

The New SNOMED CT International Medicinal Product Model

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

Julie James
Blue Wave Informatics LLP
Exeter, United Kingdom
Julie_James@BlueWaveInformatics.co.uk

Abstract— Objectives. To present the new SNOMED CT international medicinal product model. **Methods.** We present the main elements of the model, with focus on types of entities and their interrelations, definitional attributes for clinical drugs, and categories of groupers. **Results.** We present the status of implementation as of July 2018 and illustrate differences between the original and new models through an example. **Conclusions.** Benefits of the new medicinal product model include comprehensive representation of clinical drugs, logical definitions with necessary and sufficient conditions for all medicinal product entities, better high-level organization through distinct categories of groupers, and compliance with international standards.

Keywords—SNOMED CT; Medicinal products; Knowledge representation.

I. INTRODUCTION

The representation of medications in SNOMED CT [1] has long been suboptimal for reasons that are briefly analyzed below. Over the past ten years, several working groups of SNOMED International have worked at addressing various issues related to the representation of medicinal products and substances, including:

- Which levels of granularity should be represented in the international release of SNOMED CT? (e.g., do branded drugs belong in the international release?)
- How to optimally represent drug classes? (i.e., at the substance level vs. product level?)

These questions have been debated in the community, but only recently has SNOMED CT published a new model for the representation of medicinal products in the international release and engaged in retrofitting its content to this model.

The main objective of this work is to present the new SNOMED CT international medicinal product model available in the July 2018 version of SNOMED CT – a major revision of the original model, with focus on the types of entities defined and their interrelations. In addition, we discuss the benefits of the new medicinal product model for the SNOMED CT user community.

II. BACKGROUND

To set the stage for the new SNOMED CT international medicinal product model, we briefly discuss the limitations of

the original model. We also introduce the use cases the new model is designed to support, and the main international standards it intends to be compliant with.

A. Limitations of the original medicinal product model

While there is abundant literature about quality assurance issues in SNOMED CT (e.g., [2]), little work has been devoted specifically to medications. [3] have compared the hierarchical organization of medications in SNOMED CT to the National Drug File-Reference Terminology (NDF-RT), between which they reported differences. Other researchers have studied SNOMED CT medications, but mostly from the perspective of applications (e.g., clinical decision support [4] and allergy encoding [5]). However, none of these groups have provided any recommendations for a better representation of medications in SNOMED CT.

The limitations of the original medicinal product model were essentially analyzed by the SNOMED International Drug Model Working Group and the community of practice. Three main categories of issues were identified.

Incomplete representation of clinical drugs. Clinical drugs were mostly represented as primitive concepts (i.e., without necessary and sufficient conditions), and the curation of the hierarchical relations among medication entities had to be performed manually. The representation was incomplete in part because strength had not been represented (See Fig. 1 for an example). As a consequence, SNOMED CT lacked the ability to compute equivalence among clinical drugs and therefore provided limited support for interoperability among drug terminologies based on the same model as SNOMED CT.

Wrong inferences due to therapeutic role groupers. Some groupers, such as drug classes based on chemical structure, represent definitional knowledge, while others, such as drugs classes based on therapeutic roles mostly represent assertional knowledge (i.e., may not be universally valid). *Ad hoc* solutions were required to mitigate wrong inferences. (See the Discussion section below for an example).

Limited compliance with international standards. Issues include lack of standardization of dose forms (e.g., SNOMED CT used both “enteric coated” and “gastro-resistant”, which are synonymous), lack of standardization of strength for liquids (presentation strength vs. concentration strength), and lack of support for a closed world perspective on clinical drugs (i.e.,

stating exactly which active ingredients a clinical drug contains vs. at least which ingredients it contains).

B. Use cases

Unlike many aspects of clinical medicine, description of medications tends to be country-specific. For example, brand names, formulations and strengths of active ingredient substances may all differ across countries. Moreover, national regulatory agencies may have different criteria for authorizing medicinal products, and all medicinal products can only be made available for sale or supply when they have an authorization of some kind. For these reasons, many countries have developed their own medicinal product terminologies (e.g., RxNorm [6] in the U.S, the NHS Dictionary of medicines and devices (dm+d) [7] in the U.K, the Australian Medicines Terminology (AMT) [8] in Australia, and the G-Standaard [9] in the Netherlands). Although SNOMED CT is the largest clinical terminology in the world, it would be difficult for its developers to maintain a medication formulary for the entire world and support a large variety of use cases, from e-prescribing to clinical analytics.

