



Published in final edited form as:

J Biomed Inform. 2012 August ; 45(4): 667–673. doi:10.1016/j.jbi.2012.01.005.

Enabling international adoption of LOINC through translation

Daniel J. Vreeman, PT, DPT, MSc^{1,2}, Maria Teresa Chiaravalloti³, John Hook¹, and Clement J. McDonald, MD⁴

¹Regenstrief Institute, Inc., Indianapolis, IN, USA, 410 West 10th Street, Suite 2000, Indianapolis, IN USA 46202

²Indiana University School of Medicine, Indianapolis, IN, USA, 410 West 10th Street, Suite 2000, Indianapolis, IN USA 46202

³Università degli Studi della Calabria, Dipartimento di Linguistica, Laboratorio di documentazione, Via Pietro Bucci, cubo 17/b, 87036 Arcavacata di Rende (CS), Italy

⁴Lister Hill National Center for Biomedical Communications, U.S. National Library of Medicine, 8600 Rockville Pike, Bethesda, MD USA 20894

Abstract

Interoperable health information exchange depends on adoption of terminology standards, but international use of such standards can be challenging because of language differences between local concept names and the standard terminology. To address this important barrier, we describe the evolution of an efficient process for constructing translations of LOINC terms names, the foreign language functions in RELMA, and the current state of translations in LOINC. We also present the development of the Italian translation to illustrate how translation is enabling adoption in international contexts. We built a tool that finds the unique list of LOINC Parts that make up a given set of LOINC terms. This list enables translation of smaller pieces like the core component “hepatitis c virus” separately from all the suffixes that could appear with it, such “Ab.IgG”, “DNA”, and “RNA”. We built another tool that generates a translation of a full LOINC name from all of these atomic pieces. As of version 2.36 (June 2011), LOINC terms have been translated into 9 languages from 15 linguistic variants other than its native English. The five largest linguistic variants have all used the Part-based translation mechanism. However, even with efficient tools and processes, translation of standard terminology is a complex undertaking. Two of the prominent linguistic challenges that translators have faced include: the approach to handling acronyms and abbreviations, and the differences in linguistic syntax (e.g. word order) between languages. LOINC’s open and customizable approach has enabled many different groups to create translations that met their needs and matched their resources. Distributing the standard and its many language translations at no cost worldwide accelerates LOINC adoption globally, and is an important enabler of interoperable health information exchange

© 2012 Elsevier Inc. All rights reserved.

Corresponding Author Daniel J. Vreeman, PT, DPT, MSc, dvreeman@regenstrief.org, Telephone: +1 317-423-5515, Regenstrief Institute, Inc, 410 West 10th Street, Suite 2000, Indianapolis, IN USA 46202.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

LOINC; Vocabulary, Controlled; Multilingualism; Translating; Clinical Laboratory Information Systems/standards; Medical Records Systems; Computerized/standards

1. Introduction

Widespread adoption of controlled medical terminologies offers the promise of enabling efficient processing and storage of clinical data that comes from many independent sources. [1] Uniting the many varied sources that produce and store health data is challenging because it requires mapping the myriad idiosyncratic local conventions for representing clinical concepts to a standardized terminology, a process that can be complex and resource intensive.[2-5] Global adoption of terminology standards is often further challenged by language differences between local concept names and the standard terminology.

LOINC® (Logical Observation Identifiers Names and Codes) is a universal code system for identifying laboratory and clinical observations that facilitates exchange and pooling of results for clinical care, research, outcomes management, and many other purposes.[6] When used in conjunction with messaging standards such as HL7, LOINC's universal observation identifiers can be an crucial component for combining test results, measurements, and other observations from many sources.

Since its inception, LOINC has been developed by the Regenstrief Institute as an open standard and made available at no cost worldwide through the LOINC website (<http://loinc.org>). In addition to the LOINC database, Regenstrief also develops and distributes at no cost a software program called RELMA® (the REgenstrief LOINC Mapping Assistant) that facilitates searching the LOINC database, viewing detailed accessory content about the terms, and mapping local terminology to LOINC terms. The most recent LOINC version (2.36, released June 2011) contains 65,003 terms, of which 45,428 are laboratory observation terms and 19,575 are clinical observation terms. Evaluations in several domains have shown that LOINC has good content coverage in areas including clinical laboratory testing[7], radiology reports[8, 9], and clinical note titles[10, 11].

LOINC has been widely adopted, and the user community continues to grow rapidly. The worldwide LOINC community presently has more than 12,800 users in 143 countries, with about 14 new users added daily. Within the USA, LOINC has been adopted by large reference laboratories, health information exchanges, healthcare organizations, insurance companies, research applications, and several national standards. The Department of Health and Human Services has adopted LOINC as the standard across federal agencies for laboratory result names, laboratory test order names, and federally required patient assessment instruments. LOINC is a source vocabulary in the National Library of Medicine's Unified Medical Language System (UMLS) and was adopted by the National Cancer Institute's cancer Biomedical Informatics Grid through a formal review process[12]. LOINC is also commonly used in computerized decision support systems.[13] Recently, LOINC was adopted as the standard for laboratory orders and results as part of the Centers for Medicare and Medicaid Services Electronic Health Record (EHR) "Meaningful Use" incentive program in the Standards and Certification Criteria. Outside the USA, LOINC has been adopted as a national standard in Brazil, Canada, Germany, the Netherlands, Mexico, and Rwanda. Additionally, there are large data exchanges using LOINC in Spain, Singapore, and Korea.

