

OMOP exports

Guide for understanding the OMOP-CDM tables from the EBMT Registry

November 2024

EBMT Registry

EBMT Clinical Research & Registry Department

Introduction

At its core, the Observational Medical Outcomes Partnership Common Data Model (OMOP-CDM) is a standard describing:

- a) a set of tables
- b) which data items should populate which tables, and in which rows and columns
- c) a definition of how data in each of these tables should relate to each other.

In other words, it's a database schema. Alongside encouraging the use of standardised clinical coding systems, like SNOMED, to describe healthcare data items, using the OMOP-CDM should allow researchers and analysts to develop sets of tools and methods that can be deployed wherever the CDM is in place.

An organisation called Observational Health Data Sciences and Informatics (OHDSI) is responsible for maintaining and developing the OMOP-CDM standards. OHDSI are a central resource for the OMOP-CDM: They have fostered a community of researchers collaborating on open source-tooling to analyse the CDM, discussing best practices on their forum, providing training on and documentation about the CDM, alongside much more, as a visit to their [website](#) will show you.

Here are some more useful documents and links to understand the OMOP basics:

1. **Useful information for OMOP-CDM v5.3, v.5.4, v6.0:**
<https://ohdsi.github.io/CommonDataModel/index.html>
2. **Athena** (for all coding-related issues): <https://athena.ohdsi.org/search-terms/start>
3. **OHDSI query library with example queries in SQL:** <https://data.ohdsi.org/QueryLibrary/>

The rest of this document will focus on the EBMT implementation of the OMOP-CDM, which, although will currently provide functionality for many types of analyses and report generation, is not as complete a deployment of the schema as we intend. Ongoing processes in the registry to expand our use of the standard will improve the utility of our implementation with time, and continuing updates to this document will reflect those changes.

Basic Tables of OMOP

There are various tables used in **OMOP-CDM v5.4** to capture all aspects of a patient's healthcare journey, the most important ones are:

- Person;
- Condition_occurrence;
- Observation;
- Measurement;

- Procedure_occurrence.
- Drug_exposure;
- Death;
- Fact_relationship;
- Location;
- Care_site;
- Concept;

There are other tables that are not used by EBMT, including but not limited to:

- Devices;
- Episode;
- Episode_event;
- Visit;
- Visit_occurrence;

There is also one non-OMOP-CDM 5.4 table that EBMT currently uses to identify the origin of data items and link between events:

- Form_option

The relationships between tables help build a complete picture for analysis.

Tables and table columns vary in different OMOP-CDM versions, please make sure to check only the currently used by EBMT: OMOP-CDM v5.4 version.

Person

<https://ohdsi.github.io/CommonDataModel/cdm54.html#person>

The **person** table contains a single row for each person (patient or donor from the donor outcome registry) in the database. Crucial for linkage of records throughout the database, the `person_id` field is one of the most important in the CDM.

Column	Description
person_id	A unique ID for every person in the database
gender_concept_id	A concept_id code for the gender or sex of the person at birth
year_of_birth	Year in which person was born
month_of_birth	Month in which person was born
day_of_birth	Day on which person was born
birth_datetime	Not populated by EBMT

race_concept_id	A concept_id to code the race of the person
ethnicity_concept_id	A concept_id to code for the ethnicity of the person
location_id	ID for the location that links to the location_id in the location table
provider_id	Not populated by EBMT
care_site_id	ID for the care site that links to the care_site_id in the care_site table
person_source_value	An ID that links back to a unique identifier in the source database from which OMOP is populated
gender_source_value	A value from the source database representing gender or sex at birth
gender_source_concept_id	Not populated by EBMT
race_source_value	Not populated by EBMT
race_source_concept_id	Not populated by EBMT
ethnicity_source_value	Not populated by EBMT
ethnicity_source_concept_id	Not populated by EBMT
donor	Not standard OMOP-CDM, implemented for the donor outcome registry. TRUE / FALSE indicating whether the person was a donor.

Condition_occurrence

https://ohdsi.github.io/CommonDataModel/cdm54.html#condition_occurrence

The **condition_occurrence** table contains all information on diseases or syndromes that were found in a patient. Note that absence of a certain disease (e.g. the patient did not have a pulmonary comorbidity) can be found in the **observation** table. The condition occurrence only contains positive responses to whether a disease, comorbidity, or syndrome was present. Again, the **person_id** can be used to link across different tables.

