ABSTRACT
National policies in the United States require the use of standard terminology for data exchange between clinical information systems. However, most electronic health record systems continue to use local and idiosyncratic ways of representing clinical observations. To improve mappings between local terms and standard vocabularies, we sought to make existing mappings (wisdom) from health care organizations (the Crowd) available to individuals engaged in mapping processes. We developed new functionality to display counts of local terms and organizations that had previously mapped to a given Logical Observation Identifiers Names and Codes (LOINC) code. Further, we enabled users to view the details of those mappings, including local term names and the organizations that create the mappings. Users also would have the capacity to contribute their local mappings to a shared mapping repository. In this article, we describe the new functionality and its availability to implementers who desire resources to make mapping more efficient and effective.

Keywords: logical observation identifiers names and codes, crowdsourcing, controlled vocabulary, health information exchange; electronic health records, clinical laboratory information systems

ABBRVIATIONS:
IT, information technology; HIE, health information exchange; US, United States; LOINC, Logical Observation Identifiers Names and Codes; SNOMED CT, Systematized Nomenclature of Medicine—Clinical Terms; IHTSDO, International Health Terminology Standards Development Organization; EHR, electronic health record; CMS, Centers for Medicare and Medicaid Services; ELR, electronic laboratory reporting; RELMA, Regenstrief LOINC Mapping Assistant; RCMT, Reportable Conditions Mapping Table; VLDL, very-low-density lipoprotein; CDC, Centers for Disease Control and Prevention; LIS, laboratory information system; Qual, qualitative; NCnc, number concentration; ACnc, arbitrary concentration; WHO, World Health Organization; CTSI, Clinical and Translational Sciences Institute; ComMaps, field that displays the number of local test codes mapped to a given candidate LOINC; ComInst, field that displays the number of institutions that have mapped to a given candidate LOINC code; ACnc, arbitrary concentration; Pt, point in time; Ser/Plas, serum/plasma; QN, quantitative; Ord, ordinal; NCnc, number concentration; Probe amp tar, probe with target amplification; Prid, presence or identity; Imp, impression; Nom, nominal; LaCnc, log unit concentration; LnCnc, log number concentration; HIV, human immunodeficiency virus; YMDD, tyrosine-methionine-aspartate-aspartate

1Richard M. Fairbanks School of Public Health at Indiana University—Purdue University Indianapolis, Regenstrief Institute, Inc., and Center for Health Information and Communication, Department of Veterans Affairs, Veterans Health Administration, Health Services Research and Development Service, Indianapolis, IN, 2Regenstrief Institute, Inc., Indianapolis, IN, 3Indiana University School of Medicine, Regenstrief Institute, Inc., Indianapolis, IN

*To whom correspondence should be addressed. bedixon@iupui.edu

Semantic interoperability is the ability for an information technology (IT) system to receive information from another IT system and reliably apply its business rules to the information received. This definition represents a well-established, consensus-based view from the international health information exchange (HIE) community for shared messaging (syntax) and meaning (semantics) between health IT systems. The Center for IT Leadership estimates that among various health IT investments, introducing semantic interoperability would produce the greatest economic benefit to the United States (US) health system. To achieve semantic interoperability, the US health system must adopt and implement consistent clinical messaging and data standards that provide a framework and language for communicating shared meaning. Although messaging (syntax) is critically important, we focus, in this article, on the semantic aspects of interoperability—that is, how systems communicate shared meaning of clinical data.

Standard vocabularies for representing clinical data are now mature and have been internationally adopted. Logical Observation Identifiers Names and Codes (LOINC; Regenstrief Institute, Inc., Indianapolis, IN), for example, provide universal identifiers for laboratory tests and other
Laboratory QA

clinical observations. LOINC has been in development for 20 years and recently published its 53rd release. At present, LOINC has more than 36,000 registered users from 165 countries, and more than 25 countries have officially adopted it as a national standard. Similarly, the Systematized Nomenclature of Medicine—Clinical Terms (SNOMED CT) provides universal identifiers for organisms, substances, diseases, and other findings that may be recorded in the medical record or identified in test results. The International Health Terminology Standards Development Organisation (IHTSDO), which owns and develops SNOMED CT, currently has 27 member countries and has issued licenses to more than 5000 individuals and organizations. According to the certification criteria for electronic health record (EHR) systems under the Meaningful Use program administered by the Centers for Medicare and Medicaid Services (CMS) in the United States, LOINC, and SNOMED CT are required for communicating tests (LOINC) and results (SNOMED CT) in electronic laboratory reporting (ELR) for clinical operations and transmission of notifiable disease incidence to public-health authorities. Although clinical data standards are available and sometimes required when data are transmitted to another provider, most hospitals, laboratories, and physician offices continue to rely on local and idiosyncratic ways of identifying clinical observations (eg, laboratory tests and clinical measurements) and their results inside their EHR or laboratory information system. For example, an analysis of 7,000,000 ELR messages sent to health departments in 2 US states observed that less than 20% contained a LOINC code for identifying the test performed or a SNOMED CT code for identifying the test result. To users outside the assigning institution, local test codes are an enigma. Therefore, to understand the information in ELR messages, health departments must often translate inbound data into standardized LOINC and SNOMED CT codes. This translation process is often referred to as mapping and represents a set of terminology mediation strategies used by clinical and public health organizations to enable HIE within and among health enterprises.