As LOINC's international adoption has grown, so too has the desire to translate LOINC into languages other than its native English. Led by Jack Bierens de Haan, the Centre Suisse de Contrôle de Qualité in Switzerland produced the first set of LOINC term name translations that were released with LOINC version 2.00 in January 2001. Since then, many others have contributed translations. Currently, there are translation efforts underway in 18 countries to translate LOINC into 12 different languages, with translations into 9 languages included in the most recent public LOINC release.

Like Regenstrief, developers of other healthcare terminologies and classifications have also enabled and encouraged foreign language translation. The World Health Organization (WHO) has a policy that welcomes publishing partners in all languages.[14] For example, ICD-10 is available in the six official languages of WHO (Arabic, Chinese, English, French, Russian and Spanish) as well as in 36 other languages. The International Health Terminology Standards Development Organization, which develops and promotes the use of SNOMED CT, has published Guidelines for Translation of SNOMED CT@[15] and Guidelines for the Management of Translations of SNOMED CT.[16] Together these documents provide an comprehensive set of recommendations and best practices for groups producing translations of SNOMED CT. Furthermore, the IHTSDO has published a policy outlining its role as a terminology development organization with respect to the translators and translated content.[17]

Several adjunct techniques for assisting translators of controlled terminologies have been described previously. For example, preliminary work that leveraged the linkages within UMLS between concepts from SNOMED CT and other vocabularies available in French demonstrated that there were many potential matches between terms from different vocabularies. These matches could inform and accelerate the work of human translators, but this pilot work did not include an analysis of their validity. Deléger et al demonstrated detection of good quality new French translations of English terms from MeSH, SNOMED CT, and MedlinePlus Health Topics by using word alignment in parallel text corpora. By its nature, this approach depends on parallel multilingual text that contains the terms of interest. Lastly, established approaches to translation such as Newmark's four-level principle can be adapted for use in translating controlled medical terminologies.[18]

We believe that freely available translations into many languages are an important accelerator of LOINC adoption globally, and have encouraged their development in several ways. One way that Regenstrief has enabled translation of LOINC is through open development and licensing policies. Yet, even with the most advanced automated methods, translation of a controlled terminology is a complex process that still requires extensive human review. So, we have also iteratively developed tools and strategies to help translators in their work. Further, to maximize the potential benefit of the published translations for LOINC users, we also extended the features of RELMA with multilingual searching capabilities. Thus, the purpose of this paper is to describe Regenstrief's approach to foreign language translations of LOINC, including the evolution of an efficient process for constructing translations of LOINC terms names, the current state of translations in LOINC, and the foreign language functions in RELMA. To illustrate how this approach to foreign languages is enabling adoption in international contexts, we also present a case example describing the development of the Italian translation and some of the lessons learned along the way.

2. Material and methods

2.1 Overview of LOINC

LOINC names are constructed according to an established model that “fully specifies” the observation on six main axes (Table 1).[19, 20] We say that the formal LOINC name is fully specified because it contains enough information to distinguish among similar measurements that have different clinical meanings. This does not mean that it carries all possible information about an observation, but simply that it contains enough to uniquely identify it. The Method axis is the only major axis that is optional; it is only used to distinguish among observations that have clinically significant difference in interpretation (e.g. vastly different reference ranges) when made by different methods. When the method *is* specified, it is named at the most general level necessary for distinguishing among tests. The HL7 result message structure provides additional fields for carrying detailed information about how the test was performed, so it is not necessary to carry that information in the test name.

Some of the major name axes also contain minor axes, such as challenge information, adjustments, supersystem (e.g. fetus, blood product unit), and time operators (e.g. maximum, first, etc). The challenge axis has a substructure that delineates the amount, route, and timing. This substructure is used for specifying things like oral glucose tolerance tests. The full details about the LOINC axes can be found in the LOINC Users’ Guide¹⁵, which is the definitive documentation about LOINC naming conventions.

The atomic elements that comprise a fully-specified LOINC name are called LOINC “Parts”. Each LOINC Part is also assigned an identifier (that begins with the prefix “LP”), and internally Regenstrief maintains links between the full LOINC term and the Parts that comprise it. Regenstrief uses LOINC Parts in many aspects of LOINC development, such as: adding synonymy, building hierarchies, creating alternate display names, linking descriptive text, and more. The Parts and their linkages are not distributed as part of the main LOINC table, but rather are content used by the RELMA program. An example of the Parts that make up a serum hepatitis C virus antibody test is given in Table 2.

LOINC is distributed at no cost in regular releases that are available from the LOINC website. The LOINC database contains the LOINC codes, formal names, alternate names and synonyms, and a large number of additional attributes such as a category class (e.g. chemistry, serology, etc), example units of measure, external copyright information (if applicable), audit/control fields, and many more. Regenstrief also makes the RELMA desktop mapping and searching program available free for download from the LOINC website. In June 2010, Regenstrief launched a web-based LOINC search application (<http://search.loinc.org>) that makes the core search features of RELMA available from anywhere online. The LOINC website also contains a wealth of documentation and other resources, including the LOINC Users’ Guide, RELMA Users’ Manual, tutorials, and more.