Column	Description
condition_occurrence_id	This column contains the id by which this condition is known in the EBMT Registry
condition_concept_id	This column contains the concept_id for the condition
condition_start_date	This column contains the start date of the condition, i.e. the diagnosis date of the condition
condition_start_datetime	Not populated by EBMT
condition_end_date	This column contains the date on which the condition ended
condition_end_datetime	Not populated by EBMT

condition_type_concept_id	This column indicates where the data was extracted from. In the case of the EBMT Registry, it will always be 32879 (Registry)
condition_status_concept_id	This column contains the type of condition, e.g. if it was a primary diagnosis
stop_reason	Not populated by EBMT
condition_source_value	This column contains a reference to the group (within a form), field and value from which the row was generated.
condition_source_concept_id	Not populated by EBMT
condition_status_source_value	Not populated by EBMT
person_id	This column can be used to link the condition to the patient using the person table
care_site_id	Not populated by EBMT
visit_detail_id	Not populated by EBMT
visit_occurrence_id	Not populated by EBMT
provider_id	Not populated by EBMT
form_option_id	Non-OMOP-CDM column. This column contains the form option ID, that can be used to identify the form option using the form_option file
patient_event_id	Non-OMOP-CDM column. This column contains the identifier of the patient event on the application
is_date_defaulted	Non-OMOP-CDM column. This column indicates if the event date was used as a data, if no direct reference date for the condition was available
is_ongoing	Non-OMOP-CDM column. This column indicates if the end_date is marked as ongoing

Observation

<https://ohdsi.github.io/CommonDataModel/cdm54.html#observation>

The **observation** table contains a variety of data fields that cannot be covered in any of the other domains. For example, because there is no standardised test to make the observation. In EBMT's case, the observation table also contains information about donors in allogeneic HCTs or CTs.

Note: The patient consent information is in the observation table too.

Column	Description
--------	-------------

observation_id	This column contains the id by which this observation is known in the EBMT Registry
observation_concept_id	This column contains the a concept_id coding for the observation
observation_date	This column contains the date of the observation, most likely the event date of the event the observation was recorded on
observation_datetime	Not populated by EBMT
observation_type_concept_id	This column indicates where the data was extracted from. In the case of the EBMT Registry, it will always be 32879 (Registry)
value_as_number	The numerical value of the observation that was done, e.g. the chronological number of an HCT
value_as_string	This column can contain non-numerical data pertaining to the observation, providing flexibility to give extra information when the value_as_concept_id or value_as_number might benefit from additional context.
value_as_concept_id	The concept id of the observation if it was a categorical observation, e.g. the donor's sex at birth in an allogeneic HCT
qualifier_concept_id	Not populated by EBMT
unit_concept_id	This column contains the unit of the observation coded as a concept_id, if applicable
observation_source_value	This column contains a reference to the form, field and value that created the particular row in the table
observation_source_concept_id	Not populated by EBMT
unit_source_value	The unit of measure pertaining to the value_as_number, as written in the source (i.e., the form).
qualifier_source_value	Not populated by EBMT
value_source_value	Not populated by EBMT
observation_event_id	Not populated by EBMT
obs_event_field_concept_id	Not populated by EBMT
person_id	This column can be used to link the observation to the patient using the person table
visit_detail_id	Not populated by EBMT
visit_occurrence_id	Not populated by EBMT
provider_id	Not populated by EBMT
form_option_id	Non-OMOP-CDM column. This column contains the form option ID, that can be used to identify the form option using the form_option table

patient_event_id	Non-OMOP-CDM column. This column contains the identifier of the patient event in the application
is_date_defaulted	Non-OMOP-CDM column. This column indicates if the event date was used as a data, if no direct reference date for the condition was available
group_index	Non-OMOP-CDM column. Where the data row comes from a repeater group in the form, this will increment as repeater groups are added.

Measurement

<https://ohdsi.github.io/CommonDataModel/cdm54.html#measurement>

The **measurement** table contains data of all patient-related measurements. Measurements can be scores or measurements that were obtained from patient examinations.

