International Journal of Database Theory and Application
Vol.9, No.11 (2016), pp.49-60
http://dx.doi.org/10.14257/ijdta.2016.9.11.05

A Study on a Comparison of Diagnostic Domain between SNOMED CT and Korea Standard Terminology of Medicine

Mijung Kim[1*]

[1]*Dept. Of Health Administration, Cheju Halla University, Korea*
[1]*mjkim@chu.ac.kr*

Abstract

Applying to standard clinical terminologies is essential for understanding the precise meanings of the clinical terminology used in EMR systems and sharing clinical data among health providers. In Korea, Korean standard terminology of medicine was first introduced as national standard vocabulary for EMR systems in 2014. However, there is little usage yet. So we studied on a comparison of diagnosis domain between Korean standard terminology of medicine and SNOMED CT adapted in many countries. Because it is the most important in medical statistics and clinical studies. Qualitative analysis, quantitative analysis and mapping ability were studied through literature review and structure analysis as methods. As a result, Korean standard terminology of medicine was satisfied in the concept orientation, concept permanence, non- semantic concept identifier and mapping to standards terminologies, support for multiple level. But it was not in the multi-hierarchy, language independence, formal definition and so on. And some problems was raveled in the structural aspect and mapping. This paper will help to utilize and improve KOSTOM.

Keywords*: medical terminology, standards, requirements for terminology*

1. Introduction

The beginning of EHR with the enormous accumulation of medical information is increasing interest in utilizing medical information. Applying to standard clinical terminologies is essential to ensure interoperability of medical information among medical hospitals and increase the utilization value of the information [1]. Terms of various areas, for example, diagnose, procedures as well as administrative information, lifestyle, accident are recorded in EMR. So, a number of clinical data were input as free text and the terms of the limited representation possible like diagnosis, procedure, chief complaint *etc.* were input as coded data. In the past, medical records were rarely to exchange information with external hospitals. In recent years, because it activates the electronic exchange of patient information Cross hospitals and telemedicine, the demand for medical information exchange among countries and hospitals, so increased interest in domestic or international standards terms.

In the United States and Europe it has been recognized the importance of the standard terminology for a long time, and actively used to specify the standard terminology in the medical field as a national level. By contrast, in Korea, KOSTOM was first adopted as a national standard terminology for clinical terms and details like diagnosis, examination and drug *etc.* in medical records by health care providers [2]. However, the history of KOSTOM is very short compared to other standard terminologies, there is little hospital utilize this because it is not mandatory to apply for the introduction. Most of the domestic hospitals are using a self-developed terminology. Some of university hospitals are using the international standards such as SNOMED CT, UMLS and LOINC, *etc.* [3].

* Corresponding author

ISSN: 2005-4270 IJDTA

In case of disease, most healthcare computer systems use nationally prescribed coding systems such as the ICD-10 for insurance claims and calculating statistics. It uses a position dependent hierarchical coding structure. ICD-10 as representative classification of clinical area categorize similar diseases as one ICD code and allows NEC code. Categorizing is useful as administrative purpose but in case of a detailed study of the disease or medical purposes it cannot be distinguished only by ICD codes. For example, ICD-10 does not specify whether the eye is left or right or whether the fracture is simple, spiral, or compound. Therefore, Hospital systems using EMR give a unique code in hospitals to disease that is familiar and specific to user and map to ICD 10 codes. As a result, efforts and costs are needed in building interface terminology and mapping to classifications for themselves in the hospital. Reference terminology that can solve these problems was developed to cover interface terminology and classification using in the hospital. SNOMED CT is a representative reference terminology. Korean standard terminology of medicine (KOSTOM) has also the similar purpose with SNOMED CT as terminology that standardizes clinical data in EMR and map to standard terminology.

The purpose of this study is to suggest the way of improvement of KOSTOM by Comparative Analysis between KOSTOM and a health care standard terminology adapted and used in many countries.

2. Kinds of Clinical Terminology

2.1. Interface Terminology

The Terminologies can be distinguished according to their nature as interface terminology, Reference terminology and administrative terminology [4]. Clinical interface terminology is a systematic collection of health care-related terms that supports clinicians' entry of patient-related information into computer programs. It also can facilitate display of computer-stored patient information to clinician-users [5]. Hospitals developed and are using interface terminology for consistent input and output of medical information in their EMRs. These terminologies generally embody a rich set of flexible, user-friendly phrases displayed in the graphical or text interfaces of specific computer programs. Clinical interface terminologies have been used for problem list entry, clinical documentation in electronic health record systems, text generation, care provider order entry with decision support and diagnostic expert programs [5].

