ended | No date shifting  If this information is not available, set to NULL. |
| condition\_type\_concept\_id | Yes | A foreign key to the predefined concept identifier in the Vocabulary reflecting the source data from which the condition was recorded, the level of standardization, and the type of occurrence. For example, conditions may be defined as primary or secondary diagnoses, problem lists and person statuses. | Please include valid concept ids (consistent with OMOP CDMv4). Predefined value set (valid concept\_ids found in CONCEPT table where vocabulary\_id = 37)  select \* from concept where vocabulary\_id = 37 yields 67 valid concept\_ids.  If none are correct, use concept\_id = 0  If data source only identifies conditions as “primary” or “secondary” with no sequence number, use the following concept\_ids:  Inpatient primary: concept\_id = 38000199 Inpatient secondary: concept\_id = 38000201 Outpatient primary: concept\_id = 38000230 Outpatient secondary: concept\_id = 38000231 |
| stop\_reason | No | The reason, if available, that the condition was no longer recorded, as indicated in the source data. | Valid values include discharged, resolved, etc. Note that a stop\_reason does not necessarily imply that the condition is no longer occurring, and therefore does not mandate that the end date be assigned.  Leave blank for billing diagnoses. Possibly will be used for problem list diagnoses in the future. |
| associated\_provider\_id | No | A foreign key to the provider in the provider table who was responsible for determining (diagnosing) the condition. | Any valid provider\_id allowed (see definition of providers in PROVIDER table)  Make a best-guess and document method used. Or leave blank |
| visit\_occurrence\_id | No | A foreign key to the visit in the visit table during which the condition was determined (diagnosed). |  |
| condition\_source\_value | No | The source code for the condition as it appears in the source data. This code is mapped to a standard condition concept in the Vocabulary and the original code is, stored here for reference. | Condition source codes are typically ICD-9-CM diagnosis codes from medical claims or discharge status/visit diagnosis codes from EHRs. Use source\_to\_concept maps to translation from source codes to OMOP concept\_ids. |

### Additional Notes

* The 1/1/2009 date limitation that is used to define a PEDSnet active patient is \*\*NOT\*\* applied to condition\_occurrence. All conditions are included for an active patient. For PEDSnet CDM V1, we limit condition\_occurrences to final diagnoses only (not reason-for-visit, , and provisional surgical diagnoses such as those recored in EPIC OPTIME). In EPIC, final diagnoses includes both encounter diagnoses and billing diagnoses, problem lists (all problems, not filtered on “chronic” versus “provisional” unless local practices use this flag as intended).
* Condition records are inferred from diagnostic codes recorded in the source data by a clinician or abstractionist for a specific visit. In the current version of the CDM, problem list entries are not used, nor are diagnoses extracted from unstructured data, such as notes.
* Source code systems, like ICD-9-CM, ICD-10-CM, etc., provide coverage of conditions. However, if the code does not define a condition, but rather is an observation or a procedure, then such information is not stored in the CONDITION\_OCCURRENCE table, but in the respective tables instead. An example are ICD-9-CM procedure codes. For example, OMOP source-to-concept table uses the MAPPING\_TYPE column to distinguish ICD9 codes that represent procedures rather than conditions.
* Condition source values are mapped to standard concepts for cflowonditions in the Vocabulary. Since the icd9-cm diagnosis codes are not in the concept table, use the source\_to\_concept\_map table where the icd9\_code = source\_code and the source\_vocabulary\_id =2 (icd\_9) and target\_vocabulatory\_id=1 (snomed-ct) to locate the correct condition\_concept\_id value.
* When the source code cannot be translated into a Standard Concept, a CONDITION\_OCCURRENCE entry is stored with only the corresponding source\_value and a condition\_concept\_id of 0.
* Codes written in the process of establishing the diagnosis, such as "question of" of and "rule out", are not represented here.

## PROCEDURE\_OCCURRENCE

The procedure domain contains records of significant activities or processes ordered by and/or carried out by a healthcare provider on the patient to have a diagnostic and/or therapeutic purpose that are not fully captured in another table (e.g. drug\_exposure).