For this reason, SNOMED International has decided to focus on four major use cases:

- To facilitate international interoperability of medication concepts (e.g., for use in patient summaries and for cross-border care)
- To provide a strong foundation for member countries to develop their national medicinal product terminology (e.g., by adding package and branded product information)
- To support medication analytics for research purposes
- To support the development of international medication decision support, such as allergy checking and duplicate therapy checking

Additionally, since SNOMED CT is based on description logic (DL), emphasis was put on defining medication concepts with necessary and sufficient conditions, so that a DL classifier could be used for checking consistency and automatically inferring a subclass hierarchy among products.

C. International standards

In 2012, the health informatics technical committee in the International Standards Organization (ISO), under the auspices of Working Group 6, Pharmacy and Medications, published ISO 11615: Health informatics -- Identification of medicinal products -- Data elements and structures for the unique identification and exchange of regulated medicinal product information. This was updated last year (2017) with minimal changes to its core structures and definitions. This is the core of a suite of standards known as IDMP: the Identification of Medicinal Products [10], with supporting standards in the suite describing supporting concepts such as substances, dose forms, routes of administration, units of presentation and units of measure. ISO 11615 provides a conceptual model and definition of the key data items that are needed to describe a medicinal product as it is authorized by a regulatory agency, in terms of its pharmaceutical composition, its main clinical characteristics, its supporting authorization and key information from its

manufacturing processes. ISO:11615 therefore describes three types of things that are given global identifiers in the regulatory domain (the medicinal product, the packaged medicinal product and the pharmaceutical product).

Since all medicinal product terminology must be sourced and abstracted from what exists, and the authorized medicinal product information from a regulatory agency within a jurisdiction provides this, it is important that any clinical medicinal product terminology fully understands ISO:11615 and manages its own information structures in harmony with it. In addition, having the structure and population of clinical medicinal product terminology in harmony with that used in the regulatory environment facilitates the flow of information between the two domains of use. This is particularly vital for pharmacovigilance and patient safety, but also for analysis and ongoing clinical research and pharmacoepidemiology.

III. METHODS

In this section, we present the main elements of the SNOMED CT international medicinal product model and discuss its implementation.

A. Defining the new medicinal product model

1) Types and interrelations

The new SNOMED CT international medicinal product model keeps the distinction between substances (chemical or biological entities) and products (manufactured objects). The two are intimately related in that substances are the ingredients of products (e.g., the medicinal product “Product containing atorvastatin” has the substance “Atorvastatin” as its active ingredient).

However, instead of a single type of entity for medicinal products regardless of the level of granularity (“product”), SNOMED CT now has multiple types and distinguishes whether the product contains at least (existential restriction) or only (universal restriction) a given active ingredient substance. Moreover, although SNOMED CT did not introduce new types for substances, it modified the relationship between base ingredients (e.g., atorvastatin) and their salt or ester modifications (e.g., atorvastatin calcium and atorvastatin calcium trihydrate). In the past, base substances subsumed substance modifications, which is incorrect and can lead to wrong inferences (e.g., clinical drugs containing an ingredient modification are distinct from products containing the corresponding base ingredient and should not be subsumed by them). To this end, SNOMED CT introduced a new relationship “is modification of” used in replacement of “isa” between a substance and its modification (e.g., atorvastatin calcium is modification of atorvastatin, instead of atorvastatin calcium isa atorvastatin).

The main type for medications in SNOMED CT is “clinical drug” (combining precise ingredient substance, strength and dose form). In addition to “clinical drug”, SNOMED CT also provides several types for aggregation and navigation purposes, including “medicinal product form” (combining ingredient substance and dose form or dose form group) and “medicinal product” (only specifying the ingredient substance). These two types exist in two flavors, with existential restrictions (subsuming any medication containing a given ingredient

substance or its modification, which is convenient for clinical analytics purposes) or universal restriction (subsuming all medications that only contain this ingredient substance or its modification, which is necessary for decision support and for interoperability of clinical care). Optionally, SNOMED CT also supports the representation of medicinal products defined in reference to ingredient substance modifications (in the universal restriction flavor), to support use cases in which the substance modification is clinically significant (e.g., for many of the glucocorticosteroid preparations).

