CONTROLLED VOCABULARY REQUIREMENTS FOR GUIDELINE INTEROPERABILITY: A study in clinical standards

James R. Campbell MD, Karen M. Hrabak RN BSN

University of Nebraska Medical Center; Omaha, NE, USA

Abstract

We set out to evaluate the utility of contemporary controlled vocabularies in support of interoperable sharing of guideline decision logic. We employed the information model and vocabulary procedures of the SAGE[1] guideline project. We selected clinical guidelines to represent a balanced clinical focus. We found that vocabulary coverage for guideline content was: ICD-9-CM 22%, UMLS 76% and SNOMED CT 78%. Employing the compositional features of SNOMED CT and post-coordination, SNOMED CT captured 89% of guideline concepts. We further identified that 70% of guideline concepts were in fact parent concepts of aggregate lists that must be collated for complete case finding. Only SNOMED CT supports aggregation ontologies which automate this collation. Analyzing all guidelines, for problem or procedure concepts, SNOMED CT identified an average of 224 more specialized concepts for each concept encoded. Compared to evaluations of a half decade ago, UMLS and SNOMED have made considerable progress. Although work remains, we conclude that SNOMED CT offers substantial and necessary support for sharing of clinical decision support functions.

Keywords: Controlled vocabulary, Systematized nomenclature of medicine, Unified medical language system, Clinical practice guideline

Methods

A set of four clinical guidelines were selected from the National Guideline Clearinghouse(R). These included immunization practices, standards of care for diabetics, pressure ulcer assessment and prevention and management of major depression. A source document was identified for each practice guideline representing the full text guideline statement. This document was analyzed in order to identify recommendation sets within the guideline.

Individual recommendation sets were analyzed and one or more clinical action plans were formulated to implement the guideline recommendations in a theoretical but realistic clinical environment. This was accomplished employing the SAGE[1] clinical guideline model employing the PROTÉGÉ object modeling environment. The action plan and the recommended decisions and actions were formulated into a series of data queries and medical logic modules which were constructed to implement the guideline recommendations in the target clinical environment.

All data queries in support of guideline recording and decision needs were formulated as standard SQL queries in relationship to a virtual medical record information of the computerized patient record. For each data field in a vMR record, vocabulary standards were selected and restriction semantics were defined which specified the portions of the reference terminology to be allowed for the field data type. We analyzed a series of structured terminologies for their capacity to encode guideline content, including SNOMED CT (release July 2004), LOINC, ICD-9-CM and the Unified Medical Language metathesaurus (UMLS).

We further explored the usefulness of post-coordination to compositionally encode guideline content and evaluated SNOMED ontologies for their utility in providing logical aggregation of guideline concepts in support of decision logic.

Conclusions

SNOMED CT and UMLS have grown substantially in content coverage since comprehensive studies in 1997. Only these controlled terminologies have sufficient content to encode guideline decision logic. SNOMED CT, a reference terminology with an expansive ontology and compositional features, supports post-coordination of concepts and features of subsumption which are necessary to guideline decision support. The NCHS designation of SNOMED CT as core terminology for US clinical information systems is supported by these research results.