Alignment of two standard terminologies for dosage form: RxNorm from the National Library of Medicine for the United States and EDQM from the European Directorate for the Quality in Medicines and Healthcare for Europe

Natalie Karapetian\textsuperscript{a,b,}\textsuperscript{*}, Robert Vander Stichele\textsuperscript{c}, Yuri Quintana\textsuperscript{a,b}

\textsuperscript{a} Division of Clinical Informatics, Department of Medicine, Beth Israel Deaconess Medical Center, Boston, MA, USA
\textsuperscript{b} Harvard Medical School, Boston, MA, USA
\textsuperscript{c} European Institute for Innovation through Health Data (I-HD), UNICOM Innovation Action; Department of Medical Informatics, Ghent University, Ghent, Belgium

ARTICLE INFO

Keywords:
Interoperability
Global health informatics
RxNorm
EDQM
IDMP

ABSTRACT

Background: There is currently no system that aligns pharmaceutically equivalent medicinal products across nations, creating obstacles to transnational medication prescribing and medical research. EDQM has been internationally recognized as the leading system in systematic pharmaceutical product descriptions. RxNorm is a critical terminology based in the US and used widely in applications internationally that would benefit from alignment with EDQM-based dosage form descriptions.

Goal: Demonstrate a method for alignment of RxNorm dosage forms with EDQM terminologies and with EDQM dosage forms. Describe obstacles and advantages of such an alignment for ultimate application in calculating universal Pharmaceutical Product Identifiers.

Methods: A pharmaceutical sciences student and a clinical pharmacology expert in dosage forms used definitions supplied by RxNorm and EDQM technical documentation to align the 120 RxNorm dose forms to EDQM-based dosage form description terms. The alignment of RxNorm to EDQM was then used to fit the RxNorm dose forms into an ontology based on EDQM.

Results and Conclusions: The alignment of RxNorm and EDQM requires further validation but provides a potential method of establishing interoperability between the two terminologies without cumbersome manual reclassification. There remain ambiguities within each dosage form nomenclature that create obstacles to integrating medication databases rooted in EDQM and RxNorm into a single worldwide database.

1. Introduction

Conventions for identifying therapeutic drugs vary greatly between drug classification systems across the world. Standardizing the identification of medicinal products would facilitate the development of an international drug information database to support greater interoperability of prescription information, enable improved global health surveillance to assess pandemic risks, enable transnational medical trials studying multiple patient populations, and maximize interoperability of medical data related to rare diseases [1], [2]. A directive issued in the European Union in 2011 mandated cross-national recognition of prescriptions and provided measures to facilitate verifying legitimacy of prescriptions issued by other member states [3]. While this directive requires that prescriptions be written using a “common name” [3] there are also significant variations in medication names and formulations across Europe markets.

A standardized drug identification system must be established to create this global drug database. The Identification of Medicinal Products (IDMP) standards was created in 2012 by the International Organization for Standards (ISO) to outline key characteristics of medicinal product classification systems for international harmonization [4]. One key component of IDMP is ISO 11616, which defines the elements

\textsuperscript{*} Corresponding author at: Division of Clinical Informatics, Department of Medicine, Beth Israel Deaconess Medical Center, Boston, MA, USA.

E-mail address: njkarape@health.ucsd.edu (N. Karapetian).
required to construct universal Pharmaceutical Product Identifiers (PhPIDs). PhPIDs are internationally recognizable identifiers for drug products and are composed of three main attributes: the active pharmaceutical substance, dosage form, and strength [5], allowing pharmacetically equivalent drug products to be recognized as synonymous regardless of regional differences in branding. While representations of active pharmaceutical substance and strength are similar across drug nomenclature systems, representations of drug dosage forms vary significantly. This variation between regional terminologies presents a substantial obstacle to achieving interoperability between drug nomenclature systems and generating universal PhPIDs.

In response to the ISO standards, the European Directorate for the Quality of Medicines and Healthcare (EDQM) reorganized its controlled vocabulary by relating dose forms to key dosage form characteristics [6]. Additionally, the European Union commissioned a project called Up-Scaling the Global Univocal Identification of Medicines (UNICOM) to better adhere to IDMP standards and establish a unified global drug database [7,8-9]. The UNICOM project identified EDQM as the most promising nomenclature for describing dosage forms due to EDQM’s adherence to a structured ontology and to IDMP standards [10].

