Paving the way to estimate daily dose in OMOP CDM for Drug Utilisation Studies in DARWIN EU®

Theresa Burkard[1], Kim Lopez-Güell[1], Artem Gorbachev[2], Annika M Jödicke[1], Nuria Mercadé-Besora[4], Talita Duarte-Salles[3,4], Maria de Ridder[4], Mees Mosseveld[4], Dani Prieto-Alhambra[1,4], Christian Reich[4], Marti Catala[1]

[1] Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK
[2] Odysseus Inc
[3] Fundació Institut Universitari per a la recerca a l'Atenció Primària de Salut Jordi Gol i Gurina (IDIAPJGol)
[4] Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, The Netherlands

Background

Having robust and well-defined patient-level information is crucial to all research with routinely collected healthcare data. OMOP CDM provides a common structure for health data and numerous healthcare databases. Developing reliable, transparent, and standardized analytic packages can simplify and increase the analysis reliability and velocity while minimising errors from code sharing across different data sources.

Current suggestions and methods to estimate daily dose using data mapped to the OMOP CDM miss a quantitative evaluation. Furthermore, there is not a standardized way to compute daily dose that has been validated across different databases.

Therefore, we conducted a thorough investigation of the contents of the drug strength table. After cleaning up unit concept ids, we performed diagnostic checks on the drug strength table. Our aim is to share the process that will lead to the estimation of daily dose in OMOP CDM for "Drug Utilisation Studies" in DARWIN EU.

Methods

We grouped all drug concept ids in the drug strength table of OMOP CDM by amount_value (missing or present), amount_unit_concept_id, numerator_value (missing or present), numerator_unit_concept_id, denominator_value (missing or present), and denominator_unit_concept_id. The resulting combinations (further referred to as *drug strength patterns*) were independently assessed for relevant units for future dose estimations by two pharmacists (AJ, TB) and a medical doctor (AG). In a consensus meeting with a vocabulary expert (CR), relevant units and standardization were decided.

Subsequently, we performed two checks on the drug strength table that would support our reliance on the current drug strength table.

First, we assessed whether the following formula (1) held for all the drug concept ids.

(1) *Numerator_value = "drug concentration (as provided in the drug concept name)" * denominator_value*

Second, we assessed whether there were drug concentrations (as provided in the drug concept name) that may not be feasible, e.g. >1 mg/mg or >1 ml/ml (and as a sensitivity analysis: >2 mg/mg or >2 ml/ml)

For all combinations of numerator and denominator units (i.e. “concentration patterns”) we assessed the number of used unique drug concept ids (in combination with the quantity field) in the drug exposure table. This check was carried out to assess the difference between patterns with denominator value present versus those with missing denominator values. This would give us more insight into which patterns to prioritize.

Finally, we assessed coverage of the identified concentration patterns in both the drug strength table as well as drug exposure table.

All checks were performed in CPRD GOLD, CPRD AURUM, IPCI, and IQVIA PharMetrics Plus.

Results

We obtained 126 drug strength patterns in CPRD GOLD, CPRD AURUM, and IQVIA PharMetrics Plus, and 207 drug strengths patterns in IPCI which had additional source code related concepts in the drug strength table.

The review of the drug strength patterns for relevant units resulted in 31 drug strength patterns (Table 1). We identified the following relevant units for dose estimation among amount and numerator: “microgram”, “milligram”, “gram”, “international unit”, ”mega international unit”, ”liter”, ”milliliter”, ”milliequivalent”, and the following units for denominator: “milligram”, “gram”, ”hour”, ”milliliter”, ”liter”, ”actuation”, and ”square centimeter”.

Table 1. Relevant drug strength patterns for dose estimation (for better readability we display the unit concept name instead of the unit concept id)

amount	amount_unit	numerator	numerator_unit	denominator	denominator_unit
numeric	international unit	NA	NA	NA	NA
numeric	microgram	NA	NA	NA	NA
numeric	milliequivalent	NA	NA	NA	NA
numeric	milligram	NA	NA	NA	NA
numeric	milliliter	NA	NA	NA	NA
NA	NA	numeric	international unit	numeric	milligram
NA	NA	numeric	international unit	NA	milligram
NA	NA	numeric	international unit	numeric	milliliter
NA	NA	numeric	international unit	NA	milliliter
NA	NA	numeric	mega-international unit	NA	milliliter
NA	NA	numeric	microgram	numeric	hour
NA	NA	numeric	microgram	NA	hour
NA	NA	numeric	milliequivalent	NA	milligram
NA	NA	numeric	milliequivalent	numeric	milliliter
NA	NA	numeric	milliequivalent	NA	milliliter
NA	NA	numeric	milligram	numeric	Actuation
NA	NA	numeric	milligram	NA	Actuation
NA	NA	numeric	milligram	numeric	hour
NA	NA	numeric	milligram	NA	hour
NA	NA	numeric	milligram	numeric	liter
NA	NA	numeric	milligram	NA	liter
NA	NA	numeric	milligram	numeric	milligram
NA	NA	numeric	milligram	NA	milligram
NA	NA	numeric	milligram	numeric	milliliter
NA	NA	numeric	milligram	NA	milliliter
NA	NA	numeric	milligram	numeric	square centimeter
NA	NA	numeric	milligram	NA	square centimeter
NA	NA	numeric	milliliter	numeric	milligram
NA	NA	numeric	milliliter	NA	milligram
NA	NA	numeric	milliliter	numeric	milliliter
NA	NA	numeric	milliliter	NA	milliliter