Procedures records are extracted from structured data in Electronic Health Records that capture source procedure codes using CPT-4, ICD-9-CM (Procedures), HCPCS or OPCS-4 procedures as orders.

| **Field** | **Required** | **Description** | **PEDSnet Conventions** |
| --- | --- | --- | --- |
| procedure\_occurrence\_id | Yes | A system-generated unique identifier for each procedure occurrence | This is not a value found in the EHR. Sites may choose to use a sequential value for this field |
| person\_id | Yes | A foreign key identifier to the person who is subjected to the procedure. The demographic details of that person are stored in the person table. |  |
| procedure\_concept\_id | Yes | A foreign key that refers to a standard procedure concept identifier in the Vocabulary. | Valid Procedure Concepts belong to the "Procedure" domain. Procedure Concepts are based on a variety of vocabularies: SNOMED-CT (vocabulary\_id = 1), ICD-9-Procedures (vocabulary\_id = 3), CPT-4 (vocabulary\_id = 4), and HCPCS (vocabulary\_id = 5)  Procedures are expected to be carried out within one day. If they stretch over a number of days, such as artificial respiration, usually only the initiation is reported as a procedure (CPT-4 "Intubation, endotracheal, emergency procedure").  Procedures could involve the administration of a drug, in which case the procedure is recorded in the procedure table and simultaneously the administered drug in the drug table. |
| procedure\_date | Yes | The date on which the procedure was performed. |  |
| procedure\_type\_concept\_id | Yes | A foreign key to the predefined concept identifier in the Vocabulary reflecting the type of source data from which the procedure record is derived.  (OMOP vocabulary\_id = 38) | Please include valid concept ids (consistent with OMOP CDMv4). Predefined value set (valid concept\_ids found in CONCEPT table where vocabulary\_id = 38)  select \* from concept where vocabulary\_id = 38 yields 33 valid concept\_ids.  If none are correct, use concept\_id = 0 |
| associated\_provider\_id | No | A foreign key to the provider in the provider table who was responsible for carrying out the procedure. | Any valid provider\_id allowed (see definition of providers in PROVIDER table)  Document how selection was made. |
| visit\_occurrence\_id | No | A foreign key to the visit in the visit table during which the procedure was carried out. | See VISIT.visit\_occurrence\_id (primary key) |
| relevant\_condition\_concept\_id | No | A foreign key to the predefined concept identifier in the vocabulary reflecting the condition that was the cause for initiation of the procedure. | Note that this is not a direct reference to a specific condition record in the condition table, but rather a condition concept in the vocabulary.  Use OMOP vocabulary\_id = 1 |
| procedure\_source\_value | No | The source code for the procedure as it appears in the source data. This code is mapped to a standard procedure concept in the Vocabulary and the original code is stored here for reference. | Procedure\_source\_value codes are typically ICD-9, ICD-10 Proc, CPT-4, HCPCS, or OPCS-4 codes. All of these codes are acceptable source values. |

### additional notes

* The 1/1/2009 date limitation that is used to define a PEDSnet active patient is \*\*NOT\*\* applied to procedure\_occurrence. All procedures are included for an active patient. For PEDSnet CDM V1, we limit procedures\_occurrences to billing procedures only (not surgical diagnoses).
* Procedure Concepts are based on a variety of vocabularies: SNOMED-CT, ICD-9-Proc, CPT-4, HCPCS and OPCS-4.
* Procedures could reflect the administration of a drug, in which case the procedure is recorded in the procedure table and simultaneously the administered drug in the drug table.
* The Visit during which the procedure was performed is recorded through a reference to the VISIT\_OCCURRENCE table. This information is not always available.
* The Provider carrying out the procedure is recorded through a reference to the PROVIDER table. This information is not always available.

## OBSERVATION

The observation domain captures clinical facts about a patient obtained in the context of examination, questioning or a procedure. For the PEDSnet CDM version 1, the observations listed below are extracted from source data. Please assign the specific concept\_ids listed in the table below to these observations as observation\_concept\_ids. Non-standard PCORnet concepts require concepts that have been entered into an OMOP-generated vocabulary (OMOP provided vocabulary\_id = 60). ~~See Appendix for SQL INSERT statements that add the necessary rows in the CONCEPT table to support PCORnet CDM V1.0.~~

NOTE: DRG and DRG Type require special logic/processing described below.

