



Starting with the OMOP CDM

Khaleeq Mannan

Contents

- An introduction to OMOP.
- Basic structure of the OMOP CDM.
- Tutorial on how to use concepts.
- A look into how tables are structured.
- A brief introduction of derived tables.

Overview

- A OHDSI (Observational Health Data Sciences and Informatics) is a multi-stakeholder, interdisciplinary collaborative.
- The OMOP CDM (Observational Medical Outcomes Partnership Clinical Data Model) aims to standardize clinical databases around the world. It provides a schema and a dictionary.
- The OMOP CDM is designed for use in observational studies.

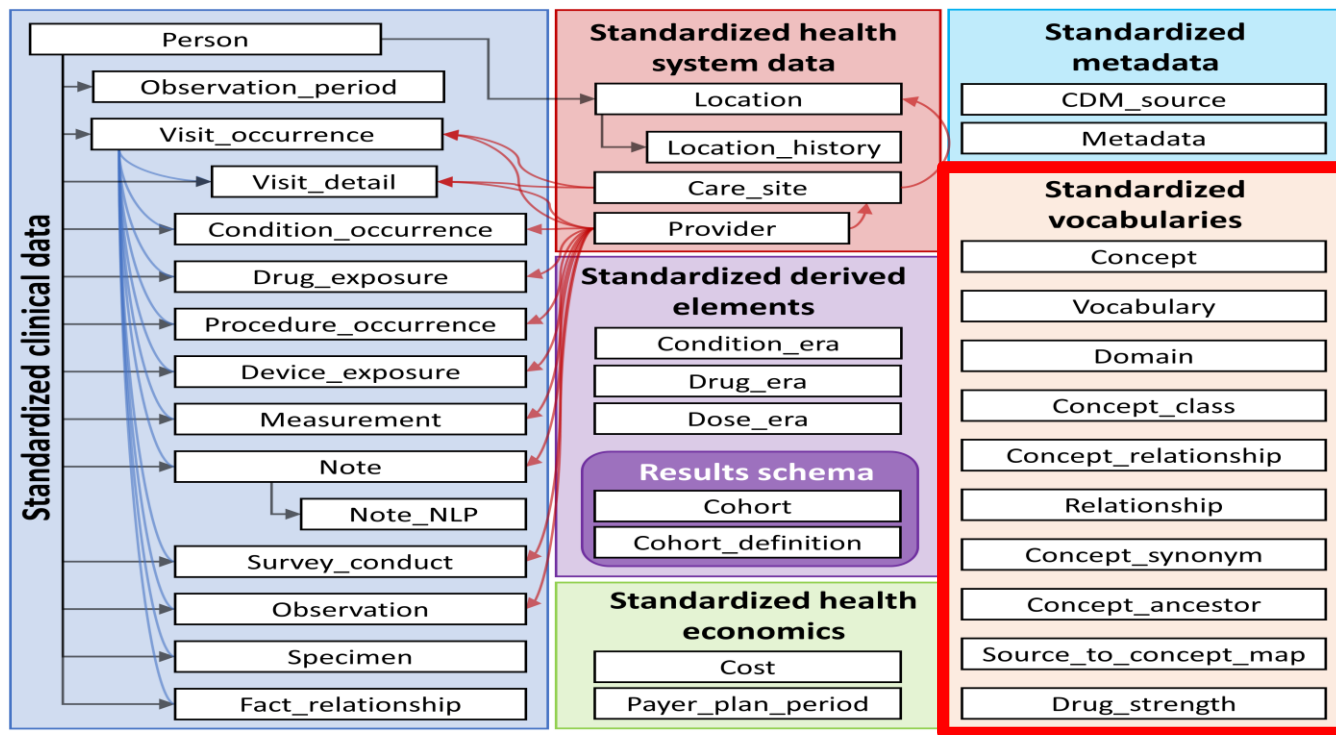
General Ideas

- Concepts – Diagnoses, drugs, procedures etc.
- Concept ID – Each concept has a concept ID. The list of all concept IDs is part of the OMOP CDM. Different concept IDs may represent the same real world concept.
- Standard concepts – The reference concept selected in OMOP. This concept will be used to represent the real world concept across the database.
- Relationships – How different concepts may relate to each other.
- Domains – The type of concept. Determines which table contains information about a given concept.

