

Proposed Algorithm with Standard Terminologies (SNOMED and CPT) for Automated Generation of Medical Bills for Laboratory Tests

Shine Young Kim, MD, MS¹, Hyung Hoi Kim, MD, PhD¹, In Keun Lee, PhD², Hwa Sun Kim, RN, PhD^{2,3}, Hune Cho, PhD^{2,3}

¹Department of Laboratory Medicine and Biomedical Informatics, College of Medicine, Pusan National University, Busan; ²MIPTH, Kyungpook National University, Daegu; ³Department of Medical Informatics, College of Medicine, Kyungpook National University, Daegu, Korea

Objectives: In this study, we proposed an algorithm for mapping standard terminologies for the automated generation of medical bills. As the Korean and American structures of health insurance claim codes for laboratory tests are similar, we used Current Procedural Terminology (CPT) instead of the Korean health insurance code set due to the advantages of mapping in the English language. **Methods:** 1,149 CPT codes for laboratory tests were chosen for study. Each CPT code was divided into two parts, a Logical Observation Identifiers Names and Codes (LOINC) matched part (matching part) and an unmatched part (unmatched part). The matching parts were assigned to LOINC axes. An ontology set was designed to express the unmatched parts, and a mapping strategy with Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) was also proposed. Through the proceeding analysis, an algorithm for mapping CPT with SNOMED CT arranged by LOINC was developed. **Results:** 75% of the 1,149 CPT codes could be assigned to LOINC codes. Two hundred and twenty-five CPT codes had only one component part of LOINC, whereas others had more than two parts of LOINC. The system of LOINC axes was found in 309 CPT codes, scale 555, property 9, method 42, and time aspect 4. From the unmatched parts, three classes, 'types', 'objects', and 'subjects', were determined. By determining the relationship between the classes with several properties, all unmatched parts could be described. Since the 'subject to' class was strongly connected to the six axes of LOINC, links between the matching parts and unmatched parts were made. **Conclusions:** The proposed method may be useful for translating CPT into concept-oriented terminology, facilitating the automated generation of medical bills, and could be adapted for the Korean health insurance claim code set.

Keywords: Logical Observation Identifiers Names and Codes, Systematized Nomenclature of Medicine, Current Procedural Terminology, Health Insurance

Received for review: September 2, 2010

Accepted for publication: September 27, 2010

Corresponding Author

Hwa Sun Kim, RN, PhD

Department of Medical Informatics & MIPTH, Kyungpook National University, 2-101 Dongin-dong, Jung-gu, Deagu 700-422, Korea. Tel: +82-53-420-4896, Fax: +82-53-423-1242, E-mail: pulala@paran.com

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

© 2010 The Korean Society of Medical Informatics

1. Introduction

Most medical billing processes are done electronically, but the manner of generating bills is still largely manual, even though the electronic medical record (EMR) and the order communication system (OCS) have been widely used. Achieving semantic interoperability not only among hospital systems (EMR, OCS, generation of bills), but also among many systems outside the hospital, requires that clinical data elements are captured in a standardized form [1]. However, several 'standards' exist, even in areas of medical terminol-

ogy such as Systematized Nomenclature of Medicine-Clinical Terms (SNOMED CT), Logical Observation Identifiers Names and Codes (LOINC), and Current Procedural Terminology (CPT), which is used only in the US. Each standard has its own purpose and the structure fits its goals, so adoption of several standard terminologies simultaneously is inevitable. Their coverage, however, often overlaps [2-4]. For interoperable health information technology to become a reality, reference mapping among standard terminologies is necessary.

SNOMED CT, LOINC, and CPT would seem to be the key terminologies for the automated generation of bills for laboratory tests in an EMR environment. Because CPT is the most widely accepted medical nomenclature used to report medical procedures and services under public and private health insurance programs, clinical data stored in SNOMED CT format should be translated to CPT for the automated generation of bills in the US. However, direct mapping between them would be difficult because SNOMED CT codes should be post-coordinated to express the CPT code and a CPT code can be expressed by numerous post-coordinated SNOMED CT code sets. The situation is similar in Korea; the only difference is the code set.

If the automated generation of billing algorithms with CPT codes and other standard terminologies were possible, then it could also work with the Korean health insurance claim code set. Because the LOINC and SNOMED CT code have not been translated into Korean, CPT is more feasible for the purpose of determining the possibility of creating algorithms for automated bill generation. Therefore, we decided to first try using the CPT code.

