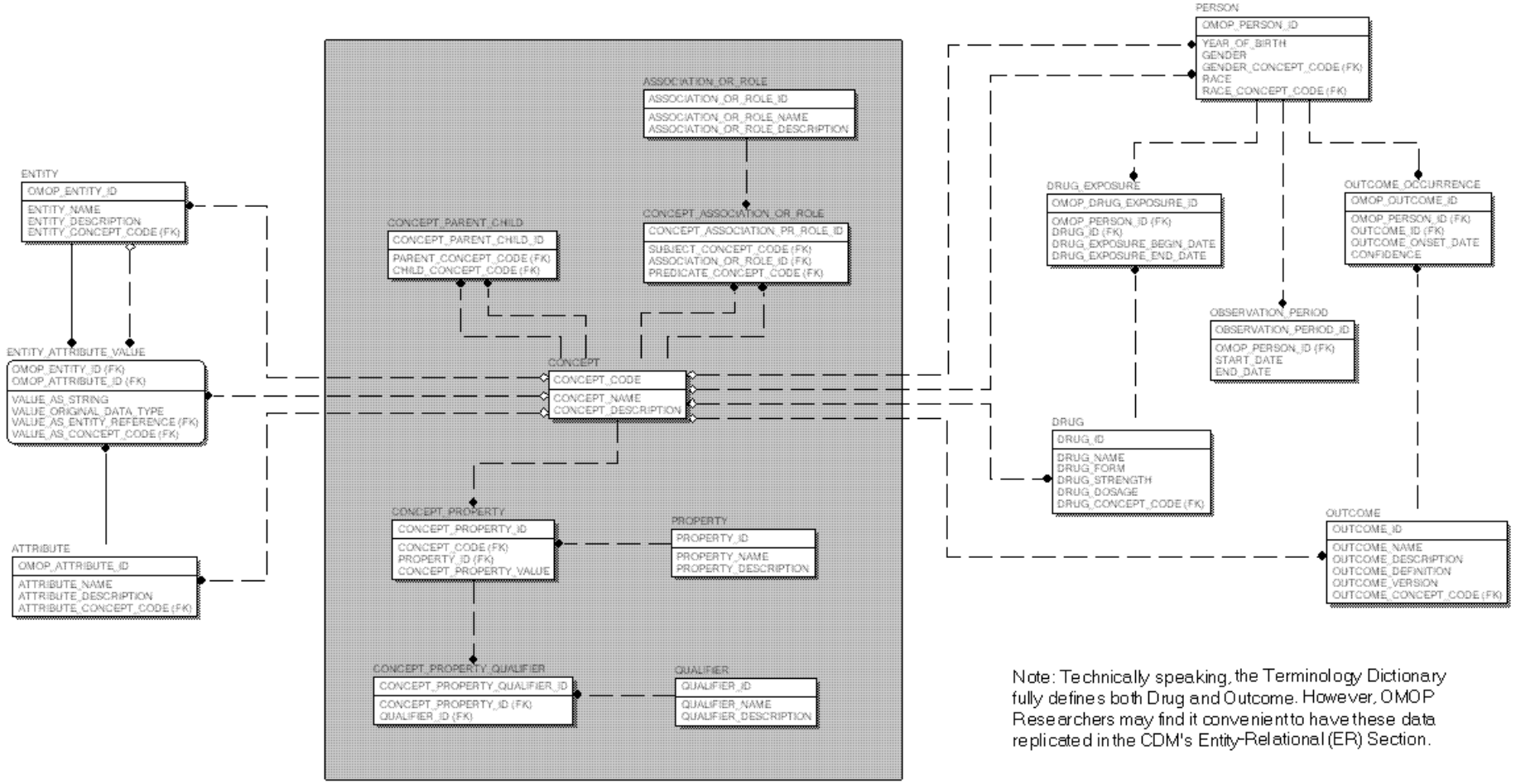
Common Data Model History and Progress

Clair Blacketer

Observational Medical Outcomes Partnership

2009

2014



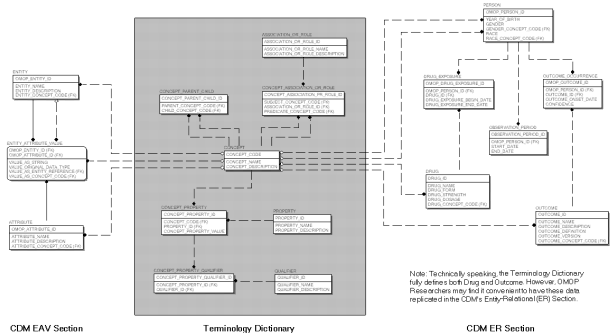
Note: Technically speaking, the Terminology Dictionary fully defines both Drug and Outcome. However, OMOP Researchers may find it convenient to have these data replicated in the CDM's Entity-Relational (ER) Section.

