Comparing Drug Classes between MED-RT and SNOMED CT

Olivier Bodenreider
U.S. National Library of Medicine
National Institutes of Health
Bethesda, Maryland, USA
olivier.bodenreider@nih.gov

Robert Wynne
U.S. National Library of Medicine
National Institutes of Health
Bethesda, Maryland, USA
robert.wynne@nih.gov

Abstract—Objectives. To compare drug classes between MED-RT and SNOMED CT. We also explore whether SNOMED CT classes could form the basis for drug value sets, and be a potential alternative to MED-RT. Methods. We compare the characteristics of the two drug classification systems for mechanism of action and chemical structure classes, with focus on drug-class membership. Results. While MED-RT provides a larger number of drug classes and deeper hierarchies, SNOMED CT provides better coverage (i.e., classifies 70-90% more drugs than MED-RT does). Conclusions. Our preliminary investigation shows that the coverage provided by SNOMED CT compares favorably to MED-RT’s. Except for open access, no other characteristic of MED-RT makes it superior to SNOMED CT for drug classification, including for use in clinical decision support and value set creation.

Keywords—SNOMED CT; MED-RT; Drug classes

I. INTRODUCTION

Clinical decision support (CDS) often relies on drug classes to formulate or analyze CDS rules related to medications (e.g., [1]). Similarly, drug classes play an important role in defining value sets for electronic clinical quality measures (eCQMs) [2]. For example, the value set “Statin Allergen” is used as part of an exclusion criterion in eCQMs mandating statin treatment in specific populations (e.g., Statin Therapy for the Prevention and Treatment of Cardiovascular Disease).

The 2019 edition of the Interoperability Standards Advisory (ISA) [3] published by the Office of the National Coordinator for Health Information Technology (ONC) lists SNOMED CT as the preferred standard for representing medications in the context of patient allergies and intolerances: “When a medication allergy necessitates capture by medication class, SNOMED CT should be used”. The Medication Reference Terminology (MED-RT), released in 2018, is listed in the same context as a standard under development. Previous versions of the ISA listed NDF-RT; MED-RT’s predecessor, as the standard for drug classes [4].

While NDF-RT has been investigated as a source of drug classes for clinical decision support and clinical quality measurement (see related work for details), MED-RT and the recently revised medicinal product hierarchy of SNOMED CT have not been studied yet.

The objective of this investigation is to compare drug classes between MED-RT and SNOMED CT. More specifically, we compare the characteristics of the two drug classification systems, with focus on drug-class membership. (Of note, a pairwise comparison of drug classes between MED-RT and SNOMED CT is out of scope for this preliminary investigation.) This comparison will help us explore whether SNOMED CT classes could form the basis for drug value sets, and a potential replacement for NDF-RT and its successor, MED-RT.

II. BACKGROUND

We review related work about the sources of drug classes under investigation before presenting the resources used in our study.

A. Related work

NDF-RT was created as a principled alternative to legacy drug classifications, formally identifying classification criteria, such as mechanism of action, chemical structure and pharmacokinetics [5, 6]. NDF-RT has been investigated from various perspectives, such as clinical decision support, including prevention of allergic reactions [7-9], the characterization of adverse events [10, 11], and basic drug classification [12-14]. Its successor, MED-RT, released just last year has not been studied yet.

Drug classes from SNOMED CT have been investigated for their support to encoding allergy information [15], and compared to NDF-RT for this purpose [7]. The hierarchical organization of SNOMED CT and NDF-RT has also been compared [16]. However, the medicinal product hierarchy of SNOMED CT was significantly modified in recent years and none of these investigations was based on a recent version of SNOMED CT.

In previous work, we have investigated drug value sets in light of drug classification systems [17], but did not consider SNOMED CT then.

The work closest to the present study is our earlier comparison of drug-class membership in NDF-RT and SNOMED CT, in which we performed a pairwise
comparison of drug classes in the two systems based on their drug members \[18\]. In our present work, we only analyze drug-class membership to assess the coverage of the drug classification systems, not to align their classes.

