
Research and Applications

Mis-mappings between a producer's quantitative test codes and LOINC codes and an algorithm for correcting them

Clement J. McDonald¹, Seo H. Baik ¹, Zhaonian Zheng¹, Liz Amos¹, Xiaocheng Luan¹, Keith Marsolo², and Laura Qualls²

¹Lister Hill National Center for Biomedical Communications, National Library of Medicine, National Institutes of Health, Bethesda, Maryland, USA and ²Department of Population Health Sciences, Duke University School of Medicine, Durham, North Carolina, USA

Corresponding Author: Clement J. McDonald, MD, Lister Hill National Center for Biomedical Communications, National Library of Medicine, National Institutes of Health, 8600 Rockville Pike, Bethesda, MD 20894, USA; clemmcdonald@mail.nih.gov

Received 15 September 2022; Revised 17 October 2022; Editorial Decision 19 October 2022; Accepted 31 October 2022

ABSTRACT

Objectives: To access the accuracy of the Logical Observation Identifiers Names and Codes (LOINC) mapping to local laboratory test codes that is crucial to data integration across time and healthcare systems.

Materials and Methods: We used software tools and manual reviews to estimate the rate of LOINC mapping errors among 179 million mapped test results from 2 DataMarts in PCORnet. We separately reported unweighted and weighted mapping error rates, overall and by parts of the LOINC term.

Results: Of included 179 537 986 mapped results for 3029 quantitative tests, 95.4% were mapped correctly implying an 4.6% mapping error rate. Error rates were less than 5% for the more common tests with at least 100 000 mapped test results. Mapping errors varied across different LOINC classes. Error rates in chemistry and hematology classes, which together accounted for 92.0% of the mapped test results, were 0.4% and 7.5%, respectively. About 50% of mapping errors were due to errors in the property part of the LOINC name.

Discussions: Mapping errors could be detected automatically through inconsistencies in (1) qualifiers of the analyte, (2) specimen type, (3) property, and (4) method. Among quantitative test results, which are the large majority of reported tests, application of automatic error detection and correction algorithm could reduce the mapping errors further.

Conclusions: Overall, the mapping error rate within the PCORnet data was 4.6%. This is nontrivial but less than other published error rates of 20%–40%. Such error rate decreased substantially to 0.1% after the application of automatic detection and correction algorithm.

Key words: PCORnet, LOINC, local test code, mapping errors

INTRODUCTION

In 1996, Clinical Chemistry introduced Logical Observation Identifiers Names and Codes (LOINC) database¹ with 6000 codes and names for laboratory, and other clinical tests. LOINC database now includes nearly 100 000 laboratory tests survey instruments, clinical assessments, narrative reports, and panels most of which

are specified for use in federal United States Core Data for Interoperability (USCDI) guidance.² At the beginning, the laboratory tests were the primary content of the LOINC database and today they still represent the plurality of today's LOINC codes. The availability of these standard codes in electronic result messages enables clinical data delivered to medical records and research databases, to

be presented as a unified whole, over time and across health care systems.

These codes are crucial to data integration within health information exchanges (HIEs), such as the Indiana Health Information Exchange,³ personal health records such as Apple's HealthKit,⁴ research networks such as the Patient-Centered Outcome Research Network (PCORnet) and Observational Health Data Science Informatics (OHDSI), and Informatics for Integrating Biology & the Besides (i2b2).⁵ LOINC codes have been translated into 19 linguistic variations and have been adopted in whole or part by 193 countries.⁶ In the United States, it is required by Food and Drug Administration (FDA), Centers for Disease Control and Prevention (CDC), Office of the National Coordinator for Health Information Technology (ONC), and Centers for Medicare & Medicaid Services (CMS) for many purposes.⁷

However, with increasing use, questions about the accuracy of the mappings between local laboratory tests and LOINC codes have arisen.

