ETL Conventions for use with PEDSnet CDM v3.2 OMOP V5.2

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 3.2 of the PEDSnet CDM reflects the ETL processes developed after several iterations of network development. As such, it proposes to align with version 4.1 of the PCORnet CDM.

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, 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 PEDSnet CDM v3.2 is based on the Observational Medical Outcomes Partnership (OMOP) data model, version 5.2.
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 5 vocabularies for PEDSnet CDM v3.2, using the complete (restricted) version that includes licensed terminologies such as CPT and others.
7. PCORnet CDM v4.1 requires data elements that are not currently considered "standard concepts". Vocabulary version 5 has a new vocabulary (vocabulary_id=PCORNet) 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=PCORNet where there are no existing standard concepts. We highlight where we are pulling concept_ids from vocabulary_id=PCORNet in the tables. While terms from vocabulary_id=PCORNet violates the OMOP rule to use only concept_ids from standard vocabularies vocabulary_id=PCORNet 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 PERSON.patient_source_value PERSON.person_id, that is used as a primary key in PERSON 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. The PCORnet CDM has specific definitons for null values (as seen below). For the PEDSNet CDM, please use the following logic on which concept value to use for `source\_concept\_id` fields where there are null values in the source `*\_source\_value`.

Null Name	Definition of each field
NULL	A data field is not present in the source system. Note. This is not a 'NULL' string but the NULL value.
'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

Guidelines for populating `'*\_concept\_id'`, `'*\_source\_concept\_id'` and `'*\_source\_value'` for flavors of null:

Null Name	'*_concept_id'	'*_source_concept_id'	'*_source_value'
'NI'	44814650	0	value as in source (leave as null)
'UN'	44814653	0	value as in source (denoting an unknown value)
'OT'	44814649	0	value as in source

10. **For populating `'*\_source\_concept\_id'` (where there exists non-null values in the source) use the following Logic :**

Populate `'*\_source\_concept\_id'` (i.e. non-zero) if the source_value is drawn from a standard vocabulary in OMOP.

Please use your local system knowledge to determine this or use the following criteria : All the values in the source_value field should be drawn from the concept_code in the concept table (for a given/relevant domain_id and a given vocabulary_id).

ELSE Use 0

(usually the case when the sites need to "manually" map the foo_source_value to foo_concept_id)

11. **For populating `*\_source\_value` please make a best effort to provide "human readable" values rather than a coded value where possible from the source.**

Example for `gender\_source\_value`, the source value at your site may be `1` for Female and `2` for Male. Please provide the label value of `Female` and `Male`.

ETL Recommendation: Due to PK/FK constraints, the most efficient order for ETL table is location, care_site, provider, person, visit_occurrence, condition_occurrence, observation, procedure_occurrence,measurement,measurement_organism,drug exposure

Table of Contents

[Appendix]

Data Extraction Guide

Please use the table headings as a guide in extracting and submitting data. These specifications are indicative of DCC and Network Requirements. All fields must be submitted to the DCC even if you are not submitting data in a field. Here are examples of how the specification should be interpreted:

Field	NOT Null Constraint	Network Requirement	Data Type	Description	PEDSnet Conventions
Field Name	• Yes	• Yes	Data Type	Description	PEDSnet Conventions

- **The above example indicates the data in this field is required by both the DCC and Network. It absolutely must be provided in the data submission.**

Field	NOT Null Constraint	Network Requirement	Data Type	Description	PEDSnet Conventions
Field Name	• No	• Provide When Available	Data Type	Description	PEDSnet Conventions

- **The above example indicates the data in this field is required by Network if it is populated or available at your site. If it is available it must provided in the data submission.**

Field	NOT Null Constraint	Network Requirement	Data Type	Description	PEDSnet Conventions
Field Name	• No	• Site Preference	Data Type	Description	PEDSnet Conventions

- **The above example indicates the data in this field is not required by the DCC or Network. A site may choose to send this information if they desire to do so.**

Field	NOT Null Constraint	Network Requirement	Data Type	Description	PEDSnet Conventions
Field Name	• No	• Optional	Data Type	Description	PEDSnet Conventions

- **The above example indicates the data in this field is truly optional for submission. A site may choose to send this information if they desire to do so.**

1.1 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 telephone encounter or a lab blood draw does not meet this definition.

For reference `Visit\_concept\_ids` that correspond to an "in person" clinical encounter are:

Visit Type	Visit_concept_id
Inpatient Hospital Stay	9201
Ambulatory/Outpatient Visit (With a Physician)	9202
Outpatient Non Physician	2000000469
Emergency Department	9203
Long Term Care Visit	42898160
Non-Acute Institutional Stay	44814710
Emergency Department Admit to Inpatient Hospital Stay (If sites are unable to split the encounter)	2000000048
Observation Stay	2000000088

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	NOT Null Constraint	Network Requirement	Data Type	Description	PEDSnet Conventions
person_id	Yes	Yes	Integer	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. **SITE RESPONSIBILITY: This field must remain a stable identifier across submissions to the DCC.** A mapping from the person_id 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.
gender_concept_id	Yes	Yes	Integer	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 CDMv5.1). Predefined value set (valid concept_ids found in CONCEPT table select * from concept where ((domain_id='Gender' and concept_class_id='Gender')or (domain_id='Observation' and vocabulary_id='PCORNet' and concept_class_id in ('Gender','Undefined'))) and concept_code not in ('Sex-F','Sex-M') and invalid_reason is null: • Ambiguous: concept_id = 44814664 • Female: concept_id = 8532 • Male: concept_id = 8507 • No Information: concept_id = 44814650 (Vocabulary_id='PCORNet') • Unknown: concept_id = 44814653 • Other: concept_id = 44814649
gender_source_concept_id	Yes	Yes	Integer	A foreign key to the gender concept that refers to the code used in the source.	**If there is not a mapping for the source code in the standard vocabulary, use concept_id = 0**
year_of_birth	Yes	Yes	Integer	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	Provide When Available	Integer	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	Provide When Available	Integer	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)
		Provide		The birth date and	Do not include timezone.