LOINC is concept-oriented terminology, and cross mapping tables exist between SNOMED CT concepts and LOINC concepts, provided by the International Health Terminology Standards Development Organization (IHTSDO) [5]. Also, the owners of SNOMED CT and LOINC have announced that they would cooperate in the generation of laboratory test terminology content [6]. Although the cross mapping table between SNOMED CT and LOINC is incomplete, the method of mapping between them can be reused to map between SNOMED CT and CPT, if the CPT codes for laboratory tests could be expressed in LOINC structure [7].

Thus our first action was to evaluate the utility of the LOINC semantic structure as a terminology model for representing CPT, by dividing CPT items into LOINC axes. Second, CPT codes for laboratory aspects have more information, which cannot be covered by LOINC, but is expressed by SNOMED CT, which has a larger coverage. We present a

proposed ontology for the categorization of the area of CPT that is less covered by LOINC [8]. Finally, a proposed mapping strategy between CPT and SNOMED CT using LOINC structure is discussed.

II. Methods

1. Analyzing Sentence of CPT Codes

In total, 1,149 codes in the 'Pathology and Laboratory' section (CPT code 80047-89356) of CPT, except the 'surgical pathology,' 'cytopathology,' and 'anatomic pathology' subsections, which are not laboratory tests, were analyzed.

Two clinical laboratory medicine doctors dissected each sentence of the CPT codes, and discrepancies were discussed and mutual agreement was reached. Each part of the CPT sentence was assigned to the six LOINC axes (component, system, property, scale, time aspect, method); remaining parts of the sentence that could not be included in LOINC axes were recorded separately.

2. Assigning CPT Codes into LOINC Axes

Because there are many synonyms in the CPT and LOINC codes, the individual parts of CPT were not mapped, but rather were manually assigned to one of the LOINC six axes, regardless of the presence of CPT words in the LOINC database.

Component, one of the axes of LOINC, was force-assigned in all CPT codes, even in cases when CPT codes did not have a real 'component.' For example, 82397 (CPT code), 'chemiluminescent assay'; this CPT code contains information only about a method, without information about component. In this case, the component was assumed to be 'any component' available, and an annotation was attached and separately recorded for mapping with SNOMED CT (e.g., 'any component could be allowed' was added as additional information as a subpart for SNOMED CT mapping).

Rules in notes under each subsection of the CPT code book were applied to all subcodes in that subsection for assignment to LOINC axes. For example, notes in the 'therapeutic drug assays' subsection state that 'examination is quantitative.' Thus, we regarded all the codes under this subsection as quantitative tests and assigned them a quantitative 'scale' (one of the LOINC axes). Another example is in the 'urinalysis' subsection; if specific codes did not define the 'system' (one of the LOINC axes), we deemed the 'system' to be urine.

CPT codes including words of 'unlisted tests' in CPT were not dissected, and the full sentence was assigned to a component of LOINC if other parts existed that could match with the axes of LOINC. For example, 85999 (CPT code), 'Unlisted

hematology and coagulation procedure’ was assigned to the component (force-assign, any component), and an annotation was recorded as a subpart for SNOMED CT mapping: ‘any component could be allowed in the hematology and coagulation section.’

The following is a general example of dissecting a CPT code into the axes of LOINC: 84156 (CPT codes); ‘Protein, total, except by refractometry: urine’ → ‘component; protein, total,’ ‘system: urine,’ ‘method: except by refractometry’ and ‘annotation for SNOMED CT mapping: (method) other than refractometry.’

3. Categorization of CPT Subpart for SNOMED CT Mapping

Some CPT codes had extra information that could not be covered by the LOINC semantic structure including some annotations created during LOINC assign. All such information was collected, analyzed, and categorized. We created three classes for categorization of subparts for SNOMED CT mapping: types, subjects, and objects. We also defined properties to be used for specifying the meaning of sentences. The ‘types’ class is defined as one of ‘allowance or restriction’ and represents the general meaning of the subpart for SNOMED CT. The ‘types’ class has one of the following properties: any, each, several, except for. The ‘object’ class is the object of the types-class expression in the sentence. The ‘subject’ class was created for presenting subparts of SNOMED CT mapping

more clearly. Most subjects of collected data were concerned with the axes of LOINC, and to make definite relationships with LOINC axes and such information, the axes of LOINC were also used as key ‘subjects.’ For example, the sentence, “This CPT code allowed for any component (within assigned LOINC axes).” can be expressed as Subjects (component) + Types (allowance, any) + Objects (component) (+ subparts of assigned LOINC axes). In this case, the subjects class is not meaningful. Here is another example: The sentence, “This CPT code could be chargeable whenever each test is performed with another system.” Can be expressed as Subjects (charge) + Types (allowance, each) + Objects (system). We simplified the subparts for SNOMED CT mapping with the rules described.

