The Science and Practice of SNOMED CT Implementation

by:
Dennis <u>Lee</u> Hon Kit
BBA, Walla Walla University, 2000
MSc, University of Victoria, 2008

A Dissertation Submitted in Partial Fulfilment of the Requirement for the Degree of

Doctor of Philosophy

in the School of Health Information Science

SUPERVISORY COMMITTEE

The Science and Practice of SNOMED CT Implementation
Dennis Lee Hon Kit
BBA, Walla Walla University, 2000
MSc, University of Victoria, 2008

Dr Francis Lau, Supervisor
School of Health Information Science, University of Victoria

Dr Ronald Cornet, Member
Department of Medical Informatics, Academic Medical Centre, University of Amsterdam
Department of Biomedical Engineering, Linköping University

Dr Jens Weber, Non-Unit Member
Department of Computer Science, University of Victoria
School of Health Information Science, University of Victoria

Abstract

Supervisory Committee

Dr Francis Lau, Supervisor
School of Health Information Science, University of Victoria

Dr Ronald Cornet, Member
Department of Medical Informatics, Academic Medical Centre, University of Amsterdam
Department of Biomedical Engineering, Linköping University

Dr Jens Weber, Non-Unit Member
Department of Computer Science, University of Victoria
School of Health Information Science, University of Victoria

The overall research question of this PhD research was: "How can the clinical value of the Systematised Nomenclature of Medicine Clinical Terms (SNOMED CT) be demonstrated in the primary health care setting to enhance patient care?" The position taken in this research is that there is clinical value in using SNOMED CT.

To inform the current state of knowledge, a literature review of SNOMED CT papers catalogued by PubMed and Embase between 2001 and 2012 was carried out, and interviews were conducted with 14 individuals from 13 health care organisations across eight countries. The results showed there was a lack of understanding of how to craft post-coordinated expressions, how to fully utilise the semantics of SNOMED CT in data retrieval, and a lack of evidence on how SNOMED CT added value.

A proposed SNOMED CT Clinical Value Framework that organised the primary and secondary uses of SNOMED CT was created and a SNOMED CT design methodology was formalised that consisted of three components to aid in auditing, encoding and retrieval through a primary health care study.

In this PhD research, the potential clinical value of SNOMED CT was demonstrated by improving the completeness of clinical records and facilitating decision support features such as alerting clinicians to potential drug-allergy interactions, and reminding clinicians to order routine tests. The realisation of the potential clinical value was based upon the accurate and unambiguous manner in which clinical terms were encoded using the encoding method, the efficient and effective retrieval of relevant concepts using the retrieval method, and to a lesser extent, the ensuring that the concepts used were consistent using the auditing method.

TABLE OF CONTENTS

List of Tables

List of Figures

ACKNOWLEDGEMENTS

After writing over 300 pages for this dissertation I find myself at a loss of words on how to express my gratitude to those who have helped me so much with conducting my research, writing my dissertation and preparing for my exam. This is my humble attempt to acknowledge those who have helped me along this challenging but enjoyable journey.

Thank you, Dr Francis Lau, my supervisor and mentor, who somehow convinced me to do my PhD, something I never thought I would do. Words are inadequate to express my gratitude for what you have done for me over the past seven years. From guiding me throughout my master's and doctorate degrees, involving me in your research projects, spending countless hours in discussion and reviewing drafts, providing generous financial support, and for believing in me.

Thank you, Dr Ronald Cornet, for taking the time to answer numerous long emails, for providing suggestions on improving my dissertation, for co-authoring papers with me, for co-presenting with me at a conference, and for taking the time to participate in committee meetings even though it was late in Amsterdam.

Thank you, Dr Jens Weber, for once again joining the ride after being on my master's committee, for always asking tough questions that helped me think and improve my methods, for participating in the committee meetings, for reviewing all the drafts and providing valuable insights.

Thank you, Dr Morgan Price, for letting me use data from your clinic, for providing your clinical expertise and for spending time to explain clinical concepts to me.

Thank you, Dr Nicolette de Keizer, for being a co-investigator in the literature review and implementation survey, and for always providing valuable feedback.

Thank you, Ed Jones, for helping me with the data extraction so that I would have data to use for my research.

Thank you, to the participants who took part in the implementation survey and expert feedback sessions, whose names I cannot mention for privacy reasons, for graciously participating and providing valuable feedback.

Thank you, to the staff at the School of Health Information Science, Sandra Boudewyn, Diane Johnston, Debbie Robertson, Shawna McNabb and Dave Hutchinson, for always being so friendly and kind and for providing any help I have needed.

Thank you, the University of Victoria, for providing financial aid through the graduate fellowships.

Thank you, Dr James Campbell, for willing to be the external examiner and for your questions in the exam and for your feedback.

Thank you, to my wife Jenna, my parents, Peng Cheong and Dorothy, my sister, Evonne, my brother-in-law, John, my nieces Little Zoe and Little Lexi, and friends, for encouraging and believing in me.

Most importantly, thank you God, for the countless blessings you have bestowed upon me, for all the individuals mentioned above, for providing wisdom just for asking, and for your continued guidance in my life.

Dedication

To my loving wife, Jenna, who left her home of Toronto where her family and close friends live, to come to Victoria with me as I pursued my studies. For always being patient when my "couple of minutes" turned into hours. Thank you for being so understanding. And to my soon to be born son, hopefully you don't come early and I will finish my oral exam before your birth so that you will have my full attention.

1. OVERVIEW

1.1 Background

The Systematised Nomenclature of Medicine Clinical Terms (SNOMED CT) is considered the most comprehensive multilingual clinical reference terminology. The use of SNOMED CT aims to improve patient care by enabling clinical data to be recorded at a granular level, and to facilitate decision support through the retrieval of encoded data at various levels of aggregation.[1] SNOMED CT has been designated as the preferred or most suitable clinical reference terminology of choice in countries such as the United States, United Kingdom and Canada, and the internationalisation of SNOMED CT over the past seven years has helped to advance its uptake. Yet there are few detailed studies on how SNOMED CT is being adopted in health care organisations as the majority of studies published in the scientific literature have focused on comparing or mapping SNOMED CT to local terms or other terminology systems.[2] Therefore there is a need to determine how SNOMED CT is being adopted and in what ways it is helping to improve patient care.

The overall research question in this PhD research is: How can the clinical value of SNOMED CT be demonstrated in the primary health care setting to enhance patient care? The position taken in this research is that there is clinical value in using SNOMED CT. For this research, the clinical value of SNOMED CT in primary health care is defined as: the ability to improve care through accurate, consistent and unambiguous use of SNOMED CT to represent patient data in ways that can enhance decision support capabilities in primary health care settings.

In 2011, Canada Health Infoway developed a set of criteria called Clinical Value Targets to measure the effective use of electronic medical records (EMRs).[3] There are currently two levels: Clinical Value Level 1 and Level 2. Clinical Value Level 1 centres on the use of EMRs in areas such as the recording of problem lists, encounter diagnoses, prescriptions, use of EMR-generated alerts and reminders and electronic exchange of laboratory results while Clinical Value Level 2 centres on linking EMRs with provincial drug information systems and electronic prescribing. In addition to the Clinical Value Targets, there are at least three ongoing initiatives in Canada (i.e., EMR Content Standards[4] and Primary Health Care Indicators[5] by the Canadian Institute for Health Information (CIHI), and Health System Use[6] Project by Canada Health Infoway) in which the objectives of the Clinical Value Targets are evident. The EMR Content Standards are a set of data element and value set specifications that identify what data should be captured electronically in primary health care with the overall aim of "improv[ing] access, quality, outcomes and chronic disease prevention and management."[7] The Primary Health Care Indicators and Health System Use demonstration projects focus on using the data in areas such as quality reporting, decision support, statistical analysis and other secondary uses.

To facilitate the implementation of SNOMED CT, Canada Health Infoway has taken the lead to develop subsets and extensions in the primary health care domain.[8] The subsets, called Primary Health Care Reference Sets, were created to help streamline the implementation of SNOMED CT by identifying relevant concepts in 41 domains (e.g., client reason for encounter; health concern; intervention) while the extensions were created to address deficiencies in terms of concept and description coverage that are needed in the Canadian health care context that were lacking in the international release of SNOMED CT. It should be noted that the subsets are also comprised of other standardised terminologies.

These Canadian initiatives all converge on at least three common areas. First, the need to capture clinical data in a standardised manner that is accurate and consistent. Second, the need to query the data captured to facilitate clinical decision support such as generating alerts and reminders as well as other secondary uses. Third, the need to improve care coordination. These areas also correspond to the stated benefits that SNOMED CT has been promoted to meet.[1]

1.2 Research Questions and Rationale

Three specific questions were addressed in this research. First, to what extent is SNOMED CT being adopted? Second, how can SNOMED CT be used to represent clinical data in primary health care settings in a manner that is accurate, consistent and unambiguous? Third, how can SNOMED CT be used to enhance decision support capabilities in primary health care settings such as identifying high-risk patients and drug-allergy alerts?

The rationale for the first research question is the lack of detailed studies on SNOMED CT use. Two surveys by Elhanan, et al.,[9] in 2010 and the International Health Terminology Standards Development Organisation (IHTSDO) centred on *what* aspects of SNOMED CT were being used (e.g., which top-level hierarchies were used; tooling; maintenance; quality assurance; reference sets; and mappings). As closed-questions are limited in answering the *how* questions (e.g., how are reference sets derived; and how are they used), a study on *how* SNOMED CT is used is needed. This part of my research focuses on conducting a literature review of all English SNOMED CT-related papers catalogued in the PubMed and Embase databases and conducting a series of interviews with individuals who have implemented SNOMED CT in health care settings in order to be informed of the current state of knowledge of how SNOMED CT is being used.

The second question draws on the proposed definition of the clinical value of SNOMED CT and the need to capture data in a standardised format. There have been many papers published on the content coverage of SNOMED CT in different domains as well as the comparison of SNOMED CT to other standardised vocabulary systems.[2] However, most of the studies have focused on using only pre-coordination. The studies that have used post-coordination did not include a detailed description of the approach used in constructing the post-coordinated expressions[10, 11, 12, 13, 14] with three exceptions. The approach taken by Wang, et al.,[15] which is based on the SNOMED CT Technical Implementation Guide, included breaking the candidate terms into atomic terms and then locating the appropriate Concept Model attribute by matching the domain with the range. Sampalli, et al.,[16] expanded on that method to include checking for qualifiers. Richesson, et al.,[17] included the use of contextual qualifiers in addition to clinical qualifiers. A review of the sample post-coordinated expressions in the literature revealed errors in relationship type, relationship target and appropriateness of anatomical structure used.[18] There were also difficulties in restricting the creation of meaningful expressions and the challenge of creating undetected duplicate concepts.[19] Therefore what is needed is a method that can facilitate the crafting of accurate, consistent and unambiguous post-coordinated expressions.[15] This part of my research focuses on developing an automated post-coordination encoding method based on an analysis of how SNOMED CT concepts are currently defined and builds on a previous encoding method developed.[20]

As subsets and extensions have been created by Canada Health Infoway to help facilitate the implementation of SNOMED CT, there is a need to validate, or audit, these subsets and extensions to ensure they do not contain any errors. Authors of auditing methods of terminology systems have reported that it is "not uncommon,"[21] "unavoidable,"[22, 23] and "inevitable"[21] that errors occur in large complex terminology systems. Therefore it is reasonable to suppose that SNOMED CT subsets and extensions are also susceptible to containing errors. While auditing the entire SNOMED CT content for quality assurance is an important step, it will be mostly beyond the scope of this research as it could very well be a single topic for a dissertation on its own. The auditing method in this case is therefore restricted to verifying the consistency of extensions and focuses on areas such as ensuring the extension concepts contain a fully specified name and preferred term, and that concept definitions conform to the Machine Readable Concept Model.

The third research question stems from the need to be able to query the data captured for clinical decision support and other reporting purposes. The use of SNOMED CT queries to retrieve patient records is still in its infancy as most studies have centred on content coverage and mapping, which is usually the first stage in determining the feasibility of implementing SNOMED CT. Queries using predecessors of SNOMED CT such as SNOMED and SNOMED RT have been compared to classification systems such as the International Statistical Classification of Diseases and Related Health Problems, Ninth Revision, Clinical Modification (ICD-9-CM)[24] and the Internal Classification of Primary Care (ICPC)[25] and have shown to improve accuracy and sensitivity. Studies that have explored SNOMED CT queries used various methods including description logic reasoning,[26] structural subsumption,[27] structured query language (SQL)[28] and full-text queries.[29] This part of my research focuses on how SNOMED CT expressions that are used for queries should be structured to ensure all relevant data is retrieved, how SNOMED CT expressions should be stored, and how to efficiently and effectively execute the queries. The encoding, auditing and retrieval methods are collectively referred to as the clinical value design methodology.

1.3 Contributions in this PhD Research

There are five main areas in which this PhD research contributes to the current state of knowledge. First, verification rules for auditing extensions. Eighty-nine verification rules for identifying errors in SNOMED CT extensions were identified, organised into six categories (i.e., identifiers; value sets; dependencies; Machine Readable Concept Model; inferred relationships; and occurrences) and represented in a machine-readable format that can be re-used to verify extensions before they are published. Second, an automated post-coordinated encoding method. This includes a sorting method to assist in selecting the appropriate concept when there is more than one exact match, the proposed Extended Concept Model to assist in selecting the appropriate body structures, the encoding templates that help in constructing a post-coordinated expression, and nine scenarios for potentially modelling expressions. The automated post-coordinated encoding method was demonstrated to be accurate between 82.0% and 84.9% of the time. Third, the optimised data storage method for storing post-coordinated expressions. This method includes storing the pre-computed long normal form and the contextual values in discrete data elements and supports the optimised structural subsumption query method developed. Fourth, the optimised structural subsumption method for data retrieval. The optimised structural subsumption is based on three types of predicate expressions to retrieve an initial set of candidate expressions before using the long normal forms. The reduction of the number of candidate expressions that needed to be compared ranged from 62.9% to 99.9% for seven chronic disease categories. Fifth, the examples of demonstrating the clinical value of SNOMED CT. They include suggestions for billing diagnostic codes, suggestions for ordering laboratory tests based on the encounter diagnosis, suggestions for adding non-transient illnesses[30] from the encounter diagnosis into the problem list, suggestions for prescribing medications based on out of range results, alerts for potential drug-allergy interactions and suggestions for improving the organisation of a clinical summary.

1.4 Road Map

The rest of this dissertation is divided into nine chapters. **Chapter Two** provides an overview of the research approach taken in this PhD research. **Chapters Three** and **Four** address the first research question ("To what extent is SNOMED CT being adopted?") by reporting on the results of a literature review of SNOMED CT use and the results of a SNOMED CT implementation survey. **Chapter Five** describes the conceptual frameworks that were used throughout this PhD research. **Chapters Six**, **Seven** and **Eight** delve into the development of the clinical value design methodology of how to ensure extensions are consistent, how to encode clinical statements with SNOMED CT, and how to retrieve the encoded data. These three chapters address the second ("How can SNOMED CT be used to represent clinical data in primary health care settings in a manner that is accurate, consistent and unambiguous?") and third ("How can SNOMED CT be used to enhance decision support capabilities in primary health care settings such as identifying high-risk patients and drug-allergy alerts?") research questions. The primary focus on these chapters should be on the methods that were used as those methods were developed in this PhD research as part of the clinical value design methodology. The results are still important as they demonstrate that the method works, but they are secondary to the method developed. **Chapter Nine** describes the application of the clinical value design methodology in an anonymised primary care dataset and demonstrates the potential added value of SNOMED CT when used in an EMR system, which ties in to the overall research question of "How can the clinical value of SNOMED CT be demonstrated in the primary health care setting to enhance patient care?" **Chapter Ten** ends with a summary of findings, lessons learned, contribution of this research to the current state of knowledge, limitations and future work. While chapters three to four and six to nine can be read as individual standalone papers, it is also important to keep in mind the context of how these chapters relate to the rest of the dissertation. Each of these chapters contains a short introduction to provide context for the material described in the chapter. It is important to note that the limitations and the conclusions from these chapters are only described in **Chapter Ten**, which provides an overall summary of the dissertation.

This roadmap is summarised below:

INTRODUCTION
- Chapter 1 – Overview
- Chapter 2 – Research Approach

CURRENT STATE OF KNOWLEDGE
- Chapter 3 – Literature Review of SNOMED CT Use
- Chapter 4 – A Survey of SNOMED CT Implementations

STUDY METHODS & FINDINGS
- Chapter 5 – Conceptual Frameworks
- Chapter 6 – Auditing Method & Results
- Chapter 7 – Encoding Method & Results
- Chapter 8 – Retrieval Method & Results
- Chapter 9 – Towards Demonstrating the Clinical Value of SNOMED CT

CONCLUSION
- Chapter 10 – Lessons Learned

1.5 References

[1] SNOMED CT Value Proposition. http://www.ihtsdo.org/snomed-ct/whysnomedct/snomedfeatures. Last accessed: September 18, 2013.

[2] Cornet, R., & de Keizer, N. (2008). Forty years of SNOMED: a literature review. BMC medical informatics and decision making, 8(Suppl 1), S2. http://www.ncbi.nlm.nih.gov/pubmed/19007439.

[3] Clinical Value: "Meaningful Use" in Canada. http://infowayconnects.infoway-inforoute.ca/blog/clinicians-and-health-informatics/220-clinical-value-meaningful-use-in-canada. Last accessed: May 25, 2011.

[4] Canadian Institute for Health Information. (2012). Draft Pan-Canadian Primary Health Care Electronic Medical Record Content Standard, Version 2.1 – Implementation Guide. https://secure.cihi.ca/free_products/PHC_EMR_CS_Implementation_Guide.pdf.

[5] Canadian Institute for Health Information. (2012). Pan-Canadian Primary Health Care Indicator Report. https://secure.cihi.ca/free_products/Pan-Canadian_PHC_Indicator_Update_Report_en_web.pdf.

[6] [PowerPoint Presentation] Canada Health Infoway. (2011). Health System Use Demonstration Projects Overview.

[7] What are EMR content standards for PHC? http://www.cihi.ca/CIHI-ext-portal/internet/EN/document/types+of+care/primary+health/faq_ph_emr_content_stdrds. Last accessed: July 24, 2013.

[8] Canada Health Infoway. SNOMED CT RefSets. https://infocentral.infoway-inforoute.ca/2_Standards/1_pan-Canadian_Standards/Terminology/4_SNOMED_CT_Terminologies. Note: Login is required.

[9] Elhanan, G., Perl, Y., & Geller, J. (2011). A survey of SNOMED CT direct users, 2010: impressions and preferences regarding content and quality. Journal of the American Medical Informatics Association, 18(Suppl 1), i36-i44. http://www.ncbi.nlm.nih.gov/pubmed/21836159.

[10] Osornio, A. L., Luna, D., Gambarte, M. L., Gomez, A., Reynoso, G., & de Quiros, F. G. (2007). Creation of a local interface terminology to SNOMED CT. Studies in health technology and informatics, 129(Pt 1), 765. http://www.ncbi.nlm.nih.gov/pubmed/17911820.

[11] Wade, G., & Rosenbloom, S. T. (2008). Experiences mapping a legacy interface terminology to SNOMED CT. BMC medical informatics and decision making, 8(Suppl 1), S3. http://www.ncbi.nlm.nih.gov/pubmed/19007440.

[12] Pathak, J., Wang, J., Kashyap, S., Basford, M., Li, R., Masys, D. R., & Chute, C. G. (2011). Mapping clinical phenotype data elements to standardized metadata repositories and controlled terminologies: the eMERGE Network experience. Journal of the American Medical Informatics Association, 18(4), 376-386. http://www.ncbi.nlm.nih.gov/pubmed/21597104.

[13] De Silva, T. S., MacDonald, D., Paterson, G., Sikdar, K. C., & Cochrane, B. (2011). Systematized nomenclature of medicine clinical terms (SNOMED CT) to represent computed tomography procedures. Computer methods and programs in biomedicine, 101(3), 324-329. http://www.ncbi.nlm.nih.gov/pubmed/21316117.

[14] Wade, G., & Rosenbloom, S. T. (2009). The impact of SNOMED CT revisions on a mapped interface terminology: terminology development and implementation issues. Journal of biomedical informatics, 42(3), 490-493. http://www.ncbi.nlm.nih.gov/pubmed/19285570.

[15] Wang, Y., Patrick, J., Miller, G., & O'Hallaran, J. (2008). A computational linguistics motivated mapping of ICPC-2 PLUS to SNOMED CT. BMC medical informatics and decision making, 8(Suppl 1), S5. http://www.ncbi.nlm.nih.gov/pubmed/19007442.

[16] Sampalli, T., Shepherd, M., Duffy, J., & Fox, R. (2010). An evaluation of SNOMED CT® in the domain of complex chronic conditions. International journal of integrated care, 10. http://www.ncbi.nlm.nih.gov/pubmed/20422022.

[17] Richesson, R. L., Andrews, J. E., & Krischer, J. P. (2006). Use of SNOMED CT to represent clinical research data: a semantic characterization of data items on case report forms in vasculitis research. Journal of the American Medical Informatics Association, 13(5), 536-546. http://www.ncbi.nlm.nih.gov/pubmed/16799121.

[18] Navas, H., Lopez, O. A., Gambarte, L., Elías, L. G., Wasserman, S., Orrego, N., ... & de Quirós, F. G. (2010). Implementing rules to improve the quality of concept post-coordination with SNOMED CT. Studies in health technology and informatics, 160(Pt 2), 1045. http://www.ncbi.nlm.nih.gov/pubmed/20841843.

[19] Rosenbloom, S. T., Miller, R. A., Johnson, K. B., Elkin, P. L., & Brown, S. H. (2006). Interface terminologies facilitating direct entry of clinical data into electronic health record systems. Journal of the American Medical Informatics Association, 13(3), 277-288. http://www.ncbi.nlm.nih.gov/pubmed/16501181.

[20] Lee, D. H., Lau, F. Y., & Quan, H. (2010). A method for encoding clinical datasets with SNOMED CT. BMC medical informatics and decision making, 10(1), 53. http://www.ncbi.nlm.nih.gov/pubmed/20849611.

[21] Ceusters, W., Smith, B., Kumar, A., & Dhaen, C. (2004). Mistakes in medical ontologies: where do they come from and how can they be detected?. Studies in Health Technolology and Informatics, 102, 145-163. http://www.ncbi.nlm.nih.gov/pubmed/15853269.

