Medication dosage and exposure duration in OMOP CDM: mapping challenges

Tatiana Banokina[1], Dmitry Dymshyts[1], Alexandra Orlova[1], Alexander Kraynov[1], Alexander Davydov[1], Eugene Paulenkovich[1]
[1]Odysseus Data Services, Inc

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. Starting from CDM v.5.2.0[1] the drug_exposure table contains the only quantity field designed for storing the amount of drug. In the latest CDM documentation[2], it is stated that "the quantity should be converted to the correct unit given in the drug_strength table". Despite the CDM documentation[2] provides examples on this topic, there is still considerable ambiguity between quantity and drug dose calculations[3].
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.

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

Conclusions

In this work we addressed 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

Methods: Calculation of quantity value

We developed an automated approach for the drug_exposure.quantity calculation by taking into account the source information, respective dosage data from the drug_strength table of a target concept and the way the source units match the units in the drug_strength table fields.

Use the formula as shown in the Table 1. Which formula to choose depends on the source drug representation, drug_strength composition of the target concept and the match between the two.

Required steps:
1. Perform mapping of the source drug description to the standard drug_concept_id
2. Create LK_UNIT table that converts source unit to the standard quantity unit(s) of the DRUG_STRENGTH fields (amount, numerator and denominator units).

Table 1. Calculation of quantity value

Source table			Concept mapping			DRUG_STRENGTH						LK_UNIT			Calculation of quantity		
source_drug_name	source_ quantity	source_quantity _unit	drug_concept_id	drug_concept_name	drug_concept_ class	amount _value	amount _unit	numerator _value	numerator _unit	denominator _value	denominator _unit	unit_concept_id	unit_concept _name	multiplier	formula	calculation	drug_exposur e. quantity
ACETAMINOPHEN	0.5	G	1125315	acetaminophen	Ingredient		8576 (milligram)					8576	milligram	1000	source quantity * unit multiplier to amount units	0.5*1000	500
ACETAMINOPHEN 250 MG TABLETS	0.5	G	19107242	acetaminophen 250 MG Oral Tablet	Clinical Drug	250	8576 (milligram)					8576	milligram	1000	source_quantity * unit multiplier to amount units/ drug_strength.amount_value	0.5*1000/250	2
FUROSEMIDE 10 MG/ML INJECTION	4	ML	35603225	furosemide 10 MG/ML Injection	Clinical Drug			10	8576 (milligram)		8587 (milliliter)	8587	milliliter	1	source quantity * unit multiplier to denominator units	4*1	4
FUROSEMIDE 10 MG/ML INJECTION	0.04	G	35603225	furosemide 10 MG/ML Injection	Clinical Drug			10	8576 (milligram)		8587 (milliliter)	8576	milligram	1000	source quantity * unit multiplier to numerator units/ drug_strength.numerator_value	0.04*1000/10	4
ROFECOXIB, VIOXX - SUSPENSION 12.5MG/5ML 150ML	150	ML	21108607	150 ML rofecoxib 2.5 MG/ML Oral Suspension [Vioxx]	Quant Branded Drug			375	8576 (milligram)	150	8587 (milliliter)	8587	milliliter	1	source quantity * unit multiplier to denominator units/ drug_strength.denominator_value	150*1/150	1
ROFECOXIB, VIOXX - SUSPENSION 12.5MG/5ML 150ML	0.375	GRAM	21108607	150 ML rofecoxib 2.5 MG/ML Oral Suspension [Vioxx]	Quant Branded Drug			375	8576 (milligram)	150	8587 (milliliter)	8576	milligram	1000	source quantity * unit multiplier to numerator units/ drug_strength.numerator_value	0.375*1000/37 5	1

Methods: Imputation of drug_exposure_end_date

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 adults. This method was discussed on the OHDSI forum[5] 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.

However, there are some limitations for ATC DDD 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.
- If ATC DDD is less than the dosage of one entity (tablet, capsule, etc.) it is unlikely that partial entities were administered (see the last three rows in Table 2).
- The method was tested for oral solid drug forms only, while liquid, inhalation forms requires much more extensive logic and calculations.
- Suitable only for adult dosages (patients at least 16 years old or older).

Table 2. ATC DDD approach implementation

Source table			Mapping table		ATC table				DRUG EXPOSURE, standard approach		DRUG EXPOSURE, suggested approach		
source_drug_name	quantity	drug_exposu re_start_dat e	drug_concept _id	drug_concept_name	atc_concept_code	atc_concept_name	ATC DDD	ATC DDD, unit	days_supply	drug_exposure _end_date	drug_exposure_end_date calculating formulas	days_supply	drug_exposure _end_date
IBUPROFEN, ADVIL – ORAL TABS 400 MG	30	01/01/2002	19019072	ibuprofen 400 MG Oral Tablet	M01AE01	ibuprofen; systemic, rectal	1200	mg	30	01/30/2002	total_quantity = quantity * dosage(drug_strength.amount_value) days_supply = total_quantity / ATC DDD drug_exposure_end_date = drug_exposure_start_date + days_supply - 1 day	10.0	01/10/2002
IBUPROFEN, ADVIL – ORAL TABS 200 MG	84	01/01/2002	19078461	ibuprofen 200 MG Oral Tablet	M01AE01	ibuprofen; systemic, rectal	1200	mg	30	01/30/2002		14.0	01/14/2002
WARFARINE 3 MG ORAL TABS	56	01/01/2002	40163540	warfarin sodium 3 MG Oral Tablet	B01AA03	warfarin; systemic	7.5	mg	30	01/30/2002		22.4	01/23/2002
SIMVASTATIN 10 MG ORAL TABS	28	01/01/2002	1539463	simvastatin 10 MG Oral Tablet	C10AA01	simvastatin; oral	30	mg	30	01/30/2002		9.3	01/10/2002
RAMIPRIL **10 MG** ORAL CAPS	28	01/01/2002	1334494	ramipril 10 MG Oral Capsule	C09AA05	ramipril; oral	**2.5**	mg	30	01/30/2002		**112.0**	**04/22/2002**
SIMVASTATIN **80 MG** ORAL TABS	28	01/01/2002	19023563	simvastatin 80 MG Oral Tablet	C10AA01	simvastatin; oral	**30**	mg	30	01/30/2002		**74.7**	**03/16/2002**
RAMIPRIL **5 MG** ORAL CAPS	28	01/01/2002	1334460	ramipril 5 MG Oral Capsule	C09AA05	ramipril; oral	**2.5**	mg	30	01/30/2002		**56.0**	**02/25/2002**

Contact: tatiana.banokina@odysseusinc.com