After standardization of units to “milligram”, “international unit”, “milliliter”, “milliequivalent”, ”hour”, “actuation”, and “square centimeter”, we were left with 25 patterns for which a dose formula should be defined.

The first check of the drug strength table, assessing whether any of the drug concentrations multiplied by the denominator value was equal to the numerator value, was confirmed for all drug concept ids in all databases.

The second check of the drug strength table, counting the number of drug concentrations with >1 mg/mg or >1 ml/ml and >2 mg/mg or >2 ml/ml, yielded 8656 (0.5%) and 7841 (0.4%) counts in all databases, respectively.

Results for the check among the concentration patterns (i.e. combinations of numerator and denominator units) with present or missing denominator value, counting the number of unique drug concept ids (in combination with quantity) used in the drug exposure table, are presented in Table 2.

Table 2. Counts of unique drug concept ids (in combination with quantity) per concentration pattern that were used in the individual drug exposure tables

pattern name	counts in CPRD GOLD	counts in CPRD AURUM	counts in IPCI	counts in IQVIA PharMetrics Plus
international unit per milligram	0	0	0	0
international unit per milligram missing denominator	0	0	0	0
international unit per milliliter	0	0	0	0
international unit per milliliter missing denominator	**0**	**0**	**0**	0
mega international unit per milliliter	0	0	0	0
mega international unit per milliliter missing denominator	0	0	0	0
microgram per hour	0	0	0	0
microgram per hour missing denominator	0	0	0	0
milliequivalent per milligram	0	0	0	0
milliequivalent per milligram missing denominator	0	0	0	**15**
milliequivalent per milliliter	0	0	0	902
milliequivalent per milliliter missing denominator	0	0	0	**4606**
milligram per actuation	0	0	**14052**	2857
milligram per actuation missing denominator	**8959**	**5369**	561	**1500**
milligram per hour	4457	2220	0	35276
milligram per hour missing denominator	**1317**	**1070**	**0**	**47**
milligram per liter	0	0	0	0
milligram per liter missing denominator	0	0	0	0
milligram per milligram	2180	1200	2851	293
milligram per milligram missing denominator	**34058**	**15167**	**5097**	**6690**
milligram per milliliter	23316	10930	**35780**	63150
milligram per milliliter missing denominator	**83959**	**24338**	10308	**104190**
milligram per square centimeter	0	0	0	0
milligram per square centimeter missing denominator	**0**	**97**	**0**	**0**
milliliter per milligram	0	0	0	0
milliliter per milligram missing denominator	0	0	0	0
milliter per milliliter	0	0	25	36
milliter per milliliter missing denominator	**105**	**0**	**139**	**285**

It seems that only few patterns would actually be deployed (according to the drug exposure table). Furthermore, most drug concept ids are part of patterns with missing denominator values.

Finally, when assessing coverage of the patterns in the individual databases, we observed that 63% and 94% of the drug concept ids in the drug strength table and drug exposure table in CPRD GOLD, respectively, and 63% and 91% of the drug concept ids in the drug strength table and drug exposure table in IQVIA PharMetrics Plus are covered with our identified patterns (no results from IPCI or CPRD AURUM yet).

Conclusion

We shared the process that would lead to daily dose estimation in OMOP CDM for "Drug Utilisation Studies" in DARWIN EU. It shall be continued in more databases including more countries and settings (claims, hospital data). Furthermore, this approach shall maximize the completeness and reliability of dose estimations in OMOP CDM through identification of relevant units and patterns among the drug strength table, through testing the reliance of relevant information in the drug strength table, and through the assessment of the importance of the identified patterns. This preparatory work shall pave the way to estimate daily dose in OMOP CDM in the future by performing the additional steps of

suggesting a dose formula per identified pattern (and potentially per newly defined route for which we have another abstract submission).