CDM EAV Section

Terminology Dictionary

CDM ER Section

Observational **M**edical **O**utcomes **P**artnership

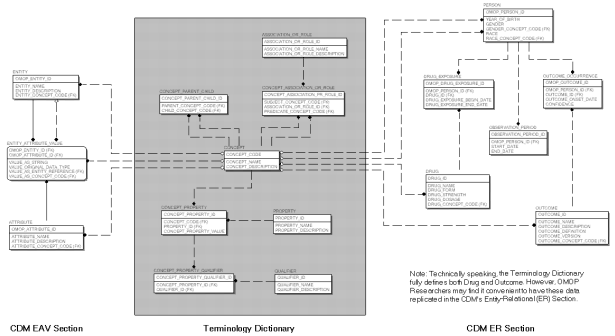


CDM v1

2009

2014

Observational Medical Outcomes Partnership

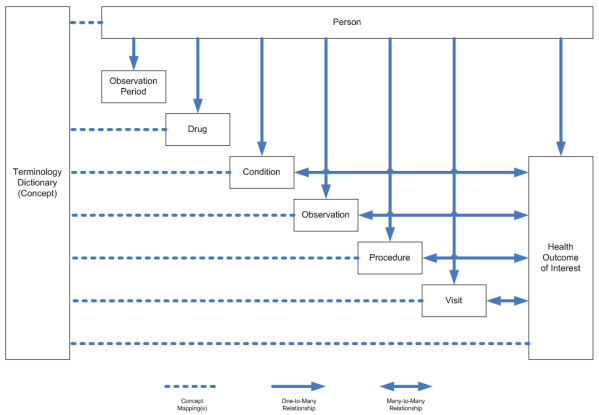


CDM v1

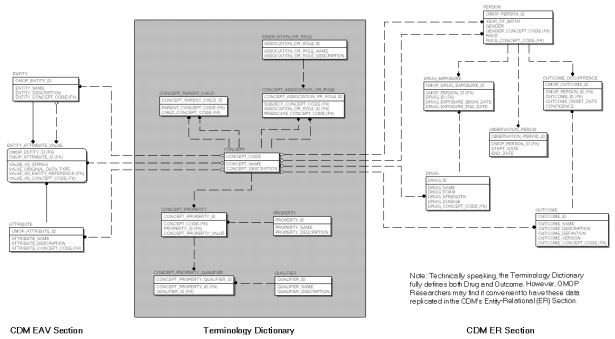
CDM v2

2009

2014



Observational Medical Outcomes Partnership



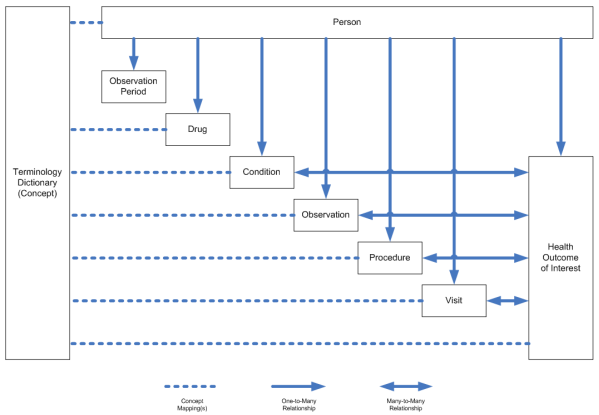
CDM v1

CDM v3

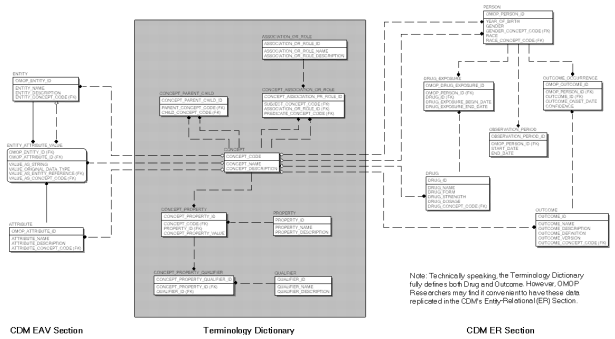
CDM v2

2009

2014



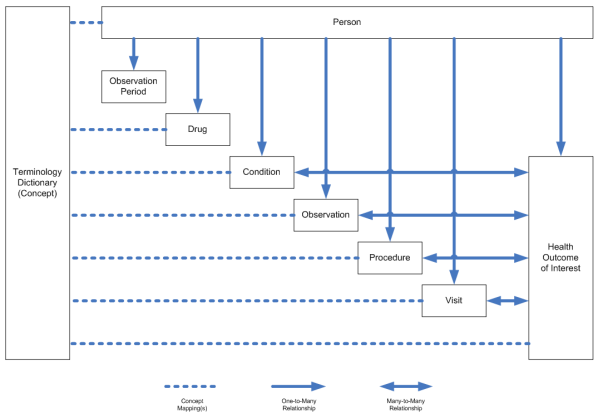
Observed Medical Outcomes Partnership



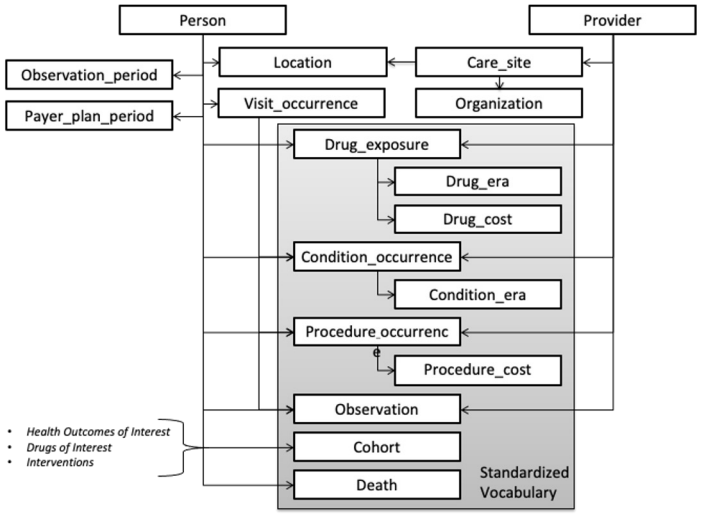
CDM v1

CDM v3

CDM v2



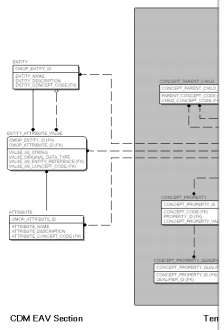
CDM v4



2009

2014

Observational Medical Outcomes Partnership



CD

2009

2014

Standardized clinical data

Person

Observation_period

Specimen

Death

Visit_occurrence

Procedure_occurrence

Drug_exposure

Device_exposure

Condition_occurrence

Measurement

Observation

Fact_relationship

Standardized health system data

Location

Care_site

Provider

Standardized health economics

Payer_plan_period

Cost

Standardized derived elements

Cohort

Cohort_attribute

Condition_era

Drug_era

Dose_era

Standardized meta-data

CDM_source

Standardized vocabularies

Concept

Vocabulary

Domain

Concept_class

Concept_relationship

Relationship

Concept_synonym

Concept_ancestor

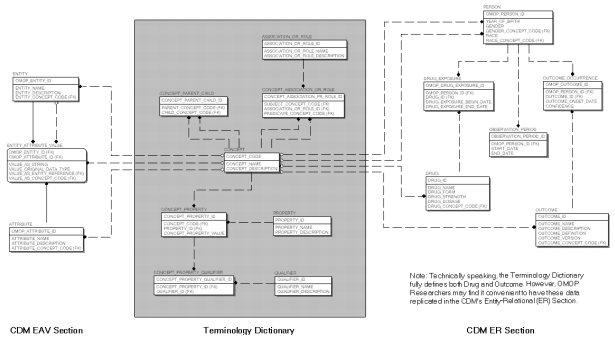
Source_to_concept_map

Drug_strength

Cohort_definition

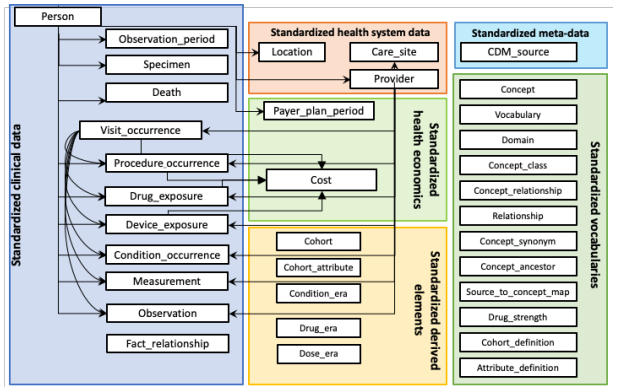
Attribute_definition

Observational Medical Outcomes Partnership



CDM v1

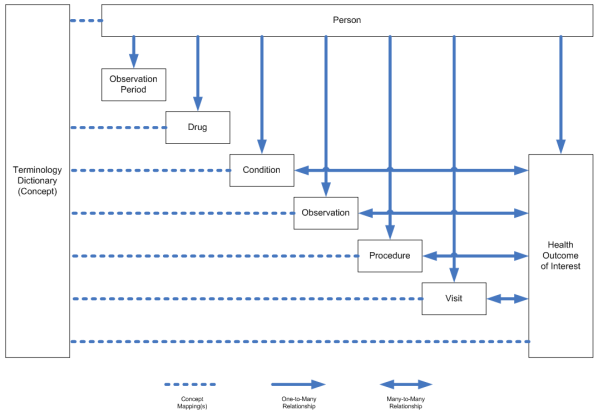
CDM v3



CDM v5.0

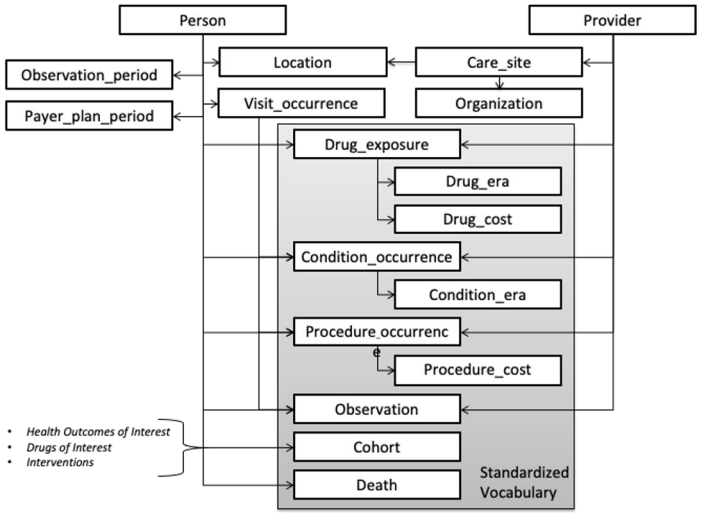
CDM v2

2009

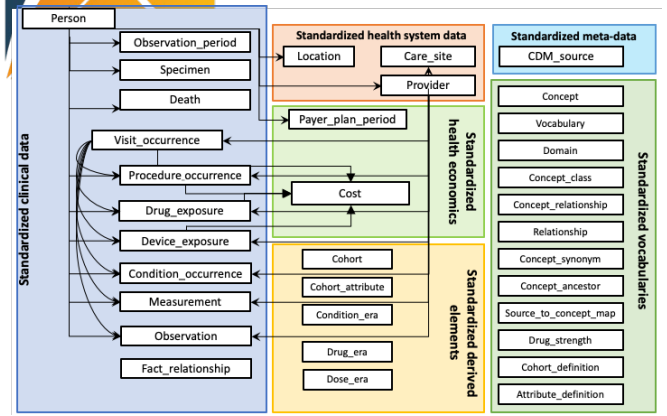


CDM v4

2014

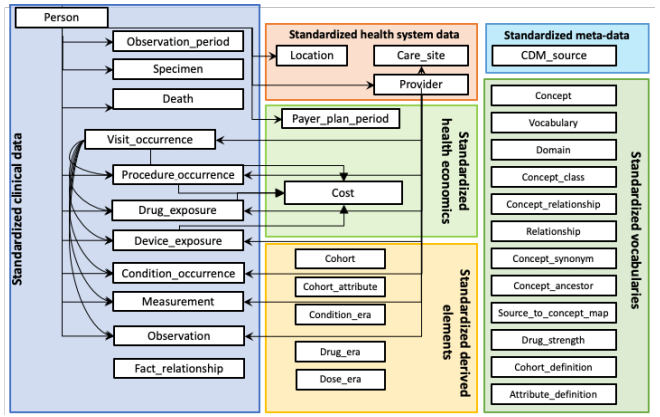


OHDSI



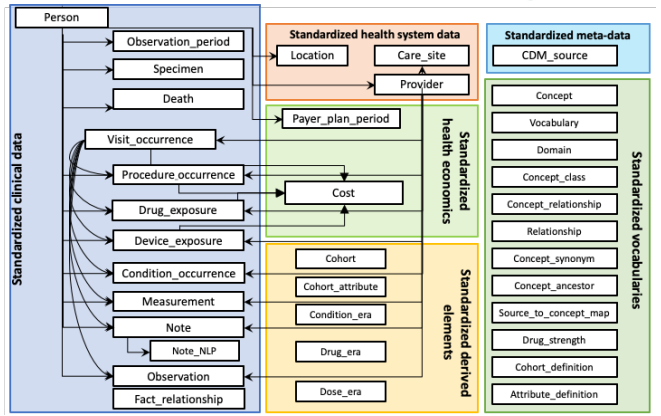
CDM v5.0.1

2016



CDM v5.1

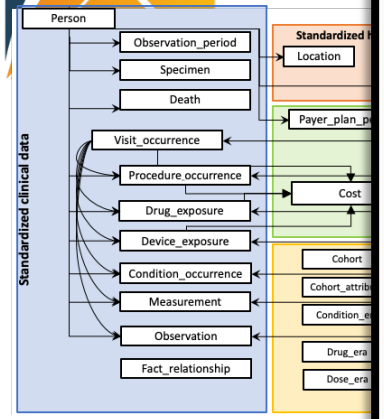
2017



CDM v5.2

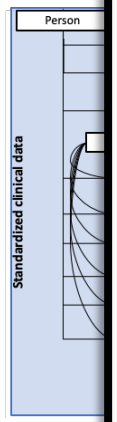
2017

2020



CDM v5.0.