The controlled medical vocabulary based on medical concepts have been developed to integrate the resources of the hospital information system and to enable precise meaning and representation of medical information [6].

2.2. Reference Terminology

Reference terminology designed to uniquely represent concepts and provide common semantics for diverse implementations. It does this by listing the concepts and specifying their structure, relationships and, if present, their systematic and formal definitions. Entry terms can then map to explicitly defined concepts in a more formal terminology, such as a reference terminology, which can then define relationships among concepts [7]. Reference terminology was made to communicate meanings of terms mutually by assigning the same concept code to the terms having the same meaning even though different descriptions. It is also possible to map the concepts of reference terminology onto the concepts defined in other standard terminologies [8]. It is a set of atomic level designations structured to support representations of both simple and compositional concepts independent of human language (within machine) [8].

The National Library of Medicine has developed the Unified Medical Language System (UMLS) as a terminology resource. UMLS is aimed at the integrated management of existing biomedical terminologies and includes more than 140 terminologies [9]. The

UMLS contains a number of knowledge sources and tools, including the UMLS Metathesaurus and Semantic Network. The UMLS Metathesaurus is a large multipurpose multilingual vocabulary database that contains over five million terms and one million concepts covering information about health and biomedical concepts, their names, and relationships between them. On the other hand, SNOMED CT is made to support vocabularies of the overall area required to assist in the recording of the clinical data [10]. In UMLS and SNOMED CT, their own concept unique identifier is assigned if terms are different but concept is same.

2.3. Administrative Terminology

Classifications allocate things into groups or classes, while coding is the allocation of identifiers, which can apply to anything. It does not of itself have to be a faithful representation of medical knowledge, but rather can be built for some administrative purpose that is fit-for-purpose. Classification is the basis for most statistical analysis, quantitative management, accountancy, and research. So classification is also referred to as administrative terminology. Classification is a set of concepts pre differentiated the terms used in a particular area in order to clean up the data in accordance with specific purposes. This classification code is selectable in predetermined. other classification cannot. All terms of the administration terms are classified into mutually exclusive categories according to a predetermined rule. Administrative terminology is called as classification. ICD-10 is representative medical standard classifications. The ICD (International Classification of Diseases) is the standard international diagnostic classification for general epidemiological, health management, and clinical uses.

3. Requirements for Terminology

3.1. Desiderata for Controlled Medical Vocabularies

In 1997, Cimino produced a paper, Desiderata for Controlled Medical Vocabularies in the Twenty-First Century, which brought together a number of common requirements for clinical terminologies [11]. The desiderata are vocabulary content, Concept Orientation, Concept Permanence, Nonsemantic Concept Identifier, Polyhierarchy, Formal Definitions, Reject “Not Elsewhere Classified”, Multiple Granularities, Multiple Consistent Views, Beyond Medical Concepts(Representing Context), Evolve Gracefully and Recognize Redundancy. In terms of scope and quality, this is paramount. Any practical clinical terminology needs to be comprehensive in terms of both domain coverage and human-readable terms. Even if a concept can be described by several terms, each concept has only one meaning. Once a concept is created its meaning is persistent. Each concept has a unique identifier, which should be meaningless. The ISO specification stated that terminologies must define their purpose and scope, quantify the extent of their domain coverage, and provide mappings to external terminologies designed for classification and to support administrative functions [12].

These desiderata were particularly influential in the design of SNOMED CT. In addition, Understandable, reproducible and useful was added for SNOMED CT. Understandable means that definitions should be understandable by general clinicians, given brief explanations. Re producible means that retrieval and representation of the same item should not vary according to the nature of the interface, user preferences, or the time of entry. Usable means that we can ignore distinctions for which there is no use in healthcare.