2.2 Regenstrief’s Approach to LOINC Translations

Consistent with its open development model, Regenstrief has always welcomed requests to translate LOINC content. The LOINC license (<http://loinc.org/terms-of-use>) encourages translation of any of the distributed materials, and articulates that the intellectual property rights of these derivative works remain with Regenstrief (and in some cases also with the LOINC Committee). This policy ensures that the translated material can be made available to the worldwide LOINC community under the same open terms as the source material. Within LOINC, translations are treated as “*linguistic variants*”, which accommodates storage of different dialects of the same language. For example, the Swiss-produced translations into French and Spanish are maintained separately from the French variant from

Canada and the Spanish variant from Spain. Of course, we encourage translators to collaborate across nations, but we recognize that there are important differences in dialects as well as in work timelines, scope, and other factors.

As we have worked with translators over the last 10 years, we have iteratively refined our processes and tools to support their effort. The collection of translated materials is collated and made available from the International page of the LOINC website (<http://loinc.org/international>). We do not prescribe a specific method for performing the translation, leaving that up to the discretion of translator. Regenstrief's approach has been to welcome all comers who are interested in translating LOINC, from dedicated individuals to national research organizations like the Consiglio Nazionale delle Ricerche (CNR) and national standards bodies such as Canada Health Infoway and Deutsche Institut für Medizinische Dokumentation und Information (DIMDI). The LOINC Translation Users' Guide[21] (which has itself been translated into Simplified Chinese and French) outlines the general process for working with Regenstrief to develop a translation and have it incorporated into the main LOINC distribution. One of the most significant milestones in the evolution of our process was the development of a mechanism to automatically generate translations for full LOINC term names based on the translation of the atomic LOINC Parts they contain.

2.2.1 Part-based Translation—The fully specified name is distributed in the LOINC table as six separate database fields. Our earliest translations were created as name strings for the whole LOINC term and stored in a single database field. We later encouraged and received translations that were broken up into the six axes, which made it easier to review for consistency. In January 2007, we began developing a new tool that would find the unique list of LOINC Parts that make up a given set of LOINCs. This tool breaks down the LOINC Parts into pieces smaller even than the six axes, so that you could translate the core component “Hepatitis C virus” separately from all the suffixes that could appear with it, such “Ab.IgG”, “DNA”, and “RNA”. Likewise, the numerators and denominators of ratio measurements (e.g. “Albumin/Creatinine”) are broken apart into separate pieces for independent translation. In addition, we support adding synonyms or alternate names tied to these Parts. We then built another tool that would generate a translation of a full LOINC name from all of these atomic pieces and link up any translated synonyms. These smaller atomic pieces and the linkages between them are not included in the public LOINC distribution, but we have leveraged these features of the internal LOINC development process in support of the translation efforts.

The efficiencies of translating a unique Part list compared to a term by term translation are illustrated in Table 3 and Table 4 using a simple example of a few Hepatitis C and A virus tests. Translating all 9 Parts from the Hepatitis C terms yields full name translations for 6 LOINC terms, but simply by adding the Part translation for “Hepatitis A virus” the full name translations for another 6 LOINC terms can be generated.

This approach can dramatically reduce the overall translation work, because the same piece can appear in many terms. For example, there are more than 670 terms that have “Glucose” as part of the Component. The atomic part only has to be translated only once, and then it can be reused in all the other term names. In total, translating approximately 12,000 individual Parts produces the translated names for about 44,000 LOINC terms. Likewise, by attaching a foreign language synonym to the Part, that synonym can be automatically linked to every term name containing that Part.

We worked closely with Canada Health Infoway who helped pilot test this approach as they developed a French translation. In the Fall of 2007, a translator from Bethune International

Peace Hospital, Shijiazhuang, People's Republic of China also used this fine-grained Part translation method for translating into Simplified Chinese. LOINC version 2.22 (released December 2007) was the first release to contain translations generated semi-automatically from this Part-based approach. Since then, we have encouraged all active translators to give primary consideration to this Part-based approach.

Differences between the linguistic syntax of English and some other languages can potentially limit the effectiveness of this Part-based translation approach. For example, the English phrase *Hepatitis C virus antibodies* could become *des anticorps contre le virus de l'hépatite C* in French. Though we did not know if such patterns would appear in test names as they do in natural language, we designed our tool with the flexibility to change ordering between the Parts and inserting strings such as punctuation or prepositions. The logic mechanism was designed to allow application of these changes at the level of individual terms or groupings of terms, such as all the terms within a particular Class.

2.2.2 Customization—LOINC translators have varied in the number and scope of content that they wish to translate, and Regenstrief has supported local customization in several ways. By storing each language as a “linguistic variant”, we support multiple dialects of the same language. We encourage collaboration among translators and make all the translations available publicly for reference, but also recognize the need for regional customization. Understanding that local needs and resources vary, Regenstrief has not specified a minimum content set that must be translated. One group started by translating only the top 300 most common laboratory order codes; another has translated more than 47,000 laboratory and clinical terms. Regenstrief works with each group to support their defined subset, which often evolves over time with new LOINC releases, expanded project scope, etc. Our Part-based translation mechanism enables foreign language synonymy at the atomic Part level. Furthermore, we also support translation of the LOINC Short Name (an alternate term name designed for optimal compactness) and the LOINC Long Common Name (an alternate term name designed for clinician-friendly viewing) as whole strings linked to the LOINC term. These translated alternate names offer great potential for tailoring the labels linked to a term. Lastly, translators also vary in the frequency with which they update their work. Some diligently update with each new LOINC release while others have longer stretches in between. To help synchronize with the latest content, Regenstrief can produce “delta” files showing what has changed or is new since their last translation.

