

Automated Mapping of Local Radiology Terms to LOINC

Daniel J. Vreeman, P.T., D.P.T. and Clement J. McDonald, M.D.

Regenstrief Institute, Inc. and Indiana University School of Medicine, Indianapolis, IN

ABSTRACT

We developed an automated tool, called the Intelligent Mapper (IM), to improve the efficiency and consistency of mapping local terms to LOINC. We evaluated IM's performance in mapping diagnostic radiology report terms from two hospitals to LOINC by comparing IM's term rankings to a manually established gold standard. Using a CPT®-based restriction, for terms with a LOINC code match, IM ranked the correct LOINC code first in 90% of our development set terms, and in 87% of our test set terms. The CPT-based restriction significantly improved IM's ability to identify correct LOINC codes. We have made IM freely available, with the aim of reducing the effort required to integrate disparate systems and helping move us towards the goal of interoperable health information exchange.

INTRODUCTION

The goal of anywhere, anytime medical information exchange is impeded by the plethora of local conventions for identifying health data in separate electronic systems. It is now widely recognized that we must adopt data and vocabulary interoperability standards to overcome the enormous entropy in integrating these disparate sources.^{1,2} Logical Observation Identifier Names and Codes (LOINC®) is a universal code system for identifying laboratory and other clinical observations^{3,4} that has been adopted in both the public and private sector, in the United States and internationally.⁴ Mapping local observation concepts to LOINC bridges the many islands of health data that exist by enabling grouping of equivalent results for a given test from isolated systems.

The work of manually mapping local terms to a standardized vocabulary is a large barrier to the adoption of such standards because of the amount of time and domain expertise it requires. Baroto et al⁵ described the process of combining laboratory data from multiple institutions where each site locally mapped to LOINC. The authors concluded that mapping to LOINC can be complex due to differences in mapping choices between sites, and asserted that a "quality standard coding procedure" is required for aggregating data without detailed human inspection. Even partially automating the mapping process has the potential to increase efficiency and consistency. Lau et al⁶ developed an automated technique for mapping local laboratory terms to LOINC. They employed parsing and logic rules,

synonymy, attribute relationships, and the frequency of mapping to a given LOINC code. Their report was a methodology paper that did not cover the technique's accuracy for finding true matches.

Integrating non-laboratory information is an important, but challenging prerequisite to a comprehensive health information exchange. Though radiology reports are often produced by electronic systems, they are frequently unavailable to providers at the time of a clinical visit.⁷ In primary care settings, providers have reported that missing clinical information, including radiology reports, was often located outside of their clinical system.⁷ A first step toward resolving the problem of missing information is the integration of disparate sources where the data reside, within and across health care organizations. Mapping radiology reports may, however, be less burdensome than mapping laboratory tests because their names are longer and more descriptive—a feature that could be exploited with automated tools.

The Regenstrief Institute, Inc. has developed a program called the Regenstrief LOINC Mapping Assistant (RELMA). RELMA is freely available (<http://loinc.org>), and contains tools for both browsing the LOINC database and mapping local tests (and other observation codes) to LOINC on a one-at-a-time basis. Regenstrief has also developed an additional RELMA program, called the Intelligent Mapper (IM), which automatically generates a ranked list of candidate LOINC codes for each local term of a submitted set. In its initial form, IM did not perform well, but we have further developed and enhanced it in many ways, including an option for narrowing the search space, based on the CPT® codes* contained in radiology system master files. Furthermore, we added synonymy and additional radiology content in the LOINC database. Here we report the functions of the IM and a formal analysis of its performance in the domain of diagnostic radiology tests.

METHODS

LOINC This study required, and was done in parallel with, an expansion of LOINC content for radiology report names. LOINC's expansion was necessary for us to accommodate HL7 messages from six new radiology sources that were flowing into our local health information infrastructure, and coincided with large submissions of radiology terms to LOINC from the Department of Defense. We used online and print

* CPT is a registered trademark of the American Medical Association

versions of radiology textbooks and expert advice to hone the LOINC radiology terms. We increased the number of radiology test terms in LOINC from 1,179 to 3,780 and added over 550 synonyms for organs, views, and methods.

Each LOINC code is linked to a name consisting of six major, and up to four minor attributes:⁴

1. Component (e.g. "View AP")
2. Property (e.g. "Finding")
3. Time (e.g. "Point in time")
4. System (e.g. "Shoulder")
5. Scale (e.g. "Narrative")
6. Method (e.g. "X-ray")

LOINC names for radiology reports are constructed similarly to the names for other clinical observations. Radiology names include entries for all six attributes, but typically the Property, Time, and Scale are constants and do not contribute to the term meaning. We limited the scope of this study and the expansion to diagnostic radiology tests; we will address interventional radiology and nuclear medicine terms in a later effort.