2016



2020

Common Data Model

- License
- Background
- The Role of the Common Data Model
  - Design Principles
  - Details of the Model
  - Data Model Conventions
  - Changes from Version 4.0
  - Glossary of Terms
- Standardized Vocabularies
  - CONCEPT
  - VOCABULARY
  - DOMAIN
  - CONCEPT_CLASS
  - CONCEPT_RELATIONSHIP
  - RELATIONSHIP
  - CONCEPT_SYNONYM
  - CONCEPT_ANCESTOR
  - SOURCE_TO_CONCEPT_MAP
  - DRUG_STRENGTH
  - COHORT_DEFINITION
  - ATTRIBUTE_DEFINITION
- Standardized meta-data
  - CDM_SOURCE
- **Standardized Clinical Data Tables**
  - PERSON (V5.1 and up)
  - OBSERVATION_PERIOD
  - SPECIMEN (V5.1 and up)
  - DEATH (V5.1 and up)
  - VISIT_OCCURRENCE (V5.1 and up)
  - PROCEDURE_OCCURRENCE (V5.1 and up)
  - DRUG_EXPOSURE (V5.1 and up)
  - DEVICE_EXPOSURE (V5.1 and up)
  - CONDITION_OCCURRENCE (V5.1 and up)
  - MEASUREMENT (V5.1 and up)
  - NOTE
  - NOTE_NLP (V5.2)

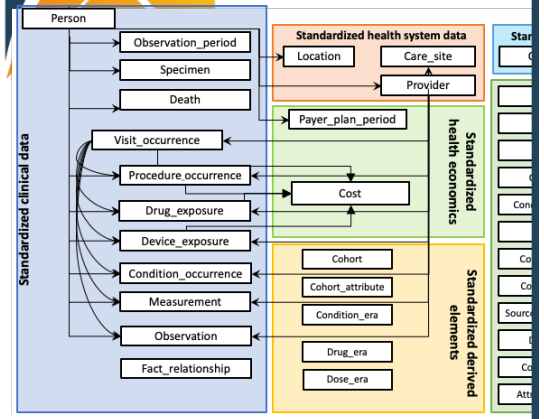
PERSON table

THIS IS OUTDATED. All documentation is now on the github wiki. Please refer there or to the CDM working group for more information

This table changed in version 5.1 of the OMOP CDM. The name of the field time_of_birth was changed to birth_datetime.

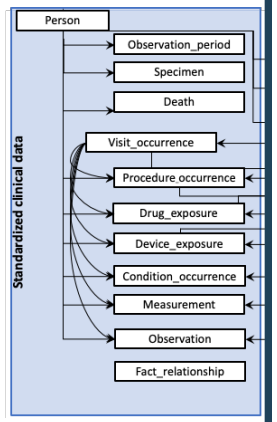
The Person Domain contains records that uniquely identify each patient in the source data who is time at-risk to have clinical observations recorded within the source systems.

Field	Required	Type	Description
person_id	Yes	integer	A unique identifier for each person.
gender_concept_id	Yes	integer	A foreign key that refers to an identifier in the CONCEPT table for the unique gender of the person.
year_of_birth	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.
month_of_birth	No	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.
day_of_birth	No	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.
birth_datetime	No	datetime	The date and time of birth of the person.
race_concept_id	Yes	integer	A foreign key that refers to an identifier in the CONCEPT table for the unique race of the person.
ethnicity_concept_id	Yes	integer	A foreign key that refers to the standard concept identifier in the Standardized Vocabularies for the ethnicity of the person.
location_id	No	integer	A foreign key to the place of residency for the



CDM v5.0.1

2016



Home

clairblacketer edited this page on Jul 12, 2017 · 26 revisions

OMOP Common Data Model v5.1.1 Specifications

Authors: Christian Reich, Patrick Ryan, Rimma Belenkaya, Karthik Natarajan, Clair Blacketer

12 July 2017

Welcome to the Common Data Model wiki! This wiki houses all of the documentation for the late
changes added with each release. You can find a pdf added to each release with a historical ver
the time of the release. You can navigate the pages using the table of contents below or the link

Table of Contents

License

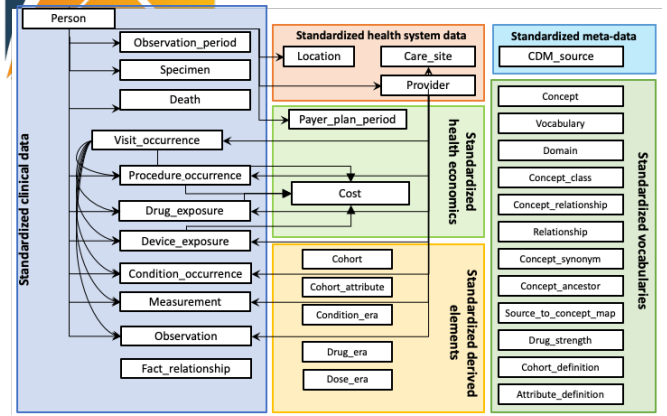
Background
The Role of the Common Data Model
Design Principles
Data Model Conventions