Laboratory results are often presented as matrices of tests belonging to a panel or in time-oriented flowsheets, both of which constrain the space available for a test name; so, contractions, acronym, and the removal of the specimen type, for example, serum or blood, from the test name, when it is the predominate specimen type for a give test, all occur. These kinds of laboratory test name shortenings make mapping more difficult and prone to error. One recent paper reported error rates as high as 20% among mappings between local laboratory tests and standard LOINC codes.⁸ This report included mappings to 10 tests, mostly from the coagulation laboratory, and the sample was taken from a laboratory survey with a low (5%) response rate, so the results may not be representative. A study with an elaborate design reported a 41% mis-mapping rate,⁹ but it did not make clear how their 9 different potential reasons for mismatch related to the 136 mis-mappings, whether each of the 136 was an independent observation or they were correlated due to errors, for example, in a master file that got propagated through many repeated instances at a given institution. To obtain a better estimate of mapping errors at large, and their genesis, we obtained a large sample of mappings from a research network which included 183 million instances of mappings from 2 independent source networks, each of which included mappings from many institutions. Of note, the mappings from these networks were not necessarily done by laboratory experts. Here, we report the overall mapping error rates and their breakdowns by multiple categories. We also report our success with correcting mis-maps broken down in different ways.

MATERIALS AND METHODS

Content obtained from PCORnet

PCORnet¹⁰ is a large research consortium which currently includes 60+ contributing organizations, called DataMarts. Each contributing organization holds and controls its own data, but the content from all DataMarts is organized and coded the same way to enable research questions to be asked across DataMarts. In early 2016, a PCORnet survey of their 34 DataMarts fed by 12 different EMR system revealed huge numbers of distinct measurement names in the DataMarts; 32% counted between 5000 and 10 000 distinct measurement names and 36% carried more than 10 000, per mart.¹¹ However, many of the names were variations on the name of one measurement. LOINC code was needed to link a given test within,

and across, the different DataMarts. We had access to their mappings from only 2 Data Marts at the start of our study. Taken together, the 2 marts included 183 million mapped test results for 4651 distinct measurements. PCORnet provided us with the local test name and local code. For privacy protection, they did not provide full data on specific mapped test results. Instead, they provided us with 1 example for each unique combination of source's dummy DataMart ID, source laboratory's test name, unit string, occasionally a separate specimen identifier, the test results' calendar year, the median value of each quantitative result in each combination and the LOINC code to which this combination was mapped. They also provided us with the count of records in each combination with which we derived error rates weighted by testing volume. They provided null counts for attributes of combinations based on fewer than 11 mapped test results to protect against possible reidentification. We imputed a value of 5 for these null counts assuming they were between 1 and 10.

We received a compressed set of tests mappings by taking only unique combinations of the attributes mentioned above. We removed duplicate records that differed only by calendar year and added their counts to the appropriate surviving record. No record that we received included any Personal Identifiable Information or individual test result value.

We converted all the raw unit of measure strings we obtained to formal Unified Code for Units of Measures (UCUM)¹² codes using tables that linked raw unit of measure strings to their correct UCUM code. Two kinds of raw unit of measure strings required extra work. Some raw unit of measure strings contain "U" alone or "U" as the numerator. "U" per IUPAC¹³ represents the amount of enzyme that will convert 1 μm of substrate per minute. However, some laboratories use "U," to mean arbitrary units. We changed UCUM unit for tests that were really measured in arbitrary units rather than the enzyme unit to [arb*U] the correct UCUM code. The second problem was the mislabeling of the units for some enzyme measures as International Units (IU). We could identify analytes that were real enzyme measures by their name which, with few exceptions, ended in "ase." We reviewed all analyte names to find enzymes that lacked a terminal "ase" (eg, renin) and to distinguish tests which were antibodies to enzymes rather than the enzymes per se.

Rules for detecting errors and correcting mis-mappings

LOINC observation terms are constructed from atoms called Parts.¹⁴ Mapping errors were usually due to an inconsistency between the content of the local test name and a single part of the LOINC term, and we classify such errors based on that part type. A few were due inconsistencies with 2 parts of a LOINC name. To simplify the classification of these, we treated them as a single part error and used the first of the 2 to classify it.

