Evaluation of the VA/KP Problem List Subset of SNOMED as a Clinical Terminology for Electronic Prescription Clinical Decision Support

Surendranath Mantena, MD\textsuperscript{1,2} and Gunther Schadow, MD\textsuperscript{1,2,3}

Regenstrief Institute\textsuperscript{1}, Indiana University School of Medicine\textsuperscript{2},
Indiana University School of Informatics\textsuperscript{3}, Indianapolis, IN

Abstract

A standardized terminology for medical indications is essential for building e-prescription applications with decision support. The FDA has adopted the Veteran Administration and Kaiser Permanente (VA/KP) Problem List Subset of SNOMED as the terminology to represent indications in electronic labels. In this paper, we evaluate the ability of this subset to represent the text phrases extracted from a medication decision support system and the indications section of existing labels. We compiled a test set of 1265 distinct indication phrases and mapped them to (1) UMLS, (2) Entire SNOMED, (3) All Precoordinated concepts from the “Clinical Finding” hierarchy of SNOMED, and (4) VA/KP Subset. 95% of the phrases mapped to concepts in UMLS, 90.3% to SNOMED, 79.5% to SNOMED Precoordinated and 71.1% mapped completely or partially to concepts in the VA/KP subset. Our study suggests that the VA/KP Subset has significant limitations for coding drug indications; however, when focusing on indications as medical conditions only, the coverage seems more adequate.

Introduction

Computerized physician order entry (CPOE) with clinical decision support (CDS) can improve medication safety and reduce medication-related expenditures\textsuperscript{1}. A study done by Bobb et al. showed that pre-defined order sentences might prevent over 75% of 1,111 dosing errors.\textsuperscript{2} Another study of outpatient prescribing by Gandhi et al determined that default dose and frequency suggestions might have eliminated 42% of prescribing errors and 53% of potential adverse drug events\textsuperscript{3}.

However, for CDS to be effective, adequate expertise must go into defining and representing medical knowledge. A report of the joint clinical decision support group cites a number of barriers impeding the optimal adoption and effectiveness of CDS interventions for medication management including limited functionality, lack of integration to important data in the EMR, uneven standards, and high costs and complexity of implementation and ongoing use. In its recommendations, the group suggests that “enhanced or new standards are required in several areas to facilitate CDS. One such area is terminology where there is a need for convenient, usable, standard dictionaries for medication ordering that support typical usage”.\textsuperscript{4}

In an effort to improve dissemination of knowledge about medications and easy integration of new knowledge into existing CPOE applications, the FDA has adopted a standard known as the Health Level-7 (HL7) Structured Product Labeling (SPL).\textsuperscript{5} Under the regulations that became effective in November 2005, the drug manufacturers are required to submit all drug related prescribing and product information (drug labels or package inserts) to the FDA electronically as XML in the SPL format.\textsuperscript{6,7} In addition, several sections of the label must contain data elements coded using standardized medical terminologies.

The SPL standard is being implemented in several phases. In the first phase, the drug manufacturers were required to submit drug labels in SPL format with free text content with coded data elements for chemical and packaging information only. In the second phase, the implementation of the Physician Labeling Rule (PLR)\textsuperscript{8} mandates that clinical data elements such as indications, contraindications and adverse effects must also be coded using a standardized terminology. The FDA selected the Veterans Health Administration and Kaiser Permanente (VA/KP) Problem List subset of SNOMED as the terminology of choice for this purpose.\textsuperscript{9}

The VA/KP Problem List is a subset of SNOMED containing precoordinated concepts from the “Clinical Finding” hierarchy. It includes only concepts that have been explicitly used or requested for use by Kaiser Permanente or VA physicians. It is available for download from NCI Enterprise Vocabulary Services and can be searched using a free browser called Yatbu. The question arises whether the VA/KP Problem List Subset might have sufficient coverage and precision to adequately code the medical conditions in the drug la-
The purpose of this study is to contribute to the answer in a systematic way.