birth_datetime	No	When Available	Datetime	time	Please keep all accurate/real dates (No date shifting). If there is no time associated with the date assert midnight.
race_concept_id	Yes	Yes	Integer	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 ((domain_id='Race' and vocabulary_id = 'Race') or (vocabulary_id='PCORNet' and concept_class_id='Undefined') or concept_id in (44814659,44814660)) and invalid_reason is null: • 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_id='PCORNet') • Refuse to answer: concept_id = 44814660 (vocabulary_id='PCORNet') • No Information: concept_id = 44814650 vocabulary_id='PCORNet') • Unknown: concept_id = 44814653 • Other: concept_id = 44814649
race_source_concept_id	Yes	Yes	Integer	A foreign key to the race concept that refers to the code used in the source.	**If there is not a mapping for the source code in the standard vocabulary, use concept_id = 0**
ethnicity_concept_id	Yes	Yes	Integer	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 CDMv5). Predefined value set (valid concept_ids found in CONCEPT table where vocabulary_id ='Ethnicity' or (vocabulary_id=PCORNet and concept_class_id='Undefined) where noted): • Hispanic: concept_id = 38003563 • Not Hispanic: concept_id = 38003564 • No Information: concept_id = 44814650 (vocabulary_id='PCORNet') • Unknown: concept_id = 44814653 (vocabulary_id='PCORNet') • Other: concept_id = 44814649 (vocabulary_id='PCORNet')
ethnicity_source_concept_id	Yes	Yes	Integer	A foreign key to the ethnicity concept that refers to the code used in the	**If there is not a mapping for the source code in the standard vocabulary, use concept_id = 0**

				source.	
location_id	No	Provide When Available	Integer	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	Provide When Available	Integer	Foreign key to the primary care provider the person is seeing in the provider table.	For PEDSnet CDM v3.2.0: Sites will use site-specific logic to determine the best primary care provider and document how that decision was made (e.g., billing provider).
care_site_id	Yes	Yes	Integer	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	Provide When Available	Integer	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	No	Provide When Available	Varchar	An encrypted key derived from the person identifier in the source data.	Insert a unique 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.
gender_source_value	Yes	Yes	Varchar	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	Yes	Yes	Varchar	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 concept_id for Multiple Race.
ethnicity_source_value	Yes	Yes	Varchar	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.
language_concept_id	Yes	Yes	Integer	A foreign key that refers to the standard concept identifier in the Vocabulary for the language of the person.	For PEDSNet, please map your source codes to acceptable language values in appendix 2 **If there is not a mapping for the source code in the network language mapping, use concept_id = 44814649 (Other PCORNet Vocabulary)**
language_source_concept_id	Yes	Yes	Integer	A foreign key to the language concept that refers to the	**If there is not a mapping for the source code in the standard**

				code used in the source.	**vocabulary, use concept_id = 0**
language_source_value	Yes	Yes	Varchar	The source code for the language of the person as it appears in the source data	The person language is mapped to a standard language concept in the Vocabulary and the original code is stored here for reference.

If a field marked as "Provide when available" for the network requirement is not available at your site, please relay this information to the DCC

1.2 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	NOT Null Constraint	Network Requirement	Data Type	Description	PEDSnet Conventions
death_cause_id	Yes	Yes	Integer	A unique identifier for each death cause occurrence	This is not a value found in the EHR. Sites may choose to use a sequential value for this field
person_id	Yes	Yes	Integer	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	Yes	Date	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_datetime	Yes	Yes	Datetime	The date the person was deceased.	**This field is custom to PEDSnet** 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. If there is no time associated with the date assert '23:59:59'.
death_type_concept_id	Yes	Yes	Integer	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 CDMv5). Predefined value set (valid concept_ids found in CONCEPT table where domain_id ='Death Type') select * from concept where concept_class_id ='Death Type' yields 9 valid concept_ids. If none are correct, use concept_id = 0 Note: Most current ETLs are extracting data from EHR. The common concept_id to insert here is • 38003569 ("EHR record patient status "Deceased") . Please assert • No information: concept_id = 44814650 where there is no information in the source **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_concept_id	No	Provide When Available	Integer	A foreign referring to a standard concept identifier in the Vocabulary for conditions.	
cause_source_value	No	Provide When Available	Varchar	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.	
cause_source_concept_id	No	Provide When Available	Integer	A foreign key to the vocabulary concept that refers to the code used in the source.	This links to the concept id of the vocabulary of the cause of death concept id as stored in the source. For example, if the cause of death is "Acute myeloid leukemia, without mention of having achieved remission" which has an icd9 code of 205.00 the cause source concept id is 44826430 which is the icd9 code concept that corresponds to the diagnosis 205.00. **If there is not a mapping for the source code in the standard vocabulary, use concept_id = 0**
death_impute_concept_id	Yes	Yes	Varchar	A foreign key referring to a standard concept identifier in the vocabulary for death imputation.	p>Please include valid concept ids (consistent with OMOP CDMv5). Predefined value set (valid concept_ids found in CONCEPT table where concept_class_id ='Death Imput Type') select * from concept where (concept_class_id ='Death Imput Type' or (vocabulary_id='PCORNet' and concept_class_id='Undefined')) and invalid_reason is null yields 8 valid concept_ids. If none are correct, use concept_id = 0 • Both month and day imputed: 2000000034 • Day imputed: 2000000035 • Month imputed: 2000000036 • Full Date imputed: 2000000038 • Not imputed:2000000037 • No Information: concept_id = 44814650 (Vocabulary_id='PCORNet') • Unknown: concept_id = 44814653 • Other: concept_id = 44814649

If a field marked as "Provide when available" for the network requirement is not available at your site, please relay this information to the DCC

1.2.1 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.

1.3 LOCATION

The Location domain represents a generic way to capture physical location or address information. Locations are used to define the addresses for Persons and Care Sites. The most important field is ZIP for location-based queries.

Field	NOT Null Constraint	Network Requirement	Data Type	Description	PEDSnet Conventions
location_id	Yes	Yes	Integer	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	Provide When Available	Varchar	The state field as it appears in the source data.	
zip	No	Provide When Available	Varchar	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	Provide When Available	Varchar	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	NO	Varchar		Do not transmit to DCC
address_2	No	NO	Varchar		Do not transmit to DCC
city	No	NO	Varchar		Do not transmit to DCC
county	No	NO	Varchar		Do not transmit to DCC

If a field marked as "Provide when available" for the network requirement is not available at your site, please relay this information to the DCC

1.3.2 Additional Notes

- Each address or Location is unique and is present only once in the table
- Locations in this table are restricted to locations that are applicable to persons and care_sites in the Pedsnet cohort at each site. When external data is implemented, valid(data containing) locations may be expanded beyond locations of those only present in clinical tables.

1.4 CARE_SITE

The Care Site domain contains a list of uniquely identified physical or organizational units where healthcare delivery is practiced (offices, wards, hospitals, clinics, etc.).