* Height/length in cm (use numeric precision as recorded in EHR)
* Height/length type
* Weight in kg (use numeric precision as recorded in EHR)
* Body mass index in kg/m2 (extracted only if height and weight are not present)
* Systolic blood pressure in mmHg
  + Where multiple readings are present on the same encounter, create observation records for \*\*ALL\*\* readings
* Diastolic blood pressure in mmHg
  + Where multiple readings are present on the same encounter, create observation records for \*\*ALL\*\* readings
* Blood pressure position is described by the selection of a concept\_id that contains the BP position as describe below. For example, in Table 1, concept\_id 3018586 is Systolic Blood Pressure, Sitting. This concept\_id identifies both the measurement (Systolic BP) and the BP position (sitting).
* Biobank availability
* Admitting source
* Discharge disposition
* Discharge status
* Chart availability
* Vital source
* DRG (requires special logic – see Note 4 below)
* Vital source (not captured in PEDSnet CDM 1.0)

Use the following table to populate observation\_concept\_ids and (where applicable) value\_as\_concept\_ids for the observations listed above. The vocabulary column is used to highlight non-standard codes from vocabulary 39 and 60 and one newly added standard concept from vocabulary 1.

Table 1: Observation concept IDs for PCORnet concepts. Concept\_ids from vocabulary\_id 99 are non-standard codes.

| Concept Name | Observation concept ID | Vocab ID | Value as concept ID | Concept description | Vocab ID |
| --- | --- | --- | --- | --- | --- |
| Biobank flag  (see Note 5) | 4001345 |  | 4188539 | Yes |  |
| Biobank flag | 4001345 |  | 4188540 | No |  |
| Biobank flag | 4001345 |  | NULL | No information |  |
| Biobank flag | 4001345 |  | 0 | Unknown/Other |  |
| Admitting source | 4145666 |  | 38004205 | Adult Foster Home |  |
| Admitting source | 4145666 |  | 38004301 | Assisted Living Facility |  |
| Admitting source | 4145666 |  | 38004207 | Ambulatory Visit |  |
| Admitting source | 4145666 |  | 8870 | Emergency Department |  |
| Admitting source | 4145666 |  | 38004195 | Home Health |  |
| Admitting source | 4145666 |  | 8536 | Home / Self Care |  |
| Admitting source | 4145666 |  | 8546 | Hospice |  |
| Admitting source | 4145666 |  | 38004279 | Other Acute Inpatient Hospital |  |
| Admitting source | 4145666 |  | 8676 | Nursing Home (Includes ICF) |  |
| Admitting source | 4145666 |  | 8920 | Rehabilitation Facility |  |
| Admitting source | 4145666 |  | 44814680 | Residential Facility | 60 |
| Admitting source | 4145666 |  | 8863 | Skilled Nursing Facility |  |
| Admitting source | 4145666 |  | 44814682 | No information | 60 |
| Admitting source | 4145666 |  | 44814683 | Unknown | 60 |
| Admitting source | 4145666 |  | 44814684 | Other | 60 |
| Discharge disposition (See Note 6) | 44813951 | 1 | 4161979 | Discharged alive |  |
| Discharge disposition | 44813951 | 1 | 4216643 | Expired |  |
| Discharge disposition | 44813951 | 1 | 44814687 | No information | 60 |
| Discharge disposition | 44813951 | 1 | 44814688 | Unknown | 60 |
| Discharge disposition | 44813951 | 1 | 44814689 | Other | 60 |
| Discharge status (see Note 6) | 4137274 |  | 38004205 | Adult Foster Home |  |
| Discharge status | 4137274 |  | 38004301 | Assisted Living Facility |  |
| Discharge status | 4137274 |  | 4021968 | Against Medical Advice |  |
| Discharge status | 4137274 |  | 44814693 | Absent without leave | 60 |
| Discharge status | 4137274 |  | 4216643 | Expired |  |
| Discharge status | 4137274 |  | 38004195 | Home Health |  |
| Discharge status | 4137274 |  | 8536 | Home / Self Care |  |
| Discharge status | 4137274 |  | 8546 | Hospice |  |
| Discharge status | 4137274 |  | 38004279 | Other Acute Inpatient Hospital |  |
| Discharge status | 4137274 |  | 8676 | Nursing Home (Includes ICF) |  |
| Discharge status | 4137274 |  | 8920 | Rehabilitation Facility |  |
| Discharge status | 4137274 |  | 44814701 | Residential Facility | 60 |
| Discharge status | 4137274 |  | 8717 | Still In Hospital |  |
| Discharge status | 4137274 |  | 8863 | Skilled Nursing Facility |  |
| Discharge status | 4137274 |  | 44814705 | Unknown |  |
| Discharge status | 4137274 |  | 44814706 | Other |  |
| Discharge status | 4137274 |  | 44814704 | No information |  |
| Chart availability  (See Note 5) | 4030450 |  | 4188539 | Yes |  |
| Chart availability | 4030450 |  | 4188540 | No |  |
| Chart availability | 4030450 |  | 0 | Unknown/Other |  |
| Chart availability | 4030450 |  | NULL | No information |  |
| Vital | 3013762 |  | See Note 1 | Weight |  |
| Vital | 3023540 |  | See Note 1 | Height |  |
| Vital | 3034703 |  | See Note 2 | Diastolic Blood Pressure - Sitting |  |
| Vital | 3019962 |  | See Note 2 | Diastolic Blood Pressure - Standing |  |
| Vital | 3013940 |  | See Note 2 | Diastolic Blood Pressure - Supine |  |
| Vital | 3012888 |  | See Note 2 | Diastolic BP – Unknown/Other |  |
| Vital | 3018586 |  | See Note 2 | Systolic Blood Pressure - Sitting |  |
| Vital | 3035856 |  | See Note 2 | Systolic Blood Pressure - Standing |  |
| Vital | 3009395 |  | See Note 2 | Systolic Blood Pressure - Supine |  |
| Vital | 3004249 |  | See Note 2 | Systolic BP – Unknown/Other |  |
| Vital | 3038553 |  | See Note 1 | BMI |  |
| Vital source | 44814721 | 39 | See Note 3 | Patient reported |  |
| Vital source | 38000280 |  | See Note 3 | Healthcare delivery setting |  |

