Transformation of an Israeli medical centers network to OMOP CDM

Nadav Rappoport, PhD[1,2], Nir Makover, MSc[2], Guy Livne, MSc[2], Binyamin Abu, MD[3], Sivan Ran[2], Robyn Robin[2], Gil Frenkel[2], Tomer Hadad[2], Naama Perry-Cohen[2]
[1]**Ben-Gurion University of the Negev, Beer Sheva, ISRAEL;** [2] **The Directorate of government medical centers, the Israeli Ministry of Health, Jerusalem, ISRAEL;** [3]**Hillel Yaffe Medical Center, Hadera, ISRAEL**

Abstract

The Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) v6.0 became de-facto one of the leading standards for clinical data CDM which enable analysis to be done behind firewall. The Directorate of government medical centers (DGMC) in the Israeli Ministry of Health own and manages over 20 medical centers including general, geriatric and psychiatric hospitals, geriatrics had decided on June 2019 to make its clinical data accessible for academic research, pharma companies, biotech and high-tech companies. DGMC chose the OMOP CDM as its CDM mainly because it's international interest and its active community.

As a pilot, the first medical center to be transformed was Hillel Yaffe Medical Center. The data is stored and managed in several systems. We designed and developed a process for ETL from these systems. In total, the pilot covered 5 years of data from 2015 to 2019. 805,000 patients, about 5,000,000 condition occurrences, 5,000 providers, 6,000,000 visits, 14,000,000 visit details records, 93,000 procedures and 2,000,000 observations. This data is now being used for retrospectives studies.

Research Category

Observational data standards and management

Introduction

The Directorate of government medical centers in the Israeli Ministry of Health is the largest medical centers network in Israel. DGMC is the owner and manager of 24 medical centers with over 13,000 hospitalization beds which is the majority of hospitalization beds in the country. Each medical center has its own authority, though resource management and allocation are centralized by DGMC. DGMC is regularly doing extensive business intelligence analysis for quality insurance, quality control and better business decision making. A strategic decision was made on June 2019 to make clinical data accessible for researchers from academia and from the medical centers as well as for pharma companies, biotech companies and high-tech companies. The vision is that it will accelerate development of new diagnostic and prognostic tools, and improve patient health and healthcare.

Multiple alternatives were considered, and an expert committee was formed. The conclusion was to use the OMOP CDM[1]. The main reasons for choosing OMOP CDM v6.0 over other alternatives are: (i) worldwide usage, (ii) Active community and forum (iii) Open source.

ETL design

We designed a process that consists of several component. (i) Concept mapping and standardizing (ii) Data mapping.

Concept mapping and standardizing

Concepts are given in the source are mostly not in standard vocabularies as defined by OMOP CDM. Some catalogues are in a common vocabulary like procedures and diagnoses which are stored in a catalogue named IC. This catalogue is an extension of the ICD-9 CM catalogue. Therefore, the ICD-9 CM concepts can be mapped to SNOMED, which is the standard vocabulary for diagnoses and procedures using concept relationships maintained by OHDSI. Other concepts need to be mapped manually directly to SNOMED. We made an extensive use of the USAGI tool[2]. For other internal catalogs in which there was no mapping to common catalog, we used a mapping that was made in a separate effort of standardizing catalogs in the network. This mapping mapped laboratory tests to the LOINC catalog which is the standard OMOP

once. But, mapped drugs to ATC catalog. As RxNorm is the standard catalog for drugs, ATC-5 was mapped using concept relationship. This is a temporary solution, as intermediate mapping to ATC-5 catalog cause a loss of information. For example, the drug *CALCIUM CARBONATE 600MG TAB* is mapped to the ATC-5 concept *A02AC01* (*calcium carbonate*) which in turn, mapped to RxNorm code 1897 (calcium carbonate). This mapping loses the route and the dosage, where a better mapping would map directly to RxNorm code 315519 (calcium carbonate 600 MG). We plan in a second phase to mapped all drugs directly to RxNorm.

Data mapping of a Network

Clinical data is not routinely shared across medical centers in the network. Therefore, the same id (visit, procedure, prescription etc.) can potentiality appear in different systems or even in the same source system, but in different medical centers represent completely different entities. Therefore, in order to unify and centralize ids we have generated global ids. These global ids are deterministic, and for each domain we have a different algorithm for how to generate the global id. In short, the global id is based on the medical center, the source system and the domain it represents.

Figure 1. Israel map with list and locations of 24 medical centers in the Israeli Ministry of Health network. 11 general medical centered encountered with light blue, 5 geriatric medical centers encountered with light green, and 8 psychiatric medical centers encountered with deep blue.

Conclusion

The Israeli Ministry of Health own the largest medical centers network in Israel in terms of hospitalization beds. There is an ongoing effort, to standardize the clinical data of 24 medical centers using the OMOP CDM. Based on the pilot in Hillel Yaffe Medical center as the first medical center in the network, we anticipate studies and collaborations to come.

References

1. Stang, Paul E., et al. "Advancing the science for active surveillance: rationale and design for the Observational Medical Outcomes Partnership." Annals of internal medicine 153.9 (2010): 600-606.
2. https://github.com/OHDSI/Usagi