Table 1. Requirements for Developing Terminologies

Attribute	Cimino	ISO/TC215
Purpose, Scope and Comprehensiveness	√	√
Concept Orientation	√	√
Concept Permanence	√	√
Nonsemantic Concept Identifier	√	√
Polyhierarchy	√	√
Formal Definitions	√	√
Reject "Not Elsewhere Classified"	√	√
Multiple Granularities (Support multiple levels of concept detail)	√	√
Multiple Consistent Views	√	√
Beyond Medical Concepts	√	√
Evolve Gracefully	√	√
Recognize Redundancy	√	√
Mapping to Other Vocabularies		√
System Definition		√
Appropriate Governing Organization		√
Language Independent		√

3.2. Comparative Studies of Terminologies

The international standards organization outlined attributes recognized as supporting high quality practices regarding terminology development, including that terminologies be formal aggregations of language-independent concept, that concepts should be represented by one favored term and appropriate synonymous terms, and that relationships among concepts should be explicitly represented. Other investigators formally evaluated several existing terminologies based on whether they supported various terminological attributes, including whether they could mapped to other terminologies, permitted compositionality and supported synonym [7]. Yun *et al*. suggested the requirements of data dictionary considering versatility and portability. They are concept orientation, a variety of expressions, relation, mapping to standards terminologies, mapping to local vocabularies, domain definition and version control [3]. These requirements are used as qualitative evaluation criteria of terminology. Boo developed quantitative evaluation model for biomedical terminologies in addition to qualitative criteria [13]. Suggested quantitative criteria by her to compare terminologies having different purpose are the number of concept instance, the number of concept type, the average and variance of the number of concepts per concept type, the number of relation type and relation instance, the average and variance of relation degree per instance, the average and variance of depth of Is-A. Table 1 shows desired attributes for a terminology. SNOMED CT and KOSTOM were evaluated based on the qualitative requirements and quantitative criteria as terminologies made to be used in EMR in this paper.

4. Comparative Analysis

4.1. General Characteristics of Subjects

SNOMED CT is the most comprehensive, multilingual clinical healthcare terminology in the world. It is used in electronic health record systems for clinical documentation and reporting. The SNOMED CT provides vocabularies of the overall area which is used to the EMR to code the means and to help recording the clinical data. It can be used to retrieve, and analyze clinical data. It is adopted and being used in 27 countries as their national standard and used in more than fifty countries. It is a resource with comprehensive, scientifically validated clinical content and enables consistent, processable representation of clinical content in electronic health records. It is mapped to other international standards. SNOMED CT has a broad coverage of health related topics. It can be used to describe a patient's medical history, the details of an orthopedic procedure, the spread of epidemics, and much more. At the same time, the terminology has an unmatched depth, which enables clinicians to record data at the appropriate level of granularity [10]. The core components of SNOMED CT are concept, description and relation. The core components are concept, description and relationship. The structure of SNOMED CT has the 19 top level of the hierarchy to include the whole area used to EMR. The concept has different hierarchy, attributes and relationships and enables logical definition and searching for meaning. It can be used as interface terminology since this contains a wealth of synonyms and various expressions even if it is reference terminology. It supports pre-coordination and post-coordination and cross-mapped with other terminologies. SNOMED CT expressions are usually presented using a notation known as compositional grammar [14]. This is a very well made terminology and used in many countries.

In this study, we used the contents that the status is active and concept type is disorder in the version in January 2016. Clinical findings in SNOMED CT represent the result of clinical observations, assessments, or judgments and include both normal and abnormal clinical states. This is covers a very broad range of concepts, with a similar range as the HL7 Observation [1]. The disorder sub-hierarchy of clinical findings covers abnormal clinical states only. The concept which is disorder is easy to be extracted because semantic tag of fully specified name is marked as disorder.

KOSTOM is designed to manage all the terms that are used in the domestic medical field comprehensively. It consists of 8 subsets tables, they are diagnosis, procedures, radiology, dentistry, public health, nursing and others and a repository for anatomical illustrations. The concepts are defined in the format "English-Korean pairs". Each concept is mapped onto different standard terminologies such as UMLS, KCD7 and ICD-9-CM. We used the KOSTOM 2.0 version in this study [15].

SNOMED CT has 415,184 concepts, 1,242,012 terms, 41 domains and 65 relationship types. One concept is expressed 2.99 terms in average. The number of concepts in KOSTOM is 184,844, term is 230,584 and domain is 8. And KOSTOM has only is-a relationships.

