Literature review of SNOMED CT use

Dennis Lee,[1] Nicolette de Keizer,[2] Francis Lau,[1] Ronald Cornet[2,3]

▶ Additional material is published online only. To view these files please visit the journal online (http://dx.doi.org/10.1136/amiajnl-2013-001636).

[1]School of Health Information Science, University of Victoria, Victoria, British Columbia, Canada
[2]Department of Medical Informatics, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands
[3]Department of Biomedical Engineering, Linköping University, Linköping, Sweden

Correspondence to
Dennis Lee, School of Health Information Science, University of Victoria, PO Box 3050 STN CSC, Victoria, BC, Canada V8W 3P5; dlhk@uvic.ca

Received 9 January 2013
Revised 10 June 2013
Accepted 19 June 2013
Published Online First 4 July 2013

To cite: Lee D, de Keizer N, Lau F, *et al*. *J Am Med Inform Assoc* 2014;**21**: e11–e19.

ABSTRACT

Objective The aim of this paper is to report on the use of the systematised nomenclature of medicine clinical terms (SNOMED CT) by providing an overview of published papers.

Methods Published papers on SNOMED CT between 2001 and 2012 were identified using PubMed and Embase databases using the keywords 'systematised nomenclature of medicine' and 'SNOMED CT'. For each paper the following characteristics were retrieved: SNOMED CT focus category (ie, indeterminate, theoretical, pre-development/design, implementation and evaluation/commodity), usage category (eg, prospective content coverage, used to classify or code in a study), medical domain and country.

Results Our search strategy identified 488 papers. A comparison between the papers published between 2001–6 and 2007–12 showed an increase in every SNOMED CT focus category. The number of papers classified as 'theoretical' increased from 46 to 78, 'pre-development/design' increased from 61 to 173 and 'implementation' increased from 10 to 34. Papers classified as 'evaluation/commodity' only started to appear from 2010.

Conclusions The majority of studies focused on 'theoretical' and 'pre-development/design'. This is still encouraging as SNOMED CT is being harmonized with other standardized terminologies and is being evaluated to determine the content coverage of local terms, which is usually one of the first steps towards adoption. Most implementations are not published in the scientific literature, requiring a look beyond the scientific literature to gain insights into SNOMED CT implementations.

INTRODUCTION

The use of free text and local terms in electronic medical records is widespread and is a source of poor data quality and a barrier to semantic interoperability, data mining, secondary use of data and computerized clinical decision support.[1] The systematised nomenclature of medicine clinical terms (SNOMED CT) is an international clinical reference terminology that has the potential to improve data quality and patient safety, and facilitate semantic interoperability by capturing clinical data in a standardized, unambiguous and granular manner.

January 2013 marked the 11th year since SNOMED CT was first released. Since January 2002, 22 new versions, released semi-annually, have been circulated. The International Health Terminology Standards Development Organisation (IHTSDO) was established 6 years ago to coordinate the maintenance and promotion of SNOMED CT as a clinical reference terminology, and 19 countries have designated SNOMED CT as the preferred clinical reference terminology for use in electronic medical records.

In this study, our objective was to investigate the use of SNOMED CT by providing an overview of published studies. Whereas the 40-year SNOMED literature review by Cornet *et al*,[2] in 2008 focused on papers published between 1966 and 2006 using any version of SNOMED, this study focused only on SNOMED CT papers published between 2001 and 2012.

METHODS

Identifying papers

Searches using PubMed (http://www.ncbi.nlm.nih.gov/pubmed) and Embase (http://www.embase.com) were performed using the terms 'SNOMED' and 'systematised nomenclature of medicine' between 2001 and 2012. Although SNOMED CT was first released in 2002, we presumed there were papers that discussed the upcoming release of SNOMED CT published in 2001. Only papers that were written in English or had an English abstract were included in this study. The search strategy is available in supplementary appendix A (available online only).

Classification criteria

We used a set of classification criteria similar to that used in the 40-year review,[2] with the addition of one new criterion, the SNOMED CT focus category. A summary of the classification criteria is available in table 1.

SNOMED CT focus category

We identified five SNOMED CT focus categories: indeterminate, theoretical, pre-development/design, implementation, and evaluation/commodity. 'Indeterminate' refers to SNOMED CT being used as an example of a terminology system without any further detail on its use or implementation, is referenced in a letter by a reader, editor or author, or is included in a survey or review. 'Theoretical' refers to SNOMED CT being discussed as a terminology system but not used in conjunction with a clinical project/study. There are likely to be no outcomes but rather descriptive work on the development of SNOMED CT or envisioned outcomes. The next three focus categories address the application of SNOMED CT. 'Pre-development/design' refers to SNOMED CT being assessed to determine if it fulfills requirements and whether it is feasible to be used in a full-scale implementation as a terminology standard. 'Implementation' refers to SNOMED CT being used in a study, pilot project or operational setting. 'Evaluation/commodity' refers to SNOMED CT being evaluated to determine the effects of the implementation and demonstrate its value (eg, how it can enhance the quality of care) or is used in an operational setting where

Table 1 Criteria used to classify SNOMED CT-related papers

No	Criteria	Definition
1.	SNOMED CT focus category	Refers to the focus of the paper (ie, indeterminate, theoretical, pre-development/design, implementation, evaluation/commodity)
2.	Usage category	Refers to how SNOMED CT is primarily used. Each usage category belongs exclusively to one focus category. Refer to table 2 for the list of usage categories and their definitions
3.	Medical domain	Refers to the medical domain of the paper
4.	Country	Refers to the country in which the study took place, if available or the country of the first author. If the study spanned multiple countries, the paper was classified as 'multiple'

SNOMED CT, systematised nomenclature of medicine clinical terms.

the focus has moved from capturing data to using the data captured in routine patient care.

