Journal of Biomedical Informatics 45 (2012) 658–666

Contents lists available at SciVerse ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

Auditing consistency and usefulness of LOINC use among three large institutions – Using version spaces for grouping LOINC codes

M.C. Lin [a], D.J. Vreeman [b,c], Clement J. McDonald [e], S.M. Huff [a,d,*]

[a] The Department of Biomedical Informatics, The University of Utah, Salt Lake City, UT, United States
[b] Regenstrief Institute, Inc., Indianapolis, IN, United States
[c] Indiana University School of Medicine, Indianapolis, IN, United States
[d] Intermountain Healthcare, Salt Lake City, UT, United States
[e] Lister Hill Center, National Library of Medicine, Washington, D.C., United States

ARTICLE INFO

Article history:
Received 3 August 2011
Accepted 18 January 2012
Available online 28 January 2012

Keywords:
Controlled vocabulary
LOINC
Evaluation research
Clinical laboratory information systems
LOINC usage
Consistency
Usefulness
Semantic interoperability
LOINC usage
Data exchange standards

ABSTRACT

Objectives: We wanted to develop a method for evaluating the consistency and usefulness of LOINC code use across different institutions, and to evaluate the degree of interoperability that can be attained when using LOINC codes for laboratory data exchange. Our specific goals were to: (1) Determine if any contradictory knowledge exists in LOINC. (2) Determine how many LOINC codes were used in a truly interoperable fashion between systems. (3) Provide suggestions for improving the semantic interoperability of LOINC.
Methods: We collected Extensional Definitions (EDs) of LOINC usage from three institutions. The version space approach was used to divide LOINC codes into small sets, which made auditing of LOINC use across the institutions feasible. We then compared pairings of LOINC codes from the three institutions for consistency and usefulness.
Results: The number of LOINC codes evaluated were 1917, 1267 and 1693 as obtained from ARUP, Intermountain and Regenstrief respectively. There were 2022, 2030, and 2301 version spaces among ARUP and Intermountain, Intermountain and Regenstrief and ARUP and Regenstrief respectively. Using the EDs as the gold standard, there were 104, 109 and 112 pairs containing contradictory knowledge and there were 1165, 765 and 1121 semantically interoperable pairs. The interoperable pairs were classified into three levels: (1) Level I – No loss of meaning, complete information was exchanged by identical codes. (2) Level II – No loss of meaning, but processing of data was needed to make the data completely comparable. (3) Level III – Some loss of meaning. For example, tests with a specific 'method' could be rolled-up with tests that were 'methodless'.
Conclusions: There are variations in the way LOINC is used for data exchange that result in some data not being truly interoperable across different enterprises. To improve its semantic interoperability, we need to detect and correct any contradictory knowledge within LOINC and add computable relationships that can be used for making reliable inferences about the data. The LOINC committee should also provide detailed guidance on best practices for mapping from local codes to LOINC codes and for using LOINC codes in data exchange.

1. Introduction

Consistency and usefulness are two important characteristics of good terminological systems (TSs), especially for information exchange. *As we use the terms in this article, consistency* means that between any two terms within TSs there is no contradictory knowledge (as represented by the implicit or explicit relationships between concepts), and *usefulness* means that there is knowledge in the terminology that allows for creation of an efficient algorithm for making inferences using the relationships in the terminology for supporting different kinds of use, e.g. information retrieval, data integration or clinical decision support [4]. Auditing TSs can be a difficult task because of the huge number of concepts, e.g. LOINC has more than 65,000 codes. In order to reduce this task to a manageable size, researchers have used semantic methods to search for similar concepts in the UMLS [9] or used semantic structures to partition SNOMED into smaller groups [28]. Previous reports have shown that most inconsistencies in LOINC mapping result from choosing codes that vary in the 'method', 'scale' and 'property' characteristics of the codes. [3,20,26]. The use of version spaces is a common technique used in machine learning for concept discovery [24]. Version spaces are used to divide all hypotheses into smaller

* Corresponding author. Address: South Office Building, 5171 South Cottonwood Street, Murray, UT 84107, United States. Fax: +1 801 507 9400.
E-mail address: stan.huff@imail.org (S.M. Huff).

1532-0464/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.jbi.2012.01.008

subspaces to make it possible to search similar concepts by a given set of constraints. This paper describes a systematic method for auditing the consistency and usefulness of LOINC use and discusses potential strategies to approach best practices in the use of LOINC for interoperable data exchange.

1.1. Auditing TSs on policy vs. use

Early papers on TSs development were focused on functional, structural and policy perspectives. These papers include Cimino's desiderata for creating controlled medical vocabularies [10], Chute et al.'s study of functional characteristics of comprehensive health terminology systems in the United States [8], and the technical specification published by International Standard Organization (ISO) – "Health informatics – Controlled health terminology – Structure and high-level indicators" [16]. As TS usage increased, discussions shifted to descriptions of practical use. These studies included analyzing coverage of the UMLS for coding of concepts in the Gene Ontology (GO) [6], comparing coding consistency of SNOMED CT among three commercial coding companies [2] and evaluating the performance of LOINC when comparing laboratory data among three hospitals [3]. To summarize all auditing methods for TSs, Zhu et al. have done a thorough literature review on different auditing methods, including manual, systematic and heuristic methods [29].