[22] Wang, Y., Halper, M., Min, H., Perl, Y., Chen, Y., & Spackman, K. A. (2007). Structural methodologies for auditing SNOMED. Journal of Biomedical Informatics, 40(5), 561-581. http://www.ncbi.nlm.nih.gov/pubmed/17276736.

[23] Min, H., Perl, Y., Chen, Y., Halper, M., Geller, J., & Wang, Y. (2006). Auditing as part of the terminology design life cycle. Journal of the American Medical Informatics Association, 13(6), 676-690. http://www.ncbi.nlm.nih.gov/pubmed/16929044.

[24] Elkin, P. L., Ruggieri, A. P., Brown, S. H., Buntrock, J., Bauer, B. A., Wahner-Roedler, D., ... & Bergstrom, L. (2001). A randomized controlled trial of the accuracy of clinical record retrieval using SNOMED-RT as compared with ICD9-CM. In Proceedings of the AMIA Symposium (p. 159). American Medical Informatics Association.

[25] Lussier, Y. A., & Bourque, M. (1997). Comparing SNOMED and ICPC retrieval accuracies using relational database models. In Proceedings of the AMIA Annual Fall Symposium (p. 514). American Medical Informatics Association.

[26] Liu, S., Ni, Y., Mei, J., Li, H., Xie, G., Hu, G., ... & Pan, Y. (2009). ismart: Ontology-based semantic query of cda documents. In AMIA Annual Symposium Proceedings (Vol. 2009, p. 375). American Medical Informatics Association.

[27] Dolin, R. H., Spackman, K. A., & Markwell, D. (2002). Selective retrieval of pre-and post-coordinated SNOMED concepts. In Proceedings of the AMIA Symposium (p. 210). American Medical Informatics Association.

[28] Lieberman, M. I., & Ricciardi, T. N. (2003). The Use of SNOMED© CT Simplifies Querying of a Clinical Data Warehouse. In AMIA Annual Symposium Proceedings (Vol. 2003, p. 910). American Medical Informatics Association. http://www.ncbi.nlm.nih.gov/pubmed/14728416.

[29] Cuggia, M., Bayat, S., Garcelon, N., Sanders, L., Rouget, F., Coursin, A., & Pladys, P. (2009). A full-text information retrieval system for an epidemiological registry. Studies in health technology and informatics, 160(Pt 1), 491-495. http://www.ncbi.nlm.nih.gov/pubmed/20841735.

[30] Poissant L, Taylor L, Huang A, Tamblyn R. Assessing the accuracy of an inter-institutional automated patient-specific health problem list. BMC Medical Informatics Decision Making. 2010 Feb 23;10:10. http://www.ncbi.nlm.nih.gov/pubmed/20178586.

2. Research Approach

2.1 Introduction

This chapter provides an overview of the research approach, design science research,[1] which guided this PhD research. An overview of the approach is shown in **Figure 2-1.** While most research focuses on understanding a phenomenon,[2] as its name implies, design science research is a type of research that focuses on the design or development of artefacts such as algorithms, human computer interfaces, design methodologies and languages. The next three sections in this chapter describe the research approach by process, methods and outputs.

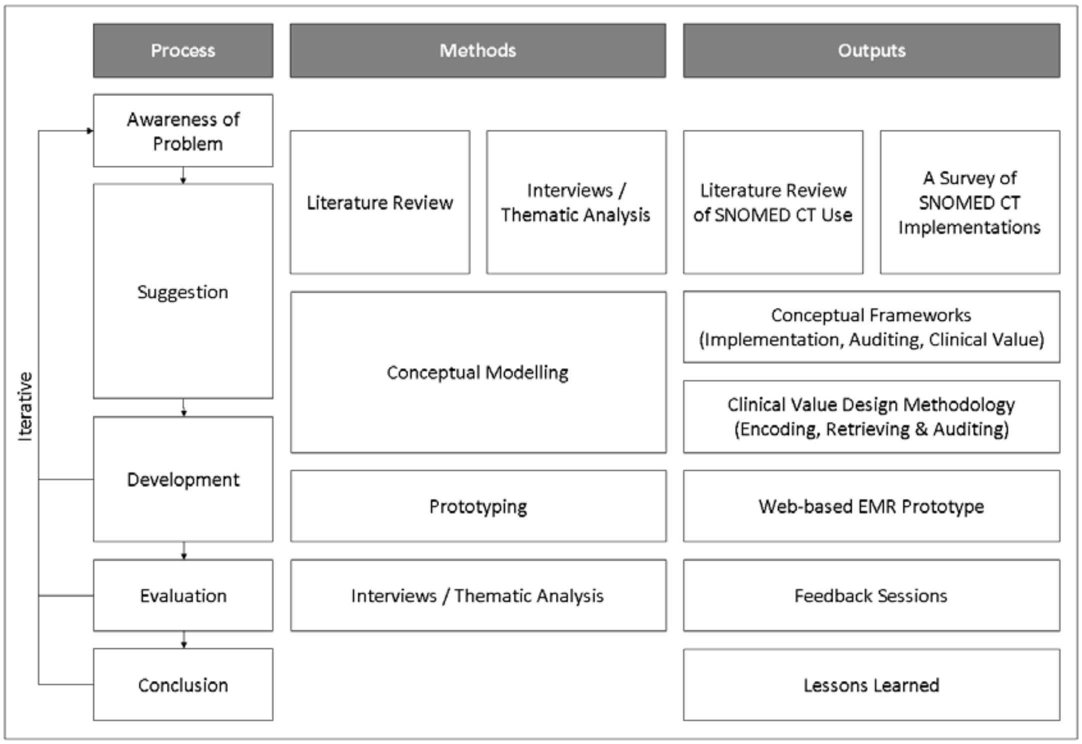


Figure 2-1. Overview of research approach.

2.2 Process

Design science research is based on reasoning in the design cycle that consists of five steps: awareness of problem, suggestion, development, evaluation and conclusion.[3] When trying to understand a problem or develop a solution for a problem, the first step is to be aware of the problem. The second step, suggestion, seeks to find solutions to the problem. The third step, development, is where a solution to the problem is being constructed based on the suggestions from the previous step and the development of artefacts and new knowledge. The fourth step, evaluation, seeks to evaluate the artefacts to determine the accuracy and relevance. The last step, conclusion, is the termination of the process where lessons learned are discussed and future work is described. The process is iterative in nature as suggestions derived from the literature, challenges encountered during development, issues raised during evaluation, and future work described in the conclusion feed back into a greater awareness of the problem or additional issues associated with the problem.

2.2.1 Awareness of Problem

In this research, the awareness of the problem came from multiple sources including past experience using SNOMED CT in research projects, and challenges identified in presentations and the scientific literature. The awareness of the problem has been raised in the previous chapter and was succinctly stated in the overarching research question, "How can the clinical value of SNOMED CT be demonstrated in the primary health care setting to enhance patient care?"

2.2.2 Suggestion

Suggestions on solutions for solving the problem identified would logically be drawn from the literature of the stated domain. This step included conducting a literature review of SNOMED CT use and an implementation survey to inform the current state of knowledge. The aim of the literature review was to investigate how SNOMED CT was being used and was based on the work by Cornet, et al.[4] While the literature review was expected to include the majority of published implementation studies, it was also expected that not all health care organisations publish their work in the scientific literature, which was why the implementation survey was conducted. The implementation survey was conducted with individuals who have implemented SNOMED CT in health care settings.

2.2.3 Development

The development step is where the solutions to the problem are developed as artefacts. In this study, the development step was further broken down in three parts. First, to formulate an implementation conceptual framework that could be used to measure the level of adoption of SNOMED CT in a health care setting. The conceptual framework was based on the Centre for Health Information Research and Development's (CHIRAD) Solutions Support Model and was refined for this research. Second, to develop a clinical value design methodology that can aid the adoption of SNOMED CT. The design methodology was developed using the conceptual modelling method and consisted of an auditing method to ensure the consistency of SNOMED CT extensions, an automated encoding method to encode free text clinical statements with SNOMED CT accurately, consistently and unambiguously, and a retrieval method to effectively and efficiently query patient data that had been encoded with SNOMED CT. Third, to develop an electronic medical record (EMR) system prototype that incorporates the clinical value design methodology to demonstrate how SNOMED CT can be implemented and the potential added value of using SNOMED CT. The EMR system prototype was developed using the prototyping method.

2.2.4 Evaluation

As with any research study, conducting an evaluation is an important step in determining the relevance and accuracy of the artefacts developed. This step involved conducting group interviews with primary care clinicians and terminology subject matter experts in order to solicit feedback on the design methodology developed and the application of the methodology in the EMR system prototype.

2.2.5 Conclusion

In design science research, the conclusion is the termination of the design project. It does not necessarily mean that the knowledge created during this process is absolute and there is no room for improvement, but rather the objectives of the research have been completed or the research questions have been adequately addressed within the defined scope.

2.3 Methods

This section provides an overview of the methods used throughout this research. They included a literature review, interviews, thematic analysis, conceptual modelling and prototyping.

2.3.1 Literature Review

In order to gain a better understanding of the current state of knowledge of SNOMED CT use, a literature review was conducted that was based on the 2008 SNOMED literature review conducted by Cornet, et al.[4] The literature review was conducted in conjunction with Dr Francis Lau, Dr Ronald Cornet and Dr Nicolette de Keizer. While the literature review by Cornet, et al.,[4] focused on all versions of SNOMED up to 2006, this literature review focused only on SNOMED CT and papers between 2001 and 2012. The databases used to identify SNOMED CT related papers were PubMed and Embase. PubMed is maintained by the United States National Library of Medicine and connects to the Medical Literature Analysis and Retrieval System Online (MEDLINE), which indexes life sciences and biomedical papers. Embase is maintained by Elsevier and focuses on drug-related publications but also includes medical and scientific papers.

The search strategy involved using the keyword "SNOMED" and phrase "systematised nomenclature of medicine" for papers published in English or had an English abstract. Only papers after 2000 were considered so as to exclude previous versions of SNOMED. Although the literature focused on the use of SNOMED CT, all SNOMED CT related papers were catalogued and classified using a similar set of criteria as in the SNOMED literature review by Cornet, et al.,[4] In all, four sets of criteria were used: focus category, usage category, medical domain and country of the location of study.

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-investigators. The results were compared and discussed until a consensus was reached on the differences and definitions on classification categories were refined. The co-investigators 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, I proceeded to classify the rest of the papers. I flagged 25 papers for review which I was uncertain of and these papers were reviewed by the other co-investigators and discussions took place to reconcile the classification.

2.3.2 Interview

The interview research method was used to uncover how individuals had implemented SNOMED CT and to solicit feedback on how SNOMED CT was used to demonstrate clinical value. The interview on how individuals had implemented SNOMED CT was once again conducted in conjunction with Dr Francis Lau, Dr Ronald Cornet and Dr Nicolette de Keizer. This method was used in the implementation survey and feedback sessions. Interviews enable researchers to try to understand phenomena from the view point of the participants and are one of the most commonly used qualitative research methods.[5] The semi-structured interview or "interview guide approach" was used.[6] In this type of interview, a list of topics and questions are prepared but the order and exact wording of the questions can vary with each interview. Advantages of interviews over observational studies include the flexibility of being able to ask questions that may not necessarily be observed, ability to focus on specific issues and to follow up on responses to gain an in-depth understanding of the subject being discussed.[7]

There are seven steps to designing and conducting an interview.[8] First, thematising, in which the purpose of the interview is identified. Second, designing, in which the study design is formalised and includes identifying potential participants, relevant questions, how and where the interview is to be conducted and how the results should be analysed and reported. Third, interviewing, in which the interview is conducted and recorded to facilitate a thorough analysis of the interview.[6] Fourth, transcribing, in which the recorded interview is transcribed verbatim to aid in the analysis. Fifth, analysing, in which different qualitative analysis methods such as thematic analysis can be used to extract the essence of the responses. Six, verifying, in which the themes identified in the analysis are verified with the transcripts and if necessary, facts are confirmed with external sources. Lastly, reporting on the results of the interview.

For the survey of SNOMED CT implementations, interview questions were compiled from previous SNOMED CT use questionnaires and questions were derived from the CHIRAD's Solutions Support Model. Questions were individually reviewed by the co-investigators and additional questions were added. The validity of the questions was discussed, duplicate questions were removed and the order of the questions was prioritised. Potential participants were identified by following up with individuals who had previously participated in an online questionnaire conducted by the International Health Terminology Standards Development Organisation (IHTSDO) and were agreeable to a follow-up interview, individuals listed on Canada Health Infoway's SNOMED CT use website,[9] and individuals who had presented to the IHTSDO Implementation Special Interest Group (SIG). Potential participants were invited by email and scheduling of the interviews was done using Doodle, a free web-based scheduling software. Phone interviews were conducted and were recorded to aid in the analysis (see section on thematic analysis).

For the feedback sessions, which were conducted after the design methodology and prototype were completed, three main open-ended questions (i.e., "does it make sense?"; "does it add value?"; and "what are the potential implementation issues?") were asked for each of the ways SNOMED CT was used. The feedback sessions were split into two types. First, a clinical feedback session that focused on the clinical value of using SNOMED CT. Second, a technical feedback session that focused on the technical aspects of how the clinical value of SNOMED CT was achieved. It should be noted that only notes for the feedback sessions were taken as the sessions themselves were not recorded.

Ethical issues of qualitative interviewing include confidentiality, informed consent, risk assessment, promises and reciprocity and interviewee mental health.[6] In both cases an ethics application was submitted to the University of Victoria Ethics Board and was approved. All participants were asked to sign a participant consent form or consent was implied by attending the interview or feedback session. There were no promises and reciprocity for participating in the interview. As this study did not focus on confidential or sensitive information, and the study setting was not conducted in any stressful situations, there was no need to debrief the interviewer or question the mental state of the interviewees. The written report did not contain any of the participant's names or their companies, or any identifying information about the participants. The ethics protocol numbers for the implementation survey and feedback sessions were 11-535 and 12-529 respectively and the approval notices are available in the **Appendix A**.

2.3.3 Thematic Analysis

The method used to analyse the interviews was thematic analysis. Thematic analysis is a process of identifying important patterns, or themes, through the analysis of a set of data that can help to understand the phenomena that is being studied.[10] Thematic analysis is widely used in qualitative studies[11] and the approach of identifying themes can be further divided into inductive ("bottom up") and theoretical ("top down"), which is also referred to as deductive. The inductive approach is akin to grounded theory whereby the themes identified are derived directly from the data[6] whereas in the theoretical approach the data tends to be fitted into the researcher's conceptual framework. Among the benefits of using thematic analysis are the flexibility and the fact that it can be applied to different research areas[12] and themes can emerge from the data.[13] However, a limitation is the reliability of interpreting the data and assigning codes to the data as it is subject to the interpretation of the researcher.[14]

Thematic analysis consists of a six step process.[12] The first step is to get familiarised with the data being analysed by a "careful reading and re-reading of the data."[15] The second step is to create an initial set of codes that can be used to represent the data. A code can be defined as "a word or short phrase that symbolically assigns a summative, salient, essence-capturing, and/or evocative attribute for a portion of a language-based or visual data."[13] The third step is to identify the candidate themes based on a collection of related codes. Themes can be defined as "conversation topics, vocabulary, recurring activities, meanings, feelings, or folk sayings and proverbs."[16] The fourth step is to review the themes identified to ensure they are coherent. The fifth step is to label and define the theme. The final step is to write the report.

After the implementation survey interviews were conducted, they were transcribed verbatim and imported into a database. An initial set of codes that were based on the implementation conceptual framework were also imported into a database. A web-based application was developed to facilitate the process of tagging the transcripts with the codes. The application also had a feature to add notes to the transcripts to aid in the analysis. Over the course of familiarisation with the data, new codes were added to the list. An inductive content analysis was conducted and the tagged transcripts for each theme were collated and reviewed to ensure they were coherent. Each theme was described and the associated data that supported each theme was summarised and written in the final report. While inductive content analysis was conducted and the themes were drawn directly from the data, it should be noted that the implementation of SNOMED CT is a very specific area of research and it is nearly impossible for a researcher to refrain from applying his or her own understanding of the phenomena being studied when assigning codes. In this study, only one researcher applied the codes which could be a concern for reliability. However, the themes and final report were reviewed by the other co-investigators in the study.

2.3.4 Conceptual Modelling

Conceptual modelling was used to develop a model that could be used to represent the level in which SNOMED CT has been adopted. Conceptual modelling is the method of abstracting what exists in the real world and to represent it in a graphical representation or model.[17] As not everything in the real world is known or understood, abstraction is necessary to simplify or to make assumption of what is unknown.[18] In the domain of information systems, the main purpose for creating conceptual models is to identify the necessary knowledge needed to perform functions in a given domain.[19] Conceptual models can be represented with one or more component lists, process flow diagrams, logic flow diagrams and activity cycle diagrams.[20]

The literature contains little information about the method of creating conceptual models as conceptual modelling is viewed as an art rather than a science[21] but conceptual modelling consists of knowledge abstraction and model abstraction.[22] Robson[20] developed a framework to aid modellers in developing their own conceptual models that consisted of four key elements: develop an understanding of the problem situation, determine the modelling objectives, design the conceptual model inputs and outputs and design the model content of the conceptual model. In terms of what to include in the model, the literature suggests keeping the model as simple as possible as long as the objectives of the study can be achieved.[20, 23]

While conceptual modelling in information systems leans towards using the model developed for simulations and software development, in this study, the objective of the conceptual model was to represent the different levels in which SNOMED CT can be adopted and to guide the process of developing a design methodology. In this study, rather than developing the implementation conceptual framework from scratch, the Solutions Support

Model developed by CHIRAD was refined. The Solutions Support Model is represented in the form of component lists and consists of three levels (i.e., no internal support for SNOMED CT; internal support for pre-coordinated SNOMED CT; and internal support for post-coordinated SNOMED CT) and seven dimensions (i.e., coded data storage; logical record architecture; extent of encoding; data entry; data retrieval; inbound communication; and outbound communication). Lessons learned on how SNOMED CT was being used in the literature review and implementation surveys added to the contents of the implementation conceptual framework. In addition, the development of the design methodology also helped to refine the content.

2.3.5 Prototyping

Prototyping was used to develop a scaled down version of an EMR system to demonstrate the clinical value of SNOMED CT. The development of a prototype can be started with just a succinct objective with few requirements as requirements gathering can be done in parallel with development. Although the final product may not be used as an operational application, the prototype should include the graphical user interface, major program modules and supporting database.[24]

There are many use cases for prototyping including requirements gathering and evaluating, determining feasibility of developing an operational system by first developing a cheaper pilot system as a proof of concept, identifying the risk involved in system development so as to reduce the risk involved, and communicating or testing ideas with end-users.[25] In this research, the main purpose for developing the prototype was to demonstrate the clinical value added aspects of SNOMED CT that are difficult to convey with just narratives. The prototype was not a full-blown EMR system and only included a limited set of features that were meant for the researcher to demonstrate to end-users and not for end-users to use. The benefits of using the prototyping method include, but are not limited to, the short turnaround time, the flexibility to make changes as requirements are identified, the ability to develop a functional but not necessarily optimised application.[26] While there are weaknesses and risks involved in using the prototyping method,[27] which includes no guarantee of an operational system, the possibility of inadequately identifying requirements, potential behavioural problems of end-users adapting to an iteration of a prototype and not adjusting to the operational system, the objective of the prototype was only to demonstrate the clinical value of SNOMED CT and was not meant to be an operational system to be used by end-users.

There are four main steps in prototyping.[28] First, the initial concept or identifying of fundamental system requirements. Second, the designing and implementing of the initial prototype. Third, the iterative rounds of refining and testing the prototype. Fourth, the completion of the prototype.

Unlike traditional systems development life cycles, not all requirements are necessary to be identified upfront, which was important in this study as not all the clinical value aspects of SNOMED CT were known in the beginning. The initial requirement was to be able to view patient records in the data repository. The prototype was a web-based application that was developed using PHP, a server-side scripting programming language, JavaScript, a client-side scripting programming language, and MySQL, a relational database management system. While it could be argued that PHP and JavaScript are not typical prototyping languages, they were sufficient to develop the prototype quickly. As more clinical value aspects of using SNOMED CT were identified, functions were programmed into the prototype. In this research, the user who was testing the prototype was the developer and the testing focused on ensuring the accuracy of the technical functionality rather than conducting usability tests with a group of users. That is because the objective of the prototype was to develop an EMR system with SNOMED CT-specific features to show to end-users but not to have end-users use the system. In this case, once the necessary clinical value features could be demonstrated, the prototype was considered complete.

2.4 Outputs

The seven outputs that were produced in this research are grouped into six categories. They include the literature review of SNOMED CT use, survey of SNOMED CT implementations, conceptual frameworks, clinical value design methodology, web-based EMR system prototype and feedback, and lessons learned.

2.4.1 Literature Review of SNOMED CT Use

The literature review included 488 papers published between 2001 and 2012 and catalogued by PubMed and Embase. It provided an overview of the current state of knowledge of SNOMED CT use by focus category (i.e., indeterminate; theoretical; pre-development/design; implementation; and evaluation/commodity), usage category (e.g., prospective content coverage; used to classify or code in a study), domain and country of publication. The literature review is available in **Chapter Three**.

2.4.2 A Survey of SNOMED CT Implementations

The survey of SNOMED CT implementations included interviews with 14 individuals in 13 organisations across eight countries and is available in **Chapter Four**. This study provided an overview of how SNOMED CT was being implemented (i.e., design, use; and maintenance), challenges encountered (e.g., quality and implementation challenges), success factors and lessons learned (e.g., simplicity, clinician involvement, expertise and collaboration) and outstanding issues (i.e., post-coordination, retrieval; and extensions).

2.4.3 Conceptual Frameworks

Three conceptual frameworks were refined or developed in this research. First, the SNOMED CT Implementation Conceptual Framework was based on a refinement of the CHIRAD's Solutions Support Model, and consists of four levels (i.e., no SNOMED CT support; pre-coordination; post-coordination; and extensions), which represent increasing levels of complexity in which SNOMED CT can be implemented, and three dimensions (i.e., auditing; encoding; and retrieval). The auditing dimension is to ensure that subsets, post-coordinated expressions and extensions are consistent; the encoding dimension is to ensure data is encoded accurately, consistently and unambiguously, while the retrieval dimension is to ensure SNOMED CT-encoded data can be retrieved effectively and efficiently.

Second, the SNOMED CT Extensions Auditing Framework. The framework was developed to organise the verification rules for identifying errors in extensions. The framework consists of three columns, which represent the three core components of SNOMED CT (i.e., concepts; descriptions; and relationships) and six rows, which represent the six categories of verification rules (i.e., identifiers; value sets; dependencies; Machine Readable Concept Model; inferred relationships; and occurrences).