**Note 1**: For height, weight and BMI observations, insert the recorded measurement into the value\_as\_numeric field.

**Note 2**: Systolic and diastolic pressure measurements will generate two observation records – one for storing the systolic blood pressure measurement and a second for storing the diastolic blood pressure measurement. Select the right SBP or DBP concept code that also represents the CORRECT recording position (supine, sitting, standing, other/unknown). To tie the two measurements together, use the observation\_id assigned to the **systolic** blood pressure measurement and insert into the value\_as\_concept\_id field of both observations records (the systolic BP measurement and the diastolic BP measurement records). This will provide a direct linkage between the SBP measurement and its associated DBP measurement.

Example: Person\_id = 12345 on visit\_occurrence\_id = 678910 had orthostatic blood pressure measurements performed as follows:

Supine: Systolic BP 120; Diastolic BP 60  
 Standing: Systolic BP 144; Diastolic BP 72

Four rows will be inserted into the observation table. Showing only the relevant columns:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Observation\_id** | **Person\_id** | **Visit\_occurrence\_id** | **Observation\_concept\_id** | **Observation\_type\_concept\_id** | **Value\_as\_Number** | **Value\_as\_String** | **Value\_as\_Concept\_ID** |
| 66661 | 12345 | 678910 | 3009395 | 38000280 | 120 |  | 66661 |
| 66662 | 12345 | 678910 | 3013940 | 38000280 | 60 |  | 66661 |
| 66663 | 12345 | 678910 | 3035856 | 38000280 | 144 |  | 66663 |
| 66664 | 12345 | 678910 | 3019962 | 38000280 | 72 |  | 66663 |

Observation\_concept\_id = 3009395 = systolic BP - supine; observation\_concept\_id = 3013940 = diastolic BP – supine  
Observation\_concept\_id = 3035856 = systolic BP – standing; observation\_concept\_id = 3019962 = diastolic BP – standing

Observation\_type\_concept\_id = 38000280 (observation recorded from EMR).

Value\_as\_concept\_id = 66661 links two observations for supine BPs to the observation ID of the supine systolic BP.  
Value\_as\_concept\_id = 66663 links two observations for standing BPs to the observation ID of the standing systolic BP.

**Note 3**: Vital source concept\_ids are used as values for the observation\_type\_concept\_id field

**Note 4**: For DRG, use the following logic (must use vocabulary version 4.5):

* The DRG value must be three digits as text. Put into “value\_as\_string” in observation
* For all DRGs, set observation\_concept\_id = 3040464 (hospital discharge DRG)
* To obtain correct value\_as\_concept\_id for the DRG:
  + If the date for the DRG < 10/1/2007, use concept\_class = “DRG”, invalid\_date = “9/30/2007”, invalid\_reason = ‘D’ and the DRG value=CONCEPT.concept\_code to query the CONCEPT table for correct concept\_id to use as value\_as\_concept\_id.
  + If the date for the DRG >=10/1/2007, use concept\_class = “MS-DRG”, invalid\_reason = NULL and the DRG value = CONCEPT.concept\_code to query the CONCEPT table for the correct concept\_id to use as value\_as\_concept\_id.

