

Medication dosage and exposure duration in OMOP CDM: mapping challenges

Tatiana Banokina, Dmitry Dymshyts, Alexandra Orlova, Alexander Kraynov, Alexander Davydov

Background:

There is a growing demand in transformation of medical data to Observational Medical Outcome Partnership (OMOP) Common Data Model (CDM). One of the main ETL challenges is to transform native data to the Drug domain retaining the exact drug dosages administered or prescribed. CDM specifications have been substantially changed in the last couple of years in terms of drug dosage and exposure length representation.

Until CDM v5.2.0, the *drug_exposure* table contained *effective_drug_dose*, *dose_unit_concept_id* and *quantity* fields. Starting from CDM v.5.2.0¹ the *drug_exposure* table contains the only *quantity* field designed for storing the amount of drug. In the latest CDM documentation², it is stated that “the *quantity* should be converted to the correct unit given in the *drug_strength* table”. Despite the CDM documentation² provides examples on this topic, there is still considerable ambiguity between *quantity* and drug dose calculations³.

Moreover, since OMOP CDM v5.2.0 *drug_exposure_end_date* field is required. It’s known that the end date of a drug intake is not always available in the source data, so OMOP CDM documentation provides the methods to infer the *drug_exposure_end_date*²: calculation based on the days supply, total dosage/daily dosage proportion, and default values (1 day for administration records, 29 days for written prescriptions and 89 days for mail-order prescriptions). Default value rule does not tend to be precise enough and there’s a need for development of alternative solutions on the data sources lacking days supply and daily dosage information.

Currently most of OHDSI studies on drugs are focused on the fact of taking medication. However, there is an increasing demand for studies which would consider therapeutic regimen, drug dosage and duration/frequency of a drug intake. It is critical to derive the *quantity* and duration of exposure. Therefore, in our work we would like to share the best practices on:

- Calculation of *quantity* comprehensively considering the source information, target Standard concept of the *drug_exposure* record and associated dosage data from the *drug_strength* table;
- Imputation of *drug_exposure_end_date* using daily dose and derived daily dose (DDD) from Anatomical Therapeutic Chemical Classification System (ATC) vocabulary.

Methods and Results:

We developed an automated approach for the *drug_exposure.quantity* calculation by taking into account the respective dosage data from the *drug_strength* table of a target concept and the way the source unit

is matched to the units in the *drug_strength* table fields. Below is the list of cases and formulas used respectively.

Table 1. Quantity calculation rules

#	Drug_strength filters	Source units can be translated to <units> from the drug_strength table	Drug_exposure.quantity formula	Related Drug classes
1	amount_value IS NULL	units of amount	source quantity * unit multiplier to amount units	Ingredient, Clinical Drug Form, Branded Drug Form
2	amount_value IS NOT NULL	units of amount	source_quantity * unit multiplier to amount units / drug_strength.amount_value	Clinical Drug, Branded Drug
3	denominator_value IS NULL and numerator_value IS NOT NULL	units of denominator	source quantity * unit multiplier to denominator units	Clinical Drug, Branded Drug, Clinical Drug Comp, Branded Drug Comp
4		units of numerator	source quantity * unit multiplier to numerator units / drug_strength.numerator_value	
5	denominator_value IS NOT NULL and numerator_value IS NOT NULL	units of denominator	source quantity * unit multiplier to denominator units / drug_strength.denominator_value	Quant Branded Drug, Quant Clinical Drug
6		units of numerator	source quantity * unit multiplier to numerator units / drug_strength.numerator_value	

Drug_exposure_end_date imputation.

If the daily dose is not available in the source data, the most frequent dose can be used: for each source drug concept or source/target drug concept combination, the most frequent dose is defined and then it is applied to those records where the dose is missing.

If daily dose is not available at all, ATC DDD (defined daily dose) can be used as the assumed average maintenance dose per day for a drug used for its main indication in adults⁴. This method was discussed on the OHDSI forum⁵ and tested on oral solid drugs. Method plausibility was assessed using another plausible guesstimate: calculation of end date based on the following prescription and assumption that most common durations of taking the drug should be 7/30/60/90 days. In addition, we reviewed results for 200 most common drugs in the source and made a conclusion that for most of cases ATC DDD method is suitable (for example, source drug is 'Amlodipine 5 MG Oral Tablet' and total quantity = 28 tablets, ATC DDD = 5 mg => calculated duration = 28 days). However, there are some limitations for this method. Some drugs are indicated in different dosages for different therapeutic purposes, e.g. aspirin is used in dosage of 3 g/day as analgesic/antipyretic and in dosage of 1 tablet per day (independent of strength) as antithrombotic agent.

Conclusion:

In this work we covered the principal use cases of calculating the *drug_exposure_quantity* values using the *drug_strength* table and imputation of *drug_exposure_end_date* based on a daily dose. Population of the *quantity* values in the *drug_exposure* table in accordance with *drug_strength* data and source units and imputing the *drug_exposure_end_date* are important steps that assure correct calculations in the *drug_era* and *dose_era* tables and make OMOP drug dosage studies possible.

The unified automated scripts that calculate *quantity* based on the *drug_strength* fields values and source unit information can be a performing solution, but more research is required to solve the limitations of this approach (conflicting source and target units, boxed medications, combined medications). One of the possible steps to improve the drug dosage conversion and representation in OMOP is drug signature standardization.

The approach of imputing *drug_exposure_end_date* using ATC DDD can be considered to be added to the CDM documentation. ATC DDD data need to be integrated into the Standardized OMOP vocabularies.

References:

1. Release CDM v.5.2.0 OHDSI/ Common Data Model.
<https://github.com/OHDSI/CommonDataModel/releases/tag/v5.2.0>
2. <https://ohdsi.github.io/CommonDataModel/cdm531.html>
3. OMOP Common Data Model. How to Calculate Drug Dose:
https://ohdsi.github.io/CommonDataModel/drug_dose.html
4. ATC DDD. <https://www.who.int/tools/atc-ddd-toolkit/about-ddd>
5. ATC DDD to impute drug_exposure_end_date. OHDSI Forum.
<https://forums.ohdsi.org/t/atc-ddd-to-impute-drug-exposure-end-date/12233>