Example

- Consider the condition, atrial fibrillation.
- Atrial fibrillation is represented by multiple different concepts depending on the dictionary used to identify it in the source data e.g. ICD, SNOMED etc.
- The concept ID for atrial fibrillation from SNOMED is 313217. The concept ID for atrial fibrillation from ICD9CM is 44821957.
- The standard concept used across the OMOP CDM will be atrial fibrillation from SNOMED so the standard concept ID is 313217.
- The domain ID for the concept ID 313217 is 'Condition'. This tells us that the concept ID 313217 refers to the condition, atrial fibrillation.

Tables



Blacketer, C. 2020. *The book of OHDSI*. [Online]. [Accessed 21/01/2021].
Available from :<https://ohdsi.github.io/TheBookOfOhdsi/StandardizedVocabularies.html>

Tables: Concept

- Contains the list of all concepts.
- Includes information on the concept IDs, concept names, standard concepts, domain IDs etc.
- Can be used to find the concept ID of a real world concept by using the real world name of the concept.

Example

Below is the list of all concepts from the Concept table with the concept name 'Atrial Fibrillation'.

'S' in the STANDARD_CONCEPT column indicates the standard concept. The concept ID in this row is the ID that is used across the database when referring to the condition atrial fibrillation.

The name of the real world concept is specified. In this case, we are looking at all concepts with concept name 'Atrial Fibrillation'.

Note that not all concepts with concept name 'Atrial Fibrillation' are conditions. This is not the standard concept for the condition atrial fibrillation, rather it is the standard concept for the atrial fibrillation clinical observation.

The standard ID of the source dictionary and the code within the dictionary are provided. In this case, the focus is on the ICD9CM code for atrial fibrillation.

	concept_id	concept_name	domain_id	vocabulary_id	standard_concept	concept_code
1	45500785	Atrial fibrillation	Condition	Read		G573000
2	44821957	Atrial fibrillation	Condition	ICD9CM		427.31
3	313217	Atrial fibrillation	Condition	SNOMED	S	49436004
4	45611600	Atrial Fibrillation	Condition	MeSH		D001281
5	45883018	Atrial fibrillation	Meas Value	LOINC	S	LA17084-7
6	45951191	Atrial Fibrillation	Condition	CIEL		148203
7	4344544	Atrial Fibrillation	Drug	NDFRT	C	N0000000507

Tables: Concept_relationship

- Contains the list of all concept relationships.
- Specifies the nature of all relationships via relationship ID.
- Can be used to easily find the standard concept for any given concept ID.

Example

Below is the list of all concept relations from the Concept_relation table for the concept of the condition atrial fibrillation taken from the ICD9CM dictionary.

	concept_id_1	concept_id_2	relationship_id
1	44821957	44834731	Is a
2	44821957	44824248	Is a
3	44821957	313217	Maps to

We search for all rows where concept_id_1 is 44821957

Note that 'Maps to' isn't the only relationship between concepts. In this case, the concept with concept ID 44824248 is the condition 'Atrial fibrillation and flutter' taken from the ICD9CM dictionary. This suggests that the concept with ID 44824248 is a parent to the concept with ID 44821957. It is worth mentioning that hierarchical relations are maintained within the source dictionary (in this case ICD9CM).

We find that the concept with concept ID 44821957 'Maps to' the concept with concept ID 313217. This concept ID represents the condition atrial fibrillation taken from SNOMED i.e. the standard concept for the condition atrial fibrillation.

Tables: Concept_ancestor

- Lists all hierarchal relationships between concepts.
- Includes information on the levels of separation between two concepts in the same hierarchy.
- Can be used to find multiple concepts that are all classified under a single concept.