Glossary of Terms

Standardized Vocabularies
CONCEPT
VOCABULARY
DOMAIN
CONCEPT_CLASS

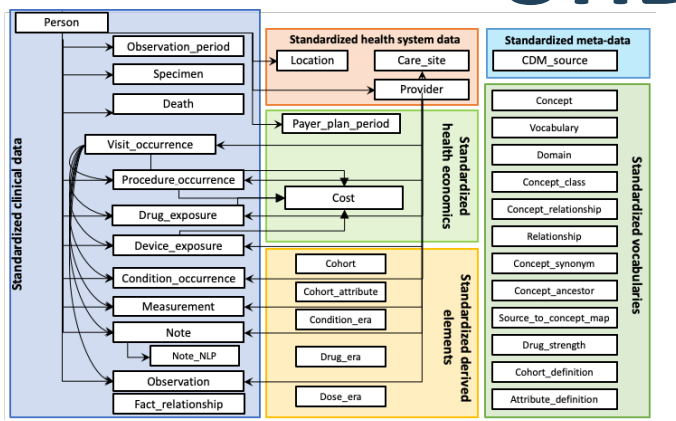
2024

OHDSI



CDM v5.0.1

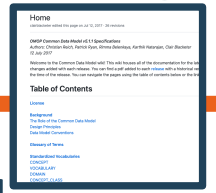
2016



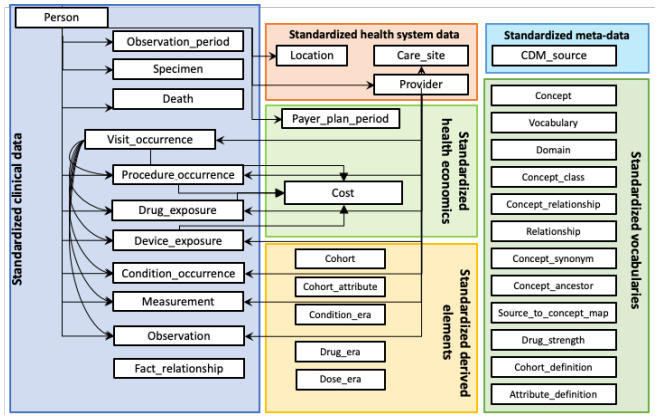
2017

CDM v5.2

2017

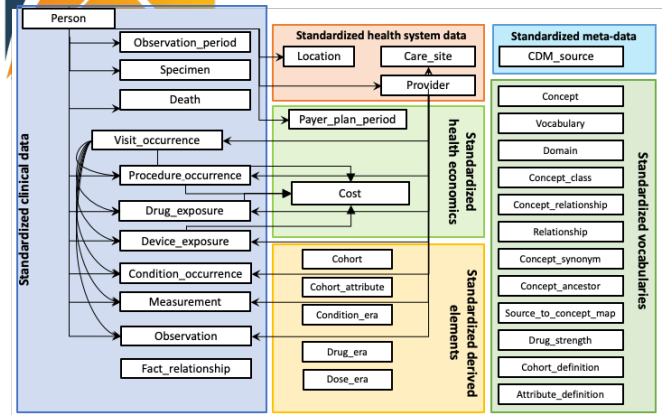


CDM v5.1



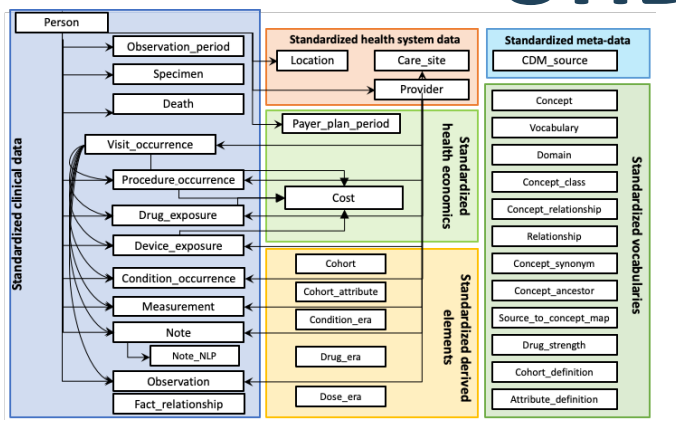
2024

OHDSI



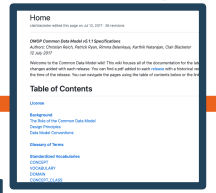
CDM v5.0.1

2016



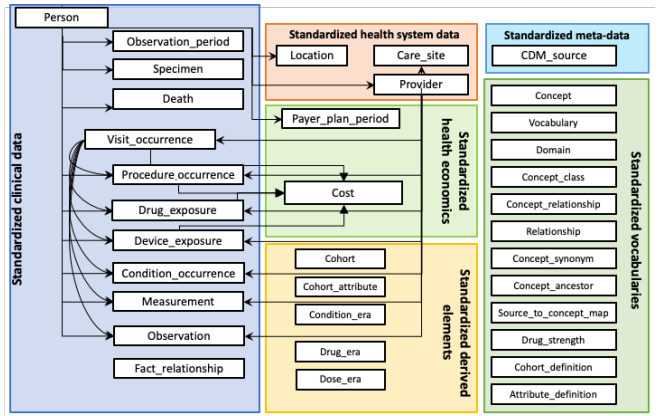
CDM v5.2

2017



CDM v5.1

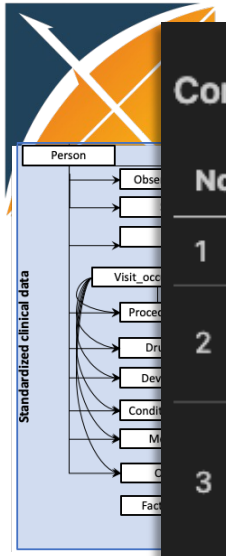
2017



CDM v5.3

2018

2024



CDM

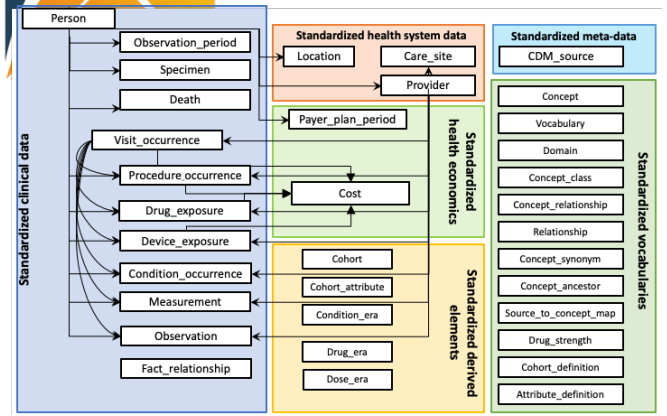
20

024

Conventions

No.	Convention Description
1	All tables representing patient-related Domains have a foreign-key reference to the person_id field in the PERSON table.
2	Each person record has associated demographic attributes which are assumed to be constant for the patient throughout the course of their periods of observation. For example, the location or gender is expected to have a unique value per person, even though in life these data may change over time.
3	The GENDER_CONCEPT_ID should store what is believed to be the biological or sex assigned at birth. If the data set does have gender identification information, this should be stored in the OBSERVATION table (using the gender concepts 8532-Female or 8507-Male in OBSERVATION_CONCEPT_ID)THEMIS issue #32.
4	If we do not know the month or day of birth, we do not guess. A person can exist without a month or day of birth. If a person lacks a birth year that person should be dropped(THEMIS issue #30).
5	Living patients should not have a value in PERSON.DEATH_DATETIME, nor should they have any records relating to death either in the CONDITION_OCCURRENCE or OBSERVATION tables
6	Only one death date per individual can be used. If a patient has clinical activity (e.g. prescriptions filled, labs performed, etc) more than 60+ days after death you may want to drop the death record as it may have been falsely reported. If multiple records of death exist on multiple days you may select the death that you deem most reliable (e.g. death at discharge) or select the latest death date.
7	If multiple death records occur, the date and the person have to be the same, but the cause can be different. Can be reported by different sources as well.
8	If PERSON.DEATH_DATETIME cannot be precisely determined from the data, the best approximation should be used.
9	The DEATH_DATETIME in the PERSON table should not be used as the way to find all deaths • `select * from PERSON where death_datetime is not null` should not be the practice • Rather, deaths should be found through the OBSERVATION table and the PERSON table is only used to determine which death date should be used in analysis

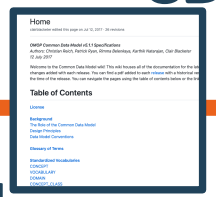
OHDSI



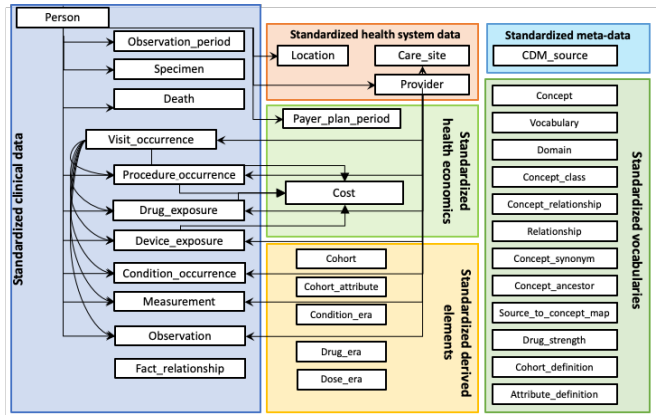
CDM v5.0.1

2016

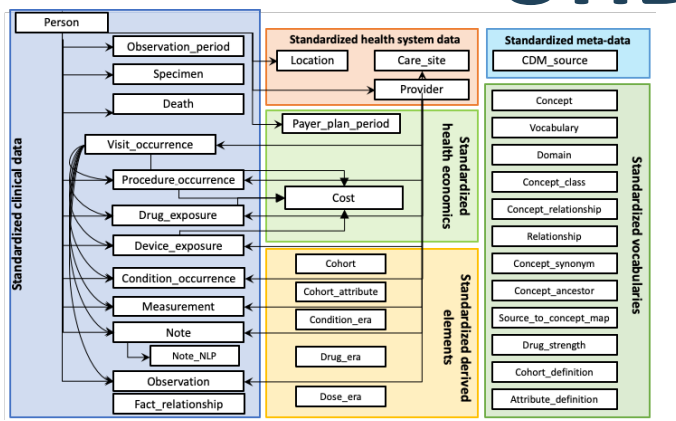
2017



CDM v5.1

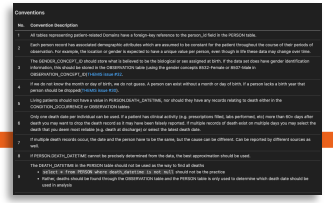


2017

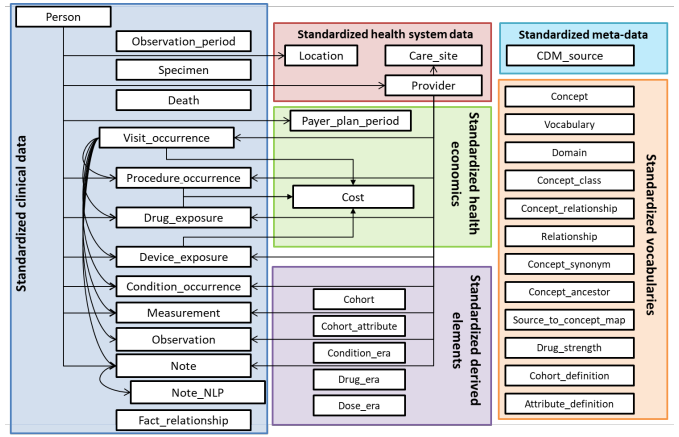


CDM v5.2

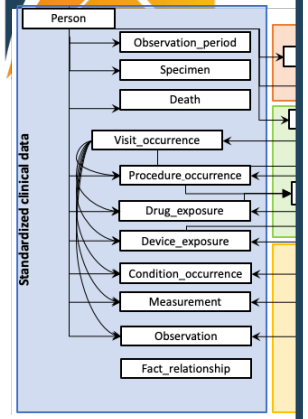
2018



CDM v5.3



2024



CDM v5

2016

OMOP Common Data Model | Background | Conventions | CDM Versions

OMOP Common Data Model

The Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) is an op
standardize the structure and content of observational data and to enable efficient analyses th
component of the OMOP CDM is the OHDSI standardized vocabularies. The OHDSI vocabularie
terms to be used across the various clinical domains of the OMOP common data model and en
knowledge base when constructing exposure and outcome phenotypes and other features wit
estimation, and patient-level prediction studies.