Radiology Naming Conventions In contrast to laboratory test names, radiology report names are typically longer and richer. Because their names are descriptive, understanding their meaning rarely requires research and dialogue with the source system as is sometimes needed to clarify the meaning of laboratory test names. The radiology codes and names we receive in our community-wide information exchange⁸ reveal that many radiology systems, especially those with geographically separate facilities, invent multiple codes for one test to distinguish those facilities. These tests usually have the same name (or a close variant), but their codes include a prefix to identify the facility location. Thus, the total number of distinct tests that need mapping may be almost one third of the number in the radiology master file. To reduce mapping effort, you can condense the files into unique report names. After mapping these terms, you can then join them back to all the location-based variants.

We have also observed that radiology systems vary in whether and how they represent anatomical laterality in report names. For studies that examine

the limbs and other bilateral body parts, LOINC includes terms for left, right, bilateral, and unspecified laterality so that it can accommodate these variations. Local systems that do not distinguish tests by laterality can map their terms to the unspecified LOINC term.

In their master files, most radiology systems include a Current Procedure Terminology⁹ (CPT®) code along with their local code and report name. Some of the terms in radiology master files (and in CPT) represent billing-specific concepts (e.g. "ea addtl vessel after basic" and "fluoro up to 1hr") that would not be used as report names. Since LOINC provides codes and names for labeling reports, we do not create LOINC codes for such concepts, and such terms were excluded from this project. Using our home institution's licensed CPT file, we created a LOINC to CPT mapping table for use in IM from linkages in our master dictionary and manual review.

Intelligent Mapper The Intelligent Mapper (IM)¹⁰, is an automated tool for producing a ranked list of candidate LOINC terms for each local term in a submission file. IM identifies candidate LOINC codes by counting the number of matches between words in the local term name and words (or synonyms) in the formal LOINC term names. Before doing the matching, it expands the words in the local term name into a tree of synonyms. For example, "CHEST MRA" becomes "CHEST, (MRI ANGIO, MRA)." IM counts exact-string word matches for all possible combinations of words and synonyms (e.g. "CHEST, MRI ANGIO" and "CHEST, MRA" are counted separately), and then uses the best count as the first part of its match score. IM ranks the candidate LOINC terms for relevance first by the number of words matched (the more the better), and second on the total number of words in the LOINC term (the fewer the better). If no words in the local term match to any in LOINC, IM does not return any candidate terms. A sample of a report that can be generated from an IM run is given in Figure 1.

We hypothesized that we could improve the accuracy of finding correct LOINC matches by using the radiology term-to-CPT linkage to narrow the search space. This concept is analogous to limiting

Figure 1. Sample of an Intelligent Mapper report showing ranked candidate LOINC codes. LOINC codes are ranked first by descending number of words matched, and then by ascending number of words in the candidate LOINC name. The (*) indicates the Gold Standard mapped LOINC. Words that appear within square brackets [] are local words that matched in the LOINC name.

LAB SECT:	HOSPITAL A	UNITS:	PROPERTIES:
OBX:	240	OBX DESC:	CT CHEST WO C
TRANS:	(NCNC, CT), CHEST, WO, (LITTLE C, CONTRAST, C)		
CPT4:	71250:CMPT TOMOGRAPHY THORAX, WITHOUT CONTRST MATERIAL		
*	4 words matched out of 5 LOINC words	29252-4:MULTISECTION^*[] [CONTRAST]:FIND:PT:[CHEST]:NAR:[CT]	
	4 words matched out of 7 LOINC words	37441-3:MULTISECTION HIGH RESOLUTION^*[] [CONTRAST]:FIND:PT:[CHEST]:NAR:[CT]	
	3 words matched out of 5 LOINC words	36534-6:MULTISECTION^*[] [CONTRAST]:FIND:PT:STERNUM:NAR:[CT]	
	3 words matched out of 6 LOINC words	37282-1:MULTISECTION^*[] [CONTRAST]:FIND:PT:STERNOCLAVICULAR JOINT:NAR:[CT]	
	2 words matched out of 3 LOINC words	24627-2:MULTISECTION:FIND:PT:[CHEST]:NAR:[CT]	

matches of tests with numeric values to LOINC terms that are consistent with the units of measure.