Usage category

The usage category refers to the primary purpose for using SNOMED CT. The 14 usage categories from the 40-year review[2] were re-examined and several categories were created, renamed and merged. Categories were created and renamed to reflect new ways in which SNOMED CT was being used and to clarify the categories. The main reason for merging the categories was due to low frequency counts. In the 40-year review,[2] there were five categories in which one paper was assigned to a category. For example, 'to prove merit in terms of costs' and 'to prove merit in terms of quality of care' were merged into 'prove merit'. Each of the 15 usage categories was linked to one and only one of the five SNOMED CT focus categories (see table 2).

As a paper could span multiple usage categories, we used the most prominent usage category in classifying the paper. For example, a paper[3] that described the comparison of a problem list with SNOMED CT or the annotation of clinical narratives with SNOMED CT to determine the content coverage was classified as 'retrospective content coverage'. If the concepts identified were used in a study (research or non-operational setting), for example, to calculate the prevalence of a disease, that paper[4] was classified as 'used to classify or code in a study'. If the setting was an operational setting in which the concepts identified were stored in actual patient records and used for patient care, that paper[5] was classified as 'implementation of SNOMED CT'.

Classifying method

A web-based application was developed that cataloged the abstracts and papers, and enabled the co-authors to classify the papers independently. Functions were also available for the authors to compare their results with each other, add comments and review the results of papers from the 40-year review. The abstracts were used to classify a paper and the full paper was referred to if details needed to classify the paper were not evident in the abstract.

To ensure interrater reliability, 10 papers were selected and classified individually by the co-authors. The results were compared and discussed until a consensus was reached on the differences and definitions on classification categories were refined.

Table 2 List of usage categories and definition, and corresponding focus category

No	Usage category	Status	Definition	SNOMED CT focus category
1.	Other	New	Includes letters submitted to journals and reports on the results of surveys, literature reviews and systematic reviews	Indeterminate
2.	As an example	Same	References SNOMED CT briefly as a standard terminology or that it is used in a study with few additional details	
3.	Illustrate terminology systems theory	Same	Describes terminology systems theory such as frameworks for describing terminologies and potential benefits of using standardized terminologies	Theoretical
4.	Description of SNOMED CT and other standards	New	Describes SNOMED CT and other terminologies including technical aspects (eg, hierarchy) and non-technical aspects (eg, potential benefits and challenges)	
5.	Terminology auditing	Renamed	Reports on auditing methods that have been applied to SNOMED CT to detect errors	
6.	Compare to or map to other terminology systems	Same	SNOMED CT is compared to other standardized terminology systems mainly in terms of content coverage	Pre-development/design
7.	Translation	New	Describes the needs for translating SNOMED CT into other languages or the progress and results of translation studies	
8.	Prospective content coverage	Same	SNOMED CT is compared to non-standardized terminology systems such as local interface terminologies for content coverage	
9.	Prospective interrater agreement	New	Similar to prospective content coverage, but the focus is on comparing the results of between two or more coders	
10.	Planned standard for electronic health records	Same	SNOMED CT is planned for use in an EHR but the focus is on the overall EHR infrastructure and not on SNOMED CT	
11.	Design considerations	Same	Describes implementation considerations such as the use of search algorithms and version management	
12.	Used to classify or code in a study	Same	SNOMED CT is used only for a study and not in a routine setting	Implementation
13.	Implementation of SNOMED CT	Same	SNOMED CT is implemented in a pilot or operational setting	
14.	Prove merit	Merged	Studies that demonstrate the benefits of using SNOMED CT in operational settings	Evaluation/commodity
15.	Retrieve or analyse patient data	Same	SNOMED CT has been in used in routine patient care and the focus has moved from capturing data with SNOMED CT to using the data captured	

Status refers to the comparison with the usage categories in the 40-year review and indicates whether the usage category is new, is the same, was renamed or was merged.
EHR, electronic health record; SNOMED CT, systematised nomenclature of medicine clinical terms.

The authors then worked in pairs to classify an additional 30 papers to ensure there was an agreement on how the criteria were to be assigned to a paper. Additional discussions took place to resolve any ambiguity, and when all differences in classification were reconciled, the first author proceeded to classify the rest of the papers. Twenty-five papers were flagged by the first author when the usage category was uncertain. These papers were reviewed by the other authors and discussions took place to reconcile the classification.

RESULTS

The searches on PubMed (n=537) and Embase (n=594) resulted in 702 unique papers (see figure 1). Two hundred and fourteen (30%) papers were excluded because the version of SNOMED was not clinical terms (n=127, 18%), the paper made no mention of SNOMED CT (n=55, 8%), an English abstract was not available for a foreign language paper (n=21, 3%), and an abstract or full paper could not be located (n=9, 1%). In all, 488 unique papers were reviewed. The list and classification of the 488 papers are available in supplementary appendix B (available online only), while a summary of the papers classified as 'pre-development/design', 'implementation' and 'evaluation/commodity' is available in supplementary appendix C (available online only).

SNOMED CT focus

The results of the classification of papers by SNOMED CT focus category and by year are shown in figure 2. The number of papers classified as 'theoretical' has remained relatively the same at between 11 and 15 papers over the past 8 years. A comparison of the papers published from 2001 to 2006, and papers published from 2007 to 2012 showed an increase in every SNOMED CT focus category. The number of papers classified as 'theoretical' increased from 46 to 78, 'pre-development/design' increased from 61 to 173, and 'implementation' increased from 10 to 34. Papers classified as 'evaluation/commodity' only started to appear in 2010.

Usage category

The results by usage category are shown in figure 1. A further breakdown of the usage categories by subcategories is shown in table 3. In this section we describe the most common usage category for each SNOMED CT focus category except for 'indeterminate'.