Materials and Methods

The Existing RELMA Software Tool

Developed by the Regenstrief Institute and distributed free of charge (in its basic version) through the LOINC website (http://loinc.org), RELMA contains a variety of tools for mapping local terms to LOINC, including a robust search function that returns candidate LOINC codes and
automated functions to suggest candidate LOINC codes. The program enables users to load in their local terms, to use its search features to identify equivalent LOINC codes, and then to save the mappings to a file that can be used by the laboratory information system and EHR. RELMA is the primary way users interact with LOINC content. Currently, RELMA is downloaded approximately 11,000 times each year and accounts for more than 66% of all LOINC downloads. Many studies evaluating mapping local terms to LOINC use RELMA-assisted mapping as the gold standard.

**Development of RELMA Functions to Learn From the Crowd**

In 2012, we commenced work on enhancements to RELMA, namely, to the RELMA Community Mappings feature, as well as creating a Community Mapping Repository to hold data and make them available. To the existing search results area of the RELMA program, we added 2 new columns (Image 1). The ComMaps field displays the number of local test codes mapped to a given candidate LOINC, and the ComInst field displays the number of institutions that have mapped to that LOINC code. In the top line of Image 1, which is highlighted in blue and refers to LOINC term 13458-5, Cholesterol in VLDL (very-low-density lipoprotein), the values of ComMaps and ComInst are 29 and 10, respectively. These values represent that 10 different institutions have mapped 29 local terms to LOINC term 13458-5. By default, the new fields appear on the far right of the grid; however, users can customize the arrangement of the fields in RELMA to show them in any order they find convenient (eg, the far left or directly after the LOINC name).

When the user clicks on values in these new fields, a details screen appears for the candidate LOINC (Image 2). On this screen, users can view a list of local codes mapped to the candidate LOINC code and other detailed information. For example, Image 2 shows 29 rows for LOINC term 13458-5, Cholesterol in VLDL, all of which contain a value in the Local Code column representing the 29 local test codes (ComMaps). Also, 10 different names are listed under the Institution column representing the 10 institutions (ComInst).

In addition to enhanced searching functionality, RELMA now includes a mechanism by which users can contribute their mappings to the Community Mapping Repository (eg, become part of the Crowd). The goal of this feature is to bootstrap LOINC mapping in the community through a dialogue in which users view mappings from others and share their own. While in RELMA and logged into their LOINC user account, users can select Upload Mappings from the main screen. The program then guides them through a few screens to upload their mappings. This
functionality is designed to make it easy for individuals to share their work with the community.

The Community Mappings Repository is a new database and set of Web pages hosted by the Regenstrief Institute on the Internet. This resource contains the master list of all Crowd-contributed mappings. RELMA users log into their LOINC user account via the program, which then loads the latest community mappings for use in RELMA. This method enables real-time access to community mappings made available by the Crowd and has the advantage of being updatable independently from installing new versions of RELMA, which is updated twice a year.

Before the first release of the RELMA Community Mappings feature, we seeded the repository with contributions from the LOINC community using a manual process. Otherwise, the first users would have had no content to view. To gather as many local term mappings as possible before release, we sent a call for submissions to the LOINC e-mail listserv, which contained 3554 addresses at the time, and to other registered LOINC users. We populated the Community Repository with 27 LOINC mapping sets from 18 organizations in 5 countries that contained 91,960 local term mappings. Since the launch of LOINC Community Repository, we have received 8 additional mapping file contributions, bringing the current total to 102,484 total local mappings.

A Case Example Illustrating the New Functionality in RELMA

Let us examine a common scenario in which the new functionality in RELMA could assist a user. Given the Meaningful Use requirement to report positive laboratory results for notifiable diseases such as shigellosis, chlamydia, and gonorrhea using ELR, a hospital laboratory desires to map its local laboratory terms to LOINC. A medical laboratory scientist is assigned the responsibility and starts by downloading and installing the RELMA software. He or she might first compare any applicable state or local public health policies that define which diseases need to be reported with the list of diseases contained in the Reportable Conditions Mapping Table (RCMT), which is published and maintained by the Centers for Disease Control and Prevention (CDC) in the United States. Next, s/he would use the laboratory information system (LIS) to identify local laboratory codes that test for the presence of those conditions and to import that list into the RELMA software. Now, s/he can begin to use RELMA to match local concepts to the most appropriate LOINC code.