A LOINC term may link to multiple CPT codes, and a broad CPT code may link to multiple LOINC codes. Because of this, IM will look for a match with any of the LOINC codes that share a CPT code with the local term. We developed a user option to loosen the strictness of the CPT matching to accept a match of the first “n” digits of the CPT code instead of a perfect match. This feature accommodates the potential to accumulate CPT codes that shift by a digit from year to year. When IM enforces CPT matching, it returns only candidate LOINC terms that match the CPT code associated with the local term. However, the match requirement is not enforced when the local term does not have a CPT code associated with it.

General Mapping Procedure Routine mapping of local test codes to LOINC involves three steps¹⁰. The first step is to establish a set of mapping rules to guide the process. The second step uses tools in RELMA to automatically scan the submission file to find all words that are not part of any LOINC term name or synonyms. Most often, these “unknown” words are unconventional abbreviations of known words. RELMA expects the user to examine each distinct unknown word and specify the LOINC word to which it is equivalent. The system then translates the unknown words into LOINC words wherever indicated. This feature helps resolve the problem of idiosyncratic abbreviations that are not likely present in a standard synonym list such as LOINC’s. In our experience, this step takes little time and improves IM’s accuracy by approximately eight percentage points. The final step involves choosing the correct LOINC code, using either RELMA’s term-by-term browser or IM to generate a short, ranked list of candidate LOINC codes from which to choose. Both RELMA’s browser and IM have a number of optional user search parameters for limiting the scope of the candidate LOINC terms returned. We followed this general procedure for both our gold standard mapping and the runs to evaluate IM’s performance.

Data Sources We used the master files of two radiology systems in central Indiana to assess IM’s performance. Both sources came from not-for-profit urban hospitals. We used hospital A’s master file as a *development set*, including debugging the IM program and enhancing LOINC’s radiology content. The mapping success for hospital A reflects successive tuning of the program and the addition of LOINC terms and synonymy to provide coverage for all of these terms. We used hospital B’s master file as our *test set* to analyze IM’s performance, because we

had not seen the terms prior to this study and did not use them in development.

Our development master file contained a total of 3,698 radiology study codes, while our test file contained 1,848 codes. Both master files contained different codes for the same test done at different facilities within their health system. We preprocessed the term files with a Perl script to remove punctuation and squeeze down to only records with unique test names. This process yielded 1,129 unique terms for our development set and 613 unique terms for our test set. Because this effort was focused on LOINC’s diagnostic radiology content, we excluded terms from the hospital master files that represented interventional radiology, nuclear medicine, or the pure billing terms described above. These exclusions reduced the development set to 539 terms and the test set to 427 terms. A CPT code was included for 95% of the terms from each hospital’s master file. CPT codes do not have a one-to-one correspondence with local radiology codes. Terms in our development set were linked to 327 unique CPT codes, with a mean of 1.6 terms per CPT code. Terms in our test set were linked to 314 unique CPT codes, with a mean of 1.3 terms per CPT code. Our test set contained test names that distinguished anatomic laterality (i.e. right and left), whereas our development set did not.

In order to characterize IM’s mapping success in relation to the volume of live clinical data exchange, we also extracted the test codes and names for one month (mid-January to mid-February 2005) of radiology messages from our clinical data repository for both of these institutions.

Establishing a Gold Standard A domain expert manually established a gold standard mapping for both hospitals against which to compare IM’s results. The process followed the recommended procedures outlined in the RELMA User’s Manual¹⁰ and described above. We mapped these terms to our in-house version of the LOINC database, version 2.14+, which contains about 300 more radiology terms than the last LOINC public release (version 2.14).

Our mapping rules stated that the gold standard mapping would be an exact correspondence from local term to LOINC. If no LOINC match was identified, we counted that local term as unmapped. Per the standard mapping procedure, we identified the words in our local submissions unknown to LOINC, and translated them into known LOINC words where possible. For example, we translated the words “tomogrn” and “tomogm” into “tomogram”, the word that LOINC knows. We provided translations for 61 words in the development set terms and 67 words in the test set terms. In the final step, the domain expert used RELMA to search for

LOINC code matches on a term-by-term basis. Reliability for manual mapping was not established.