Column	Description
measurement_id	This column contains the id by which this measurement is known in the EBMT Registry
measurement_concept_id	This column contains concept_ids of the measurement.
measurement_date	This column contains the date of the measurement, most likely the event date of the event the measurement was recorded on
measurement_datetime	Not populated by EBMT
measurement_time	Not populated by EBMT
measurement_type_concept_id	This column indicates where the data was extracted from. In the case of the EBMT Registry, it will always be 32879 (Registry)
operator_concept_id	This column contains a concept_id that references an operator like ">" (greater than) when e.g., the measurement_value_as_number requires a non-numeric prefix. This column can be joined to a concept_name on the concept_id of the concept table
value_as_number	The numerical value of the measurement that was done, e.g. the Lansky score.
value_as_concept_id	The concept id of the measurement, if it was a categorical measurement, e.g. CTCAE term.
unit_concept_id	This column contains the concept ids pertaining to the unit of the measurement, if applicable.
range_low	Not populated by EBMT
range_high	Not populated by EBMT
measurement_source_value	This column contains a reference to the form, field and value that created this OMOP-CDM row

measurement_source_concept_id	Not populated by EBMT
unit_source_value	This column indicates the unit_concept_id verbatim, e.g. 'cm'
unit_source_concept_id	Not populated by EBMT
value_source_value	Not populated by EBMT
measurement_event_id	Not populated by EBMT
meas_event_field_concept_id	Not populated by EBMT
person_id	This column can be used to link the measurement to the patient using the person table
visit_detail_id	Not populated by EBMT
visit_occurrence_id	Not populated by EBMT
provider_id	Not populated by EBMT
form_option_id	Non-OMOP-CDM column. This column contains the form option ID, that can be used to identify the form option using the form_option file
patient_event_id	Non-OMOP-CDM column. This column contains the identifier of the patient event on the application
is_date_defaulted	Non-OMOP-CDM column. This column indicates if the event date was used as a data, if no direct reference date for the condition was available
group_index	Non-OMOP-CDM column. Where the data row comes from a repeater group in the form, this will increment as repeater groups are added.

Procedure_occurrence

https://ohdsi.github.io/CommonDataModel/cdm54.html#procedure_occurrence

The **procedure_occurrence** table contains the data on treatments that patients have undergone. These treatments can be drug administrations (if any drugs were given as treatments, for the exact drugs, see the 'drug_exposure' table), HCTs, or radiotherapy.

Column	Description
procedure_occurrence_id	This column contains the primary key of the procedure_occurrence table
procedure_concept_id	This column contains the procedure coded as a concept_id.
procedure_date	This column contains the date on which the procedure began (e.g. HCT event date)
procedure_datetime	Not populated by EBMT

procedure_end_date	This column contains the date on which the procedure ended
procedure_end_datetime	Not populated by EBMT
procedure_type_concept_id	This column contains a concept_id which codes to an indication of where the data was extracted from. In the case of the EBMT Registry, it will always be 32879 (Registry).
modifier_concept_id	Not populated by EBMT
quantity	Not populated by EBMT
procedure_source_value	This column contains a reference to the form, field and value that created this OMOP-CDM row
procedure_source_concept_id	Not populated by EBMT
modifier_source_value	Not populated by EBMT
person_id	This column can be used to link the procedure to the patient using the person table
visit_detail_id	Not populated by EBMT
visit_occurrence_id	Not populated by EBMT
provider_id	Not populated by EBMT
form_option_id	Non-OMOP-CDM column. This column contains the form option ID, that can be used to identify the form option using the form_option file
patient_event_id	Non-OMOP-CDM column. This column contains the identifier of the patient event on the application
is_date_defaulted	Non-OMOP-CDM column. This column indicates if the event date was used as a data, if no direct reference date for the procedure was available
is_ongoing	Non-OMOP-CDM column. This column indicates if the treatment was marked as 'ongoing'.
group_index	Non-OMOP-CDM column. Where the data row comes from a repeater group in the form, this will increment as repeater groups are added.

Drug_exposure

https://ohdsi.github.io/CommonDataModel/cdm54.html#drug_exposure

The **drug_exposure** table contains the information about a patient's exposure to drugs, or in other words: what drugs the patient received in their treatment history that is reported to the EBMT Registry.