Third, the SNOMED CT Clinical Value Framework. The framework was based primarily on the Clinical Value Targets by Canada Health Infoway that is used to determine the value of using EMRs. The framework consists of three columns, which represent the three broad data use audience (i.e., patient; practice; and population) and three rows, which represent the three types of data uses as defined by the Clinical Value Targets (i.e., data capture; data retrieval; and data sharing).

These three frameworks are described in **Chapter Five**.

2.4.4 Clinical Value Design Methodology

Three methods were developed as part of the clinical value design methodology. First, an auditing method to ensure the consistency of extensions through 89 verification rules that were categorised according to the Extensions Auditing Framework. This auditing method was used to determine the consistency of three national SNOMED CT extensions (i.e., Canada; United States; and United Kingdom).

Second, an automated post-coordination encoding method to ensure clinical statements are encoded accurately, consistently and unambiguously that consisted of three parts (i.e., improving the quality of the search results; modeling two or more expression together; and cleaning up the post-coordinated expression. This encoding method was used to encode free text problem lists and encounter diagnoses that were extracted from a primary care clinic.

Third, an optimised data storage method that consisted of recording portions of post-coordinated expressions (e.g., focus concept, proximal primitives and contextual qualifiers) as discrete data elements and an optimised structural subsumption method for retrieving patient data that is encoded with SNOMED CT through a three step method (i.e., structuring the predicate expression; retrieving potential candidates; and comparing the potential candidates). These data storage and retrieval method were used to store the results of the encoded problem lists and encounter diagnoses, and to identify chronic conditions that were recorded in these two sets of data.

The three methods developed are available in **Chapters Six** to **Eight**.

2.4.5 Web-based Electronic Medical Records Prototype and Feedback

To demonstrate the clinical value of SNOMED CT, a web-based EMR prototype was developed using PHP, JavaScript and MySQL. Clinical decision support features were programmed to demonstrate how SNOMED CT can improve the completeness of data and clinical summaries, identify high risk patients, facilitate decision support features and potentially improve care quality. The different types of clinical value that was demonstrated were organised according to the SNOMED CT Clinical Value Framework. The description of the clinical value of SNOMED CT, prototype and the lessons learned from the feedback sessions are available in **Chapter Nine**.

2.4.6 Lessons Learned

The final chapter of this dissertation summarises the findings between **Chapters Three** to **Nine** and relates how they address the three research questions. The limitations and future work are also summarised followed by a conclusion.

2.5 References

[1] Design Science Research in Information Systems. http://desrist.org/design-research-in-information-systems. Last accessed: July 24, 2013.

[2] Babbie, E. (2012). The practice of social research. CengageBrain. com.

[3] Takeda, H., Veerkamp, P., & Yoshikawa, H. (1990). Modeling design process. AI magazine, 11(4), 37.

[4] Cornet, R., & de Keizer, N. (2008). Forty years of SNOMED: a literature review. BMC Medical Informatics and Decision Making, 8(Suppl 1), S2. http://www.ncbi.nlm.nih.gov/pubmed/19007439.

[5] Seidman, I. (1998). Interviewing as qualitative research: A guide for researchers in education and the social sciences. New York, NY: Teachers College Press.

[6] Patton, M. Q. (1990). Qualitative evaluation and research methods. Sage Publications, Incorporated.

[7] Bryman, A. (2012). Chapter 15: Interviewing in qualitative research. In Social research methods. Oxford university press.

[8] Kvale, S., & Brinkmann, S. (2009). Interviews: Learning the craft of qualitative research interviewing. Sage Publications, Incorporated.

[9] SNOMED CT in Use. https://www.infoway-inforoute.ca/standards-collaborative/snomed-ctr/snomed-ct-in-use. Last accessed: November 25, 2011.

[10] Daly, J., Kellehear, A. & Gliksman, M. (1997). The public health researcher: A methodological approach. Melbourne, Australia: Oxford University Press.

[11] Boyatzis, R. (1998). Transforming qualitative information: Thematic analysis and code development. Thousand Oaks, California: Sage Publications, Incorporated.

[12] Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2). pp. 77-101.

[13] Saldana, Johnny. (2009). The Coding Manual for Qualitative Researchers. Thousand Oaks, California: Sage Publications, Incorporated.

[14] Guest, G. (2012). Applied thematic analysis. Thousand Oaks, California: Sage Publications, Incorporated.

[15] Rice, P., & Ezzy, D. (1999). Qualitative research methods: A health focus. Melbourne: Oxford University Press.

[16] Taylor, S. J., & Bogdan, R. (1984). Introduction to qualitative research methods: The search for meanings. New York: John Wiley & Sons.

[17] Conceptual modelling for shared understanding. http://learningforsustainability.net/social_learning/conceptual_modelling.php. Last accessed: July 30, 2013.

[18] Robinson, S. (2010). Conceptual Modelling: Who Needs It? SCS M&S Magazine.

[19] Olivé, A. (2000). An introduction to conceptual modeling of information systems. Advanced database technology and design. Artech House, 25-57.

[20] Robinson, S. (2004). Simulation: The Practice of Model Development and Use. Chichester: John Wiley & Sons.

[21] Wang, W., & Brooks, R. J. (2007). Improving the understanding of conceptual modelling. Journal of Simulation, 1(3), 153-158.

[22] Kotiadis, K., & Robinson, S. (2008). Conceptual modelling: Knowledge acquisition and model abstraction. In Proceedings of the 40th Conference on Winter Simulation (pp. 951-958). Winter Simulation Conference.

[23] Brooks, R. J. (2007). Conceptual modelling: framework, principles, and future research. (Management Science Working Paper Series). Lancaster University: The Department of Management Science.

[24] Tripp, S. D., & Bichelmeyer, B. (1990). Rapid prototyping: An alternative instructional design strategy. Educational Technology Research and Development, 38(1), 31-44.

[25] Luqi, L., & Steigerwald, R. (1992,). Rapid software prototyping. In System Sciences, 1992. Proceedings of the Twenty-Fifth Hawaii International Conference on (Vol. 2, pp. 470-479). IEEE.

[26] Ghosh, T. P. (2013). Information Systems Control and Audit. Systems Development Life Cycle Methodology. The Institute of Chartered Accountants of India.

[27] Floyd, C. (1984). A systematic look at prototyping. In Approaches to prototyping (pp. 1-18). Springer Berlin Heidelberg.

[28] McConnell, S. (1996). Rapid Development: Taming Wild Software Schedules. Redmond: Microsoft Press.

3. Literature Review of SNOMED CT Use

3.1 Introduction

The use of free text and local terms in electronic medical records (EMRs) is widespread and is a source of poor data quality and a barrier to semantic interoperability, data mining, secondary use of data and computerised 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 standardised, 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 six 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 EMRs.

In this study, our objective was to investigate the use of SNOMED CT by providing an overview of published studies. Whereas the Forty 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. This literature review was conducted in conjunction with Dr Francis Lau, Dr Ronald Cornet and Dr Nicolette de Keizer.

3.2 Methods

3.2.1 Identifying Papers

Searches using PubMed[a] and Embase[b] 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 **Section 12.3.4**.

3.2.2 Classification Criteria

We used a set of classification criteria similar to that used in the Forty 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 3-1**.

Table 3-1. Criteria used to classify SNOMED CT-related papers.

No	Criteria	Definition
1.	SNOMED CT Focus Category	Refers to the focus of the paper (i.e., indeterminate, theoretical, pre-development/design, implementation, evaluation/commodity).

[a] http://www.ncbi.nlm.nih.gov/pubmed
[b] http://www.embase.com

No	Criteria	Definition
2.	Usage Category	Refers to how SNOMED CT is primarily used. Each usage category belongs exclusively to one focus category. Refer to **Table 3-2** for the list of usage categories and their definitions.
3.	Medical Domain	Refers to the primary 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 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 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 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 fulfils 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 (e.g., how it can enhance the quality of care) or is used in an operational setting where 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 Forty 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 Forty 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 to "prove merit." Each of the 15 usage categories was linked to one and only one of the five SNOMED CT focus categories (refer to **Table 3-2**).

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."

Table 3-2. List of usage categories and definition, and corresponding focus category. (Status refers to the comparison with the usage categories in the Forty Year Review and indicates whether the usage category is new, is the same, was renamed or was merged.)

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 standardised terminologies	Theoretical
4.	Description of SNOMED CT and other standards	New	Describes SNOMED CT and other terminologies including technical aspects (e.g., hierarchy) and non-technical aspects (e.g., 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 standardised 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-standardised terminology systems such as local interface terminologies for content coverage.	
9.	Prospective inter-rater 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.	

3.2.3 Classifying Method

A web-based application was developed that catalogued the abstracts and papers, and enabled the co-authors to independently classify the papers. Functions were also available for the authors to compare their results with each other, add comments and review the results of papers from the Forty Year Review. The abstract was 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. One of the authors paired up with each of the other three authors to code 10 additional papers in separate sessions to ensure there was an agreement of 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.

3.3 Results

The searches on PubMed (n=537) and Embase (n=594) resulted in 702 unique papers (refer to **Figure 3-1**). 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.

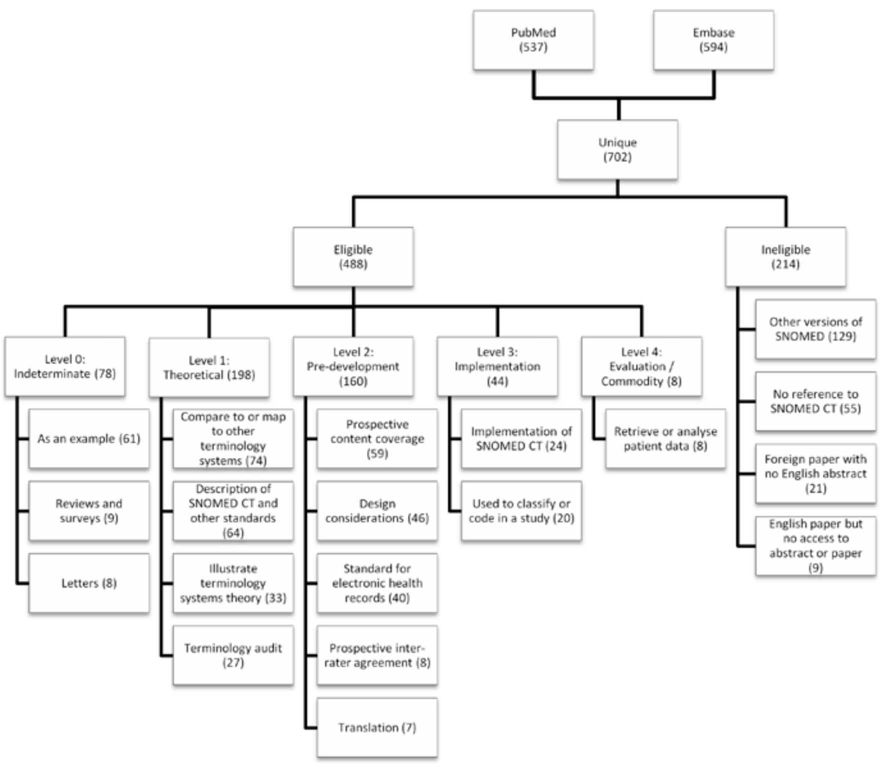


Figure 3-1. Overview of scoring of papers.

3.3.1 SNOMED CT Focus

The results of the classification of papers by SNOMED CT focus category and by year are shown in **Figure 3-2**. The number of papers classified as "theoretical" has remained relatively the same at between 11 and 15 papers over the past eight 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.

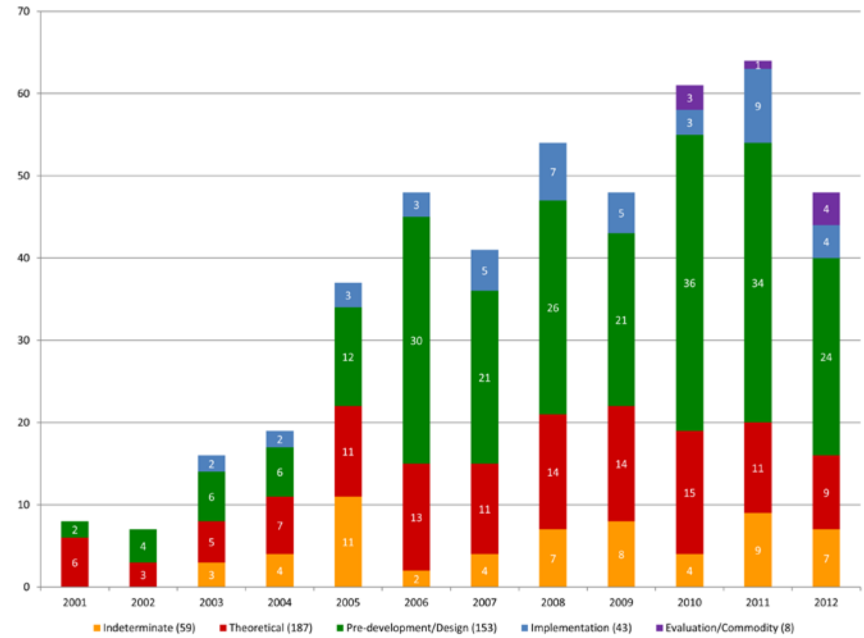


Figure 3-2. Number of papers by maturity level and year.

3.3.2 Usage Category

The results by usage category are shown in **Figure 3-1**. A further breakdown of the usage categories by subcategories is shown in **Table 3-3**. 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 standardised 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 ninth and tenth revisions (n=15), International Classification of Nursing Practice (n=6) and the Medical Dictionary for Regulatory Activities (MedDRA) (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 category in this focus category, the new usage category "terminology audit" included 24 papers, 20 of which were published in the last six 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 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 specialised 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 health care 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 was recorded in the background with SNOMED CT. Domains included cancer,[9, 10, 11, 12] pressure ulcer wounds,[13] radiology,[14] obesity,[15] and family planning.[16] Second, search boxes and auto-complete 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 Analyse 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 human immunodeficiency virus.[27] In addition to just querying for the indexed concepts, the indexed 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 (e.g., 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]

Table 3-3. 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 categorising 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 standardised terminology systems, most common were the International Classification of Diseases, Ninth and Tenth Revisions, (n=17) and International Classification for Nursing Practice (ICNP) (n=6). SNOMED CT was also compared to the Unified Medical Language System (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 1 (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

3.3.3 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 3-3.** Nursing primarily consisted of studies looking at the coverage of local nursing terms as well as standardised nursing terminologies such as ICNP.

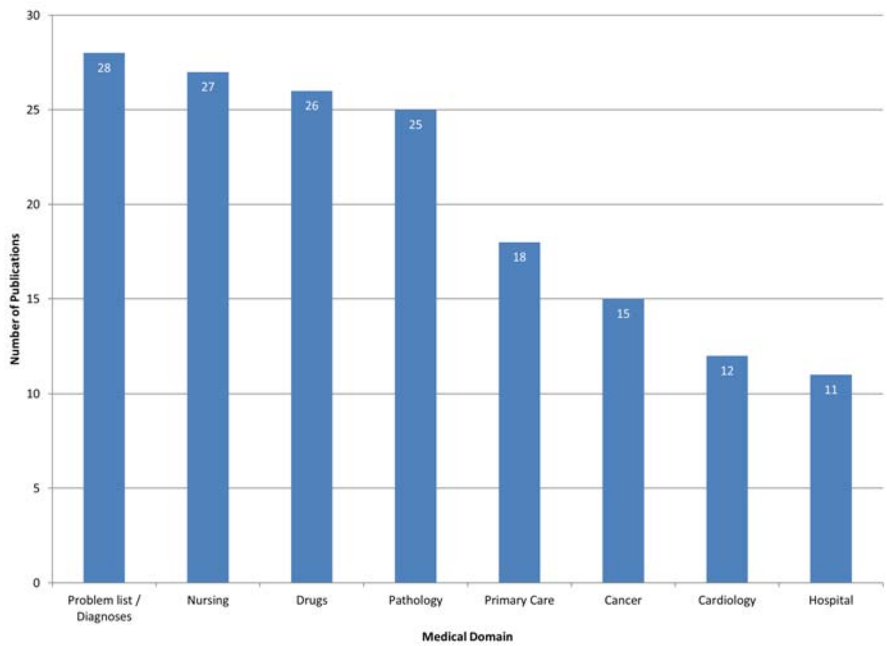


Figure 3-3. Number of papers found for each medical domain.

3.3.4 Country

The papers were from 22 countries with over half the papers coming from the United States (n=238, 53%). Refer to **Table 3-4** 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-members countries of the IHTSDO accounted for the other 13. It should be noted that the Czech Republic and Belgium have since joined the IHTSDO in 2013, after this study was conducted.

Table 3-4. Countries that belong to the IHTSDO or have published SNOMED CT-related papers in the scientific literature.

Country	Joined IHTSDO?	First Publish	# of Papers	Country	Joined IHTSDO?	First Publish	# 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

Country	Joined IHTSDO?	First Publish	# of Papers	Country	Joined IHTSDO?	First Publish	# of Papers
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	United Kingdom	2007	2001	34
Hungary		2008	1	United States	2007	2001	255
Iceland	2011	2011	2	Multiple /unknown			28
Israel	2012			**Total**			**488**
Italy		2012	1				

The number of countries that have published SNOMED CT-related papers has steadily grown over the years, with the biggest increases coming in 2007 to 2008 (refer to **Figure 3-4**). Over the past five years, papers were coming from 14 to 16 countries per year.

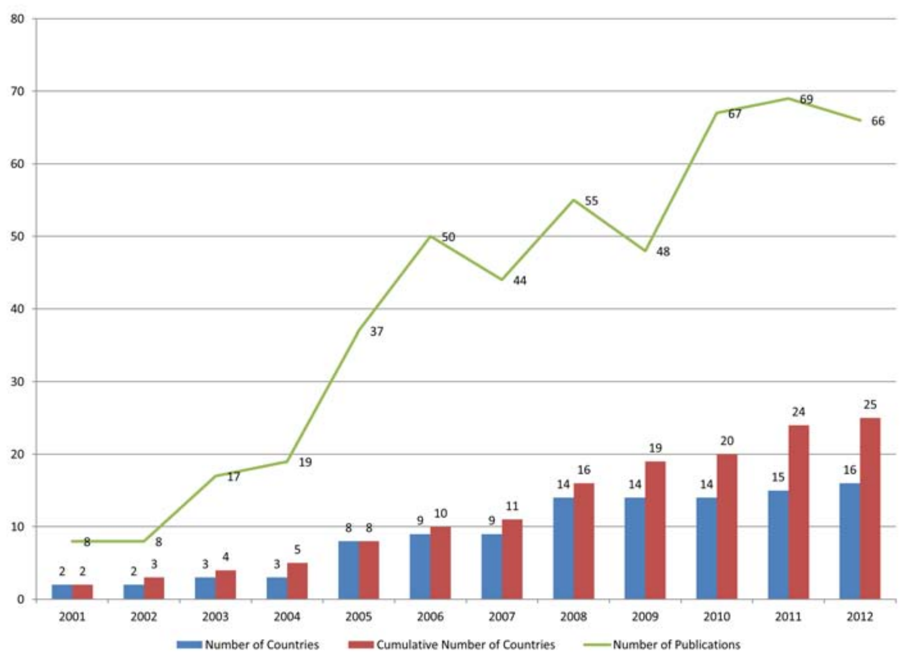


Figure 3-4. Number of papers per year by new countries, number of countries, cumulative countries and total papers.

3.4 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 six years there has been an increase in the number of SNOMED CT-related studies centring on implementation and evaluation. Thirty-seven of the forty-four papers classified as "implementation" were published over the past six years and all eight papers classified as "evaluation/commodity" were published within the past three 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 six 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 eight years and range between 11 to 15 papers each year, one usage category within this focus category has seen a steady increase. "Terminology audit," where 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.

Pre-development/Design. The use of free text is one of the barriers to computerised clinical decision support and data re-use. However, fragmented and large numbers of standardised 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 standardised terminologies is encouraging as individuals and organisations are recognising the need of harmonisation. 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 since 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 six years when SNOMED CT was released, to 34 over the last six 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 where users were only shown terms they were previously using, to the development of an interface terminology where users were exposed to over a 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 analyse patient data" 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] development of intuitive interfaces and ensuring the relevancy of search results.[20] Benefits, both realised 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 patient safety or 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 standardisation and the potential for conducting data analysis and reporting was 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 organisations are still in the process of gauging the feasibility of adopting SNOMED CT. Second, organisations 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 organisations 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 twenty-three implementations included in either or both of the IHTSDO Implementation Special Group implementation webinars[c] and Canada Health Infoway's SNOMED CT in Use website[d] have been published in the scientific literature.

It is unclear why 49 papers were retrieved when the search term "SNOMED" or "Systematised 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 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 Medical Subject Headings (MeSH) terms but the paper did not mention SNOMED. On the other hand, there are known SNOMED CT papers that are catalogued 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 X, et al.,[40] which catalogued 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 of the comparison and search strategy are available in **Appendix B**. 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.

[c] http://www.ihtsdo.org/events/conference-presentations/conference-archive/implementation-experience
[d] https://sc.infoway-inforoute.ca/standards-collaborative/snomed-ctr/snomed-ct-in-use

3.5 References

[1] Elkin, P. L., Trusko, B. E., Koppel, R., Speroff, T., Mohrer, D., Sakji, S., ... & Brown, S. H. (2010). Secondary use of clinical data. Studies in Health Technology and Informatics, 155, 14-29. http://www.ncbi.nlm.nih.gov/pubmed/20543306.

[2] Cornet, R., & de Keizer, N. (2008). Forty years of SNOMED: a literature review. BMC Medical Informatics and Decision Making, 8(Suppl 1), S2. http://www.ncbi.nlm.nih.gov/pubmed/19007439.

[3] Richesson, R. L., Andrews, J. E., & Krischer, J. P. (2006). Use of SNOMED CT to represent clinical research data: a semantic characterization of data items on case report forms in vasculitis research. Journal of the American Medical Informatics Association, 13(5), 536-546. http://www.ncbi.nlm.nih.gov/pubmed/16799121.

[4] Long, W. (2005). Extracting diagnoses from discharge summaries. In AMIA annual symposium proceedings (Vol. 2005, p. 470-4). American Medical Informatics Association. http://www.ncbi.nlm.nih.gov/pubmed/16779084.

[5] Patrick, J. D., Ryan, A., & Herkes, R. (2008). Introduction of Enhancement Technologies into the Intensive Care Service, Royal Prince Alfred Hospital, Sydney. Health Information Management Journal, 37(1), 40-5. http://www.ncbi.nlm.nih.gov/pubmed/18245864.