2.2.3 Ongoing Maintenance and Updates—With each new LOINC release we now generate a file containing the unique Parts that make up three potentially useful translation starter sets: the Top 300 Common LOINC Order terms and the Top 2000+ Common LOINC Result terms (both described at <http://loinc.org/usage>), and all laboratory LOINC terms. For those with existing translations, we also generate a listing of all of the new or changed Parts from terms since the last release. For translators who have identified specific subsets of terms they want to support, we can also create a file showing changed Parts for terms they have translated. Because LOINC follows good terminology development practices[22], many of the edits made to existing terms are just typographical (correcting typos, editing for consistency across terms, etc) and would not necessarily require a corresponding change in translation. We also produce a version of the LOINC Users' Guide showing changes since the last edition, to highlight for translators the content that may need updating.

2.3 Functions for Viewing and Searching Linguistic Variants in RELMA

Since LOINC first distributed linguistic variants, the RELMA mapping program has had features to display translations for LOINC term names. Over time, we have continued to refine and expand its capabilities. As we developed the Part-based translation mechanism we

also built functions in RELMA that enable the core search functions to work with input text in the translated languages. After discussion with the LOINC community, we decided to enable this capability for the translations that were sufficiently large and left the translations with only a few thousand terms in ‘view only’ mode. We did this so as to not give users the impression that LOINC was woefully lacking in content when the reason for returning few search results was that a limited set was translated.

RELMA version 3.25 (January 2009) was the first release that contained searchable translations. Our initial implementation relied on word indexes stored in the underlying RELMA database, as we did with the English index. Because of the size of these additional indexes it became prohibitive to distribute them all together as one package. To overcome this limitation, we built separate distributions of the RELMA program with the English index plus one additional searchable language and made them all available for download from the LOINC website. Each distribution could display all of the translations; only the foreign language search capability varied.

We later began migrating all of the search capabilities in RELMA to the widely used open source Apache Lucene (<http://lucene.apache.org>) search engine. Lucene offered a wealth of feature enhancements and speed improvements, but two important benefits of this technology shift were its multilingual search support and that it could build the word indexes for different languages on-demand. The on-demand index building made it possible to ship a compact single package that would create the additional language indexes when selected by the user. The RELMA 5.0 release (December 2010) was the first one that employed Lucene for all searching, and we were able to distribute it as a single build for all languages.

2.4 Accelerating LOINC Adoption in Italy Through Translation: A case example

To illustrate how LOINC’s approach to foreign languages is accelerating adoption in international contexts, we describe here the adoption of LOINC into the emerging Italian electronic health infrastructure. In spite of governmental regulations and recommended standards, current practice in Italian laboratories is very much like the Tower of Babel, because each laboratory uses its own local names and codes for observation identifiers. Italian laboratories have mapped their local terms to their specific region’s reimbursement codes, but these are not based on clinical utility or a structured domain classification. To achieve the goal of interoperable exchange with and across regions, and also internationally, the Consiglio Nazionale delle Ricerche (CNR) has launched a nationwide project named InFSE (FSE is the Italian acronym for EHR). The InFSE project aims to: a) translate key portions of LOINC into Italian, b) assist laboratorians in mapping their local terms to LOINC, and c) facilitate use of LOINC in the emerging Italian health infrastructure.

2.4.1 Approach to the Italian Translation—The overall approach we took in translating LOINC followed the Part-based method described in this paper and the LOINC Translation Users’ Guide.[21] We recognize that translating a terminology is not merely a matter of transposing words from one language into another, but rather that it must adapt to the context of use. Thus, before we began the actual translation we completed two preparatory activities: analyzing current laboratory test naming practice in Italy and conducting a comprehensive review of the other LOINC linguistic variants.

Our analysis of Italian laboratory practice was made possible by the formation of a ‘LOINC Italia’ group, which involved several healthcare organizations interested in LOINC adoption and was led by Molinette Hospital and School of Medicine in Turin. Within this group we gathered local test codes and names from several laboratories and evaluated the patterns and naming conventions among them. The laboratories structured their local tests lists according

to the six axes of the LOINC fully specified name, so that we could evaluate their similarities across these dimensions and inform our translations.

We also found that reviewing the translation approaches taken by the French and the Spanish LOINC linguistic variants was particularly helpful because these languages are grammatical similar to Italian. We took a similar approach to that of the Société Française d'Informatique de Laboratoire (SFIL)[23] in that we are using LOINC as a data exchange vocabulary, and are not intending to replace the test names people are using in local systems with our Italian LOINC names.

As we iteratively built and reviewed our LOINC translation, we established naming rules that were guided by the desire to balance having observations names as close as possible to the current naming conventions used by laboratorians in everyday work but yet remaining faithful to the semantic representation of the LOINC term. Our primary purpose was not, in fact, to realize an aesthetically perfect translation, but rather to ensure the best chance of finding the right LOINC code in mapping. Our review of local naming conventions revealed tremendous variability, even for the same test. Thus, in order to bridge local naming variation to the LOINC names, we also built rich synonymy that would enable the mappers to find the proper target LOINC code using the search words they were familiar with.

