MEDINFO 2007
HISA Ltd Australia
© 2007 The authors. All rights reserved.

An Analysis for the Automated Mapping of Laboratory Data with Logical Observation Identifier Names and Codes (LOINC)

Hyung Hoi Kim[a c], Bee Sung Kam[a], Shine Young Kim[a], Kyoung-Yong Jee [b], Jeong-Sik Yoon [c], Tran Tung [c] , Hwa-Sun Kim [c], Hune Cho [c], and Yun Sik Kwak [c]

[a]*Dept. of Laboratory Medicine and Biomedical informatics, Pusan National Univ., Republic of Korea*
[b]*Division of IT Services Research, Electronics and Telecommunications Research Institute, Republic of Korea*
[c] *Dept. of Medical Informatics Kyungpook National Univ., Republic of Korea*

Abstract

Logical Observation Identifiers Names and Codes(LOINC) set is the only publicly available universal standard for laboratory test names essential for Electronic Medical Record(EMR) system. Although Regenstrief Institute maintaining the LOINC database provide an automatic LOINC mapping tool (RELMA), It is not sufficient for applying LOINC code set to Laboratory Information System (LIS) of local hospital. LOINC code has up to six attributes and if one of six attributes has the different value, it means different LOINC code is defined. Because many local hospitals have used their own LIS with arbitrary codes witch is not followed by LOINC attributes and some does contain several same attributes, one code of LIS system would match with several LOINC codes. So we cannot help developing protocols and tools to match codes of LIS system to codes of LOINC as one to one. As we have been develop automatic tools to connect LIS system of Pusan National University Hospital (PNUH) with LOINC. Ant it may be valuable as facilities for applying of LOINC code set essential for EMR system to the local hospitals witch has their own LIS system.

Keywords:

logical observation identifiers names and codes; electronic medical record; laboratory information system

Introduction

Development of information technology is bringing many changes to our lives, and such processes of change can also be experienced in the medical environment. Developments are also being made in the public health areas so that data can be automatically retrieved and analyzed, and such data are already being used in the form of electronic documents and being integrated into the Public Healthcare Information System. The Hospital Information System (HIS), enhances information accessibility in the aspects of patient diagnosis, education, research, and heath care management and affects the medical profession in various ways. With Information Technology, HIS surpasses the scope of an automated tool and is evolving as Electronic Medical Record Systems, EMRS.

According to materials that studied status of hospital informationalization at approximately 122 large scale medical facilities including general hospitals, 88.1% are equipped with the Hospital Order Communication System and 9.2% are equipped with the Electronic Medical Records System. As a useful tool that can overcome the restrictions of time and space, access to EMRs is possible anywhere and any time a computer is available, therefore, it is being expanded into various fields such as remote diagnosis and remote education [1].

In order to provide appropriate information that is needed in many areas such as diagnosis and research using the Electronic Medical Records System, the significance and the essence of the data to be used must be communicated accurately to the user or to the user's computer. In reality, however, accurate expression and maintenance of accurate medical records of patients' conditions and their progress of treatment is something that is difficult to accomplish, and exchange of information is also impossible. The main reason is for these difficulties is because majority of hospitals independently operate and manage their HIS or EMRs of medical records. Majority of HIS or EMRs that are being utilized are being operated based on a Relational Database Management System, and because information and data are being expressed based on the relationship between specific key values that are unique to each hospital, accurate communication and transmission of information between disparate systems cannot be accomplished. In addition, since independent systems have been constructed using code switching, message exchange between Disparate Systems are impossible. In order to extract the massive information neglected in independently constructed systems for use as knowledge source and to communicate accurate meaning of medical records through information exchange and sharing between each facility, use of Standard Vocabulary is necessary. Currently, various studies on standard vocabulary systems including Unified Medical Language System (UMLS) [2], Systematized Nomenclature of Human and Veterinary

Medicine Clinical Term (SNOMED-CT) [3], Logical Observation Identifier Names and Codes (LOINC) [4], and International Classification of Diseases Version 10(ICD10) [5] are in progress.