This website is meant to serve as a resource describing the specification of the available versio
structure of the model itself and the agreed upon conventions for each table and field as decid
tables are part of the model and, as such, are detailed here. To download the vocabulary itself,
information about the OHDSI suite of tools designed to implement best practices in characteri
patient-level prediction, please visit https://ohdsi.github.io/Hades/.

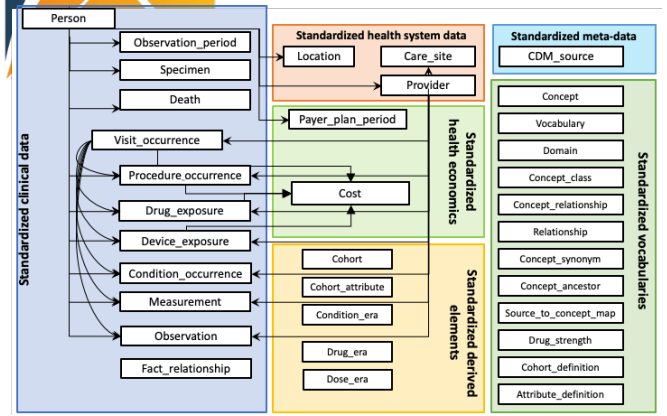
Current CDM Version

The current CDM version is CDM v5.4, depicted below. This CDM version was developed over th
sent via our issues page. The list of proposed changes was then shared with the community in
OHDSI Community calls, discussions with the OHDSI Steering Committee, and discussions with
changes were then delivered to the Community through a new R package designed to dynamic
supported SQL dialects.

- Link to DDLs for CDM v5.4
- Link to ReadMe for instructions on how to use the R package