Field	NOT Null Constraint	Network Requirement	Data Type	Description	PEDSnet Conventions
care_site_id	Yes	Yes	Integer	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
care_site_name	No	Provide When Available	Varchar	The description of the care site	
					Please include valid concept ids (consistent with OMOP CDMv5.1). Predefined value set (valid concept_ids found in CONCEPT table where concept_class_id = 'Place of Service' and invalid_reason is null)

place_of_service_concept_id	No	Provide When Available	Integer	A foreign key that refers to a place of service concept identifier in the Vocabulary	select * from concept where concept_class_id = 'Place of Service' and invalid_reason is null yields 49 valid concept_ids. Please use the following value set for PEDSnet v3.2: • **Urgent Care Facility = 8782** • Rural Health Clinic = 8761 • Outpatient (Examples: Hospital Dialysis, HOD, Day Hospital, Day Medicine) = 8756 • Office =8940 • Inpatient Psychiatric Facility =8971 • Inpatient Hospital =8717 • Independent Clinic =8716 • Emergency Room - Hospital = 8870 • Other Place of Service =8844 • Other Inpatient Care =8892 • Unknown: concept_id = 44814653 • Other: concept_id = 44814649 • No information: concept_id = 44814650
location_id	No	Provide When Available	Integer	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	Yes	Varchar	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	Provide When Available	Varchar	The source code for the place of service as it appears in the source data, stored here for reference.	
specialty_concept_id	No	Provide When Available	Integer	The specialty of the department linked to a standard specialty concept as it appears in the	Care sites could have one or more specialties or a Care site could have no specialty information. **Valid specialty concept ids for PEDSnet are found in the appendix** **Please use the following rules:** • If care site specialty information is unavailable, please follow the convention on reporting values that are unknown,null or unavailable. • If a care site has a single specialty associated with it, sites should link the specialty to the **valid specialty concepts as assigned in the appendix**. If the specialty does not correspond to a value in this listing, please use the NUCC Listing (vocabulary_id='NUCC') provided in the vocabulary as a reference. • If there are multiple specialties associated with a particular care site and sites are not able to assign a specialty value on

				Vocabulary	the visit occurrence level, sites should use the specialty concept id=38004477 "Pediatric Medicine". • If there are multiple specialties associated with a particular care site and this information is attainable, sites should document the strategy used to obtain this information and the strategy used to link the correct care site/specialty pair for each visit occurrence. Sites should also link the specialty to the **valid specialty concepts as assigned in the** appendix If the specialty does not correspond to a value in this listing, please use the NUCC Listing (vocabulary_id='NUCC') provided in the vocabulary as a reference. • If the speciality does not correspond to a value in the NUCC Listing and no value in the ABMS Listing, please use the Specialty listing (vocabulary_id='Specialty') as a reference
specialty_source_value	No	Provide When Available	Varchar	The source code for the specialty as it appears in the source data, stored here for reference.	

If a field marked as "Provide when available" for the network requirement is not available at your site, please relay this information to the DCC

1.4.1 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 Place of Service Concepts are based on a catalog maintained by the CMS (see vocabulary for values)

1.5 PROVIDER

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

Field	NOT Null Constraint	Network Requirement	Data Type	Description	PEDSnet Conventions
provider_id	Yes	Yes	Integer	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. **SITE RESPONSIBILITY: This field must remain a stable identifier across submissions to the DCC.** A mapping from the provider_id to a real provider 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 should document who they have included as a provider.
provider_name	No	NO	Varchar	A description of the provider	DO NOT TRANSMIT TO DCC
gender_concept_id	No	Provide When Available	Integer	The gender of the provider	A foreign key to the concept that refers to the code used in the source.
specialty_concept_id	No	Provide When Available	Integer	A foreign key to a standard provider's specialty concept identifier in the	Please map the source data to the mapped provider specialty concept associated with the American Medical Board of Specialties as seen in Appendix A1. Predefined value set (valid concept_ids found in CONCEPT table where domain_id='Provider Specialty' and vocabulary_id in ('Specialty', 'ABMS','NUCC','PEDsnet')) select * from concept where domain_id ='Provider Specialty' and vocabulary_id in ('Specialty', 'ABMS','NUCC','PEDsnet') and invalid_reason is null yields 1025 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

				Vocabulary.	specialty and document the logic used. For example, sites may decide to always assert the **first** specialty listed in their data source. As a first guide please use the ABMS and PEDsnet vocabulary specialty listing listing to map your specialtity values. If the specialty does not correspond to a value in these listings, please use the NUCC Listing (vocabulary_id='NUCC') provided in the vocabulary as a reference and the Specialty (vocabulary_id='Specialty') if no correspond value exists in the NUCC Listing.
care_site_id	Yes	Yes	Integer	A foreign key to the main care site where the provider is practicing.	See CARE_SITE.care_site_id (primary key)
year_of_birth	No	Provide When Available	Integer	The year of birth of the provider	
NPI	No	Site Preference	Varchar	The National Provider Identifier (NPI) of the provider.	
DEA	No	Site Preference	Varchar	The Drug Enforcement Administration (DEA) number of the provider.	
provider_source_value	Yes	Yes	Varchar	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	Provide When Available	Varchar	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.
specialty_source_concept_id	No	Provide When Available	Integer	A foreign key to a concept that refers to the code used in the source.	If providing this information, sites should document how they determine the specialty associated with the provider. **Valid specialty concept ids for PEDSnet are found in the appendix** If the specialty does not correspond to a value in this listing, please use the NUCC Listing (vocabulary_id='NUCC') provided in the vocabulary as a reference. **If there is not a mapping for the source code in the standard vocabulary, use concept_id = 0**
gender_source_value	No	Provide When Available	Varchar	The source value for the provider gender.	
gender_source_concept_id	No	Provide When	Integer	The gender of the provider as represented in the source that	**If there is not a mapping for the source code in the standard vocabulary, use concept_id = 0**

		Available		maps to a concept in the vocabulary	

If a field marked as "Provide when available" for the network requirement is not available at your site, please relay this information to the DCC

1.5.1 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.

1.6 VISIT_OCCURRENCE

The visit occurrence 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.

Exclusions:

1. Future Vists
2. Cancelled Visits (where the patient was not seen)

Note 1: Please use the following logic to assign visit concept ids:

Visit Concept Id	Concept Name	Visit Type Inclusion	In Person	Examples/Logic (includes but is not limited to)
9201	Inpatient Visit (IP)	Visits that resulted in a patient admission	Yes	Hospital Admissions
9202	Ambulatory Visit (AV)/Outpatient	In person Outpatient Visits visits where the patient was seen by a physician	Yes	Office Visits or Appointments
2000000469	Outpatient Non Physician (OP-Non Physician)	In person Outpatient Visits visits where the patient was **NOT** seen by a physician	Yes	Lab Visits, Radiology
9203	Emergency Department Visit (ED)	Emergency Department Visits and Urgent Care	Yes	Emergency Room Visits and Urgent Care
44814711	Other ambulatory Visit (OA)	Outpatient visits where the patient was not seen in person.	No	Telemedicine,Telephone, Emails, Refills and Orders Only Encounters
42898160	Long Term Care Visit	Formal or Informal long term care for chronic illness management	Yes	Site discretion
44814710	Non-Acute Institutional	Non-Acute long term management of care	Yes	Site discretion
2000000048	Emergency Department Admit to Inpatient Hospital Stay	Combination of 9203 and 9201 visits	Yes	Use only if unable to split the ED and inpatient visit.
2000000088	Observation Visit	Please discern what defines an observation visit at your site	Yes	Only map to the observation visit type if the patient leaves the hospital or is discharged from what has been determined to be an observation visit. For sites splitting visits, ED->Observation visits are only to be mapped as Observation Stay Visits. The split in this case is not required.
2000000104	Administrative Visit	Other visits that are in the source system for administrative purposes.	No	Professional Billing or Hospital Abstractions