Choosing EDQM as the standard for describing dosage forms implicated that drugs outside the EU must either be reclassified using EDQM terminology or that regional terminologies must align with EDQM. The World Health Organization, United States Food and Drug Administration (FDA), Uppsala Monitoring Centre, and UNICOM completed a pilot project to reclassify dosage form descriptions of US drug products from the FDA’s Terminology for Structured Product Labeling (SPL) to EDQM [11]. The pilot results were discussed in a June 11th, 2021 webinar identifying the main challenge in standardizing dosage form descriptions to EDQM as the large number of US drug products on the market needing manual reclassification [11]. This spurred interest in alignment of existing drug nomenclature systems with EDQM, rather than conversion, to avoid the cumbersome process of manual reclassification.

While EDQM is of growing significance for its adherence to IDMP standards, the drug terminology system operated by the United States National Library of Medicine, known as RxNorm, is of prevalent utility both nationally and internationally. RxNorm is the basis for the U.S. Department of Veterans Affairs National Drug File-Reference Terminology [12], the Drug Ontology (DrOn) [13], and the Observational and Medical Outcomes Partnerships (OMOP) open-source common data model used across the United States and Europe to conduct multidisciplinary drug studies [14]. Interoperability between the Canadian drug ontology OCRx and RxNorm has also been demonstrated, further contributing to the international significance of RxNorm [15].

While studies have evaluated interoperability between FDA SPL [11] and SNOMED-CT [16,17], to EDQM, there no studies have evaluated alignment of dosage form descriptions between RxNorm and EDQM. Overall, there are few published works on this topic as the initiative to align regional nomenclatures with EDQM is currently predominated by governing agencies rather than researchers [18], [19-20]. Given the national and international significance of RxNorm, the objective of this paper is to experimentally align RxNorm and EDQM dosage forms, using EDQM descriptors and RxNorm dosage forms, and by creating an ontology encompassing both RxNorm and EDQM dosage forms.

2. Methods

2.1. Overview of Methods

A student of pharmaceutical sciences (NK) and an international expert in the clinical pharmacology of dosage forms (RVS) compared the description of dosage form representations within RxNorm and EDQM. This process involves two steps: first, assignment of EDQM characteristics to each RxNorm Dose Form based on the definitions provided by the United States National Library of Medicine (Supplement A and Supplement B); second, using these assigned characteristics to align the RxNorm Dose Forms to a simple ontology of EDQM dose forms (Supplement C and Supplement D).

2.2. RxNorm Nomenclature

RxNorm is a pharmaceutical products naming system established by the United States National Library of Medicine. RxNorm was created to support interoperability between health-related terminologies and knowledge bases across medical applications used in the United States [21]. RxNorm provides standardized representation of pharmaceutical ingredients, strength, dose form, and brand name information.

For each unique drug product, an array of codes represents different identifiers of that product. Codes of interest to the generation of PhPID include the Ingredient (IN) code, the Dosage Form (DF) code, and the Semantic Clinical Drug (SCD) code, which is the only RxNorm code to include description of a product’s strength. For the alignment of RxNorm dosage forms to EDQM dose form descriptors, only the DF code is relevant. RxNorm dosage forms are aggregated into 42 overlapping Dose Form Groups based on the route of administration, release characteristics or product type. Each of the 120 RxNorm dose forms is contained in at least one Dose Form Group, with dose forms often belonging to several Dose Form Groups. For this reason, RxNorm does not provide an ontology but rather a list of 120 defined dosage forms in overlapping groups. A sample from the RxNorm representation of dosage forms for sublingual tablets follows, illustrating redundancies of dosage form representations between Dose Form Groups as sublingual tablet belongs to three separate dosage form groups:

**Oral product (Dosage Form Group)**

- Sublingual Tablet (Dosage Form)
- Capsule (Dosage Form)
- Tablet (Dosage Form)

**Pill (Dosage Form Group)**

- Sublingual Tablet (Dosage Form)
- Buccal Tablet (Dosage Form)
- Chewable Tablet (Dosage Form)

**Sublingual Product (Dosage Form Group)**

- Sublingual Tablet (Dosage Form)
- Sublingual Film (Dosage Form)