Some local terms represented test concepts for which no corresponding LOINC term yet existed.

We counted them separately in Table 1 and reported mis-mapped error rates without them because no correct mapping existed.

We used manual review supplemented by automated review to define mapping errors. We could automatically detect mis-mapping through 4 processes listed below:

Table 1. Counts and percent of records included and excluded from analysis because they are not quantitative value measures

	N (%) unique tests	N (%) mapped test results
Total included	3029 (65.1)	179 569 051 (94.8)
Total excluded	1622 (34.9)	9 836 382 (5.2)
Breakdown of excluded		
Nonquantitative	1486 (32.0)	9 723 848 (5.1)
Reference ranges, not test results	3 (0.1)	1622 (0.0009)
Local tests name has no corresponding term in the LOINC data base	45 (1.0)	36 044 (0.02)
Local test name all digits	86 (1.8)	74 799 (0.04)
Contradiction between local test name and local UoM	2 (0.0)	69 (0.00004)
Grand total	4651	189 405 433

Inconsistencies between qualifiers of the analyte in the local test name and qualifiers of the analyte in the mapped to LOINC term

Local name and LOINC analyte names often included qualifiers such as bound, unbound and others (see [Supplementary Table S1](#) for the list we used). If local lab name included a given qualifier and the mapped to LOINC name did not or vice versa, the mapping was wrong, and we flagged it as such.

Inconsistencies between specimen type in the local test name and the specimen type in the mapped to LOINC term (in LOINC term, the part that carries the specimen type is called the system)

The specimen's name was embedded in the local test name, for example, creatinine, dialysis fluid. We created a list of specimen names that we found embedded in the local test names (see [Supplementary Table S2](#)) and assumed this name was the correct one when it disagreed with the specimen part of the mapped to LOINC term. PCORnet also provided a column for specimen that carried a specimen identifier for 40% of the tests and 19% of the mapped test results. When either the embedded specimen name or the name in the PCORnet specimen column disagreed with the specimen in the mapped to LOINC term, we assumed the former represented the correct specimen and labeled it a mis-mapping. In the case that no specimen was asserted in either the local test name or the PCORnet specimen column, and the LOINC class was chemistry, toxicology, or drug/tox or serology, we assumed the specimen was serum or plasma, and when the class was heme/BldCt (blood count) or flow cytometry, we assumed the specimen was blood (whole blood). And when the class was urinalysis, we assumed it was urine sediment. These are the same assumptions that laboratory systems tend to make in order to minimize the test name length for tests in these classes.

Inconstancies between the property implied by the local Units of Measure (UoM) and the property of the LOINC term to which it was mapped

We had already converted all local UoM to UCUM as mentioned above. We used a preexisting large set of mappings between raw local units and properties and UCUM validator converter to find inconsistencies between the properties of the local test UoM and the property of the mapped to LOINC term.

Most UoM were one to one with a LOINC property, but local terms with UoM of percent had many to one mapping with LOINC properties. Units of percent could be used for ratio, and fraction properties, for example. Our algorithm for these could find the correct mapping by identifying terms with the same numerator as the mapped to LOINC term, then finding a term that matches all but the denominator in the mapped to LOINC term. The denominator was never a deciding factor given all of the other parts that matched with parts in the mapped to LOINC term.

Inconsistencies between the method of the local test name and the method in the mapped to LOINC term

Mix-ups between counts per high power field (HPF) and per low powered field (LPF) in urinalysis measurement were the only example of this problem. Local names would carry HPF or LPF in their local name and/or in their UoM, we would compare this unit with the HPF or LPF embedded in the method part of the mapped to LOINC term.

