Evaluation of SNOMED CT Content Coverage: A Systematic Literature Review

Farid KHORRAMI[a], Maryam AHMADI[a,1] and Abbas SHEIKHTAHERI[a]

a Dept. of Health Information Management, School of health management and
information sciences, Iran University of Medical Sciences, Tehran, Iran

Abstract. Background: One of the most important features studied for adoption of terminologies is content coverage. The content coverage of SNOMED CT as a large scale terminology system has been evaluated in different domains by various methods. Objectives: This study provided an overview of studies evaluating SNOMED CT content coverage. Methods: This systematic literature review covered Scopus, Embase, PubMed and Web of Science. It included studies in English language with accessible full-text from the beginning of 2002 to November 2017. Results: Reviewing 62 studies revealed that 76 percent of studies were carried out in the US and other countries started to study in this regard from 2007. Most of the studies focused on the comparison of SNOMED CT with disease classifications in the domain of "diagnosis and problem list". Conclusion: Studying the trend of studies in different countries shows that SNOMED CT content coverage is not limited to the early stages of SNOMED CT adoption. However, evaluation methods are likely different due to the stage of SNOMED CT implementation. Therefore, it is recommended to identify and compare evaluation methods of SNOMED CT content coverage in future studies.

Keywords: Systematized Nomenclature of Medicine, Terminology, Terminological systems, Content coverage, Evaluation

1. Introduction

Applying terminological systems (TS) at the time of data capture is one of the requirements for documentation in electronic information systems [1,2]. Terminologies have a wide range of applications such as user interface for knowledge resources including guidelines [3], critical pathways and reminders; decision support systems[4] and practice analysis support in quality improvement, epidemiology and outcomes analyses [1,5-7]. Capturing data based on TS allows for data reuse [5,6,8]. Hence, it is critical to 1) select and implement suitable TS and 2) evaluate it after adoption.

SNOMED CT has been used as both reference terminology [9-11] and interface terminology [12,13] in information systems as well as research studies. This terminology is known as the primary infrastructure for data interchange in EHR. SNOMED CT usages in different studies are classified into 15 categories and five main focus categories [14]. One of the focus categories is "evaluatio"" which is divided into two subgroups: 1) "Prove merit" dealing with studies evaluating the advantages of SNOMED CT in operational settings, and 2) "Retrieval or analysis of patient data" which deals with

---

[1] Corresponding Author: Maryam Ahmadi, School of Health Management and Information Sciences, Iran University of Medical Sciences, Address: No. 6, Rashid Yasemi St., Vali-e-asr Ave., 1995614111 Tehran, Iran. Tel.: +98 9171982334; fax: +98 21 88883334 , E-Mail: mahmadi24@gmail.com

studies evaluating SNOMED CT used for capturing data in patient care. In the latter case, the approach is changing from a focus on using SNOMED CT for capturing data to using "the captured data"[14]. As ASSESS CT project of EU studied the fitness of SNOMED CT to be used as a potential standard for EU-wide eHealth deployments [15]. One of the most important issues in studies of this subgroup is "content coverage" [7,16]. In fact, it is the "quintessential feature" of TS [17]. Having had a wide range of applications and considering its development in more than thirty countries [18], SNOMED CT evaluation methods for content coverage need to be identified and evaluated to assure its development in parallel with the growing implementation of TS in health care information systems. In a bigger study, we have focused on this matter. So as part of it, in this paper, we are going to provide an overview of studies evaluating SNOMED CT content coverage.

2. Methods

2.1. Search Strategy

Databases including PubMed, Scopus, Embase and Web of Science were searched without time limitation. The Mesh term "Systematized Nomenclature of Medicine" and all its entry terms were searched with "OR" operator. In addition, we searched different terms for evaluation such as Evaluat*, Assess*, analys*, analyz*, test* and audit* [19] using "OR" operator. Additionally, the term "coverage" was also searched. All of the three search strategies were conducted in titles, abstracts and keywords. Finally, the results of the above searches were combined together using "AND" operator. The research team also checked the references of the retrieved articles to find any related articles missed through the searching process.