2) *Definitional attributes for clinical drugs*

Presentation strength. One feature of the new SNOMED CT international medicinal product model is to provide a complete and detailed representation of medications. In the past, clinical drug entities were all primitive classes. While active ingredient substance and dose form were explicitly represented through roles, strength was not. In contrast, the new model provides necessary and sufficient conditions for all clinical drugs. Moreover, to support computation, strength is represented not as a string (e.g., “10 mg”), but as a series of discrete elements (numerator value, numerator unit, denominator value, denominator unit) and presentation strength is represented instead of concentration strength for the majority of products. The practical importance of the presentation strength is to support the distinction among iso-concentration products. For example, the distinction between the following two clinical drugs would be lost if only the concentration strength were recorded: “Product containing precisely norepinephrine (as norepinephrine bitartrate) 2 milligram/2 milliliter conventional release solution for infusion ampule” and “Product containing precisely norepinephrine (as norepinephrine bitartrate) 20 milligram/20 milliliter conventional release solution for infusion ampule”.

Basis of strength substance (BoSS). The BoSS is the substance in reference to which strength is expressed in a clinical drug. The BoSS typically corresponds to the active ingredient substance or the active moiety, but can also be an arbitrary substance selected as the reference. In the new SNOMED CT international medicinal product model, both the active ingredient and the BoSS are explicitly recorded. For example, in “Product containing precisely norepinephrine (as norepinephrine bitartrate) 2 milligram/2 milliliter conventional release solution for infusion ampule”, the active ingredient is “norepinephrine bitartrate”, but the BoSS is “norepinephrine”.

Dose form and unit of presentation. In the past, SNOMED CT only provided a Has dose form attribute. In addition to harmonizing dose forms with the international standard from IDMP (ISO:11239) implemented by the European Directorate for Quality in Medicines (EDQM), the new SNOMED CT international medicinal product model distinguishes between manufactured dose form (e.g., conventional release oral tablet) and unit of presentation (oral solid dose forms, such as tablet; containers, such as vial; and actuations, when the quantity of product released is determined by a metering valve).

3) *Categories of groupers*

Past versions of SNOMED CT have integrated groupers corresponding to different types of medication classes, including classes reflecting a mechanism of action or physiologic effect

(e.g., “Hydroxymethylglutaryl-coenzyme A reductase inhibitor”, “Antiplatelet agent”), a chemical structure (e.g., “Aminoglycosides”), an intended site of administration (e.g., “Otic dosage form product”), and a therapeutic role (e.g., “Anti glaucoma preparation”, “Antimalarial agent”, “Alopecia preparation”). These entities were just primitive classes under which the appropriate drug products were arranged manually. This ontology maintenance approach was both labor-intensive and error-prone, and did not benefit from the support of DL classifiers.

Recently, the new SNOMED CT international medicinal product model reorganized these groupers, such that all medicinal products, including the groupers, could be defined with necessary and sufficient conditions. To this end, groupers were clustered into four main categories:

Based on disposition (e.g., “Product containing 3-hydroxy-3-methylglutaryl-coenzyme A reductase inhibitor” [product grouper], defined as having active ingredient “Substance with 3-hydroxy-3-methylglutaryl-coenzyme A reductase inhibitor mechanism of action” [substance grouper], itself defined as having disposition “3-hydroxy-3-methylglutaryl-coenzyme A reductase inhibitor” [disposition]. Of note, individual substances exhibiting this disposition are (manually) subsumed by the substance grouper, while products having these individual substances as their active ingredient are automatically classified under the product grouper.

Based on chemical structure (e.g., “Product containing aminoglycoside” [product grouper], defined as having as its active ingredient the substance “Aminoglycosides” [substance grouper]. Of note, individual substances exhibiting this chemical structure are (manually) subsumed by the substance grouper, while products having these individual substances as their active ingredient are automatically classified under the product grouper.

