Real-World Data in Greece: Mapping the Papageorgiou General Hospital Data to the OMOP Common Data Model

Grigoris PAPAPOSTOLOU[a], Achilleas CHYTAS[a], Alexandros REKKAS[a],
Maria BIGAKI[b], Demetrios ZEIMPEKIS[b], Lampros DERMENTZOGLOU[b],
George TORTOPIDIS[b] and Pantelis NATSIAVAS[a,1]

[a] *Centre for Research and Technology-Hellas, Institute of Applied Biosciences, Thessaloniki, Greece*
[b] *Papageorgiou General Hospital, Thessaloniki, Greece*
ORCiD ID: Grigoris Papapostolou https://orcid.org/0009-0005-1247-980X,
Achilleas Chytas https://orcid.org/0000-0001-8486-011X,
Alexandros Rekkas https://orcid.org/0000-0002-5352-943X,
Pantelis Natsiavas https://orcid.org/0000-0002-4061-9815

Abstract. This paper showcases the results of the Extract-Transform-Load process mapping the Electronic Health Record of Papageorgiou General Hospital in Thessaloniki, Greece, to the Observational Medical Outcomes Partnership Common Data Model. We describe the staged process utilized to account for the intricate structure of the database, along with some general findings from the mapping. Finally, we investigate potential directions for future research.

Keywords. OMOP-CDM, ETL, mapping, observational data, real world data

1. Introduction

The Observational Health Data Science and Informatics initiative aims to standardize the structure and coding system of databases to the Observational Medical Outcomes Partnership Common Data Model (OMOP-CDM), promoting research across disparate observational databases in both a centralized environment and a distributed research network (1). We aimed to map the electronic health record (EHR) of Papageorgiou General Hospital (PGH) in Thessaloniki, Greece, to OMOP-CDM.

2. Methods

The eHealth Lab of the Institute of Applied Biosciences at the Centre for Research and Technology Hellas in collaboration with the PGH team planned the conversion process. The latter extracted and securely shared the data.

Due to the structure of the database we used a three-schema staged approach. An **input schema** housed the data tables extracted from various sources (SAP and LIS). A

staging schema enabled the refinement of the raw data. Finally, a **target schema** accommodated data integration to OMOP-CDM.

Non-adherence to international standards necessitated vocabulary mapping, implemented as an automated translation of Greek terms followed by manual verification. We mapped PGH database terms to OMOP concepts using the USAGI software library (2). The mapping was completed in collaboration with a physician.

3. Results

The PGH database provides information on more than 1.4 million patients living in and around Thessaloniki, with all ages being adequately represented. Among the total number of records, 64.55% fell in the domain of measurements, 10.68% in the domain of notes, 8.72% in the domain of visits, 8.06% in the domain of drug exposures, 5.71% in the domain of procedures, and the rest in the domain of death.

Most prevalent measurements concerned blood test results, all measured for more than 500,000 patients. The most prevalent conditions were angina pectoris, pain of truncal structure, fever, and cerebral infarction, all recorded for more than 10,000 patients. The most prevalent procedures were standard chest X-ray, physical examination, injection of contrast media for radiography, and electrocadiographic procedure, all performed for more than 10,000 individuals. Finally, the vast majority of visit records concerned outpatient, followed by inpatient and observation room visits.

4. Discussion

Similar to prior findings, harmonization of nation-specific terminology was challenging (3). This was resolved with a semi-automated approach, combining mapping software with expert opinion. In addition, the irregular database structure necessitated the use of a custom staged approach. With the conversion of the PGH database to OMOP-CDM, a large hospital in the Balkan region became part of an extensive network of databases. PGH also gained access to standardized tools, enabling improved healthcare services.

5. Conclusion

The successful mapping of the PGH database to OMOP-CDM translates to a huge research potential and warrants great potential for future international network studies.

References

[1] OHDSI. The Book of OHDSI: Observational Health Data Sciences and Informatics [Internet]. OHDSI; 2019. Available from: https://books.google.gr/books?id=JxpnzQEACAAJ
[2] Usagi [Internet]. [cited 2024 May 9]. Available from: https://ohdsi.github.io/Usagi/
[3] Henke E, Zoch M, Peng Y, Reinecke I, Sedlmayr M, Bathelt F. Conceptual design of a generic data harmonization process for OMOP common data model. BMC Med Inform Decis Mak. 2024 Feb 26;24(1):58. doi:https://doi.org/10.1186/s12911-024-02458-7