2.2. Study selection and eligible papers

We included English papers dealing with content coverage evaluation in SNOMED CT from 2002 (the first edition of SNOMED CT) to 2017. Papers whose full-text was not accessible and papers related to previous version of SNOMED were excluded.

Study selection steps were as per PRISMA flow diagram. After importing the selected articles into the EndNote, the duplicate items were excluded. Two members of the research team (A.ST and F.K) separately screened all the titles and the abstracts of the papers. Papers were separately grouped as 1) relevant, 2) need to more review, and 3) not relevant. The two researchers discussed over the cases of disagreement in presence of the third member (M.A) of the team. In the next step, the two researchers studied the papers in groups 1 and 2 independently. Having reached consensus, the research team selected related papers meeting the objectives of the study.

2.3. Data extraction and summarizing

Required data were extracted from the selected articles based on data elements by the two researchers (A.ST and F.K). For disagreement cases, the two researchers discussed over items extracted from the articles in presence of the third researcher (M.A). Finally, they agreed on the cases based on majority vote.
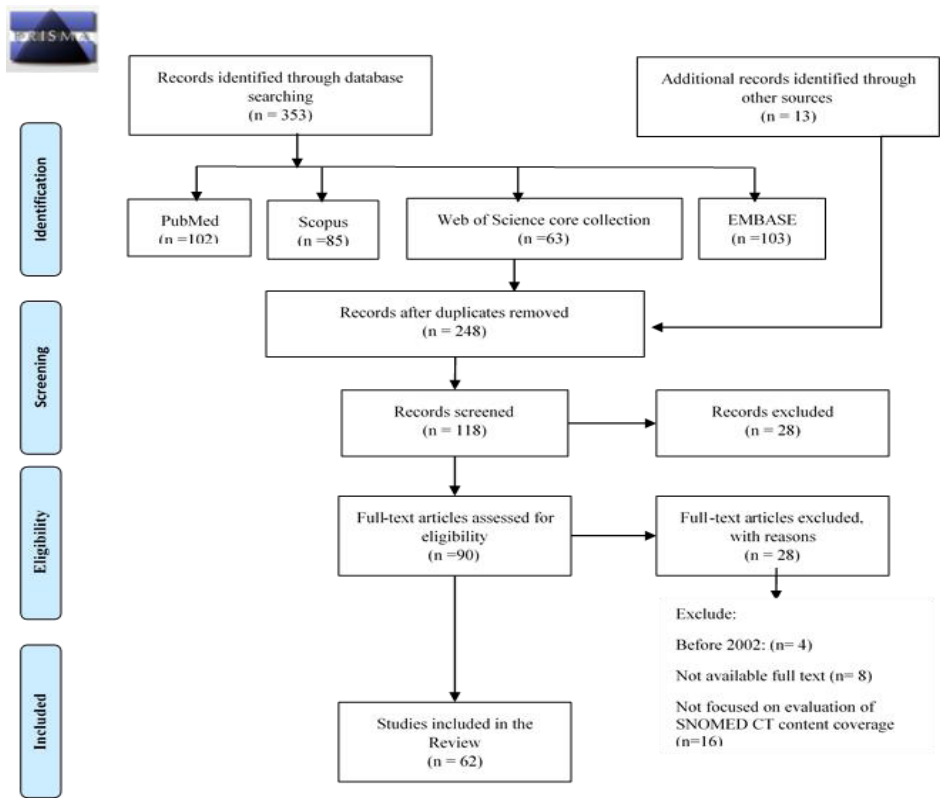
Figure 1. Details of selection of the studies based on PRISMA flow diagram.

3. Results

Out of 366 papers retrieved through the search strategy and studying the resources, we obtained 118 papers excluding duplicated items. After screening the titles and abstracts, 28 more papers were excluded from the study. Therefore, 90 papers remained for full-text study. Twenty-eight papers were also excluded from the study: 26 papers after applying the exclusion criteria and 2 papers through data extraction. Ultimately, 62 publications were selected for the final analysis. Figure 1 depicts the details of selection of the studies based on PRISMA flow diagram.
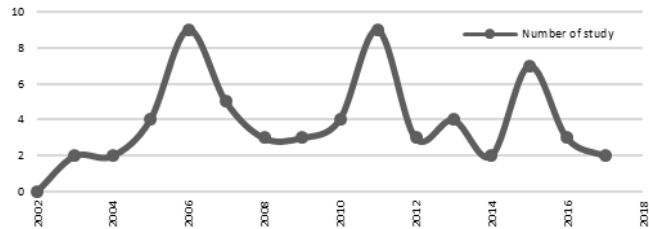