[6] Wade, G., & Rosenbloom, S. T. (2008). Experiences mapping a legacy interface terminology to SNOMED CT. BMC Medical Informatics and Decision Making, 8(Suppl 1), S3. http://www.ncbi.nlm.nih.gov/pubmed/19007440.

[7] James, A. G., & Spackman, K. A. (2008). Representation of disorders of the newborn infant by SNOMED CT. Studies in Health Technology and Informatics, 136, 833-8. http://www.ncbi.nlm.nih.gov/pubmed/18487835.

[8] Lee, S., Tsirbas, A., Goldberg, R. A., & McCann, J. D. (2006). Standardized terminology for aesthetic ophthalmic plastic surgery. Ophthalmic Plastic & Reconstructive Surgery, 22(5), 371-374. http://www.ncbi.nlm.nih.gov/pubmed/16985422.

[9] van Berkum, M. M. (2003). SNOMED CT® Encoded Cancer Protocols. In Amia Annual Symposium Proceedings (Vol. 2003, p. 1039). American Medical Informatics Association. http://www.ncbi.nlm.nih.gov/pubmed/14728542.

[10] Sherman, S., Shats, O., Fleissner, E., Bascom, G., Yiee, K., Copur, M., ... & Cowan, K. (2011). Multicenter breast cancer collaborative registry. Cancer informatics, 10, 217-26. http://www.ncbi.nlm.nih.gov/pubmed/21918596.

[11] Sherman, S., Shats, O., Ketcham, M. A., Anderson, M. A., Whitcomb, D. C., Lynch, H. T., ... & Brand, R. E. (2011). PCCR: Pancreatic cancer collaborative registry. Cancer informatics, 10, 83-91. http://www.ncbi.nlm.nih.gov/pubmed/21552494.

[12] Lusky, K. (2005). Pilot points way to speedier cancer surveillance. CAP today/College of American Pathologists, 19(2), 5-6. http://www.ncbi.nlm.nih.gov/pubmed/15787106.

[13] Kim, H. Y., & Park, H. (2012). Development and evaluation of data entry templates based on the entity-attribute-value model for clinical decision support of pressure ulcer wound management. International journal of medical informatics, 81(7), 485-492. http://www.ncbi.nlm.nih.gov/pubmed/22079242.

[14] Robinson, T. J., DuVall, S. L., & Wiggins III, R. H. (2011). Creation and storage of standards-based pre-scanning patient questionnaires in PACS as DICOM objects. Journal of Digital Imaging, 24(5), 823-827. http://www.ncbi.nlm.nih.gov/pubmed/20976611.

[15] Lee, N. J., & Bakken, S. (2007). Development of a prototype personal digital assistant-decision support system for the management of adult obesity. International Journal of Medical Informatics, 76, S281-S292. http://www.ncbi.nlm.nih.gov/pubmed/17606400.

[16] Zetterberg, C., Ahlzén, K., Ericsson, E., & Kron, B. (2012). An Example of a Multi-Professional Process-Oriented Structured Documentation Bound to SNOMED CT. Studies in Health Technology and Informatics, 180, 1215-7. http://www.ncbi.nlm.nih.gov/pubmed/22874405.

[17] Farfán Sedano, F. J., Terron Cuadrado, M., García Rebolledo, E. M., Castellanos Clemente, Y., & Serrano Balazote, P. (2009). Implementation of SNOMED CT to the medicines database of a general hospital. Studies in Health Technology and Informatics, 148, 123-30. http://www.ncbi.nlm.nih.gov/pubmed/19745242.

[18] Zaninelli, M., Campagnoli, A., Reyes, M., & Rojas, V. (2012). The O3-Vet project: Integration of a standard nomenclature of clinical terms in a veterinary electronic medical record for veterinary hospitals. Computer Methods and Programs in Biomedicine, 108(2), 760-772. http://www.ncbi.nlm.nih.gov/pubmed/22595264.

[19] Bakhshi-Raiez, F., de Keizer, N. F., Cornet, R., Dorrepaal, M., Dongelmans, D., & Jaspers, M. W. (2012). A usability evaluation of a SNOMED CT based compositional interface terminology for intensive care. International Journal of Medical Informatics, 81(5), 351-362. http://www.ncbi.nlm.nih.gov/pubmed/22030036.

[20] Liu, J., Lane, K., Lo, E., Lam, M., Truong, T., & Veillette, C. (2010). Addressing SNOMED CT implementation challenges through multi-disciplinary collaboration. Studies in Health Technology and Informatics, 160(Pt 2), 981-5. http://www.ncbi.nlm.nih.gov/pubmed/20841830.

[21] Cao, F., Sun, X., Wang, X., Li, B., Li, J., & Pan, Y. (2011). Ontology-based knowledge management for personalized adverse drug events detection. Studies in health technology and informatics, 169, 699-703. http://www.ncbi.nlm.nih.gov/pubmed/21893837.

[22] Arnot-Smith, J., & Smith, A. F. (2010). Patient safety incidents involving neuromuscular blockade: analysis of the UK National Reporting and Learning System data from 2006 to 2008. Anaesthesia, 65(11), 1106-1113.http://www.ncbi.nlm.nih.gov/pubmed/20840604.

[23] Benoit, S. R., L Clifford McDonald, M. D., English, R., & Tokars, J. I. (2011). Automated surveillance of Clostridium difficile infections using BioSense. Infection Control and Hospital Epidemiology, 32(1), 26-33. http://www.ncbi.nlm.nih.gov/pubmed/21128815.

[24] Nguyen, A., Moore, J., Zuccon, G., Lawley, M., & Colquist, S. (2012). Classification of pathology reports for cancer registry notifications. Studies in health technology and informatics, 178, 150-6. http://www.ncbi.nlm.nih.gov/pubmed/22797034.

[25] Matheny, M. E., FitzHenry, F., Speroff, T., Green, J. K., Griffith, M. L., Vasilevskis, E. E., ... & Brown, S. H. (2012). Detection of infectious symptoms from VA emergency department and primary care clinical documentation. International journal of medical informatics, 81(3), 143-156. http://www.ncbi.nlm.nih.gov/pubmed/22244191.

[26] Liaw, S. T., Chen, H. Y., Maneze, D., Taggart, J., Dennis, S., Vagholkar, S., & Bunker, J. (2012). Health reform: Is routinely collected electronic information fit for purpose? Emergency Medicine Australasia, 24(1), 57-63. http://www.ncbi.nlm.nih.gov/pubmed/22313561.

[27] Koopman, B., Bruza, P., Sitbon, L., & Lawley, M. (2012). Towards semantic search and inference in electronic medical records: An approach using concept-based information retrieval. The Australasian medical journal, 5(9), 482-8. http://www.ncbi.nlm.nih.gov/pubmed/23115582.

[28] Hussain, F., Muller, F., Husain, E. (2010). Under-reporting of invasive malignant melanomas in North East of Scotland. British Journal of Dermatology. 163: 67-.

[29] Spackman, K. A. (2005). Rates of Change in a Large Clinical Terminology: Three Years Experience with SNOMED Clinical Terms. In AMIA Annual Symposium Proceedings (Vol. 2005, p. 714-8). American Medical Informatics Association. http://www.ncbi.nlm.nih.gov/pubmed/16779133.

[30] Wade, G., & Rosenbloom, S. T. (2009). The impact of SNOMED CT revisions on a mapped interface terminology: terminology development and implementation issues. Journal of biomedical informatics, 42(3), 490-493.http://www.ncbi.nlm.nih.gov/pubmed/19285570.

[31] Ingenerf, J., Reiner, J., & Seik, B. (2001). Standardized terminological services enabling semantic interoperability between distributed and heterogeneous systems. International Journal of Medical Informatics, 64(2), 223-240. http://www.ncbi.nlm.nih.gov/pubmed/11734388.

[32] Hardiker, N. R., Casey, A., Coenen, A., & Konicek, D. (2006). Mutual enhancement of diverse terminologies. In AMIA Annual Symposium Proceedings (Vol. 2006, p. 319-23). American Medical Informatics Association. http://www.ncbi.nlm.nih.gov/pubmed/17238355.

[33] Park, H. A., Lundberg, C. B., Coenen, A., & Konicek, D. J. (2009). Evaluation of the content coverage of SNOMED-CT to represent ICNP Version 1 catalogues. Studies in health technology and informatics, 146, 303-7. http://www.ncbi.nlm.nih.gov/pubmed/9592854.

[34] Richesson, R., Young, K., Guillette, H., Tuttle, M., Abbondondolo, M., & Krischer, J. (2006). Standard terminology on demand: facilitating distributed and real-time use of SNOMED CT during the clinical research process. In AMIA Annual Symposium Proceedings (Vol. 2006, p. 1076). American Medical Informatics Association. http://www.ncbi.nlm.nih.gov/pubmed/17238695.

[35] Tvede, I., Bredegaard, K., & Andersen, J. S. (2010). Quality improvements based on detailed and precise terminology. Studies in Health Technology and Informatics, 155, 71-7. http://www.ncbi.nlm.nih.gov/pubmed/20543312.

[36] Elkin, P. L., Froehling, D., Wahner-Roedler, D., Trusko, B., Welsh, G., Ma, H., ... & Brown, S. H. (2008). NLP-based identification of pneumonia cases from free-text radiological reports. In AMIA Annual Symposium Proceedings (Vol. 2008, p. 172-6). American Medical Informatics Association. http://www.ncbi.nlm.nih.gov/pubmed/18998791.

[37] Lee, D., Cornet, R., Lau, F., & De Keizer, N. (2013). A survey of SNOMED CT implementations. Journal of Biomedical Informatics, 46(1), 87-96. http://www.ncbi.nlm.nih.gov/pubmed/23041717.

[38] Johansson, I. (2006). Bioinformatics and biological reality. Journal of Biomedical Informatics, 39(3), 274-287. http://www.ncbi.nlm.nih.gov/pubmed/16198638.

[39] Cimino, J. J. (2006). In defense of the Desiderata. Journal of biomedical informatics, 39(3), 299-306. http://www.ncbi.nlm.nih.gov/pubmed/16386470.

[40] Zhu, X., Fan, J. W., Baorto, D. M., Weng, C., & Cimino, J. J. (2009). A review of auditing methods applied to the content of controlled biomedical terminologies. Journal of biomedical informatics, 42(3), 413-425. http://www.ncbi.nlm.nih.gov/pubmed/19285571.

[41] Rajeev, D., Staes, C. J., Evans, S. R., Mottice, S., Rolfs, R., Samore, M. H., ... & Huff, S. M. (2010). In response to letter to the editor:'Concerning SNOMED-CT content for public health case reports'. Journal of the American Medical Informatics Association, 17(5), 613-614. http://www.ncbi.nlm.nih.gov/pubmed/20842802.

4. A Survey of SNOMED CT Implementations

4.1 Introduction

Countries such as the United States, United Kingdom, Canada, New Zealand and Australia have designated the Systematised Nomenclature of Medicine Clinical Terms (SNOMED CT) as the recommended clinical reference terminology for clinical information systems[a] (CIS).[1, 2, 3, 4, 5] However, despite the reported use of SNOMED CT in over 50 countries,[6] there are still few details on how SNOMED CT is implemented.

The International Health Terminology Standards Development Organisation (IHTSDO) Technical Implementation Guide defines three types of implementation: clinical records, knowledge representation, and aggregation and analysis.[7] Clinical records refer to the handling of patient data and include services such as recording, storing, retrieving and communicating SNOMED CT-enabled data in a CIS. Knowledge representation refers to expressing clinical knowledge such as clinical guidelines and care pathways in SNOMED CT. Aggregation and analysis refers to the retrieval of data from CIS for the purpose of secondary analysis. For the purpose of this study, we define implementation as the design, use and maintenance of SNOMED CT in the context of a CIS. "Design" refers to compiling subsets, developing data entry interfaces, programming search algorithms, selecting a data storage method, incorporating cross maps and developing data retrieval functions. "Use" refers to clinicians interacting with the CIS, receiving training and accepting the system. "Maintenance" addresses the continued updating of SNOMED CT-enabled CIS.

Two recent online questionnaires conducted by the IHTSDO[8] and Elhanan, et al.,[9] have shed light on SNOMED CT implementation. However, these questionnaires did not probe into the finer implementation details such as data entry, storage, retrieval methods and local maintenance. While there are many publications that describe technical aspects of SNOMED CT, compare the content coverage of SNOMED CT to other term sets and report the use of SNOMED CT in research studies,[10] our search for publications describing how SNOMED CT is implemented in clinical settings yielded few results. Implementation publications focused mostly on data capture (e.g., structured data entry templates,[11] clinical documentation,[12, 13, 14] synoptic checklists,[15] questionnaires,[16] indexing clinical notes[17]) and to a lesser extent data retrieval through the use of synonyms (e.g., clostridium difficile Infections,[18] neuromascular blockade[19]) and decision support (e.g., nursing interventions for pressure ulcer wounds,[52] detecting adverse drug events[20]). While there was a fair amount of detail on how SNOMED CT concepts were used in clinical systems, there were fewer details on how those concepts were stored and what methods were used to facilitate retrieval and decision support.

Implementing SNOMED CT is still relatively new and is a challenging proposition; therefore, the motivation behind this study was to conduct interviews with individuals who have implemented SNOMED CT to explore

---

[a] This is a generic label that includes electronic medical records (EMRs), electronic health records (EHRs) and electronic patient records (EPRs) used by the interviewees.

methodologies used to derive SNOMED CT subsets and extensions, how data is entered, stored and retrieved, the use of post-coordination, cross maps, maintenance processes and policies, and lessons learned. This study was conducted in conjunction with Dr Francis Lau, Dr Ronald Cornet and Dr Nicolette de Keizer.

The rest of this chapter is organised into four sections. First, the materials and methods used in our study. Second, we present the results of the interviews, which include descriptions of the SNOMED CT implementations, challenges, success factors and benefits of SNOMED CT described by the interviewees. Third, we discuss the issues raised and describe the steps towards a successful implementation. Finally, we end with a conclusion.

4.2 Materials

Survey Instrument. We used the Delphi method to select the interview questions. First, we compiled questions from previous SNOMED CT use questionnaires[8, 9] and derived questions pertaining to the Centre for Health Information Research and Development's (CHIRAD) Solutions Support Model.[21] Second, each co-investigator individually reviewed the questions, added questions of interest and prioritised the questions by importance. Third, we discussed together the validity of the questions, combined duplicate questions and after another round of individual prioritisation, pared down the number of questions to 36, spread across 10 sections based on consensus. The first three sections dealt with background information, lessons learned and usability while the last seven sections centred on specific areas of design, use and maintenance.

The questions were submitted to the IHTSDO Chief Implementation and Innovation Officer (CIIO) for feedback and I conducted pilot tests with the other three co-authors, all of whom have been involved in SNOMED CT implementation in clinical settings.

4.3 Method

4.3.1 Recruitment

Our process to recruit individuals involved in SNOMED CT implementation was via email and was carried out in four stages.

First, we invited individuals who responded to the "IHTSDO Survey to Gather the Use, Benefits and Tools of SNOMED CT."[8] One question in the survey was "Would you be willing to accept a follow up regarding the topics of this survey?" As the survey included a confidentiality clause, we contacted the IHTSDO CIIO, who sent out email invitations to the survey respondents to participate in our study.

Second, we invited individuals who were listed on Canada Health Infoway's SNOMED CT in use website.[22] The publicly accessible website provides a summary of 18 initiatives and contact information. The statuses of the initiatives listed included two that were "on hold/research initiative", seven that were "in progress/ongoing/development" and nine that were "implemented." As the catalogue of initiatives was last updated in June 2011, we contacted five individuals whose initiatives were not listed as "implemented" to check on the statuses but they responded that their initiatives were still in progress.

Third, we invited six individuals who made presentations at the IHTSDO Implementation Special Interest Group webinars.[23] The presentations ranged from theoretical demonstrations of SNOMED CT to pre-development work (e.g., the development of reference sets) to the implementation of pilot projects and production systems.

Lastly, we invited other individuals whom we knew had implemented SNOMED CT in clinical settings but were not part of the first three stages of recruitment. The main inclusion criterion was that the individuals had implemented SNOMED CT in a clinical setting.

4.3.2 Interviews and Analysis

Participants agreed to the terms of the participation consent form as part of the human ethics requirement at the University of Victoria (Protocol Number 11-535). Doodle was used to schedule individual interviews and participants were interviewed by phone or through Skype.

All interviews were digitally recorded and transcribed to aid in the inductive content analysis. We developed a secure website whereby the co-investigators could review the transcripts, mark-up the transcripts using a colour coding scheme for the inductive content analysis and to add notes.

4.4 Results

4.4.1 Subject Characteristics

Of the 50 invitations sent out (IHTSDO survey: 30; Infoway catalogue: 9; IHTSDO webinars: 6; Other individuals: 5), 13 interviews were conducted with 14 individuals (IHTSDO survey: 9; Infoway catalogue: 2; IHTSDO webinars: 2; Other individuals: 3) for a 28% response rate. Six other individuals initially expressed interest in participating but did not respond to follow-up emails. One other organisation also expressed interest but their current use of SNOMED CT was in mapping and not implementation. The interviews were conducted over a seven week period between February and April 2012. The participants were from eight countries and included physicians, academics, clinical terminologists, software developers and vendors.

4.4.2 Description of Implementations

As most interviewees had worked on multiple SNOMED CT implementations, we focused on the most mature implementation or one that the interviewee thought was most beneficial to this study. Refer to **Table 4-1** for a summary of the projects. Of the 13 projects discussed, 2 were pilot projects, 2 were under development, and 2 were currently being implemented while 7 were production systems. The production systems had been in place between three and ten years. It should be noted that the two pilot projects and one production system are no longer in use. The first pilot project was discontinued upon completion and funding was not available to proceed to a second phase while the second pilot project was discontinued after the trial period due to disappointing usability results. In the third project, the vendor had left the country and the organisation had to switch to a new vendor.

The domains included ambulatory care, intensive care, palliative care, primary care and specialist care. The extent of encoding included problems and complaints, signs and symptoms, past medical history, patient summary, allergies, metastasis, reason for admission, procedures, various reports (e.g., radiology, pathology) and partial laboratory results.

Table 4-1. Summary of results of interviews.

Inter-view	Type	Use	Clinical Domain	Extent of Encoding	Develop Own Subsets?	Uses Cross Maps?	Data Storage Design	Retrieval	Post-coordination
#1	Academic	Pilot (1 year)	Palliative care	Consult letters (i.e., problems, metastasis and past medical history)	Palliative care (2k concepts)	No	Interface Terminology	Structural subsumptio n for study purpose only	Limited qualification for end-users
#2	Vendor	Production (6 years)	Hospital-wide	Problem list	Hierarchy subsets, 4k concepts	ICD-10	SNOMED CT Concept Id	For reports, not using hierarchy	Not needed
#3	Vendor	Implemen-tation	Personal health records	Patient summary, problem list, allergies	No, but uses CORE Problem List, VA/KP Problem List	ICD-9, UMLS	Interface Terminology	Semantic search using SNOMED CT relationship s	Not used
#4	Health Care Enterpris e	Developmen t	Hospital-wide	Problem list, laboratory	5,000 subsets (range five to several thousand)	No	Interface Terminology	Enumerated list	For terminology team only
#5	Vendor	Implemen-tation	Hospital-wide	Problem list, allergies	No	No	Interface Terminology	No	Not used
#6	Govern-ment	Developmen t	Primary care EMR	Problem list, allergies	Yes, unknown, diagnosis up to 70,000	Plan on using ICD-10	SNOMED CT Concept Id	Enumerated list	Not used
#7	Academic	Production		Hospital-wide reports (e.g., radiology, pathology)	Not this particular project	No	SNOMED CT Concept Id	Class expansion, enumerated list	Not used
#8	Health Care Enterpris e	Production (6 years)	Ambulatory Care, Hospital-wide	Complaints, past medical history, signs and symptoms	1.5k concepts	Plan on using ICD-10	SNOMED CT Concept Id	Hierarchy	Yes, but no further details
#9	Academic	Pilot (1 year)	Intensive care	Reason for admission	Intensive care unit (~83k concepts)	No	SNOMED CT Expression, description id	No	Refinement and qualification for end-users
#10	Vendor	Production	Hospital-wide		8 main subsets	No	SNOMED CT Extension Description Id	Hierarchy, attributes, enumerated list	For terminology team only
#11	Health Care Enterpris e	Production (3 years)	Primary care EMR	Problem list, diagnosis, procedures	No	Would like to use ICD-9-CM	SNOMED CT Description Id	Enumerated lists	Limited qualification for end users

Inter-view	Type	Use	Clinical Domain	Extent of Encoding	Develop Own Subsets?	Uses Cross Maps?	Data Storage Design	Retrieval	Post-coordination
#12	Health Care Enterprise	Production (3 years)	Inter-disciplinary Practice	Chief complaint, past medical history, social history, family history, physical exam	No	ICD-9-CM	Probably SNOMED CT Description Id	Enumerated lists	Limited qualification for end users
#13	Health Care Enterprise	Production (>10 years)	Hospital-wide	Diagnoses	8 main subsets	ICD-9-CM	Interface Terminology	Subsumption	Not used

The reasons for using SNOMED CT included pilot implementation for proof of concept, replacing an interface terminology, a mandate to migrate from a previous standardised terminology, SNOMED CT being the best available terminology for the use case, complying with government requirements for meaningful use and certification and to facilitate decision support.

We report the results of the implementations according to our definition of implementation: design, use and maintenance. It should be noted that there are some overlaps between design and use. For example, data entry interfaces need to be developed for and used by clinicians. In this case, the emphasis is on the use as opposed to designing the interface.

4.4.2.1 Design

4.4.2.1.1 Subsets

Two types of subsets were identified: data entry and data retrieval (see **Section 4.4.2.2.3**). Data entry subsets were used to help constrain the concepts to a specific domain for use in recording patient encounters. They were used in eight projects and derived in four main ways: SNOMED CT hierarchy (n=2), local or standardised terminologies (n=3), patient records (n=3) and expert selection (n=2). Refer to **Table 4-2** for a summary of the data entry subsets and extensions. In addition to developing subsets, one project used the Clinical Observations Recording and Encoding (CORE) Problem List subset[b] and the Veterans Health Administration and Kaiser Permanente (VA/KP) Problem List subset.[c]

SNOMED CT Hierarchy. This refers to limiting the concepts to a certain hierarchy or to subtypes of a (set of) concept(s). Examples include limiting the scope to **64572001|Disease (disorder)|** and **71388002|Procedure (procedure)|**.

