Introduction of a Hierarchy to LOINC to Facilitate Public Health Reporting

Steven J. Steindel, PhD, John W. Loonsk, MD
Information Resource Management Office
Alan Sim, MS, Timothy J. Doyle, MPH, Robb S.Chapman, Samuel L. Groseclose, DVM, MPH
Division of Public Health Surveillance and Informatics, Epidemiology Program Office
Centers for Disease Control and Prevention, Atlanta, GA

Abstract

Public health reporting of laboratory results requires unambiguous identification of the test performed and the result observed. Some laboratories are currently using Logical Observation Identifier Names and Codes (LOINC) for the electronic reporting of laboratory tests and their results to public health departments. Initial use revealed inconsistent identification and use of LOINC concepts by laboratories and public health agencies and an inability to systematically extend, for public health use, the tables when adding new concepts. We applied simple, logical rules to existing LOINC concepts to facilitate the creation of a hierarchy of concepts and to allow the identification and specification of appropriate terms for public health reporting and subsequent data aggregation. The hierarchy also allows the systematic addition of new concepts further supporting public health reporting. Application of the hierarchy is illustrated by using all laboratory LOINC concepts assigned to the subset of microbiology test types (CLASS MICRO).

Introduction

A reportable disease is one for which regular, frequent and timely information regarding individual cases is considered necessary for the prevention and control of the disease. Medical providers and laboratories have a statutory responsibility to notify public health agencies when a reportable disease is observed.[1] Recently Effler et.al. demonstrated that electronic reporting from laboratories to public health departments can increase the timeliness and completeness of case reporting.[2] The Centers for Disease Control and Prevention (CDC) is investigating, through the National Electronic Disease Surveillance System (NEDSS)[3], expanded electronic reporting of laboratory results from public and private laboratories to the public health system and has published guidelines for the process.[4] Those guidelines call for the use of a Health Level 7 (HL7) Version 2.3 Observation Report – Unsolicited (ORU) message as the format for laboratory data transmission from clinical facilities to public health entities. Laboratory results are found in ≥1 Observation Result (OBX) segments. Field 3 of the OBX segment contains an identifier for the test performed, and the CDC guidelines request that it be transmitted using Logical Observation Identifier Names and Codes (LOINC).[5;6].

As part of previous work in this area, the Council of State and Territorial Epidemiologists (CSTE), in an effort lead by Diane Dwyer, M.D., and with the cooperation of the Regenstrief Institute, prepared a list of suggested LOINC codes that represent laboratory tests to diagnose diseases reportable to state and local public health departments.[7] Laboratory-based public health reporting requires a mechanism for routine maintenance that auto-classifies new LOINC concepts, which represent new laboratory tests, methods, or classifications, corresponding to public health conditions.

Additionally, the CDC observed during electronic laboratory reporting demonstration projects that certain tests performed by private health care organizations were not represented in the 1999 list of public health relevant LOINC codes.[4] Hence, a method of associating laboratory-assigned LOINC concepts to those desired for reporting and aggregation by the public health system was needed to support surveillance and response activities.

We hypothesized that these goals could be achieved through specification of the parent-child relationship inherent in a hierarchal structure. Preliminary investigation of the six main axes of LOINC (Component, Property, Time Aspect, System, Scale Type, and Method Type) along with the fields of CLASS and CLASSTYP revealed a simple hierarchy of concepts. We observed, however, that further organization was needed to achieve our goals of concept-based aggregation and monitoring of laboratory tests by

the public health system, ease of concept identification and assignment at the local level, and eventual knowledge discovery (e.g., representation of infectious disease case definitions using concept-based terminologies). Review of the sources used to develop LOINC revealed no evident prior work on hierarchies.[6] Based on the concepts contained in the MICRO CLASS of the CLASSTYP 1 (Laboratory), a set of rules for grouping of LOINC concepts was developed. This report describes those rules, and illustrates the effect they could have on using LOINC for public health laboratory testing and reporting.

Methods

An abridged version of the LOINC database was prepared, containing the fields LOINC_NUM, COMPONENT, PROPERTY, TIME_ASPECT, SYSTEM, SCALE_TYP, METHOD_TYP, CLASS, and CLASSTYP. Although the LOINC Users' Guide [8] notes only Clinical and Laboratory CLASSTYPs, four categories were identified in the database and were assigned the descriptive names of Attachment, Clinical, Lab and Survey, corresponding to their perceived use. Within each CLASSTYP, the field CLASS tended to organize similar LOINC identifiers into usable groupings (e.g., CHEM, MICRO) as noted in the User's Guide. On the basis of these observations, we decided to create a simple hierarchy based on CLASSTYP, CLASS, and an alphabetical list of COMPONENT within those groupings.