While working through the imported list of local laboratory terms, s/he attempts a search for “hepatitis B PCR.” RELMA initially returns a list of 36 LOINC codes; s/he is unsure which of them is the best fit. S/he notices that some of the LOINC codes are for nonserum specimens, so s/he adds “serum” to the search and narrows the list to 15 candidate LOINC codes. Eleven of these LOINC codes have at least 1 local code mapping in the Community Repository, so s/he sorts the results grid to see the most frequently chosen LOINC codes at the top (Table 1).

Looking at the top of the list helps the scientist weed out specialized tests, including those that look for specific mutations. At first glance, the first 3 candidate LOINC codes (42595-9, 29610-3, and 29615-2) look very similar. S/he clicks on the details page for the first one, LOINC term 42595-9, and notices that most of the local test names include viral load “quant” or “quantification” and have units of IU per mL. The laboratory at which this scientist works (hereinafter, the home laboratory) reports this specific hepatitis B test in IU per mL, and the local test names from most of the other institutions look similar. Hence, the scientist believes that this code is correct but wants additional verification. When s/he clicks on the details page for the second candidate LOINC, some of the local test names say “Qual” (qualitative). The test performed by the home laboratory quantifies the viral load, so the scientist knows that LOINC term 29610-3 is incorrect in this context. Comparing the first LOINC term 42595-9 and the third code in the list, LOINC term 29615-2, s/he observes that the LOINC names are identical except that term 42595-9 has a Property of “ACnc” and term 29615-2 has a Property of “NCnc”. Unsure of what these terms mean, s/he reviews the details page for term 29615-2 and notices that many of the local test names possess the qualifier “copies” or the units are “copies/mL”. Therefore, the term with LOINC Property of “NCnc” (number concentration) is for reporting number of copies per mL, whereas the “ACnc” (arbitrary concentration) is for reporting results with the international units established by the World Health Organization (WHO). In a final stage of confirmation, s/he verifies that the home laboratory uses a different test code for reporting the viral load in log10
IU per mL (represented by LOINC 48398-2, which has a Property of “LaCnc” for log unit concentration). S/he now feels very confident that the first LOINC code in the list is the correct code for how the home laboratory reports this hepatitis B test. S/he assigns the mapping and moves on to the next test code from the service catalog.

## Evaluating the Community Mapping Functionality

The new functionality was first released to the public with RELMA version 6.0 in December 2012; we then began promoting it at meetings and presentations and on the LOINC website. Just before launch, we conducted a convenience survey of the LOINC community about its perceptions of the proposed Crowd-based functionality. Since the launch, we have monitored adoption and use of the new functionality. New community mappings submitted to the LOINC team were collected, and we are conducting a follow-up convenience survey of users. Analysis of surveys conducted before and after the release of the new functionality, along with submitted mappings, are in progress.

Evaluation of the new functionality will focus on the perceived need among users for such functionality, their planned and actual use of community mappings, and the willingness of their organizations to contribute mappings to the LOINC community. Feedback from users will be critical for determining the evolution of the functionality within RELMA and for generating new ideas for how best to support the LOINC community. In addition, we will assess the validity of the community mappings using a combination of automated- and human-review processes.

## Discussion

Mapping local terms to standard vocabularies is necessary to enable semantic interoperability; however, it is complex, time consuming, and often costly. The findings of previous studies and our experience in supporting the LOINC standard have demonstrated that healthcare organizations need help in maintaining mappings from local terms to standard ones. In 2012 and 2013, the number of registered LOINC users grew by more than 14,000. Of those new users, 74% of them were from the United States. National policies in the United States requiring the use of LOINC not only contribute to this growth but also exert pressure on organizations to provide the mappings quickly because of short timelines. Our goal is to further support the LOINC community through enhanced functionality in RELMA, the primary software tool for mapping to LOINC. We believe that a Crowd-sourced repository of mappings will be valuable to users who are mapping their local terms to LOINC.