Based on intended site of administration (e.g., “Product manufactured as parenteral dosage form” [product grouper], defined as having manufactured dose form “Parenteral dose form” [dose form grouper], itself defined as having intended site “Parenteral” [intended site]. Of note, individual parenteral dose forms are (manually) subsumed by the dose form grouper, while products having these individual manufactured dose forms are automatically classified under the product grouper.

Based on therapeutic role. In some cases, therapeutic role and mechanism of action are intimately related and both are definitional properties for the drug (e.g., “Product containing antimalarial”). In this case, the therapeutic role is simply treated as a disposition, i.e., a definitional characteristic. In other cases, the therapeutic role reflects the culture and practice of healthcare and may not be universally acknowledged by all regulatory agencies. Here, the therapeutic role is a non-definitional characteristic. The organization of medications under therapeutic roles in SNOMED CT is still under development.

B. *Implementation*

New definitions. As mentioned earlier, one major difference from the original medicinal product model is that all medicinal products are now defined with necessary and sufficient conditions in the new SNOMED CT international medicinal

product model, which enables the SNOMED CT developers to detect inconsistencies and infer hierarchies automatically.

New types. The implementation of the medicinal product model has required the creation of several new types of medications, initially all represented as “product”. These types include Clinical drug, Medicinal product form, and Medicinal product. Similarly, types had to be created for the new definitional characteristics (e.g., Disposition, Dose form, Intended site). Finally, the Substance hierarchy was significantly revised to accommodate these changes.

New attributes. Similarly, new attributes had to be created to support the definition of medications in relation to these new types. These attributes include Has disposition, Has manufactured dose form, Has active ingredient, Has precise active ingredient, Has basis of strength substance, and Has presentation strength (for unit and value, and for numerator and denominator).

The vast majority of the changes described here will be available as part of the July 2018 release of SNOMED CT (international release).

IV. RESULTS

A. Quantitative results

Table 1 shows the approximate number of concepts for the various types of entities discussed above projected to be available in the July 2018 international release of SNOMED CT, along with examples for each type.

TABLE 1. TYPES OF ENTITIES USED FOR THE REPRESENTATION OF MEDICINAL PRODUCTS IN SNOMED CT

Type	# concepts (approx.)
Medicinal product 398731002 <i>Product containing sulfamethoxazole and trimethoprim (medicinal product)</i>	3750
Medicinal product form 392419009 <i>Product containing ofloxacin in ocular dosage form (medicinal product form)</i>	2100
Clinical drug 317335000 <i>Product containing precisely esomeprazole 20 milligram/1 each conventional release oral tablet (clinical drug)</i>	4850
Disposition 768681000 <i>Antibacterial (disposition)</i>	300
Dose form 385219001 <i>Conventional release solution for injection (dose form)</i>	290
Intended site 738989005 <i>Vaginal (intended site)</i>	25
Groupers based on disposition or structure Disposition: 703645005 <i>Product containing B-Raf inhibitor (product)</i> Structure: 763875007 <i>Product containing sulfonamide (product)</i>	560
Groupers based on intended site of administration 440131009 <i>Product manufactured as oral dosage form (product)</i>	20
Groupers based on therapeutic role 57952007 <i>Antilipemic agent (product)</i>	378

B. Extended example

In this example, we contrast the representation of the clinical drug “Amlodipine 10 mg and atorvastatin 10 mg oral tablet” in the original SNOMED CT medicinal product model (circa 2015) and in the new SNOMED CT international medicinal product model (preview of the July 2018 release).

1) Axioms

As shown in Fig. 1, the original model represented clinical drugs as primitive concepts (i.e., only with necessary conditions and *SubClassOf* axioms), while all clinical drugs (and most medication entities) are fully defined (with necessary and sufficient conditions and *EquivalentTo* axioms) in the new model. In other words, the original model only provided a partial representation of the clinical drugs (with no representation of the strength, for example), while the new model now represents strength. In practice, it also means that the product hierarchy is entirely inferred by the DL classifier in the new model, whereas it was created manually in the original model.

Moreover, in addition to strength, the new model offers a more complete representation of clinical drugs, including unit of presentation and basis of strength substance.