2024

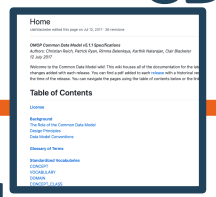
OHDSI



CDM v5.0.1

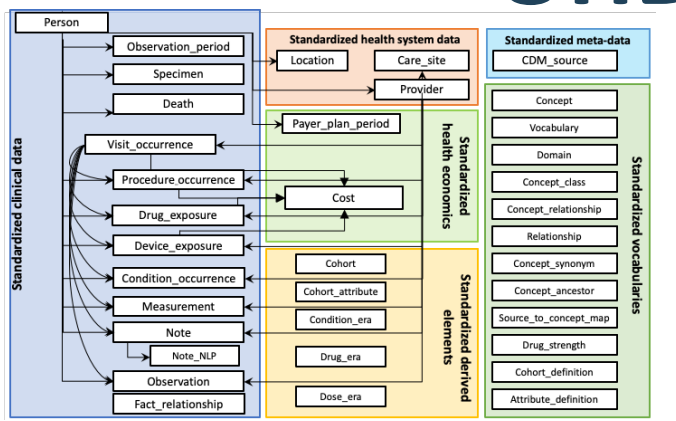
2016

2017



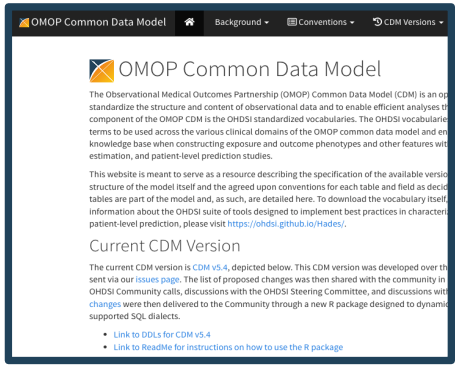
CDM v5.1

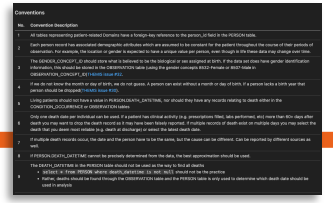
2017



CDM v5.2

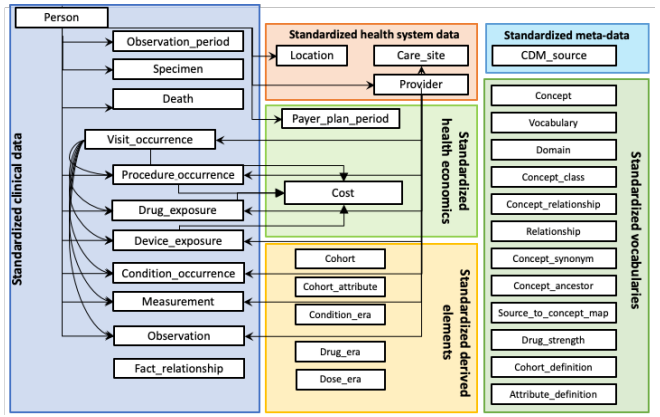
2018

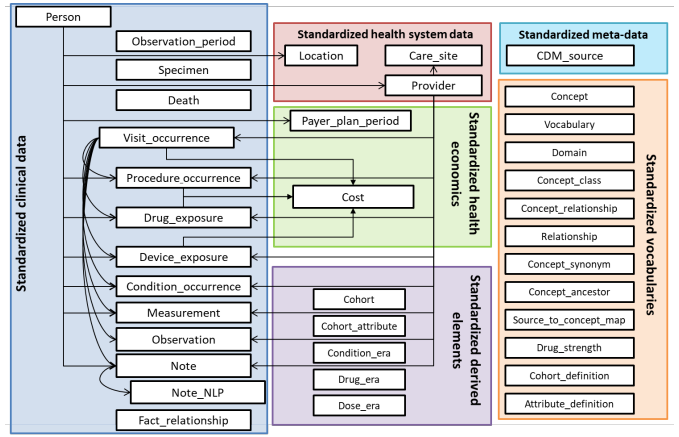


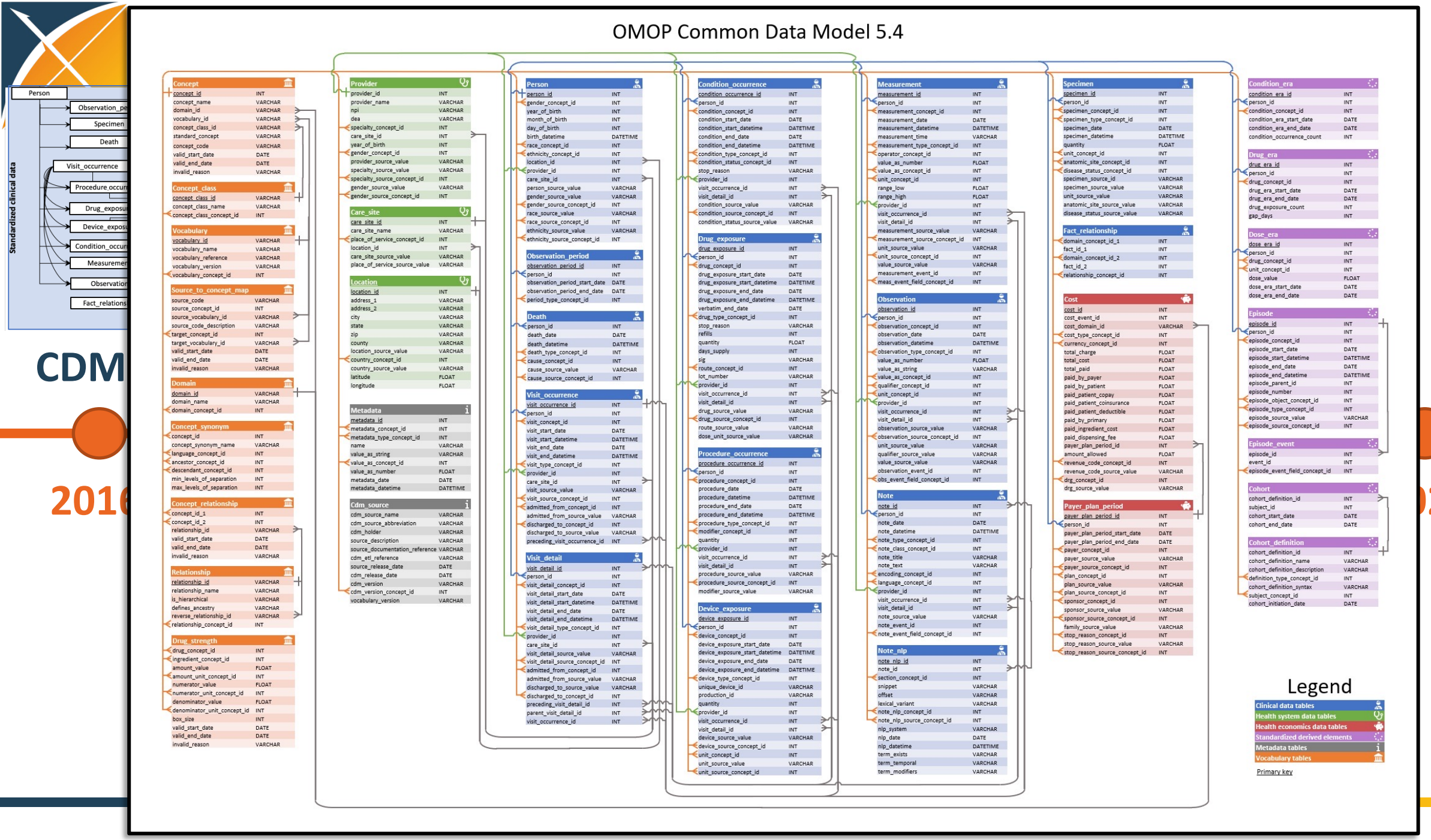


CDM v5.3

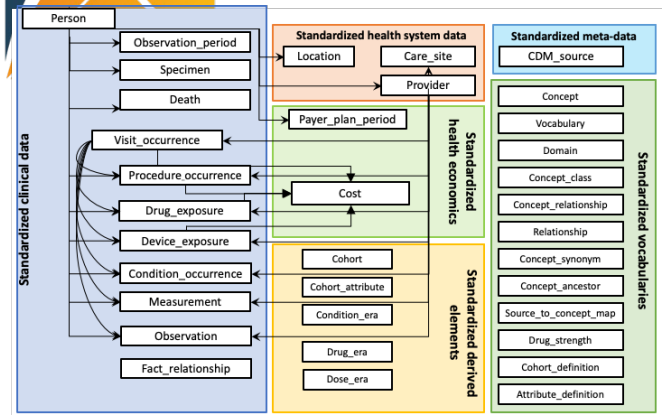
2024





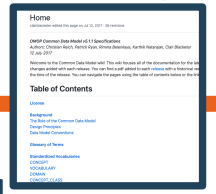

OMOP Common Data Model 5.4
Person
Observation_pe
Specimen
Death
Visit_occurrence
Procedure_occur
Drug_exposu
Device_exposu
Condition_occur
Measuremer
Observatior
Fact_relations
Standardized clinical data
CDM
2016
024
Concept
concept_id INT
concept_name VARCHAR
domain_id VARCHAR
vocabulary_id VARCHAR
concept_class_id VARCHAR
standard_concept VARCHAR
concept_code VARCHAR
valid_start_date DATE
valid_end_date DATE
invalid_reason VARCHAR
Concept_class
concept_class_id VARCHAR
concept_class_name VARCHAR
concept_class_concept_id INT
Vocabulary
vocabulary_id VARCHAR
vocabulary_name VARCHAR
vocabulary_reference VARCHAR
vocabulary_version VARCHAR
vocabulary_concept_id INT
Source_to_concept_map
source_code VARCHAR
source_concept_id INT
source_vocabulary_id VARCHAR
source_code_description VARCHAR
target_concept_id INT
target_vocabulary_id VARCHAR
valid_start_date DATE
valid_end_date DATE
invalid_reason VARCHAR
Domain
domain_id VARCHAR
domain_name VARCHAR
domain_concept_id INT
Concept_synonym
concept_id INT
concept_synonym_name VARCHAR
language_concept_id INT
ancestor_concept_id INT
descendant_concept_id INT
min_levels_of_separation INT
max_levels_of_separation INT
Concept_relationship
concept_id_1 INT
concept_id_2 INT
relationship_id VARCHAR
valid_start_date DATE
valid_end_date DATE
invalid_reason VARCHAR
Relationship
relationship_id VARCHAR
relationship_name VARCHAR
is_hierarchical VARCHAR
defines_ancestry VARCHAR
reverse_relationship_id VARCHAR
relationship_concept_id INT
Drug_strength
drug_concept_id INT
ingredient_concept_id INT
amount_value FLOAT
amount_unit_concept_id INT
numerator_value FLOAT
numerator_unit_concept_id INT
denominator_value FLOAT
denominator_unit_concept_id INT
box_size INT
valid_start_date DATE
valid_end_date DATE
invalid_reason VARCHAR
Provider
provider_id INT
provider_name VARCHAR
npi VARCHAR
dea VARCHAR
specialty_concept_id INT
care_site_id INT
year_of_birth INT
gender_concept_id INT
provider_source_value VARCHAR
specialty_source_value VARCHAR
specialty_source_concept_id INT
gender_source_value VARCHAR
gender_source_concept_id INT
Care_site
care_site_id INT
care_site_name VARCHAR
place_of_service_concept_id INT
location_id INT
care_site_source_value VARCHAR
place_of_service_source_value VARCHAR
Location
location_id INT
address_1 VARCHAR
address_2 VARCHAR
city VARCHAR
state VARCHAR
zip VARCHAR
county VARCHAR
location_source_value VARCHAR
country_concept_id INT
country_source_value VARCHAR
latitude FLOAT
longitude FLOAT
Metadata
metadata_id INT
metadata_concept_id INT
metadata_type_concept_id INT
name VARCHAR
value_as_string VARCHAR
value_as_concept_id INT
value_as_number FLOAT
metadata_date DATE
metadata_datetime DATETIME
Cdm_source
cdm_source_name VARCHAR
cdm_source_abbreviation VARCHAR
cdm_holder VARCHAR
source_description VARCHAR
source_documentation_reference VARCHAR
cdm_etl_reference VARCHAR
source_release_date DATE
cdm_release_date DATE
cdm_version VARCHAR
cdm_version_concept_id INT
vocabulary_version VARCHAR
Person
person_id INT
gender_concept_id INT
year_of_birth INT
month_of_birth INT
day_of_birth INT
birth_datetime DATETIME
race_concept_id INT
ethnicity_concept_id INT
location_id INT
provider_id INT
care_site_id INT
person_source_value VARCHAR
gender_source_value VARCHAR
gender_source_concept_id INT
race_source_value VARCHAR
race_source_concept_id INT
ethnicity_source_value VARCHAR
ethnicity_source_concept_id INT
Observation_period
observation_period_id INT
person_id INT
observation_period_start_date DATE
observation_period_end_date DATE
period_type_concept_id INT
Death
person_id INT
death_date DATE
death_datetime DATETIME
death_type_concept_id INT
cause_concept_id INT
cause_source_value VARCHAR
cause_source_concept_id INT
Visit_occurrence
visit_occurrence_id INT
person_id INT
visit_concept_id INT
visit_start_date DATE
visit_start_datetime DATETIME
visit_end_date DATE
visit_end_datetime DATETIME
visit_type_concept_id INT
provider_id INT
care_site_id INT
visit_source_value VARCHAR
visit_source_concept_id INT
admitted_from_concept_id INT
admitted_from_source_value VARCHAR
discharged_to_concept_id INT
discharged_to_source_value VARCHAR
preceding_visit_occurrence_id INT
Visit_detail
visit_detail_id INT
person_id INT
visit_detail_concept_id INT
visit_detail_start_date DATE
visit_detail_start_datetime DATETIME
visit_detail_end_date DATE
visit_detail_end_datetime DATETIME
visit_detail_type_concept_id INT
provider_id INT
care_site_id INT
visit_detail_source_value VARCHAR
visit_detail_source_concept_id INT
admitted_from_concept_id INT
admitted_from_source_value VARCHAR
discharged_to_source_value VARCHAR
discharged_to_concept_id INT
preceding_visit_detail_id INT
parent_visit_detail_id INT
visit_occurrence_id INT
Condition_occurrence
condition_occurrence_id INT
person_id INT
condition_concept_id INT
condition_start_date DATE
condition_start_datetime DATETIME
condition_end_date DATE
condition_end_datetime DATETIME
condition_type_concept_id INT
condition_status_concept_id INT
stop_reason VARCHAR
provider_id INT
visit_occurrence_id INT
visit_detail_id INT
condition_source_value VARCHAR
condition_source_concept_id INT
condition_status_source_value VARCHAR
Drug_exposure
drug_exposure_id INT
person_id INT
drug_concept_id INT
drug_exposure_start_date DATE
drug_exposure_start_datetime DATETIME
drug_exposure_end_date DATE
drug_exposure_end_datetime DATETIME
