

ADAPTION OF THE OMOP CDM FOR RHEUMATOLOGY: A Portuguese experience.

Harmonization of Biobanco-iMM Rheumatology Collection data to OMOP CDM

Background: The OMOP Common Data Model (OMOP CDM) is an option to store patient data and to use these in an international context.

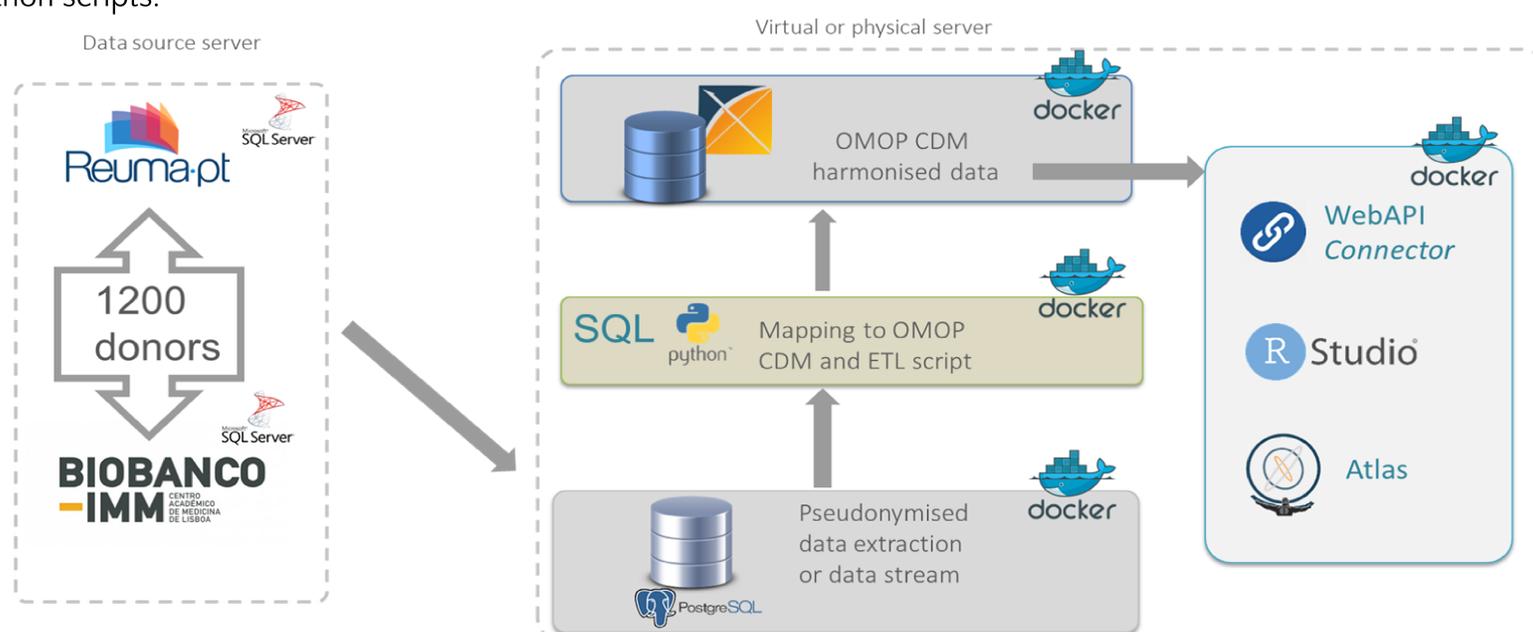
Biobanco-iMM includes biological samples (from surgery, biopsies, blood samples) which are voluntarily donated with permission for preservation and future use in biomedical research. Biobanco-iMM has a collection for rheumatology that is connected with Reuma.pt – the Rheumatic Diseases Portuguese Register from Portuguese Society of Rheumatology (SPR).

Reuma.pt protocols include structured information about socio-demographic data and information about standard diagnostic criteria and clinical features of each disease.

This project aimed to convert the Biobanco-iMM Rheumatology Collection data to the OMOP Common Data Model (CDM) version 5.4, using data from Biobanco-iMM and Reuma.pt databases.

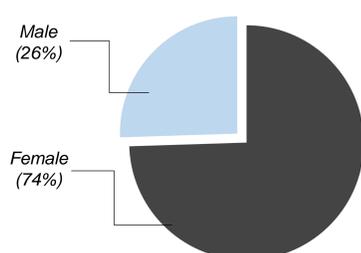
Methods

Source data was extracted from the Biobanco-iMM and Reuma.pt databases to be loaded onto a PostgreSQL DBMS instance (postgresql 10+190ubuntu0.1), where CDM was also implemented. ETL was implemented through SQL procedures and orchestrated through Python scripts.

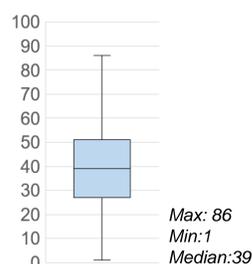


Results

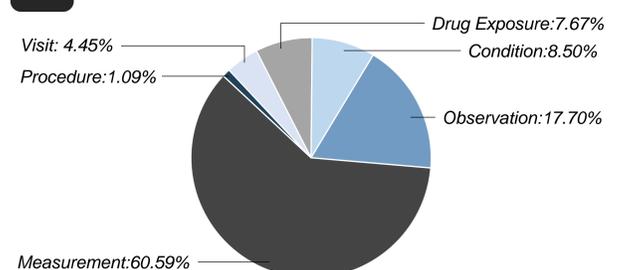
1 Data distribution *per gender*



2 Data distribution *per age*



3 Data distribution *per concept*



Conclusion: Data harmonization, besides to contribute to the data quality improvement, also facilitates the development of relevant clinical projects. In addition, the implementation of country nodes strongly improves its outcome. In the future, CDM materialization will be manually updated based on the needs expressed by Biobanco-iMM and Rheumatology Service.

Limitation: A few Biobanco-iMM Rheumatology Collection data was not standardized due to the lack of appropriate standard vocabulary.