The abridged LOINC database was converted to an XML file for import into the Apelon Inc. (Ridgefield CT) Terminology Development Environment (TDE). This tool allowed easy manipulation of the hierarchical relationships, linkage to other concepts of public health importance such as infectious disease case definitions, and application of inferential logic to the developing terminology to reveal new relationships, providing possible knowledge extension.

The TDE requires three unique fields for a concept: NAME, CODE and ID. The LOINC identifier was used as a CODE and a sequential nondefining integer was developed for the ID. We choose the LOINC field COMPONENT as the TDE NAME. However, LOINC is not rigid in assigning COMPONENT definitions and multiple concepts are assigned the same term. To facilitate use as the TDE NAME, each COMPONENT term was made unique by assigning it an appending term "(#)," where the number assigned was generally the arbitrary sequence number in an alphabetical list. Concepts were added to the XML file to correspond to the CLASS and CLASSTYP and these concepts were then used as TDE hierarchical identifiers (i.e., super concepts). Finally, COMPONENT, PROPERTY, TIME_ASPECT, SYSTEM, SCALE_TYP, and METHOD_TYP were added as concept properties in the TDE.* After import, the LOINC concepts could readily be displayed in a hierarchy that consisted of CLASSTYP, CLASS, and alphabetical list by appended COMPONENT.

Proposed Hierarchy Rules

The simple hierarchical display resulting from the above organization allowed us to more readily locate similar LOINC concepts, and resulted in the development of the following proposed grouping rules:

1. Group similar concepts by CLASSTYP and within CLASSTYP by CLASS.
2. Group all concepts within the CLASS by COMPONENT alphabetical order.
3. Group all concepts with the same COMPONENT by METHOD_TYP. If a null METHOD_TYP exists, it forms the parent under which all other child METHOD_TYPs appear. Topic experts are then needed to further define METHOD_TYP hierarchies as the following example illustrates:
   ACID FAST STAIN
   ACID FAST STAIN: ZIEHL-NELSON
   ACID FAST STAIN: KINYOUN
   ACID FAST STAIN: KINYOUN MODIFIED
4. Group concepts having a different COMPONENT but can be considered the same on the basis of other properties by METHOD_TYP. For example when the COMPONENT is a specific microorganism

* The TDE has a limitation on NULL property names and these were assigned the name of "blank" to facilitate import and later manipulation. We present the LOINC nomenclature of NULL in this paper.

determined by the method of SPECFIC ORGANISM CULTURE it can be grouped under one of the general culture method concepts associated with the COMPONENT "MICROORGANISM IDENTIFIED."

5. Within each METHOD_TYP, group concepts by SCALE_TYP and PROPERTY. The ordinal (ORD) and ordinal/quantitative ORDQN scale types are considered children of the SCALE_TYP parent quantitative (QN). Accepting these relationships is reasonable if the operator assumes ordinal tests (POS, NEG, 1+, etc.) are quantitative results on the basis of method defined zero or cut-off points and units. The nominal (NOM), narrative (NAR) and multiple (MULTI) scale types are left unchanged. While the PROPERTY attribute is not widely used in microbiology tests, similar reasoning can be used. For example, concepts having the property Arbitrary Concentration (ACNC) could be considered as parents of the more specific Mass Concentration (MCNC).
6. Group concepts within the same level by SYSTEM on the basis of an assigned hierarchy by using "XXX (To be specified in another part of the message)" as the base with the other LOINC SYSTEM designations

Figure 1: *Bacillus anthracis* testing LOINC hierarchy developed using proposed rules

BACILLUS ANTHRACIS AB:NULL:ACNC:QN:XXX (22109-3)* **[Rules 1,3,6]****
- BACILLUS ANTHRACIS AB:NULL:ACNC:QN:SER (7814-7) **[Rule 2,6]**
  - BACILLUS ANTHRACIS AB:IB:ACNC:QN:SER (11467-8) **[Rule 3,6]**
  - BACILLUS ANTHRACIS AB:HAI:ACNC:QN:SER (5055-9) **[Rule 3,6]**
  - BACILLUS ANTHRACIS AB:NULL:TITR:QN:SER (22859-3) **[Rule 3,5,6]**
    - BACILLUS ANTHRACIS AB:ID:TITR:QN:SER (22865-0)
  - BACILLUS ANTHRACIS AB:NULL:ACNC:ORD:SER (22860-1) **[Rule 5,6]**
    - BACILLUS ANTHRACIS AB:ID:ACNC:ORD:SER (22861-8)
    - BACILLUS ANTHRACIS AB:AGGL:ACNC:ORD:SER (22862-7)
    - BACILLUS ANTHRACIS AB:EIA:ACNC:ORD:SER (22863-5)
    - BACILLUS ANTHRACIS AB:CF:ACNC:ORD:SER (22864-3)
- BACILLUS ANTHRACIS AB:IF:ACNC:QN:XXX (11468-6) **[Rule 3]**