**Methods**

**Creating a Test Set**
We created a test set by extracting clinical text phrases from the Regenstrief Medical Gopher and from the existing drug labels available from DailyMed [http://dailymed.nlm.nih.gov], merging them, and removing all duplicates.

Gopher is a CPOE application provided by the Regenstrief Institute and used at Wishard hospital for over 20 years. It incorporates several medication-related clinical decision support functions, including allergy checking, basic dosing guidance for individual medications and differential dosing guidance for medications with more than one indication, formulary decision support, drug interaction checking, and medication associated laboratory testing. Dosing guidance is provided in the form of instruction templates (“SIG templates”) which contain text headings of medical conditions relevant to the choice of dosage. We extracted these text headings by initially converting all of the Gopher knowledge base into an XML document and then transforming it using XSLT.

We downloaded all the available 2218 SPL labels from DailyMed as of February 2007. We extracted the clinical text phrases from the drug indications sections using XSLT in two steps: (1) extracting the text lead by the phrases “is/are indicated in/for” (a pattern found to be predominantly used across labels) and (2) manually examining the extracted phrases and deleting all the non clinical text.

**Mapping Text Phrases to Coded Concepts**

To compare the coverage and precision of the VA/KP Problem List Subset we mapped each phrase in our test set to concepts in the Unified Medical Language System (UMLS) using the NLM’s Java-based release of MetaMap Transfer Application (MMTx). MMTx maps arbitrary text to concepts in the Unified Medical Language System (UMLS) although the default MMTx dataset maps to the entire UMLS, it is not possible for a concept to exist in the subset but not in the superset. We therefore created three custom MMTx datasets:

1. SNOMED CT terms in UMLS (henceforth referred to as Entire SNOMED),
2. A subset of SNOMED containing all the pre-coordinated concepts of the SNOMED “Clinical Finding” hierarchy (henceforth referred to as SNOMED Precoordinated), and
3. The VA/KP Problem List subset of SNOMED (henceforth referred to as the VA/KP Subset).

We included all UMLS synonyms so that each included concept had the same synonyms in all datasets. We previously created a web-service wrapper around MMTx which codes phrases in “term processing mode” where it tries to match the entire input string as a single phrase rather than dividing the input into multiple phrases.

The output of the mapping process was an XML document containing the phrase from the test set followed by the mapped concepts from the four vocabularies, a mapping score indicating the nearness of the match, the Concept Unique Identifier (CUI) for the concept, and semantic type of the concept. See Exhibit 1. For each dataset (C) represents mapping to a precoordinated concept and (R) represents mapping to a postcoordinated concept. To identify mapped phrases, we first looked at the UMLS mapping. We recorded the phrase as (1) a precoordinated mapping if a single concept best represented the indication phrase; (2) postcoordinated if more than one concept was required to represent the phrase in its entirety; and (3) non match if the concepts incorrectly or incompletely represented the phrase. If a phrase mapped to both a precoordinated concept and postcoordinated concepts, we recorded it as a precoordinated mapping. For example the phrase “acute coronary syndrome” mapped to both a single concept and three concepts with separate CUIs.

For all the phrases that had an UMLS mapping, we recorded the mapping to concepts in the other datasets. Since the VA/KP is a subset of SNOMED Precoordinated which, in turn, is a subset of Entire SNOMED, which is a subset of UMLS, it is not possible for a concept to exist in the subset but not in the superset. However, some phrases with a precoordinated mapping in UMLS mapped to postcoordinated concepts in Entire SNOMED. For example, the phrase “familial hypophosphatemia” mapped to a single concept in UMLS but to two concepts in Entire SNOMED, “familial” of semantic type “qualifying concept” and “hypophosphatemia” of semantic type “disease or syndrome”.