Figure 2-Trend of studies from 2002 to 2017

Table 1. The frequency and percentage of published papers by country

Country	n	%
USA	47	76%
Netherlands	4	6%
Sweden	3	5%
Korea	2	3%
Canada	1	2%
France	1	2%
Germany	1	2%
Denmark	1	2%
Argentina	1	2%
Kenya	1	2%
Total	**62**	**100%**

Figure 2 shows the trend of studies from 2002 to 2017. Most of studies were carried out in 2006 and 2011.Twenty-four percent of the selected studies were presented in conferences. AMIA Annual Symposium Proceedings and Journal of the American Medical Informatics Association had the maximum share of published articles (45 percent).

3.1. Countries

Most of the studies were carried out in high income countries. The extracted data showed that 76 percent were published in the US and 16 percent in Europe. However, the Netherlands had the highest share (50%) in Europe (Table 1).

Having in mind the first edition of SNOMED CT published in 2002 and its internationalization in 2007, Table 2 shows the trend of the studies in different years by country in three periods. Seventy-three percent of the total studies were carried out after internationalization of the system (2007), and this figure was 86 percent in other countries apart from the US.

3.2. Domain

Table 3 shows the results by the evaluated domains. The domain with maximum shares included Diagnosis and Problem List and Radiology. The focus of all studies out of the US was Diagnosis and Problem List, Nursing and Oncology.

3.3. Terminological systems

Studies show that the evaluation of SNOMED CT content coverage was carried out either by itself or in comparison with one or more terminology systems. The findings of this research are:
- Evaluating the content coverage of SNOMED CT by itself: 29% (18 cases)
- Evaluating the content coverage of SNOMED CT with another system: 34% (21 cases)
- Comparing the content coverage of several terminology systems including SNOMED CT: 37% (23 cases)

Table 2. The frequency of papers published in three different time periods by grouping of countries

Years range	Total		USA		Other countries	
	n	%	n	%	n	%
2002-2006	17	27%	15	32%	2	13%
2007-2011	24	39%	16	34%	8	53%
>2012	21	34%	16	34%	5	33%
Total	**62**		**47**		**15**	

Table 3. The frequency of domain evaluation in SNOMED CT content coverage

Domains groups	n	%
Diagnosis and Problem list	34	51%
Nursing	4	6%
Radiology	4	6%
Ophthalmology	2	3%
Medical Guidelines	2	3%
Adverse drug reactions	1	2%
Dental finding	1	2%
Oncology	4	6%
Clinical research concepts	4	6%
Allergy	2	3%
Disability	2	3%
Herbal terms	1	2%
Phonemics	1	2%
Total	**62**	

Table 4 shows the terminology systems had been studied. Most of evaluation studies were carried out against disease classification, UMLS and LOINC with 26, 24, and 10 percent, respectively.

Fourteen studies also addressed the comparison of SNOMED CT with interface terminologies which; in fact, was a combination of various terminology systems based on the user needs.

Thirty-nine percent of papers did not mention the version of SNOMED CT or UMLS clearly.

Table 4. The frequency of other terminology systems in SNOMED CT content coverage

Category title (total)	
Classification of diseases (16)	**n**
ICD-9-CM	7
ICD-10-CM	2
ICD-10	4
ICD-11	1
ICD-O-3	1
ICF	1
UMLS (12)	
Medication and Allergy terminologies (9)	
MedDRA	3
RxNorm	2
NDF-RT	2
UNII	2
LOINC (6)	
Imaging terminologies (4)	
RadLex	4
Interface terminologies(14)	
Medcin	4
Veterans Benefit Administration codes (VBA)	3
Cancer Data Standards Registry and Repository (CaDSR)	1
Categorical Health Information Structured Lexicon (CHISL)	1
Healthcare Data Dictionary (HDD)	1
Phenoslim terminology (PT)	1
Mayo clinic vocabulary	1
Diagnoses for Intensive Care Evaluation (DICE)	1
Nursing terminologies (3)	
ICNP	3
Other terminologies (4)	
CPT	2
MESH	2

4. Discussion

The aim of this review is to provide a general overview of articles addressing evaluation of SNOMED CT content coverage. The obtained results can generally be classified into four main categories: 1) Place of the studies (countries), 2) Trend of studies in the time limit of this study, 3) Terminology systems compared with SNOMED CT in studies of content coverage evaluation and 4) Domains evaluated.