Column	Description
--------	-------------

drug_exposure_id	This column contains the id by which this drug is known in the EBMT Registry
drug_concept_id	This column contains the drug coded as a concept_id.
drug_exposure_start_date	This column contains the start date of the drug exposure, i.e. the treatment start date
drug_exposure_start_datetime	Not populated by EBMT
drug_exposure_end_date	This column contains the end date of the drug exposure, i.e. the maintenance treatment end date
drug_exposure_end_datetime	Not populated by EBMT
verbatim_end_date	Not populated by EBMT
drug_type_concept_id	This contains a concept_id coding indicating the source of the data. In the case of the EBMT Registry, it will always be 32879 (Registry)
stop_reason	Not populated by EBMT
refills	Not populated by EBMT
quantity	This column contains the numeric component of the the dose of the drug that was administered - see dose_unit_source_value
days_supply	Not populated by EBMT
sig	Not populated by EBMT
route_concept_id	This column contains a concept id that indicates the route through which the drug was administered, e.g. oral or intravenous. This column can be joined to a concept_name on the concept_id of the concept table
lot_number	Not populated by EBMT
drug_source_value	This column contains a reference to the form, field and value that created this OMOP-CDM row
drug_source_concept_id	Not populated by EBMT
route_source_value	Not populated by EBMT
dose_unit_source_value	This unit contains the unit to which the dose refers.
person_id	The ID of the person to whom this row pertains. This column can be used to link the drug to the patient using the person table
visit_detail_id	Not populated by EBMT
visit_occurrence_id	Not populated by EBMT
provider_id	Not populated by EBMT
form_option_id	Non-OMOP-CDM column. This column contains the form option ID, that can be used to identify the form option using the form_option file

patient_event_id	Non-OMOP-CDM column. This column contains the identifier of the patient event on the application
is_date_defaulted	Non-OMOP-CDM column. This column indicates if the event date was used as a data, if no direct reference date for the drug was available
is_ongoing	Non-OMOP-CDM column. This column indicates if the end_date is marked as ongoing
group_index	Non-OMOP-CDM column. Where the data row comes from a repeater group in the form, this will increment as repeater groups are added.

Death

Where a person is known to have died, the **death** table will be populated with a row. At the EBMT if the cause of death isn't specifically related to the treatment or types of disorders about which we primarily collect data, the cause of death will be recorded as "Other" or "Unknown". The **person_id** can be used to link records across OMOP tables.

<https://ohdsi.github.io/CommonDataModel/cdm54.html#death>

Column	Description
person_id	The ID of the person to whom this row pertains. This column can be used to link the drug to the patient using the person table
death_date	Date on which the person became deceased
death_datetime	Populated in EBMT, but the time component is always midnight on the date of death
death_type_concept_id	A concept ID telling us the source of the death information. Currently the only code is for the concept "Registry", however, as more data sources are brought into our OMOP-CDM this may expand.
cause_concept_id	A concept ID telling us the cause of death. The current range of codes covers EBMT-related causes, anything else is coded to "Other".
cause_source_value	Not populated by EBMT
cause_source_concept_id	Not populated by EBMT
form_option_id	Non-OMOP-CDM column. Allows re-linking of forms via the non-standard form_option table.
patient_event_id	Non-OMOP-CDM column. Allows re-linking of forms via the non-standard form_option table.

Fact_relationship

https://ohdsi.github.io/CommonDataModel/cdm54.html#fact_relationship

The fact_relationship table enables linkage between different “facts” in the OMOP-CDM via each table’s primary ids (n.b., not the concept_id of the row, but the primary key id of the table). Understanding the fact_relationship table can be a little tricky.

An example of how the fact_relationship table works; fact_id_1 might reference an id in the procedure_occurrence table (i.e., the procedure_occurrence_id of that table), whilst fact_id_2 - in that same row of the fact_relationship table - might reference an id in the condition_occurrence table (analogously, in the condition_occurrence_id of that table). These two records are thus linked. The fact-relationship in this case might be that the person to whom this observation belongs had a procedure because of a condition.

The field domain_concept_id_1 for this row will have the concept_name “procedure_occurrence”, whilst the domain_concept_id_2 for this row will have the concept_name “condition_occurrence”. If we were to perform the join to the procedure_occurrence table on fact_id_1 = procedure_occurrence_id, and look at the procedure_concept_id for this row, we might see a value representing something like “Allogeneic bone marrow transplantation”, whilst the condition_concept_id for the row in condition_occurrence pertaining to fact_id_2 = condition_occurrence_id might have a value representing something like “Acute leukaemia”.