Example

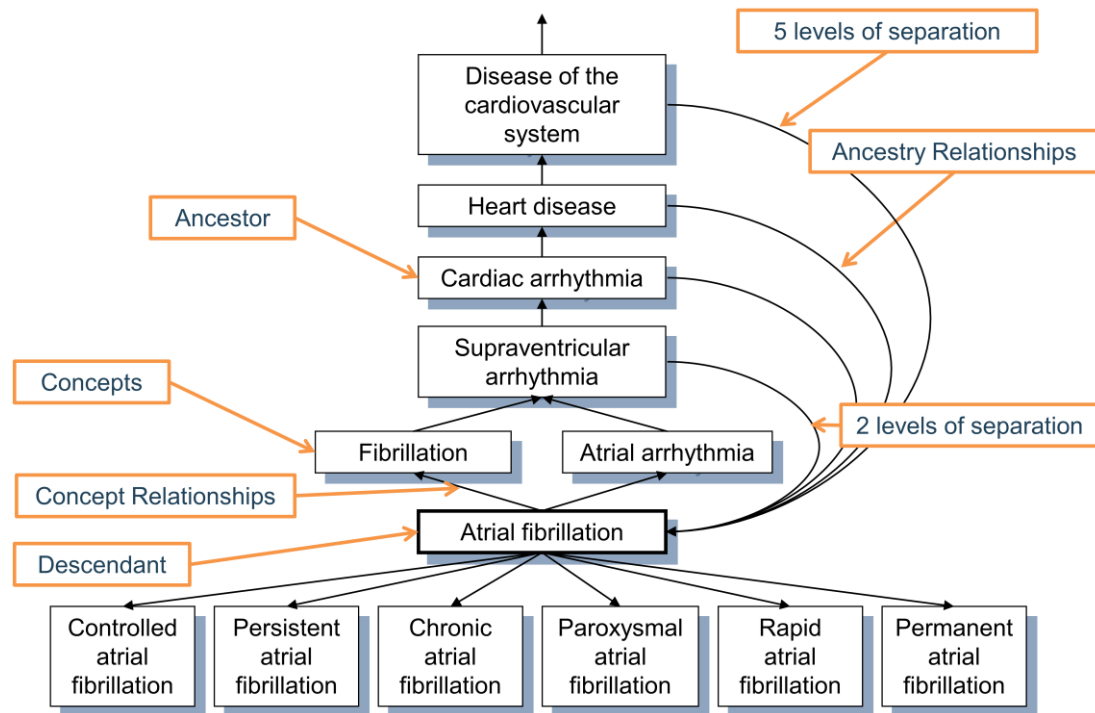
I have searched for all records where the descendant concept has concept ID '313217' (The concept names of ancestors and descendants have been joined on).

I now have all concept IDs for all concepts that are ancestors of atrial fibrillation that have a minimum of 4 levels of separation with atrial fibrillation.

The levels of separation indicate how many links in the hierarchy diagram are between the descendant and ancestor concept.

	ancestor_concept_id	descendant_concept_id	min_levels_of_separation	max_levels_of_separation	ancestor_concept_name	descendant_concept_name
1	313217	313217	0	0	Atrial fibrillation	Atrial fibrillation
2	4226399	313217	1	1	Fibrillation	Atrial fibrillation
3	500001801	313217	1	1	OMOP Qt Prolongation/Torsade De Pointes 1	Atrial fibrillation
4	4068155	313217	1	1	Atrial arrhythmia	Atrial fibrillation
5	500002401	313217	1	1	OMOP Atrial Fibrillation 1	Atrial fibrillation
6	4248028	313217	2	2	Supraventricular arrhythmia	Atrial fibrillation
7	44784217	313217	2	3	Cardiac arrhythmia	Atrial fibrillation
8	321588	313217	3	4	Heart disease	Atrial fibrillation
9	440142	313217	4	5	Disorder of mediastinum	Atrial fibrillation
10	4103183	313217	4	5	Cardiac finding	Atrial fibrillation
11	4266169	313217	4	5	Disorder of body cavity	Atrial fibrillation
12	134057	313217	4	5	Disorder of cardiovascular system	Atrial fibrillation

