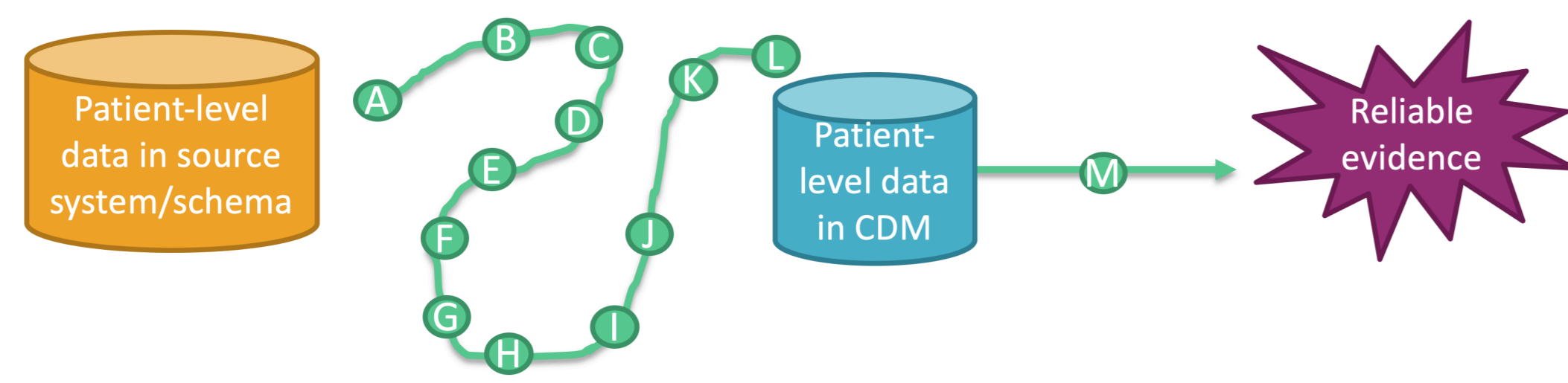


INTRODUCTION

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 EDHEN 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.



The screenshot shows the course landing page for 'OMOP CDM and Standardised Vocabularies'. It includes a video player with a presentation slide titled 'PERSON' listing requirements for creating a unique record per person. The slide content is as follows:

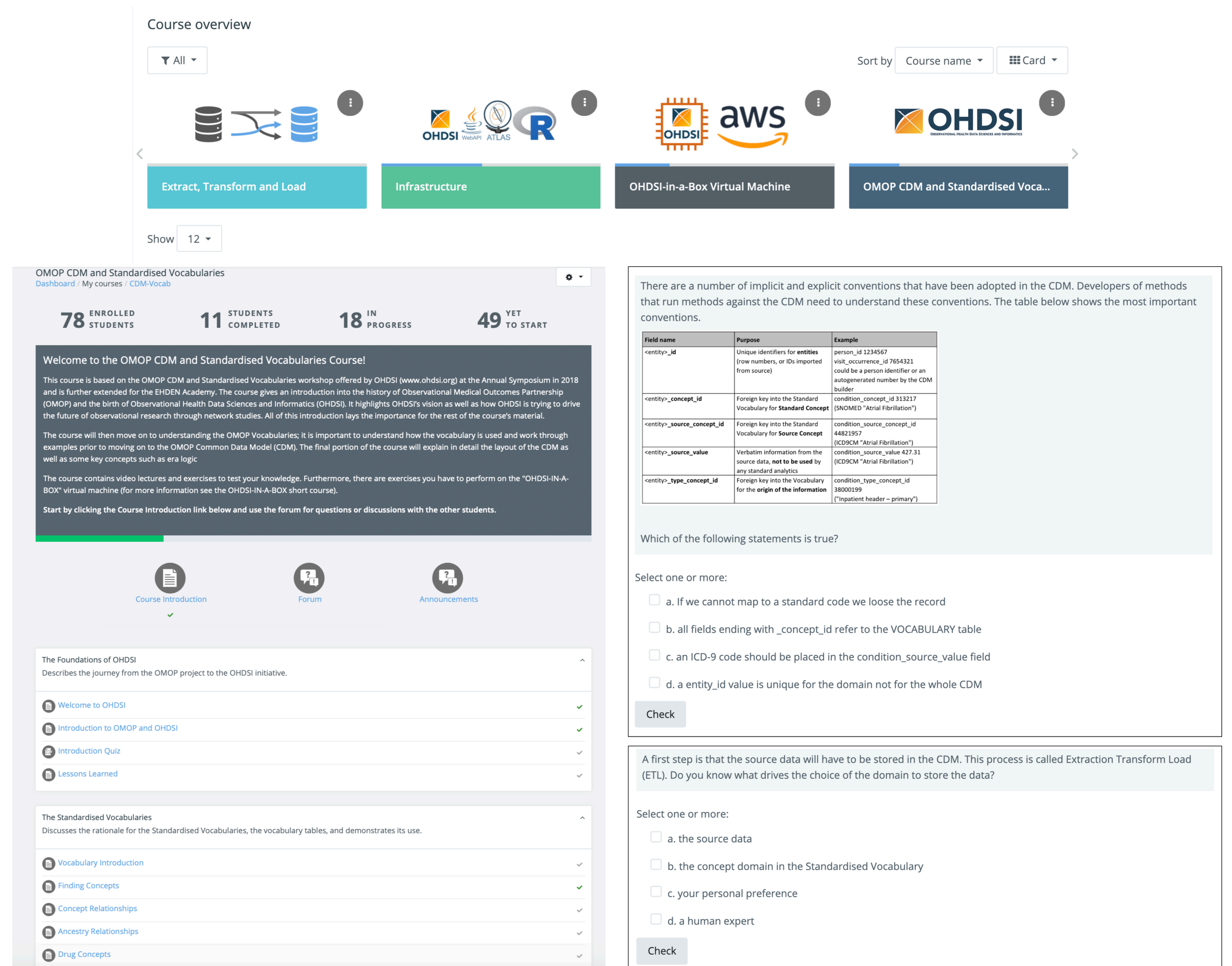
PERSON
<ul style="list-style-type: none">Need to create one unique record per personNo history of location/demographics: need to select latest availableYear of birth required...day/month optionalForeign key to the LOCATION, PROVIDER, and CARE_SITE table that contains one record