Vreeman et al previously demonstrated that a relatively small number of tests account for the vast majority of laboratory data. Anecdotally, RELMA users (especially novices) tell us they appreciate the functionality that

<table>
<thead>
<tr>
<th>LOINC</th>
<th>Component</th>
<th>Property</th>
<th>Timing</th>
<th>System</th>
<th>Scale</th>
<th>Method</th>
<th>ComMaps</th>
<th>ComInst</th>
</tr>
</thead>
<tbody>
<tr>
<td>42595-9</td>
<td>Hepatitis B virus DNA</td>
<td>ACnc</td>
<td>Pt</td>
<td>Ser/Plas</td>
<td>Qn</td>
<td>Probe.amp.tar</td>
<td>49</td>
<td>12</td>
</tr>
<tr>
<td>29610-3</td>
<td>Hepatitis B virus DNA</td>
<td>ACnc</td>
<td>Pt</td>
<td>Ser/Plas</td>
<td>Ord</td>
<td>Probe.amp.tar</td>
<td>30</td>
<td>8</td>
</tr>
<tr>
<td>29615-2</td>
<td>Hepatitis B virus DNA</td>
<td>NCnc</td>
<td>Pt</td>
<td>Ser/Plas</td>
<td>Qn</td>
<td>Probe.amp.tar</td>
<td>12</td>
<td>8</td>
</tr>
<tr>
<td>32366-7</td>
<td>Hepatitis B virus genotype</td>
<td>Prid</td>
<td>Pt</td>
<td>Ser/Plas</td>
<td>Nom</td>
<td>Probe.amp.tar</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>48398-2</td>
<td>Hepatitis B virus DNA</td>
<td>LaCnc</td>
<td>Pt</td>
<td>Ser/Plas</td>
<td>Qn</td>
<td>Probe.amp.tar</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>45161-7</td>
<td>Hepatitis B virus DNA</td>
<td>LnCnc</td>
<td>Pt</td>
<td>Ser/Plas</td>
<td>Qn</td>
<td>Probe.amp.tar</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>54210-0</td>
<td>Hepatitis B virus basal core promoter mutation</td>
<td>Prid</td>
<td>Pt</td>
<td>Ser</td>
<td>Nom</td>
<td>Probe.amp.tar</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>59052-1</td>
<td>HIV 1 + hepatitis C virus RNA + hepatitis B virus DNA</td>
<td>ACnc</td>
<td>Pt</td>
<td>Ser/Plas</td>
<td>Ord</td>
<td>Probe.amp.tar</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>43279-9</td>
<td>Hepatitis B virus YMDD mutation</td>
<td>Pr</td>
<td>Pt</td>
<td>Ser/Plas</td>
<td>Ord</td>
<td>Probe.amp.tar</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>33633-9</td>
<td>Hepatitis B virus precore TAG mutation</td>
<td>Pr</td>
<td>Pt</td>
<td>Ser</td>
<td>Ord</td>
<td>Probe.amp.tar</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>42322-8</td>
<td>Hepatitis B virus S + Pol gene</td>
<td>Prid</td>
<td>Pt</td>
<td>Ser/Plas</td>
<td>Nom</td>
<td>Probe.amp.tar</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
limits search results to only the most common tests by volume. We designed the RELMA Community Mappings feature because we hypothesized that the number of organizations and local test codes mapped to a particular LOINC code would provide another frequency-based statistic that could inform the mapping process. Planned evaluations of this new functionality over the next year will provide evidence on the perceptions and usage of these enhancements by LOINC community members.

A potential limitation of a Crowd-driven repository of LOINC mappings is the ambiguity surrounding whether a given local term to LOINC pair is appropriate. The complexities of mapping make it difficult to assess whether the mappings submitted to the community repository are accurate. For example, it is well known that local laboratory test names often lack information important for LOINC mapping.\textsuperscript{15} In current work, we are analyzing the validity of the community mappings using a combination of automated- and human-review processes.

**Conclusion**

Mapping local terms to standard vocabularies remains a challenge but is necessary to enable semantic interoperability between the myriad health information systems used across hospitals, laboratories, clinics, and other health care facilities. Mapping activities require people, processes, and informatics tools. Crowd-driven knowledge on the most appropriate standard terms may provide value for the many individuals who create and maintain mappings to these standards. Future work is necessary to tap into the wisdom of the Crowd and to harness collective knowledge to make creation and maintenance of mappings easier and more efficient.

**Acknowledgments**

This work was made possible in part by support from the Indiana Clinical and Translational Sciences Institute (CTSI); funded in part by grant no. TR000006 from the National Institutes of Health, National Center for Advancing Translational Sciences, Clinical and Translational Sciences Award [B.E.D. and D.J.V.], a contract (HHSN2762008000006C) from the National Library of Medicine (NLM) (D.J.V. and J.H.), and by the Department of Veterans Affairs, Veterans Health Administration, Health Services Research and Development Service CIN 13-416 (B.E.D.). Dr. Dixon is currently a Health Research Scientist at the Richard L. Roudebush Veterans Affairs Medical Center in Indianapolis, Indiana. The content is solely the responsibility of the authors and does not necessarily represent the official views of the CTSI, the NLM, the Department of Veterans Affairs, or the US Government. **LM**

**References**