III. Results

1. Assigning CPT Codes into LOINC Axes

All of the analyzed CPT codes were forced to be assigned into a component part of LOINC, as described in the Methods. The system of LOINC axes was found in 309 CPT codes, scale 555, property 19, method 412, and time aspect 4. The scale was usually expressed in CPT codes as ‘quantitative,’ ‘qualitative,’ ‘qualitative or semiquantitative,’ and ‘ratio.’ The scale of the LOINC axes could be analogized from the properties of LOINC. For example, test results with a ratio property were assumed to be on a quantitative scale. Table

Table 1. Pattern and number of CPT codes assigned with the axes of LOINC

Component	Six axes of LOINC					Cases of CPT codes
	System	Scale	Property	Time aspect	Method	
X		X				302
X						225
X					X	194
X		X			X	112
X	X	X				109
X	X					83
X	X				X	72
X	X	X			X	29
X			X			7
X	X		X			7
X	X		X		X	5
X	X	X		X		3
X	X			X		1
					Total	1,149

CPT: Current Procedural Terminology, LOINC: Logical Observation Identifiers Names and Codes.

1 shows how the CPT codes were dissected into the six axes of LOINC. CPT codes containing component and scale only were the most common.

2. Categorization of CPT Subpart for SNOMED CT Mapping

Of the 1,149 CPT codes we analyzed, 351 had additional information that did not match the axes of LOINC. We categorized the remaining sentences as described. The 'types' class was defined as 'allowance or restriction.' The contents (instances) of the object were determined, such as component, system, method, time, number, purpose, diagnosis, and XXX (other) in our study. We also determined the 'subjects' class to be subject to combinations of type and object. The class of 'subjects' was supposed to have firm instances in the six axes of LOINC and 'charge' by definition. Table 2 presents the extracted combination of type, property, object, and subject to classes from the remaining sentences of CPT codes. All combinations of 'Type + Object + Subject to' would be possible, at least theoretically, but those shown in Table 2 were the only combinations extracted in our study. The instance

number and properties are flexible and changeable, as are combinations of them.

3. Schema for Mapping CPT with SNOMED CT

SNOMED CT provides an integration table with LOINC. It contains concept identifiers from SNOMED CT that relate specific components of the LOINC test to the SNOMED CT hierarchy [9]. 'RelationshipType' in SNOMED CT is a concept identifier (ConceptID) from the SNOMED CT concepts table and defines the relationship between the LOINC name and the target SNOMED concept. Because components of the 'subject to' class, except 'charge,' are the same as the six axes of LOINC, the relationship among components of the 'subject to' class can be readily expressed by SNOMED CT RelationshipType. Then, if the relationship among 'types,' 'property,' and 'object' and the relationship between 'charge' (extracted from CPT, which is a part unmatched with LOINC) and the LOINC axes could be defined with SNOMED CT ConceptID, the sentences of CPT code could be translated into SNOMED CT without ambiguity (Figure 1).

Table 2. A conceptual structured combinations of CPT sentences that were not directly dissected into LOINC axes

Types	Classes		Description	No.
	Objects	Subjects		
Allowance (any)	Component	Component	Any component would be allowed	22
Allowance (any)	System	System	Any system would be allowed	43
Allowance (any, NOS)	Component	Component	Any component would be allowed, not other specific	20
Allowance (each)	Component	Charge	Charge whenever each component performed	41
Allowance (each)	System	Charge	Charge whenever each test performed with other system	65
Allowance (each)	Method	Charge	Charge whenever each test performed with other method	23
Allowance (each)	Time	Charge	Charge per time	11
Allowance (several)	Component	Component	Several component	78
Restriction (except)	Component	Component	Other than some component	7
Restriction (except)	Method	Method	Other than some method	12
Restriction (except)	System	System	Other than some system	13
Restriction (for)	NOS	Charge	Charge only when meeting XXX restriction	2
Restriction (for)	System	Charge	Charge only when meeting restriction for system	4
Restriction (for)	Number	Component	Restriction for component number is present	22
Restriction (for)	NOS	Component	Other restriction for component is present	3
Restriction (for)	Purpose	Component	Restriction for component purpose is present	21
Restriction (for)	Diagnosis	Component	The component would not be allowed if it would not meet to specific diagnosis, disease or clinical condition	20
Restriction (for)	System	System	Restriction for system of specific condition is present	13
Total				420

CPT: Current Procedural Terminology, LOINC: Logical Observation Identifiers Names and Codes.