Local or Standardised Terminologies. Subsets were derived by mapping a local interface terminology or (inter)national classification to SNOMED CT. For example, an intensive care unit (ICU) classification that contained

[b] http://www.nlm.nih.gov/research/umls/Snomed/core_subset.html
[c] http://www.nlm.nih.gov/research/umls/Snomed/snomed_problem_list.html

over 450 terms for reasons for ICU admission was mapped to SNOMED CT. In addition, all subtype concepts of those mapped concepts were also included in the subset. For example, the term "heart attack" was mapped to **22298006|Myocardial infarction (disorder)|**, which contained nearly 60 subtype concepts.

Patient Records. Historical electronic patient records recorded using free text were analysed and mapped to SNOMED CT. In one project, over 10 million patient records were analysed. Normalisation techniques and encoding algorithms were applied and terms that occurred at least 10 times. The resulting subset contained over 20,000 unique descriptions.

Expert Selection. Clinicians listed terms that they wanted and worked with terminology experts, who identified corresponding SNOMED CT concepts and suggested additional terms, such as sibling or child concepts. For example, clinicians listed the commonly used body sites for measuring blood pressure.

Table 4-2. Subsets and extensions.

Inter-view	Domain	Source	Hierarchies	Concepts	Extensions
#1	Palliative care	IT, encoded historical patient records for problems at referral and diagnosis	Clinical findings	~2,000	Local only
#2	Hospital	SNOMED CT hierarchical subsets	Clinical findings, disorders, procedures and events		None
#4	Problem list and 5,000 value sets	IT, historical records and expert feedback	All hierarchies	Each range from 5 to several thousand	Local only
#8	Various	Interviewed own doctors, reviewed paper forms, proposed sibling and children concepts	Clinical findings	~1,500	Local only
#9	Intensive care unit reason for admission	Mapped ~450 reasons for intensive care unit admission to SNOMED CT and included all subtype concepts	Clinical findings and procedures	~83,000 plus the defining attributes of those concepts	Local only
#10	Diagnosis, signs & symptoms, drugs and substances, surgical procedures, diagnostic procedures, medical devices, obstetrics & gynaecology	Encoded over 10 million historical patient records with frequency greater than 10	All hierarchies	~20,000	Yes, using namespace
#12	Haematology and oncology, cardiology, ophthalmology, musculoskeletal, neurology, mental health, other	Clinician selection		~75,000	Yes, using namespace

4.4.2.1.2 Extensions

SNOMED CT extensions are formal additions to the SNOMED CT core using a namespace, which is a unique identifier that is assigned to individuals or organisations.[d] Two organisations created SNOMED CT extensions that were used to replicate components in the SNOMED CT core as an interface terminology and to create new components

[d] http://www.ihtsdo.org/develop/namespaces

that did not exist in the core. One organisation submitted extensions to the National Library of Medicine for possible inclusion into the national extension and SNOMED CT core while the other felt it was unnecessary as the extensions they created were highly localised. Both organisations planned to send the parent concept that was part of the SNOMED CT core instead of extensions concepts when sending SNOMED CT data beyond their enterprise.

Five organisations used their interface terminology to create new descriptions that did not exist in SNOMED CT and therefore did not see the need to create formal extensions. The organisations used SNOMED CT extensions or IT as they could not afford to wait for six months for a new concept to be released. The example cited was during the outbreak of the severe acute respiratory syndrome (SARS) whereby clinicians needed to code the disorder immediately. In one project, clinicians needed pre-coordinated concepts that included laterality so the vendor worked with the IHTSDO to develop pre-coordinated body structure concepts that included laterality.

4.4.2.1.3 Data Storage

There were four ways in which SNOMED CT-enabled data was stored: concept ids (n=4), description ids (n=4), post-coordinated expressions (n=1) and interface terminology codes (n=5). In three projects the free text descriptions entered or selected were recorded in addition to the codes.

Concept Ids. Concept ids were recorded directly into patient records. As description ids were not recorded, the terms that were displayed back to the clinicians were the preferred terms even though a synonym may have been selected during data entry.

Description Ids. Only description ids were recorded directly into patient records. In this case, the actual description used could be displayed in the record. One project only stored description ids from a formal SNOMED CT extension.

Post-coordinated Expressions. The close-to-user form of post-coordinated expressions was stored as a text string directly into patient records.

Interface Terminology Codes. Interface terminology codes, which were mapped to SNOMED CT, were stored in the patient record. Reasons for using interface terminology codes include organisations had been using their own codes for several decades and it would be difficult to change over, and to shield themselves from the updates to SNOMED CT.

4.4.2.2 Use

4.4.2.2.1 Data Entry

SNOMED CT data was captured in five ways with projects using multiple data entry methods: drop down lists (n=3), browsing the hierarchy (n=2), auto-complete (n=12), free text with ad-hoc coding (n=1) and free text with post-hoc coding (n=1).

Drop Down List. Drop down lists were used when the number of items were small, generally less than 20. Examples included positions and body sites for measuring blood pressure.

Browsing the Hierarchy. Clinicians were able to browse through a section of the SNOMED CT hierarchy and select concepts. For example, clinicians first selected whether they wanted to record a disorder or procedure, after which the selection was filtered to the hierarchy.

Auto-Complete. Clinicians would type the first few letters of a word or words and the system would retrieve potential matches. A wide range of indexed tables and algorithms such as extensive keyword search mechanisms, spell check, the expansion of abbreviations and acronyms, word equivalency and synonym substitution (e.g., "lung" and "pulmonary") were used. The order of the search results included displaying the terms by relevancy, frequency of term used in historical records or a hotlist. The "hotlist" referred to a set of terms selected by clinicians as the terms they wanted to see first. If clinicians were unable to find the term they wanted, free text was used. Five organisations had processes in place to retrieve free text entries, which were reviewed by terminology analysts and clinicians in order to be coded. Two organisations stated that in the six years of using SNOMED CT, there were only a handful of terms that could not be found as clinicians were trained to use different synonyms in order to locate the terms they needed.

Free text with Ad-hoc Coding. This is similar to the auto-complete but instead of recording a single phrase in a single data element, the auto-complete was used in a narrative. The clinician would type a narrative and suggestions for SNOMED CT concepts would be prompted, which could be selected where needed.

Free text with Post-hoc Coding. Narratives were recorded and natural language processing algorithms were used to index the narratives with SNOMED CT concepts.

Post-coordination was used in six projects and can be grouped into limited qualification (n=3), qualification and refinement (n=1), and full post-coordination (n=2).

Limited Qualification. Clinicians were allowed to use limited qualification that centred on laterality, severity, episodicity and clinical course. The qualifiers were used as discrete data elements though it is unclear whether post-coordinated expressions were constructed in the background.

Qualification and Refinement. Clinicians were allowed to qualify and refine concepts and were shown the defining attributes of concepts once a concept was selected from the auto-complete.

Full Post-coordination. Post-coordination, including the use of Concept Model attributes, was used only by the technical team to map the interface terminology to SNOMED CT as it was felt that post-coordination was too complex for clinicians. It should be noted that an extension concept was created for each post-coordinated expression.

In one CIS, post-coordination was also done in the background for family history. For example, when a clinician typed in "acute myocardial infarction" in the family history section, a look up was performed to determine if there was a pre-coordinated concept for "family history of acute myocardial infarction." If it existed, that concept was used. If it did not, an extension was created automatically and the new concept was queued for validation by a clinical modeller. This process was transparent to clinicians.

4.4.2.2.2 Cross Maps

Cross maps are mappings between SNOMED CT to another terminology and were used by four organisations. First, the ICD-10-CM cross map, which was supplied by the National Health Services in the United Kingdom, was used to generate statistics as part of government requirements. Second, two organisations used the ICD-9-CM cross map to generate billing codes. Third, a cross map to APACHE II & APACHE IV codes was used for calculating mortality risks and for benchmarking quality of care.

Two organisations were planning on using cross maps. The first was still in the process of mapping their interface terminology while the second was in the process of implementing SNOMED CT for diagnosis (in addition to the complaints, past medical history and signs and symptoms). One other organisation wanted to use the ICD-9-CM cross map to aid the billing process but the vendor declined to implement this feature.

4.4.2.2.3 Data Retrieval

The use of retrieval functions on SNOMED CT-encoded data varied across the projects and included retrieving concepts by an enumerated set of concepts (n=7), subsumption through the SNOMED CT hierarchy (n=5) and defining attributes (n=2). Enumerated sets of concepts, or data retrieval subsets, refers to users (either clinicians or technical analysts) selecting concepts individually either by descriptions or concept/description ids. Transitive closure tables were not used when testing for subsumption as built-in database functions such as Oracle's hierarchical queries were sufficient. Description logic was also not used when testing for equivalency and subsumption. Three projects did not allow for data retrieval, one pilot project only had data retrieval functions for the investigators while clinicians in one project needed the terminology team to run the queries. The reasons for including data retrieval functionality were the need to report statistics for government purposes, to identify patients with chronic diseases, or to conduct research for clinical or educational purposes.

4.4.2.2.4 User Acceptance

Four interviewees felt that clinicians were generally in favour of using SNOMED CT as long as it did not interfere with their workflow. The use of SNOMED CT helped to demonstrate the importance of using standardised codes and consistent processes. In one project where historical records were encoded with SNOMED CT to form the basis of their interface terminology, clinicians were surprised at the poor quality of data. In most cases, SNOMED CT had been so seamlessly integrated that users were unaware that they were using SNOMED CT through an interface terminology.

4.4.2.2.5 Training

Five interviewees reported that it was difficult to ascertain the amount of training that was needed to use a SNOMED CT-enabled system because training sessions focused on CIS as a whole as opposed to just a specific segment where SNOMED CT was used. In the case where a hospital was changing from paper records to electronic records, the training centred on computer skills and general acquaintance with the CIS interface. With respect to organisations that were already using a CIS, drop down lists and auto-complete functionality were already commonplace; therefore no additional training was required. In terms of post-coordination, the qualifier values were kept as discrete data elements so clinicians did not have to select a Concept Model attribute to link the two concepts together. Only one project chose to display the full set of defining attributes for a concept to allow end-users to refine the concept. Although training was provided, clinicians still found it challenging to use the system and it was concluded that the interface was too complex. In general, organisations that had implemented SNOMED CT via interface terminology did not require any additional training for clinicians.

4.4.2.3 Maintenance

Three interviewees updated SNOMED CT every six months almost immediately, three had an allotted time period to make the switch, one made the switch once a year while two were using the same version for the past three years as the vendors had stopped supporting the product or declined to update the version. The other four projects were either pilot projects and did not encounter multiple versions, did not know what version they were using or when it was last updated.

The organisations that had updated SNOMED CT mainly checked to see if any concepts in their subsets or concepts mapped to their interface terminology were inactivated. If they were inactive, terminology analysts searched for alternative concepts and suggested them to clinicians. It is unclear how and if the organisations updated the historical patient records as the interviewees did not have that information. It was mentioned that several countries such as the United Kingdom have some form of clinical safety and legal requirements that the original code should never be deleted or altered in order to preserve them in perpetuity.

In the project that recorded patient records using description ids from an extension, if the concept linked to the description was inactivated, the description id was re-linked to another active concept. The rationale was that the textual description of the description should never change but it was permissible for the description to be re-linked to a more appropriate concept.

4.4.3 Challenges

The challenges indentified by the interviewees can be categorised as SNOMED CT quality challenges; design, use and maintenance challenges; and other challenges.

4.4.3.1 SNOMED CT Quality Challenges

The SNOMED CT quality challenges fall into four main categories: content coverage (n=5), hierarchical relationships (n=4), ambiguity of terms (n=3) and syntactic consistency (n=1).

Content Coverage. While two interviewees mentioned that they had only encountered a handful of missing terms since implementing SNOMED CT over five years ago, another interviewee estimated that SNOMED CT was missing between 1% and 15% of terms needed for any given domain. For example, there was a concept for **302203004|Wife unable to cope (finding)|** but not for "husband unable to cope." Another area pointed out by three interviewees was the lack of medications and ingredients.

Hierarchical Relationships. Interviewees expressed challenges with using the hierarchy because of missing relationships or inconsistent intermediate relationships. For example, the concepts **69973000|Vascular anomaly of eyelid (disorder)|** and **193966008|Eyelid vascular anomalies (disorder)|** both refer to a vascular anomaly of the eyelid structure, with the latter concept referring to a congenital occurrence. The two concepts, however, do not have a subsumption relationship. Traversing the hierarchy and aggregating data using the hierarchy was difficult because of variations in recording the same encounter with different concepts that may reside in different locations in the hierarchy. The complexity of the hierarchy as well as missing concepts led one user to admit that they did not completely trust the hierarchies and that a healthcare organisation cannot solely depend on the hierarchies to develop their decision support features.

Ambiguity of Terms. There were synonyms in SNOMED CT that had the same description but referred to different concepts. One organisation admitted to making coding mistakes in the early stages because they did not know how to distinguish between the nuances in SNOMED CT but had improved through training. Incidentally two interviewees cited the same example of "cold," whereby it could refer to a "common cold" (**82272006|Common cold (disorder)|**) or "cold injury" (**11925005|Effects of reduced temperature (disorder)|**). While only the former includes a synonym of "cold," this can still cause confusion to clinicians who interpret that to mean the latter. Another organisation dealt with this issue by creating an exclusion subset of ambiguous descriptions.

Syntactic Consistency. There were inconsistencies in which punctuation such as hyphens, full stops and commas were used in SNOMED CT. In addition, there were different linguistic styles and a mixture of acronyms in descriptions. For example, there was no single description of "lung cancer." The fully specified name was **755174012|Malignant tumor of lung (disorder)|**, the preferred term was **482515017|Malignant tumor of lung|** and there was a synonym of **1228498010|CA - Lung cancer|** but there was no synonym for "lung cancer" by itself. This inconsistency was challenging when developing auto-complete and natural language processing algorithms.

Other SNOMED CT Quality Challenges. Other challenges mentioned include the lack of translation to other languages, the challenge of handling metonymy and relationships using the current flavour of description logic and the lack of cross maps for ICD-10-CM, which is currently under development.

4.4.3.2 Implementation Challenges

There were three main types of SNOMED CT-related implementation challenges mentioned by the interviewees: post-coordination (n=7), subsets (n=4), and data retrieval (n=4).

Post-coordination. First, the interviewees did not have a good strategy on how to design a post-coordination interface that was intuitive and unobtrusive. Second, clinicians were not willing to split their input into separate terms. The example cited was "respiratory failure due to pneumonia." Separating it into "respiratory failure" and "pneumonia" would lose the context that "respiratory failure" was the result of "pneumonia." Requiring the clinician to select the Concept Model attribute **42752001|Due to (attribute)|** to bridge the two concepts was deemed cognitively taxing and time-consuming. Third, there were concerns as to whether post-coordinated expressions could be tested for equivalency and subsumption with pre-coordinated concepts and how ICD codes could be retrieved for post-coordinated expressions.

Subsets. The main challenge was how to craft a subset for domains that were broad (e.g., reason for admission) as concepts could not be easily restricted to a hierarchy or parts of a hierarchy. For example, not all reasons for admission were diagnoses; they could also include events and monitoring after procedures. Interviewees expressed concern for the lack of clear subset development methodologies and felt that the IHTSDO should provide more guidance. Interviewees who have had experience developing subsets suggested starting from a domain as restricted as possible and working towards more complex ones.

Data Retrieval. Data retrieval using the hierarchy was challenging for three reasons. First, the hierarchy is constantly changing with each release version of SNOMED CT; therefore the answers to clinicians' queries could change over time. Second, the hierarchies were not always conductive for data aggregation as there were missing hierarchical relationships and intermediate concepts which made it difficult to select appropriate concepts for "roll up." Therefore clinicians could end up with unexpected results.

4.4.3.3 Other Challenges

Other challenges identified by the interviewees that were not directly related to SNOMED CT content are described here.

Change Resistance. There were clinicians who balked at using SNOMED CT as they feared it would interfere with their patients consults. For clinicians who had not heard of SNOMED CT or had only heard negative comments, it was important to introduce them to SNOMED CT, expose them to the benefits of using SNOMED CT, and made aware of the deficiencies of their current coding scheme.

Coding Granularity. There was a need to recognise that not everything could be coded and that there was still a place for free text. For example, consult or referral letters had to include narratives and could not just be a list of codes as they needed to be used for legal or insurance purposes. As one interviewee stated, "the whole world is not a template you tick off" (*Interviewee #12*).

Policies. The lack of policies was a barrier as vendors did not necessarily see the benefits of incorporating a new complex terminology when there was no clear government mandate. The example cited was in a country where SNOMED CT was identified as the most suitable clinical terminology for 24 sub-domains but there was a lack of promotion of SNOMED CT or how it should be used. The provincial EMR adoption certification programs bear little references to how SNOMED CT should be implemented.

4.4.4 Success Factors

The success factors described by the interviewees fall into five categories: simplicity, clinician involvement, expertise and collaboration, demonstrate value and training.

Simplicity. The most common success factor was to keep the user interface simple for clinicians by hiding the complexities of SNOMED CT. As one interviewee described it, "our model is simplicity" (*Interviewee #2*). In one pilot project that did not adequately hide the complexities of SNOMED CT, clinicians were confused with the number of options available. Multiple references were made about how simple Google was to use and how suggestions and relevant results were provided quickly.

Clinician Involvement. Engaging clinicians during the development phase to solicit their input and act upon their suggestions and concerns were important to gain their support. Interviewees felt that migrating to SNOMED CT was easier when there were clinicians who understood the value of having a longitudinal electronic health records and had positions of decision making influence.

Expertise and Collaboration. Having a terminology team comprised of terminology experts, analysts, clinicians and programmers was necessary to ensure both clinical and technical viewpoints were represented. One the interviewees pointed to the lack of a reference implementation as one of the pitfalls of implementation and that it was important to contact other organisations that have implemented SNOMED CT to learn from their experience.

Demonstrate Value. There was a need to demonstrate immediate value to clinicians for using SNOMED CT and the value depended on the maturity level of the implementation. Having a legible patient record was adequate for first time CIS users while experienced CIS users required more functionality such as decision support features.

Training. The amount of training needed varied depending on the stage an organisation was with their CIS implementation. The organisation that changed from a paper-based system to a CIS required more training and was a central success factor compared to other organisations that already has been using a CIS. For organisations that had just made the switch to CIS, the key to training was to start in very small areas and add complexities later on once the foundation was there.

4.4.5 Benefits of Using SNOMED CT

None of the initiatives had carried out extensive evaluations to determine the benefits of using SNOMED CT. Reasons included SNOMED CT applications were still being developed or implemented and they had other priorities or had no capacity to carry out evaluations. The interviewees were, however, able to describe some of the benefits they had observed: direct data entry (n=4), data reuse (n=4), content coverage and subset development (2) and legibility (n=1).

Direct Data Entry. The large number of synonyms in SNOMED CT enabled clinicians to record the exact diagnosis they had in mind in contrast to post-coding whereby a terminology analyst reviewed the free text entry and selected a concept that he/she thought was appropriate. An interviewee believed that direct data entry through the intuitive interface increased the accuracy and speed in which records were coded.

Data Reuse. Organisations that used SNOMED CT in conjunction with ICD cross maps were able to re-use the data that was captured via SNOMED CT to generate ICD codes for billing and statistical reports. One organisation profited from using SNOMED CT by being able to identify patients with certain clinical conditions for referral to clinical trials conducted by pharmaceutical companies and universities. Another way in which SNOMED CT helped to generate revenue was by enabling clinicians to identify patients with chronic diseases such as diabetes mellitus and congestive heart failure in chronic disease management programs that included remuneration for registration.

Content Coverage and Subset Development. Interviewees felt that SNOMED CT provided them with the best content coverage for their use cases compared to other terminologies and that it was still a good starting point for developing subsets despite the issues of missing concepts and relationships.

Legibility. While not directly related to SNOMED CT, the legibility of the patient record was one of the immediate benefits in one project that made the switch from a paper-based system to an electronic system using SNOMED CT. In a project where SNOMED CT was used in an existing CIS, the problem list was standardised.

4.5 Discussion

There are few publications that describe the implementation of SNOMED CT in clinical settings in detail, whereas as the majority of publications focus on comparing SNOMED CT to other terminologies and to illustrate terminology systems theory.[10] The lack of publications that span the full scope of design, use and maintenance has made it challenging for organisations that are looking for reference SNOMED CT implementations. This study attempted to fill the gap through a series of interviews that looked that the benefits, success factors, challenges and implementation approaches.

4.5.1 Towards a Successful SNOMED CT Implementation

In this section we attempt to enumerate the steps towards a successful SNOMED CT implementation drawn on lessons from this study as well as other publications. Additional details are available in **Appendix C**.

(1) Understand Process of SNOMED CT Implementation. A multi-disciplinary team[13,24] that includes terminology experts, clinicians, technical experts and project managers should be formed to understand the technical details of SNOMED CT as well as the challenges, success factors and benefits of using SNOMED CT such as those described in this study. They should also be made aware of limitations such as data quality issues[25] and the inability to adequately represent negation and disjunction.[26] In addition, SNOMED CT should be used in conjunction with an information model to represent data types such as dates and numeric values.[27]

(2) Encode Local Terms to SNOMED CT. Allowing clinicians to use an interface terminology they are familiar with may make the transition to SNOMED CT smoother.[14, 28, 29, 30] The most common source of local terms may be from currently coded value sets but historical patient records recorded with free text should also be considered as they contain a rich source of terms.[31] The interface terminology layer will help to shield end-users from changes made to SNOMED CT although the technical team should review any change made to SNOMED CT.[32, 33]

(3) Create Extensions. Where applicable, extensions should be created for concepts that do not exist in SNOMED CT.[34] While informal extensions can be created using an interface terminology, formal extensions that are defined can contribute to the semantics. Extensions should be submitted to national release centres for possible inclusion into national extensions or into the SNOMED CT core.

(4) Compile Subsets. Subsets should be compiled to constrain the relevant concepts to use cases and may be compared to publicly available subsets (e.g., VA/KP, CORE and Convergent Medical Terminology (CMT)) for content coverage. Effort should be still spent on developing search algorithms as there is still a need to be able to retrieve terms within a subset.