Error correction

In general, error corrections were aligned with one of the four detection process. We assumed that the content of the local test name that disagreed with a LOINC part was the correct content, and we replaced that LOINC part with that content in the mapped to LOINC term to define a hypothetical LOINC term. For example, if the local test name had a unit mg/dL implying a LOINC property of mass concentration (MCnc) and it had been mapped to a term with substance (molar) concentration (SCnc) as property, we replaced SCnc with MCnc in the LOINC term to which it had been mapped and then looked for a LOINC term that matched all of the parts of this hypothetical LOINC. If it found one, it submitted it as the correction and classed the previous mapping as an error. If not, it was also classed the mapping as an error, but could not propose a LOINC term with which to correct the mapping. When the algorithm found more than 1 match, we applied rules to find the best match. The JavaScript code for the detection and correction process can be found in GitHub.¹⁵

For qualifiers and specimens, the algorithm depended on lists of specimen strings and qualifier strings derived from the PCORnet content ([Supplementary Tables S1 and S2](#)). Users who wished to apply our algorithms to their local set of mapped test results would have to create their own sets of specimen names and qualifiers based on specimen, or qualifiers, strings found in the local test names of their local data set. The tables used to detect and correct property errors came from a broad set of independent sources, not the PCORnet data set, so should be applicable to the detection and correction of property errors in any arbitrary set of mapped test results.

Note, however that we could not declare an error when no LOINC term that corresponded to the local test term existed. Mappers tended always to assert a mapping, typically on close to the local term. So, we tallied them separately in [Table 1](#) but did not count such mappings as errors per se.

STATISTICAL ANALYSIS

Our aim was to find the mapping error rates overall and how they varied with other factors such as the class, and the frequency of a given test in the database.

Nonquantitative results were a small fraction, less than 20%, of the tests results in the full sample of 7 billion PCORnet mapped test

Table 2. Among included mapping, unweighted and weighted counts and percent of right and corrected wrongs

N (%*) unique tests	N (%) correctly mapped tests	N (%*) mapped test results	N (%) correctly mapped test results	N (%) of corrected test mappings	N (%) of corrected mapped test results
3029 (65.1)	2766 (91.3)	179 569 051 (94.8)	171 294 723 (95.4)	153 (5.1)	8 092 964 (4.5)

Note: % * : percentage among all tests (instances).

Table 3. Weighted counts and percent of right and fixable wrong by the number of records

N mapped test results	N (%*) unique tests	N (%) correctly mapped tests	N (%*) mapped test results	N (%) correctly mapped test results	N (%) corrected test mappings	N (%) corrected mapped test results
(1) 1–10	780 (25.8)	704 (90.3)	3900 (0.0)	3520 (90.3)	39 (5.0)	190 (4.9)
(2) 10–100	596 (19.7)	538 (90.3)	25 836 (0.0)	23 314 (90.2)	34 (5.7)	1490 (5.8)
(3) 100–1000	749 (24.7)	692 (92.4)	312 424 (0.2)	287 028 (91.9)	26 (3.5)	12 263 (3.9)
(4) 1000–10 000	527 (17.4)	483 (91.7)	1 780 626 (1.0)	1 584 829 (89.0)	30 (5.7)	92 456 (5.2)
(5) 10 000–100 000	198 (6.5)	180 (90.9)	7 690 738 (4.3)	7 152 090 (93.0)	14 (7.1)	402 975 (5.2)
(6) 100 000–1 000 000	138 (4.6)	130 (94.2)	49 904 905 (27.8)	48 109 187 (96.4)	8 (5.8)	1 795 718 (3.6)
(7) 1 000 000–10 000 000	41 (1.4)	39 (95.1)	119 922 622 (66.8)	114 134 755 (95.2)	2 (4.9)	5 787 867 (4.8)
Total	3029	2766 (91.3)	179 569 051	171 294 723 (95.4)	153 (5.1)	8 092 964 (4.5)

Notes: % *: column percentage; %: row percentage.

results (we obtained this full set after we had manually identified errors in the smaller 180 million mapped test results). For quantitative tests, the supplied UoM provided an independent indicator of the property of the LOINC term which greatly narrowed the search space for the correct mapping. Due to privacy concerns about possibility of identifiers lurking in answers, PCORnet did not provide us answer lists for nonquantitative tests, which information would have allowed us to verify the correctness of such tests, so, we decided to focus the study solely on quantitative tests, which were in the large majority and for which the units of measures could provide some independent validating information.