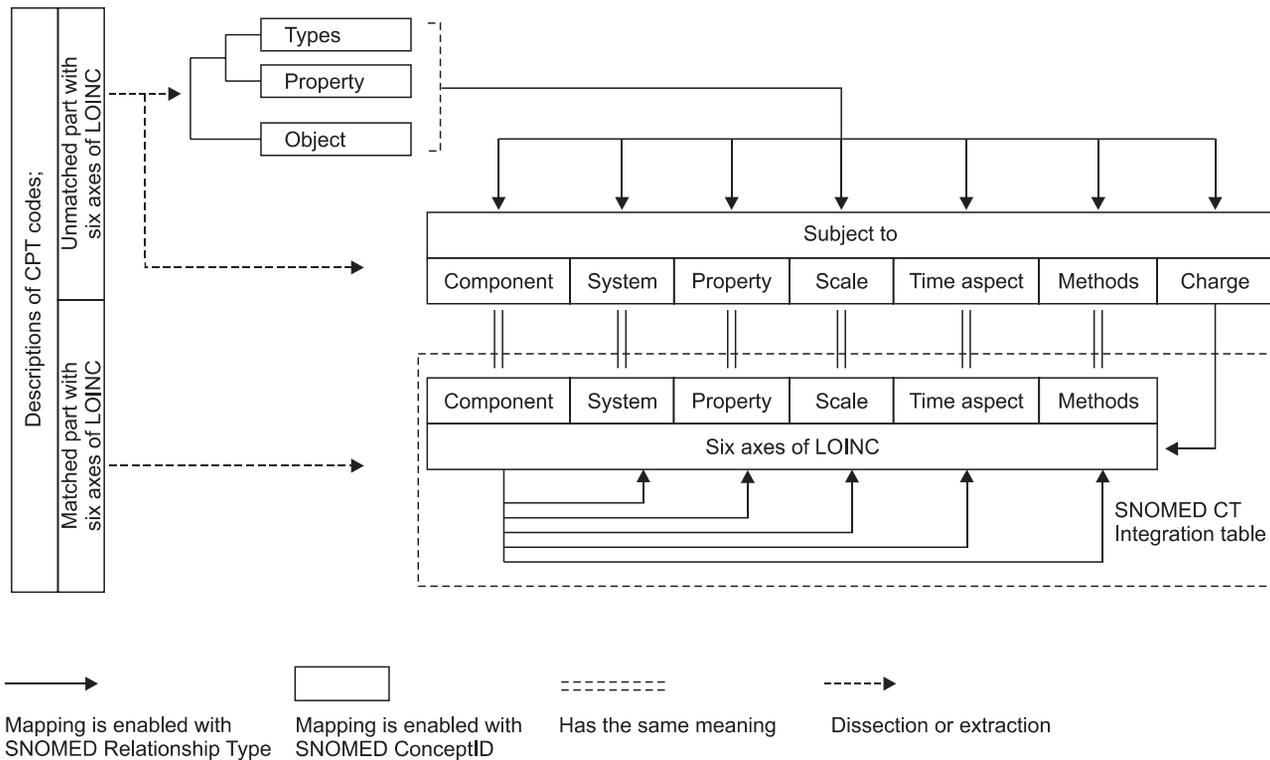


Figure 1. Schema mapping CPT with SNOMED CT.

CPT: Current Procedural Terminology, LOINC: Logical Observation Identifiers Names and Codes, SNOMED CT: Systematized Nomenclature of Medicine Clinical Terms.

In this schema, the instances of parts are flexible. New instances in the property or object class did not break down the structure of the coordinated SNOMED CT codes system. We supposed that every instance (voluntarily created) of each class could be matched with SNOMED CT ConceptID or post-coordinated codes.

IV. Discussion

For the automated generation of bills for laboratory tests, we adapted the LOINC structure to CPT. A draft version of a mapping table between LOINC and CPT was published by the National Library of Medicine in 2006 [5]. However, it did not contain all possible matches, and it was just a mapping table, not a mapping method or algorithm. Thus, it was limited in that it could not keep up with a new version of LOINC or CPT, which is why we developed a mapping algorithm that would be less influenced by the contents of LOINC or CPT.

In some studies, the relationship between the axes of LOINC has been analyzed for each test and has a different SNOMED CT code, case by case [10]. However what we wanted was not a real-world mapping with SNOMED CT, but CPT with SNOMED CT using the LOINC structure.