LOINC is a general standard vocabulary system that is used for expressing medical diagnosis results, and of all forms of medical information, a medical diagnosis result is the most objective form of information that expresses and displays the condition of the patient. Also, by providing basic materials with which treatments are made, it is also extremely important in a clinical environment. Standardization of medical diagnosis results not only saves waiting period for patients when transferring to a different hospital or when sharing information between medical facilities it also serves as the basis for Evidence Based Medicine (EBM) [6].

Materials and methods

We have been mapping the LOINC for Local code of 183 blood chemistry in LIS. We need Link code's generation to connect Local code and LOINC code for Automated LOINC mapping.

Logical Observation Identifier Names and Codes

The Logical Observation Identifier Names and Codes (LOINC) code set is the only publicly available universal standard for laboratory test names [7-9]. The current version of the LOINC code set, maintained by the Regenstrief Institute, contains more than 33,000 observations. Each LOINC record corresponds to a single test result and includes fields for specifying component (analyte), property measured, timing, type of scale, and, where relevant, the method used to property measured, timing, type of sample, type of sample, type of scale, and, where relevant, the method used to produce the result. Therefore, each LOINC code has up to six attributes and belongs to a class or group (e.g., Chemistry, Microbiology). The purpose of the LOINC is to facilitate the exchange and pooling of results, such as blood hemoglobin, serum potassium, or vital signs, for clinical care, outcomes management, and research [10,11].

Local Terminology

The Local Terminology (LT) being utilized is organized in the form in which alphabetic code or acronym of classification codes are documented and expanded into a numeric code based on an independent classification system (Figure 1)

Table 1 – LOINC specimen abbreviation to available in LOINC database

Specimen code	Specimen code name	LOINC specimen abbreviation
A0	Blood, Plain tube	SER/PLAS
C0	Body fluid	FLD
C4	Cerebrospinal fluid(CSF)	CSF
C7	Body fluid, gastrointestinal	GAST
C8	Joint(synovial) fluid	SNV
CB	Peritoneal(ascetic) fluid	PRT
CC	Pleural fluid	PLR
Y0	Urine	UR

This is the form most commonly used by majority of HIS being operated based on RDBMS. LT is linked with specimen code more than one in single laboratory code [12]. However, LOINC must recognize by laboratory test that same test is different if specimen is different. Therefore, we created link code so that compose of LT and specimen code. As a result, LT could correspond with LOINC code (Fig. 1).

Lab. Code	Specimen code	Specimen code name	Tube code	tube	unit
LCLDH	A0	Blood, Plain tube	55	SST tube	IU/L
LCLDH	C0	Body fluid	37	Plain tube	IU/L
LCLDH	C4	Cerebrospinal fluid	37	Plain tube	IU/L
LCLDH	CC	Pleural fluid	37	Plain tube	IU/L
LCLDH	Y0	Urine	28	Culture cup	IU/L

LCLDHY0

Figure 1- LT of PNUH and formation of Link code for LOINC

Development tools

For the purpose of this study, in order to map and simulate the LT being used at a university hospital with the LOINC, Oracle 10g DBMS Version 10.0.2, a relational database management system by Oracle was installed on a PC. In addition, for development and management of an efficient database, Toad for Oracle Freeware Version 8.0.5.5 by QuestSoft was used as a Database Managing Tool and for development to User Interface of the application for viewing LIS (Laboratory Information System) results, Visual Basic 6.0 Enterprise by Microsoft was utilized.

