Case Report ■

Standardizing Laboratory Data by Mapping to LOINC

Agha N. Khan, MD, MPH, Stanley P. Griffith, MD, Catherine Moore, Dorothy Russell, Arnulfo C. Rosario Jr., MD, MPH, Jeanne Bertolli, PhD

Abstract The authors describe a pilot project to standardize local laboratory data at five Indian Health Service (IHS) medical facilities by mapping laboratory test names to Logical Observation Identifier Names and Codes (LOINC). An automated mapping tool was developed to assign LOINC codes. At these sites, they were able to map from 63% to 76% of the local active laboratory tests to LOINC using the mapping tool. Eleven percent to 27% of the tests were mapped manually. They could not assign LOINC codes to 6% to 19% of the laboratory tests due to incomplete or incorrect information about these tests. The results achieved approximate other similar efforts. Mapping of laboratory test names to LOINC codes will allow IHS to aggregate laboratory data more easily for disease surveillance and clinical and administrative reporting efforts. This project may provide a model for standardization efforts in other health systems.

■ **J Am Med Inform Assoc.** 2006;13:353–355. DOI 10.1197/jamia.M1935.

Public health is increasingly moving toward automated capture and analysis of data, use of data that are already in electronic form, and integration of public health and health care information systems. Laboratory data are becoming an increasingly valuable tool for public health agencies,[1] but a major challenge to automated aggregation of these data from different facilities is that code sets for laboratory test names are different from one information system to another. A solution to this problem is to map the local test names from these systems to a standard set of codes.

The Logical Observation Identifier Names and Codes (LOINC) code set is the only publicly available universal standard for laboratory test names.[2–4] The current version of the LOINC code set, maintained by the Regenstrief Institute, contains more than 33,000 observations. Each LOINC record corresponds to a single test result and includes fields for specifying component (analyte), property measured, timing, type of sample, type of scale, and, where relevant, the method used to produce the result. Therefore, each LOINC code has up to six attributes and belongs to a class or group (e.g., Chemistry, Microbiology). LOINC syntax looks like this:

(analyte):(property):(time aspect):(specimen):(scale):(method)

Example (10351-5): HIV 1 RNA:ACNC:PT:SER/PLAS:QN:PROBE.AMP

The terms in the above example are HIV 1, human immunodeficiency virus 1; RNA, ribonucleic acid; ACNC, arbitrary concentration; PT, point in time; SER/PLAS, serum/plasma; QN, quantitative; and PROBE.AMP, probe with amplification.

In this manuscript we describe a pilot project, The Map to LOINC Project, which was a collaborative effort undertaken by the Centers for Disease Control and Prevention and the Indian Health Service (IHS) to design and test a semiautomated process to map local laboratory test names at five IHS medical facilities to LOINC. IHS is a federal agency of the United States Department of Health and Human Services that provides health care services to American Indians and Alaska Natives. IHS medical facilities use an integrated clinical and administrative information system, the Resource and Patient Management System (RPMS). The system was developed in "MUMPS" or "M" programming language and uses VA FileMan, developed by the Veterans Administration (VA), as the database engine. The RPMS consists of more than 35 different applications including a Laboratory Package application, which is used at about 33% of IHS medical facilities for laboratory records, but there are some variations in its application from one medical facility to another. The major task of our project was to eliminate these variations in laboratory test names data by mapping these data to LOINC.

Methods

Facilities using the RPMS laboratory package were eligible to participate in the pilot project. The process for mapping laboratory test names to LOINC was designed to accommodate future changes in laboratory test names/codes, to meet all data security and confidentiality standards, and to be easily expandable to other IHS medical facilities in the future.

The laboratory test names and synonyms along with the test result measurement units and type of specimen tested were exported from the five participating IHS medical facilities.

Affiliations of the authors: Centers for Disease Control and Prevention, Atlanta, GA (ANK, ACR, JB); Indian Health Service, Albuquerque, NM (SPG, CM); Cimarron Medical Informatics, Tucson, AZ (DR).

The authors thank their colleagues in the five Indian Health Service facilities that participated in the pilot phase of this project: Albuquerque Indian Hospital, Rapid City PHS Sioux San Indian Hospital, Ignacio PHS Indian Health Center, Phoenix Indian Medical Center, and W. W. Hastings Indian Hospital. They also thank Russell Pitman, Patrick Beatty, and Tim Frazier from IHS and Lori Butcher and Mark Williams from Cimarron Medical Informatics.