Table 2. The General Characteristics for the Contents of SNOMED CT and KOSTOM

Heading level	Number of Concepts	Number of Terms	Average of Synonyms	Concept Type (Domain)	Relationship Type	concept structure
SNOMED CT	415,184	1,242,012	2.99	41	65	formal definition
KOSTOM	184,844	230,584	1.25	8	1	English-Korean pair

4.2. Comparison by Evaluation Criteria of Terminology

KOSTOM was satisfied in the concept orientation, concept permanence, Nonsemantic concept Identifier and mapping to standards. Clinical diagnostic area of KOSTOM consisted of terms used in the medical field and the actual terms of the ICD-10. So KOSTOM which takes over as the classification of the ICD-10 has a mutually exclusive hierarchy. The meaning of concepts can be defined only by synonyms, but it is impossible to define the descriptive and formal definition. It contains the content of the ICD 10 as it is, NEC exist. The representation of the context in the diagnostic area is not possible. This lack of content for a variety of expressions has no relationship between concepts. Is-a relationship just only exist according to the categorization of the ICD. Version control is possible in KOSTOM. But the term code that existed in the version1 is removed in version2. The final version cannot show the code inactive in the previous version. These results show that this did not satisfy both the requirements of CMV and reference terminology. It can only support the roles of metathesaurus that has synonyms and mapping to ICD classification.

SNOMED CT is equipped to meet all the requirements of the terminology listed above. And also is understandable, reproducible and usable. SNOMED CT is a multinational, multilingual terminology. It has a built-in framework to manage different languages and dialects. The International Release includes a set of language independent concepts and relationships. Today, SNOMED CT is available in US English, UK English, Spanish, Danish and Swedish. Partial translations into Canadian French, Lithuanian, and several other languages are currently taking place, and further language translations are being planned by IHTSDO Members. But SNOMED CT does not currently support the languages of a small number of population-countries, such as Korean language. In addition, because of the different medical environment and preferred term, it is difficult to apply in the field of Korea.

4.3. Structure of Contents

SNOMED CT is composed of components, which include concepts, relationships, descriptions, Subsets and cross maps. The core component types in SNOMED CT are concepts, descriptions and relationships. Every component is identified by a SNOMED CT Identifier. Figure 1 shows SNOMED CT logical model [14]. The logical model therefore specifies a structured representation of the concepts used to represent clinical meanings, the descriptions used to refer to these, and the relationships between the concepts. A relationship represents an association between two concepts. Relationships are used to logically define the meaning of a concept in a way that can be processed by a computer. For example, we can express that diabetes mellitus type2 is a diabetes mellitus and its finding site is structure of endocrine system using some relationships. Figure 1 is a SNOMED CT logical model.

