Digital Personalized Health and Medicine
L.B. Pape-Haugaard et al. (Eds.)

doi:10.3233/SHTI200198

The LOINC Content Model and Its Limitations of Usage in the Laboratory Domain

Cora DRENKHAHN[a,1] and Josef INGENERF[a,b]
[a] *Institute of Medical Informatics (IMI), University of Lübeck, Germany,*
[b] *IT Center for Clinical Research, Lübeck (ITCR-L), University of Lübeck, Germany*

Abstract. To unambiguously encode the semantic meaning of laboratory tests, the LOINC terminology is widely used. With regard to the constantly changing and diverse requirements of the laboratory domain, LOINC's long-established content model and related publications are reviewed conjointly, revealing some obstacles for flexible adaptation in terms of new or varying application needs as well as issues regarding the comprehensive reusability of lab data. In a concise overview, four specific limitations are identified that require adaptation or the usage of other terminologies.

Keywords. Terminologies, Semantic Interoperability, Clinical Laboratory

1. Introduction

Logical Observation Identifiers Names and Codes (LOINC) is one of the most notable and well-known terminologies in medicine and widely used for standardized coding of laboratory data in electronic health records [1]. LOINC's content model has remained unchanged since its beginnings [2] but is confronted with varying needs and requirements in modern healthcare. Here, the significance of laboratory diagnostics entails a great diversity of tasks in digitalized lab environments including order/entry of tests and reporting of their results. Distributed laboratory data often needs to be integrated as well, for example to compare a patient's test results for clinical monitoring or to aggregate and analyze data on a larger scale in a data warehouse. In all of these aspects, economic considerations are highly relevant as well, with the billing of lab procedures being part of everyday clinical practice [3]. Like all subdomains of medicine, the laboratory field is subject to constant change, but here the rapid emergence of new diagnostic tests and methods, for example in molecular genetic analysis, speeds up this process in particular [4].

In this paper, LOINC's properties are analyzed regarding the current challenges in the laboratory domain to identify present limitations of its appliance. Related recent research is incorporated to bring together individual insights for a thorough and accurate overview. This evaluation shall not be misunderstood as criticism of LOINC – its usefulness in coding lab tests is undeniable; instead we identify potential for

1 Cora Drenkhahn, Institute of Medical Informatics, University of Lübeck, Ratzeburger Allee 160, 23538 Lübeck, Germany; E-mail: c.drenkhahn@uni-luebeck.de.

improvement and cases in which another coding system such as the comprehensive ontology SNOMED CT is needed to fully cover the complex demands of laboratory medicine.

2. Material and Methods

Based on the authors' extensive knowledge gained in previous projects, the LOINC content model and the features of its current version 2.66, as well as related literature, are critically reviewed in the context of requirements in modern laboratory diagnostics.

2.1. LOINC content model

Since its inception, LOINC is published biannually by the Regenstrief Institute as a database of unique codes associated with human-readable terms. Each code corresponds to a distinct lab test (currently 55,844 out of 91,388 codes) or another clinical observation explicitly defined by five mandatory main dimensions plus an optional method part. Therein, LOINC's referenced properties are restricted to those required for a correct test result interpretation so that each combination represents the minimal set of values needed to distinguish this test from all others. For most axes, some further information can be included in specific subcomponents if needed for a full definition.

Component describes the kind of substance or analyte measured stating the test's most basic information with an immense variety of possible values (currently 24,707 part codes) and complexity of contents (up to three subcomponents).

Property denotes in which characteristic or type of property the component was measured, e.g. as number, as substance concentration or the qualitative presence. Obviously, this value is highly correlated with the test result's unit which is *not specified* in the LOINC term [5]. With 135 different options, the property axis is known to cause misunderstandings and therefore issues while mapping to LOINC [6].

Time is used to differentiate measurements at a single point of time from those over a (specified) interval of time with an optional subcomponent.

System defines the specimen or material sample used for testing as the second most distinct characteristic. Further refinement is typically needed by other properties (e.g. 32 LOINC codes for "creatinine" in "urine") or a subcomponent.

Scale broadly differentiates between quantitative, ordinal and nominal measurements, forming an interrelation to the *Property* part. Answer options for qualitative tests shall be coded with SNOMED CT; for quantitative results exemplary value ranges and units are partially mentioned but not explicitly defined.

Method information is only included if the measurement technique is clinically significant by affecting the test's result or reference range. So, most terms (63.32%) lack this dimension.

Table 1. Some exemplary LOINC codes with their defining characteristics of six dimensions. Subcomponents are divided by the ^ sign.