Field	NOT Null Constraint	Network Requirement	Data Type	Description	PEDSnet Conventions

visit_occurrence_id	Yes	Yes	Integer	A unique identifier for each person's visits or encounter at a healthcare provider.	This is not a value found in the EHR. VISIT_OCCURRENCE_ID must be unique for all patients within a single data set. **SITE RESPONSIBILITY: This field must remain a stable identifier across submissions to the DCC.** A mapping from the visit occurrence id to a real patient encounter 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. Do not use institutional encounter ID.
person_id	Yes	Yes	Integer	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	Yes	Date	The start date of the visit.	No date shifting. Full date.
visit_end_date	No	Provide When Available	Date	The end date of the visit.	No date shifting. Full date. 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.
visit_start_datetime	Yes	Yes	Datetime	The start date of the visit.	No date shifting. Full date and time. **If there is no time associated with the date assert midnight for the start time**
visit_end_datetime	No	Provide When Available	Datetime	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. Full date and time. **If there is no time associated with the date assert 11:59:59 pm for the end time**
provider_id	No	Provide When Available	Integer	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 required in OMOP CDM v4, but appears in OMOP CDMv5.**
care_site_id	No	Provide When Available	Integer	A foreign key to the care site in the care site table that was visited.	See CARE_SITE.care_site_id (primary key)
visit_concept_id	Yes	Yes	Integer	A foreign key that refers to a place of service concept identifier in the vocabulary.	**In PEDSnet CDM v1, this field was previously called place_of_service_concept_id** Please include valid concept ids (consistent with OMOP CDMv5). Predefined value set (valid concept_ids found in CONCEPT table where (domain_id='Visit' and ~ 'Enc
visit_type_concept_id	Yes	Yes	Integer	A foreign key to the predefined concept identifier in the standard vocabulary reflecting the type of source data from which the visit	select * from concept where concept_class_id='Visit Type' yields 3 valid concept_ids. If none are correct, user concept_id=0. The majority of visits should be type 'Visit derived from EHR record'

				record is derived.	which is concept_id=44818518
visit_source_value	No	Provide When Available	Varchar	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.	
visit_source_concept_id	No	Provide When Available	Integer	A foreign key to a concept that refers to the code used in the source.	If a site is using HCPS or CPT for their visit source value, the standard concept id that maps to the particular vocabulary can be used here. **If there is not a mapping for the source code in the standard vocabulary, use concept_id = 0**
preceding_visit_occurrence_id	No	NO	Integer	A foreign key to the VISIT_OCCURRENCE table record of the visit immediately preceding this visit.	Do not transmit to DCC
admitted_from_concept_id	No	Provide When Available	Integer	A foreign key to the predefined concept in the Place of Service Vocabulary reflecting the admitting source for a visit.	Please use the following valid concept id set for Admitting source: • Adult Foster Home=44814670 • Assisted Living Facility=44814671 • Ambulatory Visit=44814672 • Emergency Department=8870= • Home Health=44814674 • Home / Self Care=44814675 • Hospice=8546 • Other Acute Inpatient Hospital=38004279 • Nursing Home (Includes ICF)=44814678 • Rehabilitation Facility=44814679 • Residential Facility=44814680 • Skilled Nursing Facility=8863 • No information=44814650 • Unknown=44814653 • Other=44814649 This should be populated for inpatient encounters in the source but may vary for emergency department (ED) visits and outpatient encounters (AV,OA).
discharge_to_concept_id	No	Provide When Available	Integer	A foreign key to the predefined concept in the Place of Service Vocabulary reflecting the discharge disposition (destination) for a visit.	Please use the following valid concept id set for Discharge Destination: • Adult Foster Home=38004205 • Assisted Living Facility=38004301 • Against Medical Advice=4021968 • Absent without leave=44814693 • Expired=4216643 • Home Health=38004195 • Home / Self Care=8536 • Hospice=8546 • Other Acute Inpatient Hospital=38004279 • Nursing Home (Includes ICF)=8676 • Rehabilitation Facility=8920 • Residential Facility=44814701 • Still In Hospital=8717 • Skilled Nursing Facility=8863 • No information=44814650 • Unknown=44814653 • Other=44814649 This should be populated for inpatient encounters in the

					source but may vary for emergency department (ED) visits and outpatient encounters (AV,OA).
admitted_from_source_value	No	Provide When Available	Varchar	The source code for the admitting source as it appears in the source data.	This should be populated for inpatient encounters in the source but may vary for emergency department (ED) visits and outpatient encounters (AV,OA).
discharge_to_source_value	No	Provide When Available	Varchar	The source code for the discharge disposition as it appears in the source data.	This should be populated for inpatient encounters in the source but may vary for emergency department (ED) visits and outpatient encounters (AV,OA).

**If a field marked as "Provide when available" for the network requirement is not available at your site, please relay this information to the DCC

1.6.1 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 visit_concepts (ED -> inpatient) sites may opt to split the record into separate visit_occurrence records**.

To show the relationship of the split (ED -> inpatient) encounter, use the FACT_RELATIONSHIP table.

An example of this is below:

VISIT_OCCURRENCE

visit_occurrence_id	person_id	visit_start_date	visit_end_date	provider_id	care_site_id	place_of_service_concept_id	place_of_service_source_value
35022489	209846	2011-11-14 17:36:00-05	2011-11-14 22:25:00-05	2238	322	9203	Emergency
35022490	209846	2011-11-14 22:25:00-05	2011-11-15 16:33:00-05	2238	43	9201	Emergency

FACT_RELATIONSHIP

Domain_concept_id_1	fact_id_1	Domain_concept_id_2	fact_id_2	relationship_concept_id
Visit	35022489	Visit	35022490	Occurs before
Visit	35022490	Visit	35022489	Occurs after

Because the domain_concept_id and relationship_concept_id are actually numeric values the following is an example of how the table is stored:

Domain_concept_id_1	fact_id_1	Domain_concept_id_2	fact_id_2	relationship_concept_id
8	35022489	8	35022490	44818881
8	35022490	8	35022489	44818783

- Operating and Anesthesia encounters that occur as apart of the Inpatient stay should be rolled up into one Inpatient encounter.
- 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.

1.7 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 or ICD-10-CM codes, but may not have a way to capture out-of-system conditions.
- EHRs may also capture External Injury codes in different place in the source system. These types of codes are also to be included.

For the PEDSNet network, please provide **clinical physician based diagnosis** as opposed to billing or claim based diangosis data.

Note 1: For the PEDSNet network, we are coding all diagnosis codes to the SNOMED-CT Vocabulary. Research has showed that the IMO to SNOMED native mapping and IMO to ICD to SNOMED OMOP mapping produces highly variable results. For a particular IMO Code, when comparing the two mapping options, the same SNOMED concept id is only produced 25% of the time. See below examples of the mapping differences (IMO-SNOMED, ICD10 and ICD9):

IMO Description	Direct SNOMED	Via ICD
Numbness of Toes	Numbness of toe	Altered Sensation of Skin
Cerebellar ataxia/dyskinesia	Cerebellar Disorder	Cerebellar Ataxia
Choking episode	Choking sensation	Finding of head and neck region
Intestional malrotation	Congenital malrotation of intestine	Congenital anomaly of fixation of intestine
Genetic disease carrier status testing	Genetic finding	Genetic disorder carrier
Duchenne muscular dystrophy	Duchenne muscular dystrophy	Hereditary progressive muscular dystrophy

For diagnosis codes, please provide the IMO to SNOMED mapping where it exists in the source system.