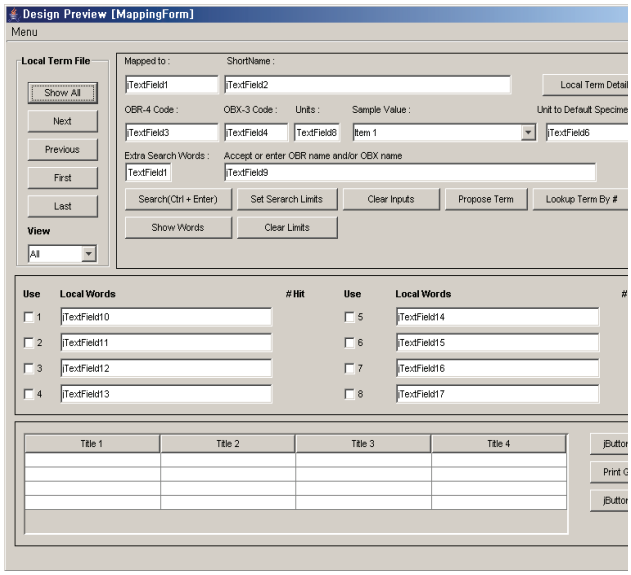

Figure 2- Screen capture of original mapping program, in mapping process

Natural Language Processing

The above figure represents the natural language processing. In this process, there are two steps such as normalization process and words index. Normalization process is get local terminology as input, it convert this string to uppercase string. And then process split string into array of words, and eliminate all of special characters such as ‘, “, ?, etc. Eventually, the process will give array of words as result [12, 13].

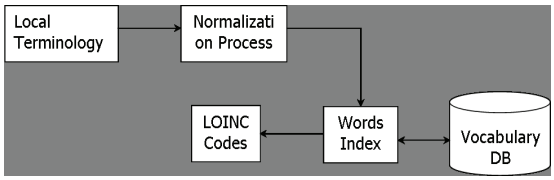

Figure 3- Natural language process

The words Index process gets result of Normalization process as input data. Each word is used in order to look for LOINC within Vocabulary database. Eventually, the output is LOINC that related to LT. In the summary, the natural language processing is used for looking for all of LOINC that related to Local terminologies.

KEYWORD	LOINC_NUM
ABXBACT	91-9
ABXBACT	92-7
ABXBACT	93-5
ABXBACT	94-3
ABXBACT	95-0
ABXBACT	96-8
ABXBACT	97-6
ABXBACT	98-4
ABXBACT	99-2
ABY	1000-9
ABY	1001-7
ABY	1002-5
ABY	1010-8
ABY	1011-6
ABY	1012-4
ABY	1016-5
ABY	1017-3
ABY	1018-1

Figure 4- Results of Natural language process

Vocabulary database includes LOINC table, keywords table, hierarchy data, and so on. The keywords table is called dictionary. The words index process use keywords table in order to look for LOINC.

Results

137 codes (excepted battery item of 183 Link code) did automatic mapping by results of 133 (97%). The failed Link code did Mapping in LOINC manually. Table 2 is result for automated LOINC mapping tools.

Table 2- Result for automated LOINC mapping tools

Specimen code name	Code Link	LOINC
Body fluid	LCLDHC0	22989-1
Cerebrospinal fluid(CSF)	LCLDHC4	22958-7
Body fluid, gastrointestinal	LCLDHC7	22989-1
Joint(synovial) fluid	LCLDHC8	23141-3
Peritoneal(ascetic) fluid	LCLDHCB	22989-1
Pleural fluid	LCLDHCC	23019-5
Urine	LCLDHY0	22958-7

This study was conducted in order to identify a method for efficient information exchange between disparate systems through informationalization of massive medical data that are stored in independent HIS. For the purpose of the study, 183 chemistry codes being utilized at a university hospital were randomly selected and mapped with LOINC, and as can be seen in Table 2, in a single examining room UI, independent use of LT or LOINC is possible, and mixed use is also possible. At this time, items that are displayed to the user are only partially shown, and because information that are internally hidden includes various types of information as seen possibility of efficient information exchange with a disparate system was verified.

Therefore, user convenience in an independent HIS can be secured by utilizing a LT and information exchange with disparate systems can be accomplished by utilizing the HL7 Interface Engine of the LOINC.