LOINC code	Component	Property	Time	System	Scale	Method
14682-9	Creatinine	SCnc	Pt	Ser/Plas	Qn	
14684-5	Creatinine	SRat	24H	Urine	Qn	
5802-1	Nitrite	PrThr	Pt	Urine	Ord	Test strip

14754-6	Glucose^1H post 50 g glucose PO	SCnc	Pt	Ser/Plas	Qn
14578-9	AB0 group	Type	Pt	Bld^BPU	Nom

Structure and content of terms and subcomponents as shown in Table 1 reveal LOINC's origin to work conjunctively with the syntactic interoperability standard HL7 v2. Omitted information like result value, unit, reference range or context are meant to be communicated in other fields of the HL7 v2 message or the (newer) HL7 FHIR Observation resource. Another design principle states to only add sensible combinations required in practice [2,7], limiting availability to existing pre-coordinated terms.

2.2. Related research

In a previous project, we investigated the mapping of laboratory services contained in an internal catalog of a German hospital to a standardized terminology. Internal catalogs capture the local range of services and are thus used for billing purposes. This leads to a primary focus on methods and environmental information, whereas measurement parameters like specimen or property are seldom mentioned. Because of this, we found it impossible to map the heterogenous contents to LOINC due to the small overlap of included information and the prerequisite to use complete pre-coordinated terms. As a solution, SNOMED CT could be utilized based on its concepts of diverse granularity and ability for post-coordination [8].

In a literature review, we found 929 publications since 2017 mentioning the terms 'LOINC', 'laboratory' plus 'limitations' or 'challenges' and further investigated 25 of these classified as most relevant by Google Scholar. Two articles covered the most relevant aspects for our review. Bietenbeck et al. assessed three terminologies including LOINC and SNOMED CT for their ability to encode laboratory results of different complexity and correlated evaluation parameters. Whereas LOINC was found to be easily usable, interpretive comments like measurement uncertainty or reference intervals could not be expressed adequately. With SNOMED CT more content could be encoded although full coverage remained elusive. Another issue concerning LOINC was detected in the non-existent formal definitions – and explicit hierarchies in particular – resulting in unclear (subclass) relations between terms and limited possibilities of computational analysis [9]. Stram et al. conducted a comprehensive evaluation of current challenges in LOINC usage in pathology laboratories. Among other results, an increased risk for mismatches between laboratories due to inconsistent mapping, especially of method properties, was revealed. Additionally, molecular genetic diagnostics were identified to pose serious problems on the current content model regarding the vast quantities of singular tests with heterogeneous features involved, e.g. for 22,000 genes. Rapidly evolving subdomains like this were also mentioned to be challenging for keeping LOINC up to date [10].

3. Results

Putting LOINC content model characteristics into the context of current challenges in lab diagnostics reveals four types of limitations and their underlying causes:

3.1. Billing

Although the billing of laboratory procedures is highly correlated with the creation of test results and thus combined evaluations may be of interest, both applications require different sets of attributes and can not easily be integrated with the same terminology as shown in [8]. These results can be transferred to the general scope because LOINC's usage-specific content model offers no means to reasonably represent billing data, primarily because of its limitation to precoordinated combinations. The enforced usage of five axes that are partially irrelevant for economic purposes whereas intentionally omitting billing-relevant methods lead to an unsuitable data model.

3.2. Context and interpretive comments

By examining the historic origin as an addition to HL7 standards and the coverage of its accordingly developed content model, LOINC turns out to be not intended as a standalone format to fully describe laboratory test results. Additional information referring to test environment, device characteristics or result specific comments are meant to be communicated with relation to but outside of the term. So, the attempt to interpret laboratory data exclusively based on LOINC code and result value is inherently futile.

3.3. Aggregation of test results

Test results sharing the same LOINC code cannot ensure a simple and correct integration for combined analysis due to some distorting factors. As explained in section 3.2, LOINC terms are lacking some crucial information for complete result interpretation, e.g. different devices and reference ranges leading to incomparable results. Furthermore, seemingly identical tests specified by the same LOINC code may actually refer to different observations caused by an inaccurate mapping of idiosyncratic terms in the first place. Here, the (necessary) complexity of the content model with subcomponents and obscure *Property* values can be identified as a source of error. During mapping, another issue arises from non-existent formal definitions and explicit hierarchies so that easily locating related terms of different granularity (e.g. with a specified method) is prevented. The missing formal conceptualization further restricts advanced aggregation as well. Otherwise, LOINC terms differing only in the *Property* axis may be principally evaluated together when their units of measurement are convertible into one another but both unit defining relations and conversion support are currently missing.

3.4. Molecular genetics and other rapidly changing domains

Because of the limitation to reviewed pre-coordinated terms, LOINC can hardly keep up with newly developing tests and methods. Restricting contents to sensible combinations has obvious advantages in terms of data quality and error prevention but sacrifices any flexibility on the other hand. This rigid content model does not scale well for the numerous and diverse tests required for genetic diagnostics.