In summary, the specific contribution of this work is to investigate current versions of these drug classification systems (MED-RT and the new medicinal product hierarchy of SNOMED CT), from the perspective of their characteristics, including drug-class membership. To our knowledge, this is the first investigation of these two drug classification systems since they have undergone important changes in recent years. Aligning classes between MED-RT and SNOMED CT is out of scope for this preliminary investigation.

B. Materials

1) MED-RT/DailyMed

The Medication Reference Terminology (MED-RT) is developed by the Veterans Health Administration at the U.S. Department of Veterans Affairs \[19\]. MED-RT replaced the National Drug File – Reference Terminology (NDF-RT) in 2018. NDF-RT provided a vocabulary for drugs and for drug classes, as well as drug-class membership assertions. In contrast, MED-RT asserts drug-class membership relations between RxNorm drugs (see below) and drug classes expressed with the legacy NDF-RT vocabulary (e.g., for mechanism of action and physiologic effect) or with external vocabularies (e.g., chemical concepts and disease concepts from the Medical Subject Headings (MeSH)).

While MED-RT provides drug-class membership assertions, other sources, such as the U.S. Food and Drug Administration (FDA), link drugs to MED-RT classes for mechanism of action, physiologic effect and chemical structure. These assertions are released as “SPL indexing files” on the DailyMed website \[20\]. We use these assertions over those in MED-RT, because we consider FDA a more authoritative source.

The versions of MED-RT and SPL indexing files used in this investigation were the versions available for download in April 2019.

2) SNOMED CT

SNOMED CT is the largest clinical terminology in the world \[21\]. In 2018, SNOMED International made significant changes to the representation of medicinal products in SNOMED CT, including in the representation of drug-class membership information. More specifically, medicinal products inherit properties, such as “dispositions” (roughly equivalent to mechanism of action) and “structure” (i.e., chemical structure), ascribed to the substances these medicinal products have as their active ingredient(s).

For example, the substance clopidogrel (386952008) has disposition Platelet aggregation inhibitor (771451006), which makes it a subclass of Substance with platelet aggregation inhibitor mechanism of action (771452004). Similarly, the medicinal product Product containing clopidogrel (108979001), which has active ingredient the substance clopidogrel, is a subclass of the drug class Product containing platelet aggregation inhibitor (773388004), defined as having active ingredient Substance with platelet aggregation inhibitor mechanism of action.

Of note, SNOMED CT also provides the vocabulary for therapeutic role classes, namely Medicinal product categorized by therapeutic role (763087004) and its descendants. However, therapeutic roles, such as indications, age generally defined by regulatory agencies in the member countries of SNOMED International. For this reason, there are currently no drug-class membership relations (with very few exceptions) between medicinal products and drug classes for therapeutic roles.

The version of SNOMED CT used in this investigation is the U.S. Edition released in March 2019.

3) RxNorm

RxNorm is the standard drug terminology in the U.S. developed by the National Library of Medicine \[22\]. It is used by MED-RT as the vocabulary of reference for drugs, with which drug classes are associated. RxNorm also integrate substances and medicinal products from SNOMED CT, making it possible to compare drugs in MED-RT (i.e., RxNorm concepts) to drugs in SNOMED CT (mapped to RxNorm concepts in the RxNorm dataset). Finally, RxNorm distinguishes between base ingredients (e.g., atorvastatin) and “precise ingredients”, often salts and esters (e.g., atorvastatin calcium). Through RxNorm, precise ingredients can be normalized to their base ingredient to facilitate comparisons across sources of drug-class membership information, such as MED-RT and SNOMED CT.

The version of RxNorm used in this investigation is the April 2019 version. It is consistent with the set of drugs from the March 2019 version of SNOMED CT and with the April 2109 version of MED-RT.