We classified all local tests with no associated UoM and no median value, as nonquantitative. We excluded from analysis all local tests whose name, inexplicably, consisted of all digits, those whose corresponding LOINC terms did not exist and those which had contradictions between their local name and their local UoM, for example, local test name =NOREPINEPHRINE MCG/DAY, local test units (a mass rate)= $\mu\text{g/L}$ (a mass concentration), assuming they were transcription errors by the original mappers.

These PCORnet test results were pulled from a local medical record and subjected to a number of manipulations including, extraction from the medical and aggregated, but these do not account for the some of the weird content such as local test names represented as all digits.

We provided simple counts of included unique tests and mapped test results. We also provided the count and percentage of correctly mapped unique tests and the same information weighted by mapped test results. Further, we present the number and percent of wrong mappings, that our algorithm could correct, before and after weighted by mapped test results. We provided similar statistics for breakdowns by logarithmic ranges, for example: 1–10, 10+–100, 100+–1000, and so on up to 10^7 . We also provide breakdown by 2 DataMarts, by class, property, and specimen and by grouped issues that characterized the mis-mapping.

RESULTS

We refer to each of the 4651 records carrying a unique combination of dimensions as a “test” to simplify the discourse and refer to the 189 405 433 discrete mappings, as mapped test results.

We presented unweighted mis-mapping rates for tests and weighted rates for mapped test results.

We excluded 34.9% of the tests and 5.2% of the mapped test results from our analyses for reasons given in the method section (Table 1). Table 1 also presents the breakdown of the excluded records, including 45 unique tests (or 36 004 mapped test results) for which the LOINC database included no matching terms.

Proportion of correct/corrected mappings by test and by mapping instance

Of included tests, 91.3% were mapped correctly (unweighted) implying an 8.7% incorrect mapping rate. Of mapped test results (weighted), 95.4% were mapped correctly (4.6% incorrectly). Importantly, our correction algorithm fixed 5.1% of the mis-mapped tests (unweighted) and 4.5% of the mapped test results (weighted). The sum of correct and corrected test mappings reached 96.4% and comparable figure for mapped test results reached 99.9% (Table 2). We hasten to add that the correction algorithm was tuned to the PCORnet content for specimen and, qualifiers. So, we should not expect the same success for specimens and qualifiers with different data set. However, our correction algorithm for properties did not depend on the PCORnet data set; so is likely to be as successful with a different data set. Importantly, the LOINC part that accounted for the greatest proportion of mis-mappings was the property.

Proportion of unweighted and weighted correct mappings by logarithmic ranges of mapped test results

Table 3 shows the proportion correctly mapped tests (unweighted) and correctly mapped test results (weighted) by ranges from 1–10

Table 4. Unweighted and weighted counts and percent of right and fixable wrong by DataMart

DataMart	N (%*) unique tests	N (%) correctly mapped tests	N (%*) mapped test results	N (%) correctly mapped test results	N (%) corrected test mappings	N (%) corrected mapped test results
DM102	1825 (60.3)	1710 (93.7)	149 391 626 (83.2)	141 622 427 (94.8)	74 (4.1)	7 707 125 (5.2)
DM62	1204 (39.7)	1056 (87.7)	30 177 425 (16.8)	29 672 296 (98.3)	79 (6.6)	385 839 (1.3)

Notes: % *: column percentage; %: row percentage.