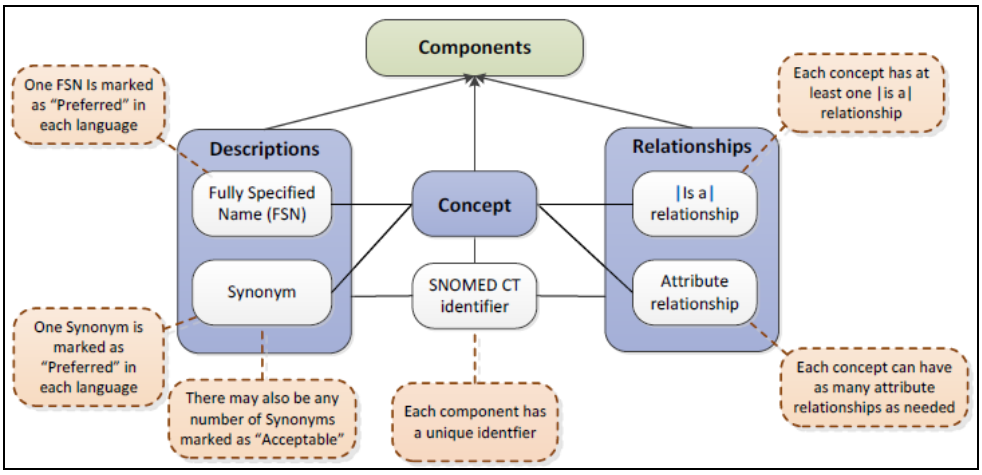

Figure 1. The SNOMED CT Logical Model

The concept is composed of concept id, concept status, fully specified name, tag, CTV3 id, SNOMED id and Is Primitive. In 1999, the English National Health Service and the College of American Pathologists agreed to merge SNOMED RT with the Read codes Version 3 that known as the NHS Clinical Terms Version 3 (CTV3), to produce a single joint clinical terminology, SNOMED CT. Every Read code and previous SNOMED code ever released was incorporated into SNOMED CT so that migration would not result in loss of information [1]. The description is description id, description status, concept id, term Unicode, initial capital status, description type, and language code. The description type can be specified with various terms that describe a concept as fully specified name, preferred, synonym, unspecified. The FSN is a term used to describe the concept fully. Relationship shows the relationship between the concepts and is composed of relationship id, concept id1, relationship type, concept id2, characteristic type, refinability and relationship group [Figure 2].

• Concepts

Concept ID	Concept Status	FULLY SPECIFIED NAME	Tag	CTV3 ID	SNOMED ID	IS PRIMITIVE
102452004	0	Abscess of thyroid (disorder)		C0502	DB-80612	0

• Descriptions

Description ID	Description STATUS	Concept ID	Term	Initial Capital Status	Description Type	Language Code
165687010	0	102452004	Abscess of thyroid	0	1	en
541214011	0	102452004	Abscess of thyroid (disorder)	0	3	en

• Relationships

Relationship ID	Concept ID1	Relationship TYPE	Concept ID2	Characteristic Type	Refinability	Relationship Group
1293244027	102452004	246112005	272141005	1	2	0
1293245026	102452004	246456000	288526004	1	2	0
2528981022	102452004	116680003	363171009	0	0	0

Figure 2. Structure of Contents in SNOMED CT

Diagnosis domain of the KOSTOM has a term table and an optional item table. The concepts and terms are not separated in KOSTOM. It consists of one term table. Mapping

table not exist but mapping codes are included in term table. Figure 3 shows an illustration of components of KOSTOM.

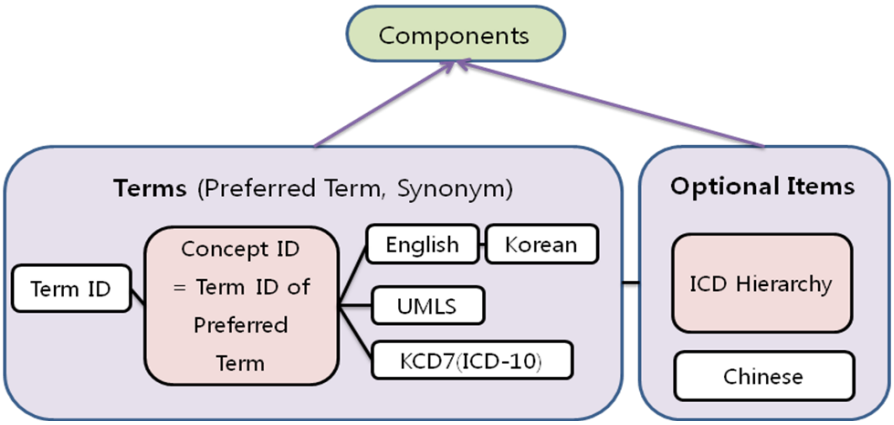

Figure 3. Illustration of Components of KOSTOM

The term table is composed of a term code, concept code, English, Korean, UMLS concept code, version and KCD code equal to ICD [Figure 4]. The term table is available in an Excel file. We can see that some terms are the same term code and concept code.

• Diagnosis

Term ID	Concept ID	English	Korean	UMLS	Ver.	KCD
H00011460	H00011460	Abscess of thyroid	갑상선의농양	C0342174	[2.0]	E06.0
H00880629	H00880629	Intra-abdominal hernia	복부내탈장	C0178282	[2.0]	K46|K46.0|K46.1
H01769799	H00880629	Unspecified abdominal hernia	상세불명의복부탈장	C0178282	[2.0]	K46|K46.0|K46.1
H02492131	H00880629	Unspecified abdominal hernia	상세불명의복벽탈장	C0178282	[2.0]	K46
H02492866	H00880629	Abdominal hernia NOS	복벽탈장NOS	C0178282	[2.0]	K46.9
H00003580	H00003580	Abdominal pregnancy	복강임신	C0032984	[2.0]	O00.0|P01.4
H00003595	H00003580	Abdominal pregnancy	복부임신	C0032984	[2.0]	O00.0|P01.4
H00003234	H00003580	Abdominal gestation	복강임신	C0032984	[2.0]	O00.0|P01.4