Studies addressing evaluation of SNOMED CT content coverage were carried out in a few countries with high income. Seventy-six percent of the studies were in the US. The contribution of this country before 2007 (before internalization of SNOMED CT) was 88%. However, in an earlier review article, Lee et al (2014) who studied the usage of SNOMED CT from 2001 to 2012 showed that the contribution of the US was 53%[14]. Our study showed that the contribution of European countries was 16%. Fifty percent of the studies in Europe were published by the Netherlands. The significant point was that there were no studies from England and Australia in this subject. However, Lee et al (2014) reported the two countries along with France had maximum number of articles after the US [14]. The countries did not have any studies in the evaluation domain.

Studying the trend of studies carried out within the period of our study shows that content coverage evaluation has grown since 2007. Ninety-three percent of the total studies in all countries excluding the US were carried out after 2007. Since content coverage studies are included in SNOMED CT predevelopment [14], it is natural for new country members of IHTSDO to study on this domain at initial years of membership. For example, combination of some studies [3; 18-24] in the Netherlands led to the development of an evaluation framework for characterizing terminological systems[17]. It is a good example of specialized follow-up for SNOMED CT adoption in this country.

In some of studies evaluating SNOMED CT content coverage, other TS were also used. Studying compared to SNOMED CT showed that they were either carried out alone or in comparison with other TS. They include SNOMED CT content coverage evaluation compared to medical record texts or clinical cases (29%), and content coverage compared to other TS (81%). The systems for diseases classification had the highest portion (26%). It can be the weakness of classification systems in covering specific terms at the time of clinical data capturing [25]. Moreover, most of studies in this topic introduce SNOMED CT terminology systems as the best terminology system for improving content coverage in different domains [9; 13; 26-29]. Some of the studies compared SNOMED CT system with interface terminologies such as Medcin [12; 13; 30], CHISL [13] and Mayo clinic [16; 31]. Since such TS are used for systematic recording of data in most IHTSDO member and non-member countries, development of a framework for their evaluation is necessary. This issue is technically discussed in [32]. UMLS – as a terminological system and also a controlled meta-thesaurus – is known for creating interoperability between computerized systems [33]. Twenty-four percent of studies directly mentioned that they had used UMLS tools in different stages of their study. It can, therefore, be implied that UMLS tools not only affect on interoperability but also on terminological systems.

SNOMED CT covers a wide range of domains. The findings of this research shows that most of the studies were carried out in "Diagnosis and problem list"(51%). This confirms Studying [14] the evaluated domains shows that selecting them gets more specialized through the years and the studies focus on a disease advanced findings. For example, studies [34-36] concentrate on content coverage of a specific disease.

We just reviewed articles in English with accessible full-text. However, there might been studies addressing content coverage of SNOMED CT in other languages which we missed.

5. Conclusion

Our review showed that the number of studies addressing evaluation of SNOMED CT content coverage is growing and the topic is considered by different countries which most of them are high income countries. Although content coverage is basically carried out at the preliminary stages of SNOMED CT adoption, our study revealed that such studies do not stop; however, the evaluation methods may change. Therefore, the identification and comparison of SNOMED CT content coverage – which is being carried out as a future study.