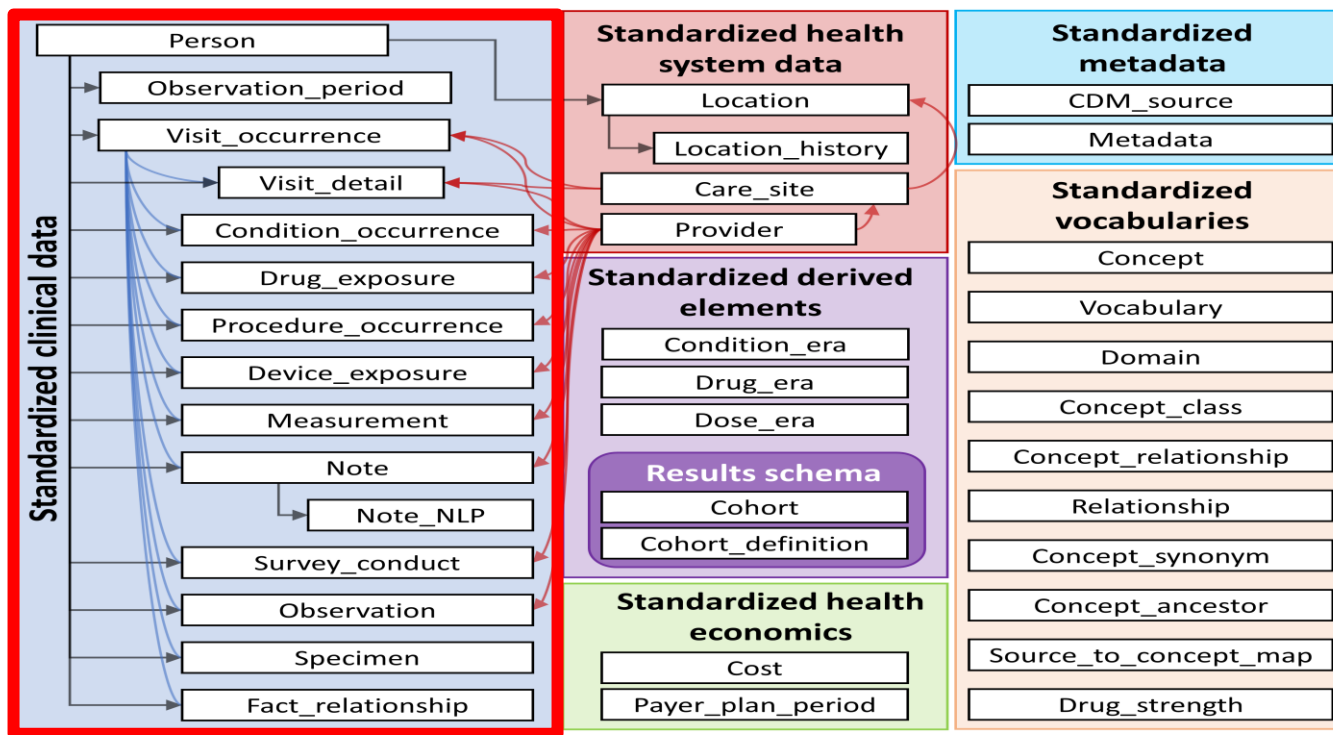
Example



This diagram represents parts of the table from the previous slide. Levels of separation indicates the number of 'links' between two conditions in the tree.

Ostropolets, A & Reich, C. 2020. *The book of OHDSI*. [Online]. [Accessed 11/12/2020]. Available from :<https://ohdsi.github.io/TheBookOfOhdsi/StandardizedVocabularies.html>

Tables



Blacketer, C. 2020. *The book of OHDSI*. [Online]. [Accessed 21/01/2021].
Available from :<https://ohdsi.github.io/TheBookOfOhdsi/StandardizedVocabularies.html>

Tables

- Person
 - Lists all people that can be identified in the database.
 - Includes basic demographic information like birthdate, gender, race etc.
- Observation_period
 - Lists any period in the database where a person is observed.
 - Includes the observation type of each period.
- Visit_occurrence
 - Lists all medical encounters.
 - Contains information about the type of visit.

Tables

- Condition_occurrence
 - Lists all signs, symptoms and diagnoses that are reported.

- Drug_exposure
 - Lists all instances of exposure to drugs.

- Procedure_exposure
 - Lists all instances of exposure to procedures.

Clinical Tables Structure

```
129. CONDITION_OCCURRENCE_ID
129. PERSON_ID
129. CONDITION_CONCEPT_ID
129. CONDITION_START_DATE
129. CONDITION_START_DATETIME
129. CONDITION_END_DATE
129. CONDITION_END_DATETIME
129. CONDITION_TYPE_CONCEPT_ID
Aa STOP_REASON
129. PROVIDER_ID
129. VISIT_OCCURRENCE_ID
129. CONDITION_STATUS_CONCEPT_ID
129. CONDITION_SOURCE_CONCEPT_ID
Aa CONDITION_SOURCE_VALUE
Aa CONDITION_STATUS_SOURCE_VALUE
```

