

Background

The Main Association of Austrian Social Security Institutions (HVB) provides health claims data for the ongoing research project “Adverse drug events in relation to inappropriate medication of geriatric patients with renal insufficiency - a retrospective register-based cohort study” (ADE-PIM).

Cohort:

- 11,547 patients from one Austrian province
- between the years 2008 and 2011
- Inclusion criteria: renal insufficiency

Data:

- medication data from pharmacies
- overnight hospital stays
- hospital diagnoses

To evaluate the feasibility of the OMOP common data model for Austrian health claims data, the source data is transformed into the CDM. The transformed data is tested for equivalence to the source. Selected statistical calculations from the ADE-PIM project are applied to the transformed data and the statistical results are compared.

Methods

Preparatory work

1. Import of the source data into a dedicated PostgreSQL database schema
2. Installation and set up of an OMOP CDM v6.0 instance into a dedicated schema on the database
3. Import of the necessary OMOP vocabularies (e.g. Gender, SNOMED, ATC, ICD10, ICD10CM)

Vocabulary mapping

1. Comparison of local vocabularies to OMOP vocabularies
2. Extension of Vocabulary, Concept and Concept_Relationship tables with local terms
3. Definition of rules for mapping local terms to concepts in the CDM

Extract-Transform-Load (ETL) process

1. Selection of data to be transformed
2. Definition of rules for mapping local tables/data fields to tables/data fields in the CDM
3. Development of SQL scripts to handle the ETL process according to prior rules

Assessment

1. Development of Python/SQL scripts to test the transformed data for integrity and equivalence
2. Adaption of statistical calculations from the ADE-PIM project to work with the transformed data
3. Comparison of the results to the originals



Results

Local vocabularies have been mapped to OMOP vocabularies as depicted in Table 1. Four local ATC codes do not match the OMOP ATC vocabulary. 60 codes from the Austrian ICD10 modification occurring in the data are not included in the OMOP ICD10 vocabulary. 31 of these are part of OMOPs ICD10CM vocabulary.

Table 1: Vocabulary mapping from source to OMOP

Source Domain	Source Vocabulary	OMOP Vocabulary	OMOP Domain
Medication	PZN	ATC	Drug
Diagnoses	ICD10 BMSG	ICD10, ICD10 CM	Condition, Procedure, Measurement, Observation
Specialty	GAPDRG2 FGB	ABMS, Medicare Specialty	Provider Specialty

PZN: Pharmazentralnummern, **ICD10 CM:** ICD10 Clinical Modification, **ICD10 BMSG:** ICD10 Austrian Modification, **GAP-DRG2 FGB:** Local list of provider specialties, **ABMS:** American Board of Medical Specialties

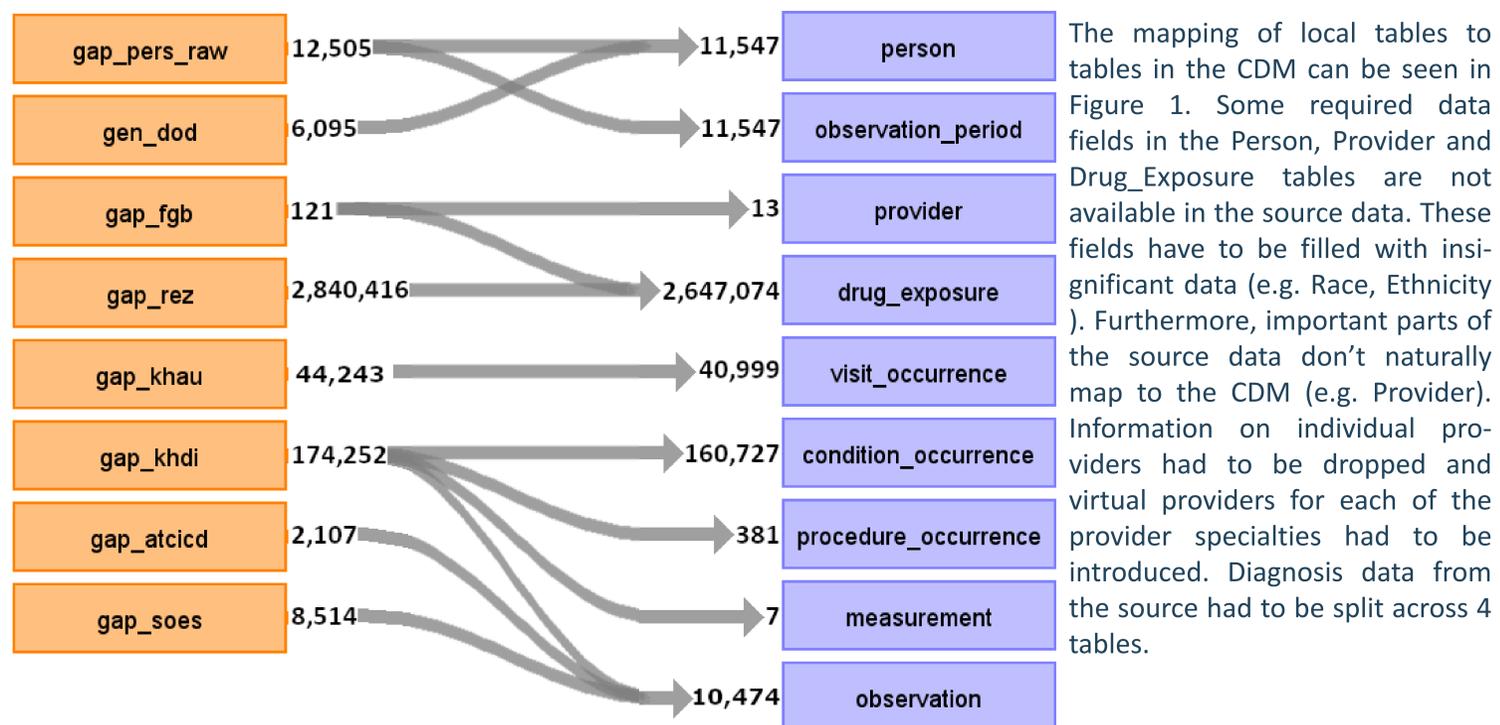


Figure 1: Table mapping from source to OMOP with row counts

Conclusions

Not all of the local vocabularies used within the data provided by the HVB are available in OMOP. Additionally, even if a vocabulary is available, local versions differ slightly. There is no one-to-one relationship between the provided data model and the OMOP CDM. Nevertheless, a strategy was found to preserve most of the significant information for the ADE-PIM project in OMOP.