Figure 4. Structure of Contents in KOSTOM

One of characteristics in KOSTOM is that the term code is used as the concept code. The 7 concepts contrary to the principle that the preferred term code should be used as concept code has been found. A single term has not synonymous in itself is a representative term and concept. However, term code is different from concept code in the 7 single terms. The preferred term code change is not possible in this structure even if preferred term that used very often because the language is changed as times goes by is changed in the different representation.

KOSTOM guarantee the permanence of the concept, but it is impossible to change the preferred terms. The second characteristic is composed of an English-Korean pair. The English-Korean pair structure is useful, but the other hand, results in a duplication of the code, the term that occurs in one of the English word is translated into different word. In other words, it does not meet the requirements for language independence. Because there is no distinguish in active and inactive in this contents, we should confirm inactive terms and concepts in every version. In KOSTOM there are separate detailed table in addition to the term table, which may specify whether to use a hierarchical structure of ICD and the

interface terms or not. The detailed table is configured in terms codes, subcategory, category, clinical usage and Chinese character. The concept relationship of the KOSTOM may be connected to the "is-a" hierarchy relation of ICD.

SNOMED CT is able to support pre-coordination and post-coordination because it has the compositional grammar but, KOSTOM is not. In fact KOSTOM is a form of classification that a content is added only to the ICD classification system and KOSTOM code assigned to the similar concept. If it includes a lot of terms used in the field, it will be enough useful, but if not, a manual mapping should be done with the building an interface terminology in every hospital. Since we cannot fully cover clinical terminologies that used in the hospital with a reference terminology, we have to do those efforts. These problems can be reduced if it includes many diagnostic term often used in clinical.

4.4. Quantitative Evaluation

There are scope of coverage, complexity ratio as concreteness of meaning, and granularity, semantic connectivity and expressiveness of semantics for quantitative evaluation. SNOMED CT concepts which marks a disorder in FSN of the entire contents are 79,252, the concepts whose status of 79,252 are active (current, limited, pending move) are 69,277. The terms that can express these concepts are 209,149. The concept of the SNOMED CT is expressed as at least more than 2 terms, because it is including the FSN. Therefore, it was expressed in 2 terms in minimum and up to 35 in maximum. Except for the FSN, the case that one of the concepts is expressed as a term is 63% (49,708), that expressed in two is 25% (19,718), that is represented as three is 6% (4,900), more than four is 6% (4926). Diffuse spasm of esophagus has 34 terms. One concept is expressed in 3.0 terms in average. The number of diagnostic domain- concepts in KOSTOM is 58,607, the content-status in that version which is noticed is considered as active without differentiating.

The number of terms is 81,402 and one concept is expressed in one in minimum and 17 terms in maximum. The case that the number of terms of concept is 1 is 70% (41,225), two is 25% (14,721), three is 2.3% (1,345) and more than 4 is 2% (1,316).

One concept is expressed 1.4 terms in average. This is lacking diversity of terminology compared to SNOMED CT. Four concepts are expressed in 17 terms which are Deuteranomaly, Multiple delivery, Anaphylactic reaction and Epilepsy. They are the same Korean expression, but regard the spacing of the English letters which is pair, plurals and NOS *etc*. as synonymous. That is, this shows that more synonyms than needed are listed in specific terms. Disorder allows the attributes. The number of relation type of in the disorder domain is 19, whose relationships are 505,387. There are Is a, Finding site, Episodicity, Clinical course, Severity, Associated morphology, Causative agent, WAS A, Pathological process, Occurrence, Has definitional manifestation, After, Due to, Associated with, SAME AS, Interprets, Finding informer, Finding method and Has interpretation. All of concepts have more than one is-a relationship at least. Table 2 shows the results by quantitative criteria on component of the two terminologies.