Correspondence and reprints: Agha N. Khan, MD, MPH, Centers for Disease Control and Prevention, 1600 Clifton Road NE, MS 08, Atlanta, GA 30333; e-mail: <akk3@cdc.gov>.

Received for review: 08/15/05; accepted for publication: 12/18/05.

Only the details of tests in active use at the facilities were gathered. These data sets were combined to create an IHS LOINC Master File. Two scientists, through mutual agreement, manually assigned LOINC codes to the tests in the master file. The Regenstrief LOINC Mapping Assistant was used as a browser to identify the LOINC codes for the tests. The tests, which had incomplete or incorrect information, were marked as "uncodable." Panel tests that had no corresponding LOINC codes were excluded from the master file.

An automated mapping tool was developed for mapping local laboratory test names to LOINC using the mappings specified in the master file. The mapping tool incorporates a two-step process. LOINC codes are assigned to the laboratory tests if there is an exact match (case insensitive) between the laboratory test name in the master file and the local laboratory test name. Otherwise, all leading and trailing typographic characters (nonnumeric alpha characters) are removed and the matching process tried a second time. If a match is found during the second pass, a LOINC code is assigned; otherwise no LOINC code is assigned to the test.

The mapping tool and master file were combined in an RPMS software patch, the LR patch. The VA's parallel effort to map laboratory data to the LOINC standard (VA patch) provided a starting place. The LR patch consists of the automated mapping tool, an option to manually assign LOINC codes to local laboratory tests (for tests for which codes could not be assigned by the mapping tool) and the VA patch containing a LOINC look-up file, which may be used to identify LOINC codes for manual mapping. Two other RPMS patches, the IHS Patient Care Data Entry (APCD) patch and the Generic Interface System (GIS) patch, were also developed by IHS. The main function of the APCD patch is to create a field in the Patient Care Component of RPMS to receive LOINC codes. The GIS patch contains a file transfer protocol utility to transfer selected laboratory data to the IHS server in Health Level Seven (HL7) format.

All three RPMS patches were applied to the databases at the five pilot sites. Tests not assigned a code by the mapping tool were reviewed and codes were assigned manually, if possible. After the mapping process was completed at each participating medical facility, the laboratory tests with manually assigned LOINC codes were added to the IHS master LOINC file.

Results

At the five sites, we were able to map from 63% to 76% of the local active laboratory tests to LOINC by using the mapping tool. Eleven percent to 27% of the tests were mapped manually. We could not assign LOINC codes to 6% to 19% of the laboratory tests due to incomplete or incorrect information about these tests (Table 1). Most of the tests with incomplete information were missing the laboratory test's unit of measure or site of specimen. For example, at one of the medical facilities, the test name "Creatinine" was in use without a unit of measure or site of specimen.

To validate the performance of our mapping tool, we tested it on a laboratory test file from a facility that did not participate in the pilot project and therefore had not contributed to the LOINC master file. Of 703 local laboratory tests in this file, we were able to map 569 (81%) of the tests to the LOINC standard using only the mapping tool (no manual mapping).

Table 1 ■ Results from Mapping Laboratory Data at Indian Health Service Medical Facilities to LOINC

Site	Total Active Tests No.	Automated Mapping No. (%)*	Manual Mapping No. (%)*	Uncodable Tests No. (%)*
1	1,050	800 (76)	111 (11)	139 (13)
2	1,098	687 (63)	204 (19)	207 (19)
3	1,315	872 (66)	360 (27)	83 (6)
4	1,213	765 (63)	244 (20)	204 (17)
5	291	205 (70)	36 (12)	50 (17)
Total	4,967	3,329 (67)	955 (19)	683 (14)

LOINC = Logical Observation Identifier Names and Codes.
*Percentages do not add to 100% due to rounding.

Discussion

In this pilot project, IHS laboratory data were successfully mapped to the LOINC standard using a combination of an automated mapping tool and manual mapping. Although we were able to map the majority of the laboratory tests to LOINC, having complete information about all laboratory tests would have improved performance. We did not find any significant mapping errors in our validation process, as the mapping tool was designed to look for an exact match between the test names in the master file and the local file of laboratory test names from each facility.

