An Efficient Approach To Map LOINC Concepts To Notifiable Conditions

Wei Li, MD[1], Jerome I. Tokars, MD, MPH[1], Nikolay Lipskiy, DrPH[1], Sundak Ganesan, MD[2]

1. National Center for Public Health Informatics (NCPHI), Centers for Disease Control and Prevention (CDC). 2. SAIC Consultant to NCPH, CDC

OBJECTIVE

The objective of this project was to develop an efficient method to map Logical Observation Identifier Names and Codes (LOINC) to notifiable conditions to accommodate future changes in standard laboratory test codes.

BACKGROUND

Electronic reporting of laboratory results has been shown to increase timeliness and completeness of reporting of notifiable conditions. Such electronic data interchange depends on tables that relate coded laboratory or clinical findings to the notifiable conditions under surveillance [1]. The Public Health Information Network Notifiable Condition Mapping Tables (PHIN NCMT) are a set of resources for public health reporting. They can be used to filter the output of clinical labs for test results of public health importance, to provide a framework for the development and maintenance of a controlled vocabulary for reportable events, and to provide a reusable domain knowledge component for intelligent surveillance information system architecture. PHIN NCMT relating LOINC (version 2.12) to notifiable conditions were produced in May 2004. The data in the tables is based on three standard coding systems that are publicly available in the United States: LOINC® for lab test names; SNOMED® -CT for organism names; and the CDC Nationally Notifiable Disease Surveillance System (NNDSS) Event Code List for notifiable conditions. However, the CDC NNDSS Event Code List is updated annually and new LOINC concepts are added semi-annually. For example, there are 46,812 terms in the December 2006 release of LOINC (version 2.19), which included 11,972 new terms that had been added since May 2004. Also, 6,735 terms were retired from LOINC since May 2004. Therefore, an efficient method to add new concepts to, and delete obsolete concepts from, mapping tables is needed.

METHODS

A list of 155 notifiable conditions obtained from the NNDSS event code list (2007) for the national and/or state levels was used. A list of 118 disease agents and agent groups that cause these conditions was produced. Four methods were used to map NNDSS to LOINC. First, the Regenstrief LOINC Mapping Assistant (RELMA v2.19) was used to identify appropriate LOINC [2]. Second, a SAS (v9.1.3) program was used to perform a text search of the component field of the LOINC database to find words that matched notifiable condition names. Third, the National Library of Medicine (NLM) RRF browser was used to identify the microorganism and navigate the relationships (hierarchical and non-hierarchical) present in UMLS 2007AA Metathesaurus subset. The Metamorphosys tool was used to create a UMLS 2007AA metathesaurus subset with just LOINC and SNOMED CT. Fourth, to check our results, we also used a manual process.

RESULTS

The existing mapping table produced in May 2004 contained 88 notifiable conditions mapped to 3,204 LOINC entries. Using the RELMA tool and UMLS RRF Browser, we mapped the 155 notifiable conditions to 5,352 LOINC entries; this process took approximately 200 person-hours. The manual method also found 5,352 LOINC entries and required approximately 1,050 person-hours. The SAS text parsing program found 4901 LOINC entries that mapped to notifiable condition names. Of the 451 LOINC entries not found by the SAS method, most had a non-specific test name, e.g., code 33700-6 (a spore identification test) is used to identify Anthrax, but the component field does not contain the word "Anthrax." Our 2007 mapping table included 2,199 LOINC entries not present in the 2004 table; also, 257 entries present in the 2004 table were deleted. The RELMA, UMLS RRF Browser, and manual results matched completely.

CONCLUSIONS

Compared with the 2004 NCMT, our 2007 NCMT added >2000 entries and deleted >200. The methods we used were labor intensive but currently necessary, since the fully automated SAS text search method is not currently accurate enough to be dependable. Efforts to improve the SAS method by adding additional search terms such as "spore" may enable more frequent automated updates. In the interim, the use of RELMA or the UMLS RRF Browser provides an accurate and efficient semi-automated process to update the Notifiable Condition Mapping Table

REFFERENCES

[1] Doyle TJ, Ma H, Samuel L Groseclose SL, and Hopkins RS: PHSkb: A knowledgebase to support notifiable disease surveillance. *BMC Medical Informatics and Decision Making* 2005, 5:27

[2] McDonald CJ, et al: LOINC, a Universal Standard for Identifying Laboratory Observations: A 5-Year Update. Clinical Chemistry 2003, 49:4: 624-633

Further Information: Wei Li, wli1@cdc.gov