Mapping Custom Source Codes to Standard Concepts: A Comparison of Two Approaches

Melanie Philofsky, RN, MS

INTRODUCTION:

There are two methods to map custom source codes to standard concepts within the OMOP CDM. Only the Source to Concept Map (STCM) method is described in the Book of OHDSI. Both methods and their strengths and weaknesses will be presented here.

WHY IS IT IMPORTANT?

Source codes in a dataset may be custom codes not supported by OHDSI, necessitating the data holder to create a mapping from the source code to a standard concept in order to utilize the data in network research and with the OHDSI tools.

METHODS

There are two methods to map custom source codes:

1. The STCM method. Insert mappings into the Source to Concept Map table.
2. Create a custom source concept in the Concept table and the appropriate relationships in the Concept Relationship (CR) table.

METHOD #1:

Source_to_Concept_Map

Example source code = Pediatric interventional cardiologist

Field	Source_code	Source_concept_id	Source_vocabulary_id	Source_code_description	Target_concept_id	Target_vocabulary_id	Valid_start_date	Valid_end_date	Invalid_reason
Required Field?	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	No
Value	Pediatric interventional cardiologist	0	prov_specialty		903276	Medicare Specialty	01/01/1970	12/31/2099	

Figure 1. A local source code is mapped to a standard concept_id in the Source to Concept Map table

METHOD #2:

Concept

Example source code = Pediatric interventional cardiologist

Field	Concept_id	Concept_name	Domain_id	Vocabulary_id	Concept_class_id	Standard_concept	Concept_code	Valid_start_date	Valid_end_date	Invalid_reason
Required Field?	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	No
Value	2100000000	Pediatric interventional cardiologist	Provider	prov_specialty	Physican Specialty		Pediatric interventional cardiologist	01/01/1970	12/31/2099	

Figure 2. A custom concept is created for the local source code, ‘Pediatric interventional cardiologist’

Concept Relationship

Field	Concept_id_1	Concept_id_2	Relationship_id	Valid_start_date	Valid_end_date	Invalid_reason
Required Field?	Yes	Yes	Yes	Yes	Yes	No
Record #1 Value	2100000000	903276	Maps to	01/01/1970	12/31/2099	
Record #2 Value	903276	2100000000	Mapped from	01/01/1970	12/31/2099	

Figure 3. The custom concept created for Pediatric interventional cardiologist is mapped to a standard concept

Comparison of the two methods

	STCM	Concept & CR
Creation	Create a mapping from the source code to the standard concept_id in the STCM table.	Create a non-standard concept_id AND bi-directional mappings between the source concept_id and the standard concept_id.
Loading	Less complex since the table is designed for insertion of records.	Requires alteration of the Vocabulary tables provided by Athena.
ETL logic	Requires additional logic to lookup source codes in the STCM table when the source code isn’t found in the Concept table.	Utilizes the standard lookup for a concept in the Concept table.
Research use	**None**; This data is not visible in Atlas or the Concept table. Network researchers are unable to assess the quality of the mapping.	Researchers are able to view the local source code mapped to a standard concept_id in Atlas. Allows for the inclusion/exclusion of data
Summary	Efficient and effective. However, the original source code is not visible to CDM users.	The increased complexity for the mapping and loading processes are well worth the tradeoff .

Conclusion

While both methods used to custom map non-OHDSI supported source codes have advantages and disadvantages, only the creation of concepts and concept relationships allow the source code and it’s mapping to a standard concept to be visible in Atlas.