Discussion

In this study, by mapping diagnosis results codes and their expression methods that are currently being utilized in hospitals with standardized codes, a fundamental study for standardized information sharing of medical information and Evidence Based Medicine was conducted. Of all medical records, medical diagnosis result is the most objective information that provides reliable data. LT being independently used in hospitals can be applied and utilized various ways within each hospital, but because the vocabulary system of each hospital are completely different, diagnosis results of one facility cannot be used directly at another facility. In order to extract and exchange the massive amount of information stored within medical information systems as knowledge information, a study on a method for efficiently utilizing a standardized vocabulary system must be conducted.

An applicability of a standard vocabulary system was evaluated by mapping an LT of a hospital system with LOINC, an international standard coding system; and rather than terminating the study at standardization of codes, a database was designed to maintain their 6 axial forms so that the information to be exchanged can remain their compatibility. Driven by development of information technology and specialization of medical technology, importance and need for real time remote diagnosis and exchange of medical information between hospitals are increasing. This study is scientifically significant because it serves as substantial basic work for electronic data exchanges in the medical environment.

Acknowledgments

This article is based on research supported by grant No. (R-2006-000-10009-0) from the Basic Research Program of the Korea Science & Engineering Foundation, with funds provided by the second stage of BK21.

References

[1] Kyoung-Jin L, Jieun C, Sung-Chul H, et al. Development of a Medical Ontology Library: Analysis of the Clinical Terms in the Medical Records of a COPD Patient. *J Korean Med Inform* 2006; 12(1): 21-29.

[2] UMLS Knowledge Source Server, U.S. National Library of Medicine, 8600 Rockville Pike, Bethesda, MD 20894, National Institutes of Health, Available from: http://umlsks.nlm.nih.go.

[3] The Systemized Nomenclature of Medicine (SNOMED) and College of American Pathologists (CAP), Available from: http://www.sNomed.org,

[4] Loinc, Avalilable from: http://www.regenstrief.org/loinc.

[5] International Classification of Diseases version 10, World Health Organization (WHO), Available from: http://www.who.org/.

[6] Hsu C, Goldberg HS. Knowledge-mediated retrieval of laboratory observations. *Proc AMIA Symp* 1999:809-13.

[7] Bakken S, Cimino JJ, Haskell R, et al. Evaluation of the clinical LOINC (Logical Observation Identifiers, Names, and Codes) semantic structure as a terminology model for standardized assessment measures. *J Am Med Inform Assoc* 2000; 7(6): 529-38.

[8] Matney S, Bakken S, Huff SM. Representing nursing assessments in clinical information systems using the logical observation identifiers, names, and codes database. *J Biomed Inform* 2003; 36(4-5): 287-93.

[9] McDonald CJ, Huff SM, Suico JG, et al. LOINC, a universal standard for identifying laboratory observations: a 5-year update. *Clin Chem* 2003; 49(4): 624-33.

[10] McDonald CJ. Need for standards in health information. *Health affairs* 1998; 17(6): 44-46.

[11] Khan AN, Russell D, Moore C, et al. The map to LOINC project. *AMIA Annu Symp Proc* 2003:890.

[12] Meystre S, Haug PJ. Natural language processing to extract medical problems from electronic clinical documents: performance evaluation. *J Biomed Inform* 2006; 39(6):589-99.

[13] Chapman WW, Christensen LM, Wagner MM, et al. Classifying free-text triage chief complaints into syndromic categories with natural language processing. *Artif Intell Med* 2005; 33(1):31-40.

Address for correspondence

Hune Cho, ,PhD
Department of Medical Informatics
Kyungpook National University School of Medicine, Jung-Gu
Dongin-Dong 2-101, Daegu Korea, 700-422
Phon: +82-53-420-4896
Fax: +82-53-423-1242
e-mail: hunecho@knu.ac.kr