Intelligent Mapper Processing We processed the terms from both hospitals with IM to identify candidate LOINC terms. In order to evaluate the relative value of our CPT-restriction feature, we ran IM on each hospital's file four times: (1) without CPT restriction enabled, (2) CPT restriction for all five digits, (3) CPT restriction for the first four digits, and (4) CPT restriction for the first three digits. For all runs, we used the vocabulary translations for unknown words identified in the gold standard mapping and selected the user option to limit the search to only LOINC codes in the "radiology studies" class. All analyses ran on a 1600 MHz computer with 1.0 GB of RAM.

Measures For each IM run, we calculated IM's ability to include the correct LOINC code in its top rankings and recorded computational costs. We chose to limit the list of candidate LOINC codes to the top five because preliminary analyses indicated that no additional matches were found in LOINC codes ranked between five and ten. We evaluated IM's accuracy in identifying correct matches for the entire set of local terms and in the context of one month of clinical messages for each institution. For incorrect matches, we classified the reasons for failure.

RESULTS

The gold standard mapping identified a true LOINC match for all 539 terms in our development set and for 393 of 427 terms in our test set. The 34 unmatched terms in our test set included those with laterality variants currently not in LOINC (n=14), and other tests currently not in LOINC (n=20).

The summary results for each IM run on the two data sets are given in Table 1. The fact that *all* of the terms in our development set had a LOINC match is an artifact of our using it to expand LOINC. We calculated IM's performance on our test set both for all terms and for only terms that had a LOINC match.

Overall, when IM used the full digit CPT-based restriction, it ranked the correct LOINC code first in 90% of development set terms and in 80% of test set terms, versus 71% of development set terms and 72% of test set terms when the CPT restriction was *not* used. The difference in the success of matching between using and not using the CPT restriction was significant for both the development set ($\chi^2 P < 0.0001$) and the test set ($\chi^2 P < 0.01$). There was no significant difference in success of matching between ranking the correct LOINC code in the top three versus top five for either set, whether using the CPT restriction or not ($\chi^2 P > 0.05$). In the subset of our test set terms for which a LOINC code exists, IM ranked the correct LOINC first in 87% of terms using the full

CPT restriction, and in 78% of terms without using the CPT restriction. These results represent IM's recall (correct matches made/correct matches possible). Loosening the CPT restriction to match on four or three digits decreased IM's recall by one to two percentage points. When IM used the CPT restriction for our test set, correctly it did not return any candidate LOINC codes for 10 of the 34 terms that had no LOINC match. Thus, IM's precision (correct matches/total matches made) was 82%. Employing the CPT restriction almost tripled the computational cost; the processing time for our larger (development) set increased from 29 minutes (no CPT restriction) to 89 minutes (full CPT restriction).

Table 1. Performance of Intelligent Mapper (IM) for identifying correct LOINC matches.

IM Parameter	Correct LOINC Codes Identified		
	Top Ranked % (n)	Rank in Top 3 % (n)	Rank in Top 5 % (n)
<i>Hospital A* (539 terms, all with a LOINC match)</i>			
Full CPT restriction	90 (485)	97 (522)	97 (525)
No CPT restriction	71 (380)	85 (460)	89 (482)
<i>Hospital B* (393 terms with a LOINC match)</i>			
Full CPT restriction	87 (340)	93 (367)	94 (368)
No CPT restriction	78 (307)	90 (355)	92 (364)
<i>Hospital B* (all 427 terms)</i>			
Full CPT restriction	80 (340)	86 (367)	86 (368)
No CPT restriction	72 (307)	83 (355)	85 (364)

*Hospital A represents our development set and Hospital B represents our test set.

After manual review, we categorized IM's failures for both term sets when using the full digit CPT restriction. We found that 65% were due to vocabulary discrepancies between the local terms and LOINC, 32% were due to discrepancies in linkages between local terms, CPT codes, and LOINC codes, and three percent were due to required external knowledge that was not contained in the term name.

The one-month extraction of radiology messages for hospital A contained 21,357 reports which included 47% (n=254) of the distinct diagnostic radiology studies listed in its master file. The monthly total was similar for hospital B; it contained 20,922 reports which included 59% (n=253) of its diagnostic radiology studies. If the LOINC code that IM ranked first (with the full CPT restriction) was assigned to the radiology report codes in hospital A, 95% of the reports (n=20,315) in the message extract would be correctly mapped. For hospital B, if IM assigned LOINC codes in this way, 91% of the reports (n=19,094) in the message extract would be correctly mapped. The correct LOINC code was ranked in the top three for the names of 99% (n=21,139) of the reports from hospital A, and in 95% (n=19,797) of the reports from hospital B.