Of note, the classic approach to restricting a multi-ingredient drug to specific ingredients is through closure axioms, for example, the last axiom (underlined) in: *has_ingredient SOME amlodipine and has_ingredient SOME atorvastatin and has_ingredient ONLY (amlodipine or atorvastatin)*. However, the flavor of DL used in SNOMED CT, \mathcal{EL}^{++} , does not support universal restrictions (needed to express *has_ingredient ONLY* ...). The workaround implemented in SNOMED CT consists in adding an axiom for the count of active ingredients instead, to distinguish this clinical drug from other clinical drugs containing additional ingredients. While somewhat unorthodox, this workaround is effective for the representation of medicinal products.

2) Subclass hierarchy

As mentioned earlier, the subclass hierarchy consisted of stated *SubClassOf* relations in the original model, whereas it is entirely inferred in the new model. Moreover, as shown in Fig. 2A, different types of groupers were entangled in the original model (e.g., for dispositions and therapeutic roles). While dispositions are definitional properties (e.g., atorvastatin has always the disposition of being an inhibitor of a specific enzyme), therapeutic roles are not (i.e., atorvastatin may not be recognized as an antilipemic agent by all regulatory agencies). In the new model (Fig. 2B), therapeutic roles are no longer part of the subclass hierarchy. In the past, having therapeutic roles in the subclass hierarchy could lead to incorrect inferences (e.g., timolol is an antiglaucoma drug only when administered in eye drops), and *ad hoc* solutions were required to avoid such wrong inferences (e.g., link the therapeutic role to the medicinal product form rather than the medicinal product).

V. DISCUSSION

A. Benefits of the new medicinal product model

The new SNOMED CT international medicinal product model provides a more comprehensive representation of clinical drugs (e.g., explicit representation of strength, unit of

presentation, basis of strength substance). Providing necessary and sufficient conditions for each clinical drug not only supports automatic inference of the subclass hierarchy, but also supports the identification of equivalent classes. The latter will be required for interoperability between the international release of SNOMED CT and national medicinal product extensions, in which clinical drugs will be defined, along with country-specific medication entities, such as packages and branded products. In other words, the new model provides a strong foundation for national medication extensions. Member countries without a medicinal product extension can enrich the international model with the information required for their use cases (e.g., packaging information and brand names for e-prescribing purposes). Member countries that already have medicinal product terminologies (e.g., RxNorm, NHS dm+d, AMT) should be able to align the clinical drugs in these terminologies to SNOMED CT's and will benefit from the high-level organization in SNOMED CT to organize their content. Once national extensions are aligned with SNOMED CT, medication entities in various extensions are *de facto* interoperable through their shared relations to SNOMED CT.

The distinction introduced in the new model among the various types of groupers (related to dispositions, to chemical structure, to intended site, and to therapeutic roles) not only facilitates the organization and navigation of drug entities, but also avoids wrong inferences without having to use *ad hoc* solutions (such as linking the medicinal product form rather than the medicinal product to a therapeutic role). In other words, the new model offers better distinction between definitional and assertional knowledge.

Finally, the new model is compliant with international standards, such as IDMP, which is required to support pharmacovigilance beyond national boundaries. Features of the new model required for compliance with IDMP include providing an exact list of ingredients for clinical drugs (i.e., "Product containing ONLY..."), providing the manufactured dose form (not only the administrable dose form) and presentation strength (not only concentration strength). Related to the representation of strength, the substance in reference to which strength is expressed (i.e., the basis of strength substance) must also be explicitly represented. Similarly for the unit of presentation (in reference to the EDQM standard). In addition, the SNOMED CT model extends and complements the IDMP model in a compatible and harmonious way, providing international identifiers for types of representation of medicines for clinical use not required in the regulatory domain: in particular the Clinical Drug, in its representation and identification of the product using its manufactured dose form. This is of significant value for international interoperability of medication information for individual patients.

B. Limitations and future work

As mentioned earlier, the absence of universal restrictions in \mathcal{EL}^{++} , the flavor of description logic used by SNOMED CT, requires a workaround to ensure that a clinical drug does not have any other active ingredients than the active ingredients listed as part of the definition. This is achieved through the addition of an axiom for the count of active ingredients in the clinical drug. For example, as shown in Fig. 1, "Amlodipine 10

mg and atorvastatin 10 mg oral tablet" is made distinct from (hypothetical) clinical drugs containing amlodipine, atorvastatin and another ingredient, because it is also specified that the count of active ingredients of this clinical drug is 2 – the count would be 3 for the other drug. When SNOMED CT transitions to a more expressive DL, the use of universal restrictions will make it possible to express proper closure axioms, rendering this ingredient counting axiom unnecessary.