III. METHODS

In our investigation of drug classes in MED-RT and SNOMED CT, we first describe drug classes in both terminologies, before we describe drug-class membership (i.e., relations between drug entities and class entities). Finally, we briefly discuss implementation.

A. Acquiring drug classes and drug-class membership information for MED-RT and SNOMED CT

1) MED-RT

We use the restful RxClass application programming interface (API) \[23\] to obtain the list of all drug classes in MED-RT (/allClasses). For each class, we retrieve the class name (/class/byId), its ancestors (/classGraph) and its drug members (/classMembers).

2) SNOMED CT

We use the SNOMED CT relational database \[21\] to obtain the list of all isa relations within the Medicinal products hierarchy (766779001), as well as the fully specified name for each concept from this hierarchy. Since SNOMED CT does not explicitly distinguish between drug classes and their drug members, we rely on the semantic tags
(i.e., the parenthetical expression at the end of fully specified names). More specifically, we identify as a class any descendant, direct or not, of the three class types under investigation (see below) with a semantic tag of “(product)”. Analogously, we identify as drug member any direct descendant of a drug class with a semantic tag of “(medicinal product)”. Of note, drug members that are direct descendants of the top-level of the Medicinal products hierarchy or direct descendants of any of the three class types are ignored, since no specific class is assigned to them.

3) RxNorm

We use the restful RxNorm API [24] to map SNOMED CT drugs identifiers to RxNorm (/rxcui/{rxcui}/related?tty=IN). SNOMED CT drugs identifiers are mapped to corresponding RxNorm concepts (where available) and to associate drugs from MED-RT and SNOMED CT, with their base ingredient (/rxcui/{rxcui}/related?tty=IN).

B. Describing drug classes

1) Types of drug classes

MED-RT provides vocabulary for the following types of drug classes: Mechanisms of Action, Physiologic Effects, FDA Established Pharmacologic Classes, Pharmacokinetics, and Therapeutic Categories. In addition, MED-RT relies on the MeSH vocabulary for the Structural Classifications of Ingredients and for Diseases (e.g., for indications).

SNOMED CT includes three drug class types: Medicinal product categorized by disposition (766779001), Medicinal product categorized by structure (763760008), and Medicinal product categorized by therapeutic role (763087004).

The class types common to MED-RT and SNOMED CT include mechanism of action [MOA] (Mechanisms of Action / Medicinal product categorized by disposition), chemical structure [CHEM] (Structural Classifications of Ingredients / Medicinal product categorized by structure) and therapeutic aspects [THER] (Therapeutic Categories / Medicinal product categorized by therapeutic role). However, since SNOMED CT does not currently assert drug-class membership for therapeutic aspects, we limit our investigation to MOA and CHEM classes.

2) Number of classes

We count the number of classes under the two class types under investigation, MOA and CHEM. In practice, since drug classes are rooted by a class type concept, we simply count the descendants of the class type concept that correspond to drug classes, regardless of whether or not these classes have drug members.

3) Hierarchical organization

Traversing the hierarchical structure of the drug classes, we calculate the maximal length of the path between any drug class and the root of the hierarchy (i.e., the class type).

C. Describing drug-class membership

For the initial coverage investigation, we simply consider substances as they appear MED-RT and SNOMED CT (i.e., not normalized to base ingredients). For the comparison, however, we consider substances after normalization to their base ingredient. Of note, no multi-ingredient drugs are included in this investigation, because only single-ingredient drugs are linked to MED-RT and SNOMED CT classes directly.

1) Coverage in MED-RT and SNOMED CT from the perspective of drug classes

For each drug class, we count the number of drug entities (not normalized to their base ingredient) that are a member of this class, directly or indirectly (i.e., through subclasses) and report whether the drug class has at least one drug member. For example, in MED-RT, the MOA class Monoamine Oxidase Inhibitors has direct member Isocarboxazid and indirect member safinamide (through its subclass Monoamine Oxidase-B Inhibitors).