Table 2. Results by Quantitative Criteria on Component

Criteria	SNOME CT	KOSTOM
The number of concepts	69,277	58,607
The number of terms	209,149	81,402
The average of number of terms per concept	3.0	1.4
The number of concept having one term (%)	49,708(63%)	41,225(70%)
The number of concept having two terms (%)	19,718(25%)	14,721(25%)

The number of concept having more than 3 terms (%)	9,826(12%)	2,661(4%)
Maximum of the number of descriptions per concept	34	17
The Number of relation types	19	1

4.5. Mapping to Classifications

The SNOMED CT to ICD-10 Map is included in the SNOMED CT International release as a Baseline. The SNOMED CT to ICD-10 Map was created to support the epidemiological, statistical and administrative reporting needs of IHTSDO member countries and WHO Collaborating Centers. The SNOMED CT to ICD-10 Map is released as Reference sets. The map is a directed set of relationships from SNOMED CT source concepts to ICD-10 target classification codes. The SNOMED CT source domains for the MAP are limited to subtypes of clinical finding, event and situation with explicit context. The target classification codes are ICD-10 2010 release. SNOMED CT provides map rules.

Contents of KOSTOM are mapped to KCD7 namely Korea version of the ICD-10. UMLS concept code is assigned to them in order to ensure the compatibility at home and abroad. There is no mapping table separately, it has the KCD7 and UMLS code as a single attribute value each in the diagnostic term table. KCD code is assigned to 49,639 concepts that is the 85% of the total concepts. If KCD code can be assigned to a certain term, UMLS code can be assigned to this term also since UMLS has all of contents of ICD and KCD. However, in KOSTOM, UMLS concept code is assigned to only 21,456 concepts that is corresponding to 36% of 58,607.

For each mapping, even though they are the same concept, different ICD-10 codes or many other codes are mapped to it. In the case of abdominal pregnancy, there are two codes in one concept, namely the code "O00.0" granted to the mother and "P01.4" granted to the Newborn [Figure 4]. Some terms have a blocks of codes within ICD-10. For example, The ICD-10 mapping cord of acute upper respiratory infections is “J00-J06”. This mapping in the state that does not support the formal definition suggests that a single concept can be interpreted in two ways.

5. Discussion

SNOMED has a long history but KOSTOM has not. By the comparison SNOMED CT with KOSTOM, I suggest the following for developing national standard terminology of diagnosis. First of all coverage of domain specific content is the most important. The importance of vocabulary content cannot be over stressed. Even if KOSTOM includes diagnostic terms which are used in the medical field, the number of terms is very small that expresses a single concept in the state of including all ICD-10 terminology.

Second problem is no formal definition or descriptive definition for expressing explicitly the meaning of concepts. The current level of KOSTOM is just a starting level of thesaurus that has one more synonyms. The meaning of the concepts should be acknowledged by preferred term or synonyms, but human being as well as computer cannot understand the meaning of concepts exactly in the situation that the 70% of the concepts is expressed in a single term. So the hierarchy and the relationship definition among concepts is need. The aim of KOSTOM should be the sharing clinical data among health care providers and it can be utilized easily. So the meaning of each concept is able to be understood exactly through formal definition.

Third, the ambiguity of concept codes of UMLS is being raised steadily as a comprehensive vocabulary including more than 140 nomenclatures. Therefore, we should consider the mapping validity on UMLS. UMLS Mapping Errors as well as the problems of UMLS itself were also found many. Forth, the language independence should be ensured in the structure of the content. We should consider the profit and loss gained by

English-Korean pair structure. English-Korean pair structure results in a duplication of the term and increase instance. One English name can be expressed as the number of Korean, one Korean may be expressed as a number of English also. Actually, Deuteranopia can be expressed to 6 English synonyms and to 5 Korean, but 17 term codes were made in KOSTOM because of the structure of English-Korean pair. This is an irrational and heavy wasteful structure.

Fifth, as preferred terms can be changed according to time and habits, we should consider problems occurred by equality of concept code and preferred term code.

If you apply a national standard terminology to the hospital information system and it cannot show a certain merits, the national standard terminology cannot be activated in Korean environment in which a private hospital leads to patient care. Above all, we should improve national terminology accurately operated in hospital information systems in order to support the exchange of clinical data, clinical decision support, domain and so on. In addition, because the national standard terminology cannot completely cover the hospital-specific interface terminology, the existing code system and the mapping of standard terminology code should be supported.