Table 5. Unweighted and weighted counts and percent of right and fixable wrong by LOINC Class (table includes only the 9 classes with more than 90 tests)

Class	N (%*) unique tests	N (%) correctly mapped tests	N (%*) mapped test results	N (%) correctly mapped test results	N (%) of cor- rected test mappings	N (%) of corrected mapped test results
Chemistry	1263 (41.7)	1192 (94.4)	84 894 593 (47.3)	84 539 305 (99.6)	50 (4.0)	348 275 (0.4)
Drugs and toxicology	420 (13.9)	395 (94.0)	727 149 (0.4)	661 483 (91.0)	16 (3.8)	45 439 (6.2)
Microbiology	244 (8.1)	216 (88.5)	352 074 (0.2)	343 220 (97.5)	20 (8.2)	6586 (1.9)
Hematology/blood counts	240 (7.9)	215 (89.6)	80 314 529 (44.7)	74 316 442 (92.5)	14 (5.8)	5 927 369 (7.4)
Allergen testing	223 (7.4)	216 (96.9)	110 508 (0.1)	101 993 (92.3)	2 (0.9)	5811 (5.3)
Coagulation	160 (5.3)	139 (86.9)	2 628 943 (1.5)	2 622 925 (99.8)	4 (2.5)	381 (0.0)
Serology	121 (4.0)	110 (90.9)	185 912 (0.1)	180 856 (97.3)	7 (5.8)	603 (0.3)
Urinalysis	100 (3.3)	72 (72.0)	9 669 455 (5.4)	7 934 630 (82.1)	27 (27.0)	1 734 757 (17.9)
Cell markers (Flow cytometry)	95 (3.1)	77 (81.1)	60 463 (0.0)	51 471 (85.1)	3 (3.2)	3537 (5.8)
Miscellaneous (count < 90)	163 (5.4)	134 (82.2)	625 425 (0.3)	542 398 (86.7)	10 (6.1)	20 206 (3.2)
Grand total	3029	2766 (91.3)	179 569 051	171 294 723 (95.4)	153 (5.5)	8 092 964 (4.5)

Notes: % *: column percentage; %: row percentage.

up to 1 million to 10 million mapped test results. The mapping accuracy was highly skewed toward the more common test and the accuracy for the more common tests (ie, mapped test results $\geq 100\,000$) became almost 100% when mapping errors were corrected (ie, proportion of correct and corrected mapped test results).

Mapping success by data mart

Weighted mapping error rate of DataMart 62 was 3 times better than that of Datamart 102 (1.7% vs 5.2%, Table 4), suggesting important differences between 2 DataMarts in either expertise, or care, applied to the mapping effort. These differences were not explained by differences in class content of the 2 DataMarts in Table 5 (P value $< .001$).

Breakdown by LOINC class

Nine classes carry more than 90 unique tests (Table 5).¹⁶ Chemistry was the dominant class with 47.3% of all mapped test results (weighted), with hematology a close second with 44.7% of the mapped test results. Though microbiology carried more than 100 unique tests, it represents the tiny proportion (0.2%) of all mapped test results, at least as observed within in the PCORnet data set. The mapping accuracy varied importantly by class. The mapping accuracy by mapped test results was above 90% for most classes. For microbiology and hematology/blood counts, it was 97.5% and 92.5%, respectively. By mapped test results, the mapping accuracy of flowcytometry tests (Class CellMark) and urinalysis were poor, 85.1% and 82.1%, respectively. In general, the names of cell markers are poorly standardized in clinical labs and laboratory

records. Their test names rarely indicate the many possible dimensions of a flow cytometry cell marker, such as the initial gating, stem cell versus mature cell, the type of cells blasts, monocytes, etc. So, with a few exceptions, cell markers are very difficult to map and are likely to be unreliable until laboratories apply more formal standardization to their names. The urinalysis testing had similar error rates to cell markers. However, they were easy to detect and correct. All of the urinalysis errors were mix-ups between HPF and LFP.