4. Discussion

In a complex field of application like the laboratory domain, the presented list of limitations is obviously non-exhaustive, instead we focused on the apparently most urgent issues. An application of LOINC for billing purposes is clearly out of its predefined scope but has previously been discussed and is not far-fetched. For most of the above-mentioned challenges (billing, context, and aggregation), SNOMED CT provides an approach for improvement based on its larger coverage and ontological features, but no complete solution either [8,9]. A combined usage of both terminologies is explicitly favorable. Of course, improvements to LOINC are continually proposed and implemented as well, e.g. we recently developed an ontological representation including LOINC's implicit hierarchies [11], Hauser et al. introduced an automated process for unit conversion of compatible codes [12], and the HL7 Clinical Genomics Work Group published an implementation guide for genetic test reporting with LOINC [13].

5. Conclusion

The long-established content model of LOINC partially restricts the coding system's usefulness for derived, originally not-intended applications needs, for changing demands as well as for test result interpretation and aggregation. This is due to its inflexibility, lacking formalization and incomplete information. LOINC is no all-purpose, standalone format but can benefit largely from further specification, extension and combined usage with other terminologies, primarily SNOMED CT.

References

[1] O. Bodenreider, R. Cornet, D.J. Vreeman, Recent developments in clinical terminologies - SNOMED CT, LOINC, and RxNorm, Yearbook of medical informatics. 27.01 (2018), 129-139.

[2] A.W. Forrey, C.J. Mcdonald, G. DeMoor, S.M. Huff., D. Leavelle, D. Leland, T. Fiers, L. Charles, B. Griffin, F. Stalling, A. Tullis, Logical observation identifier names and codes (LOINC) database: a public use set of codes and names for electronic reporting of clinical laboratory test results, Clinical chemistry. 42.1 (1996), 81-90.

[3] K.D. Pagana, T.J. Pagana, Mosby's Manual of Diagnostic and Laboratory Tests, Elsevier Health Sciences, St. Louis, 2017.

[4] S. Roy, W.A. LaFramboise, Y.E. Nikiforov, M.N. Nikiforova, M.J. Routbort, J. Pfeifer, R. Nagarajan, A.B. Carter; L. Pantanowitz, Next-generation sequencing informatics: challenges and strategies for implementation in a clinical environment, Archives of pathology & laboratory medicine. 140.9 (2016), 958-975.

[5] D.J. Vreeman, J. Hook, B.E. Dixon, Learning from the crowd while mapping to LOINC. Journal of the American Medical Informatics Association. 22.6 (2015), 1205-1211.

[6] D.J. Vreeman, Top 10 Tips for Mapping to LOINC, (2016). https://danielvreeman.com/top-10-tips-for-mapping-to-loinc/ (accessed September 24, 2019).

[7] C. McDonald, S. Huff, J. Deckard, S. Armson, S. Abhyankar, D.J. Vreeman, LOINC Users' Guide. Regenstrief Institute, Indianapolis, 2019.

[8] C.M. Drenkhahn, H. Ulrich, P. Duhm-Harbeck, J. Ingenerf, Mapping of Laboratory Services included in an Internal Catalog to SNOMED CT using the UMLS, Student Conference 2018, Medical Engineering Science, Medical Informatics and Biomedical Engineering. 2018, 337-340.

[9] A. Bietenbeck, M. Boeker, S. Schulz, NPU, LOINC, and SNOMED CT: a comparison of terminologies for laboratory results reveals individual advantages and a lack of possibilities to encode interpretive comments, LaboratoriumsMedizin. 42.6 (2018), 267-275.

[10] M. Stram, T. Gigliotti, D. Hartman, A. Pitkus, S.M. Huff, M. Riben, W.H. Henricks, N. Farahani, L. Pantanowitz, Logical Observation Identifiers Names and Codes for Laboratorians: Potential Solutions and Challenges for Interoperability, Archives of pathology & laboratory medicine. 2019.
[11] C. Drenkhahn, P. Duhm-Harbeck, J. Ingenerf, Aggregation and Visualization of Laboratory Data by Using Ontological Tools Based on LOINC and SNOMED CT, Studies in health technology and informatics, 264 (2019), 108-112.
[12] R.G. Hauser, D.B. Quine, A. Ryder, S.Campbell, Unit conversions between LOINC codes, Journal of the American Medical Informatics Association. 25.2 (2017), 192-196.
[13] J. Deckard, C.J. McDonald, D.J. Vreeman, Supporting interoperability of genetic data with LOINC. Journal of the American Medical Informatics Association. 22.3 (2015), 621-627.