2.3. EDQM Nomenclature

EDQM maintains a set of controlled vocabularies to describe six characteristics of pharmaceutical dosage forms [6] (without including ingredient and strength). These characteristics include Basic Dose Form (BDF), State of Matter (SOM), Transformation (TRA), Release Characteristics (RCA), Intended Site (ISI), and Administration Method (AME). The BDF and SOM refer to the drug’s form (such as a cream, tablet) and physical state of matter (solid, semi-solid, liquid, or gas). Transformation refers to whether the product requires modification before administration, such as dilution or mixing. RCA refers to alteration of the drug release timing (such as prolonged or delayed-release), ISI refers to the anatomical site of drug administration (such as oral or ocular), and AME refers to the method of drug administration (such as via swallowing or inhalation). The ISI organizes EDQM dosage forms. Sublingual tablet only appears once in the EDQM ontology, as demonstrated in the EDQM representation below. The four characteristics used for alignment with RxNorm are bolded:

**Oral (Site of Administration)**

- Sublingual Tablet (Dosage Form)
  - State of Matter (SOM): Solid
  - Basic Dose Form (BDF): Tablet
  - Transformation (TRA): No Transformation
  - Release Characteristics (RCA): Conventional
2.4. Proposed Generation of PhPID from RxNorm-EDQM Alignment

As stated in the introduction, generation of universally recognized PhPID requires input of universal codes describing the product’s main ingredient, strength, and dosage form (Fig. 1). RxNorm-supplied IN and SCD codes can be mapped to universal ingredient and strength codes, respectively, to complete the generation of PhPID; the decision of which ingredient and strength codes should be used for universal PhPID input is outside the scope of this paper, which is focused on using EDQM as the standard for dosage forms. We thereby focus on alignment of RxNorm DF codes and EDQM to generate the dosage form component of PhPID.

To align RxNorm-supplied DF codes and EDQM codes for PhPID generation, all US products could be manually reclassified to EDQM dosage forms (which was deemed unfavorable during the FDA pilot study) or unique combinations of a limited number of EDQM dosage form characteristics could be generated to represent US dosage forms. For each RxNorm drug product, the Dose Form (DF) code can be aligned to RCA, ISI, TRA, and AME codes. These four dosage form characteristics, combined with ingredient and strength, can generate a PhPID to identify pharmaceutical products independent of regional naming conventions (Fig. 1).

2.5. Aligning RxNorm Dose Forms with EDQM Descriptors

A list of all RxNorm dose forms was downloaded from Supplement B of the RxNorm Technical Documentation (version reviewed July 6, 2020) [21]. This list was uploaded into a single column of a Microsoft Excel spreadsheet. Four columns were created corresponding to the four key characteristics described previously: TRA, RCA, ISI, and AME. To accommodate for lacking delineation between manufactured dose forms and administrable dose forms, two additional columns were created representing the basic dose form of the manufactured product (prior to any transformation) and of the administrable drug (after any required transformation). The SOM of the administrable dose form (ADF) was also generated for completion but was not needed for our proposed method of alignment (example in Table 1).

A student of pharmaceutical sciences (NK) and an international expert in clinical pharmacology of dosage forms (RVS) reviewed the definitions for RxNorm dosage forms, as provided by Supplement B of the United States National Library of Medicine RxNorm Technical Documentation. They also reviewed the EDQM Standard Terms and Internal Controlled Vocabularies for Pharmaceutical Dose Forms (Version 1.2.0), which defines EDQM descriptors. The reviewers consulted these resources to independently assign the six EDQM characteristics to all 120 RxNorm dosage forms according to the guidelines presented in Supplement A (example in Table 1). For dose forms not requiring transformation, the basic dose form of the Manufactured Dose Form and the Administrable Dose Form were the same. The RCA, ISI, and AME were subsequently assigned in reference to the administrable dose form. The expert in pharmaceutical dosage forms reviewed each alignment for adherence to EDQM and RxNorm definitions. Any discrepancies were discussed by the reviewers until a consensus was reached. A third medical informatics expert (YQ) was available to resolve differences, which did not occur. A complete file of these assignments is provided in Supplement B. The authors are members of UNICOM, a European government-funded consortium focused on drug interoperability, which provided the broader expert opinion related to this initiative.

2.6. Construction of an ontology

We re-organized RxNorm dosage forms by fitting them into a simple
EDQM-based ontology created by the UNICOM project [19] (Supplement C and Supplement D). This ontology was based on the EDQM ISI as the primary grouping mechanism and provides a structured method of representing RxNorm dosage forms. This method of organizing dosage forms eliminates the redundancies of RxNorm Dosage Form Groups and allows direct comparison of EDQM and RxNorm dosage forms. EDQM ISI was chosen as the primary grouping mechanism as it permits intuitive searching through the ontology and only has 21 types. Additionally, intermediate level groupings are possible based on the RxNorm ISI by looking at the unique combinations of the characteristics.