If the IMO to SNOMED mapping is not available in the system, utilize the IMO to ICD to SNOMED OMOP mapping in the vocabulary.

Please use the following logic to populate the `condition\_concept\_id`, `condition\_source\_concept\_id` and `condition\_source\_value` based on what is available in your source system:

You have in your source system	condition_concept_id	condition_source_concept_id	condition_source_value
Any diagnosis that was captured as a term or name (e.g. IMO to SNOMED)	Corresponding SNOMED concept id	Corresponding concept for site diagnosis captured (must correspond to ICD9/ICD10 concept mapping)	Diagnosis Name "|" IMO Code "|" Diagnosis Code
Any diagnosis that was captured directly as a code (e.g. ICD9/10) by a coder	Corresponding SNOMED concept id	Corresponding concept for site diagnosis code (must correspond to ICD9/ICD10 concept mapping)	Diagnosis Name "|" IMO Code "|" Diagnosis Code

Note 2: For the PEDSNet network, please provide clinical physician based diagnosis as opposed to billing or claim based diangosis data. The clinical physician based diagnosis corresponds to the "Order origin" concept ids for `condition\_type\_concept\_id`. If you are providing billing or claim diagnosis data, please use the "Billing" or "Claim" concept_ids for `condition\_type\_concept\_id`.

Use the following logic to determine the correct `condition\_type\_concept\_id` as it pertains to the visit the diagnosis stems from:

Visit_concept_id	Condition_type_concept_id
9201 (Inpatient)	Inpatient header
9202 (Outpatient)	Outpatient header
9203 (Emergency)	Emergency header
2000000048 (ED to Inpatient)	Inpatient header
2000000088 (Observation)	Inpatient header

Note 3: We have been made aware that there are a significant amount of conditions that route to a domain of Procedure, Measurement etc. Please **DO NOT** route these conditions to those domains or tables (i.e. Procedure_Occurrence, Measurement). Instead, include all records coming out of our source tables for diagnosis data in the Condition_Cccurrence table.

Field	NOT Null Constraint	Network Requirement	Data Type	Description	PEDSnet Conventions
condition_occurrence_id	Yes	Yes	Integer	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	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.	
					Please include valid concept ids (consistent with OMOP CDMv5). Predefined value set (valid concept_ids found

condition_concept_id	Yes	Yes	Integer	A foreign key that refers to a standard condition concept identifier in the Vocabulary.	in CONCEPT table where vocabulary_id ='SNOMED') select * from concept where vocabulary_id ='SNOMED' yields ~440,000 valid concept_ids. If none are correct, use concept_id = 0
condition_start_date	Yes	Yes	Date	The date when the instance of the condition is recorded.	No date shifting.
condition_end_date	No	Provide When Available	Date	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_start_datetime	Yes	Yes	Datetime	The date and time when the instance of the condition is recorded.	No date shifting. Full date and time. **If there is no time associated with the date assert midnight for the start time**
condition_end_datetime	No	Provide When Available	Datetime	The date and time when the instance of the condition is considered to have ended	No date shifting. If this information is not available, set to NULL. Full date and time. **If there is no time associated with the date assert 11:59:59 pm for the end time**
condition_type_concept_id	Yes	Yes	Integer	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 CDMv5). Predefined value set (valid concept_ids found in CONCEPT table where concept_class_id ='Condition Type' and vocabulary_id='PEDSnet') select * from concept where concept_class_id ='Condition Type' and vocabulary_id='PEDSnet' yields 21 valid concept_ids. If none are correct, use concept_id = 0 **For the primary diagnosis for the inpatient, outpatient or emergency setting (may be identified as Dx#1 in a source system), Please use concepts the following concepts:** • Outpatient header - 1st position - Order Origin=2000000095 • Outpatient header - 1st position - Billing Origin=2000000096 • Outpatient header - 1st position - Claim Origin=2000000097 • Inpatient header - primary - Order Origin=2000000092 • Inpatient header - primary - Billing Origin =2000000093 • Inpatient header - primary - Claim Origin= 2000000094 • Emergency Header - 1st Position - Order Origin=2000001280 • Emergency Header - 1st Position - Claim Origin=2000001281 • Emergency Header - 1st Position - Billing Origin=2000001282 **All other diagnosis that is not the primary (or Dx#1) in the inpatient, outpatient or emergency setting should correspond to the following concept ids:** • Inpatient header - 2nd position - Order Origin=2000000098 • Inpatient header - 2nd position - Billing Origin = 2000000099 • Inpatient header - 2nd position - Claim Origin =

					2000000100 • Outpatient header - 2nd position - Order Origin=2000000101 • Outpatient header - 2nd position - Billing Origin =2000000102 • Outpatient header - 2nd position - Claim Origin =2000000103 • Emergency Header - 2nd Position - Order Origin=2000001283 • Emergency Header - 2nd Position - Claim Origin=2000001284 • Emergency Header - 2nd Position - Billing Origin=2000001285 **For diagnosis from the problem list, please use the following concept ids:** • EHR problem list entry - Order Origin = 2000000089 • EHR problem list entry - Billing Origin =2000000090 • EHR problem list entry - Claim Origin =2000000091 **SEE NOTE 2 for further guidance on determining the origin**
stop_reason	No	Provide When Available	Varchar	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.
provider_id	No	Provide When Available	Integer	A foreign key to the provider in the provider table who was responsible for determining (diagnosing) the condition.	**In PEDSnet CDM v1, this field was previously called associated_provider_id** 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	Provide When Available	Integer	A foreign key to the visit in the visit table during which the condition was determined (diagnosed).	
condition_source_value	Yes	Yes	Varchar	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 or ICD-10-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. **Please include the diagnosis name and source code when populating this field, by using the pipe delimiter "|" when concatenating values.** Example: Diagnosis Name "|" IMO Code "|" Diagnosis Code
condition_source_concept_id	No	Provide When Available	Integer	A foreign key to a condition concept that refers to the code used in the source	As a standard convention this code must correspond to the ICD9/ICD10 concept mapping of the source value only. For example, if the condition is "Acute myeloid leukemia, without mention of having achieved remission" which has an icd9 code of 205.00 the condition source concept id is 44826430 which is the icd9 code concept that corresponds to the diagnosis 205.00. **If there is not a mapping for the source code in the standard vocabulary, use concept_id = 0**
condition_status_concept_id	No	Optional	Integer	A foreign key to the predefined concept in the standard vocabulary reflecting the condition status.	For PEDSnet v3.2 we are only reporting final diagnosis, please use the following concept id: • Final Diagnosis=4230359

condition_status_source_value	No	Optional	Varchar	The source code for the condition status as it appears in the source data.	
poa_concept_id	No	Optional	Integer	A foreign key to value in the source for that determines if the diagnosis is present on admission	For Pedsnet CDM v3.2, please use the following: • Yes=4188539 • No=4188540 • No Information: concept_id = 44814650 • Unknown: concept_id = 44814653 • Other: concept_id = 44814649 If none are correct, use concept_id = 0.