Acknowledgements

6. References

[1]   D. Bronnum, A.R. Hojen, K.R. Goeg, and P.B. Elberg, Terminology-Based Recording of Clinical Data for Multiple Purposes Within Oncology, Stud Health Technol Inform 228 (2016), 267-271.
[2]   M. Dugas, F. Fritz, R. Krumm, and B. Breil, Automated UMLS-Based Comparison of Medical Forms, Plos One 8 (2013).
[3]   L. Ahmadian, N.F. De Keizer, and R. Cornet, The use of SNOMED CT for representing concepts used in preoperative guidelines, Stud Health Technol Inform 150 (2009), 658-662.
[4]   Y. Lin, C.J. Staes, D.E. Shields, V. Kandula, B.M. Welch, and K. Kawamoto, Design, Development, and Initial Evaluation of a Terminology for Clinical Decision Support and Electronic Clinical Quality Measurement, AMIA Annu Symp Proc 2015 (2015), 843-851.
[5]   C.M. Cusack, G. Hripcsak, M. Bloomrosen, S.T. Rosenbloom, C.A. Weaver, A. Wright, D.K. Vawdrey, J. Walker, and L. Mamykina, The future state of clinical data capture and documentation: a report from AMIA's 2011 Policy Meeting, J Am Med Inform Assoc 20 (2013), 134-140.
[6]   A. Randorff Hojen and K. Rosenbeck Goeg, Snomed CT implementation. Mapping guidelines facilitating reuse of data, Methods Inf Med 51 (2012), 529-538.
[7]   S.H. Brown, P.L. Elkin, E. Lumakovska, M. Rallins, T. Rosenbloom, J. Trajkovski, and M.S. Tuttle, Terminology and Terminological Systems, Springer, London, 2012.
[8]   A. Randorff Hojen, E. Sundvall, and K. Rosenbeck Goeg, Visualizing sets of SNOMED CT concepts to support consistent terminology implementation and reuse of clinical data, Stud Health Technol Inform 192 (2013), 1160.
[9]   S.H. Brown, C.S. Husser, D. Wahner-Roedler, S. Bailey, L. Nugent, K. Porter, B.A. Bauer, and P.L. Elkin, Using SNOMED CT as a reference terminology to cross map two highly pre-coordinated classification systems, Stud Health Technol Inform 129 (2007), 636-639.
[10]   K.A. Spackman, Integrating sources for a clinical reference terminology: experience linking SNOMED to LOINC and drug vocabularies, Medinfo 9 Pt 1 (1998), 600-603.
[11]   S. Schulz, J.M. Rodrigues, A. Rector, and C.G. Chute, Interface Terminologies, Reference Terminologies and Aggregation Terminologies: A Strategy for Better Integration, Stud Health Technol Inform 245 (2017), 940-944.
[12]   S.H. Brown, S.T. Rosenbloom, B.A. Bauer, D. Wahner-Roedler, D.A. Froehling, K.R. Bailey, M.J. Lincoln, D. Montella, E.M. Fielstein, and P.L. Elkin, Direct comparison of MEDCIN and SNOMED CT for representation of a general medical evaluation template, AMIA Annu Symp Proc 11 (2007), 75-79.

[13] S.T. Rosenbloom, S.H. Brown, D. Froehling, B.A. Bauer, D.L. Wahner-Roedler, W.M. Gregg, and P.L. Elkin, Using SNOMED CT to represent two interface terminologies, J Am Med Inform Assoc 16 (2009), 81-88.

[14] D. Lee, N. de Keizer, F. Lau, and R. Cornet, Literature review of SNOMED CT use, J Am Med Inform Assoc 21 (2014), e11-19.

[15] ASSESS CT - Assessing SNOMED CT for Large Scale eHealth Deployments in the EU, in, 2018.

[16] P.L. Elkin, S.H. Brown, A. Balas, Z. Temesgen, D. Wahner-Roedler, D. Froehling, M. Liebow, B. Trusko, S.T. Rosenbloom, and G. Poland, Biosurveillance evaluation of SNOMED CTś terminology (BEST Trial): Coverage of chief complaints, in: Stud Health Technol Inform, 2008, pp. 797-802.

[17] R. Cornet, N.F. de Keizer, and A. Abu-Hanna, A framework for characterizing terminological systems, Methods Inf Med 45 (2006), 253-266.

[18] R. Cornet and N. de Keizer, Forty years of SNOMED: a literature review, BMC Med Inform Decis Mak 8 Suppl 1 (2008), S2-S2.

[19] D.G.T. Arts, R. Cornet, E. De Jonge, and N.F. De Keizer, Methods for evaluation of medical terminological systems: A literature review and a case study, Methods Inf Med 44 (2005), 616-625.