Breakdown of incorrectly mapped tests and instances by part that was wrong in mapping

Table 6 provides the breakdown of wrong mapping according to the part in the mapped to LOINC name that was wrong. The property part of the LOINC name was responsible for the majority, 50.2%, of LOINC mapping errors. This is good news, because mis-maps due to wrong properties are “easy” to correct and our mapping algorithm is likely to function well for property errors on any mapping data set.

DISCUSSION

Recently published 20%⁸ and 40%⁹ mapping error rates might not be representative of the error rates of a typical local code to LOINC code mapping effort. The 20% error rate was based on a small sample of underpowered survey results. The second paper reported 40% error rate, because they did not exclude special categories, such as, mapping to nonquantitative tests as we did.

Table 6. Unweighted and weighted counts and percent of fixable wrong by part mis-mapped

What LOINC part was wrong problem	N (%*) wrong unique tests	N (%*) wrong mapped test results	N (%) of corrected test mappings	N (%) of corrected mapped test results
Not specified	16 (6.1)	4598 (0.1)	2 (12.5)	446 (9.7)
Invalid LOINC code	15 (5.7)	66 986 (0.8)	15 (100.0)	66 986 (100.0)
Properties	132 (50.2)	2 658 804 (32.1)	103 (78.0)	2 532 043 (95.2)
Specimens	23 (8.7)	4 377 477 (52.9)	8 (34.8)	4 352 848 (99.4)
Timing point vs 24	1 (0.4)	293 (0.0)	0 (0.0)	0 (0.0)
Analyte name	59 (22.4)	72 654 (0.9)	8 (13.6)	47 125 (64.9)
HPF LPF	17 (6.5)	1 093 516 (13.2)	17 (100)	1 093 516 (100)
Total (wrong)	263	8 274 328	153 (58.2)	8 092 964 (97.8)

Notes: % *: column percentage; %: row percentage.

We agree with most of Cholan et al's conclusion about mapping problems and their thoughtful recommendations. But we believe stronger requirements or encouragement from ONC, CLIA, or CMS might be needed to solve the problem of mapping errors. Regardless, automatic detection and correction algorithms, applied as exemplified in this report, could shrink the error rates substantially. Our algorithm certainly performed well on the PCORnet data. Starting with data that had been mapped by humans and fixing their mapping errors is an easier task than automatic mapping without human involvement as other investigators have tried.^{17,18}

One might quibble with our decision to exclude mappings that had no valid mapped to term in the LOINC database. However, they represent less than 1% of the test mappings and its inclusion in the analysis has little effect on our results.

Finally, both papers suggested possible safety concerns assuming that the results would be displayed to provider labeled with a wrong LOINC name. However, as it turns out neither of the 2 major EMRs show any LOINC content in their test displays. Indeed, my researcher friends have a tough time finding LOINC codes within these systems. EMR vendors should store the standard codes with the results and allow users to review them as needed, for example, with click on an icon. Apple Health's personal health record⁴ allows users to see the full FHIR structure and all of the standard codes with one click.

CONCLUSIONS

We found that the overall test mis-mapping rate in the PCORnet data set was 4.6%. This is an important error rate but much less than other published rates of 20%–40%. Importantly, mapping errors can be reduced substantially by using automatic error detection and correction algorithms, as we have shown in this study. We cannot be *certain* such the corrections will be absolutely correct in all cases, however for those based on property mix-ups, the reported UoM provide point strongly to the correct LOINC term. Furthermore, users would not have to accept the corrections as certain but could use them to identify the mappings that require further review.

Clinical users do not see LOINC names, and/or, codes on the result displays of the 2 major EMR systems in the US. Indeed, researcher friends describe challenge in even finding the mapping though they are present in such systems. However, EMR vendors should make the standard codes available to routine administrative and clinical users, through mouse-over or click, so that interested users and administrators could identify inconsistencies between the local name and the LOINC name and correct them to the benefit of

researchers and administrators who may collect such data across multiple system. The mantra of opensource software is, when a thousand eyes can see the source code, they will see all errors. If mappings are always hidden in EMRS, no eyes will see the errors.