3. Results

3.1. Results of RxNorm-EDQM Descriptor Alignment

Overall, 119 of the 120 RxNorm dosage forms were aligned to all four of key EDQM characteristics. The outstanding RxNorm dosage form, Mucosal Spray, could not be assigned an ISI as mucosal tissue can be found throughout the body. Additionally, 72 of the RxNorm dosage forms had an equivalent EDQM dosage form, 48 dosage forms were exclusive to RxNorm, and 356 dosage forms were exclusive to EDQM. Assignment of manufactured drug basic form to RxNorm dosage forms revealed ambiguity of some RxNorm dosage forms, creating disagreement. For example, the dosage form “Nasal Inhalant” could contain a liquid or a gas. The RxNorm dosage forms “auto-injector,” and “jet injector” are analogous to delivery devices and could contain a “suspension,” “solution,” or “reconstituted powder.” Additionally, there are no EDQM equivalents for some RxNorm terms, including “oral flake” and “oral wafer.” These factors complicated assignment of Basic Form of ADF and SOM of ADF, which are contingent on accurate assignment of Manufactured Drug Basic Form.

The assignment of transformation was more linear, with ambiguity only regarding the RxNorm dosage form “topical liquefied gas.” This transformation could not be described using “dispersion”, “dissolution”, “dilution,” or “mixing” so it was described using the “unknown” descriptor. There were no issues assigning the basic form of ADF once the manufactured drug basic form and transformation were established and there were no issues assigning the state of matter of ADF once the basic form of the ADF was established.

Two issues appeared when assigning the RCA of ADF. First was that the EDQM descriptor “modified” release is ambiguous as any non-conventional release form is encompassed by the “modified” term. We therefore did not use the “modified” classification and opted instead for the more specific “delayed” or “prolonged” descriptors. Additionally, some dosage forms, such as drug implant, are innately “prolonged” release. This leads to the question of being assigned “conventional” compared to other drug implants or assigned “prolonged” release due to the innate prolonged release patterns of a drug implant. The authors, as members of the UNICOM project, have brought these issues to the leaders of EDQM and a discussion among EDQM experts is being planned.

One major issue with both EDQM and RxNorm descriptor systems is they do not distinguish between systemic and locally active dosage forms. For example, the “nasal” intended site can act locally (such as allergy medications) or systemically (such as naloxone). Similarly, ambiguity was present in assigning “pulmonary” and “nasal” administration methods to some inhalant-related products as these products were inserted nasally but inhaled for delivery to the pulmonary system. These issues have been brought to the attention of the ISO committees responsible for dose form standards to be included in the next revision of the standards.

3.2. Comparing RxNorm and EDQM with a Common Ontology

Creation of the ontology allows direct alignment of EDQM and RxNorm dosage forms coupled with the associated EDQM descriptors as object properties. This process permitted direct comparison of the EDQM and RxNorm dosage forms, revealing most RxNorm dosage forms could be aligned with an EDQM dosage form. However, this process highlighted a problematic lack of representation of many EDQM dose forms in RxNorm and that the same combinations of EDQM characteristics described unique RxNorm dosage forms.

Substantial differences between EDQM and RxNorm include EDQM specification of drops, cutaneous dosage forms beyond topical liquids, endocervical dosage forms, dialysis dosage forms, gastroenteral dosage forms, specification of dosage forms for injection and infusion, and specification of nebulized dosage forms. Additionally, some RxNorm dose forms are more reminiscent of drug delivery devices, such as jet injector, metered dose inhaler, injection cartridge, and enema. While drug delivery devices as described above may be key in correct dispensing of drugs, there is no mention of drug delivery devices in EDQM.