All concepts in the VA/KP Subset are from the “Clinical Finding” hierarchy of SNOMED. In UMLS this includes the semantic types “disease or syndrome”, “sign or symptom”, “clinical finding”, and “pathological function”. If we define indication for a treatment as a...
medical condition which can be a reason for administering that treatment, then these semantic types are appropriate and indeed expected for all indications. To validate this assumption, we took specific note of the phrases mapped to UMLS concepts with semantic types which are none of the above.

If precoordinated concepts are expected to be of a semantic type appropriate for indications, the same must hold true for post-coordinated concepts. Most post-coordinated concepts consist of a lead concept (e.g. “hypophosphatemia”) and one or several qualifiers (e.g. “familial”). The lead concept is the carrier of the semantic type of the overall concept and so is expected to be of the semantic type mentioned above.

Since in SNOMED Precoordinated and the VA/KP Subset post-coordination of lead concepts with qualifiers is impossible, those phrases would only match on their lead term. We call these “partial matches”. Example of such a partial match is “drug-induced anemia”, where “anemia” is the lead term and “drug-induced” a qualifier and only “anemia” is mapped. Only phrases with a mapping of the lead term were considered partial matches, the others were considered non-matches.

Results

After merging the Gopher and Dailymed lists, the test set contained 1850 clinical phrases. Elimination of redundant and misspelled phrases resulted in 1265 distinct clinical phrases. The results of the mapping to UMLS are summarized in Table 1.

<table>
<thead>
<tr>
<th>Count</th>
<th>% of total</th>
<th>% of mapped</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Set</td>
<td>1265</td>
<td>100.0%</td>
</tr>
<tr>
<td>Phrases mapped successfully</td>
<td>1201</td>
<td>94.9%</td>
</tr>
<tr>
<td>Phrases mapped to a precoordinated concept in UMLS</td>
<td>1056</td>
<td>83.5%</td>
</tr>
<tr>
<td>Phrases represented using postcoordination</td>
<td>145</td>
<td>11.5%</td>
</tr>
<tr>
<td>Phrases with either no match or incorrect match</td>
<td>64</td>
<td>5.0%</td>
</tr>
</tbody>
</table>

Five percent of the phrases (64) either did not have a mapping concept or were incorrectly mapped in all four datasets. These phrases were often complex or ambiguously constructed. An example of an unmapped phrase is listed in Exhibit 2, other examples include: “secondary amenorrhea of undetermined etiology”, “steroid-responsive inflammation of the palpebral and bulbar conjunctiva”, “liver structural abnormalities”, “parkinsonism which may follow injury to nervous system by carbon monoxide intoxication”, and “contrast enhancement of computed tomographic body imaging”.

Of the 1056 phrases that mapped to a precoordinated concept in UMLS, Entire SNOMED contained a precoordinated concept for 950. A further 47 mapped to postcoordinated concepts and 59 phrases did not have an exact match (Table 2). All the 145 phrases that had postcoordinated concept mapping in UMLS were also represented as postcoordinated concepts in Entire SNOMED.

<table>
<thead>
<tr>
<th>Count of Phrases</th>
<th>% of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precoordinated concepts in UMLS but not in SNOMED</td>
<td>106</td>
</tr>
<tr>
<td>Precoordinated UMLS concepts represented in SNOMED with postcoordination</td>
<td>47</td>
</tr>
<tr>
<td>Mapped to UMLS but not to SNOMED concepts</td>
<td>59</td>
</tr>
</tbody>
</table>

The VA/KP Subset had completely matching concepts for 712 phrases and partially matching concepts for 187 phrases. SNOMED Precoordinated had completely matching concepts for 829 phrases and partially matching concepts for 176 phrases. See Table 3 for a breakdown of the results. The VA/KP Subset had no match for 366 (28.9%) phrases overall, 302 of which mapped to UMLS. SNOMED Precoordinated had no match for 260 (20.5%) phrases, 198 of which mapped to UMLS.

When compared to VA/KP Subset, SNOMED Precoordinated contained 117 more precoordinated concepts. Only 86% of the concepts in SNOMED Precoordinated were represented in the VA/KP Subset.