verbatim_end_date DATE
drug_type_concept_id INT
stop_reason VARCHAR
refills INT
quantity FLOAT
days_supply INT
sig VARCHAR
route_concept_id INT
lot_number VARCHAR
provider_id INT
visit_occurrence_id INT
visit_detail_id INT
drug_source_value VARCHAR
drug_source_concept_id INT
route_source_value VARCHAR
dose_unit_source_value VARCHAR
Procedure_occurrence
procedure_occurrence_id INT
person_id INT
procedure_concept_id INT
procedure_date DATE
procedure_datetime DATETIME
procedure_end_date DATE
procedure_end_datetime DATETIME
procedure_type_concept_id INT
modifier_concept_id INT
quantity INT
provider_id INT
visit_occurrence_id INT
visit_detail_id INT
procedure_source_value VARCHAR
procedure_source_concept_id INT
modifier_source_value VARCHAR
Device_exposure
device_exposure_id INT
person_id INT
device_concept_id INT
device_exposure_start_date DATE
device_exposure_start_datetime DATETIME
device_exposure_end_date DATE
device_exposure_end_datetime DATETIME
device_type_concept_id INT
unique_device_id VARCHAR
production_id VARCHAR
quantity INT
provider_id INT
visit_occurrence_id INT
visit_detail_id INT
device_source_value VARCHAR
device_source_concept_id INT
unit_concept_id INT
unit_source_value VARCHAR
unit_source_concept_id INT
Measurement
measurement_id INT
person_id INT
measurement_concept_id INT
measurement_date DATE
measurement_datetime DATETIME
measurement_time VARCHAR
measurement_type_concept_id INT
operator_concept_id INT
value_as_number FLOAT
value_as_concept_id INT
unit_concept_id INT
range_low FLOAT
range_high FLOAT
provider_id INT
visit_occurrence_id INT
visit_detail_id INT
measurement_source_value VARCHAR
measurement_source_concept_id INT
unit_source_value VARCHAR
unit_source_concept_id INT
value_source_value VARCHAR
measurement_event_id INT
meas_event_field_concept_id INT
Observation
observation_id INT
person_id INT
observation_concept_id INT
observation_date DATE
observation_datetime DATETIME
observation_type_concept_id INT
value_as_number FLOAT
value_as_string VARCHAR
value_as_concept_id INT
qualifier_concept_id INT
unit_concept_id INT
provider_id INT
visit_occurrence_id INT
visit_detail_id INT
observation_source_value VARCHAR
observation_source_concept_id INT
unit_source_value VARCHAR
qualifier_source_value VARCHAR
value_source_value VARCHAR
observation_event_id INT
obs_event_field_concept_id INT
Note
note_id INT
person_id INT
note_date DATE
note_datetime DATETIME
note_type_concept_id INT
note_class_concept_id INT
note_title VARCHAR
note_text VARCHAR
encoding_concept_id INT
language_concept_id INT
provider_id INT
visit_occurrence_id INT
visit_detail_id INT
note_source_value VARCHAR
note_event_id INT
note_event_field_concept_id INT
Note_nlp
note_nlp_id INT
note_id INT
section_concept_id INT
snippet VARCHAR
offset VARCHAR
lexical_variant VARCHAR
note_nlp_concept_id INT
note_nlp_source_concept_id INT
nlp_system VARCHAR
nlp_date DATE
nlp_datetime DATETIME
term_exists VARCHAR
term_temporal VARCHAR
term_modifiers VARCHAR
Specimen
specimen_id INT
person_id INT
specimen_concept_id INT
specimen_type_concept_id INT
specimen_date DATE
specimen_datetime DATETIME
quantity FLOAT
unit_concept_id INT
anatomic_site_concept_id INT
disease_status_concept_id INT
specimen_source_id VARCHAR
specimen_source_value VARCHAR
unit_source_value VARCHAR
anatomic_site_source_value VARCHAR
disease_status_source_value VARCHAR
Fact_relationship
domain_concept_id_1 INT
fact_id_1 INT
domain_concept_id_2 INT
fact_id_2 INT
relationship_concept_id INT
Cost
cost_id INT
cost_event_id INT
cost_domain_id VARCHAR
cost_type_concept_id INT
currency_concept_id INT
total_charge FLOAT
total_cost FLOAT
total_paid FLOAT
paid_by_payer FLOAT
paid_by_patient FLOAT
paid_patient_copay FLOAT
paid_patient_coinsurance FLOAT
paid_patient_deductible FLOAT
paid_by_primary FLOAT
paid_ingredient_cost FLOAT
paid_dispensing_fee FLOAT
payer_plan_period_id INT
amount_allowed FLOAT
revenue_code_concept_id INT
revenue_code_source_value VARCHAR
drg_concept_id INT
drg_source_value VARCHAR
Payer_plan_period
payer_plan_period_id INT
person_id INT
payer_plan_period_start_date DATE
payer_plan_period_end_date DATE
payer_concept_id INT
payer_source_value VARCHAR
payer_source_concept_id INT
plan_concept_id INT
plan_source_value VARCHAR
plan_source_concept_id INT
sponsor_concept_id INT
sponsor_source_value VARCHAR
sponsor_source_concept_id INT
family_source_value VARCHAR
stop_reason_concept_id INT
stop_reason_source_value VARCHAR
stop_reason_source_concept_id INT
Condition_era
condition_era_id INT
person_id INT
condition_concept_id INT
condition_era_start_date DATE
condition_era_end_date DATE
condition_occurrence_count INT
Drug_era
drug_era_id INT
person_id INT
drug_concept_id INT
drug_era_start_date DATE
drug_era_end_date DATE
drug_exposure_count INT
gap_days INT
Dose_era
dose_era_id INT
person_id INT
drug_concept_id INT
unit_concept_id INT
dose_value FLOAT
dose_era_start_date DATE
dose_era_end_date DATE
Episode
episode_id INT
person_id INT
episode_concept_id INT
episode_start_date DATE
episode_start_datetime DATETIME
episode_end_date DATE
episode_end_datetime DATETIME
episode_parent_id INT
episode_number INT
episode_object_concept_id INT
episode_type_concept_id INT
episode_source_value VARCHAR
episode_source_concept_id INT
Episode_event
episode_id INT
event_id INT
episode_event_field_concept_id INT
Cohort
cohort_definition_id INT
subject_id INT
cohort_start_date DATE
cohort_end_date DATE
Cohort_definition
cohort_definition_id INT
cohort_definition_name VARCHAR
cohort_definition_description VARCHAR
definition_type_concept_id INT
cohort_definition_syntax VARCHAR
subject_concept_id INT
cohort_initiation_date DATE
Legend
Clinical data tables
Health system data tables
Health economics data tables
Standardized derived elements
Metadata tables
Vocabulary tables
Primary key