Below the video player, there are navigation links for 'Previous Activity' and 'Next Activity', and a 'Jump to...' search bar.



<https://academy.ehden.eu>

UNLIMITED POSSIBILITIES



The screenshot shows the OHDSI course overview and content. It includes a 'Course overview' section with a 'Show 12' dropdown. Below this, there are four main course components: 'Extract, Transform and Load', 'Infrastructure', 'OHDSI-in-a-Box Virtual Machine', and 'OMOP CDM and Standardised Vocabularies'. The 'OMOP CDM and Standardised Vocabularies' section is expanded, showing a welcome message and a list of course topics.

OMOP CDM and Standardised Vocabularies

78 ENROLLED STUDENTS | 11 STUDENTS COMPLETED | 18 IN PROGRESS | 49 YET TO START

Welcome to the OMOP CDM and Standardised Vocabularies Course!

This course is based on the OMOP CDM and Standardised Vocabularies workshop offered by OHDSI (www.ohdsi.org) at the Annual Symposium in 2018 and is further extended for the EHDEN Academy. The course gives 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 have to perform on the "OHDSI-IN-A-BOX" virtual machine (for more information see the OHDSI-IN-A-BOX short course).

Start by clicking the Course Introduction link below and use the forum for questions or discussions with the other students.

Course Introduction | **Forum** | **Announcements**

The Foundations of OHDSI
Describes the journey from the OMOP project to the OHDSI initiative.

- Welcome to OHDSI
- Introduction to OMOP and OHDSI
- Introduction Quiz
- Lessons Learned

The Standardised Vocabularies
Discusses the rationale for the Standardised Vocabularies, the vocabulary tables, and demonstrates its use.

- Vocabulary Introduction
- Finding Concepts
- Concept Relationships
- Ancestry Relationships
- Drug Concepts

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 (row 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 (OHMED "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 3800019 ("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

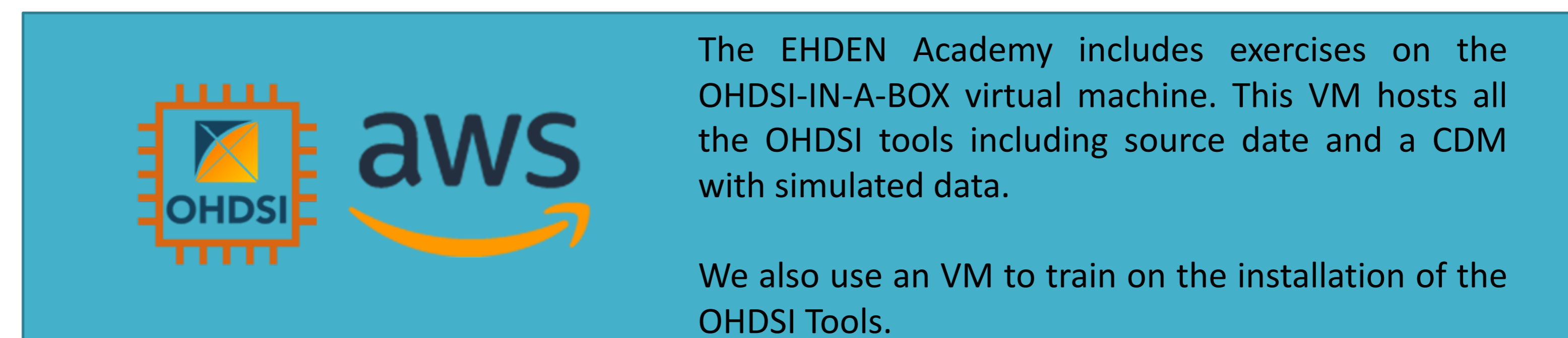
A first step is that the source data will have to be stored in the CDM. This process is called Extraction Transform Load (ETL). Do you know what drives the choice of the domain to store the data?

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. The courses contain quizzes, final exams and exercises to test the students.



The image shows the OHDSI logo (a stylized 'O' with a grid pattern) and the AWS logo (the Amazon logo). The text next to the logos reads: 'The EHDEN Academy includes exercises on the OHDSI-IN-A-BOX virtual machine. This VM hosts all the OHDSI tools including source data and a CDM with simulated data. We also use an VM to train on the installation of the OHDSI Tools.'

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. We will continue this development and will enable large-scale community training by opening the Academy also to the full OHDSI community. We hope you are interested and if you would like to contribute to this effort, do not hesitate to reach out to us!