[20] R. Cornet, Do SNOMED CT relationships qualify?, Stud Health Technol Inform 136 (2008), 785-790.

[21] L. Ahmadian, R. Cornet, W.A. van Klei, and N.F. de Keizer, Diversity in preoperative-assessment data collection, a literature review, Stud Health Technol Inform 136 (2008), 127-132.

[22] F. Bakhshi-Raiez, R. Cornet, and N.F. de Keizer, Development and Application of a Framework for Maintenance of Medical Terminological Systems, J Am Med Inform Assoc 15 (2008), 687-700.

[23] R. Cornet and S. Schulz, Relationship groups in SNOMED CT, Stud Health Technol Inform 150 (2009), 223-227.

[24] R. Cornet, Definitions and qualifiers in SNOMED CT, Methods Inf Med 48 (2009), 178-183.

[25] K.W. Fung and J. Xu, An exploration of the properties of the CORE problem list subset and how it facilitates the implementation of SNOMED CT, J Am Med Inform Assoc 22 (2015), 649-658.

[26] H. Wasserman and J. Wang, An applied evaluation of SNOMED CT as a clinical vocabulary for the computerized diagnosis and problem list, AMIA Annu Symp Proc (2003), 699-703.

[27] T. Sampalli, M. Shepherd, J. Duffy, and R. Fox, An evaluation of SNOMED CT (R) in the domain of complex chronic conditions, Int J Integr Care 10 (2010).

[28] P.L. Elkin, S.H. Brown, A. Balas, Z. Temesgen, D. Wahner-Roedler, D. Froehling, M. Liebow, B. Trusko, S.T. Rosenbloom, and G. Poland, Biosurveillance evaluation of SNOMED CT's terminology (BEST Trial): coverage of chief complaints, Stud Health Technol Inform 136 (2008), 797-802.

[29] D.A. Travers, S.W. Haas, D.A. Travers, and S.W. Haas, Unified Medical Language System Coverage of Emergency-medicine Chief Complaints, Acad Emerg Med 13 (2006), 1319-1323.

[30] D. Montella, S.H. Brown, P.L. Elkin, J.C. Jackson, S.T. Rosenbloom, D. Wahner-Roedler, G. Welsh, B. Cotton, O.D. Guillamondegui, H. Lew, K.H. Taber, L.A. Tupler, R. Vanderploeg, and T. Speroff, Comparison of SNOMED CT versus Medcin terminology concept coverage for mild Traumatic Brain Injury, AMIA Annu Symp Proc 2011 (2011), 969-978.

[31] P.L. Elkin, S.H. Brown, B.A. Bauer, C.S. Husser, W. Carruth, L.R. Bergstrom, and D.L. Wahner-Roedler, A controlled trial of automated classification of negation from clinical notes, BMC Med Inform Decis Mak 5 (2005).

[32] S.T. Rosenbloom, R.A. Miller, K.B. Johnson, P.L. Elkin, and S.H. Brown, A model for evaluating interface terminologies, J Am Med Inform Assoc 15 (2008), 65-76.

[33] S.T. Wu, H. Liu, D. Li, C. Tao, M.A. Musen, C.G. Chute, and N.H. Shah, Unified Medical Language System term occurrences in clinical notes: a large-scale corpus analysis, J Am Med Inform Assoc 19 (2012), e149-156.

[34] D. Brammen, H. Dewenter, K.U. Heitmann, V. Thiemann, R.W. Majeed, F. Walcher, R. Rohrig, and S. Thun, Mapping Equivalence of German Emergency Department Medical Record Concepts with SNOMED CT After Implementation with HL7 CDA, Stud Health Technol Inform 243 (2017), 175-179.

[35] M.S. Park and Z. He, Consumers' Use of UMLS Concepts on Social Media: Diabetes-Related Textual Data Analysis in Blog and Social Q&A Sites, JMIR Med Inform 4 (2016), e41.

[36] T. Oluoch, N. de Keizer, P. Langat, I. Alaska, K. Ochieng, N. Okeyo, D. Kwaro, and R. Cornet, A structured approach to recording AIDS-defining illnesses in Kenya: A SNOMED CT based solution, J Biomed Inform 56 (2015), 387-394.