Regarding the semantic types we found that of the 1201 phrases that mapped to either precoordinated or post-coordinated concepts in UMLS, 1064 (88.6%) were of the expected semantic types for medical conditions (i.e. “disease or syndrome”, “sign or symptom”, “clinical finding”, and “pathological function”). Among
these 1064 phrases that mapped to a UMLS concept of appropriate semantic type, coverage was much improved in both SNOMED Precoordinated (1005, 94.5%) and VA/KP Subset (899, 84.5%)

Among the remaining 137 phrases that mapped to UMLS concepts of other semantic types, 97 were of type “diagnostic procedure”, “laboratory procedure”, “pharmaceutical substance”, or “therapeutic or preventive procedure”, and 40 were of type “bacterium”, “virus”, “fungus”, or “invertebrate”.

**Discussion**

In our data, 95% of the indication phrases map to some UMLS concept, either pre-coordinated or post-coordinated. However, just 75% of these concepts were mapped to the VA/KP Subset. If all of SNOMED Precoordinated were allowed instead, nearly 84% of phrases could have been mapped. In other words, the VA/KP Subset covers only 86% of the phrases of SNOMED Precoordinated.

In order to rule out any systematic defect in the VA/KP Subset, we compared it to SNOMED Precoordinated in detail. Omissions in the VA/KP Subset seemed to be evenly distributed across all levels and branches of the SNOMED findings hierarchy and we were not able to identify any pattern for the missing concepts (e.g., “erosive esophagitis”, “scabies”, “digoxin toxicity”, “hypomania”, and “primary dysmenorrhea”). This is consistent with the VA/KP Subset’s nature being an empirical collection of terms actually used or requested by clinicians for problem list coding.14

Thus, our findings seem to cast doubt on the FDA’s decision to use the VA/KP Subset as the vocabulary for representing indications. However, there are other factors involved in this decision. According to Humphreys et al15 adoption of standards by end users requires (1) a reasonable base of controlled terms (2) an open and sustainable process for enhancing and updating the vocabulary, (3) an efficient and low-cost electronic distribution method and (4) mandates and incentives to use the vocabulary. The FDA’s adoption of the VA/KP subset exemplifies all these requirements, for obviously the FDA (4) mandates, but also (3) provides free electronic access to all involved standards and terminologies from its SPL resources web site. Its initial adoption of the VA/KP subset under a special agreement with SNOMED gave access to (1) a reasonable base of controlled terms, and (2) to a process of extending the VA/KP Subsets with existing or even newly created concepts as necessary.

Our expectation to find indication terms coded in the semantic types for medical conditions have held true in 88.6% of phrases, and if focusing on these semantic types only, we find that coverage even of the VA/KP Subset is much better. Conversely, even if successfully coded with some UMLS concept, the utility of those indications not coded as a medical condition is questionable, as those cannot be compared against the patient’s problem list. For example, preparatory agents, contrast agents, or anesthetics are given not to treat a medical condition but to induce an unnatural condition. While these indications may be hard to code, they may not be important to code as indications as their purpose is different in both human reasoning and decision support functions.

A sizable number of phrases also coded to a concept of a microorganism. These appear to be systematically deficient mappings, as the correct mapping should have been to a medical condition introduced by the implied phrase “infection with” before the name of the microorganism. Microorganisms are also found as further qualifiers. These are clearly important for distinguishing indications and should be incorporated in the VA/KP Subset by defining appropriate precoordinated terms for these infectious medical conditions.

