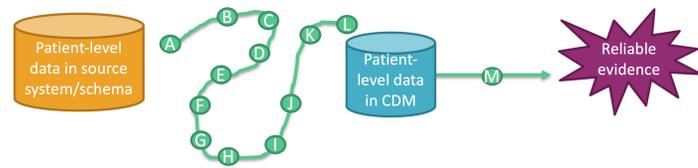


INTRODUCTION

The Observational Health Data Sciences and Informatics (OHDSI) initiative has proven its value for the generation of reliable and reproducible evidence to improve patient care. The adoption of the OMOP Common Data Model (OMOP-CDM) is expanding on a global level and is becoming a *de facto* standard for capturing health data in a standardised way.

The European Health Data and Evidence Project (EHDEN) has as aim to build an active and sustainable federated data network across Europe standardised to the OMOP-CDM. In EHDEN a large harmonization fund has been made available for data sources to be mapped to the OMOP-CDM. This process will be supported by Small to Medium Enterprises (SMEs) that will be trained and certified in the EHDEN project.



The community is growing strongly, and we have a challenge but also an obligation to train all the stakeholders. We need to assure they are trained on all the important steps in the journey from source data to reliable evidence.

The EHDEN project is investing heavily in building up a curriculum to train our stakeholders and we would like to invite you to contribute to building this exciting training curriculum.

THE EHDEN ACADEMY

We have been working on the EHDEN Academy, an online learning platform that will allow for self-paced training and blended training by incorporating it in onsite training activities.



The EHDEN Academy is implemented in the world's most popular open source learning platform called Moodle (www.moodle.com). This tool is used in many universities and companies around the world. It is a highly flexible platform to create a fully customized learning environment.

OMOP Common Vocabulary Model

What it is

- Standardized structure to house existing vocabularies used in the public domain
- Compiled standards from disparate public and private sources and some OMOP-grown concepts

What it's not

- Static dataset – the vocabulary updates regularly to keep up with the continual evolution of the sources
- Finished product – vocabulary maintenance and improvement is ongoing activity that requires community participation and support



We will incorporate all the currently available training material from the OHDSI Community, such as the CDM and Standardised Vocabularies Tutorial as shown above, and we will actively collaborate with OHDSI to extend the curriculum.

We will develop questions for quizzes and tests, but also for example SQL exercises that will require access to a CDM instance in the cloud.

UNLIMITED POSSIBILITIES

OMOP CDM and Standardised Vocabularies
Dashboard / My courses / CDM-Vocab

4 ENROLLED USERS 0 STUDENTS COMPLETED 3 IN PROGRESS 1 YET TO START

Welcome to the OMOP CDM and Standardised Vocabularies Course!

This course is based on the full day OMOP CDM and Standardised Vocabularies workshop offered by OHDSI at the Annual Symposium in 2018. This course will give an introduction into the history of Observational Medical Outcomes Partnership (OMOP) and the birth of Observational Health Data Sciences and Informatics (OHDSI). It highlights OHDSI's vision as well as how OHDSI is trying to drive the future of observational research through network studies. All of this introduction lays the importance for the rest of the course's material.

The course will then move on to understanding the OMOP Vocabularies; it is important to understand how the Vocabulary is used and work through examples prior to moving on to the OMOP Common Data Model (CDM). The final portion of the course will explain in detail the layout of the CDM as well as some key concepts such as era logic.

The course contains video lectures and exercises to test your knowledge. Furthermore, there are exercises you can perform on a Virtual Machine made available in the Amazon Community called "OHDSI-IN-A-BOX"

- Course Introduction
- Announcements
- CDM and Vocab Forum

The Foundations of OHDSI

- Intro Video
- Introduction to OHDSI and the OMOP-CDM
- Introduction Quiz

The Standardised Vocabularies

- Lecture 1
- Lecture 2
- Lecture 3
- Vocabulary Quiz

There are a number of implicit and explicit conventions that have been adopted in the CDM. Developers of methods that run methods against the CDM need to understand these conventions. The table below shows the most important conventions.

Field name	Purpose	Example
<entity_id>	Unique identifiers for entities (from numbers, or IDs imported from source)	person_id 1234567 visit_occurrence_id 7654321 could be a person identifier or an autogenerated number by the CDM builder
<entity_concept_id>	Foreign key into the Standard Vocabulary for Standard Concept	condition_concept_id 313217 (SNOMED "Atrial Fibrillation")
<entity_source_concept_id>	Foreign key into the Standard Vocabulary for Source Concept	condition_source_concept_id 4482197 (ICD9CM "Atrial Fibrillation")
<entity_source_value>	Verbatim information from the source data, not to be used by any standard analytics	condition_source_value 427.31 (ICD9CM "Atrial Fibrillation")
<entity_type_concept_id>	Foreign key into the Vocabulary for the origin of the information	condition_type_concept_id 38000199 ("inpatient header - primary")

Which of the following statements is true?

Select one or more:

- a. If we cannot map to a standard code we loose the record
- b. all fields ending with <_concept_id> refer to the VOCABULARY table
- c. an ICD-9 code should be placed in the condition_source_value field
- d. a entity_id value is unique for the domain not for the whole CDM

Check

A first step is that the source data will have to be stored in the CDM. This process is called Extraction Transform Load (ETL). Do you know what drives the choice of the domain to store the data?

Select one or more:

- a. the source data
- b. the concept domain in the Standardised Vocabulary
- c. your personal preference
- d. a human expert

Check

As shown above, each course will have a landing page that allows the trainee to keep track of the progress. A library of questions will be made to test the acquired knowledge of the trainee.

Question 1

Correct

Marked out of 1.00

Flag question

Edit question

We like to know how many different vocabularies are available in this CDM? Write a query to get that number.

Answer: (penalty regime: 10, 20, ..., %)

Reset answer

```
1 select count(*) from vocabulary
```

Check

Expected	Got
count(*)	count(*)
46	46

Passed all tests!

Next page

SQL Exercises against the OMOP-CDM

A powerful feature in Moodle is that we can create exercises in which the trainee is asked to write SQL queries against an actual OMOP-CDM instance. This is a fantastic approach to test the knowledge about the CDM and especially the use of the vocabularies.



The EHDEN Academy will also include tests that will leverage the OHDSI-IN-A-BOX virtual machines that will allow trainees to spin-up a virtual environment that hosts all the OHDSI tools including a CDM with simulated data.

CONCLUSION

As EHDEN, we believe that the investment in building an e-learning environment for EHDEN and OHDSI is a crucial next step to enable and further stimulate the adoption of our revolutionary approach to the analysis of health data and generation of reliable evidence. In the upcoming months we will take the next major step in enabling large-scale community training. We hope you are interested and if you would like to contribute to this effort, do not hesitate to reach out to us!