2) Coverage in MED-RT and SNOMED CT from the perspective of drugs

For each drug entity (not normalized to its base ingredient), we count the number of drug classes that have this drug as a member, directly or indirectly (i.e., through ancestor classes). For example in SNOMED CT, the drug Amoxicillin is a member of the CHEM class Product containing aminopenicillin and an indirect member of its parent class, Product containing broad spectrum penicillin.

3) Comparing coverage between MED-RT and SNOMED CT

To facilitate the comparison between MED-RT and SNOMED CT, we normalize all drugs to their base ingredient in RxNorm as described earlier. For example, Etidronic Acid is normalized to Etidronate.

For each drug entity (normalized to its base ingredient), we count the number of drug classes (for each class type, respectively) that have this drug as a direct member in MED-RT, SNOMED CT or both. This simple analysis does not ascertain that the classes asserted for a given drug in MED-RT and SNOMED CT are equivalent, but simply that such an assertion exists for a given class type. For example, the drug Atenolol has direct MOA class Adrenergic beta-Antagonists in MED-RT and Product containing beta-1 adrenergic receptor antagonist in SNOMED CT. In contrast, efavirenz has CHEM class Non-Nucleoside Analog in MED-RT, but none in SNOMED CT, while clopidogrel has CHEM class Product containing chlorinated hydrocarbon in SNOMED CT, but none in MED-RT.

D. Implementation

Investigating relations between drugs and classes in MED-RT and SNOMED CT can be thought of as a reachability problem over linked data resources. Namely from a base ingredient in RxNorm, which corresponding drugs can we find in MED-RT and SNOMED CT, and which drug classes are they a member of? To facilitate this exploration, we first transform RxNorm, MED-RT and SNOMED CT to RDF – the Semantic Web “Resource Description Framework” – and use SPARQL queries to test reachability. All queries were processed using the triple store Virtuoso (open source edition, version 7.20).
IV. RESULTS

A. Describing drug classes

1) Types of drug classes
As mentioned earlier, we limit our investigation to MOA and CHEM classes.

2) Number of classes
The number of MOA and CHEM classes in MED-RT and SNOMED CT is shown in Table 1.

<table>
<thead>
<tr>
<th>Class type</th>
<th>MED-RT</th>
<th>SNOMED CT</th>
</tr>
</thead>
<tbody>
<tr>
<td>MOA</td>
<td>642</td>
<td>339</td>
</tr>
<tr>
<td>CHEM</td>
<td>9867</td>
<td>306</td>
</tr>
</tbody>
</table>

3) Hierarchical organization
The maximal depth of the MOA class hierarchy is 7 in MED-RT and 5 in SNOMED CT, while the maximal depth of the CHEM class hierarchy is 12 in MED-RT and 6 in SNOMED CT.

B. Describing drug-class membership

1) Coverage in MED-RT and SNOMED CT from the perspective of drug classes
The number of MOA and CHEM classes with direct and indirect drug members in MED-RT and SNOMED CT is shown in Table 2.

<table>
<thead>
<tr>
<th>Class type</th>
<th>Direct</th>
<th>Direct + indirect</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>MOA</td>
<td>291</td>
<td>445</td>
<td>815</td>
</tr>
<tr>
<td>CHEM</td>
<td>271</td>
<td>516</td>
<td>2942</td>
</tr>
</tbody>
</table>

2) Coverage in MED-RT and SNOMED CT from the perspective of drugs
The number of drugs with (direct) MOA and CHEM classes in MED-RT and SNOMED CT is shown in Table 3.

<table>
<thead>
<tr>
<th>Class type</th>
<th>MED-RT</th>
<th>SNOMED CT</th>
</tr>
</thead>
<tbody>
<tr>
<td>MOA</td>
<td>820</td>
<td>1391</td>
</tr>
<tr>
<td>CHEM</td>
<td>1267</td>
<td>2435</td>
</tr>
</tbody>
</table>

3) Comparing coverage between MED-RT and SNOMED CT
After normalizing the drugs to their base ingredients, there are 3239 drugs in common between MED-RT and SNOMED CT. Of these, 1536 (47%) have a MOA class and 2942 (92%) have a CHEM class in either terminology.