Most of the failures of the automated mapping tool were due to local naming choices. Either there was incomplete information about the test names or the test names differed because of the facility's naming convention. For example, the names of the test for blood platelets in use in different facilities was Plt, Plat, and Platelets. One of the major inconsistencies was seen in the units of measure of a laboratory test. Either the units were missing or were described differently (e.g., micrograms were referred to as mcg or mg). We consulted with the laboratory managers in the pilot facilities to get more details about the units of measure of tests not mapped by the mapping tool. Even after consultation with the laboratory managers, it was not possible for us to map all laboratory test names to LOINC. However, we were able to map more than two-thirds of laboratory test names using this semiautomated process. The results are consistent with other reported efforts.[5]

In the IHS information system, "blood" is listed as the specimen for all tests using serum or plasma. We assigned codes for "serum or plasma" to such tests unless the specimen was specified as arterial, venous, or capillary blood. We found a number of tests for which the test name in a facility's laboratory file was not in agreement with the site/specimen (e.g., test name, "Urine Creatinine"; site specimen, "Serum"). We did not assign a LOINC code to such tests. If a test had no units assigned and the only LOINC code available for that test name specified concentration in units/volume, the LOINC code with units was assigned (e.g., "Magnesium–mg/dl"). Some of the laboratory tests on urine or serum in active use at the five IHS facilities were screening tests. However, if screening was not mentioned in the method component of the test names, we assigned these tests LOINC codes "with unspecified method." Some of the tests in use at participating IHS medical facilities had incorrect synonyms in the facilities' laboratory files, which made it impossible to determine the

specific test being used and which LOINC code should be assigned. For example, one medical facility had named one of the hepatitis tests as "HepE Ab" and had used "HepBe Ab" as the synonym of the test. It was difficult for us to determine whether this local name was used for hepatitis E virus antibody or for the "e" antibody of hepatitis B virus.

Recent evaluation studies have revealed that there are gaps in content covered by existing terminologies[6] and that no single terminology can meet all needs. The scope of LOINC codes is intentionally restricted to the names of observations and does not include names for all microorganisms. To fully standardize the laboratory data, IHS will need to incorporate codes for test results, especially for microorganisms.

Following the success of the pilot phase of The Map to LOINC Project, the standardization effort has been expanded to other IHS medical facilities. Mapping of laboratory data to the LOINC code set will allow IHS software applications to extract and aggregate data for analysis and reporting of data from different medical facilities. For example, the IHS Government Performance and Reporting Act Plus (GPRA+) project is designed to develop disease-related taxonomies for data analysis and reporting using the LOINC code set.

Conclusions

At each of the five participating facilities, we mapped two thirds or more of the laboratory test names to LOINC using the automated mapping process. These results approximated other similar efforts.[5] The mapping tool's performance on a file from a facility whose laboratory data had not been used to create the master file suggests that this semiautomated process will achieve comparable results if expanded to other IHS medical facilities. Improvement in quality of data in the RPMS system will increase the percentage of tests mapped in the future. Standardization of laboratory names will allow IHS to aggregate laboratory data more easily for disease surveillance and clinical and administrative reporting efforts. The process that we successfully piloted may provide a model for standardization efforts in other health systems.

References ■

1. Emons MF. Integrated patient data for optimal patient management: the value of laboratory data in quality improvement. Clin Chem. 2001;47:1516–20.
2. Overhage MJ, Suico J, McDonald CJ. Electronic laboratory reporting: barriers, solutions and findings. J Public Health Manag Pract. 2001;7:60–6.
3. Lau LM, Johnson K, Monson K, Lam SH, Huff SM. A method for the automated mapping of laboratory results to LOINC. Proc AMIA Symp. 2000;472–6.
4. Forrey AW, McDonald CJ, DeMoor G, 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.
5. Zollo KA, Huff SM. Automated mapping of observation codes using extensional definitions. J Am Med Inform Assoc. 2000;7:586–92.
6. Bakken S, Campbell KE, Camino JJ, Huff SM, Hammond WE. Toward vocabulary domain specifications for Health Level 7-coded data elements. J Am Med Inform Assoc. 2000;7:333–42.