4. Discussions and Conclusions

Our experimental alignment of RxNorm and EDQM dosage forms demonstrated that 72 RxNorm and EDQM dosage forms are equivalent in EDQM. And 119 of the 120 RxNorm dosage forms could be assigned EDQM TRA, AME, ISI, and RCA terms for the calculation of PhPID (Supplement C and Supplement D). We encountered many of the same obstacles detailed by the FDA pilot project [11]. A key issue to the proposition of using calculated PhPID based on the four characteristics to identify dosage forms is there are many unique dosage forms with the same combination of the four EDQM descriptors. This could be largely overcome by including the basic dose form (BDF), of which there are 52 descriptive terms, in PhPID calculations. We have already begun
aligning RxNorm Dosage Forms with the BDFs (Supplement B). Ambiguity in RxNorm dosage form names is another contributor to unreliable assignment of descriptors, particularly for manufactured dosage forms. These issues contributed to variable assignment of 19 RxNorm terms to EDQM characteristics between the two assigning investigators (NK and RVS) and required discussion to reach consensus. The disparities were mostly related to assigning ISI and AME codes. Lacking specificity of RxNorm dosage form descriptions, such as hard versus soft capsule, was highlighted in the ontology and contributes to the high number of EDQM dosage forms compared to RxNorm. Due to the inconsistent granularity of RxNorm dosage form groups and redundant dose form representations across groups, it is recommended that ontologies and data models founded on RxNorm be structured on the list of dose forms, rather than the dose form groups. This reflects the structure of OMOP [22].

It remains debatable whether generating PhPID codes based on a combination of EDQM characteristics is a justified approach over manually assigning EDQM dosage forms to US products. As all US medicinal products are currently described using RxNorm nomenclature, a clear advantage of using the EDQM characteristic alignment we presented is that it is more automated compared to manually assigning EDQM dosage forms to all US medicinal products. However, this convenience comes risks potentially misrepresenting dosage forms. The importance of accurate EDQM descriptor assignment was emphasized during a UNICOM Community of Experts Webinar during discussion of nations variably assigning basic dose form characteristics dispersion, suspension, and solution to the Pfizer coronavirus vaccine, leading to the generation of three unique PhPID for the same product [8].

One major hindrance to clinical utility of this approach is that there is no reference to whether the drug product is systemically or locally active, an important aspect of understanding the drug’s actions. The quality issues and redundancies in dose form representations discussed here have also been described as challenges using RxNorm for clinical decision support [23,24]. RxNorm lacks the attributes of a traditional ontology required to facilitate universal interpretability, further supporting the proposal to describe RxNorm dose forms with structured EDQM characteristics or using the EDQM dose form ontology proposed by the UNICOM project [19] (Supplement C and Supplement D).

Given the lack of literature on this topic to formulate our alignment strategy, our alignment of RxNorm terms and EDQM characteristics should now be subjected to further validation. One potential method of validating this technique includes leveraging a mapping hosted by BioPortal using SNOMED-CT as an intermediary between RxNorm and EDQM to identify additional points of misalignment. Additionally, the presented alignment could be applied to a set of real medicinal products from the US and EU to assess reliable expression with both terminologies. This paper provides foundational contributions to improving the interoperability of pharmacy systems and mobile apps used internationally by patients for refilling drugs when traveling or relocating. Future interoperability efforts will continue under the initiatives of the UNICOM consortium and discussion with governments and standards agencies.

5. Summary points

What was already known on the topic?

• FDA has expressed a need to establish interoperability between US-based dose form descriptions and European-based EDQM. This has been piloted with the FDA SPL nomenclature.
• Interoperability has been established between the drug-classification systems of other countries, such as Canada, and RxNorm.

What did this study add to our knowledge?

• RxNorm dose forms can be aligned to EDQM dose form descriptors using RxNorm DF codes and EDQM RCA, ISI, TRA, and AME codes.

• This alignment process requires further validation but is a promising bridge in harmonizing drug descriptions internationally.

Authors’ contributions

This paper was written by Natalie Karapetian under the guidance and revision of Robert Vander Stichele and Yuri Quintana. Natalie Karapetian and Robert Vander Stichele independently conducted the alignment process. Robert Vander Stichele also crafted the ontology to which the RxNorm and EDQM dosage forms were fitted, with explanation of the ontology creation currently pending publishing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The UNICOM Innovation Action has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 875299.

Statement on conflicts of interest

Natalie Karapetian participated in an educational Informatics internship program at the Division of Clinical Informatics at Beth Israel Deaconess Medical Center with financial support from Pfizer, but Pfizer had no role in this study or article preparation. Yuri Quintana and Natalie Karapetian declare no competing interests. Robert Vander Stichele works in the UNICOM Innovation Action and declares no competing interests.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijmedinf.2022.104826.

References


[16] Bodenreider O. 201907 Comparing the representation of medicinal products in RxNorm and SNOMED CT – Consequences on interoperability. :5.