Theoretical: compare to or map to other terminology systems (n=74)

SNOMED CT was compared to or mapped to at least 40 standardized terminologies. The exact number is unknown as not all papers listed all the terminologies used, and therefore we are uncertain of the number of unique terminologies compared. The most common terminologies SNOMED CT was compared or mapped to were the International Classification of Diseases, both the 9th and 10th revisions (n=15), International Classification of Nursing Practice (n=6) and the Medical Dictionary for Regulatory Activities (n=5). SNOMED CT was also compared to the unified medical language system (UMLS) metathesaurus directly (n=6) and indirectly (n=12). The direct comparisons occurred when a terminology system was mapped to SNOMED CT and other terminology systems including the UMLS metathesaurus. The indirect comparisons occurred when the UMLS metathesaurus was primarily used to look up mappings to other terminologies. While 'compare to or map to other terminology systems' was the most common usage

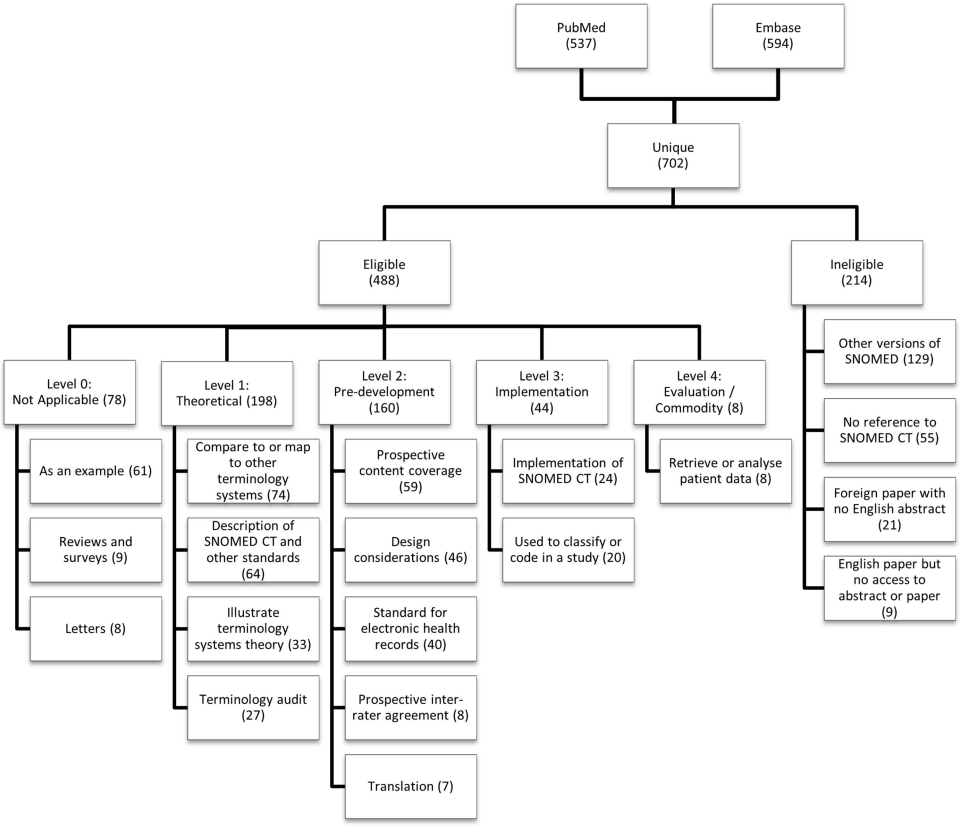

Figure 1 Overview of scoring of papers.

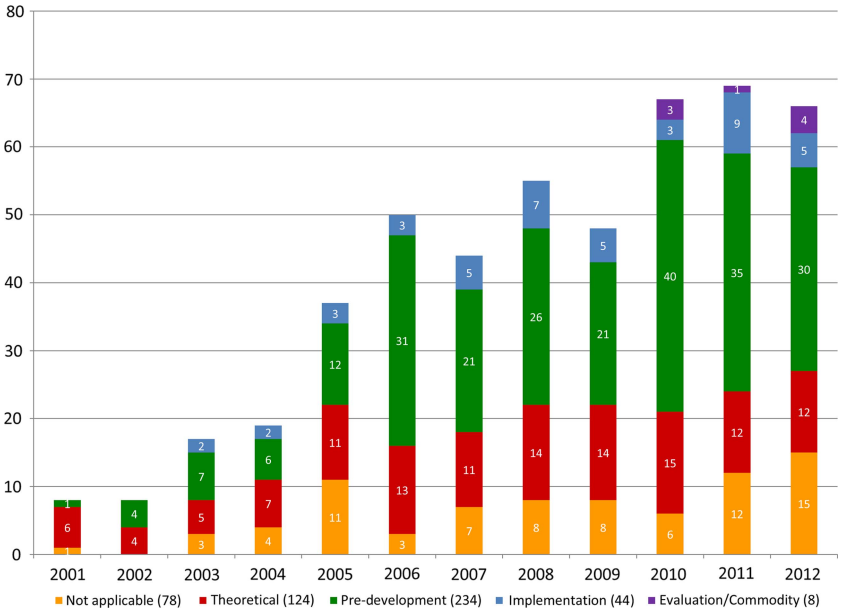

Figure 2 Number of papers by maturity level and year.

category in this focus category, the new usage category 'terminology audit' included 24 papers, 20 of which were published in the past 6 years.

Pre-development/design: prospective content coverage (n=59)
SNOMED CT was used in 59 studies to determine the degree to which SNOMED CT could provide content coverage for local terms. The content coverage included comparing SNOMED CT against larger enterprise interface terminologies and data dictionaries such as the Vanderbilt EHR interface terminology[6] and Mayo mastersheet index,[1] as well as to smaller sets of terms in the domains of problem lists and diagnosis (n=7), care planning and guidelines (n=6) and nursing (n=4). Content coverage was usually assessed using exact matches, partial matches, no