(5) Design Intuitive Data Entry Interfaces. Auto-complete is well-suited data entry method as opposed to long drop down lists or browsing the hierarchy.[35] Depending on the context of the data elements, hotlists and most frequently used terms should be made available. Search algorithms should include partial multi-word searches,[36] spelling corrections, word equivalency, and abbreviation and acronym expansion. The use of templates can also help to facilitate the recording of data into the correct data elements.[37, 38]

(6) Select a Data Storage Method. Interface terminology codes and SNOMED CT ids are the two main data storage methods and each has its advantages and disadvantages. The use of interface terminology codes in patient records is less invasive method as minimal changes will be made to the database backend and historical patient records can remain unchanged. Any changes made to SNOMED CT will only need to be reflected in the mapping tables. On the other hand, storing SNOMED CT ids directly will enable decision support systems to be supported directly without the need of interface terminology mapping tables and will not require the maintenance of the interface terminology.

(7) Incorporate Cross Maps. Cross maps from SNOMED CT to classification systems such as ICD-9-CM are needed to facilitate reimbursement processes and the generation of statistical reports without the need to code multiple times[39] and to allow for the continuity of historically records coded with classification systems.[40]

(8) Design Retrieval Functions. Data retrieval functions should be included to harness the rich semantics in SNOMED CT and enable clinicians to run patient case queries and to facilitate decision support. Organisations that do not have the capacity to develop advanced retrieval functionality can consider using third party terminology services or application programming interfaces (APIs) such as the open-source Apelon Distributed Terminology System (DTS),[e] IHTSDO Workbench[f] or National Health Services Snofyre[g] to compute the aggregation of SNOMED CT concepts that can then be used in CIS.

---

[e] http://apelon-dts.sourceforge.net
[f] http://www.b2international.com/portal/ihtsdo-workbench

(9) Conduct Training Sessions. Training sessions should be conducted as not all clinicians have the same technical capabilities[41] and it provides a forum where clinicians can be educated on new features and have the opportunity to have their questions answered. If data entry interfaces are intuitive and search algorithms are effective and efficient, the amount of training needed may be reduced.[42]

(10) Develop Maintenance Policies. Two types of maintenance policies should be in place. First, unencoded terms should be routinely extracted for review to determine if they can be coded. Second, when a new version of SNOMED CT is released, the mappings between SNOMED CT and the interface terminology and classification systems should be reviewed to ensure that only active concepts are used.

4.5.2 Incremental Value of SNOMED CT

The promoted benefits of using SNOMED CT[43] were generally not realised to its full potential in the projects we examined. It should be noted that none of the organisations have conducted full-scale evaluations so it is premature to come to a decisive conclusion. Granted two were pilot projects and a few were still early in the implementation stages, but the organisations that had been using SNOMED CT for at least three years have not reported significant increment value.

While considerable benefits have been realised such as being able to query patients with specific clinical conditions, record clinical records at a granular level and automatically generate ICD codes for statistical reports, it can be argued that these benefits may have been realised without the use of SNOMED CT. For example, the organisation that was using an interface terminology that was compiled and refined over several decades was already able to capture data at a granular level and facilitate decision support functionality, thus limiting the incremental value of using SNOMED CT. One of the reasons for using standardised terminologies is to enable interoperability; however, none of the organisations were communicating SNOMED CT data beyond their enterprise as part of routine operations. The benefit of using standardised terminologies can be achieved once (inter)national bodies formalise their guidelines and quality indicates based on SNOMED CT concepts.

We suggest some reasons for the limited benefits. First, large organisations already have robust interface terminology and will see limited incremental value until information is exchanged with external organisations. As there are very few SNOMED CT implementations in clinical care settings, it is unlikely that an organisation will transmit SNOMED CT-encoded data with other organisations in the near future. Second, the organisations that have been developing their own SNOMED CT solutions have thus far focused on data entry as opposed to data retrieval and decision support functionality and therefore are unable to demonstrate improvement in areas such as the quality of care. Third, organisations that rely on off-the-shelf vendor solutions have been unable to convince vendors to add new functionality due to the limited demand for SNOMED CT and therefore use SNOMED CT in a very basic manner.

---

[g] http://code.google.com/p/snofyre

4.5.3 Outstanding Issues

From this study, there remain three outstanding issues that require further study: how to implement post-coordination, retrieval and extensions.

Post-Coordination. Implementing post-coordination continues to be a challenge both from a graphical user interface design and clinical terminology point of view (e.g., creating clinically nonsensical concepts, concept duplication and inefficiency of concept composition). In some cases there are some alternatives to post-coordination, such as using an information model. For example, instead of including the subject relationship context in a single post-coordinated expression, the relationship is stored as a separate data element.

Retrieval. While retrieving pre-coordinated concepts and defining attributes are relatively straightforward, retrieval of post-coordinated expressions is relatively new. There are also unresolved issues as to how to retrieve corresponding ICD codes from post-coordinated expressions. A possibility would be to test if the concept used in the mapping subsumes the post-coordinated expression but this method has been untested. Tools such as the IHTSDO Workbench and National Health Services Snofyre have the potential to simplify the retrieval process by computing complex SNOMED CT-related calculations such as the testing for equivalency and subsumption outside of a CIS, but to date there are no published studies on the effectiveness or efficiency of these tools.

Extensions. As only one of the 13 organisations consistently submitted extensions for possible inclusion into a national extension set or the SNOMED CT core, there needs to be a streamlined method of submitting extensions and to monitor the progress. It remains to be seen how the creation of extensions will affect interoperability. Interviewees who stated they created formal extensions were asked this question and the answer given was that the parent concept from the SNOMED CT core would be transmitted instead. The implications of using the parent concept instead of the extension concept are unclear.

4.6 References

[1] SNOMED CT. http://www.nlm.nih.gov/research/umls/Snomed/snomed_main.html. Last accessed: May 6, 2012.

[2] SNOMED CT: A standard clinical terminology is essential for the interoperability of electronic health records across care settings. http://www.connectingforhealth.nhs.uk/systemsandservices/data/uktc/snomed. Last accessed: April 11, 2012.

[3] SNOMED CT Canada. https://sl.infoway-inforoute.ca/content/disppage.asp?cw_page=snomedct_e. Last accessed: March 6, 2012.

[4] SNOMED Endorsement. http://ww.ithealthboard.health.nz/sites/all/files/SNOMED CT Endorsement.doc. Last accessed: April 11, 2012.

[5] SNOMED CT Australia. http://www.nehta.gov.au/connecting-australia/terminology-and-information/clinical-terminology/snomed-ct-au. Last accessed: May 6, 2012.

[6] Meaningful Use Quality Performance Measures Benefit from New SNOMED CT "Public Good" Use Policy. http://www.ihtsdo.org/news/article/article/meaningful-use-quality-performance-measures-benefit-from-new-snomed-ct-public-good-use-policy. Last accessed: May 6, 2012.

[7] International Health Terminology Standards Development Organisation. Technical Implementation Guide. http://www.ihtsdo.org/fileadmin/user_upload/doc/tig/index.html.

[8] Results of the Survey to Gather the Use, Benefits and Tools of SNOMED CT. http://www.ihtsdo.org/fileadmin/user_upload/Docs_01/Publications/Implementation/Results_of_Survey_on_Use_of_SNOMED_V1.0._Binder.pdf. January 19, 2011.

[9] Elhanan, G., Perl, Y., & Geller, J. (2011). A survey of SNOMED CT direct users, 2010: impressions and preferences regarding content and quality. Journal of the American Medical Informatics Association, 18(Suppl 1), i36-i44. http://www.ncbi.nlm.nih.gov/pubmed/21836159.

[10] Cornet, R., & de Keizer, N. (2008). Forty years of SNOMED: a literature review. BMC Medical Informatics and Decision Making, 8(Suppl 1), S2. http://www.ncbi.nlm.nih.gov/pubmed/19007439.

[11] Kim, H. Y., & Park, H. (2012). Development and evaluation of data entry templates based on the entity-attribute-value model for clinical decision support of pressure ulcer wound management. International Journal of Medical Informatics, 81(7), 485-492. http://www.ncbi.nlm.nih.gov/pubmed/22079242.

[12] Farfán Sedano, F. J., Terron Cuadrado, M., García Rebolledo, E. M., Castellanos Clemente, Y., & Serrano Balazote, P. (2009). Implementation of SNOMED CT to the medicines database of a general hospital. Studies in Health Technology and Informatics, 148, 123-30. http://www.ncbi.nlm.nih.gov/pubmed/19745242.

[13] Liu, J., Lane, K., Lo, E., Lam, M., Truong, T., & Veillette, C. (2010). Addressing SNOMED CT implementation challenges through multi-disciplinary collaboration. Studies in Health Technology and Informatics, 160(Pt 2), 981-5. http://www.ncbi.nlm.nih.gov/pubmed/20841830.

[14] Bakhshi-Raiez, F., de Keizer, N. F., Cornet, R., Dorrepaal, M., Dongelmans, D., & Jaspers, M. W. (2012). A usability evaluation of a SNOMED CT based compositional interface terminology for intensive care. International Journal of Medical Informatics, 81(5), 351-362. http://www.ncbi.nlm.nih.gov/pubmed/22030036.

[15] Lusky, K. (2005). Pilot points way to speedier cancer surveillance. CAP today/College of American Pathologists, 19(2), 5-6, 8. http://www.ncbi.nlm.nih.gov/pubmed/15787106.

[16] Robinson, T. J., DuVall, S. L., & Wiggins III, R. H. (2011). Creation and storage of standards-based pre-scanning patient questionnaires in PACS as DICOM objects. Journal of Digital Imaging, 24(5), 823-827. http://www.ncbi.nlm.nih.gov/pubmed/20976611.

[17] Patrick, J. D., Ryan, A., & Herkes, R. (2008). Introduction of Enhancement Technologies into the Intensive Care Service, Royal Prince Alfred Hospital, Sydney. Health Information Management Journal, 37(1), 40-5. www.ncbi.nlm.nih.gov/pubmed/18245864.

[18] Benoit, S. R., L Clifford McDonald, M. D., English, R., & Tokars, J. I. (2011). Automated surveillance of Clostridium difficile infections using BioSense. Infection Control and Hospital Epidemiology, 32(1), 26-33. http://www.ncbi.nlm.nih.gov/pubmed/21128815.

[19] Arnot - Smith, J., & Smith, A. F. (2010). Patient safety incidents involving neuromuscular blockade: analysis of the UK National Reporting and Learning System data from 2006 to 2008. Anaesthesia, 65(11), 1106-1113. http://www.ncbi.nlm.nih.gov/pubmed/20840604.

[20]Cao, F., Sun, X., Wang, X., Li, B., Li, J., & Pan, Y. (2011). Ontology-based knowledge management for personalized adverse drug events detection. Studies in health technology and informatics, 169, 699-703. http://www.ncbi.nlm.nih.gov/pubmed/21893837.

[21] Centre for Health Information Research and Development Implementing SNOMED CT within national electronic record solutions. http://www.chirad.org.uk/paper_one.htm. Last accessed: March 22, 2012.

[22] SNOMED CT in use. https://www.infoway-inforoute.ca/standards-collaborative/snomed-ctr/snomed-ct-in-use. Last accessed: March 6, 2012.

[23] Implementation experience. http://www.ihtsdo.org/snomed-ct/snomed-present/implementation-experience. Last accessed: March 6, 2012.

[24] Windle, J., Van-Milligan, G., Duffy, S., McClay, J., & Campbell, J. (2003). Web-based physician order entry: an open source solution with broad physician involvement. In AMIA Annual Symposium Proceedings (Vol. 2003, p. 724-7). American Medical Informatics Association. http://www.ncbi.nlm.nih.gov/pubmed/14728268.

[25] Rector, A. L., Brandt, S., & Schneider, T. (2011). Getting the foot out of the pelvis: modeling problems affecting use of SNOMED CT hierarchies in practical applications. Journal of the American Medical Informatics Association, 18(4), 432-440. http://www.ncbi.nlm.nih.gov/pubmed/21515545.

[26] Rector, A. L., & Brandt, S. (2008). Why do it the hard way? The case for an expressive description logic for SNOMED. Journal of the American Medical Informatics Association, 15(6), 744-751. http://www.ncbi.nlm.nih.gov/pubmed/18755993.

[27] Bird, L., Brooks, C., Cheong, Y. C., & Tun, N. N. (2011). A logical approach to semantic interoperability in healthcare. Studies in Health Technology and Informatics. v168, 1-9. http://www.ncbi.nlm.nih.gov/pubmed/21893905.

[28] Osornio, A. L., Luna, D., Gambarte, M. L., Gomez, A., Reynoso, G., & de Quiros, F. G. (2007). Creation of a local interface terminology to SNOMED CT. Studies in Health Technology and Informatics, 129(Pt 1), 765-9. http://www.ncbi.nlm.nih.gov/pubmed/17911820.

[29] Dolin, R. H., Mattison, J. E., Cohn, S., Campbell, K. E., Wiesenthal, A. M., Hochhalter, B., ... & Zingo, C. (2004). Kaiser Permanente's Convergent Medical Terminology. Studies in Health Technology and Informatics, 107(Pt 1), 346-50. http://www.ncbi.nlm.nih.gov/pubmed/15360832.

[30] Rosenbloom, S. T., Miller, R. A., Johnson, K. B., Elkin, P. L., & Brown, S. H. (2006). Interface terminologies facilitating direct entry of clinical data into electronic health record systems. Journal of the American Medical Informatics Association, 13(3), 277-288. http://www.ncbi.nlm.nih.gov/pubmed/16501181.

[31] Lee, D. H., Lau, F. Y., & Quan, H. (2010). A method for encoding clinical datasets with SNOMED CT. BMC Medical Informatics and Decision Making, 10(1), 53. http://www.ncbi.nlm.nih.gov/pubmed/20849611.

[32] Wade, G., & Rosenbloom, S. T. (2009). The impact of SNOMED CT revisions on a mapped interface terminology: terminology development and implementation issues. Journal of Biomedical Informatics, 42(3), 490-493. http://www.ncbi.nlm.nih.gov/pubmed/19285570.

[33] Lee, D., Cornet, R., & Lau, F. (2011). Implications of SNOMED CT versioning. International Journal of Medical Informatics, 80(6), 442-453. http://www.ncbi.nlm.nih.gov/pubmed/21450517.

[34] Hurrell M.J., Monk T.G., Nicol A., Norton A.N., Reich D.L., Walsh J.L.. (2011). Implementation of a standards-based, CDA-compliant anesthesia record. Journal of Clinical Monitoring and Computing, 25, 5-16.

[35] de Lusignan, S., Chan, T., & Jones, S. (2011). Large complex terminologies: more coding choice, but harder to find data--reflections on introduction of SNOMED CT (Systematized Nomenclature of Medicine--Clinical Terms) as an NHS standard. Informatics in Primary Care, 19(1), 3. http://www.ncbi.nlm.nih.gov/pubmed/22118330.

[36] Sevenster, M., van Ommering, R., & Qian, Y. (2012). Algorithmic and user study of an autocompletion algorithm on a large medical vocabulary. Journal of Biomedical Informatics, 45(1), 107-119. http://www.ncbi.nlm.nih.gov/pubmed/22019376.

[37] Robinson, T. J., DuVall, S. L., & Wiggins III, R. H. (2011). Creation and storage of standards-based pre-scanning patient questionnaires in PACS as DICOM objects. Journal of Digital Imaging, 24(5), 823-827. http://www.ncbi.nlm.nih.gov/pubmed/20976611.

[38] Bennett, K. J., & Steen, C. (2010). Electronic medical record customization and the impact upon chart completion rates. Family Medicine, 42(5), 338-42. http://www.ncbi.nlm.nih.gov/pubmed/20461565.

[39] Tvede, I., Bredegaard, K., & Andersen, J. S. (2010). Quality improvements based on detailed and precise terminology. Studies in Health Technology and Informatics, 155, 71-7. http://www.ncbi.nlm.nih.gov/pubmed/20543312.

[40] Nadkarni, P. M., & Darer, J. A. (2010). Migrating existing clinical content from ICD-9 to SNOMED. Journal of the American Medical Informatics Association, 17(5), 602-607. http://www.ncbi.nlm.nih.gov/pubmed/20819871.

[41] Pole, D. (2010). Electronic Patient Records in Sri Lankan Hospitals. Sri Lanka Journal of Bio-Medical Informatics, 1(1):43-45.

[42] Øvretveit, J., Scott, T., Rundall, T. G., Shortell, S. M., & Brommels, M. (2007). Improving quality through effective implementation of information technology in healthcare. International Journal for Quality in Health Care, 19(5), 259-266. http://www.ncbi.nlm.nih.gov/pubmed/17717038.

[43] SNOMED CT Benefits. http://www.ihtsdo.org/snomed-ct/whysnomedct/benefits. Last accessed: March 6, 2012.

5. CONCEPTUAL FRAMEWORKS

5.1 Introduction

This chapter describes the three conceptual frameworks that were used in this PhD research. First, the SNOMED CT Implementation Framework, which was used to guide the development of the clinical value design methodology. Second, the SNOMED CT Extensions Auditing Framework, which was used to organise the verification rules for the Systematised Nomenclature of Medicine Clinical Terms (SNOMED CT) extensions auditing. Third, the SNOMED CT Clinical Value Framework, which was used to categorise the value added benefits of using SNOMED CT.

The motivation for developing these frameworks came in part from the discussion section and outstanding issues identified in the literature review in **Chapter Three** and implementation survey in **Chapter Four** respectively. The literature review identified "terminology audit," "prospective content coverage" and "retrieve and analyse patient data" as three of the most prominent upcoming or most common usage categories, while the survey identified three outstanding issues, which were post-coordination, retrieval and extensions.

For the SNOMED CT Implementation Framework, the "terminology audit" and "extensions" aided in the formation of the auditing dimension, the "prospective content coverage" and "post-coordination" aided in the formation of the encoding dimension, while the "retrieve and analyse patient data" and "retrieval" aided in the formation of the retrieval dimension. The "extensions" and "post-coordination" also aided in supporting the need for implementation levels. The SNOMED CT Extensions Auditing Framework was developed as the result of the need to organise in a coherent manner the verification rules that were identified while the lack of a conceptual framework to categorise the clinical value of using SNOMED CT in both the literature review and implementation survey prompted the development of the SNOMED CT Clinical Value Framework. The rest of this chapter is devoted to describing these three conceptual frameworks.

5.2 SNOMED CT Implementation Framework

The SNOMED CT Implementation Framework used in this research was adapted and refined from the Centre for Health Informatics Research and Development's (CHIRAD) Solutions Support Model.[1] The Solutions Support Model consists of three levels (i.e., Level 0: no internal support for SNOMED CT; Level 1: internal support for pre-coordinated SNOMED CT; and Level 2: internal support for post-coordinated SNOMED CT) and seven dimensions (i.e., coded data storage; logical record structure; extent of encoding; data entry; data retrieval; inbound communication; and outbound communication). The levels are similar to the Health Information and Management Systems Society (HIMSS) Electronic Medical Records (EMR) Adoption Model[2] in the sense that the higher levels indicate an adoption of increasingly advanced functionality.

The SNOMED CT Implementation Framework in this research expands on the levels by introducing a new level (i.e., level 3: extensions) and consolidates the dimensions into three dimensions (i.e., auditing; encoding; and retrieval). The data entry and extent of encoding dimensions were relabelled to encoding, the data retrieval and coded data storage dimensions were relabelled to retrieval and a new dimension, auditing, was added. The inbound and outbound communication dimensions, and logical record structure were declared out of scope of this research. An overview of the implementation conceptual framework is shown in **Figure 5-1**. The rest of this section describes the four levels and three dimensions.

Scalable Conceptual Model		Auditing Method	Encoding Method	Retrieval Method
Purpose		To ensure the subsets, post-coordinated expressions and extensions are consistent.	To ensure data is encoded accurately, consistently and unambiguously.	To ensure SNOMED CT-encoded data can be retrieved effectively and efficiently.
Rationale		There are known errors in SNOMED CT that can hinder the accuracy of encoding data and retrieval.	The literature indicates a need for guidance on creating post-coordinated expressions.	There are currently very few publications on SNOMED CT queries.
LEVEL 3 Extensions	Non-Leaf Concepts Leaf Concepts Description Extensions	• Ensure new SNOMED CT concepts are modelled consistently • Ensure appropriate location in the hierarchy • Ensure new SNOMED CT concepts do not currently exist	• New concepts • New descriptions • New relationships	• Regenerate canonical table • Regenerate transitive closure table
LEVEL 2 Post-Coordination	Concept Model Refinement Qualification	• Ensure expressions do not contain syntax errors • Ensure expressions conform to the Concept Model • Ensure appropriate pairing of concepts are used	• Negation • Certainty • Finding/procedure context • Temporal context • Subject context • Merging algorithms	• Context-sensitive retrieval • Structural subsumption • Long and short normal form • Optimised data storage for post-coordinated expressions
LEVEL 1 Pre-Coordination	Subsets (RefSets) Hierarchy Unrestricted	• Ensure concepts from appropriate hierarchy are used • Ensure only active concept are used • Version control	• Data cleaning • Search algorithms (e.g., exact match, match all) • Sorting algorithms (e.g., handling multiple exact matches) • Exclusion concepts and descriptions	• Equivalency • Subsumption • Defining attributes • Hierarchy-specific • Subset-specific • Historical relationships
LEVEL 0 No SNOMED CT	Standardised Terminologies Local Codes Free Text	• Not applicable since SNOMED CT is not used at this level	• Not applicable since SNOMED CT is not used at this level	• Not applicable since SNOMED CT is not used at this level

Figure 5-1. SNOMED CT Implementation Framework.

5.2.1 Four Levels

The four levels represent an increasing complex level of adoption of SNOMED CT from no SNOMED CT support in level 0 to the use of SNOMED CT with extensions in level 3. The SNOMED CT Implementation Framework can be extended to include level 4: Information Model, which describes how clinical data encoded with SNOMED CT should be used in an EMR and how it relates to other data elements, but that is beyond the scope of this study.

5.2.1.1 Level 0: No SNOMED CT

As the name implies, SNOMED CT is not used at this level and represents the level prior to SNOMED CT implementation. Clinical statements recorded at this level may be captured in the form of free text, local codes and other standardised terminologies other than SNOMED CT. As SNOMED CT is not used, there are no applications in terms of auditing, encoding or retrieval.

5.2.1.2 Level 1: Pre-coordination

At this level, only pre-coordinated concepts are used. This limits the usefulness of SNOMED CT but represents a basic implementation of SNOMED CT. This level was further defined as unrestricted, hierarchy and subsets. In the unrestricted sub-level, there are no constraints as to which concepts should be used. In the hierarchy sub-level, only concepts from certain hierarchies are allowed. In the subsets sub-level, concepts are constrained to a list of pre-selected concepts.