If a field marked as "Provide when available" for the network requirement is not available at your site, please relay this information to the DCC

1.7.1 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 v3.2, 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). Medical History diagnosis are optional.
- 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, diagnoses extracted from unstructured data (such as notes) are not included.
- 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 conditions in the Vocabulary. For mapping ICD9 Codes to SNOMED, use the concept_relationship table where the icd9_code = concept_id_1 and relationship_id='Maps to'. Concept_id_2 will be the SNOMED concept_id mapping you need to populate the condition_concept_id.
- 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.

1.8 PROCEDURE_OCCURRENCE

The procedure occurrence 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),ICD-10 (Procedures), HCPCS or OPCS-4 procedures as orders.

More specifically the procedure occurrence domain is intended to stores information about activity or processes involving a patient that has a billable code. This includes but is not limited to the following: - LOS Codes ((Eg. 99123) This code may not Not necessarily be a CPT and could require local mapping) - Lab Procedures (including a Lab Panel Order and Culture Orders) - Surgery Procedures - Imaging Procedures - Ancilliary Therapies (Speech, Physical, Occupational etc)

Notes: **Only instantiated procedures are included in this table. Please exclude cancelled procedures For CPT Codes, only include codes that are included in the standard CPT4 vocabulary from the distributed vocabulary**

Note 1: Please use the following logic to populate the `procedure\_concept\_id` , `procedure\_source\_concept\_id` and `procedure\_source\_value` based on what is available in your source system:

Site Information	procedure_concept_id	procedure_source_concept_id	procedure_source_value
Procedure codes using CPT-4, ICD-9-CM (Procedures),ICD-10 (Procedures), HCPCS or OPCS-4 procedures as orders	Corresponding CPT-4, ICD-9-CM (Procedures),ICD-10 (Procedures), HCPCS or OPCS-4 concept id	Corresponding CPT-4, ICD-9-CM (Procedures),ICD-10 (Procedures), HCPCS or OPCS-4 concept id	Procedure Name I Procedure Source Code
Custom Procedure Coding (That a site has knowledge of corresponding to a standard code but requires manual mapping)	Corresponding CPT-4, ICD-9-CM (Procedures),ICD-10 (Procedures), HCPCS or OPCS-4 concept id	0	Procedure Name I Custom Procedure Code

Field	NOT Null Constraint	Network Requirement	Data Type	Description	PEDSnet Conventions

procedure_occurrence_id	Yes	Yes	Integer	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	Yes	Integer	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	Yes	Integer	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 ='SNOMED'), ICD-9-Procedures (vocabulary_id ='ICD9Proc'),ICD-10-Procedures (vocabulary_id ='ICD10PCS' **NOT YET AVAILABLE**), CPT-4 (vocabulary_id ='CPT4'), and HCPCS (vocabulary_id ='HCPCS') 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.
modifier_concept_id	No	Provide When Available	Integer	A foreign key to a standard concept identifier for a modifier to the procedure (e.g. bilateral)	Valid Modifier Concepts belong to the "Modifier" concept class. select /* from concept where concept_class_id like '%Modifier%'.
quantity	No	Provide When Available	Float	The quantity of procedures ordered or administered.	
procedure_date	Yes	Yes	Date	The date on which the procedure was performed.	
procedure_datetime	Yes	Yes	Datetime	The date and time on which the procedure was performed. If there is no time associated with the date assert midnight.	**This field is a custom PEDSnet field**
procedure_type_concept_id	Yes	Yes	Integer	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 = 'Procedure Type')	Please include valid concept ids (consistent with OMOP CDMv5). Predefined value set (valid concept_ids found in CONCEPT table where vocabulary_id = 'Procedure Type') select * from concept where vocabulary_id ='Procedure Type' yields 93 valid concept_ids. For procedures coming from billing records please map to the following concepts: • Primary Procedure: 44786630 • Secondary Procedure: 44786631 If you are unable to distinguish between primary and secondary procedures. Please map to the following: • Secondary Procedure: 44786631 For procedures coming from physician orders and all other types, please map to the following: • EHR order list entry: 38000275
				A foreign key to the provider in	

provider_id	No	Provide When Available	Integer	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	Provide When Available	Integer	A foreign key to the visit in the visit table during which the procedure was carried out.	See VISIT.visit_occurrence_id (primary key)
procedure_source_value	Yes	Yes	Varchar	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.Please also include the procedure name. See Note 1.
procedure_source_concept_id	No	Provide When Available	Integer	A foreign key to a procedure concept that refers to the code used in the source.	For example, if the procedure is "Anesthesia for procedures on eye; lens surgery" in the source which has a concept code in the vocabulary that is 2100658. The procedure source concept id will be 2100658. **If there is not a mapping for the source code in the standard vocabulary, use concept_id = 0**
modifier_source_value	No	Provide When Available	Varchar	The source code for the modifier as it appears in the source data.	

If a field marked as "Provide when available" for the network requirement is not available at your site, please relay this information to the DCC

1.8.1 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 v3.2, 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, ICD-10-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.

1.9 OBSERVATION

The observation domain captures clinical facts about a patient obtained in the context of examination, questioning or a procedure. The observation domain supports capture of data not represented by other domains such as unstructured measurements. For the PEDSnet CDM version 3.2, 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 ='PCORNet').

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

- Discharge status (Inpatient and outpatient visit types where available)
- DRG (requires special logic - see Note 1 below)
- Tobacco Information (see Note 4)

Use the following table to populate observation_concept_ids for the observations listed above. The vocabulary id 'PCORNet' contains concept specific to PCORNet requirements and standards.

Table 1: Valid Observation concept IDs and Value as concept IDs for PEDSNet v3.2.