BACILLUS ANTHRACIS AG:IF:ACNC:ORD:XXX (22867-6) **[Rule 1]**

MICROORGANISM IDENTIFIED:CULTURE:PRID:NOM:XXX (11475-1) **[Rules 1,3,6]**
- MICROORGANISM IDENTIFIED:ANAEROBIC+AEROBIC CULTURE:PRID:NOM:XXX (21020-3) **[Rule 3]**
  - MICROORGANISM IDENTIFIED:AEROBIC CULTURE:PRID:NOM:XXX (634-8) **[Rule 2,3]**
    - BACILLUS ANTHRACIS IDENTIFIED:ORGANISM SPECIFIC CULTURE:ACNC:ORD:XXX (11469-4) **[Rule 4]**
      - BACILLUS ANTHRACIS:ORGANISM SPECIFIC CULTURE:ACNC:ORD:XXX (20691-2) **[Duplicate – see text]**
  - MICROORGANISM IDENTIFIED:AEROBIC CULTURE:PRID:NOM:NOS (10353-1) **[Rule 3,6]**
- MICROORGANISM IDENTIFIED:ENVIRONMENTAL CULTURE:PRID:NOM:XXX (14325-5) **[Rule 3]**

MICROSCOPIC OBSERVATION:NULL:PRID:NOM:XXX (11545-9) **[Rules 1,3,6]**
- MICROSCOPIC OBSERVATION:GRAM STAIN:PRID:NOM:XXX (664-3) **[Rules 3,6]**
  - MICROSCOPIC OBSERVATION:GRAM STAIN:PRID:NOM:FLU (27112-3) **[Rule 6]**
    - MICROSCOPIC OBSERVATION:GRAM STAIN:PRID:NOM:CSF (14357-8)
    - MICROSCOPIC OBSERVATION:GRAM STAIN:PRID:NOM:PLR (14360-2)
- MICROSCOPIC OBSERVATION:INDIA INK PREPARATION:PRID:NOM:XXX (666-9) **[Rules 3,6]**
  - MICROSCOPIC OBSERVATION:INDIA INK PREPARATION:PRID:NOM:CSF (638-7)

*Key: COMPONENT:METHOD_TYP:PROPERTY:SCALE_TYP:SYSTEM (LOINC Number)
Abbreviations are those defined by the LOINC Users' Guide(7)
**See text for rules used to classify parent and child concepts.

following. In developing the SYSTEM hierarchy synonyms are grouped with the most common term as the parent and the others as children [i.e.: CALC (Calculus) as parent of STON (Stone)]. Dual parents are assigned to those terms with ambiguous roots (e.g.: Wound Abscess assigned to parents Wound and Abscess). In certain cases, the role of the specimen in laboratory testing was considered in assigning a parent. For example FLU was restricted to bodily fluids other than blood components such as serum or plasma. Blood components are a fluid, but are tested in a much larger volume than other fluids [9] and, therefore, are considered separately.

Note that these rules do not use TIME_ASPECT. Establishing rules for TIME_ASPECT use in the hierarchy added no particular advantage over rules proposed thus far, presented greater conceptual difficulty and was not pursued.

These rules were applied to the entire Laboratory CLASSTYP CLASS MICRO to impose a hierarchy.

Results

While a display of the entire hierarchy is beyond the scope of this paper, as an example Figure 1 indicates the LOINC concepts defined for *Bacillus anthracis* testing arranged in hierarchical order using the proposed rules.

Discussion

Figure 1 illustrates the value of a hierarchical order to LOINC. Note that testing falls into four general groups: antibody tests, antigen tests, traditional culture, and microscopic methods. For public health purposes, previous exposure, but not proof of current disease generally associates with the antibody tests whereas current disease would generally be defined by positive antigen or culture results. However, antibody results that increase a specified amount over a defined timeframe are indicative of recent exposure, a concept that is not well represented in LOINC or our hierarchy. Also, immunoglobin M and A antibodies represent recent exposure and are considered diagnostic.

By using this hierarchy we can describe the laboratory testing performed to identify *B. anthracis* contamination and subsequent individual exposure, which is not possible with the distributed non-hierarchical LOINC concept structure. Presumptive testing for the organism involves a Gram stain of contaminated fluid (27112-3, 14357-8, or 14360-2) sometimes followed by an India Ink preparation (666-9 or 638-7) to detect encapsulation. Confirmation testing involves a culture, and depending on laboratory practice, the hierarchy indicates suggested codes. For all culture tests with the exception of the organism-specific culture, the results would be the organism identified. Results for an organism-specific culture are ordinal (Figure 1). Further confirmation testing from culture or directly from blood might involve antigen testing by immunofixation (22867-6) or indication of the presence of the organism by using a specific probe and subsequent identification through a polymerase chain reaction (PCR) procedure. Workplace contamination could be detected by using an organism-specific culture, but the preferred method would be that involving an environmental culture (14325-5). Presumptive exposure might involve antibody testing (e.g., the tree defined by 22109-3) or nasal swab cultures (10353-1). Testing laboratories can report using LOINC concepts at any appropriate point in the hierarchy. Public health department staff can use the hierarchy to guide the aggregation of laboratory test results to directly support their surveillance and intervention activities.