5.2.1.3 Level 2: Post-coordination

At this level, post-coordinated expressions are used in addition to pre-coordinated concepts. This level was further defined as qualification, refinement and Concept Model as they represent an increasing level of complexity of forming post-coordinated expressions. The sub-level qualification involves selecting concepts that have been defined using the qualifying relationships (**CharacteristicType=1**) and are limited to eight main sets of concepts from the **362981000|Qualifier value (qualifier value)|** hierarchy (e.g., **24484000|Severe (severity modifier) (qualifier value)|**). The sub-level refinement involves the selection of a subtype concept of an existing non "is a" defining attribute so as to represent a clinical statement with a greater level of granularity. In both qualification and refinement, the constraints are tightly controlled while in the sub-level Concept Model, the range of post-coordination are broadly defined using the Machine Readable Concept Model (MRCM).

Combination post-coordination is not included as combination expressions can either be encoded as separate pre-coordinated concepts or Concept Model post-coordination. The example in the SNOMED CT Technical Implementation Guide refers to "gallstones with cholecystitis," which was represented as **235919008|Gallbladder calculus (disorder)|+76581006|Cholecystitis (disorder)|**. It is also permissible to record them as two separate concepts (i.e., **235919008|Gallbladder calculus (disorder)|** and **76581006|Cholecystitis (disorder)|**), or a post-coordinated expression such as **64572001|Disease (disorder)|:{116676008|Associated morphology (attribute)|=23583003|Inflammation (morphologic abnormality)|,363698007|Finding site (attribute)|=28231008|Gallbladder structure (body structure)|}{116676008|Associated morphology (attribute)|=56381008|Calculus (morphologic abnormality)|,363698007|Finding site (attribute)|=28231008|Gallbladder structure (body structure)|}**. It should be noted that there is also a pre-coordinated concept **25924004|Calculus of gallbladder with cholecystitis (disorder)|**.

5.2.1.4 Level 3: Extensions

Organisations, jurisdictions and countries are encouraged to develop SNOMED CT extensions when concepts and descriptions are not available to suit their needs. Extensions can take the form of new synonyms to existing concepts to completely new concepts. Only organisations that have obtained a namespace identifier can create valid

SNOMED CT identifiers. Although extensions are pre-coordinated concepts, they can be considered a more advanced implementation of SNOMED CT because extension concepts that are created need to be defined (which can be considered a form of post-coordination), checked to ensure they are not duplicate and are situated in the appropriate location in the hierarchy. In addition, new concepts and relationships that are added introduce an increased level of complexity especially if the new concepts are non-leaf concepts (i.e., they are added in the subsumption relationship between two core concepts). This level was further defined as descriptions, leaf concepts and non-leaf concepts. The sub-level description extensions refer to adding new synonyms for existing concepts. The sub-level leaf concepts refer to adding extension concepts as child concepts and not parent concepts to existing core concepts. The sub-level non-leaf concepts refer to adding extension concepts as child and parent concepts to existing core concepts.

5.2.2 Three Dimensions

The three dimensions represent different methods that are needed in order to implement SNOMED CT. They are briefly described in this section but more information is available in the next three chapters as the process of developing and testing the methods are described.

5.2.2.1 Auditing

This dimension centres on quality assurance and version control. Auditing methods can and should be applied to the entire SNOMED CT content but in this research it refers only to SNOMED CT content used in an implementation, and more specifically, subsets and extensions. Auditing subsets occurs at the pre-coordination and post-coordination level while auditing extensions occurs at the extensions level.

At the pre-coordination level, auditing checks need to ensure subsets that have been compiled contain concepts from the appropriate hierarchy. For example, a problem list subset should not contain concepts from the **363787002|Observable entity (observable entity)|** hierarchy. Auditing checks need to ensure only active concepts are used. SNOMED CT undergoes changes with each release version that affects both textual descriptions and semantic relationships. These changes can have an impact in terms of whether concepts are still encoded accurately with active concepts and whether post-coordinated expressions still conform to the Concept Model.

At the post-coordination level, auditing checks need to ensure that post-coordinated expressions do not contain syntax errors. For example, nested refinements should be enclosed with parentheses **()** and relationship groups should be enclosed with curly brackets **{}**. Auditing checks also need to ensure that post-coordinated expressions conform to the Concept Model. For example, **272741003|Laterality (attribute)|** should only be used with the domain **123037004|Body structure (body structure)|** and not directly with concepts from the **404684003|Clinical finding (finding)|** and **71388002|Procedure (procedure)|** hierarchies.

At the extensions level, auditing checks need to ensure that extension concepts are modelled correctly similar to the checks done for post-coordinated expressions and to ensure that the inferred relationships of parent and child concepts are compatible (i.e., the inferred defining attributes of the child concept is subsumed by the inferred defining attributes of the parent concept). Finally, auditing checks need to ensure that the extension concepts do not already exist in the core.

5.2.2.2 Encoding

The purpose of this dimension is to ensure data is encoded accurately, consistently and unambiguously. There are about 300,000 active SNOMED CT concepts and some concepts have the same or very similar descriptions that can make selecting the appropriate concept challenging. There are also a wide range of body structure concepts such as general, bone, muscle and skin structures that are used in different diseases and findings, and selecting an incorrect concept can lead to unexpected results when testing for equivalency and subsumption.

At the pre-coordination level, the encoding method should incorporate data cleaning algorithms (e.g., spelling corrections, expanding acronyms and abbreviations),[3] a variety of search algorithms (e.g., exact match, match all) and sorting algorithms to ensure the most relevant results are retrieved first.

At the post-coordination level, the encoding method needs to be able to handle how and when to combine two or more pre-coordinated concepts into a post-coordinated expression that is technically and semantically correct. As the MRCM guidelines are broad, the encoding method also needs to ensure that the post-coordinated expressions are not nonsensical. It needs to be able to represent negation, certainty and time duration through the different contextual qualifiers (e.g., **408729009|Finding context (attribute)|** and **408731000|Temporal context (attribute)|**).

At the extensions level, commonly used clinical statements that are represented using post-coordinated expressions may be suggested for inclusion as a pre-coordinated concept. In addition, clinical statements that could not be encoded using post-coordination may be manually reviewed and a new extension concept may need to be manually crafted. The auditing check should ensure that no equivalent pre-coordinated concept exists, and the retrieval method can suggest where the new concept should be located in the hierarchy.

5.2.2.3 Retrieval

This dimension centres on retrieving SNOMED CT-encoded patient records. While a large portion of the SNOMED CT literature thus far has focused on prospective content coverage, only a small portion has focused on retrieving SNOMED CT-encoded patient records using SNOMED CT's semantic relationships. Retrieval is an important dimension because decision support functionality, patient case queries and the reporting of quality indicators all centre on being able to retrieve what has been recorded.

At the pre-coordination level, retrieval can be conducted for equivalency (exactly the same identifier), subsumption (using the transitive closure table) and defining attributes (using the relationships table). As only pre-coordinated concepts are used, queries can be conducted with structured query language (SQL) statements using the relationships and transitive closure tables. Retrieval can be constrained to a portion of the hierarchy or to a specific subset. Active concepts can be located for inactive concepts using the historical relationships.

At the post-coordination level, retrieval is more complex as the short and long normal forms need to be generated when using structural subsumption. As structural subsumption involves testing each pair of expressions individually, which is usually not feasible or practical when a dataset contains thousands of unique expressions, there is a need to optimise structural subsumption. This can be done by storing the post-coordinated expressions in a format that is optimised for retrieval, and by developing algorithms to first retrieve a set of candidate expressions before using the normal forms for more detailed comparisons.

At the extension level, retrieval is not much different from retrieval at the pre-coordination level as extensions are pre-coordinated concepts. The complexity of the retrieval method will depend on whether the concepts and relationships are imported into the same tables as the core concepts and relationships or kept separately. The canonical table will need to be regenerated to ensure the correct proximal primitive concepts are used, and the transitive closure table will need to be regenerated to improve the efficiency of testing whether two concepts have a supertype/subtype relationship.

5.3 SNOMED CT Extensions Auditing Framework

The SNOMED CT Extensions Auditing Framework was developed after the verification rules for auditing extensions were identified as part of the study in **Chapter Six** that focused on auditing three country level extensions. As it was unwieldy to present 89 verification rules as a long list of rules without any structure, the framework was developed to organise the verification rules by the type of errors they were designed to track and which SNOMED CT component it belonged to.

SNOMED CT consists of three core components: concepts, descriptions and relationships (refer to **Figure 5-2**), and these three components formed one of the dimensions (i.e., columns) of the extensions auditing framework. Each component consists of a variety of SNOMED CT identifiers, value sets, external identifiers and textual descriptions. For example, **SCT_Concepts** includes a SNOMED CT identifier (i.e., **ConceptId**), codes from value sets (i.e., **ConceptStatus** and **IsPrimitive**), external identifiers (i.e., **CTV3ID** and **SNOMEDID**) and textual descriptions (i.e., **FullySpecifiedName**). The verification rules that were compiled were based on these data elements that formed the basis of how concepts, descriptions and relationships are modelled.

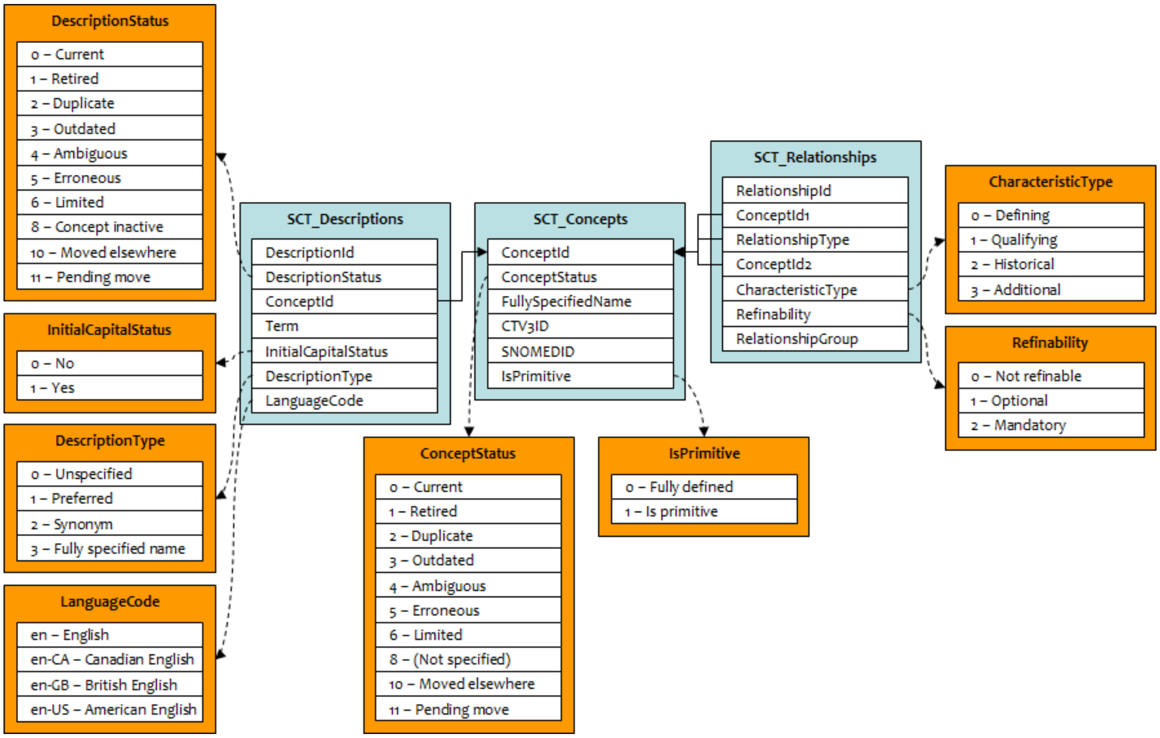


Figure 5-2. SNOMED CT Release Format 1 schema with associated value sets.

In the auditing extension study that is reported in **Chapter Six**, a total of 89 verification rules were identified. Upon closer review of the rules, it was found that they could be grouped into six categories, which formed the second dimension (i.e., rows) of the extensions auditing framework. The six categories are as follows. First, verification of extension identifiers, which include ensuring component identifiers use the correct namespace, partition identifier and check digit and there are no duplicate identifiers. For example, the **RelationshipId** in the Relationships column indicates that the **RelationshipId** is being verified to ensure the namespace, partition identifier and check digit are correct. Second, verification of the data elements that use value sets, which ensure that coded data elements use the appropriate value sets. For example, the **CharacteristicType** in the Relationships column indicates the **CharacteristicType** is being verified to ensure it contains one of the four values (i.e., 0, 1, 2, and 3). Third, verification of dependency values, including intra- and inter-component dependencies, which ensure that components are defined correctly. For example, **DescriptionStatus → DescriptionType → LanguageCode** in the Relationships column indicates that the description status, description type and language code have an intra dependency between each other. Fourth, verification that defining attributes assigned to each concept conforms to the Machine Readable Concept Model (MRCM), which ensures that concepts are being modelled consistently. For example, **Attribute-domain** in the Relationships column indicates that a domain is being checked to ensure the appropriate attribute is used with it. Fifth, verification that the relationships distributed contains the inferred relationships, which ensures that all the necessary defining attributes are included in the release. Sixth, verification of the number of occurrences a component may have, which ensures that mandatory descriptions (e.g., fully specified name) and relationships (e.g., "is a" relationship) exist. The auditing extensions framework is shown in **Table 5-1**Error! Reference source not found. and additional details are available in **Chapter Six** that describe the verification rules in detail.

Table 5-1. Framework for auditing method for SNOMED CT Release Format 1 for consistency.

Category / Component		Concepts	Descriptions	Relationships
Identifiers		▪ ConceptId ▪ Identifier Duplication	▪ DescriptionId ▪ Identifier Duplication	▪ RelationshipId ▪ Identifier Duplication
Value Sets		▪ ConceptStatus ▪ IsPrimitive	▪ DescriptionStatus ▪ DescriptionType ▪ InitialCapitalStatus ▪ LanguageCode	▪ CharacteristicType ▪ Refinability ▪ RelationshipGroup
Dependencies	**Intra**	▪ IsPrimitive →ConceptStatus	▪ DescriptionStatus → DescriptionType → LanguageCode	▪ CharacteristicType → Refinability → RelationshipType → RelationshipGroup
	Inter	▪ N/A	▪ DescriptionId → Concept.ConceptId ▪ DescriptionStatus → Concept.ConceptStatus	▪ ConceptId1 → Concept.ConceptId ▪ ConceptId2 → Concept.ConceptId ▪ RelationshipType → Concept.ConceptId ▪ RelationshipType → Concept.ConceptStatus
Machine Readable Concept Model		▪ N/A	▪ N/A	▪ Attribute-domain ▪ Range-attribute ▪ Co-dependency of findings and procedures
Inferred Relationships		▪ N/A	▪ N/A	▪ Inferred relationships exist
Occurrences		▪ N/A	▪ Fully specified name ▪ Preferred term	▪ "Is a" relationships ▪ Duplication relationships ▪ Redundant relationships ▪ Location of inactive concepts ▪ Presence of historical relationships

5.4 SNOMED CT Clinical Value Framework

In **Chapter Nine**, the SNOMED CT Clinical Value Framework was developed to categorise the different clinical values of using SNOMED CT. This framework is based primarily upon the Clinical Value Targets by Canada Health Infoway but is also supported by three other initiatives in Canada.

As mentioned in **Chapter One**, Canada Health Infoway developed a set of criteria called Clinical Value Targets to measure the effective use of EMRs in 2011.[4] The Clinical Value Targets were grouped into two levels. First, Clinical Value Level 1, which focused on the use of EMRs in areas such as the recording of diagnoses and prescriptions, use of EMR generated alerts and reminders and the receiving of laboratory results electronically. Second, Clinical Value Level 2, which focused on the linking of EMRs with provincial drug information systems and electronic prescribing. The three other initiatives in Canada that indirectly support the objectives of the Clinical Value

Targets are the EMR Content Standards[5] and Primary Health Care Indicators[6] by the Canadian Institute for Health Information (CIHI), and Health System Use[7] Project by Canada Health Infoway. The EMR Content Standards are a set of data element and value set specifications that identify what data should captured electronically in primary care with the overall aim of "improve[ing] access, quality, outcomes and chronic disease prevention and management."[8] The Primary Health Care Indicators and Health System Use demonstration projects focus on using the data in areas such as quality reporting, decision support, statistical analysis and other secondary uses.

These Canadian initiatives all converge on at least three common areas. First, the need to capture clinical data in a standardised manner that is accurate and consistent. Second, the need to query the data captured to facilitate clinical decision support such as generating alerts and reminders as well as other secondary uses. Third, the need to improve care coordination. These areas also correspond to the stated benefits that SNOMED CT has been promoted to meet.[9] To facilitate the implementation of SNOMED CT, Canada Health Infoway has taken the lead to develop subsets and extensions in the primary health care domain.[10]

In **Chapter Nine**, the incremental clinical value of using SNOMED CT in primary health care is demonstrated from multiple perspectives. That study contributes to SNOMED CT knowledge by showing how: (a) SNOMED CT concepts can be incorporated into published inference methods to identify patients with specific health conditions; (b) SNOMED CT semantic features such as subsumption and defining relationships can be used to create SNOMED CT-based query tools; and (c) SNOMED CT may be implemented in ways that add tangible value in clinical settings.

A review of the proposed criteria[3, 11] (refer to **Table 5-2**) revealed they are targeted at three areas. First, data capture (e.g., "enter encounter notes," "enter problem lists"). Problem lists have been defined as "a list of current and active diagnoses as well as past diagnoses relevant to the current care of the patient"[12] or "non-transitive illnesses" [13] (non-transient). On the other hand, an encounter diagnosis is specific to an encounter or appointment. For example, a patient who complaints of nausea will have nausea recorded as part of the encounter diagnoses but probably not the problem list as the condition is more likely to be a transitive illness. A chronic disease such as diabetes would be considered a non-transitive illness and is suitable for inclusion in the problem list. Second, data retrieval (e.g., "generate automated alerts from within the EMR," "generate automatic reminders from within the EMR"). Third, data sharing (e.g., "receive laboratory results into the EMR," "access hospital clinical reports and encounter summaries from the EMR"). These three areas of data use (i.e., "data capture," "data retrieval" and "data sharing") therefore made up the first dimension of the conceptual framework.

Table 5-2. Canada Health Infoway Clinical Value Targets.

Clinical Value Level 1	Clinical Value Level 2
▪ Enter encounter notes ▪ Enter problem lists ▪ Enter allergies ▪ Enter immunisations	▪ Clinician's use of an EMR ▪ Interoperability with drug information systems ▪ Clinician's use of electronic prescribing capability from the EMR

Clinical Value Level 1	Clinical Value Level 2
▪ Enter vital stats ▪ Enter new or renewal prescriptions and print the prescription ▪ Generate automated alerts from within the EMR ▪ Generate automated reminders from within the EMR ▪ Receive laboratory results into the EMR (electronic interface, scanning or data entry) ▪ Receive diagnostic imaging results into the EMR ▪ Access hospital clinical reports and encounter summaries from the EMR ▪ Create referral letters or consultation reports	

While the Clinical Value Targets are aimed at primary use of data (i.e., patient), the Primary Health Care Indicators and Health System Use project are aimed at the secondary use of data (i.e., practice, population). Therefore these target uses of data made up the second dimension (i.e., patient; practice; and population). These data use and data use targets were combined into a 3x3 table, which formed the proposed conceptual framework for demonstrating the clinical value of SNOMED CT (refer to **Table 5-3**). It should be noted that the Clinical Value Targets do not indicate that SNOMED CT should be used and **Table 5-3** only contains examples of how SNOMED CT can be used.

Table 5-3. SNOMED CT Clinical Value Framework for demonstrating the clinical value of SNOMED CT and examples of SNOMED CT use.

Data Use / Data Use Target Audience	Patient	Practice	Population
Data Capture	▪ Suggest billing diagnostic code based on encounter diagnoses ▪ Suggest encounter diagnoses based on billing diagnostic code ▪ Suggest adding encounter diagnosis to problem list	▪ Audit (critique) completeness of problem list and encounter diagnoses ▪ Audit (critique) accuracy of billing diagnostic codes	▪ Compare the completeness of problem lists, encounter diagnoses and billing diagnostic codes between practices and by population
Data Retrieval	▪ Alert clinicians to potential drug-allergy interactions ▪ Alert clinicians to out-of-range values ▪ Remind clinicians to order routine tests ▪ Suggest medications based on out-of-range test results	▪ Audit (critique) evidence-based guideline adherence ▪ Patient case queries and patient recall ▪ Generate statistics on key primary health care indicators	▪ Compare clinical outcomes between practices and by population
Data Sharing	▪ Generate standardised patient clinical summary in the form of the Continuity of Care Document (CCD) for referrals that are encoded with SNOMED CT	▪ Generate standardised practice quality reports in the form of the Quality Reporting Document Architecture (QDRA) Category II/III that are encoded with SNOMED CT	▪ Generate standardised practice quality reports in the form of the Quality Reporting Document Architecture (QDRA) Category III that are encoded with SNOMED CT

5.5 References

[1] Centre for Health Information Research and Development Implementing SNOMED CT within national electronic record solutions. http://www.chirad.org.uk/paper_one.htm. Last accessed: June 7, 2010.

[2] HIMSS EMR Adoption Model. http://www.himssanalytics.org/hc_providers/emr_adoption.asp. Last accessed: July 24, 2013.

[3] Lee, D. H., Lau, F. Y., & Quan, H. (2010). A method for encoding clinical datasets with SNOMED CT. BMC Medical Informatics and Decision Making, 10(1), 53. http://www.ncbi.nlm.nih.gov/pubmed/20849611.

[4] Clinical Value: "Meaningful Use" in Canada. http://infowayconnects.infoway-inforoute.ca/blog/clinicians-and-health-informatics/220-clinical-value-meaningful-use-in-canada. Last accessed: May 25, 2011.

[5] Canadian Institute for Health Information. (2012). Draft Pan-Canadian Primary Health Care Electronic Medical Record Content Standard, Version 2.1 – Implementation Guide. https://secure.cihi.ca/free_products/PHC_EMR_CS_Implementation_Guide.pdf.

[6] Canadian Institute for Health Information. (2012). Pan-Canadian Primary Health Care Indicator Report. https://secure.cihi.ca/free_products/Pan-Canadian_PHC_Indicator_Update_Report_en_web.pdf.

[7] [PowerPoint Presentation] Canada Health Infoway. (2011). Health System Use Demonstration Projects Overview.

[8] What are EMR content standards for PHC? http://www.cihi.ca/CIHI-ext-portal/internet/EN/document/types+of+care/primary+health/faq_ph_emr_content_stdrds. Last accessed: July 24, 2013.