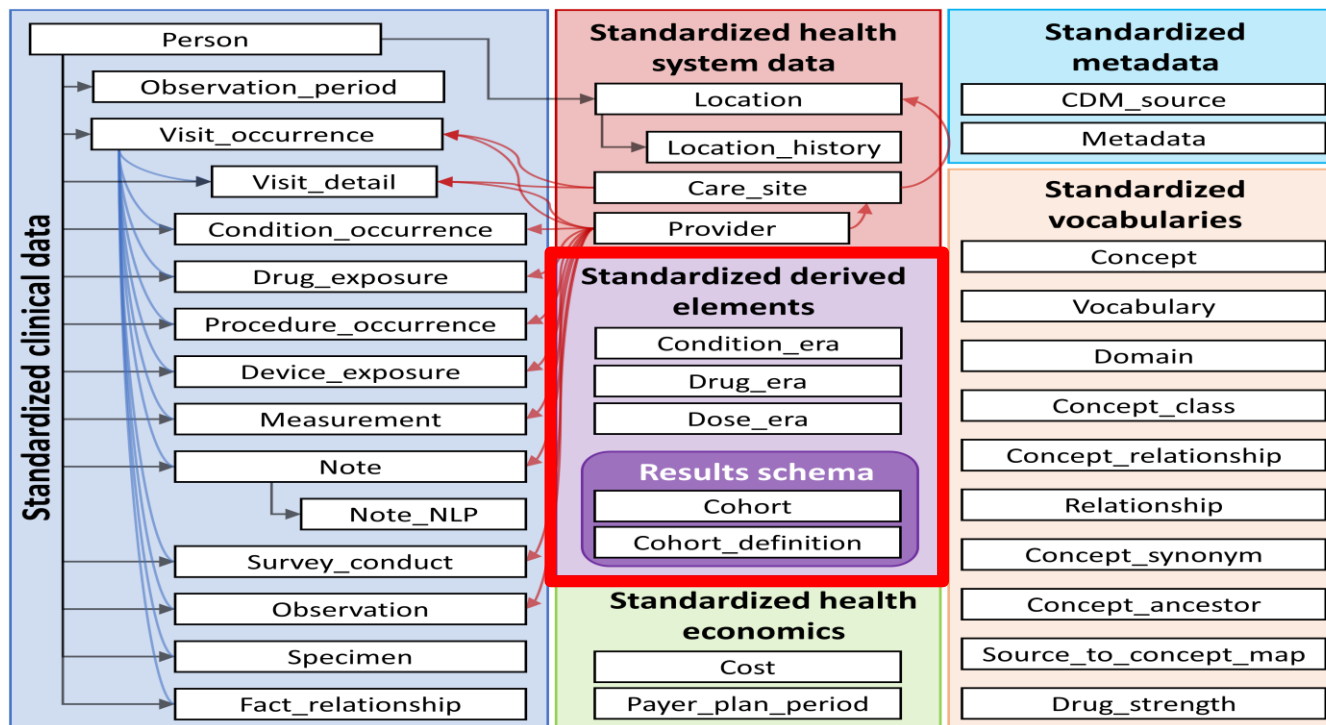
- Unique key is always *TABLE NAME*_ID.
- The primary concept type concept ID is always *PRIMARY CONCEPT TYPE*_CONCEPT_ID.
- The source concept is also available in the column with name *PRIMARY CONCEPT TYPE*_SOURCE_CONCEPT_ID.

Clinical Tables Structure

```
123. CONDITION_OCCURRENCE_ID
67.. PERSON_ID
123. CONDITION_CONCEPT_ID
67.. CONDITION_START_DATE
123. CONDITION_START_DATETIME
67.. CONDITION_END_DATE
123. CONDITION_END_DATETIME
67.. CONDITION_TYPE_CONCEPT_ID
Aa STOP_REASON
123. PROVIDER_ID
67.. VISIT_OCCURRENCE_ID
123. CONDITION_STATUS_CONCEPT_ID
67.. CONDITION_SOURCE_CONCEPT_ID
Aa CONDITION_SOURCE_VALUE
Aa CONDITION_STATUS_SOURCE_VALUE
```

```
123. DRUG_EXPOSURE_ID
67.. PERSON_ID
123. DRUG_CONCEPT_ID
67.. DRUG_EXPOSURE_START_DATE
123. DRUG_EXPOSURE_START_DATETIME
67.. DRUG_EXPOSURE_END_DATE
123. DRUG_EXPOSURE_END_DATETIME
67.. DRUG_TYPE_CONCEPT_ID
Aa STOP_REASON
123. REFILLS
67.. QUANTITY
123. DAYS_SUPPLY
67.. SIG
123. ROUTE_CONCEPT_ID
67.. EFFECTIVE_DRUG_DOSE
```

Tables



Blacketer, C. 2020. *The book of OHDSI*. [Online]. [Accessed 21/01/2021].
Available from :<https://ohdsi.github.io/TheBookOfOhdsi/StandardizedVocabularies.html>

References

Observational Health Data Sciences and Informatics. 11/01/2021 *The book of OHDSI* [Online]. Accessed 20/01/2021. Available from: <https://ohdsi.github.io/TheBookOfOhdsi/>