While a majority of oral solid dose form drugs will be available if the new model in the upcoming international release (July 2018), the conversion effort will continue over the next few releases for liquids (oral solutions and parenteral drugs) and topical drugs.

While SNOMED CT is expected to provide a strong foundation for the development or anchoring of national drug extensions, this remains to be evaluated in practice. No national drug extensions have been developed based on the international model yet. Following earlier attempts (e.g., [11]), efforts are underway to harmonize existing drug terminologies (RxNorm, dm+d, AMT) with SNOMED CT, but no metrics of success have been reported yet. Assuming these efforts succeed, the national terminologies harmonized with SNOMED CT could provide a mechanism for the long-term maintenance of the SNOMED CT drug content, i.e., new clinical drugs in national extensions could be automatically promoted to the international release of SNOMED CT, and ancillary entities (medicinal product and medicinal product form) could be automatically derived from the definition of clinical drugs.

VI. CONCLUSIONS

The new SNOMED CT international medicinal product model represents a major revision of the original model in use over the past ten years. Benefits of the new model include comprehensive representation of clinical drugs, logical definitions with necessary and sufficient conditions for all medicinal product entities, better high-level organization through distinct categories of groupers, and compliance with international standards.

ACKNOWLEDGMENT

This work was supported in part by the Intramural Research Program of the NIH, National Library of Medicine. The authors thank the contributors to the SNOMED International Drug Model Working Group for useful discussion, particularly Jim Case, Yongsheng Gao, Emma Melhuish, Toni Morrison, Guillermo Reynoso, and Phuong Skovgaard. Special thanks to Yongsheng Gao and Kai Kewley for sharing the OWL preview of SNOMED CT used in the Protégé screenshots.

REFERENCES

- [1] SNOMED International, "SNOMED CT," 2018.
- [2] A.L. Rector, et al., "Getting the foot out of the pelvis: modeling problems affecting use of SNOMED CT hierarchies in practical applications," *J Am Med Inform Assoc*, vol. 18, no. 4, 2011, pp. 432-440.
- [3] A. Zakharchenko and J. Geller, "Auditing of SNOMED CT's Hierarchical Structure using the National Drug File - Reference Terminology," *Stud Health Technol Inform*, vol. 210, 2015, pp. 130-134.

- [4] K. Greibe, "Development of a SNOMED CT based national medication decision support system," *Stud Health Technol Inform*, vol. 192, 2013, pp. 1147.
- [5] F.R. Goss, et al., "Evaluating standard terminologies for encoding allergy information," *J Am Med Inform Assoc*, vol. 20, no. 5, 2013, pp. 969-979.
- [6] U.S. National Library of Medicine, "RxNorm," 2018.
- [7] NHS Business Service Authority, "Dictionary of medicines and devices (dm+d)," 2018.
- [8] Australian Digital Health Agency, "Australian Medicines Terminology," 2018.
- [9] Royal Dutch Pharmacists Association (KNMP), "The G-Standaard: the medicines standard in healthcare," 2018.
- [10] European Medicines Agency, "Introduction to ISO Identification of Medicinal Products," 2018.
- [11] S.J. McBride, et al., "Using Australian medicines terminology (AMT) and SNOMED CT-AU to better support clinical research," *Stud Health Technol Inform*, vol. 178, 2012, pp. 144-149.

