

Background

- The Information System for Research in Primary Care (SIDIAP; www.sidiap.org) contains data of anonymized patient's healthcare records for approximately seven million people (80% of the Catalan population) registered at approximately 287 primary care practices throughout Catalonia (Figure 1) since 2006.
- It includes **data collected by health professionals** during routine visits in **primary care**, such as clinical diagnoses and measurements (weight, blood pressure, etc.), laboratory tests, treatments, referrals, demographic and lifestyle information **and discharge records** from the 8 biggest hospitals of Catalonia. SIDIAP can be linked to other sources of data such as hospital discharge from other providers, population registries, etc.
- **Mapping of SIDIAP to OMOP Common Data Model (CDM)** has been recently performed as part of the European Medical Information Framework (EMIF) project.
- However, the overall quality of the mapping needs to be evaluated in order to assess the readiness of the mapped SIDIAP to perform research projects.

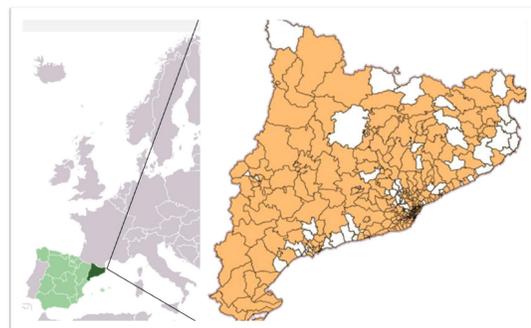


Figure 1. SIDIAP population coverage in Catalonia, Spain.

Objective

- Deploy the stack of OHDSI tools to test and analyze SIDIAP data conformance to the OMOP-CDM.
- Perform quality control of the CDM implementation in SIDIAP through the use of the provided reports.

Methods

- SIDIAP data from 2006 to 2017 was mapped to the OMOP-CDM v5.0. **WhiteRabbit** was used to profile source data, **Rabbit-in-a-Hat** helped design the ETL mapping logic and code mappings were reviewed with **Usagi**.
- The **OHDSI tools** were deployed by creating a Docker multi-container application (Figure 2).
- ETL was developed with SQL running over a MariaDB database where SIDIAP source data rested.

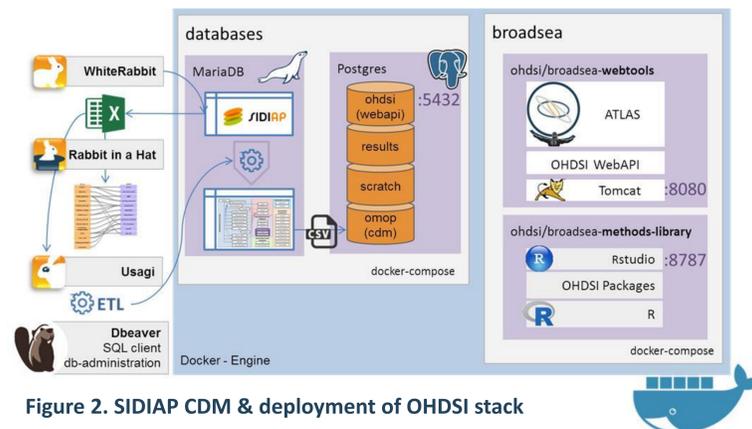


Figure 2. SIDIAP CDM & deployment of OHDSI stack

- CSV files of CDM transformed tables (person, observation_period, death, condition_occurrence, procedure_occurrence) were captured in a PostgreSQL schema.
- Condition_era table was derived from condition_occurrence with constructEras function of cohortMethods package.
- **Achilles Heel Results Viewer** quality checks and their printed queries helped to iteratively review and address the generated errors and warnings.

Results

- **+24,700 terms and +1,400 million registers** were mapped to standard codes on clinical domains: person, observation_period, death, condition_occurrence, procedure_occurrence, drug_exposure, and measurement.
- The ATLAS Reports for the person domain showed the SIDIAP population includes **7,5 million individuals** (Figure 3).
- **3,8 million followed for ≥ 10 years**. These findings are in line with those found in the SIDIAP source data.

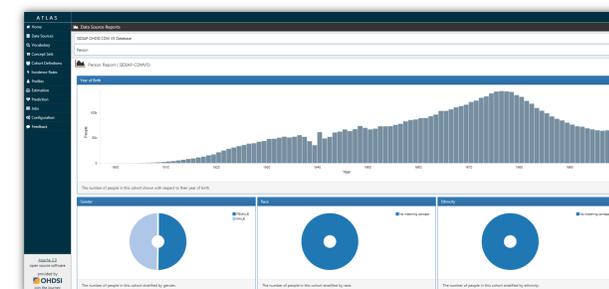


Figure 3. Person dashboard from Atlas



Figure 4. Death dashboard from Atlas

- **Death prevalence** increased with age,
- higher in men as expected and in accordance with what is observed in the SIDIAP source data (Figure 4).

- 10,057 target conditions.
- **Infectious diseases and hypertension** were most common (Table 1).

Table 1. Top five conditions in SIDIAP OMOP-CDM

Concept Id	Name	Person Count	Prevalence	Records per person
260427	Common cold	2,742,567	36.22%	2.08
198062	Infectious colitis, enteritis and gastroenteritis	1,526,959	20.17%	1.50
320128	Essential hypertension	1,511,828	19.97%	1.38
194133	Low back pain	1,481,098	19.56%	1.28
25297	Acute pharyngitis	1,422,280	18.78%	1.46

- 3,179 target **procedures** terms from the hospital were identified, representing a median of **3 procedure per person** with top procedures:
 - **injections** (either of antibiotics or of other therapeutic or prophylactic substances),
 - **imaging** (heart, head or abdomen),
 - **cataracts** related (either removal and/or replacement by intraocular lens).

Conclusions

- The OHDSI tools were successfully implemented in SIDIAP and were useful for quality control of the SIDIAP mapping to the OMOP-CDM.
- Further steps, including the quality control of other mapped and transformed domains and the replication of a scientific study, are needed in order to insure the quality of SIDIAP mapping.
- Other implementers could benefit from a fully tested and reproducible Docker container deployment strategy.