1.2. The development of LOINC

1.2.1. Rapid evolution of LOINC model

LOINC provides a universal terminology for reporting laboratory tests and other clinical observations. Since 1994, LOINC has grown from about 6000 codes to more than 65,000 in the current version. As Cimino noted in his desiderata [10], an important characteristic of TSs is to "Evolve Gracefully", and LOINC tries to adhere to this principle [15]. The LOINC committee has emphasized practical experience in using LOINC to improve its design. Whenever the original design of LOINC is not sufficient, the design is enhanced or a new model is created. Before migrating to the current six-axis model, at least four different earlier models were created (Table 1). For example, the first design of LOINC was a four-axis model, but with more implementation, the original model was insufficient for specifying some tests, e.g. it lacked the ability to specify timing (24 h, 12 h, or 4 h, etc.). Therefore, "⟨timing⟩" was added to create a new model.

1.2.2. LOINC in action

Many places adopted LOINC in their daily operations, including large commercial laboratories, hospitals, health care provider networks, insurance companies, and public health departments [21]. Recently, LOINC was adopted as the terminology standard for certification of laboratory orders and results, including electronic reporting of lab results to public health agencies as part of the Centers for Medicare and Medicaid Services (CMS) Electronic Health Record (EHR) "Meaningful Use" incentive program. LOINC was also used in a German Hospital Information System (HIS) to identify the document type of reports sent as Clinical Document Architecture (CDA) documents [14] and to retrieve laboratory data of adverse events automatically from clinical trial databases [7]. LOINC is also frequently used in computerized clinical decision support systems [1]. Although the scope of LOINC covers both clinical and laboratory observations, for the purpose of this paper we focus exclusively on laboratory content.

1.2.3. Evaluations of LOINC

Evaluating LOINC performance in actual practice can help to improve LOINC design. McDonald et al. summarized LOINC development and worldwide use [21]. Lau et al. reported LOINC coverage for the laboratory test dictionary in the US Department of Defense (DoD) [18] and Vreeman et al. reported LOINC coverage for tests in the Indiana Network for Patient Care (INPC) [27]. We also conducted a series of studies about LOINC usage among three large institutions. First, we reported that LOINC codes can cover more than 99% of the volume of every day laboratory tests among two institutions and 79% of tests in a reference laboratory [19]. Second, we evaluated the correctness of LOINC mapping and reported that there were 0.45% (4/884) tests mapped to totally unrelated LOINC codes and 4% (36/884) tests containing at least one error in mapping to the 6 axis model of LOINC [20]. An earlier study by Baorto et al. also evaluated LOINC performance when combining laboratory results associated with congestive heart failure patients among three teaching hospitals [3].

1.2.4. Requirements for ideal LOINC use

According to Devanbu et al.'s defintion of a good knowledge system [13], the best practice of LOINC should have the following characteristics: (1) *Completeness*: it should have all the necessary LOINC codes to cover the domain of interest, (2) *Correctness*: mapped LOINC codes should be faithful to the original meaning of the tests, (3) *Consistency*: the knowledge implied by different LOINC codes should be consistent, e.g. if two different codes have identical meanings, the codes are duplicates and the consistency principle is violated, and (4) *Competence*: usefulness is the fundamental goal of LOINC for supporting use of laboratory data in different fields. Support for the use of ontologic relationships is one of the important competencies of TSs [25]. LOINC should define the relations between codes and combinations of codes that allow users to infer equivalence, if their meanings in data instance representation are interoperable. That is, if the combination of two codes has the same meaning as a single code (a difference in the use of pre- or post-coordination), relationships should exist between the codes that support the assertion of equivalence. Previous evaluations have described LOINC with respect to the first two characteristics, completeness [18,19] and correctness [20]. The focus of this paper is on the evaluation of consistency and usefulness.

1.3. Definition of consistency and usefulness of TSs

1.3.1. Consistency

Consistency in a system implies that the system does not contain contradictory knowledge. Consistency of TSs could be discussed from two perspectives: (1) *Internal consistency:* Inconsistency can

Table 1
Evolution of LOINC model.

	LOINC model	Explanation
1	⟨analyte⟩:⟨specimen⟩:⟨precision⟩:⟨method⟩	Initial model
2	⟨analyte⟩:⟨timing⟩:⟨specimen⟩:⟨precision⟩:⟨method⟩	Adding 'timing' axis
3	⟨analyte⟩.⟨subspecies⟩:⟨property⟩:⟨timing⟩:⟨system⟩:⟨precision⟩:⟨method⟩	Adding chemical subspecies and kind of property
4	⟨analyte⟩.⟨subspecies⟩^⟨chall⟩:⟨property⟩:⟨timing⟩:⟨system⟩:⟨precision⟩:⟨method⟩	Adding 'challenge' information