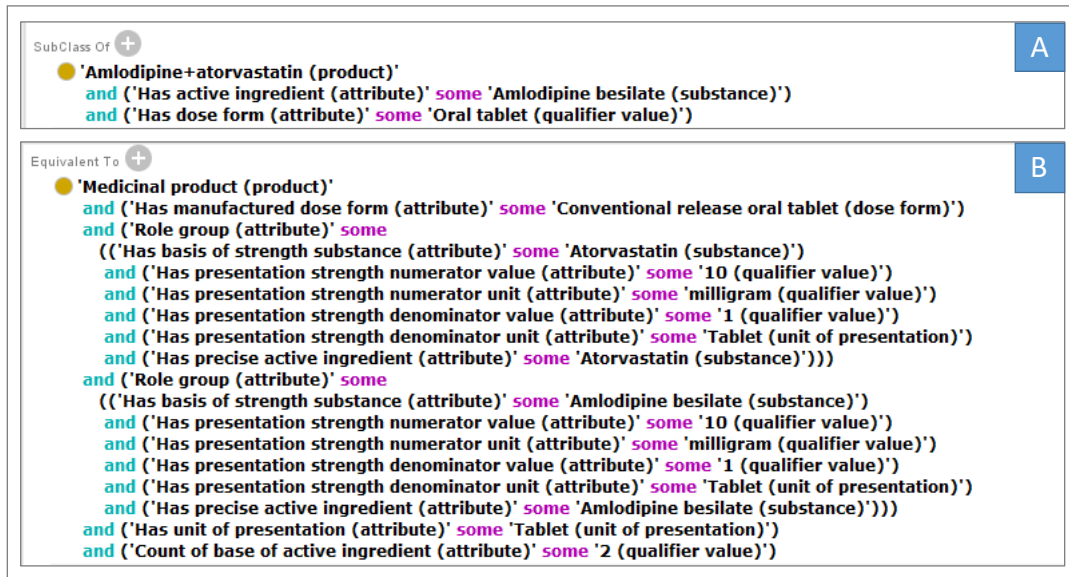


Fig. 1. Class description for the clinical drug “Amlodipine 10 mg and atorvastatin 10 mg oral tablet” in the original (A) and new (B) SNOMED CT international medicinal product model

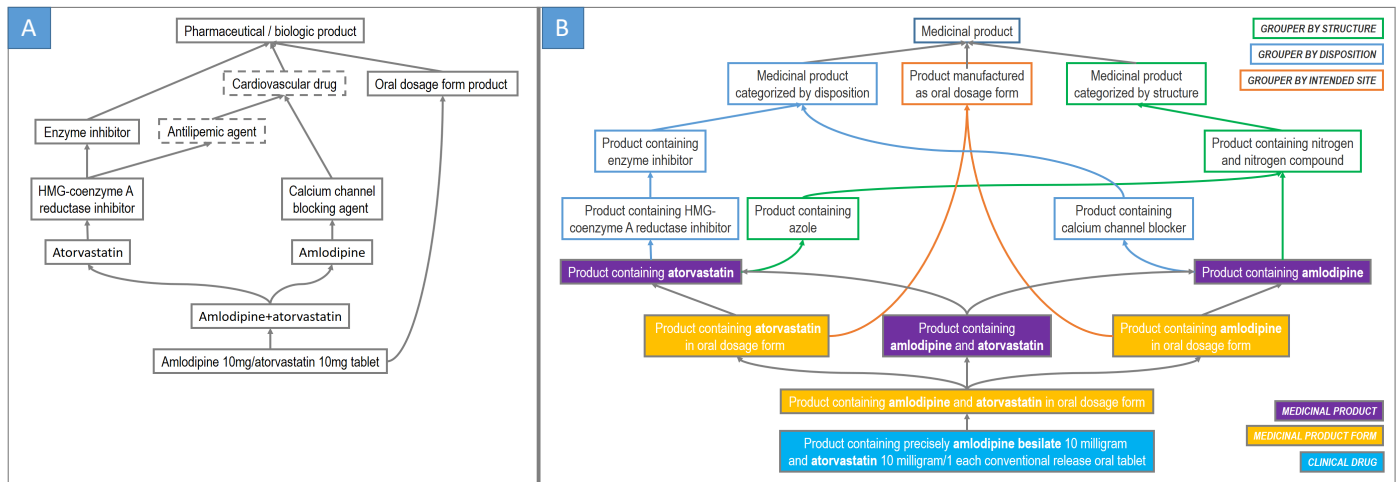


Fig. 2. Subclass hierarchy for the clinical drug “Amlodipine 10 mg and atorvastatin 10 mg oral tablet” in the original (A) and new (B) SNOMED CT international medicinal product model