Using SNOMED CT in Building a Database for Comparative Effectiveness Research

Ramkiran Gouripeddi, MBBS, MS, Ryan Butcher, MS, MBA, University of Utah, Salt Lake City UT

Audience

This presentation will be of interest to those interested in implementing SNOMED CT for clinical research and within informatics platforms, and to terminologists mapping local coding schemas to SNOMED CT.

Objectives

1. Mapping local coding schemata to SNOMED CT – Decisions to be made.
2. Implementation of SNOMED CT in clinical research informatics platforms.

Abstract

The PHIS+ database is being developed on the Federated Utah Research and Translational Health electronic Repository (FURTHeR) platform, which federates heterogeneous data on demand, from multiple data sources, and provides syntactic and semantic data interoperability for clinical and translational research purposes [1]. This database is being used to conduct pediatric comparative effectiveness research (CER). The six data contributing hospitals use different electronic source systems for their clinical data and implement different local coding schemata. Therefore, we had to map the local terminologies using metadata from each site, to standard terminologies based on national recommendations, expert discussions among project personnel, and availability of local metadata. This knowledge of how local systems define, use, and store their data is utilized by FURTHeR to process the local terminology and data models to populate a harmonized database.

We are using SNOMED CT as the standard terminology in mapping local coding schemata for microbiology culture results [1]. The PHIS+ database currently has approximately 1.8 million microbiology culture results mapped to 122 unique SNOMED CT culture codes, 851 unique SNOMED CT specimen codes, 250 unique SNOMED CT organism codes and 8 anti-microbial sensitivity interpretation codes. We also have approximately 120 unique units of measure codes associated with 142 million laboratory results.

We developed and used a metadata collection specification for obtaining the necessary detail to map local coding schemata to SNOMED CT which led to an improved accuracy of the mappings. The decisions that we made during the mapping process were influenced by discussions with domain experts and the requirements of the CER studies. These included issues such as the granularity of mappings, use of negations and qualifiers, and microbiological classification concerns. We are currently developing methods to capture versioning of local terminologies and SNOMED CT and to semi-automate some of the mappings, especially for free text entries. While we were able to use SNOMED CT for most cases, we developed methods to manage those few instances of local codes that could not be mapped to a standard. In this presentation we will share our processes and lessons learned in utilizing SNOMED CT in developing a standard terminologies-based database that will be used to perform CER.

References

1. Gouripeddi R, Warner PB, Mo P, Levin JE, Srivastava R, et al. Federating Clinical Data from Six Pediatric Hospitals: Process and Initial Results for Microbiology from the PHIS+ Consortium. AMIA Annu Symp Proc. 2012;2012:281–90.