Concept Name	Observation concept ID	Vocab ID	Value as concept ID	Concept description	Vocab ID	PCORNet Mapping
Discharge status(See Note 3)	44813951	SNOMED	4161979	Discharged alive		
	44813951	SNOMED	4216643	Expired		

Discharge status						
Discharge status	44813951	SNOMED	44814650	No information	PCORNet	
Discharge status	44813951	SNOMED	44814653	Unknown	PCORNet	
Discharge status	44813951	SNOMED	44814649	Other	PCORNet	
Tobacco	4005823		4005823	Tobacco User		01 = Current user
Tobacco	4005823		45765920	Never used Tobacco		02 = Never
Tobacco	4005823		45765917	Ex-tobacco user		03 = Quit/Former Smoker
Tobacco	4005823		4030580	Non-smoker's second hand smoke syndrome		04 = Passive or environmental exposure
Tobacco	4005823		2000000040			06 = Not asked
Tobacco	4005823		44814650	No information	PCORNet	NI
Tobacco	4005823		44814653	Unknown	PCORNet	OT
Tobacco	4005823		44814649	Other	PCORNet	UN
Tobacco Type	4219336	Multiple Response allowed	4298794	Smoker		01 = Smoked tobacco only
Tobacco Type	4219336	Multiple Response allowed	4224317	Pipe smoking tobacco		01 = Smoked tobacco only
Tobacco Type	4219336	Multiple Response allowed	4282779	Cigarette smoking tobacco		01 = Smoked tobacco only
Tobacco Type	4219336	Multiple Response allowed	4132133	Cigar smoking tobacco		01 = Smoked tobacco only
Tobacco Type	4219336	Multiple Response allowed	4218197	Snuff tobacco		02 = Non-smoked tobacco only
Tobacco Type	4219336	Multiple Response allowed	4219234	Chewing tobacco		02 = Non-smoked tobacco only
Tobacco Type	4219336		45765920	Never used tobacco		04 = None
Tobacco Type	4219336		45765917	Ex tobacco user		04 = None
Tobacco Type	4219336		4030580	Non-smoker's second hand smoke syndrome		04 = Passive or environmental exposure/None
Tobacco Type	4219336		44814650	No information	PCORNet	NI
Tobacco Type	4219336		44814653	Unknown	PCORNet	OT
Tobacco Type	4219336		44814649	Other	PCORNet	UN
Smoking	4275495		42709996	Smokes tobacco daily		01 = Current everyday smoker
Smoking	4275495		2000000039	Occasional tobacco smoker - SNOMED International Code	PEDSNet	02 = current some day smoker
Smoking	4275495		4310250	Ex-smoker		03 = Former smoker
Smoking	4275495		4144272	Never smoked tobacco		04 = Never smoker
Smoking	4275495		4298794	Smoker		05 = Smoker, current status unknown
Smoking	4275495		4141786	Tobacco smoking consumption(status) unknown		06 = Unknown if ever smoked

Smoking	4275495	**USE AS DEFAULT FOR CATEGORY**	4044778	Chain smoker		07 = Heavy tobacco smoker
Smoking	4275495		4209006	Heavy smoker (over 20 per day)		07 = Heavy tobacco smoker
Smoking	4275495	**USE ONLY IF QUANTITY OF CIGARETTES IS KNOWN**	4209585	Moderate smoker (20 or less per day)		08 = Light tobacco smoker
Smoking	4275495		44814650	No information	PCORNet	NI
Smoking	4275495		44814653	Unknown	PCORNet	OT
Smoking	4275495		44814649	Other	PCORNet	UN
Delivery Mode (see note 5)	40760190	SNOMED	4192676	Born by cesarean section	SNOMED	
Delivery Mode	40760190	SNOMED	4212794	Born by elective cesarean section	SNOMED	
Delivery Mode	40760190	SNOMED	4250010	Born by emergency cesarean section	SNOMED	
Delivery Mode	40760190	SNOMED	4216797	Born by normal vaginal delivery	SNOMED	
Delivery Mode	40760190	SNOMED	4217586	Born by forceps delivery	SNOMED	
Delivery Mode	40760190	SNOMED	4236293	Born by ventouse delivery	SNOMED	
Delivery Mode	40760190	SNOMED	4250009	Born by breech delivery	SNOMED	

Note 1: For DRG, use the following logic (must use vocabulary version 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_id = "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_id = "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.

- If your site has **APR-DRGs** please include these in the observation table. We have requested the APR-DRG vocabulary to be incorporated as apart of the OMOP standard vocabulary.
- Please use the following in the qualifier_concept_id:
  - Primary/Principal: concept_id = 4269228
  - Secondary: concept_id = 4093903

  If you are unable to distinguish between primary and secondary DRG type. Please map to the following:

  - Secondary: concept_id = 4093903

  If none are correct, use concept_id = 0.

Note 2: - For each inpatient encounter or in some cases the outpatient encounter, there can be 1 discharge status and 1 or more DRG (May not be 1:1 if patients still admitted (therefore no discharge disposition, discharge details or DRG yet)) - There should **NOT** be discharges without admission.

Note 3: Please provide tobacco information from the primary source of data capture at your site. If tobacco information is available at the visit level, please provide this information. If it is not, sites are welcomed to make a high level assertion about tobacco use and tobacco type information for individuals in the cohort.

Note 4: Below are examples of how the observation table and the fact relationship table would be populated for tobacco, smoking and tobacco type scenarios. In the case where tobacco information is recorded at a visit but there is missing information for tobacco, smoking or tobacco type please assert. The PEDSnet standard relationship concept id for linking tobacco items will be *0*. This concept id was chosen as there was not a specific concept id that exists in the standard vocabulary that adequately defined an appropriate relationship for linking the tobacco items.

Example 1:

Patient 1 smokes 5 cigarettes per day and does not use non-smoked tobacco

Observation table:

Observation ID	Person ID	Observation concept id	Value as concept id
0001	1	4005823	4005823
0002	1	4219336	4282779
0003	1	4275495	4209585

Fact relationship:

Domain_concept_id_1	Fact_id_1	Domain_concept_id_2	Fact_id_2	relationship_concept_id
27	0001	27	0002	0
27	0001	27	0003	0

Example 2: Patient 2 smokes 25-40 cigarettes per day and also chews tobacco

Observation table:

Observation ID	Person ID	Observation concept id	Value as concept id
0004	2	4005823	4005823
0005	2	4219336	4282779
0006	2	4219336	4219234
0007	2	4275495	4209006

Fact relationship:

Domain_concept_id_1	Fact_id_1	Domain_concept_id_2	Fact_id_2	relationship_concept_id
27	0004	27	0005	0
27	0004	27	0006	0
27	0004	27	0007	0

For more examples, or if you have a specific scenario that you have a question about, please contact the DCC.

Note 5: - For delivery mode, if you are unable to discern between elective (concept_id = 4212794) and emergency (concept_id = 4250010) cesarean, please default to the born by cesarean section (concept_id = 4192676).

