Collaboration across Disciplines for the Health of People, Animals and Ecosystems
L. Stoicu-Tivadar et al. (Eds.)

doi:10.3233/SHTI241097

The Creation of Intensional Medication Lists Using the NHS Dictionary of Medicines and Devices

Gavin JAMIE[a,1], Rachel BYFORD[a], Rashmi WIMALARATNA[a] and Simon DE LUSIGNAN [a]
[a] *Nuffield Department of Primary Care Health Sciences, University of Oxford, UK*
ORCiD ID: Gavin JAMIE https://orcid.org/0000-0001-9147-7784, Rachel BYFORD https://orcid.org/0000-0002-4792-8995, Rashmi WIMALARATNA https://orcid.org/0009-0000-3244-9679, Simon DE LUSIGNAN https://orcid.org/0000-0002-8553-2641

Abstract. The identification of medications prescribed to patients in routinely collected health records is an important part of the identification of cohorts for surveillance and research. Preparations available for prescription can change frequently and this presents challenges to the maintenance of extensional or "flat lists" of medications, particularly in ongoing studies such as disease surveillance. The NHS publishes a Dictionary of Medicines and Devices weekly, listing almost all the medications available in the UK as an extension to the UK edition of SNOMED CT. We developed a method of creating intensional specifications of medications using specified active ingredients and the form of the medication. The specifications can be expressed using the SNOMED CT Expression Constraint Language, and can be used to form a library which may be used across multiple projects. We have developed intensional definitions of medication groups for all drugs likely to be used in primary care. We have shown that these can be shared as FHIR valuesets using the NHS Terminology Server. Here we show examples of expressions about medications used for neuropathic pain. We have created expressions which improve the specificity of the extraction by filtering on the form and number of ingredients.

Keywords. SNOMED CT, dm+d, Medication, Valuesets

1. Introduction

1.1. The need for medication valuesets

The RCGP (Royal College of General Practitioners) Research Surveillance Centre (RSC) is one of Europe's oldest sentinel networks and has been collecting information about the incidence of acute infective disease since 1967 [1]. The ORCHID database contains information from nearly 2000 primary care practice electronic health record (EHR) systems throughout England coded using SNOMED CT. There is a frequent need to

[1] Corresponding author: Gavin Jamie, Nuffield Department of Primary Care Health Sciences, University of Oxford, United Kingdom; E-mail:gavin.jamie@phc.ox.ac.uk.

identify cohorts of patients using prescribed medication records. Examples of this use include identification of patients who may be immunosuppressed or have a diagnosis such as diabetes or dementia. Medication use may also signify an endpoint such as the use of antibiotics or steroids in respiratory infection.

Intensional valuesets have been shown to be faster and more accurate to develop as well as requiring less time to maintain than static, or extensional valuesets [2]. This is particularly valuable in disease surveillance where data is extracted over a period of years. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement requires that these should be published alongside research outputs although this has been found to happen in only a minority of cases [3,4].

1.2. The dictionary of medications and devices

Medications are presented in data extracts using codes from the NHS Dictionary of Medications and Devices (dm+d) which is updated weekly [5]. Whilst the use of a new drug in primary care is not a particularly common event, pharmaceutical presentations change on a regular basis, particularly where generic manufacturers start and stop production of medications which are no longer under patent. Static lists of medications can rapidly become out of date and would need to be refreshed on a regular basis to take account of new products. This limits the ability to share valuesets between projects and would require almost constant review in surveillance programmes.

Prescriptions in primary care are presented as either Virtual Medical Products (VMP), such as “paracetamol 500mg tablets” or Actual Medical Products (AMP) e.g. “Panadol ActiFast 500mg tablets”. Dispensing data uses Actual Medical Product Packs (AMPP) or Virtual Medical Product Packs (VMPP) e.g. 28 Paracetamol 500mg tablets. The dictionary is presented as an extension to the UK Edition of SNOMED CT with all of the features of SNOMED implemented. There is the usual “is a” hierarchy and each AMP has a VMP. Above that sits a Virtual Therapeutic Moiety, such as “paracetamol”.

There are several attributes associated with each concept in dm+d. The most significant are the active ingredients and the dispensed dose form. Active ingredients are coded using the substance hierarchy in SNOMED and the form using concepts under the Pharmaceutical Dose Form qualifier value. The concentration, or amount of active ingredient in each unit (e.g. pill) is not recorded in the attributes in dm+d. A VMPP or AMPP will have a VMP or AMP as an attribute.