The relationship between these two “facts” is revealed by the relationship_concept_id field, which in this case might have a concept_id representing “Has due to”, and constructing the relationship we reveal that the allogeneic bone marrow transplantation (from the procedure_occurrence table, with procedure_occurrence_id = fact_id_1) is occurring due to (from the relationship_concept_id) acute leukaemia (from the condition_occurrence table, with condition_occurrence_id = fact_id_2).

Column	Description
_id	Non-OMOP-CDM, a primary ID for the fact_relationship table that has no bearing on anything else
domain_concept_id_1	This column contains the concept_ids representing the “domain” (table) to which fact_id_1 pertains. This column can be joined to a concept_name on the concept_id column in the concept table
fact_id_1	This column references the primary keys from the tables referenced in domain_concept_id_1, which link to records containing the first “fact”
domain_concept_id_2	This column contains the concept_ids representing the domain of the second “fact”. Similarly, this column can be joined to a concept_name on the concept_id column in the concept table
fact_id_2	This column references the primary key of the domain referenced in domain_concept_id_2, which link to records containing the second “fact”

relationship_concept_id	This column contains the concept_ids of the relationships between facts one and two. This column can also be joined to a concept_name on the concept_id in the concept table.
-------------------------	---

Care_site

https://ohdsi.github.io/CommonDataModel/cdm54.html#care_site

The **care_site** table contains information about all the centres at which care was provided with which the EBMT interacts.

Column	Description
care_site_id	The primary key field of the care_site table
care_site_name	The name of the care site. Not the address, just the name e.g., of the hospital.
place_of_service_concept_id	Not populated by EBMT
location_id	An id which links to the location_id in the location table, which can be used to ascertain further information about the location of the care site
care_site_source_value	This column contains strings which correspond to the data item in the application.
place_of_service_source_value	Not populated by EBMT

Location

<https://ohdsi.github.io/CommonDataModel/cdm54.html#location>

The **location** table contains information about all the locations of the care sites with which the EBMT interacts.

Column	Description
location_id	The primary key field of the location table. Links to the location_id in the care_site table
address_1	Not populated by EBMT
address_2	Not populated by EBMT
city	Not populated by EBMT
state	Not populated by EBMT
zip	Not populated by EBMT

county	Not populated by EBMT
location_source_value	This column contains an id from the application that references the country data item as stored in the source
country_concept_id	This column contains concept ids for countries
country_source_value	This column contains the country names in human-readable form
latitude	Not populated by EBMT
longitude	Not populated by EBMT

Concept

<https://ohdsi.github.io/CommonDataModel/cdm54.html#concept>

Last, but by no means least for the standard OMOP-CDM 5.4 tables: The **concept** table is one of the most important tables in the OMOP-CDM. A “concept” in the context of the OMOP-CDM is a piece of information, which can be anything from “Not performed” (concept_id 4118638) in the observation table’s observation_concept_id column, to “Proteinuria” (concept_id 75650) in the condition_occurrence table’s condition_concept_id.

More generally speaking, where x_concept_id columns are referenced across the many other OMOP-CDM tables, the **concept** table may be joined to this table on the concept_id column. The joined concept_name field will then provide human-readable information pertaining to the row from the table to which you have joined the **concept** table.

Column	Description
concept_id	This column contains the concept id. Where EBMT can pick a standard SNOMED concept_id from Athena, this will be used, otherwise a custom concept id will be created, though we do use other coding systems, such as RxNorm, LOINC, and OMOP Genomic amongst others.
concept_name	This row contains a short, human readable description of the concept
domain_id	EBMT populates this with a human-readable string that denotes which table the concept generally belongs to, according to the OMOP standard, though this relationship is not strict, especially regarding data in the observation table.
vocabulary_id	This column shows us from which coding system the concept_id was picked.
concept_class_id	This column shows us what type of concept the row pertains to. For example, concept_id 24006, coding for concept_name “Sickle cell-haemoglobin C disease” has the concept_class “Disorder”.
standard_concept	This column indicates whether the concept is standard (“S”, meaning that the concept conforms to standard OMOP mapping rules), a classification concept (“C”,

	which is used for broader grouping of concepts), or, if null, indicates that it is a concept outwith the OMOP-CDM.
concept_code	This column contains the source code from which the concept_id (as used throughout the rest of the OMOP tables) was generated.
valid_start_date	This column shows the date on which the code was first used
valid_end_date	The date on which the code was retired (because it became superseded by another code, for example).
invalid_reason	If the code has been retired, the reason given: Can be "U" for "updated", or "D" for "deleted", and is null where the valid_end_date has not been passed.