Field	NOT Null Constraint	Network Requirement	Data Type	Description	PEDSnet Conventions
observation_id	Yes	Yes	Integer	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	Yes	Integer	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	Yes	Integer	A foreign key to the standard observation concept identifier in the Vocabulary.	Lab results and vitals are not stored in this table in V5 but are stored in the Measurement table.
observation_date	Yes	Yes	Date	The date of the observation.	No date shifting. Full date and time. If there is no time associated with the date assert midnight.
observation_datetime	No	Provide When Available	Datetime	The time of the observation.	No date shifting. Full date and time. If there is no time associated with the date assert midnight.

observation_type_concept_id	Yes	Yes	Integer	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 CDMv5). Predefined value set (valid concept_ids found in CONCEPT table where vocabulary_id ='Observation Type') select * from concept where vocabulary_id = 'Observation Type' yields 11 valid concept_ids. FOR PEDSnet CDM v3.2, 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 concept_id = 44814721
value_as_number	No (see convention)	Provide When Available	Float	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 PCORNet where all three value_as_* fields are NULL.
value_as_string	No (see convention)	Provide When Available	Varchar	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 PCORNet where all three value_as_* fields are NULL.
value_as_concept_id	No (see convention)	Provide When Available	Integer	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 PCORNet where all three value_as_* fields are NULL.
qualifier_concept_id	No	Provide When Available	Integer	A foreign key to standard concept identifier for a qualifier (e.g severity of drug-drug interaction alert)	Predefined value set (valid concept_ids found in CONCEPT table where domain_id='Observation' and concept_class_id ='Qualifier Value') select * from concept where domain_id='Observation' and concept_class_id ='Qualifier Value' yields 10496 valid concept_ids. For **DRG VALUES**, please use the following: • Primary/Principal: concept_id = 4269228 • Secondary: concept_id = 4093903 If you are unable to distinguish between primary and secondary DRG type. Please map to the following: • Secondary: concept_id = 4093903 If none are correct, use concept_id = 0.
unit_concept_id	No	Provide When Available	Integer	A foreign key to a standard concept identifier of observation units in the Vocabulary.	Please include valid concept ids (consistent with OMOP CDMv5). Predefined value set (valid concept_ids found in CONCEPT table where domain_id='Unit' and vocabulary_id ='UCUM') select * from concept where domain_id='Unit' and vocabulary_id ='UCUM' yields 971 valid concept_ids. If none are correct, use concept_id = 0.

provider_id	No	Provide When Available	Integer	A foreign key to the provider in the provider table who was responsible for making the observation.	
visit_occurrence_id	No	Provide When Available	Integer	A foreign key to the visit in the visit table during which the observation was recorded.	
observation_source_value	No	Provide When Available	Varchar	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.	
observation_source_concept_id	No	Provide When Available	Integer	A foreign key to a concept that refers to the code used in the source.	**If there is not a mapping for the source code in the standard vocabulary, use concept_id = 0**
unit_source_value	No	Provide When Available	Integer	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.	
qualifier_source_value	No	Provide When Available	Varchar	The source value associated with a qualifier to characterize the observation	For DRG Values, please populate information pertaining to "Primary" or "Secondary" DRG Status as it corresponds to the concept id value at your site.

If a field marked as "Provide when available" for the network requirement is not available at your site, please relay this information to the DCC

1.9.1 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 v3.2, 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.
- Observations obtained using standardized methods (e.g. laboratory assays) that produce discrete results are recorded by preference in the MEASUREMENT table.

1.10 OBSERVATION PERIOD

The observation period domain 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 a clinical fact represented in the PEDSNet version 3.2 data model. 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. This logic can be found here

Field	NOT Null Constraint	Network Requirement	Data Type	Description	PEDSnet Conventions
Observation_period_id	Yes	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	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	Yes	Date	The start date of the observation period for which data are available from the data source	Use the earliest clinical fact date available for this patient. No date shifting.
Observation_period_end_date	Yes	Yes	Date	The end date of the observation period for which data are available from the source.	Use the latest clinical fact 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.
Observation_period_start_time	Yes	Yes	Datetime	The start date of the observation period for which data are available from the data source	Use the earliest clinical fact time available for this patient. No date shifting. Full date and time. **If there is no time associated with the date assert midnight for the start time**
Observation_period_end_time	Yes	Yes	Datetime	The end date of the observation period for which data are available from the source.	Use the latest clinical fact time 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. Full date and time. **If there is no time associated with the date assert 11:59:59 pm for the end time**
period_type_concept_id	Yes	Yes	Integer	A unique identifier for each observation period.	
person_id	Yes	Yes	Integer	A foreign key identifier to the person for whom the observation period is defined. The demographic details of that person are stored in the person table.	

If a field marked as "Provide when available" for the network requirement is not available at your site, please relay this information to the DCC

1.10.1 Additional Notes

- Because the 1/1/2009 date limitation for "active patients" is not used to limit visit_occurrence, the start_date of an observation period for an active PEDSnet patient may be prior to 1/1/ 2009.

1.11 DRUG EXPOSURE

The drug exposure domain captures any biochemical substance that is introduced in any way to a patient. This can be evidence of prescribed, over the counter, administered (IV, intramuscular, etc), immunizations or dispensed medications. These events could be linked to procedures or encounters where they are administered or associated as a result of the encounter.

EHRs may store medications in different vocabularies (GPI,NDC etc).

Exclusions:

1. Cancelled Medication Orders
2. Missed Medication administrations

Note 1: The `effective\_drug\_dose` is the dose basis.(E.g. 45 mg/kg/dose). This is the discrete dose value from the source data if available. If the discrete dose value is **not** available from the source data, then compute the dose basis by looking for a weight observation **+/- 60 days of the date of the medication**. (E.g. Total Amount/**(divided by)**Weight) (Dose per kg)

The dose_unit_concept_id is the unit of the effective dose.

Please use the following logic to populate the effective_dose and dose unit based on what is available in your source system:

Site Information	Effective Drug Dose	Dose Unit Concept Id	Dose Unit Source Value
Pre-calculated effective dose available (E.g. 90 mg/kg)	90	Corresponding concept for unit (E.g. mg/kg = 9562)	mg/kg
Site is able to compute effective dose (E.g. Dose 500 mg and Available Weight +/- 60 days is 54.43 kg)	9.18	Corresponding concept for unit (E.g. mg/kg = 9562)	mg
Site is not able to compute effective dose(E.g. Site Only has dose (E.g. 450 mg))	450	Corresponding Concept for unit (E.g. mg = 8576)	mg
No discrete dosing information		0	

Note 2: The quantity is the actual dose given. (E.g. 450 mg for 10 kg patient) Extract numbers as much as possible , full value should be a part of the xml sig field.

Note 3: For dispensing records, compute the dose basis by looking for a weight observation +/- 60 days of the dispensed date.

Note 4: For the sig, encode the value using XML.

- Element 1: Actual SIG from source data
- Element 2: Raw "Supply/Quantity" (Examples: "1 bottle" "10 ml Bottle" "1 pack"
- Element 3: Refills

```
<XML>
<SIG>1/2 capful in 4 oz clear liquid</SIG>
<QUANTITY>1 jar</QUANTITY>
<REFILLS>2</REFILLS>
</XML>
```

Note 5: If there are multiple RxNorm mappings associate with a mapping, choose the mapping in the following order and stop when you find your first match.

1. BPCK (Branded Pack)
2. GPCK (Clinical Pack)
3. SBD (Branded Drug, Quant Branded Drug)
4. SCD (Clinical Drug, Quant Clinical Drug)
5. SBDF (Branded Drug Form)
6. SCDF (Clinical Drug Form)
7. MIN (Ingredient)
8. SBDC
9. SCDC
10. PIN (Ingredient)
11. IN (Ingredient)

Note 6: Please use the following table as a guide to determine how to populate the `drug\_source\_value` , `drug\_source\_concept\_id` and `drug\_concept\_id` for Drug Exposure Values