The proposed hierarchy also clearly identifies similar or missing LOINC concepts. Figure 1 describes two apparently identical concepts for organism specific culture (11469-4 and 20691-2). Review of the details in the LOINC database reveals that in the past they had differing PROPERTY/SCALE_TYP codes. Recently, the concepts properties were redefined by LOINC to be equivalent. Two confirmatory methods for *B. anthracis*, PCR and detection in tissue by immunohistochemical (IHC) stains, need to be added to LOINC as new concepts with appropriate specimen codes. Presently, no primitive LOINC concepts exist for these methods for *B. anthracis,* and laboratories cannot use LOINC to represent them in electronically transmitted information. Although isolate (ISLT) is a widely used SYSTEM for susceptibility

testing, it is not defined often within LOINC for organism identification confirmation testing. A concept for isolate testing exists for anaerobic cultures and is perhaps needed at the aerobic and general levels as well, illustrating the use of the proposed hierarchy for identifying "missing" LOINC concepts. Testing on isolates is common for reference, especially public health, laboratories. Specific SYSTEM concepts (e.g., vesicular fluid for cutaneous anthrax testing) do not exist and, if added, would extend leaf nodes of the tree. Meanwhile, reporting under the generic fluid code that exists in the AEROBIC CULTURE tree (610-6, not shown) could substitute.

While this investigation confined itself to microbiological testing, we feel that a similar approach will work in all areas of the Laboratory CLASSTYP and perhaps the other CLASSTYPs as well. The major challenge will be the extension of the rules by topic experts as they evaluate the value assigned within each axis.

Conclusion

Figure 1 illustrates that LOINC codes can be organized in a hierarchy. Concept-based hierarchical terminologies support the public health system's move to informatics-based surveillance activities that take advantage of increasing availability of electronic health information to support public health's need to detect and respond to emerging health threats. We anticipate that the other areas of laboratory LOINC have similar benefits. Although this hierarchy was developed using only existing LOINC concepts, improvement could come if true primitives, LOINC concepts that do not represent tests that could be ordered, were defined. If LOINC is extended in this fashion, the potential ambiguity of using certain LOINC concepts as hierarchical headers is eliminated.

References:

(1) Roush S, Birkhead G, Koo D, Cobb A, Fleming D. Mandatory reporting of diseases and conditions by health care professionals and laboratories. JAMA 1999; 282:164-170.

(2) Effler P, Ching-Lee M, Bogard A, Ieong MC, Nekomoto T, Jernigan D. Statewide system of electronic notifiable disease reporting from clinical laboratories: comparing automated reporting with conventional methods. JAMA 1999; 282:1845-1850.

(3) Centers for Disease Control and Prevention. Supporting Public Health Surveillance through the National Electronic Disease Surveillance System (NEDSS). US Department of Health and Human Services, CDC, 1999. Available at http://www.cdc.gov/od/hissb/docs/NEDSS%20Intro.pdf (date last accessed: 2/19/2002)

(4) Centers for Disease Control and Prevention. Electronic reporting of laboratory information for public health January 7-8, 1999: summary of meeting proceedings. US Department of Health and Human Services, CDC, 1999. Available at http://www.cdc.gov/od/hissb/docs/elr-1999.pdf (date last accessed: 2/19/2002)

(5) Forrey AW, McDonald CJ, DeMoor G, Huff SM, Leavelle D, Leland D et al. Logical observation identifier names and codes (LOINC) database: a public use set of codes and names for electronic reporting of clinical laboratory test results. Clin Chem 1996; 42:81-90.

(6) Huff SM, Rocha RA, McDonald CJ, De Moor GJ, Fiers T, Bidgood WD, Jr. et al. Development of the Logical Observation Identifier Names and Codes (LOINC) vocabulary. J Am Med Inform Assoc 1998; 5:276-292.

(7) McDonald CJ, Overhage JM, Dexter P, Takesue BY, Dwyer DM. A framework for capturing clinical data sets from computerized sources. Ann Intern Med 1997; 127(8 Pt 2):675-682.

(8) Regenstrief Institute. Logical Observation Identifier Names and Codes (LOINC) User's Guide. January 5, 2001 ed. Indianapolis, IN.

(9) Steindel SJ, Rauch WJ, Simon MK, Handsfield J. National Inventory of Clinical Laboratory Testing Services (NICLTS). Arch Pathol Lab Med 2000; 124:1201-1208.