A perpetual point of contention is whether to use post-coordination to increase the coverage of the vocabulary. Indeed, according to our data, if post-coordination was permitted, 95% of the phrases could have been mapped to Entire SNOMED. In our test set, the most frequent qualifiers were of the following semantic types:

<table>
<thead>
<tr>
<th>Phrases mapped to a precoordinated concept</th>
<th>Phrases represented completely with postcoordination</th>
<th>Phrases represented by postcoordination in UMLS with a partially matching concept</th>
<th>Phrases represented by precoordinated concept in UMLS with a partially matching concept</th>
<th>Phrases with no match</th>
</tr>
</thead>
<tbody>
<tr>
<td>VA/KP Subset</td>
<td>712 (56.3%)</td>
<td>102 (8.1%)</td>
<td>85 (6.7%)</td>
<td>366 (28.9%)</td>
</tr>
<tr>
<td>SNOMED Precoordinated</td>
<td>829 (65.5%)</td>
<td>104 (8.2%)</td>
<td>72 (5.7%)</td>
<td>260 (20.5%)</td>
</tr>
<tr>
<td>Entire SNOMED</td>
<td>950 (75.1%)</td>
<td>192 (15.2%)</td>
<td>-</td>
<td>123 (9.7%)</td>
</tr>
</tbody>
</table>

Table 3: Results of indication phrase mapping to the VA/KP, SNOMED Precoordinated, and Entire SNOMED datasets.
1. *functional concept*, e.g., “active crohn’s disease”, and “resistant mycobacterial infection”.
2. *temporal concept*, e.g., “acute gout”, “chronic gout”, and “postpartum urinary retention”.
3. *qualitative concept*, e.g., “advanced carcinoma of prostate”, “drug-induced extrapyramidal disorders” and “severe bradycardia”.
4. *intellectual product*, e.g., “stage V chronic kidney disease”, and “stage D2 carcinoma of prostate”.
5. *bacterium, virus, fungus, invertebrate* (microorganism), e.g., *streptococcus viridans endocarditis*, “listeria meningitis”, and “pseudomonas UTI”.

These qualifiers seem to serve an important purpose. For instance, when providing differential dosage guidance, it is important to distinguish between acute vs. chronic gout. In addition an indication for which a certain medication is approved may be just a subclass of a medical condition. While these findings would seem to advocate for using post-coordination there are practical difficulties with it which existing systems have not yet widely overcome, and hence, appear difficult for FDA to mandate at this time.

In order for our study to be applicable in a real world situation, we extracted the phrases for our test set from a real world application and from existing drug labels. Thus we tested the ability of the various terminologies to represent phrases currently used by existing applications for describing indications. However, the question arises whether indication terms used in practice are the ones that should be necessary. For example, pharmaceutical companies have incentives for inventing special indication terms for the purpose of product differentiation. So it may not be important to represent these fully. On the other hand, some distinctions that are medically important in choosing a therapy may not be expressed in our sources. Scrutinizing indication terms in this way is beyond the scope of our study, but could be a worthwhile follow-up to it.

**Conclusion**

Our study suggests that the VA/KP Subset has significant limitations for coding indications; however, when focusing on indications as medical conditions only, the coverage seems more adequate. The VA/KP Subset had been selected with many practical considerations in mind and it may be sufficient if the process of extending this Subset works and is indeed used, particularly for the coding of microorganisms and stages of diseases where it makes a difference in indication and dosage.

**Acknowledgements**

This work was performed at the Regenstrief Institute, Indianapolis, IN, and was supported in part by the National Library of Medicine (NLM) grant T15 LM07117, the Agency for Healthcare Research and Quality (AHRQ) grant R01 HS15377 and the Food and Drug Administration (FDA).

**References**

5 Schadow G, Gitterman S, Boyer S, Dolin RH eds. HL7 v3.0 structured product labeling, release 2 [standard]. Ann Arbor, MI, Health Level Seven, 2005
6 http://www.fda.gov/bbs/topics/NEWS/2005/NEW01232.html
9 http://www.fda.gov/oc/datacouncil/term.html#med
14 Dolin RH. Personal communication.
15 Humphreys BL, McCray AT, Cheh M. Evaluating the Coverage of Controlled Health Data Terminologies: Results of the NLM/AHCRP Large Scale Vocabulary Test. J Am Med Inform Assoc. 1997;4:484-500