FUNDING

This research was supported in part by the Intramural Research Program of the National Library of Medicine, National Institutes of Health.

AUTHOR CONTRIBUTIONS

CM initialed the study and developed the study cohort. CM, SB, AL, and XL conducted literature reviews. ZZ and XL extracted the data. CM, SB, ZZ, and XL designed and conducted data analysis. CM, SB, KM, and LQ interpreted the results. CM, SB, AL, KM, and LQ were involved in study supervision and provided critical revision of the manuscript. All authors participated in manuscript development and are accountable for the integrity of this work.

SUPPLEMENTARY MATERIAL

Supplementary material is available at *Journal of the American Medical Informatics Association* online.

CONFLICT OF INTEREST STATEMENT

None declared.

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

- 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.
- The Office of the National Coordinator for Health Information Technology. United States Core Data for Interoperability (USCDI) – January 2022 – Version 3; 2022.
- Indiana Health Information Exchange – Your Healthcare Records Matter. <https://www.ihie.org/>. Accessed May 26, 2022.

4. HealthKit | Apple Developer Documentation. <https://developer.apple.com/documentation/healthkit>. Accessed May 26, 2022.
5. i2b2: Informatics for Integrating Biology & the Bedside. <https://www.i2b2.org/>. Accessed May 26, 2022.
6. Atlas – LOINC. <https://loinc.org/atlas/>. Accessed May 26, 2022.
7. The Office of the National Coordinator for Health Information Technology. United States Core Data for Interoperability (USCDI) – July 2021 – Version 2.
8. Stram M, Seheult J, Sinard JH, *et al.*; Members of the Informatics Committee, College of American Pathologists. A survey of LOINC code selection practices among participants of the College of American Pathologists Coagulation (CGL) and Cardiac Markers (CRT) Proficiency Testing Programs. *Arch Pathol Lab Med* 2020; 144 (5): 586–96.
9. Cholan RA, Pappas G, Rehwoldt G, *et al.* Encoding laboratory testing data: case studies of the national implementation of HHS requirements and related standards in five laboratories. *J Am Med Inform Assoc* 2022; 29: 1372–80.
10. Home – The National Patient-Centered Clinical Research Network. <https://pcornet.org/>. Accessed May 26, 2022.
11. Smerek M, Priest E, Rosenbloom, *et al.* Assessment of factors and approaches to mapping laboratory results in PCORnet. | Scholars@Duke. In: AMIA annual symposium proceedings. AMIA Symposium. AMIA; 2016; Washington, DC. <https://scholars.duke.edu/display/pub1309384>. Accessed June 17, 2022.
12. Schadow G, McDonald CJ, Suico JG, *et al.* Units of measure in clinical information systems. *J Am Med Inform Assoc* 1999; 6 (2): 151–62.
13. International Union of Pure and Applied Chemistry. <https://iupac.org/>. Accessed May 25, 2022.
14. LOINC Term Basics – LOINC. <https://loinc.org/get-started/loinc-term-basics/>. Accessed July 26, 2022.
15. loinc-mapping-validation-correction/README.md at main lhncbc/loinc-mapping-validation-correction GitHub. <https://github.com/lhncbc/loinc-mapping-validation-correction/blob/main/README.md>. Accessed July 28, 2022.
16. Overton JA, Vita R, Dunn P, *et al.* Reporting and connecting cell type names and gating definitions through ontologies. *BMC Bioinform* 2019; 20 (S5): 259–64.
17. Kelly J, Wang C, Zhang J, *et al.* Automated mapping of real-world oncology laboratory data to LOINC. *AMIA Annu Symp Proc* 2021; 2021: 611–20.
18. Ai D, He Y, Jin S, *et al.* A novel deep learning model for automated mapping of Chinese laboratory test terminologies to logical observation identifiers names and codes (Loinc) [published online ahead of print May 2022]. SSRN. doi:10.2139/SSRN.4092365