In addition, the following observations are derived via the DCC (concept\_ids to be assigned in future version of this document. However, concept\_ids are not needed for ETL since these observations will be derived/calculated using scripts developed by DCC):

* Body mass index in kg/m2 if not directly extracted
* Height/length z score for age/sex using NHANES 2000 norms for measurements at which the person was <240 months of age. In the absence of a height/length type for the measurement, recumbent length is assumed for ages <24 months, and standing height thereafter.
* Weight z score for age/sex using NHANES 2000 norms for measurements at which the person was <240 months of age.
* BMI z score for age/sex using NHANES 2000 norms for visits at which the person was between 20 and 240 months of age.
* Systolic BP z score for age/sex/height using NHBPEP task force fourth report norms.
* Diastolic BP z score for age/sex/height using NHBPEP task force fourth report norms.

**Note 5**: In the Observation table, the biobank flag and chart availability concept\_ids can appear multiple times capturing changes in patient consent over time. The temporally most recent observation will be used to determine the current consent status.

**Note 6:** Discharge disposition and discharge status appear only once per visit\_occurence. These vales can change across different visit\_occurrences. Use the visit\_occurrence\_id to tie these observations to the corresponding visit.

|  |  |  |  |
| --- | --- | --- | --- |
| Field | Required | Description | PEDSnet Conventions |
| observation\_id | Yes | A unique identifier for each observation. | This is not a value found in the EHR. Sites may choose to use a sequential value for this field |
| person\_id | Yes | A foreign key identifier to the person about whom the observation was recorded. The demographic details of that person are stored in the person table. |  |
| observation\_concept\_id | Yes | A foreign key to the standard observation concept identifier in the Vocabulary. | Valid Observation Concepts belong to the "Observation" domain. Observation Concepts are based mostly on the LOINC vocabulary, with some additions from SNOMED-CT.  Observations must have an object represented as a concept, and a finding. A finding (see below) is represented as a concept, a numerical value or a verbatim string or more than one of these.  There are three Standard Vocabularies defined for observations:  Laboratory tests and values: Logical Observation Identifiers Names and Codes (**LOINC**) (vocabulary\_id 6).  (FYI: Regenstrief also maintains the "**LOINC Multidimensional Classification**" – vocabulary\_id 49)  Qualitative lab results: A set of SNOMED-CT Qualifier Value concepts (vocabulary\_id 1)  Laboratory units: Unified Code for Units of Measure (**UCUM**( )vocabulary\_id 11)  All other findings and observables: SNOMED-CT – (vocabulary\_id 1).  For vital signs, pull information from flow sheet rows (EPIC sites only) |
| observation\_date | Yes | The date of the observation (UTC). | No date shifting |
| observation\_time | No | The time of the observation (UTC). | No date shifting |
| observation\_type\_concept\_id | Yes | A foreign key to the predefined concept identifier in the Vocabulary reflecting the type of the observation. | Please include valid concept ids (consistent with OMOP CDMv4). Predefined value set (valid concept\_ids found in CONCEPT table where vocabulary\_id = 39)  select \* from concept where vocabulary\_id = 39 yields 7 valid concept\_ids.  FOR PEDSnet CDM V1, all of our observations are coming from electronic health records so set this field to concept\_id = 38000280 (observation recorded from EMR). When we get data from patients, we will include the non-standard concept\_id = 44814721 from vocabulary 99 |
| value\_as\_number | No\*  (see convention) | The observation result stored as a number. This is applicable to observations where the result is expressed as a numeric value. | Value must be represented as at least one of {value\_as\_number, value\_as\_string or values\_as\_concept\_id}. There are a few exceptions in vocabulary id 99 where all three value\_as\_\* fields are NULL. |
| value\_as\_string | No\*  (see convention) | The observation result stored as a string. This is applicable to observations where the result is expressed as verbatim text. | Value must be represented as at least one of {value\_as\_number, value\_as\_string or values\_as\_concept\_id}. There are a few exceptions in vocabulary id 99 where all three value\_as\_\* fields are NULL. |
| value\_as\_concept\_id | No\*  (see convention) | A foreign key to an observation result stored as a concept identifier. This is applicable to observations where the result can be expressed as a standard concept from the Vocabulary (e.g., positive/negative, present/absent, low/high, etc.). | Value must be represented as at least one of {value\_as\_number, value\_as\_string or values\_as\_concept\_id}. There are a few exceptions in vocabulary id 99 where all three value\_as\_\* fields are NULL. |
| unit\_concept\_id | No | A foreign key to a standard concept identifier of measurement units in the Vocabulary. | Please include valid concept ids (consistent with OMOP CDMv4). Predefined value set (valid concept\_ids found in CONCEPT table where vocabulary\_id = 11)  select \* from concept where vocabulary\_id = 11 yields 766 valid concept\_ids.  If none are correct, use concept\_id = 0  For the PEDSnet observation listed above, use the following concept\_ids:  Centimeters (cm): concept\_id = 8582  Kilograms (kg): concept\_id = 9529  Kilograms per square meter (kg/m2): concept\_id = 9531  Millimeters mercury (mmHG): concept\_id = 8876 |
| associated\_provider\_id | No | A foreign key to the provider in the provider table who was responsible for making the observation. |  |
| visit\_occurrence\_id | No | A foreign key to the visit in the visit table during which the observation was recorded. |  |
| relevant\_condition\_concept\_id | No | A foreign key to the condition concept related to this observation, if this relationship exists in the source data (*e.g.* indication for a diagnostic test). |  |
| observation\_source\_value | No | The observation code as it appears in the source data. This code is mapped to a standard concept in the Vocabulary and the original code is, stored here for reference. |  |
| unit\_source\_value | No | The source code for the unit as it appears in the source data. This code is mapped to a standard unit concept in the Vocabulary and the original code is, stored here for reference. |  |
| range\_low | No | Optional - Do not transmit to DCC  The lower limit of the normal range of the observation. It is not applicable if the observation results are non-numeric or categorical, and must be in the same units of measure as the observation value. |  |
| range\_high | No | Optional - Do not transmit to DCC  The upper limit of the normal range of the observation. It is not applicable if the observation results are non-numeric or categorical, and must be in the same units of measure as the observation value. |  |