3. Results

3.1 Present State of LOINC Translations

As of version 2.36 (June 2011), LOINC terms have been translated into 9 languages from 15 linguistic variants other than its native English. Table 5 gives the details for each of the currently available linguistic variants. The five largest linguistic variants have all used the Part-based translation mechanism. Since January 2009 when we first distributed RELMA with searchable translations, the RELMA program has been downloaded more than 17,400 times (about 580 times per month) by users worldwide.

In addition to term translations, LOINC users have produced translations of other reference material. The LOINC Users' Guide has been translated into Estonian, German, Simplified Chinese, and Spanish. Peking University Medical Press published a "Book of LOINC"[24] that included Simplified Chinese translations of the LOINC Users' Guide, RELMA Users' Manual, and all of the terms from LOINC version 2.19. The Laboratory LOINC and RELMA Tutorial (a slide presentation) has been translated into Estonian and Simplified Chinese.

There are also several current translation efforts in process that have not yet completed a version for distribution. LOINC translators are working in the Netherlands to translate terms into Dutch, as a joint effort between translators in Russia and the Ukraine to translate terms into Russian, and there are plans in Spain to translate terms into Catalan.

3.2 Case Example: Teaching LOINC to Speak Italian

Starting in the Fall of 2010, we did a rapid initial translation of 11,748 LOINC Parts to evaluate if the resulting full names could work in the Italian context. These Parts generated fully translated names for 43,152 LOINC codes that were first distributed by Regenstrief in the LOINC 2.34 and RELMA 5.0 release in December 2010. In creating our Italian LOINC translation, the main issue we encountered using Regenstrief's Part-based approach was reconciling the syntactical and semantic rightness of the strings it produced. These problems arose because of the deep difference between English and Italian grammatical rules. An additional challenge we encountered was the use of acronyms and abbreviations. Some of

the issues we identified related to many terms, but others were specific to a single term. We addressed these issues by conducting several cycles of review and refinement.

After producing and reviewing the initial translation, we recognized that we needed to define more precise translational rules that could comply with both Regenstrief's naming conventions and Italian grammatical rules. Over the next several months, we collaborated closely with the LOINC development team at Regenstrief Institute to better understand LOINC's naming conventions, establish sound translational rules, find linguistic solutions to issues we encountered, and to update the translated LOINC Parts based on this more comprehensive strategy.

As a result of this work, we produced a new version of Italian linguistic variant that was distributed in the LOINC 2.36 and RELMA 5.3 release (June 2011) and contained translations for 43,989 LOINC terms based on revised translations for 11,748 unique atomic Parts. The 837 additional term translations came not from translating additional Parts, but from the routine expansion of LOINC content (based on user requests) that remixed those previously translated Parts into new terms. We also focused on building up the synonymy for Parts; we added an additional 2,918 since the December 2010 release. Many of these new synonyms were typical Italian phrasings (with prepositions, etc) that we gleaned from our hospital test lists but that differed from the conventions we used in the formal translated name of the term.

3.2.1 Using the Italian LOINC Translation—We have incorporated the Italian LOINC translation into a web-based tool that we are using to help five regions in Italy map their laboratory tests to LOINC and the region-specific reimbursement codes. The mapping tool stores the translation and details of each LOINC term, provides term search capabilities, and also shows the mapping choices already made by others in the community of laboratories. In the future, we hope to supplement the view of peer-based mappings with empiric analysis of test ordering and result frequency that may help improve searching and mapping. We are also translating the LOINC® Users' Guide into Italian as a resource for laboratorians in Italy.

4. Discussion

Regenstrief has developed an efficient process and set of tools that have been used successfully to translate LOINC into 9 languages from 15 linguistic variants other than its native English. The open and customizable approach has enabled many different groups to create translations that met their needs and matched their resources. We have developed fast and efficient tools in RELMA that enable multi-lingual searching and mapping to LOINC. Distributing the standard and its translations into many languages at no cost worldwide accelerates LOINC adoption globally, and is an important enabler of interoperable health information exchange.

Even with efficient tools and processes, translation of standard terminology is a complex undertaking. Two of the prominent linguistic challenges that translators have faced include: a) the approach to handling acronyms and abbreviations, and b) the differences in linguistic syntax (e.g. word order) between languages.

Acronyms and abbreviations present special challenges in a multilingual context because they are often based on expression in the source language. LOINC naming conventions generally prohibit use of abbreviations and acronyms in the Component, but there are a few enumerated exceptions allowed and some of the other LOINC axes use them extensively

(e.g. Property). The Consiglio Nazionale delle Ricerche resolved this issue with a decision tree:

- Use the Italian equivalent acronym or abbreviation, if available. E.g. CSF (Cerebro Spinal Fluid) was translated into LCS (*Liquido Cerebro Spinale*);
- Use the full form, if an equivalent acronym doesn't exist in Italian. E.g. BPU (Blood Product Unit) was translated into *Unità di prodotto sangue*;
- Keep the English acronym or abbreviation if it is used and recognized in Italy too. E.g. DNA, HIV, GnRH, Ab;
- Use LOINC-specific acronyms and abbreviations in four of the main axes (Property, Time, System, Scale), but provide explanations of what they mean and how they should be used

This last point of using LOINC-specific acronyms (e.g. SCnc for substance concentration) was discussed at length. CNR's first inclination was to fully expand these acronyms in their translation, but they later concluded that some of these Parts were so integral to the LOINC system that they would not have made sense if extrapolated from that context. In this way, it would be easier for an Italian laboratorian to go through the LOINC conventions than to try making sense of some "abstractly" translated names.