OHDSI



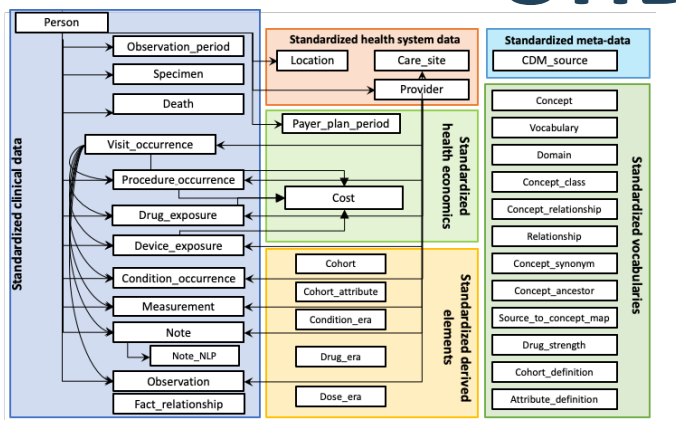
CDM v5.0.1

2016



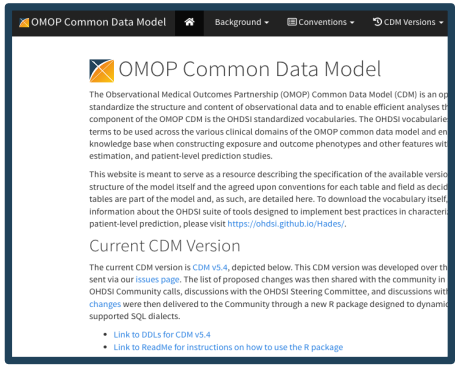
CDM v5.1

2017

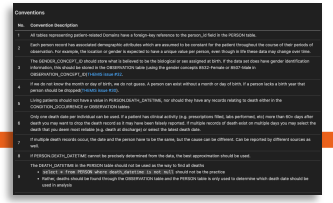


CDM v5.2

2017



2018

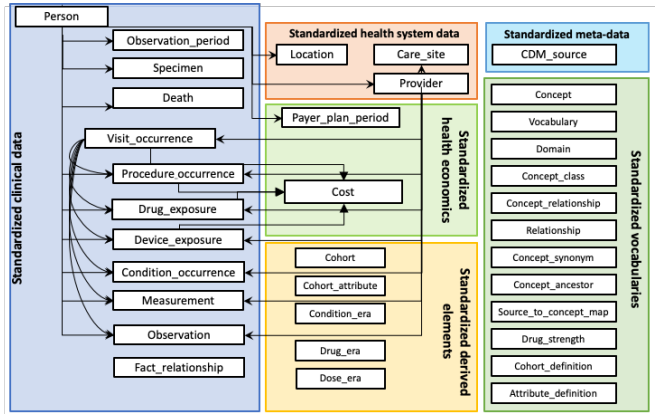


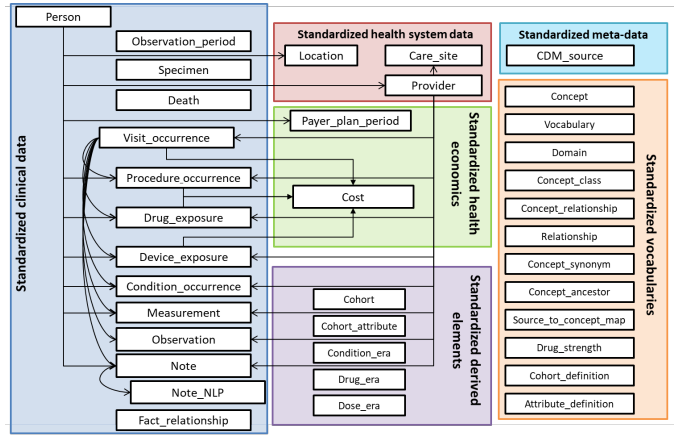
CDM v5.3

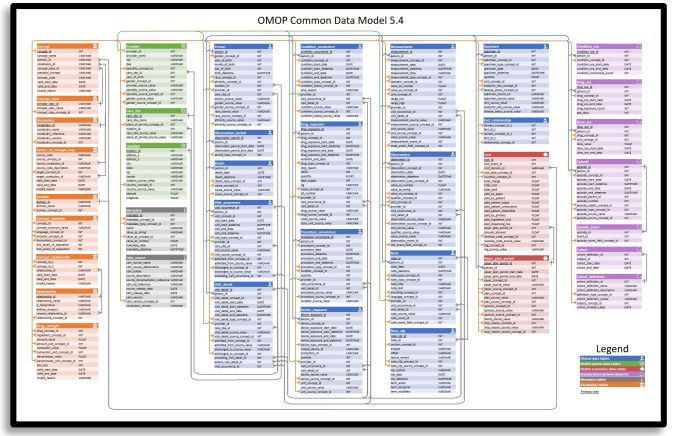
2022

CDM v5.4

2024







THEMIS Repository | Nav | THEMIS GitHub | CDM Docs | DQD Docs | Feedback

THEMIS Conventions

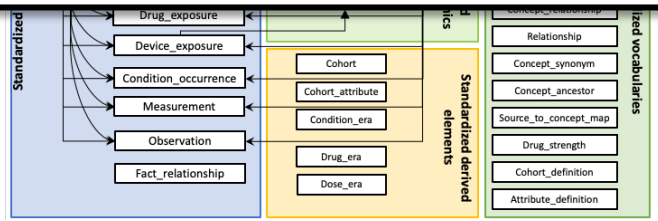
- General Conventions
- CDM Tables
- Tag Browser

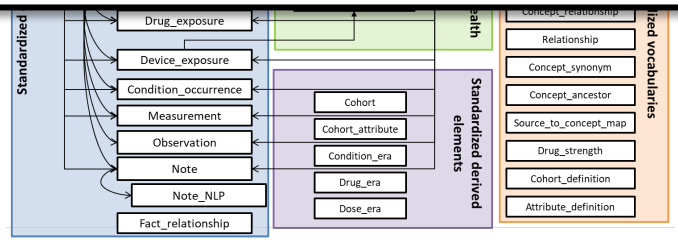
THEMIS

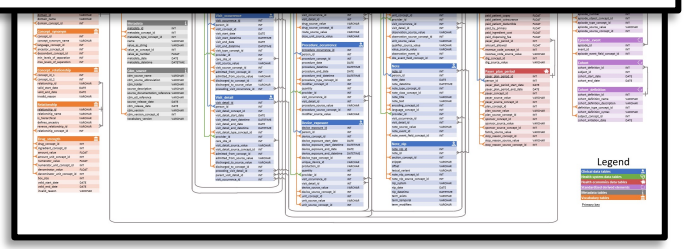
Summary: This page provides an overview of the THEMIS project.

Table of Contents

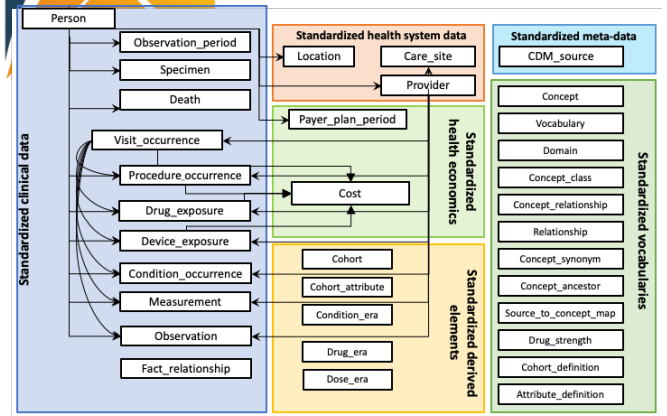
- Introduction
  - What falls under Themis and what does not
- Topic Processing
- Getting Involved





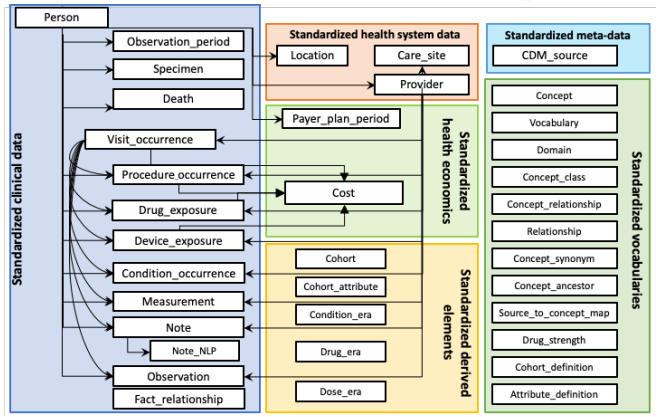


OHDSI



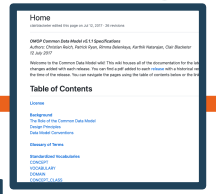
CDM v5.0.1

2016



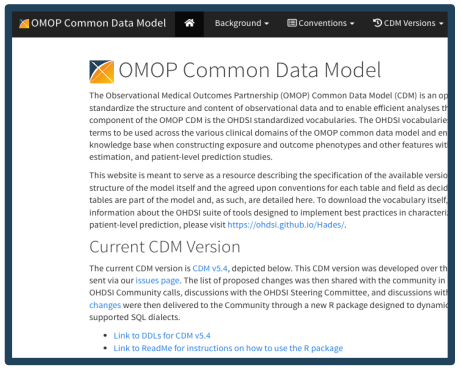
CDM v5.2

2017

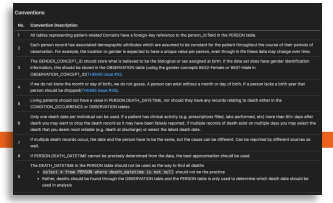


CDM v5.1

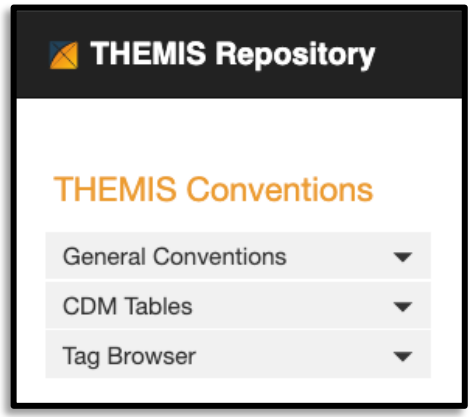
2017



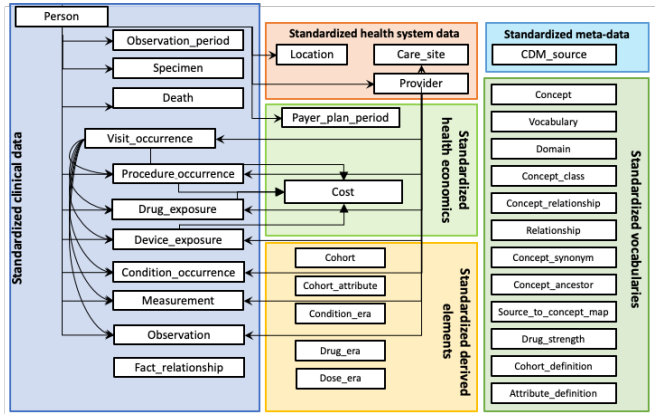
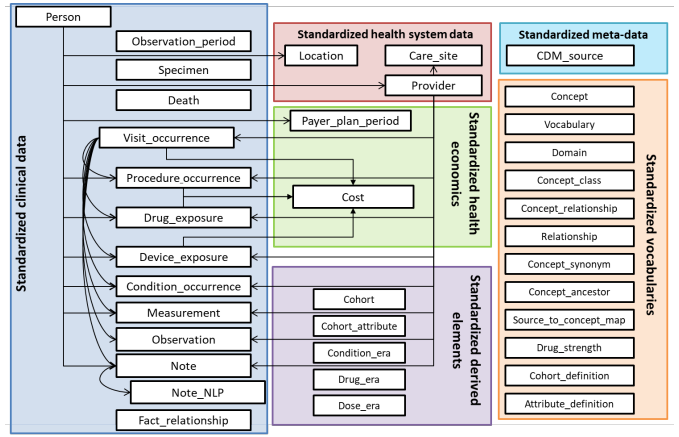
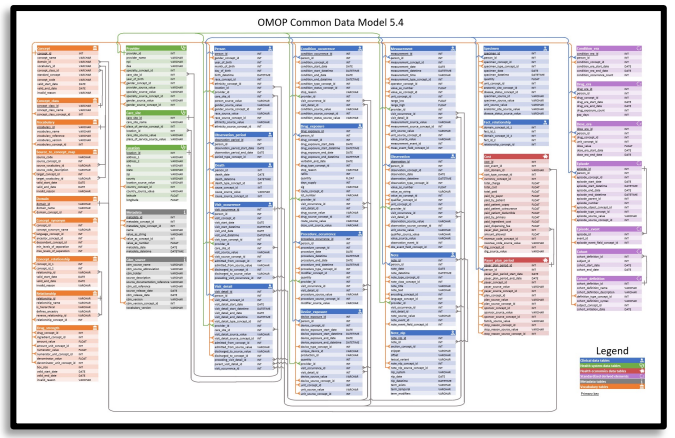
2018



CDM v5.3



2022

CDM v5.4

2024