In the case of foreign countries, the commercial CMV products with most of the functionalities that services terms needed in the electronic medical records has shown a lot. There are no this kind of commercial products in domestic and in order to activate national standard terminology the government should support the environment in which this kind of commercial CMV service can be developed. Managing Director will be supported by a national environment that can be developed in terms of these commercial services in order enabled the national standard nomenclature. These commercial products in domestic and Managing Director will be supported by a national environment that can be developed in terms of these commercial services in order enabled the national standard nomenclature. In case of applying a national standard terminology to the hospital information system, it will be difficult to activate, if does not merit the attention. In other words, it must first be improved so that it can be exactly the terminology which operates data exchanges, clinical decision support and providing domain *etc.* in the domestic hospital information system.

6. Conclusion

Here we compared and analyzed diagnosis domain of the international standard terminology and terminology system in Korea by qualitative and quantitative criteria of the terminology and a literature review and structural analysis. Application of the diagnosis standard terminology is an important basic factor in the utilization and sharing of medical information among the medical institutions. We found several problems of terminologies in this comparative study. In Korea we are discussing various measures for dissemination and activation of KOSTOM. However, studies on effectiveness and refinement of KOSTOM are very insufficient. It is useless no matter how well made where there is no means to apply it. It is necessary to apply it to hospital information system and then to analyze the effectiveness and refine it in the future. Users try to use it for themselves and should take into account the real problems.

References

[1] T. Benson, “Principles of Health Interoperability HL7 and SNOMED”, Springer-Verlag, London, **(2012)**.
[2] Medical law enforcement regulations of 14, Section, Korea, **(2014)**.
[3] J. H. Yun, M. J. Kim, S. J. Ahn, M. S. Kwak, Y. Kim and H. K. Kim, “The Development of Clinical Terminology Dictionary for Integration and Management of Clinical Terminologies in EMR Systems”, Journal Kor Soc Med Informatics., vol. 15, no. 4, **(2009)**, pp. 411-421
[4] P. Scott, P. Macisaac and P. Saad, “An introduction to health terminologies”, Brisbane: National Centre for Classification in Health, **(2002)**.

[5] S. T. Rosenbloom, M. P. H., R. A. Miller, K. B. Johnson, P. L. Elkin and S. H. Brown, "Interface Terminologies: Facilitating Direct Entry of Clinical Data into Electronic Health Record Systems", Journal Am Med Inform Assoc., vol. 13, no. 3, **(2006)**, pp. 277-288.
[6] R. A. Rocha, S. M. Huff, P. J. Haug and H. R. Warner, "Designing a controlled medical vocabulary server: the voser project", Comput Biomed Res., vol. 27, no. 6, **(1994)**, pp. 472-507.
[7] Interface Terminologies: Facilitating Direct Entry of Clinical Data into Electronic Health Record Systems, Journal of the American Medical Informatics Association., vol. 13, no. 3, **(2006)**, pp. 277-288
[8] ISO/TC 215 business plan. Version 3, **(2003)**.
[9] Unified Medical Language System, http://www.nlm.nih.gov/research/umls/
[10] SNOMED CT, http://www.ihtsdo.org/snomed-ct/
[11] J. J. Cimino, "Desiderata for Controlled Medical Vocabularies in the Twenty-First Century", Methods Inf. Med., vol. 37, no. 4-5, **(1998)**, pp. 394-403.
[12] Health informatics-Criteria for the categorization and evaluation of clinical terminologies, working draft, ISO/TC215 WG3, **(2006)**.
[13] Y. K. Boo, "A Study on Developing Quantitative Evaluation Model for Biomedical Terminologies", The Catholic University of Korea, **(2009)**.
[14] SNOMED CT Starter Guide, 2014, SNOMED CT, **(2014)**.
[15] KOSTOM guideline, social security information service, ministry health & welfare, **(2015)**.

Authors

Mijung Kim, received her M.S. in Health Information Management from Yonsei University in 2005 and completed a PhD. course in Computer Engineering from Jeju National University. She is currently a professor in the department of health administration, Cheju Halla University in Korea. She is interested in artificial intelligence, clinical decision support system and clinical terminology.