SNOMED CT Expression Constraint Language (ECL) allows a formal specification based on concept attributes [6]. This specification can then be run on any terminology server which has been loaded with the dm+d. The NHS Terminology Server is such a server. It is the official valueset server for the NHS England and uses the commercial Ontoserver software with a Fast Healthcare Interoperability Resources (FHIR) compatible application programming interface [7].

2. Methods

We developed a tool to allow the construction of a valueset containing AMPs and VMPs based on the active ingredients. There were further options to restrict the forms of each

ingredient. This was used particularly where a drug had indications for both systemic and local delivery. Examples include oral and topical corticosteroids or timolol eye drops and tablets. Restrictions could be to specific physical forms or broad categories of routes of administration such as oral, topical or injections.

A member of our clinical team would search for substances listed under the "drug or medicament" section of SNOMED CT. We chose to include drugs even when they had been marked as inactive in dm+d as these codes may continue to appear in historic extracts from practice EHRs. Equally some preparations (AMP/VMPs) may no longer be active but should be included.

Multiple ingredients could be included in one valueset which would be combined such that a product with any of the ingredients would be included in the output. In this way we could create general valuesets such as "penicillins" or "antihypertensive drugs".

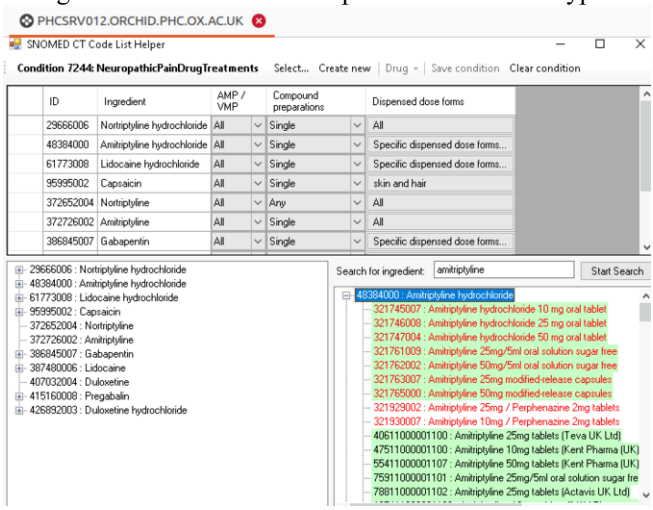

Figure 1. The helper tool that has been developed to allow clinicians to create medication valuesets

We have used this method within the research group several hundred times. We have developed a method to translate these to a FHIR valueset using the SNOMED CT Expression Constraint Language (ECL) that, with the addition of metadata, can be uploaded to a terminology server. We have used the "history" modification to include medications that are equivalent to the ones that we have selected but have been made inactive. The stages of clinical curation and checking are similar to those we use for non-drug valuesets [8].

3. Results

We present criteria to identify prescriptions issued for medications which may be used to treat neuropathic pain in English primary care data in the ORCHID database. These have been converted to SNOMED CT expression constraint language and uploaded to the NHS Terminology Server.

Our specification contains ten different ingredients, with seven of these appearing in medications currently listed in the dm+d. The three other ingredients are chemical variations of the others (e.g. amitriptyline versus amitriptyline hydrochloride). All of these were restricted to preparations including a single ingredient only in order to

increase the specificity of the list. Combination medications tended to have indications other than neuropathic pain e.g. lidocaine and hydrocortisone ointment is licenced and used for haemorrhoids and anal fissure. Five preparations were also restricted by form, again to remove medications used for other causes. Lidocaine, for example, is a commonly used, injected, local anaesthetic for minor surgical procedures. This specification was run on the NHS Terminology Server and produced a list of 1060 medications and preparations.

The expression constraint language for single ingredient topical lidocaine is below. We have added colours to illustrate the clauses. Similar expressions are used for each of the ingredients.

```
<10363801000001108 | Virtual medicinal product (product) | {{+ HISTORY-MOD }}
        :10362801000001104 |Has specific active ingredient (attribute)| = 61773008 |
Lidocaine hydrochloride(substance) |,
        10362901000001105 |Has dispensed dose form (attribute)| = (421628006 |
        Conventional release cutaneous cream (dose form) | OR 385100002 | Gel (basic dose
        form) | OR 385101003 | Ointment (basic dose form) | ),
        [1..1] 10362801000001104 |Has specific active ingredient (attribute)| = < 105590001
    |Substance|
```

The four parts of the expression are: firstly, that this should be a descendant of virtual medicinal product, that is either a virtual or actual product. This part also contains the term to include inactive but equivalent terms. In orange we specify that it must contain lidocaine hydrochloride. In green we restrict the forms of the medication to cream, gel or ointment. This phrase can be omitted if we wish to include all forms of a medication.