Following the same principle, the CNR also maintained Latin names where LOINC used them, because they are universally understood. Italian translations for the Latin words *pre* and *post* (that are used in challenge tests) were provided as synonyms to assist in searching.

Differences in linguistic syntax are a more prevalent issue. Although we built functionality in the tool that generates full name translations from the Parts to change the order between the Parts and insert strings such as punctuation or prepositions, in practice, we have not had to make much use of it. Translators have largely opted to keep the LOINC-like word order because it facilitates sorting by the core component, e.g. Hepatitis C virus. Even in a language like German that has more inflection, derivation, and compounding than English, translators from DIMDI decided that the term name translations did not necessarily need to exist in natural German language in order to be clear and perfectly understandable by Germans. CNR developed a translation rule called "words order exchange" that always lists the identifier of the substance being measured first, as is consistent with the usual LOINC naming conventions. As an example, the regular Italian translation for *Mouse epithelium* would be *Epitelio di topo*. In order to place *mouse* first, the regular order of the words was switched and a comma inserted between them so as to provide a meaningful representation in Italian according to the Italian grammatical and syntactical rules. So the translation becomes *Topo, epitelio*. The more natural Italian phrasing was stored as a synonym, so that users can find the test using prepositions too.

Any large translation effort will uncover specific peculiarities between the source and target language. In some cases, the repository of existing translations can be informative. For example, in considering a *challenge test*, translators from CNR debated about the Italian representation because there is not a unique translation, even if the meaning is clear. Both *administration* and *stimulation* were considered, but after reviewing the French and Spanish translations (*test de provocation* and *test de provocaciòn*, respectively), *test di provocazione* emerged as the best solution in Italian. Another interesting example from the Italian experience was *dander*, because there is not an Italian word that covers its exact semantic area. After review and discussion *cat dander* was translated to *forfora di gatto*, even though the usual translation of *forfora* is *dandruff*.

There are many potential opportunities for continuing to refine the LOINC translation tools and processes. We have already begun prototype work that allows full internationalization of the online LOINC search application (<http://search.loinc.org>), including translations for all of the program labels as well as capabilities for multilingual searching. We are currently pilot testing a full Italian translation of the application. Although the complexity of the RELMA program is much greater than the online search program, we hope to fully internationalize the RELMA program in the future as well. These features could further lower the language barrier to mapping local test names to LOINC. With additional mapping work, the promising approach described by Joubert et al[25] for finding potential translations from linkages to other source terminologies in the UMLS could be incorporated into our process of generating a unique Part list. An automatically generated “starter” translation of high quality might dramatically accelerate the work of human translators. Another resource that could potential benefit translators are the hierarchies of LOINC Parts that Regenstrief creates for each axis of the formal name. These hierarchies are used in RELMA for restricting searches, but they could also assist in the translation process. For example, organizing the translations for “Estradiol.bioavailable”, “Estradiol.albumin bound”, and “Estradiol.unconjugated” under the node of “Estradiol” might help detect errors and inconsistencies.

5. Conclusions

Regenstrief has developed an efficient process and set of tools that have been used successfully to translate LOINC into 9 languages from 15 linguistic variants other than its native English. The open and customizable approach has enabled many different groups to create translations that met their needs and matched their resources. We plan to continue refining our tools and processes to further accelerate and assist the work of translators. Distributing a multi-lingual vocabulary standard at no cost worldwide accelerates LOINC adoption globally, and is an important enabler of interoperable health information exchange.

Acknowledgments

The authors thank Mark Fisher for technical assistance; Lorie Carey, Beverly Knight, and Lin Zhang for providing feedback on the LOINC translation process, Giuseppe Alfredo Cavarretta for his support of the Italian LOINC effort, and all of the other individuals who have contributed to a LOINC translation. This work was performed at the Regenstrief Institute, Indianapolis, IN, USA and Laboratorio di Documentazione, Dipartimento di Linguistica dell'Università della Calabria, Italy, in partnership with Azienda Ospedaliero-Universitaria Molinette, Torino, Italy. It was supported by Consiglio Nazionale delle Ricerche (CNR), Information and Communication Technology Department, Italy and by contract HHSN276200800006C from the National Library of Medicine.

References

1. McDonald CJ. The barriers to electronic medical record systems and how to overcome them, *Journal of the American Medical Informatics Association*. JAMIA. 1997; 4:213–221. [PubMed: 9147340]
2. Baorto DM, Cimino JJ, Parvin CA, Kahn MG. Combining laboratory data sets from multiple institutions using the logical observation identifier names and codes (LOINC). *International Journal of Medical Informatics*. 1998; 51:29–37. [PubMed: 9749897]
3. Lau LM, Banning PD, Monson K, Knight E, Wilson PS, Shakib SC. Mapping Department of Defense laboratory results to Logical Observation Identifiers Names and Codes (LOINC), AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium. 2005:430–434. [PubMed: 16779076]
4. Lin MC, Vreeman DJ, McDonald CJ, Huff SM. A characterization of local LOINC mapping for laboratory tests in three large institutions. *Methods of information in medicine*. 2011; 50:105–114. [PubMed: 20725694]