This means that the role of SNOMED CT is not to express real things at this time. For the automated generation of bills, the role of SNOMED CT is cross bridging between CPT and LOINC. If the LOINC codes were used to order OCS, additional information for matching CPT is required. This seeking process would be possible by defining SNOMED RelationshipType of the 'type' class, 'object' class, and 'subject' class [9]. Each LOINC code fully describes its own meaning; the relationship among the six axes need not 'describe a real meaning,' but 'describe only which axis is concerned,' with SNOMED CT. This algorithm can be applied with any code set that represents laboratory tests, including the Korean code set, due to the excellence of the LOINC axes structure.

In the results of dissecting CPT into LOINC axes, almost half of the CPT codes were dissected into more than three axes of LOINC (Table 1). If CPT codes have more matching parts for the axes of LOINC, the less LOINC codes can be mapped to CPT code (Generally, one CPT code can be mapped with multiple LOINC codes, which is also true for the Korean code set). About 25% of CPT codes had only one matching part of LOINC, 'component,' and some concepts of CPT could not be found among LOINC items. At this time, like the procedure code, the SNOMED CT ConceptID could be used instead of LOINC to express concepts of CPT.

Table 2 shows our suggestion for categorizing subparts for SNOMED CT mapping. It is just one option for categorization. We hope this can be developed further by experts in the field.

The algorithm developed in this study has some limitations. First, to express instances of each class, post-coordination of SNOMED CT may be needed. Some relationships between type class and object class or charge and subject to class and other classes may not be found in SNOMED CT ConceptID. To minimize the use of post-coordination, more and more delicate categorizations of classes would be required. Second, this method can be applied only in the area of laboratory tests. Finally, the generated SNOMED CT code in our algorithm could not fully describe real things, and is only useful for generating bills.

Our study presents a method for mapping between SNOMED CT and CPT laboratory test concepts through LOINC for automated coding to CPT from EMR data recorded with SNOMED CT for billing purposes. The CPT codes are used in the US, but not in Korea. Nonetheless, we suggest that the algorithm for mapping could be widely used, and would also work with the Korean health insurance claim code set.

Conflict of Interest

No potential conflict of interest relevant to this article was reported.

Acknowledgments

This article is based on research supported by the R&D Program of MKE/KEIT (KI10033576, KI10033545).

References

1. Garde S, Knaup P, Hovenga E, Heard S. Towards semantic interoperability for electronic health records. *Methods Inf Med* 2007; 46: 332-343.
2. Cimino JJ. Review paper: coding systems in health care. *Methods Inf Med* 1996; 35: 273-284.

3. American Medical Association. CPT: Current Procedural Terminology [Internet]. Chicago: American Medical Association; 2009 [cited 2010 Feb 23]. Available from: <http://www.ama-assn.org/ama/pub/physician-resources/solutions-managing-your-practice/coding-billing-insurance/cpt.shtml>.
4. 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.
5. National Library of Medicine (US). LOINC to CPT Mapping [Internet]. Bethesda, MD: National Library of Medicine (US); 2006 [cited 2010 Feb 23]. Available from: http://www.nlm.nih.gov/research/umls/mapping_projects/loinc_to_cpt_map.html.
6. Regenstrief Institute, LOINC Committee, IFCC, IUPAC, IHTSDO. Owners of LOINC, NPU, and SNOMED CT begin trial of cooperative terminology development [Internet]. Indianapolis: The Regenstrief Institute, Inc.; 2009 [cited 2010 Feb 23]. Available from: <http://loinc.org/news/owners-of-loinc-npu-and-snomed-ct-begin-trial-of-cooperative-terminology-development-joint-press-release.html>.
7. Donnelly K. SNOMED-CT: the advanced terminology and coding system for eHealth. *Stud Health Technol Inform* 2006; 121: 279-290.
8. Zhang S, Bodenreider O. Experience in aligning anatomical ontologies. *Int J Semant Web Inf Syst* 2007; 3: 1-26.
9. International Health Terminology Standards Development Organisation (IHTSDO). SNOMED Clinical Terms[®] Technical Reference Guide [Internet]. Copenhagen: IHTSDO; 2008 [cited 2010 Feb 23]. Available from: <http://web.his.uvic.ca/Research/HTG/index.php?ContentFileId=57>.
10. Bodenreider O. Issues in mapping LOINC laboratory tests to SNOMED CT. *AMIA Annu Symp Proc* 2008: 51-55.