Table 3 Number of papers by subcategories

No	Usage category and subcategory	Number
1.	As an example—no subcategories	61
2.	Other—letters to editor (n=3), reply from authors (n=2), literature reviews (n=5), surveys (n=4)	17
3.	Illustrate terminology systems theory—terminology theory and ontological principles (n=14), semantic similarity (n=8), frameworks and models for categorizing terminology systems (n=6), need for mapping (n=5)	33
4.	Description of SNOMED CT and other standards—general description of SNOMED CT (n=35), development process and milestones of SNOMED CT (n=7), changes, improvements and advancement of SNOMED CT (n=7), use of definitions and qualifiers (n=5), use of relationship groups (n=4), use of description logic (n=3), potential benefits of SNOMED CT (n=3)	64
5.	Terminology auditing—abstraction network (n=8), ontological principles (n=4), lexical/linguistic (n=5), combination of methods (n=2), other methods with frequency of one each (n=8)	27
6.	Compare to or map to other terminology systems—39 other standardized terminology systems, most common were the International Classification of Diseases, 9th and 10th Revisions, (n=17) and International Classification for Nursing Practice (n=6). SNOMED CT was also compared to the UMLS directly (n=6) and indirectly through the UMLS metathesaurus (n=12)	74
7.	Translation—languages included French (n=5), Swedish (n=1) and Chinese (n=1)	7
8.	Prospective content coverage—interface terminologies, data dictionaries and medical corpora (n=7), chief complaints/problem lists (n=6), care planning and guidelines (n=6), newborn disorders (n=3), drugs (n=3), nursing (n=4), cardiovascular disorders (n=2), complex chronic conditions (n=2), ophthalmology (n=2), reason for visit/chief complaint for emergency department (n=2), pathology diagnoses (n=2), allergies (n=2) and others with frequency of one (n=21)	59
9.	Prospective inter-rater agreement—number of reviews were two (n=1), three (n=6) and 10 (n=1)	8
10.	Standard for electronic health records—electronic health records frameworks/infrastructure and integration with information models (n=24), binding to clinical models, templates or archetypes (n=14)	40
11.	Design considerations—search and retrieval algorithms (n=18), general implementation challenges (n=8), process and challenges related to the development of subsets (n=8), version control, management and migration (n=5), the role and use of interface terminologies in conjunction with SNOMED CT to facilitate data capture (n=3), encoding methodologies or comparison of coding techniques (n=3)	46
12.	Used to classify or code in a study—identifying and extracting mainly from free text narratives and reports, general medical conditions (n=6), cancer characteristics (n=4), emergency room (n=2), pneumonia and influenza cases (n=3), medications and drug concerns (n=2), intensive care (n=1), pathology (n=1) and negation (n=1).	20
13.	Implementation of SNOMED CT—terminology servers and services to support data entry (n=10), use of data entry templates (n=10), use of search boxes and auto-complete (n=3), use of natural language processing (n=1)	24
14.	Prove merit—no subcategories	0
15.	Retrieve or analyse patient data—use of SNOMED CT synonyms against free text (n=2), indexed free text with SNOMED CT concepts using natural language processing and queried indexed concepts (n=4), unclear if synonyms or concepts were used (n=1), subject matter experts encoded queries (n=1)	8

SNOMED CT, systematised nomenclature of medicine clinical terms; UMLS, unified medical language system.

matches, and matches using post-coordination. Exact or complete matches were as high as 90% in areas such as the representation of disorders of newborn infants[7] and as low as 19% in areas such as aesthetic ophthalmic plastic surgery.[8] Post-coordination was required in over 40% of domains such as cardiovascular diseases, computed tomography procedures, and clinical phenotype data.

Implementation: implementation of SNOMED CT (n=24)

This usage category can be further divided into the development of SNOMED CT terminology servers and services to support data entry (n=10) and the implementation of SNOMED CT in clinical settings in both pilot projects and operational settings (n=14). The terminology servers and services included visual exploration of terminologies and specialized search algorithms to navigate the hierarchy and retrieve relevant concepts for data entry (n=6), search for publications using SNOMED CT concepts (n=1), search for healthcare providers using consumer terms mapped to SNOMED CT and clinician expertise (n=1). Two other papers listed the features of their own terminology servers (n=1) and that of vendors (n=1).

The user interfaces in which SNOMED CT was implemented can be further classified into three categories. First, items in checklists, questionnaires and data entry templates were mapped to SNOMED CT. In those cases, the options in the forms were fixed and did not require users to search for SNOMED CT descriptions directly (n=8). Local terms were presented to users in the form of pick lists and radio buttons while the data were recorded in the background with SNOMED CT. Domains included cancer,[9–12] pressure ulcer wounds,[13] radiology,[14] obesity,[15] and family planning.[16] Second, search boxes and autocomplete fields were used to display results based on user input (n=5). SNOMED CT subsets were developed based on historical patient records so as to constrain the concepts used in the results rather than search against the entire SNOMED CT content. Domains included drugs,[17] veterinary,[18] intensive care,[19] ambulatory care[20] and general patient records.[21] Third, natural language processing algorithms were used to locate potentially relevant SNOMED CT concepts from clinical narratives (n=1). Clinicians were shown the candidate concepts for review before the concepts were indexed to the patient record.[5]

Evaluation/commodity: retrieve and analyze patient records (n=8)

Two papers used SNOMED CT to identify synonyms for neuromuscular blockade[22] and *Clostridium difficile* infections[23] as keywords for searching against clinical narratives. Four papers used natural language processing to index clinical narratives with SNOMED CT concepts followed by a query against those concepts. The queries were for cancer,[24] infectious symptoms,[25] and diabetes mellitus, cardiovascular diseases, asthma and congestive obstructive pulmonary disease,[26] and 54 diseases such as esophageal reflux and HIV.[27] In addition to just querying for the index concepts, the index concepts' children in the SNOMED CT hierarchies were included in search queries although the value of querying for children concepts was not reported. One paper used SNOMED CT to identify occurrences of melanoma,[28] but it was unclear whether synonym or concept matching of melanoma was used. In one paper, subject matter experts encoded 10 queries (eg, patients who had acute myocardial infarction and were on aspirin), which were then executed against a SNOMED CT-encoded patient database. Searches using SNOMED CT concepts were also shown to have better precision than keyword searches.[27]

Medical domain

The papers spanned 36 medical domains and specialties. Problem list/diagnoses, nursing, drugs and pathology were the most common medical domains. The medical domains and specialties that occurred in at least 10 papers are shown in figure 3. Nursing primarily consisted of studies looking at the coverage of local nursing terms as well as standardized nursing terminologies such as International Classification for Nursing Practice.

Country

The papers were from 22 countries with over half the papers coming from the USA (n=238, 53%) (see table 4 for the full list of countries). SNOMED CT-related papers originated from 10 of the 19 countries that are members of the IHTSDO while affiliates and non-member countries of the IHTSDO accounted for the other 13.

The number of countries that have published SNOMED CT-related papers has steadily grown over the years, with the

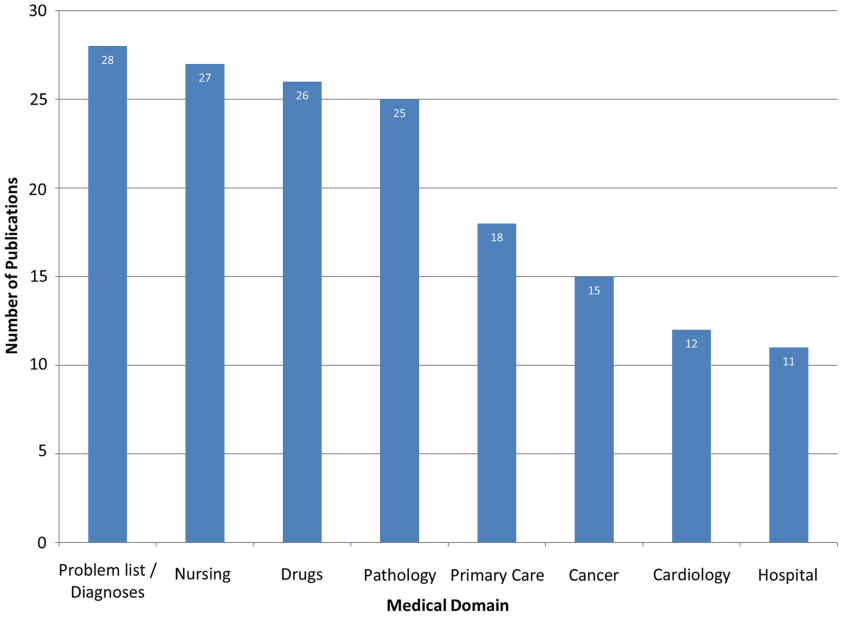

Figure 3 Number of publications found for each medical domain.

Table 4 Countries that belong to the IHTSDO or have published SNOMED CT-related papers in the scientific literature

Country	Joined IHTSDO?	First publish	No of papers	Country	Joined IHTSDO?	First publish	No of papers
Argentina		2007	4	Lithuania	2007		
Australia	2007	2005	26	Malta	2011		
Austria		2011	2	New Zealand	2007		
Belgium		2011	1	Poland	2011		
Brazil		2009	1	Singapore	2008	2011	2
Canada	2007	2009	15	Slovak Republic	2009		
China		2009	3	Slovenia	2010		
Cyprus	2009			South Korea		2008	10
Czech Republic		2010	2	Spain	2009	2008	14
Denmark	2007	2006	7	Sweden	2007	2006	12
Estonia	2010			Switzerland		2008	2
France		2002	26	The Netherlands	2007	2005	21
Germany		2005	19	UK	2007	2001	34
Hungary		2008	1	USA	2007	2001	255
Iceland	2011	2011	2	Multiple/unknown			28
Israel	2012			Total			488
Italy		2012	1				

IHTSDO, International Health Terminology Standards Development Organisation; SNOMED CT, systematised nomenclature of medicine clinical terms.

biggest increases coming in 2007–8 (see figure 4). Over the past 5 years, papers were coming from 14 to 16 countries per year.

DISCUSSION

In this study, we searched for SNOMED CT-related papers in PubMed and Embase and classified the papers by SNOMED CT focus category, usage category, medical domain and country. Over the past 6 years there has been an increase in the number of SNOMED CT-related studies centering on implementation and evaluation. Thirty-seven of the 44 papers classified as 'implementation' were published over the past 6 years, and all eight papers classified as 'evaluation/commodity' were published within the past 3 years. Nevertheless, the majority of the papers were classified as 'pre-development/design', which means SNOMED CT was mainly used in non-operational settings. The proportion of studies by focus category over the past 6 years, with the exception of 'evaluation/commodity', has remained roughly the same.

Theoretical

While the number of papers classified as 'theoretical' has been steady over the past 8 years and range between 11 and 15 papers each year, one usage category within this focus category has seen a steady increase. 'Terminology audit', in which auditing methods such as the abstraction network and ontological principles have been developed and used to check SNOMED CT for consistency, has been steadily increasing since 2005. As SNOMED CT undergoes significant changes with each new release version[29 30] we expect that these auditing methodologies will play a larger role in ensuring that SNOMED CT is consistent.

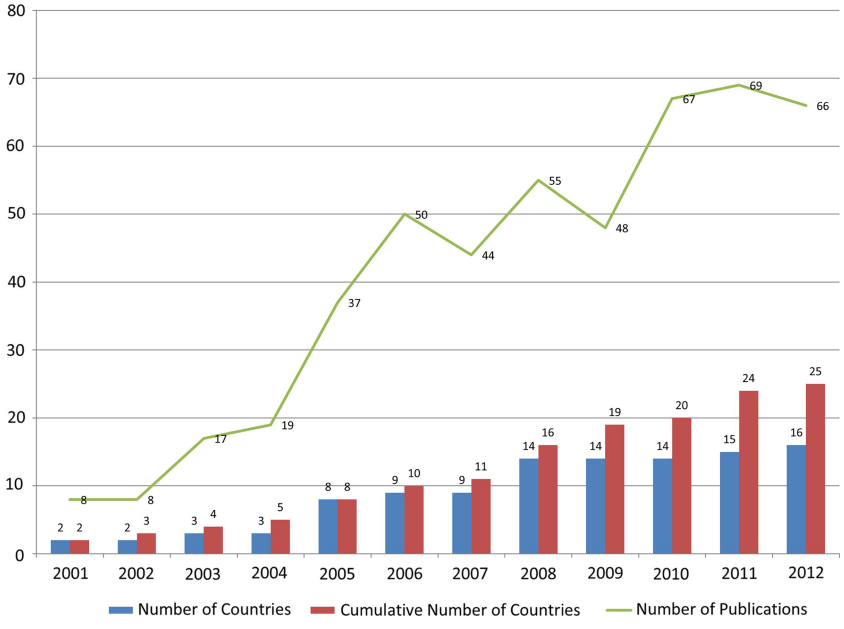

Figure 4 Number of papers per year by new countries, number of countries, cumulative countries and total papers.

Pre-development/design

The use of free text is one of the barriers to computerized clinical decision support and data re-use. However, fragmented and large numbers of standardized terminologies with partial and overlapping domain coverage is also a barrier.[31] The large number of studies involved in comparing and mapping SNOMED CT to other standardized terminologies is encouraging as individuals and organizations are recognizing the need for harmonization. For example, nursing terminologies were one of the most frequently used terminologies that were compared to or mapped to SNOMED CT. Gaps in concept and synonym coverage identified in those studies can help to improve the completeness of nursing terms in SNOMED CT.[32 33] After the usage categories of 'description of SNOMED CT' and 'compare to or map to other terminology systems', the third highest usage category was 'prospective coverage'. In this category, SNOMED CT was evaluated to determine the content coverage of local terms. The high number of studies in this area is also encouraging because determining the content coverage was usually one of the first steps in the implementation studies identified in this study. The use of post-coordination in content coverage studies also indicates that while SNOMED CT may not include every pre-coordinated concept to represent a local term, it is possible to create semantically equivalent terms. As the crafting of post-coordinated expressions is more complex than just using pre-coordinated concepts, potential implementers will require additional training.

Implementation

The number of studies classified as 'implementation' has more that tripled from 10 during the first 6 years when SNOMED CT was released to 34 over the past 6 years. Although SNOMED CT is reportedly used in over 50 countries and the number of studies classified as 'implementation' has been steadily increasing, there are still few papers that describe how SNOMED CT is being used in operational settings. Excluding the development of terminology servers and services, which are important and provide generic search and browsing capabilities, we encountered 14 studies of SNOMED CT in operational clinical settings and pilot projects. The sophistication of SNOMED CT implementations for data capture varied widely. Data entry ranged from mapping terms in data entry forms, templates and checklists to SNOMED CT in the background when users were only shown terms they were previously using, to the development of an interface terminology in which users were exposed to over thousands of descriptions and used auto-complete functionality to retrieve relevant terms, to the automatic indexing of clinical narratives using natural language processing techniques.

Evaluation/commodity

We were only able to identify studies in the 'retrieve and analyze patient data' usage category. Data retrieval functionality ranged from very rudimentary use, such as the use of synonyms to search clinical narratives, to complex queries, such as the use of subsumption and querying against post-coordinated expressions. Unfortunately, the value of using subsumption queries was not reported.

Success factors for implementing SNOMED CT included the development and use of tools that enabled SNOMED CT to be searched effectively and efficiently,[34] usability and ease of use of clinical applications,[19] the constraining of relevant concepts to create subsets in applicable domains,[19] the incorporating of terms familiar to clinicians, and collaboration among clinical users and technical developers.[20] Challenges included the management of subsets and extensions,[19] the development of intuitive interfaces and ensuring the relevancy of search results.[20] Benefits, both realized and anticipated, included improved quality of documentation,[16] improved efficiency and consistency of encoding,[5] improved patient safety,[17] reduced time and costs for transcribing, post-coding and quality management,[5 16 35] ability to conduct biosurveillance monitoring,[36] ability to audit patient records,[26] support patient case queries,[5] support integration with clinical practice guidelines,[17] enable international benchmarking,[35] and facilitate decision support systems.[13 21]

We did not encounter any studies that described the value of SNOMED CT in terms of improved outcomes. The three systems that developed decision support capabilities for detecting adverse drug events,[21] managing wounds[13] and obesity[15] did not report on patient outcomes. While improved data standardization and the potential for conducting data analysis and reporting were frequently cited as benefits, these benefits have not been quantified and we have not found any studies that demonstrate the value of SNOMED CT from a clinical perspective in an operational setting (as opposed to a study). We suggest three reasons. First, a large proportion of the studies have been on prospective coverage, therefore organizations are still in the process of gauging the feasibility of adopting SNOMED CT. Second, organizations that have implemented SNOMED CT have been focusing on data capture and therefore have not reached the stage of using the captured data. In a separate survey we conducted, we found that most organizations that have implemented SNOMED CT have been focused on the implementation and have not had the time or resources to conduct full-scale evaluations.[37] Third, we compared the papers in this study with two implementation inventories and found only five of the 23 implementations included in either or both of the IHTSDO implementation special group implementation webinars (http://www.ihtsdo.org/events/conference-presentations/conference-archive/implementation-experience) and Canada Health Infoway's SNOMED CT in use website (https://sc.infoway-inforoute.ca/standards-collaborative/snomed-ctr/snomed-ct-in-use) have been published in the scientific literature.

It is unclear why 49 papers were retrieved when the search term 'SNOMED' or 'systematized nomenclature of medicine' was used but neither the abstract nor paper made any reference to SNOMED. For example, 'bioinformatics and biological reality'[38] was retrieved via PubMed and Embase but neither the medical subject headings (MeSH) terms, abstract nor paper contained any references to SNOMED. In another example, 'in defense of the desiderata'[39] included 'systematised nomenclature of medicine*' as one of the MeSH terms but the paper did not mention SNOMED. On the other hand, there are known SNOMED CT papers that are cataloged within PubMed and Embase that were not retrieved using those keywords. For example, the literature review, 'A review of auditing methods applied to the content of controlled biomedical terminologies', by Zhu *et al*,[40] which cataloged the types of auditing methods applied to SNOMED CT (and other terminologies) was not retrieved using the keywords. To check the completeness of our search results, we compared the search results for papers published by the *Journal of American Medical Informatics Association* (*JAMIA*) using *JAMIA*'s website and PubMed. The results are available in supplementary appendix D (available online only). Refer to the appendix for the search strategy and full results. PubMed produced 27 results while *JAMIA* produced 24 results when searching in the title and abstract, and 167 results when searching the full text. A comparison of the 27 and

24 papers by PubMed and *JAMIA* showed that 23 papers overlapped. The one paper that was not retrieved by PubMed was a letter response from the authors.[41] It should be noted that the letter was retrieved using Embase. The 143 difference between the search in the title and abstract versus the full text was usually the result of SNOMED CT being briefly mentioned as an example of a terminology system or the title in one of the references. Therefore, while it is possible that our search strategy missed some papers, it is unlikely to have missed substantial numbers.

Limitations

We only reviewed papers cataloged in PubMed and Embase and only included papers that were published in English or had an English abstract. Our review of two inventories of SNOMED CT use and the papers included in our study showed that the majority of implementations are not published in the scientific literature or are not captured in PubMed or Embase. Therefore, a limitation of this study is that it includes a publication bias. A second limitation is that the majority of the papers were reviewed only by the first author. To ensure consensus in the classification of the papers, 40 (9%) papers were reviewed by at least two authors to ensure a high level of agreement on how to assign the usage categories. In addition, 25 (6%) papers that the first author flagged were reviewed by a second author.

CONCLUSION

Our literature review of 488 SNOMED CT-related papers showed that the majority of studies focused on theoretical and pre-development/design. This is still encouraging as work is being done to harmonize SNOMED CT with other standardized terminologies, and SNOMED CT is being evaluated to determine the content coverage of local terms, which is usually one of the first steps towards adopting SNOMED CT. The number of implementation studies has increased steadily although not many are in operational settings. We found that most implementations are not published in the scientific literature; therefore, a look beyond the scientific literature is needed to gain insights into SNOMED CT implementations.

Contributors All authors reviewed and classified papers. DL drafted the initial manuscript, which was edited by the other authors. All authors participated in reviewing the comments by the associate editor and reviewers and contributed to addressing the concerns raised. The final version was approved by all authors.

Competing interests None.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Elkin PL, Trusko BE, Koppel R, *et al*. Secondary use of clinical data. *Stud Health Technol Inform* 2010;155:14–29. http://www.ncbi.nlm.nih.gov/pubmed/20543306.
2 Cornet R, de Keizer N. Forty years of SNOMED: a literature review. *BMC Med Inform Decis Mak* 2008;8 (Suppl. 1):S2. http://www.ncbi.nlm.nih.gov/pubmed/19007439.
3 Richesson RL, Andrews JE, Krischer JP. Use of SNOMED CT to represent clinical research data: a semantic characterization of data items on case report forms in vasculitis research. *J Am Med Inform Assoc* 2006;13:536–46. http://www.ncbi.nlm.nih.gov/pubmed/16799121.
4 Long W. Extracting diagnoses from discharge summaries. *AMIA Annu Symp Proc* 2005:470–4. http://www.ncbi.nlm.nih.gov/pubmed/16779084.
5 Ryan A, Patrick J, Herkes R. Introduction of enhancement technologies into the intensive care service, Royal Prince Alfred Hospital, Sydney. *HIM J* 2008;37:40–5. http://www.ncbi.nlm.nih.gov/pubmed/18245864.
6 Wade G, Rosenbloom ST. Experiences mapping a legacy interface terminology to SNOMED CT. *BMC Med Inform Decis Mak* 2008;8 (Suppl. 1):S3. http://www.ncbi.nlm.nih.gov/pubmed/19007440.
7 James AG, Spackman KA. Representation of disorders of the newborn infant by SNOMED CT. *Stud Health Technol Inform* 2008;136:833–8. http://www.ncbi.nlm.nih.gov/pubmed/18487835.
8 Lee S, Tsirbas A, Goldberg RA, *et al*. Standardized terminology for aesthetic ophthalmic plastic surgery. *Ophthal Plast Reconstr Surg* 2006;22:371–4. http://www.ncbi.nlm.nih.gov/pubmed/16985422.
9 Van Berkum MM. SNOMED CT encoded cancer protocols. *AMIA Annu Symp Proc* 2003:1039. http://www.ncbi.nlm.nih.gov/pubmed/14728542.
10 Sherman S, Shats O, Fleissner E, *et al*. Multicenter breast cancer collaborative registry. *Cancer Inform* 2011;10:217–26. http://www.ncbi.nlm.nih.gov/pubmed/21918596.
11 Sherman S, Shats O, Ketcham MA, *et al*. PCCR: pancreatic cancer collaborative registry. *Cancer Inform* 2011;10:83–91. http://www.ncbi.nlm.nih.gov/pubmed/21552494.
12 Lusky K. Pilot points way to speedier cancer surveillance. *CAP Today* 2005;19:5–6, 8. http://www.ncbi.nlm.nih.gov/pubmed/15787106.
13 Kim HY, Park HA. Development and evaluation of data entry templates based on the entity-attribute-value model for clinical decision support of pressure ulcer wound management. *Int J Med Inform* 2012;81:485–92. http://www.ncbi.nlm.nih.gov/pubmed/22079242.
14 Robinson TJ, DuVall SL, Wiggins RH III. Creation and storage of standards-based pre-scanning patient questionnaires in PACS as DICOM objects. *J Digit Imaging* 2011;24:823–7. http://www.ncbi.nlm.nih.gov/pubmed/20976611.
15 Lee NJ, Bakken S. Development of a prototype personal digital assistant-decision support system for the management of adult obesity. *Int J Med Inform* 2007;76 (Suppl. 2):S281–92. http://www.ncbi.nlm.nih.gov/pubmed/17606400.
16 Zetterberg C, Ahlzén K, Ericsson E, *et al*. An example of a multi-professional process-oriented structured documentation bound to SNOMED CT. *Stud Health Technol Inform* 2012;180:1215–17. http://www.ncbi.nlm.nih.gov/pubmed/22874405.
17 Farfán Sedano FJ, Terrón Cuadrado M, García Rebolledo EM, *et al*. Implementation of SNOMED CT to the medicines database of a general hospital. *Stud Health Technol Inform* 2009;148:123–30. http://www.ncbi.nlm.nih.gov/pubmed/19745242.
18 Zaninelli M, Campagnoli A, Reyes M, *et al*. The O3-Vet project: integration of a standard nomenclature of clinical terms in a veterinary electronic medical record for veterinary hospitals. *Comput Methods Programs Biomed* 2012;108:760–72. http://www.ncbi.nlm.nih.gov/pubmed/22595264.
19 Bakhshi-Raiez F, de Keizer NF, Cornet R, *et al*. A usability evaluation of a SNOMED CT based compositional interface terminology for intensive care. *Int J Med Inform* 2012;81:351–62. http://www.ncbi.nlm.nih.gov/pubmed/22030036.
20 Liu J, Lane K, Lo E, *et al*. Addressing SNOMED CT implementation challenges through multi-disciplinary collaboration. *Stud Health Technol Inform* 2010;160:981–5. http://www.ncbi.nlm.nih.gov/pubmed/20841830.
21 Cao F, Sun X, Wang X, *et al*. Ontology-based knowledge management for personalized adverse drug events detection. *Stud Health Technol Inform* 2011;169:699–703. http://www.ncbi.nlm.nih.gov/pubmed/21893837.
22 Arnot-Smith J, Smith AF. Patient safety incidents involving neuromuscular blockade: analysis of the UK National Reporting and Learning System data from 2006 to 2008. *Anaesthesia* 2010;65:1106–13. http://www.ncbi.nlm.nih.gov/pubmed/20840604.
23 Benoit SR, McDonald LC, English R, *et al*. Automated surveillance of *Clostridium difficile* infections using BioSense. *Infect Control Hosp Epidemiol* 2011;32:26–33. http://www.ncbi.nlm.nih.gov/pubmed/21128815.
24 Nguyen A, Moore J, Zuccon G, *et al*. Classification of pathology reports for cancer registry notifications. *Stud Health Technol Inform* 2012;178:150–6. http://www.ncbi.nlm.nih.gov/pubmed/22797034.
25 Matheny ME, Fitzhenry F, Speroff T, *et al*. Detection of infectious symptoms from VA emergency department and primary care clinical documentation. *Int J Med Inform* 2012;81:143–56. http://www.ncbi.nlm.nih.gov/pubmed/22244191.
26 Liaw ST, Chen HY, Maneze D, *et al*. Health reform: is routinely collected electronic information fit for purpose? *Emerg Med Australas* 2012;24:57–63. http://www.ncbi.nlm.nih.gov/pubmed/22313561.
27 Koopman B, Bruza P, Sitbon L, *et al*. Towards semantic search and inference in electronic medical records: an approach using concept-based information retrieval. *Australas Med J* 2012;5:482–8. http://www.ncbi.nlm.nih.gov/pubmed/23115582.
28 Hussain F, Muller F, Husain E. Under-reporting of invasive malignant melanomas in North East of Scotland. *Br J Dermatol* 2010;163:67.
29 Spackman KA. Rates of change in a large clinical terminology: three years experience with SNOMED Clinical Terms. *AMIA Annu Symp Proc* 2005:714–18. http://www.ncbi.nlm.nih.gov/pubmed/16779133.
30 Wade G, Rosenbloom ST. The impact of SNOMED CT revisions on a mapped interface terminology: terminology development and implementation issues. *J Biomed Inform* 2009;42:490–3. http://www.ncbi.nlm.nih.gov/pubmed/19285570.
31 Ingenerf J, Reiner J, Seik B. Standardized terminological services enabling semantic interoperability between distributed and heterogeneous systems. *Int J Med Inform* 2001;64:223–40. http://www.ncbi.nlm.nih.gov/pubmed/11734388.
32 Hardiker NR, Casey A, Coenen A, *et al*. Mutual enhancement of diverse terminologies. *AMIA Annu Symp Proc* 2006:319–23. http://www.ncbi.nlm.nih.gov/pubmed/17238355.

33 Park HA, Lundberg CB, Coenen A, *et al*. Evaluation of the content coverage of SNOMED-CT to represent ICNP version 1 catalogues. *Stud Health Technol Inform* 2009;146:303–7. http://www.ncbi.nlm.nih.gov/pubmed/9592854.

34 Richesson R, Young K, Guillette H, *et al*. Standard terminology on demand: facilitating distributed and real-time use of SNOMED CT during the clinical research process. *AMIA Annu Symp Proc* 2006:1076. http://www.ncbi.nlm.nih.gov/pubmed/17238695.

35 Tvede I, Bredegaard K, Andersen JS. Quality improvements based on detailed and precise terminology. *Stud Health Technol Inform* 2010;155:71–7. http://www.ncbi.nlm.nih.gov/pubmed/20543312.

36 Elkin PL, Froehling D, Wahner-Roedler D, *et al*. NLP-based identification of pneumonia cases from free-text radiological reports. *AMIA Annu Symp Proc* 2008:172–6. http://www.ncbi.nlm.nih.gov/pubmed/18998791.

37 Lee D, Cornet R, Lau F, *et al*. A survey of SNOMED CT implementations. *J Biomed Inform* 2013; 46:87–96.

38 Johansson I. Bioinformatics and biological reality. *J Biomed Inform* 2006;39:274–87. http://www.ncbi.nlm.nih.gov/pubmed/16198638.

39 Cimino JJ. In defense of the desiderata. *J Biomed Inform* 2006;39:299–306. http://www.ncbi.nlm.nih.gov/pubmed/16386470.

40 Zhu X, Fan JW, Baorto DM, *et al*. A review of auditing methods applied to the content of controlled biomedical terminologies. *J Biomed Inform* 2009;42:413–25. http://www.ncbi.nlm.nih.gov/pubmed/19285571.

41 Wilcke JR, Green JM, Spackman KA, *et al*. Concerning SNOMED-CT content for public health case reports. *J Am Med Inform Assoc.* 2010;17:613; author reply 613–4. http://www.ncbi.nlm.nih.gov/pubmed/20842802.