5. Vreeman DJ, Stark M, Tomaszewski GL, Phillips DR, Dexter PR. Embracing change in a health information exchange, AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium. 2008:768–772. [PubMed: 18999242]
6. McDonald CJ, Huff SM, Suico JG, Hill G, Leavelle D, Aller R, Forrey A, Mercer K, DeMoor G, Hook J, Williams W, Case J, Maloney P. LOINC, a universal standard for identifying laboratory observations: a 5-year update. *Clin Chem*. 2003; 49:624–633. [PubMed: 12651816]
7. Vreeman DJ, Finnell JT, Overhage JM. A rationale for parsimonious laboratory term mapping by frequency, AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium. 2007:771–775. [PubMed: 18693941]
8. Vreeman DJ, McDonald CJ. Automated mapping of local radiology terms to LOINC, AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium. 2005:769–773. [PubMed: 16779144]
9. Vreeman DJ, McDonald CJ. A comparison of Intelligent Mapper and document similarity scores for mapping local radiology terms to LOINC, AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium. 2006:809–813. [PubMed: 17238453]
10. Hyun S, Shapiro JS, Melton G, Schlegel C, Stetson PD, Johnson SB, Bakken S. Iterative evaluation of the Health Level 7--Logical Observation Identifiers Names and Codes Clinical Document Ontology for representing clinical document names: a case report, *Journal of the American Medical Informatics Association*. JAMIA. 2009; 16:395–399. [PubMed: 19261945]
11. Dugas M, Thun S, Frankewitsch T, Heitmann KU. LOINC codes for hospital information systems documents: a case study, *Journal of the American Medical Informatics Association*. JAMIA. 2009; 16:400–403. [PubMed: 19261942]
12. Cimino JJ, Hayamizu TF, Bodenreider O, Davis B, Stafford GA, Ringwald M. The caBIG terminology review process. *Journal of Biomedical Informatics*. 2009; 42:571–580. [PubMed: 19154797]
13. Ahmadian L, van Engen-Verheul M, Bakhshi-Raiez F, Peek N, Cornet R, de Keizer NF. The role of standardized data and terminological systems in computerized clinical decision support systems: literature review and survey. *International Journal of Medical Informatics*. 2011; 80:81–93. [PubMed: 21168360]
14. World Health Organization, Publishing translations of WHO information materials.
15. International Health Terminology Standards Development Organisation, Guidelines for Translation of SNOMED CT®. 2010.
http://www.ihtsdo.org/fileadmin/user_upload/Docs_01/About_IHTSDO/Publications/IHTSDO_Translation_Guidelines_v2.00_20100407.pdf
16. International Health Terminology Standards Development Organisation, Guidelines for Management of Translation of SNOMED CT®. 2010.
http://www.ihtsdo.org/fileadmin/user_upload/Docs_01/About_IHTSDO/Publications/IHTSDO_Translation_Guidelines_v2.00_20100407.pdf
17. International Health Terminology Standards Development Organisation, Policy on IHTSDO's Role in Translation. 2010.
http://www.ihtsdo.org/fileadmin/user_upload/Docs_01/About_IHTSDO/Governance_and_Advisory/Policies/IHTSDO_Role_in_Translation_May_2010.pdf
18. Reynoso GA, March AD, Berra CM, Strobietto RP, Barani M, Iubatti M, Chiaradio MP, Serebrisky D, Kahn A, Vaccarezza OA, Leguiza JL, Ceitlin M, Luna DA, Bernaldo de Quiros FG, Otegui MI, Puga MC, Vallejos M. Development of the Spanish version of the Systematized Nomenclature of Medicine: methodology and main issues, *Proceedings / AMIA ... Annual Symposium*. AMIA Symposium. 2000:694–698.
19. Forrey AW, McDonald CJ, DeMoor G, Huff SM, Leavelle D, Leland D, Fiers T, Charles L, Griffin B, Stalling F, Tullis A, Hutchins K, Baenziger J. 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. [PubMed: 8565239]
20. McDonald, CJ.; Huff, SM.; Mercer, K.; Hernandez, J.; Vreeman, DJ. Logical Observation Identifiers Names and Codes (LOINC®) Users' Guide. 2011. <http://loinc.org/downloads>

21. Logical Observation Identifiers Names and Codes (LOINC®). Translation Users' Guide, Regenstrief Institute. 2009
22. Cimino JJ. Desiderata for controlled medical vocabularies in the twenty-first century. *Methods of information in medicine*. 1998; 37:394–403. [PubMed: 9865037]
23. Domas G, Marchand M. Priorité au contenu, *Spectra Biologie*. 2007; 158:3.
24. McDonald, CJ.; Huff, SM. Logical Observation Identifiers Names and Codes v2.19. U.S.: Peking University Medical Press, B; 2008.
25. Joubert M, Abdoune H, Merabti T, Darmoni S, Fieschi M. Assisting the translation of SNOMED CT into French using UMLS and four representative French-language terminologies, AMIA ... Annual Symposium proceedings / AMIA Symposium. *AMIA Symposium*. 2009; 2009:291–295. [PubMed: 20351867]

Highlights

- We developed an open, efficient, and customizable approach for translating LOINC.
- LOINC has been translated into 9 languages from 15 linguistic variants.
- Distributing LOINC as a no cost, multi-lingual standard accelerates global adoption.