DISCUSSION

IM identified correct LOINC matches for diagnostic radiology terms with accuracy comparable to the best reported automated techniques for mapping between laboratory systems.¹¹ Because this tool so often includes the correct match first in its list of ranked candidates, domain experts can review the results and choose the right term much faster than with a term-by-term search. We found no significant increase in LOINC code matches past an IM rank of three, thus domain experts can focus their attention on a short list of candidate terms. Furthermore, these benefits are achieved with freely available software.¹⁰

Employing the CPT restriction significantly improved the accuracy of IM for both term sets. IM's accuracy was highest when we used all digits of the CPT code. We had hypothesized that loosening the match criteria to fewer CPT digits would improve its performance, but in net, it worsened it. Manual review of the ranking results for these scenarios showed that loosening the restriction improved the matching for some terms and worsened it for others. The increased accuracy of IM with the CPT restriction came at a higher, but affordable computational cost because IM runs unattended. We have not yet optimized the IM algorithm for speed, and hope to improve on its efficiency with further tuning. Eight percent of our test set terms had no LOINC match. The percent of terms in other systems with no LOINC match will likely decrease as LOINC continues to expand its coverage of radiology reports.

This study has several limitations. We restricted our analysis to diagnostic radiology tests, but as LOINC expands to more exhaustively cover nuclear medicine and interventional radiology reports, we will apply this analysis to those domains as well. We chose to use radiology system master files as the starting point of the mapping process, yet only about half of the terms from each file appeared in the one month extraction of clinical messages. This suggests that a sensible approach to mapping may be to use the clinical message stream as the first source, rather than attempting to map the entire master file at once.

It was not surprising that the IM identified a greater percentage of correct matches in our development set than in our test set. IM's performance on terms from our test set is likely more representative of what we would expect at other institutions. Presently, we are not distributing CPT code mapping in the RELMA package because of licensing restrictions. The National Library of Medicine is pursuing a project to create a validated LOINC to CPT mapping for broader dissemination. When such a mapping becomes available, we will distribute it under the specified terms and conditions.

CONCLUSION

This study demonstrates the feasibility of using freely-available software for automatically identifying a ranked list of candidate LOINC codes. Because vocabulary mapping is a time-consuming step in interfacing systems, reduced effort here may help us move more quickly towards the goal of interoperable health information exchange.

Acknowledgements

The authors thank Kathy Mercer for LOINC development and John Hook and Brian Dixon for programming support. This work was performed at the Regenstrief Institute, Inc. in Indianapolis, IN and was supported in part by the National Library of Medicine (T15 LM-7117 and N01-LM-3-3501). The publication was supported in part by Grant/Cooperative Agreement H75/CCH520501-01 from the Centers for Disease Control and Prevention. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the Centers for Disease Control and Prevention.

References

1. McDonald CJ. The barriers to electronic medical record systems and how to overcome them. *J Am Med Inform Assoc.* 1997;4(3):213-221.
2. Stead WW, Kelly BJ, Kolodner RM. Achievable steps toward building a National Health Information infrastructure in the United States. *J Am Med Inform Assoc.* 2005;12(2):113-120.
3. 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(1):81-90.
4. McDonald CJ, Huff SM, Suico JG, et al. LOINC, a universal standard for identifying laboratory observations: a 5-year update. *Clin Chem.* 2003;49(4):624-633.
5. Baorto DM, Cimino JJ, Parvin CA, Kahn MG. Combining laboratory data sets from multiple institutions using the logical observation identifier names and codes (LOINC). *Int J Med Inform.* 1998;51(1):29-37.
6. Lau LM, Johnson K, Monson K, Lam SH, Huff SM. A method for the automated mapping of laboratory results to LOINC. *Proc AMIA Symp.* 2000:472-476.
7. Smith CP, Araya-Guerra R, Bublitz C, et al. Missing clinical information during primary care visits. *JAMA.* 2005;293(5):565-571.
8. Overhage JM, Tierney WM, McDonald CJ. Design and implementation of the Indianapolis Network for Patient Care and Research. *Bull Med Libr Assoc.* 1995;83(1):48-56.
9. American Medical Association. *Current Procedural Terminology CPT 2004.* Chicago, IL: American Medical Association; 2003.
10. LOINC Committee. *RELMA® users' manual.* Indianapolis, IN:Regenstrief Institute, 2004. Available at: <http://loinc.org>. Accessed March 2005.
11. Zollo KA, Huff SM. Automated mapping of observation codes using extensional definitions. *J Am Med Inform Assoc.* 2000;7(6):586-592.