The final, blue, phrase of the expression states that there should be only one active ingredient. This will eliminate compound preparations. This phrase can be omitted if all preparations are to be included. We can also include only compound preparations by starting the phrase with "[2..*]".

4. Discussion

4.1. Strengths and Limitations

Running intensional specifications on a terminology server allows updates to medicines to be automatically included in results. The DM+D is much more dynamic than the International Edition and UK Extension of SNOMED CT and the advantages compared to static, extensional lists are more apparent here.

The specification is very flexible, and we have been able to use in for most cases. There have been a small number of cases where the model could not be applied. Some medications do not have a specified active ingredient. This is common in dermatological emollients and lubricating eye drops where the ingredients have a physical rather than chemical action. The majority of devices do not have an active ingredient attribute for the same reason. We also had some difficulty where medications may have different indicators depending on the dose – this occurred most commonly in sex hormones. For example, norethisterone can be used as a contraceptive at 350mcg daily. It may also be used to reduce abnormal uterine bleeding at a dose of 5mg three times a day. We could not distinguish between these doses using the defining attributes of dm+d.

4.2. Possible further development

Although we have not done so, it would be possible to use the "is a" hierarchy to select ingredients containing any of a group of similar chemicals e.g. opioids, using a single expression. This could simplify the definition and reduce maintenance further. This technique is, in practice, more limited than it first appears as SNOMED groups drugs more by chemical structure than therapeutic group. Novel therapeutic agents, particularly those used in primary care, are sufficiently uncommon that they add little to the maintenance requirements.

5. Conclusions

We have been able to use SNOMED Expression Constraint Language to create intensional specifications for medications in the DM+D and have been actively using these. We have shown that these can be run and shared using the NHS Terminology Server and once published, they can be used by anyone searching NHS data. Maintenance of these valuesets is anticipated to be much simpler than keeping extensional lists up to date. We plan to increase the valuesets published alongside research papers as part of our commitment to open science.

Acknowledgements

John Williams devised the medication list curation which was written by RB.

References

[1] de Lusignan S, Correa A, Smith GE, Yonova I, Pebody R, Ferreira F, Elliot AJ, Fleming D. RCGP Research and Surveillance Centre: 50 years' surveillance of influenza, infections, and respiratory conditions. Br J Gen Pract. 2017 Oct;67(663):440-441. doi: 10.3399/bjgp17X692645.

[2] Chu L, Kannan V, Basit MA, et al. SNOMED CT concept hierarchies for computable clinical phenotypes from electronic health record data: Comparison of intensional versus extensional value sets. . 2019;7(1):e11487. https://www.ncbi.nlm.nih.gov/pubmed/30664458. doi: 10.2196/11487.

[3] Benchimol EI, Smeeth L, Guttmann A, et al. The REporting of studies conducted using observational routinely-collected health data (RECORD) statement. PLOS Medicine. 2015;12(10):e1-e1001885. https://www.ncbi.nlm.nih.gov/pubmed/26440803. doi: 10.1371/journal.pmed.1001885

[4] Massen GM, Stone PW, Kwok HHY, et al. Review of codelists used to define hypertension in electronic health records and development of a codelist for research. Open Heart. 2024;11(1):e002640. http://dx.doi.org/10.1136/openhrt-2024-002640. doi: 10.1136/openhrt-2024-002640.

[5] NHS England. Dictionary of medicines and devices (dm+d). https://www.nhsbsa.nhs.uk/pharmacies-gp-practices-and-appliance-contractors/dictionary-medicines-and-devices-dmd. Accessed 18 June, 2024.

[6] International Health Terminology Standards Development Organisation. Expression constraint language - specification and guide. https://confluence.ihtsdotools.org/display/DOCECL. Accessed 25th July, 2023.

[7] NHS England. Terminology servers - navigation and service status. https://ontology.nhs.uk/. Accessed 11 June 2024

[8] Jamie G, Elson W, Kar D, et al. Phenotype execution and modelling architecture to support disease surveillance and real-world evidence studies: English sentinel network evaluation. JAMIA open. 2024;7(2):ooae034. https://www.ncbi.nlm.nih.gov/pubmed/38737141. doi: 10.1093/jamiaopen/ooae034