Table 1

Formal model for constructing LOINC fully specified names

Axis Name	Description/Example
Component	The analyte or attribute being measured or observed. E.g., sodium, body weight.
(Kind of) Property	Differentiates kinds of quantities relating to the same substance. E.g. mass concentration, catalytic activity.
Time (Aspect)	Identifies whether the measurement is made at a point in time or a time interval. E.g. 24H for a urine sodium concentration.
System	The specimen, body system, patient, or other object of the observation. E.g. cerebral spinal fluid, urine, radial artery.
(Type of) Scale	The scale or precision that differentiates among observations that are quantitative, ordinal (ranked choices), nominal (unranked choices), or narrative text.
(Type of) Method	An optional axis that identifies the way the observation was produced. It is used only when needed to distinguish observations that have clinically significant differences in interpretation if made by different methods.

Table 2

LOINC Parts from term 5198-7, Hepatitis C virus Ab [Units/volume] in Serum by Immunoassay

Part Type	Part Number	Part Name
Component (full)	LP38332-0	Hepatitis C virus Ab
Component (core)	LP14400-3	Hepatitis C virus
Suffix	LP20667-9	Ab
Property	LP6773-8	ACnc [Arbitrary Concentration]
Time	LP6960-1	Pt [Point in time (spot)]
System	LP7567-3	Ser [Serum]
Scale	LP7753-9	Qn
Method	LP6241-6	EIA [Immunoassay (EIA)]

Table 3

Unique list of LOINC Parts necessary to translate terms 5198-7, 22327-1, 5198-7, 16936-7, 57006-9, 53376-0, 51824-1, 5183-9, 22312-3, 5179-7, 22313-1, 5181-3, and 22315-6.

Part Number	English Part Name	Italian Part Name
LP14400-3	Hepatitis C virus	Epatite C, virus
LP16708-7	Hepatitis A virus	Epatite A, virus
LP20667-9	Ab	Ab
LP32401-9	Ab.IgG	Ab.IgG
LP32403-5	Ab.IgM	Ab.IgM
LP6773-8	ACnc	ACnc
LP6960-1	Pt	Pt
LP7567-3	Ser	Siero
LP7753-9	Qn	Qn
LP6241-6	EIA	EIA

Table 4
Full Italian translated names of 12 LOINC terms based on translations of the 10 LOINC Parts shown in Table 3.

LOINC	Component	Property	Timing	System	Scale	Method
22327-1	Epatite C, virus	ACnc	Pt	Siero	Qn	
5198-7	Epatite C, virus, Ab	ACnc	Pt	Siero	Qn	EIA
16936-7	Epatite C, virus, Ab,IgG	ACnc	Pt	Siero	Qn	
57006-9	Epatite C, virus, Ab,IgG	ACnc	Pt	Siero	Qn	EIA
53376-0	Epatite C, virus, Ab,IgM	ACnc	Pt	Siero	Qn	
51824-1	Epatite C, virus, Ab,IgM	ACnc	Pt	Siero	Qn	EIA
22312-3	Epatite A, virus, Ab	ACnc	Pt	Siero	Qn	
5183-9	Epatite A, virus, Ab	ACnc	Pt	Siero	Qn	EIA
22313-1	Epatite A, virus, Ab,IgG	ACnc	Pt	Siero	Qn	
5179-7	Epatite A, virus, Ab,IgG	ACnc	Pt	Siero	Qn	EIA
22315-6	Epatite A, virus, Ab,IgM	ACnc	Pt	Siero	Qn	
5181-3	Epatite A, virus, Ab,IgM	ACnc	Pt	Siero	Qn	EIA

Table 5

LOINC translations available in RELMA 5.3 (June 2011)

Linguistic Variant	Producer ^d	Searchable in RELMA ^b	Terms Translated	Parts Translated ^c	Synonyms Translated
Chinese (CHINA)	Bethune International Peace Hospital	Yes	48,422	20,417	90,546
Estonian (ESTONIA)	Estonian E-Health Foundation	Yes	44,005	11,992	1,036
French (CANADA)	Canada Health Infoway Inc.	Yes	39,551	8,903	10
French (FRANCE)	Société Française d'Informatique de Laboratoires		4,408		
French (SWITZERLAND)	CUMUL, Switzerland		4,940		
German (GERMANY)	Institute for Medical Documentation and Information (DIMDI)	Yes	11,059		
German (SWITZERLAND)	CUMUL, Switzerland		4,941		
Greek (GREECE)	Efsratia Kontaxi, MD, MSc; Evripidis Stefanidis, MD; and Panagiotis Kontaxis		2,094	307	38
Italian (ITALY)	Consiglio Nazionale delle Ricerche	Yes	43,989	11,748	6,268
Italian (SWITZERLAND)	CUMUL, Switzerland		4,939		
Korean (KOREA, REPUBLIC OF)	Korean Ministry for Health, Welfare, and Family Affairs	Yes	26,893		
Portuguese (BRAZIL) -- Draft	Brazilian Federal Agency for Health Plans and Insurance; Brazilian Clinical Analysis Society; Brazilian Clinical Pathology Society and Diagnóstico das Américas (DASA)		2,780		
Spanish (ARGENTINA)	Conceptum Medical Terminology Center	Yes	38,308		
Spanish (SPAIN)	Clinical Laboratory Committee of Servicio Extremoño de Salud, with the support of BITAC MAP	Yes	44,352	12,071	10,283
Spanish (SWITZERLAND)	CUMUL, Switzerland		4,940		

^a Additional details are available at <http://loinc.org/international>.^b Per Regenstrief policy, only translations of sufficient size (e.g. > 10,000 terms) are enabled for searching in RELMA.^c Translations generated from our Parts-based method have values in this column. Translations developed at the whole term level have null entries in this column.