Non-OMOP-CDM:

Form_option

The form_option table is not in the OMOP-CDM but has been loaded into MicroStrategy as currently our OMOP-CDM implementation, lacking the visit_occurrence and visit_detail tables, make it difficult to provide context to reports. The form_option_id field, which is found in the most important clinical tables, can be joined on the id field of the form_option table, allowing access to fields such as the event_type_name from the form_option table.

Column	Description
id	The primary key field of the form_option table. In the condition_occurrence, observation, measurement, procedure_occurrence, drug_exposure, and death tables this field can be joined to the form_option_id
event_type_name	This column contains the name of the event type to which the id pertains. Values in this field will be familiar to users of the EBMT as they often reference the form names, e.g., "HCT Day 100", "Allogenic HCT", etc.
group_label	This column contains the name of the sub-section from the form to which the id pertains. Typical values will again be familiar to users of the registry: Where the event_type_name is "Allogenic HCT" for example, group_labels might be "Patient serological status" or "Main treatment description"
field_label	This column contains the name of the individual fields within the groups, within the forms. Running with the example of "Allogenic HCT", field_labels within the "Patient serological status" would be "Patient CMV status" and "Patient EBV status"
date_field	This column contains the name, in human readable format, of the date field associated with the field itself. This can be useful for determining whether the date of the particular row (from one of the clinical tables, to which you would join the form_option table) is associated with e.g., a date of diagnosis, or a date of onset of a certain condition.

date_field_id	This column contains a string that references the date field to which the field itself is associated.
data_service_name	This column contains the name of the data service to which the field is linked, if any. A data service contains the mappings between options selected on the forms, (from, for example, a drop-down list) their names as displayed in the forms, and their values as represented in the databases in terms of concept_ids and source_ids
service_option_label	This column contains the labels as displayed in the forms. For example, under the event_type_name "Acute leukaemia", with the group_label "Chromosomal abnormalities" the data_service_name we call upon for that would be called "XXX_Chromosomal abnormalities" and the service_option_labels would have individual abnormalities listed, like "del(3q) / 3q-" or "-16 / monosomy 16".
service_option_value	This column contains the values corresponding to the service_option label as they are coded in the data service.
concept_id	This column contains a concept_id that maps to the row. For example, the service_option_label "Acute myeloid leukaemia (AML)" will map to the concept_id 140352: the code for AML
concept_name	Currently, very few concept names have been populated in the form_option table, so this field is of limited use. Joining the concept table on the concept_id is recommended if you want to view the concept_names
omopcdm_table	This column tells you where in the OMOP-CDM the data associated with the form_option row would be found
value_as_concept_id	Where there are values from the forms that can be represented as concept ids, the value_as_concept_id column will contain concept ids that correspond to those values
unit_concept_id	Where the form options pertain to e.g., numerics, this column will contain a concept_id that corresponds to the unit of measurement. For example, with the field_label "Age in months", the user will input a numeric value, and the unit_concept_id will be "9580", which is the concept_id code for "month"
field_id	This column contains a string that identifies the specific field into which the data was input.
event_type_id	This column contains a string that identifies the specific event type the row is found in. Again, the values in this column will be familiar to users of the EBMT as they are the names of the forms ("Status at HCT/CT/IST", "HCT Day 100", etc)
data_service_option_id	This column contains a string that identifies the specific data service option in the application.
field_version_id	This column contains a string that identifies the version of the field_id. As it is specific to the field_id, and as no other versions are populated into the database, this field currently has a 1:1 correspondence with the field_id - i.e., every unique value of field_id has a corresponding unique value in field_version-id

data_service_option_version_id	Similarly to the field_version_id, this column contains a string that identifies the version of the data_service_option_id, and again has a 1:1 correspondence.
--------------------------------	---

Linking between tables

Linking between tables is essential for building meaningful reports from the OMOP-CDM. Whilst it will be possible to ascertain simple summary statistics without linkage, e.g., counts of conditions by counting the number of condition_concept_ids, anything more complex will require linkage between tables. As an example, one of the most-used linkages will be with the concept table: All the clinical tables in the OMOP-CDM (and most of the non-clinical tables) will have one or more fields with the suffix “_concept_id”, indicating that this field can be joined on (or looked-up against) the concept_id in the concept table, which (amongst other things) lets you access the human-readable concept_name.