The number of drugs with MOA classes in MED-RT and SNOMED CT is shown in Table 4 and the number of drugs with CHEM classes in MED-RT and SNOMED CT is shown in Table 5.

<table>
<thead>
<tr>
<th>Class</th>
<th>Yes</th>
<th>No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>MED-RT</td>
<td>553</td>
<td>262</td>
<td>815</td>
</tr>
<tr>
<td>SNOMED CT</td>
<td>721</td>
<td>--</td>
<td>721</td>
</tr>
<tr>
<td>Total</td>
<td>1274</td>
<td>262</td>
<td>1536</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Class</th>
<th>Yes</th>
<th>No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>MED-RT</td>
<td>404</td>
<td>851</td>
<td>1255</td>
</tr>
<tr>
<td>SNOMED CT</td>
<td>1687</td>
<td>--</td>
<td>1687</td>
</tr>
<tr>
<td>Total</td>
<td>2091</td>
<td>851</td>
<td>2942</td>
</tr>
</tbody>
</table>

V. DISCUSSION

A. Findings

Types of classes. From the perspective of types of drug classes, both MED-RT and SNOMED CT offer distinct classification systems for various types of classification criteria (e.g., MOA, CHEM). This is in contrast to other drug classification systems, such as ATC (the Anatomical Therapeutic Chemical drug classification) and the legacy classification system included in the Veterans Health Administration National Drug File (VANDF), where multiple classification criteria are integrated and contribute to creating a mixed hierarchy of drug classes.

MED-RT offers a larger number of classification criteria than SNOMED CT does. More specifically, both MED-RT and SNOMED CT cover mechanism of action (under “disposition” in SNOMED CT), and chemical structure. The therapeutic use of drugs is also covered in both MED-RT (“Diseases for indications”) and SNOMED CT (“therapeutic role”), although SNOMED CT provides virtually no drug-class membership information for this type of classes. Some types of classes are covered only in MED-RT, including Physiologic Effects, FDA Established Pharmacologic
Classes, Pharmacokinetics, and Therapeutic Categories. In effect, MED-RT is richer than SNOMED CT. However, MOA and CHEM are arguably the most important categories for clinical decision support and are covered in both drug classification systems. And SNOMED CT is investigating plans for therapeutic role to be covered (in the international release or, more likely, in national extensions, given the regulatory nature of this information).

**Number of classes, coverage and hierarchical organization.** MED-RT offers a larger number of classes than SNOMED CT (Table 1). More specifically, MED-RT has roughly twice as many MOA classes as SNOMED CT and 30 times more CHEM classes. However, the extent to which “more is better” is arguable. In fact, as shown in Table 2, the number of classes effectively used for drug classification (i.e., classes with direct drug members) is similar in both systems (and even slightly larger in SNOMED CT). And even considering the number of indirect classes, the difference between the two systems is less important than suggested by the total number of classes. What this means is that MED-RT provides a large number of classes that do not play any role in drug classification, especially for CHEM. The profusion of classes can create difficulties in navigating the class hierarchy. For example, MED-RT only categorizes 17 drugs under the relatively large Inorganic Chemicals hierarchy borrowed from the MeSH vocabulary.

Of note, MED-RT uses MeSH concepts rather than MeSH descriptors for chemical classes – descriptors are aggregates of concepts in MeSH. However, MeSH does not provide any hierarchical relations among their concepts, only among their descriptors. As a result, the MeSH descriptor hierarchy is used in RxClass to elicit hierarchical relations among CHEM classes and to make the MED-RT CHEM hierarchy browsable. A similar challenge exists with the MED-RT Disease hierarchy, also borrowed from MeSH, but does not exist with other MED-RT hierarchies or in SNOMED CT.