### additional notes

* The 1/1/2009 date limitation that is used to define a PEDSnet active patient is \*\*NOT\*\* applied to observations. All observations are included for an active patient. For PEDSnet CDM V1, we limit observations to only those that appear in Table 1.
* Observations have a value represented by one of a concept ID, a string, \*\*OR\*\* a numeric value.
* The Visit during which the observation was made is recorded through a reference to the VISIT\_OCCURRENCE table. This information is not always available.
* The Provider making the observation is recorded through a reference to the PROVIDER table. This information is not always available.

## OBSERVATION PERIOD

The observation period table is designed to capture the time intervals in which data are being recorded for the person. An observation period is the span of time when a person is expected to have the potential of drug and condition information recorded. This table is used to generate the PCORnet CDM enrollment table.

While analytic methods can be used to calculate gaps in observation periods that will generate multiple records (observation periods) per person, for PEDSnet, the logic has been simplified to generate a single observation period row for each patient.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Field | Required | Data Type | Description | PEDSnet Conventions |
| observation\_period\_id | Yes | Integer | A system-generate unique identifier for each observation period | This is not a value found in the EHR. Sites may choose to use a sequential value for this field. |
| person\_id | Yes | Integer | A foreign key identifier to the person who is experiencing the condition. The demographic details of that person are stored in the person table. |  |
| observation\_period\_start\_date | Yes | Date | The start date of the observation period for which data are available from the data source | Use the earliest encounter date available for this patient.  No date shifting |
| observation\_period\_end\_date | No | Date | The end date of the observation period for which data are available from the source. | Use the last encounter date available for this patient. If there exists one or more records in the DEATH table for this patient, use the latest date recorded in that table.  For patients who are still in the hospital or ED or other facility at the time of data extraction, leave this field NULL. |

### Additional Notes

* Because the 1/1/2009 date limitation for “active patients” is not used to limit visit\_occurrance, the start\_date of an observation period for an active PEDSnet patient may be prior to 1/1/ 2009.

**APPENDIX**

**PEDSnet-specific Vocabulary 99 has been supplanted by OMOP-supported Vocabulary 60, which contains all of the additional concept\_id codes needed in PEDSnet for PCORnet CDM V1.0**

The INSERT statements that created Vocabulary 99 have been removed from this Appendix based on the current use of Vocabulary 60.

**Elements for future versions**

|  |  |  |  |
| --- | --- | --- | --- |
| Date requested | Requestor | Data request | Target PEDSnet  DM Version |
| 10/24/2014 | Chris Forrest | Prescription meds | 2 |
| 10/24/2014 | Chris Forrest | Lab results: A1C, TC, HDL, TG, LDL, glucose insulin | 2 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |