Use of Automated SNOMED CT Clinical Coding in Clinical Decision Support Systems for Preventive Care

Save to myBoK

by Bader Al-Hablani, MS, RHIA, CHPS

Abstract

Objective: The objective of this study is to discuss and analyze the use of automated SNOMED CT clinical coding in clinical decision support systems (CDSSs) for preventive care. The central question that this study seeks to answer is whether the utilization of SNOMED CT in CDSSs can improve preventive care.

Methods: PubMed, Google Scholar, and Cochrane Library were searched for articles published in English between 2001 and 2012 on SNOMED CT, CDSS, and preventive care.

Outcome Measures: Outcome measures were the sensitivity or specificity of SNOMED CT coded data and the positive predictive value or negative predictive value of SNOMED CT coded data. Additionally, we documented the publication year, research question, study design, results, and conclusions of these studies.

Results: The reviewed studies suggested that SNOMED CT successfully represents clinical terms and negated clinical terms.

Conclusions: The use of SNOMED CT in CDSS can be considered to provide an answer to the problem of medical errors as well as for preventive care in general. Enhancement of the modifiers and synonyms found in SNOMED CT will be necessary to improve the expected outcome of the integration of SNOMED CT with CDSS. Moreover, the application of the tree-augmented naïve (TAN) Bayesian network method can be considered the best technique to search SNOMED CT data and, consequently, to help improve preventive health services.

Key terms: clinical decision support system (CDSS); preventive care; Systematized Nomenclature of Medicine–Clinical Terms (SNOMED CT); clinical terminology; medical errors

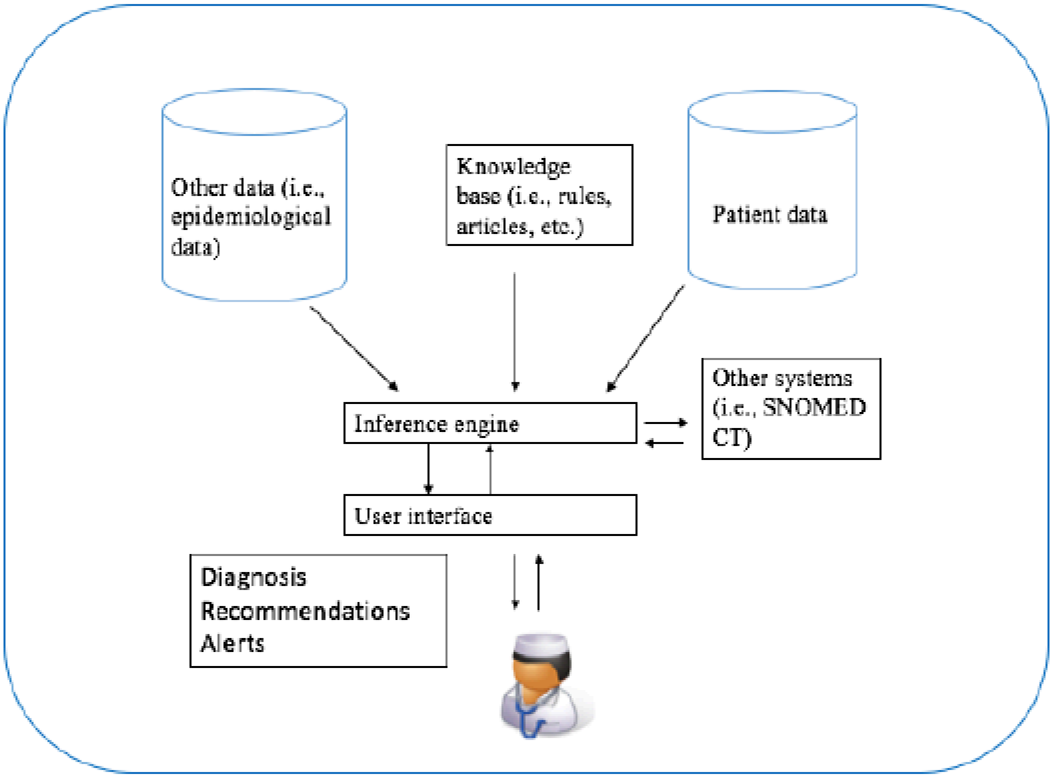
Introduction

Systematized Nomenclature of Medicine–Clinical Terms (SNOMED CT) is a clinical reference terminology standard that allows for semantic interoperability and gives meaning to raw medical data.[1] SNOMED CT was originally developed by the College of American Pathologists and was the result of merging of two systems: SNOMED-RT and Clinical Terms version 3.[2] In April 2007, the International Health Terminology Standards Development Organization (IHTSDO) acquired the intellectual property rights to SNOMED CT and started to be responsible for its maintenance and distribution.[3] SNOMED CT is the most comprehensive and multilingual clinical terminology; it functions as a taxonomy for concepts such as signs and symptoms and includes approximately 357,000 concepts.[4]

Clinical decision support systems (CDSSs) are a significant concept in the field of biomedical informatics. A CDSS is defined as any computer system that is developed to aid healthcare practitioners in making medical decisions. It is an expert system that deals with clinical data and information about patients or “with the knowledge of medicine necessary to interpret such data.”[5]

The minimum essential technical architecture for CDSS consists of a communication engine to retrieve heterogeneous data, a health information database (patient database), an inference engine to “interpret and filter patient data and knowledge,” a knowledge base, and a vocabulary engine to attain semantic interoperability. Healthcare practitioners receive reminders and alerts from CDSSs through different methods, such as clinical portals, e-mails, and paging systems.[6] Figure 1 illustrates the main elements of a CDSS, demonstrating the interaction and flow of data and diagnoses in the medical decision support system.[7]

Figure 1: Core Components of a Clinical Decision Support System



Source: Adapted from Tsui, R. "Core Components in CDSS: Clinical Decision Support Systems." Lecture, University of Pittsburgh, Department of Biomedical Informatics, September 26, 2012.

A major issue that this study aims to address is whether use of SNOMED CT in CDSSs can improve preventive health services. Preventive care is crucial in this era of highly expensive and harmful medical errors. According to a report published by the Institute of Medicine in 1999, medical errors led to the deaths of between 44,000 and 98,000 people in the United States annually.[8] Moreover, a study conducted by HealthGrades found that in-hospital medical mistakes were potentially preventable but led to approximately 195,000 deaths between 2000 and 2002.[9] Between 2003 and 2005, about 1.16 million incidents related to patient safety occurred in more than 40 million hospital admissions for Medicare patients.[10] These patient safety events led to $8.6 billion of unnecessary costs during the same years.[11] The use of SNOMED CT in CDSS is one likely answer to the costly problem of medical errors. It is critical to evaluate if SNOMED CT can represent clinical concepts and, as a result, be used in CDSSs for preventive care. The results of this research will ultimately guide healthcare professionals in the advancement of SNOMED CT and CDSSs.[12]

Methodology

The search strategy aimed to identify articles related to SNOMED CT, CDSSs, and preventive care. The databases accessed and searched were PubMed, Google Scholar, and Cochrane Library. Reference lists were examined to find relevant articles, and researchers in the field of health information management (HIM), clinical coding, and biomedical informatics were contacted to identify additional research studies and reports. The literature review include studies published in English from January 1, 2001, to December 31, 2012. PubMed was searched using the following Medical Subject Headings (MeSH) and search terms: *clinical terms, SNOMED*; *Clinical Decision Support Systems*; and *preventive health services*. For Google Scholar and Cochrane Library, the following search was used: *SNOMED CT AND Clinical Decision Support System AND preventive care*. Outcome measures were the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of SNOMED CT coded data. Data extraction was conducted, and year of publication, research question, study design, results, and conclusions were documented. The articles were selected on the basis of the following criteria: (1) the researchers measured the effectiveness and ability of SNOMED CT to enhance CDSSs, which in turn would advance preventive care, and (2) the authors investigated the sensitivity and specificity of SNOMED CT in representing clinical concepts.

Review of Literature

The ability of a clinical terminology or a classification system to represent and encode clinical concepts can be considered one of the important factors that would lead to better interoperability between electronic health records and CDSSs, not to mention other important applications. The granularity of SNOMED CT is critical in the ability of this clinical terminology to represent clinical concepts within problem lists, for example. Evaluation of this knowledge representation can be done with different analysis methods. Sensitivity and specificity rates as well as PPV and NPV can be used to measure the reliability of SNOMED CT in representing clinical concepts.

Sensitivity is the percentage that includes "all true cases correctly identified."[13] It can be calculated with the following equation: True Positive / (True Positive + False Negative). PPV is the "probability that a positive is a true positive given a specified probability threshold for the variable of interest."[14] Specificity is "the percentage of all true noncases correctly identified."[15] The following equation is used to calculate specificity: True Negative / True Negative + False Positive). NPV is the probability that a negative result is a true negative given a specified probability threshold for the variable of interest.[16]

Elkin et al.[17] evaluated the sensitivity, specificity, and PPV of SNOMED CT in representing 4,996 problem lists found in the Mayo (Rochester) Sheet Index. Of the problems related to inpatient and outpatient episodes of care, they found that SNOMED CT succeeded in representing most of the terms. Their findings suggest that SNOMED CT had a sensitivity of 92.3 percent, specificity of 80 percent, PPV of 99.8 percent, and NPV of 8.6 percent.

In that study, the Mayo Clinic Vocabulary Server (MCVS) was utilized as a tool to map free text into encoded data. Results showed that 4,312 clinical concepts were correctly mapped to SNOMED CT out of the 4,996 terms that were automatically mapped to SNOMED CT; 256 clinical concepts out of the 4,568 terms that were represented correctly by SNOMED were not completely mapped by the MCVS application; and 9 clinical concepts were inappropriately mapped by MCVS and SNOMED CT. For the MCVS mapping, these results suggest that without correction for synonymy, the specificity was 97.9 percent, sensitivity was 94.4 percent, and PPV was 99.8 percent. On the other hand, when correcting for synonymy, sensitivity was improved by 5.3 percent (99.7 percent), specificity was 97.9 percent, and PPV was 99.8 percent. Thus, this experiment suggested that SNOMED CT is a powerful and effective controlled terminology that can be used in CDSSs to advance patient safety in addition to preventive care. However, enhancements to synonymy as well as the addition of missing modifiers would advance the effectiveness of SNOMED CT to represent common problems found in problem lists.[18]

The use of CDSSs to advance preventive care depends on the effectiveness of SNOMED CT in representing clinical procedures. An important study conducted by De Silva et al.[19] evaluated the effectiveness of SNOMED CT in representing computed tomography exam findings. They investigated the sensitivity, specificity, PPV, and NPV of SNOMED CT for precoordinated expressions (i.e., "computed tomography guidance for needle biopsy expressed as a single Concept ID: 14211004"), and postcoordinated expressions (i.e., "computed tomography of abdomen for renal colic can be represented as:

> 169070004?|?CT???of???abdomen?|:?363702006?|?has???focus?|?=7093002?|?renal???colic?|"). Their results suggested that SNOMED CT had 56 percent sensitivity, 75 percent specificity, 81.1 percent PPV, and 47.6 percent NPV for precoordination, whereas for postcoordination, SNOMED CT had 98 percent sensitivity, 75.4 percent specificity, 88.2 percent PPV, and 95.2 percent NPV.

The application of CDSSs to enhance preventive medicine depends on the use of compositional terminologies—such as SNOMED CT—that can capture and represent negated concepts. Elkin et al.[20] investigated the ability of SNOMED CT to capture and cover negated concepts (i.e., *no*, *denies*, *ruled out*). The researchers used SNOMED CT to provide concept representation for 14,791 concepts found in 41 history and physical examination reports. Furthermore, they automatically assigned, to each of these concepts, an attribute documenting that the concept is a positive, negative, or uncertain assertion. To explain these assertions, the authors offered the following example: "The patient denied [operator] a history of previous cardiac disease [negative assertions] other than [operator] palpitations, which he experienced while giving a presentation resulting in syncope [positive assertions]."[21] A medical terminologist reviewed the documents and was able to identify 1,823 concepts—out of the total concepts reviewed—as negative. The findings of this experiment suggested that the sensitivity and specificity of the assignment of negation were 97.2 percent and 98.8 percent, respectively. The PPV was 91.2 percent, and

SNOMED CT successfully covered 88.7 percent of the clinical concepts that were automatically negated. Missed negation can lead to excess medical testing and, consequently, increased medical errors. Similarly, clinicians who wrongly assign negations would miss vital health data (e.g., medication allergies) that can negatively affect patient health and safety.

The integration of medical decision support rules with SNOMED CT enables the creation of a powerful CDSS. Elevitch[22] stated that an important factor that triggers decision support rules is the mapping of different kinds of concepts to each other (i.e., drugs to allergies, operations to medical devices, and disorders to contraindications). Additionally, several of the recognized rules are based on data and information that may be documented, during the course of the patient journey, by different healthcare practitioners at different times and places. For instance, alerts or clinical reminders can be constructed on the basis of disease coding and allergy medical coding with regard to specific drug codes. CDSS rules built on SNOMED CT support the creation of a high-quality decision support system for preventive care.

The use of CDSSs to advance the quality of healthcare services as well as preventive care depends on the application of clinical terminologies such as SNOMED CT that can provide up- to-date patient data and information. The CDSSs that have been created to assist healthcare practitioners in the diagnostic procedure sometimes are "based on static data [i.e., already known data and information to the physician such as drug allergies] which may be out of date."[23] Ciolko et al.[24] compared four machine-learning techniques in terms of their effectiveness to work with SNOMED CT. These methods include artificial neural networks (ANNs), Bayesian networks, decision trees and random Forests, and Gaussian processes. The researchers compared these machine-learning systems in terms of their advantages, disadvantages, uses in healthcare, and applicability to work with SNOMED CT. They found that ANNs and Bayesian networks are the most suitable methods that can be used with SNOMED CT; more specifically, they concluded that the tree-augmented naïve (TAN) Bayesian network technique is the most effective technique to apply to SNOMED CT data.

Clinicians could benefit from CDSSs by applying a TAN Bayesian classifier—an extension of a naïve Bayesian network that "approximates the interactions between attributes by using a tree structure imposed on the naïve Bayesian structure"[25] —to SNOMED CT data. Moreover, Chinnasamy et al.[26] stated that a TAN Bayesian classifier "consists of a class node connecting to all child nodes each representing a feature. Moreover, each child node can have at most one other feature node as parent. [An] attractive property of the TAN Bayesian classifier is that it learns the probabilities from the data in polynomial time." The application of a TAN Bayesian classifier to SNOMED CT data would allow clinicians to access timely and current diagnostic data and information, which is of significant importance in advancing preventive care.[27]

Discussion

The main question in this study was whether the utilization of SNOMED CT in CDSSs can improve preventive medicine. A key result was that SNOMED CT could be regarded as a powerful clinical vocabulary within EHRs to advance preventive care in particular and CDSSs in general. A controlled vocabulary—such as SNOMED CT—aids clinicians in developing a data warehouse for use in CDSSs.[28] The integration of a comprehensive ontology such as SNOMED CT with CDSSs would ultimately enhance healthcare services. As a result, it would allow for better patient outcomes as well as enable early detection of vital diseases, advance population health, reduce costs, and improve the quality of healthcare services in general.[29] However, CDSSs need to be enhanced through the utilization of quality clinical terminology systems and other standards.

As our study shows, the use of SNOMED CT in CDSSs has positive implications in preventive health services, such as for the identification of drug interactions, abnormal laboratory results, medication allergies, and other concepts. Whether representing clinical concepts within problem lists or clinical procedures, SNOMED CT successfully encoded and represented these concepts. However, some clinical or negated concepts were not represented by SNOMED CT. Enhancements of the synonyms and modifiers in SNOMED CT would ultimately help in improving the ability of SNOMED CT to code and represent clinical concepts and consequently would result in better CDSSs. Moreover, if sensitivity and specificity rates are lower than expected, institutions should consider developing a clinical documentation improvement program.

It is critical for a CDSS to have access to and be able to relate all of the patient's diagnoses, symptoms, laboratory results, history, and physical examination data. The application of the TAN Bayesian method would allow clinicians to have a system that learns relationships (whether old or new) within a database in which SNOMED CT was used to encode clinical concepts; the probabilities of these relationships would create a Bayesian tree. As a result, this application would enhance critical decisions within a patient encounter. However, further study of the application of other artificial intelligence techniques in SNOMED CT coded and raw patient data is required.[30]

This exploratory research study has several limitations. First, this study discusses the use of SNOMED CT in CDSSs for preventive care. Other classifications and coding systems could also be utilized to enhance the outcome of preventive medicine. Important ontologies that should be considered in future research include the Medical Dictionary for Regulatory Activities (MedDRA) and the Common Terminology Criteria for Adverse Events (CTCAE) classifications. CTCAE specifically covers adverse events, while SNOMED CT and MedDRA extend beyond adverse events.[31] Second, this study did not review the use of the Unified Medical Language System (UMLS), namely, its Metathesaurus, to allow for data and information exchange between SNOMED CT and the previously mentioned coding systems. It is estimated that 284,859 concepts from UMLS contain preferred terms from SNOMED CT—one concept from UMLS can contain more than one term from SNOMED CT.[32] Finally, this research did not review the integration of Logical Observation Identifiers Names and Codes (LOINC) with SNOMED CT in CDSSs for preventive care. Careful attention should be given when integrating the two ontologies into CDSSs because SNOMED CT and LOINC overlap in the domain of laboratory procedures. However, the utilization of UMLS would help in overcoming this issue.[33]

Conclusion

Our study provided important findings on the utilization of SNOMED CT in CDSSs for preventive health services. SNOMED CT had high sensitivity and specificity plus good results for PPV and NPV in representing clinical and negated terms. However, enhancement of modifiers and synonymy in SNOMED CT is needed. SNOMED CT lacks learning elements because it is a standard terminology that consists of lists and tables of clinical terms. This characteristic causes SNOMED CT to have errors that are sometimes difficult to eliminate or correct. By using a learning component such as the TAN Bayesian network method, clinicians could mine for possible and new relations in SNOMED CT data. Finally, information system ontologists, bioinformaticians, healthcare providers, and other applicable specialists should collaborate to create stated principles that allow for better use and higher quality of CDSSs. More specifically, they should cooperate to develop a unified pattern that would allow them to eliminate, for instance, system incompatibility issues. Such cooperation would allow for the development of a standard CDSS that uses standard clinical terminology for a variety of health, medical, and preventive care services.

Bader Al-Hablani, MS, RHIA, CHPS, is a research associate at King Faisal Specialist Hospital and Research Centre in Riyadh, Saudi Arabia.

Notes

[1] Pestontnik, S. L. "Expert Clinical Decision Support Systems to Enhance Antimicrobial Stewardship Programs: Insights from the Society of Infectious Diseases Pharmacists." *Pharmacotherapy* 25, no. 8 (2005): 1116–25.

[2] National Institutes of Health. "SNOMED CT." 2012.

[3] Ibid.

[4] Goldstein, D. E. Medical informatics 20/20: Quality and electronic health records through collaboration, open solutions, and innovation. Sudbury, MA: Jones and Bartlett, 2007 D.

[5] Shortliffe, E. H. "Computer Programs to Support Clinical Decision Making." *JAMA* 258, no. 1 (1987): 61–66.

[6] Pestontnik, S. L. "Expert Clinical Decision Support Systems to Enhance Antimicrobial Stewardship Programs: Insights from the Society of Infectious Diseases Pharmacists."

[7] Tsui, R. "Core Components in CDSS: Clinical Decision Support Systems." Lecture, University of Pittsburgh, Department of Biomedical Informatics, September 26, 2012.

[8] Kohn, L. T., Corrigan, J., & Donaldson, M. S. (1999). To err is human: Building a safer health system. Washington, D.C.: National Academy Press.

[9] Jao, C. S., & Hier, D. B. Clinical Decision Support Systems: An Effective Pathway to Reduce Medical Errors and Improve Patient Safety. INTECH Open Access Publisher, 2010.

[10] Ibid.

[11] Ibid.

[12] Ibid.

[13] Forrestal, E., and V. J. Watzlaf. *Health Informatics Research Methods: Principles and Practice*. Chicago, IL: AHIMA, 2009.

[14] Ibid.

[15] Ibid.

[16] Ibid.

[17] Elkin, P. L., S. H. Brown, C. S. Husser, B. A. Bauer, D. Wahner-Roedler, S. T. Rosenbloom, and T. Speroff. "Evaluation of the Content Coverage of SNOMED CT: Ability of SNOMED Clinical Terms to Represent Clinical Problem Lists." *Mayo Clinic Proceedings* 81, no. 6 (2006): 741–48.

[18] Ibid.

[19] De Silva, T. S., D. MacDonald, G. Paterson, K. C. Sikdar, and B. Cochrane. "Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) to Represent Computed Tomography Procedures. *Computer Methods and Programs in Biomedicine* 101, no. 3 (2011): 324–29.

[20] Elkin, P. L., 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 Medical Informatics and Decision Making* 5 (2005): 13.

[21] Ibid.

[22] Elevitch, F. R. "SNOMED CT: Electronic Health Record Enhances Anesthesia Patient Safety." *AANA Journal* 73, no. 5 (2005): 361–66.

[23] Ciolko, E., F. Lu, and A. Joshi. "Intelligent Clinical Decision Support Systems based on SNOMED CT." *Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (2010): 6781–84.

[24] Ibid.

[25] Friedman, N., D. Geiger, and M. Goldszmidt. "Bayesian Network Classifiers." *Machine Learning* 29 (1997): 131–63.

[26] Chinnasamy, A., W. K. Sung, and A. Mittal. "Protein Structure and Fold Prediction Using Tree-Augmented Naïve Bayesian Classifier." *Pacific Symposium on Biocomputing* 9 (2004): 387–98.

[27] Ciolko, E., F. Lu, and A. Joshi. "Intelligent Clinical Decision Support Systems based on SNOMED CT."

[28] Amatayakul, M. *Electronic Health Records: A Practical Guide for Professionals and Organizations.* 2nd ed. Chicago, IL: AHIMA, 2004.

[29] Bowman, S. "Coordination of SNOMED-CT and ICD-10: Getting the Most out of Electronic Health Record Systems." *Perspectives in Health Information Management* (Spring 2005).

[30] Ciolko, E., F. Lu, and A. Joshi. "Intelligent Clinical Decision Support Systems based on SNOMED CT."

[31] Richesson, R. L., K. W. Fung, and J. P. Krischer. "Heterogeneous But 'Standard' Coding Systems for Adverse Events: Issues in Achieving Interoperability between Apples and Oranges." *Contemporary Clinical Trials* 29, no. 5 (2008): 635–45.

[32] Ibid.

[33] Bodenreider, O. "Issues in Mapping LOINC Laboratory Tests to SNOMED CT." *AMIA Annual Symposium Proceedings* (2008): 51–55.

Article citation:
Al-Hablani, Bader. "Use of Automated SNOMED CT Clinical Coding in Clinical Decision Support Systems for Preventive Care" *Perspectives in Health Information Management* (Winter, January 2017).