The difference in maximal depth of the MOA and CHEM hierarchies between MED-RT and SNOMED CT is more important for CHEM, in part due to use of the entire chemical tree from MeSH.

**Coverage (from the perspective of drugs).** SNOMED CT provides markedly better coverage for drug classification than MED-RT does. Overall, SNOMED CT classifies 1391 drugs for MOA and 2435 for CHEM, compared to 820 and 1267 for MED-RT, respectively. In other words, SNOMED CT classifies 70% more drugs for MOA and over 90% more for CHEM.

For the 3239 drugs in common between MED-RT and SNOMED CT (after normalization), the proportion of drugs classified by both systems is 36% (553/1536) for MOA and 14% (404/2942) from CHEM. As shown in Table 4 and Table 5, the proportion of drug-class membership associations provided solely by SNOMED CT (47% (721/1536) for MOA and 57% (1687/2942) for CHEM) is larger than that provided solely by MED-RT (17% (262/1536) for MOA and 29% (851/2942) for CHEM).

**Interpretation.** Coverage is one of the most important desirable characteristics for a drug classification system [25]. The coverage provided by SNOMED CT is largely superior to that of MED-RT. From this perspective, SNOMED CT is a better choice than MED-RT for defining value sets for medication allergy, which is consistent with the recent change in recommendation from the 2019 Interoperability Standards Advisory publication.

Regarding the other desiderata [25], notable differences between MED-RT and SNOMED CT include public availability (restricted to users from member countries of SNOMED International for SNOMED CT). Granularity differences have been discussed earlier (number of classes). Both MED-RT and SNOMED CT are developed by stable institutions and regularly updated. Both have a hierarchical structure, with drugs attached to various levels of the hierarchy, not only leaf classes.

Of note, from a practical perspective, one major difference between MED-RT and SNOMED CT is that MED-RT makes it clear which entities are drug classes (represented in reference to the MOA or MeSH vocabulary) vs. drugs (represented in reference to RxNorm). In contrast, in SNOMED CT, both drug classes and drugs are hierarchically related entities in the Medicinal products hierarchy and indirect cues, such as the semantic tag, are necessary to distinguish between the two, which makes it unnecessarily difficult to extract drug-class membership information.

**B. Limitations and future work**

This preliminary investigation is limited to two types of drug classes, MOA and CHEM. In future work, we will investigate therapeutic role classes (as this information becomes available in SNOMED CT).

This preliminary investigation does not explore the impact of drug classification on clinical drugs (only ingredients), although clinical drugs are important to consider for clinical decision support and clinical quality measurement (in the context of electronic health records). However, the extension to clinical drugs should be straightforward. In MED-RT, RxNorm will provide associations between ingredients and clinical drugs. In SNOMED CT, the association can be derived from either RxNorm or SNOMED CT itself.

This preliminary investigation is purely quantitative and does not attempt to align drug classes between MED-RT and SNOMED CT. For example, assessing whether the class monoamine oxidase inhibitors contains the same drug members in MED-RT and SNOMED CT is out of scope, and so is the comparison of, say MOA classes in MED-RT and SNOMED CT, for the drug Isocarboxazid. We intend to perform this evaluation in future work.

**VI. Conclusions**

We compared drug classes between MED-RT and SNOMED CT and explored whether SNOMED CT classes could form the basis for drug value sets, and be a potential
alternative to MED-RT. Ours is the first investigation of these two drug classification systems since they have undergone important changes in recent years. Our preliminary investigation shows that the coverage provided by SNOMED CT compares favorably to MED-RT’s for MOA and CHEM classes. Except for open access, no other characteristic of MED-RT makes it superior to SNOMED CT for drug classification.

ACKNOWLEDGMENT

This work was supported by the Intramural Research Program of the NIH, National Library of Medicine.

REFERENCES