[9] SNOMED CT Benefits. http://www.ihtsdo.org/snomed-ct/whysnomedct/benefits. Last accessed: July 24, 2013.

[10] Canada Health Infoway. SNOMED CT. https://infocentral.infoway-inforoute.ca/2_Standards/1_pan-Canadian_Standards/Terminology/4_SNOMED_CT_Terminologies. Last accessed: July 24, 2013. Note: Login is required.

[11] Manitoba e-Health. EMR Adoption Program: Program Guidelines and Detailed Requirements V2.1. http://www.manitoba-ehealth.ca/commPhysicians/files/Prog_Guidelines.pdf.

[12] Centres for Medicare and Medicaid Services. Eligible Professional Meaningful Use Core Measures Measure 3 of 15. http://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/downloads/3MaintainProblemList.pdf.

[13] Holmes, C. (2011). The problem list beyond meaningful use. Part I: The problems with problem lists. Journal of AHIMA/American Health Information Management Association, 82(2), 30-3. http://www.ncbi.nlm.nih.gov/pubmed/21337850.

6. AUDITING METHOD & RESULTS

6.1 Introduction

A large portion of the Systematised Nomenclature of Medicine Clinical Terms (SNOMED CT) publications focus on prospective content coverage, in which local terms and data dictionaries are mapped to SNOMED CT.[1] When the content coverage of SNOMED CT is inadequate, organisations are encouraged to create their own extension concepts, descriptions and relationships to meet their terminology requirements.[2, 3] This is done by using a SNOMED CT namespace, a seven-digit identifier that is used to construct SNOMED CT component identifiers. In order to obtain a namespace, organisations need to submit written requests to their National Release Centres (NRC), which will in turn forward the request to the International Health Terminology Standards Development Organisation (IHTSDO). To date, over 165 namespaces have been allotted to health care organisations, vendors and academic institutions.[4] Extensions that have been published at country levels include Australia (AU),[a] Canada (CA),[b] the United States (US)[c] and the United Kingdom (UK)[d] while publicly available extensions at a health care organisation level include Kaiser Permanente's Convergent Medical Terminology.[5]

Authors of auditing methods of terminology systems have reported that it is "not uncommon,"[6] "unavoidable,"[7, 8] and "inevitable"[21] that errors occur in large complex terminology systems. Therefore it is reasonable to suppose that SNOMED CT extensions are also susceptible to containing errors. Errors in extensions can lead to incorrect descriptions being displayed to clinicians, incorrect inferences made and inaccuracies when testing for equivalency and subsumption. Therefore the objective of this study is to develop an auditing method that can be used to verify extensions so as to improve the quality of extensions. In this context, verification refers to whether the extension concepts, descriptions and relationships adhere with the requirements and guidelines as specified in the User Guide (UG) and Technical Implementation Guide (TIG) as opposed to whether the extensions fulfil user requirements and expectations. In another words, the auditing method focuses on consistency ("the knowledge should not be self-contradictory")[9] as opposed to completeness ("it should have the necessary knowledge")[9] or correctness ("the knowledge should be faithful to the real world").[9] The auditing method developed adds to the current state of knowledge by identifying the verification rules, developing a framework of categorising the errors and representing the verification rules in a machine-readable format that can be easily used to verify extensions. As more countries and health care organisations release extensions, it is expected that the auditing method will play a greater role in ensuring extensions are error-free.

While the auditing method in this chapter focuses on auditing SNOMED CT extensions for consistency, there are other issues that can affect the accuracy of encoding and retrieval. Although these issues can be classified as

---

[a] https://nehta.org.au/aht
[b] https://infocentral.infoway-inforoute.ca
[c] http://www.nlm.nih.gov/research/umls/Snomed/us_extension.html
[d] https://www.uktcregistration.nss.cfh.nhs.uk/trud3/user/guest/group/2/pack/8/subpack/12/releases

within the scope of auditing, the auditing in this chapter deals only with the consistency of extensions and therefore those issues are addressed in the appendices of the **Encoding Method Chapter** and **Retrieval Method Chapter**.

6.2 Materials

The UG and TIG are published semi-annually by the IHTSDO with each release version of SNOMED CT and are available as PDF files that are distributed by the NRCs and on the IHTSDO's website.[10, 11] The UG and TIG were used to compile the verification rules.

In this study, three extensions (i.e., CA, US and UK) were analysed and audited. The AU extension was not included in this study as only Australian-based organisations can obtain a copy of the extension. At the time of this study, the CA and UK extensions were only available via Release Format 1 (RF1) while the US extension was available in both RF1 and Release Format 2 (RF2). While the US and UK extensions have undergone at least five revisions, the CA extension has only been released once (December 21, 2012) using the July 2012 release version of SNOMED CT. Therefore in this study, the July 2012 RF1 extensions for the three countries were used as that is the only release version and format that was available across the three countries.

The number of concepts, descriptions and relationships contained in the three extensions are shown in **Table 6-1** and shows that the UK extension was significant larger than the CA and US extensions.

Table 6-1. Summary of CA, US and UK extensions.

No	Component	CA	US	UK
1.	Concepts	1,378	1,327	71,871
2.	Descriptions	2,878	4,228	172,937
3.	Relationships	3,519	2,267	222,415

Excluding inactive concepts, the majority of the CA extension concepts were from the **71388002|Procedure (procedure)|** hierarchy while the majority of the US and UK extension concepts were from the **404684003|Clinical finding (finding)|** hierarchy (refer to **Table 6-2**). Additional frequency counts of the three extensions are available in **Appendix D**.

Table 6-2. Frequency counts of CA, US and UK extensions by top-level hierarchy.

No	Hierarchy	Canada		United States		United Kingdom	
		Total	%	Total	%	Total	%
1.	123037004|Body structure (body structure)|	5	0.4%	27	2.0%	512	0.7%
2.	404684003|Clinical finding (finding)|	54	3.9%	499	37.8%	7,643	10.6%
3.	308916002|Environment or geographical location (environment / location)|	127	9.2%	2	0.2%	56	0.1%
4.	272379006|Event (event)|			4	0,3%	3,452	4.8%
5.	106237007|Linkage concept (linkage concept)|					7	0.0%
6.	363787002|Observable entity (observable entity)|	2	0.1%	39	3.0%	660	0.9%
7.	410607006|Organism (organism)|	79	5.7%	214	16.2%	2	0.0%

No	Hierarchy	Canada		United States		United Kingdom	
		Total	%	Total	%	Total	%
8.	373873005|Pharmaceutical / biologic product (product)|	8	0.6%	6	0.5%		
9.	78621006|Physical force (physical force)|					17	0.0%
10.	260787004|Physical object (physical object)|			11	0.8%		
11.	71388002|Procedure (procedure)|	613	44.5%	199	15.1%	5,903	8.2%
12.	362981000|Qualifier value (qualifier value)|	290	21.0%	10	0.8%	976	1.4%
13.	419891008|Record artifact (record artifact)|			2	0.2%	892	1.2%
14.	243796009|Situation with explicit context (situation)|	4	0.3%	166	12.6%	3,085	4.3%
15.	48176007|Social context (social concept)|	55	4.0%	9	0.7%	432	0.6%
16.	370115009|Special concept (special concept)|	40	2.9%	62	4.7%	47,961	66.7%
17.	123038009|Specimen (specimen)|			23	1.7%	60	0.1%
18.	254291000|Staging and scales (staging scale)|					194	0.3%
19.	105590001|Substance (substance)|	101	7.3%	48	3.6%	19	0.0%

6.3 Methods

As the IHTSDO does not publish the guidelines for checking extensions in a machine-readable format, the verification rules that were identified in this study were derived from the UG and TIG. Any ambiguities in the guides or assumptions were tested against the July 2012 international release version of SNOMED CT to determine if the assumptions were accurate. After the verification rules were identified, they were grouped into categories based on the type of error they were designed to track.

SNOMED CT consists of three core components: concepts, descriptions and relationships (refer to **Figure 6-1**), and these three components were the focus of the verification rules. Some of these components are defined with value sets and interdependencies between each other. Verifications of textual descriptions (i.e., **SCT_Concepts.FullySpecifiedName** and **SCT_Descriptions.Term**) such as the checking for the appropriate suffix of a fully specified name, appropriateness of language codes for preferred terms and synonyms, spelling mistakes, grammar or consistency of descriptions were beyond the scope of this study. In addition, legacy identifiers **SCT_Concepts.CTV3ID** and **SCT_Concepts.SNOMEDID** were omitted as there is little added value in verifying legacy identifiers. Except for the four data elements mentioned, each of the data elements in the three components were analysed to determine the range of applicable values.

In this study, a total of 89 verification rules were identified. Where applicable, structured query language (SQL) statements were constructed to ensure the verification rules could be used to automatically audit an extension. These verification rules were organised into the SNOMED CT Extensions Auditing Framework and are shown in **Table 6-3**, which consists of columns that refer to the three main components of SNOMED CT (i.e., concepts; descriptions; and relationships) and rows that refer to the categories of verification rules. The six categories are as follows. First, verification of extension identifiers, which ensures component identifiers use the correct namespace, partition identifier and check digit and there are no duplicate identifiers. Second, verification of the data elements that use value sets, which ensures coded data elements use the appropriate value sets. Third, verification of dependency values, including intra- and inter-component dependencies, which ensures components are defined correctly. Fourth, verification that defining attributes assigned to each concept conforms to the Machine Readable Concept Model (MRCM), which ensures that concepts are being modelled consistently. Fifth, verification that the relationships distributed contains the inferred relationships, which ensures that all the necessary defining attributes are included in the release. Sixth, verification of the number of occurrences a component may have, which ensures that mandatory descriptions (e.g., fully specified name) and relationships (e.g., "is a" relationship) exist. A violation of verification rules is considered an error, and will result in different levels of consequences, which are discussed in the following subsections.

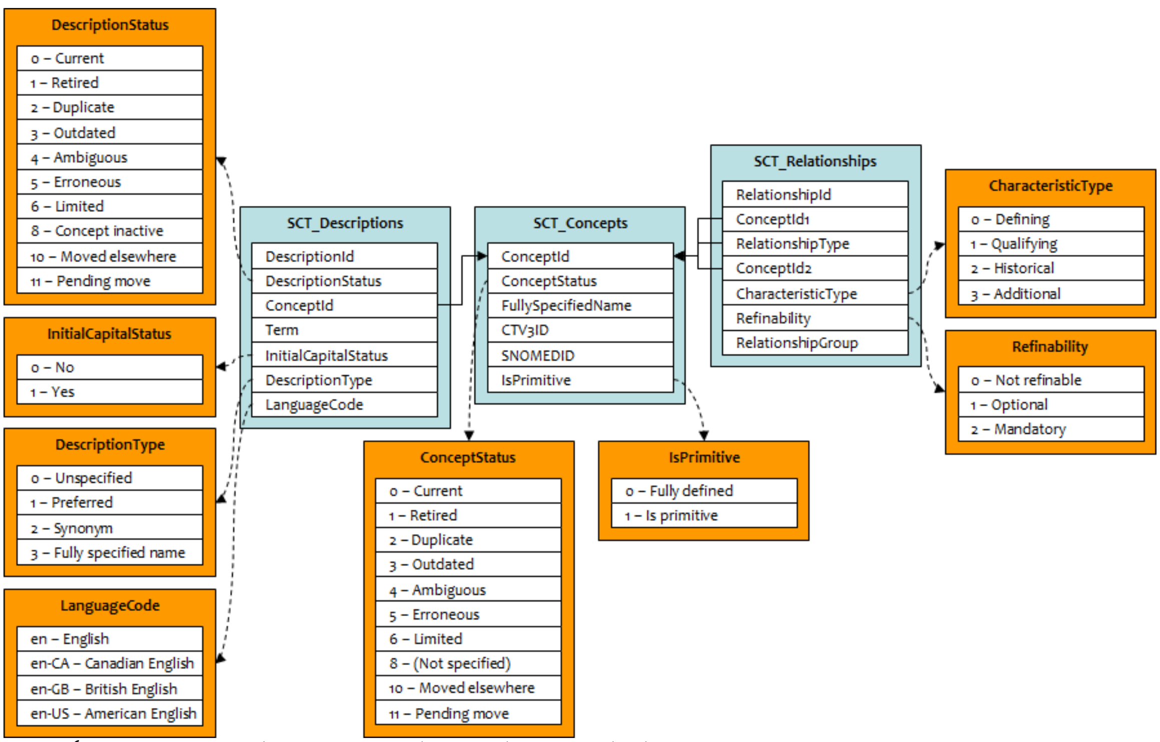


Figure 6-1. SNOMED CT Release Format 1 schema with associated value sets.

After the verification rules were compiled, the CA, US and UK concepts, descriptions and relationships extensions were imported into the tables **TBL_MyExtensionConcepts**, **TBL_MyExtensionDescriptions** and **TBL_MyExtensionRelationships** respectively rather than the core tables. The transitive closure table was generated and stored in the table **TBL_MyExtensionTransitiveClosure** to improve the performance of auditing the extensions. The verification rules were then applied to these extensions.

A summary of all 89 verification rules are shown in **Table 6-3** while detailed descriptions are available in **Appendix D**.

Table 6-3. Summary of all verification rules.

Category / Components		Concepts	Descriptions	Relationships
Identifiers		▪ ConceptId uses correct namespace ▪ ConceptId uses correct partition identifier ▪ ConceptId has correct check digit ▪ ConceptId does not exist in the core	▪ DescriptionId uses correct namespace ▪ DescriptionId uses correct partition identifier ▪ DescriptionId has correct check digit ▪ DescriptionId does not exist in the core	▪ Relationship Id uses correct namespace ▪ RelationshipId uses correct partition identifier ▪ RelationshipId has correct check digit ▪ RelationshipId does not exist in the core
Value Sets		▪ ConceptStatus uses one of 10 permissible values ▪ IsPrimitive uses one of 2 permissible values	▪ DescriptionStatus uses one of 10 permissible values ▪ DescriptionType uses one of 4 permissible values ▪ InitialCapitalStatus uses one of 2 permissible values ▪ LanguageCode uses one of 3 permissible values	▪ CharacteristicType uses one of 4 permissible values ▪ Refinability uses one of 2 permissible values ▪ RelationshipGroup is numeric
Dependencies	**Intra**	▪ Fully defined concepts must be active concepts	▪ Active and pending move fully specified name descriptions must have EN as the language code	▪ "Is a" attribute must be defining, not refinable and must be ungrouped ▪ Concept model attributes (with the exception of "associated finding," "associated procedure," "is a" and "part of") not linked to qualifier value ranges must have be defining and have an optional refinability ▪ Concept model attributes linked the qualifier value ranges (plus "associated finding" and "associated procedure", minus "surgical approach" and "using device") must be defining with optional refinability or qualifying with mandatory refinability ▪ "Surgical approach" attribute must be defining or qualifying or and optional refinability ▪ "Using device" attribute must have defining (with optional refinability) or qualifying (with either non refinable or optional refinability) ▪ "Laterality" attribute must have a defining (with refinability optional) or qualifying (with refinability mandatory) characteristic type and relationship group must be 0 ▪ Concept history attributes must have a historical characteristic type, refinability is not refinable and relationship group must be 0 ▪ "Part of" attribute must have an additional characteristic type, refinability is not refinable and relationship group must be 0
	Inter	▪ N/A	▪ ConceptId in the Descriptions	▪ ConceptId1 in the Relationships

Category / Components	Concepts	Descriptions	Relationships
		table must exist in the Concepts table ▪ Current descriptions can only belong to current concepts ▪ Retired, duplicate, outdated, erroneous, and inappropriate descriptions can belong to any concept ▪ Limited descriptions can only belong to limited concepts ▪ Concept inactive descriptions can belong to any inactive concept except limited ▪ Moved elsewhere descriptions can only belong to moved elsewhere concepts ▪ Pending move descriptions can only belong to pending move concepts	table must exist in the Concepts table ▪ RelationshipType in the Relationships table must exist in the Concepts table ▪ ConceptId2 in the Relationships table must exist in the Concepts table ▪ All destination concepts in the relationships table are current ▪ 370124000|REPLACED BY (attribute)| should be used to link retired, outdated or erroneous to current concepts ▪ 168666000|SAME AS (attribute)| should be used to link duplicate concepts to current concepts, or limited concepts to other limited concepts ▪ 149016008|MAY BE A (attribute)| should be used to link ambiguous concepts to current concepts ▪ 159083000|WAS A (attribute)| should be used to link limited concepts to current concepts ▪ 370125004|MOVED TO (attribute)| should be used to link moved elsewhere concepts to current concepts.
Machine Readable Concept Model	▪ N/A	▪ N/A	▪ Appropriateness of attribute with domain ▪ Appropriateness of range with attribute ▪ Concepts should either be defined with (associated finding and finding context) or (associated procedure and procedure context) but not a combination of the two ▪ Concepts should either be defined with only one each of the following: associated finding, finding context, associated procedure, procedure context, temporal context and subject relationship context
Inferred Relationships	▪ N/A	▪ N/A	▪ Inferred relationships exist
Occurrences	▪ N/A	▪ Each current extension concept should have one current fully specified name in English ▪ Each current extension concept should have only one current fully specified name in English ▪ Each pending move extension concept should have one pending move fully specified name in English	▪ Each extension relationship added to a core concept should not duplicate what already exists in the core ▪ Each extension "is a" relationship should not be redundant to what already exists in the core ▪ Each extension concept should belong to only one top-level

Category / Components	Concepts	Descriptions	Relationships
		▪ Each pending move extension concept should have only one pending move fully specified name in English ▪ Each inactive extension concept should have one concept inactive fully specified name in English ▪ Each inactive extension concept should have only one concept inactive fully specified name in English ▪ Each limited extension concept should have one pending move fully specified name in English ▪ Each limited extension concept should have only one pending move fully specified name in English ▪ Each current extension concept should have at least one current preferred term ▪ Each current extension concept should have at only one current preferred term in each language code ▪ Each pending move extension concept should have at least one pending move preferred term ▪ Each pending move extension concept should have at only one pending move preferred term in each language code ▪ Each inactive extension concept should have at least one concept inactive preferred term ▪ Each inactive extension concept should have only one concept inactive preferred term in each language code ▪ Each limited extension concept should have at least one current preferred term ▪ Each limited extension concept should have only one current preferred term in each language code	hierarchy ▪ Each current extension concept should have at least one "is a" relationship ▪ Concepts should only have one of the following: associated finding, associated procedure, finding context, procedure context, temporal context and subject relationship context ▪ Each active extension concept should not have a historical relationship except for 384598002|MOVED FROM (attribute)| ▪ Each inactive extension concept should only have an "is a" relationship and concept history attributes ▪ Retired without a stated reason concepts should be a subtype of 363661006|Reason not stated concept (inactive concept)| ▪ Duplicate concepts should be a subtype of 363662004|Duplicate concept (inactive concept)| ▪ Outdated concepts should be a subtype of 363663009|Outdated concept (inactive concept)| ▪ Ambiguous concepts should be a subtype of 363660007|Ambiguous concept (inactive concept)|. ▪ Erroneous concepts should be a subtype of 363664003|Erroneous concept (inactive concept)| ▪ Limited concepts should be a subtype of 443559000|Limited status concept (inactive concept)| ▪ Moved elsewhere concepts should be a subtype of 370126003|Moved elsewhere (inactive concept)| ▪ Each retired concept must be 370124000|REPLACED BY (attribute)| by one, and only one current concept ▪ Each duplicate concept must be the 168666000|SAME AS (attribute)| one, and only one current concept ▪ Each outdated concept must be 370124000|REPLACED BY (attribute)| by one, and only one current concept ▪ Each ambiguous concept 149016008|MAY BE A (attribute)| one or more current concepts ▪ Each erroneous concept must be 370124000|REPLACED BY (attribute)| by one, and only one current concept ▪ Each limited concept

Category / Components	Concepts	Descriptions	Relationships
			159083000|WAS A (attribute)| current concept ▪ Each moved elsewhere concept must be 370125004|MOVED TO (attribute)| to one and only current concept

6.4 Results

The results are divided into two parts. First, an analysis of the three extensions is provided in **Section 6.4.1**, which focused on the relationships between the extension and core concepts. Second, the results of applying the auditing method against the three extensions are reported in **Section 6.4.2** by the six auditing categories (i.e., identifiers; value sets; dependencies; Machine Readable Concept Model; inferred relationships; and occurrence).

6.4.1 Analysis of Extensions

The analysis of the extensions are divided into a comparison of "is a" and non "is a" relationships to demonstrate how the hierarchy is change and how concepts are being defined.

6.4.1.1 "Is a" Relationships

The results of the "is a" relationships analysis is shown in **Table 6-4**. The majority (71.5% to 83.8%) of new "is a" relationships are between the extension (child) and core (parent) concepts. Only the US extension included new relationships between "core to core" and "core to extension" concepts. The examples in the rest of this "is a" relationships section will focus on the US Extension as there are examples of the four types of "is a" relationships.

Table 6-4. Summary of "is a" relationships in the Canadian, United States and United Kingdom extensions.

No	Defining Is A Relationships (Total, Not By Concept)	Canada		United States		United Kingdom	
		Total	%	Total	%	Total	%
1.	Core to core			38	2.5%		
2.	Core to extension			5	0.3%		
3.	Extension to core	1,203	74.0%	1,047	71.5%	60,345	83.8%
4.	Extension to extension	422	26.0%	376	25.7%	11,682	16.2%
	Total	**1,625**	**100.0%**	**1,465**	**100.0%**	**72,027**	**100.0%**

6.4.1.1.1 Core Concept to Core Concept

As reported in **Table 6-4**, only the US Extension created new "is a" relationships between core concepts. Thirty out of the 38 concepts were to link "Enteropathogenic Escherichia coli serogroup" concepts to **112283007|Escherichia coli (organism)|** (refer to **Table 12-32** in the appendix). It should be noted that these relationships are redundant as they currently already exist in SNOMED CT. There was one inactive concept (**1111000119100|Sebaceous nevus (disorder)|**) was that was linked to an active concept (**239107007|Epidermal nevus**

(disorder)|). The remaining seven concepts are shown in **Table 6-5** and **Figure 6-2**. The dotted lines represent the core-to-core relationships in the extension.

Table 6-5. Examples of core concepts to core concepts United States extension.

No	Relationship Id	Child Concept	Concept Status	Parent Concept	Concept Status
1.	441721000124121	279953009|Familial neonatal seizures (disorder)|	0	32895009|Hereditary disease (disorder)|	0
2.	441251000124124	442111003|Intermittent pneumatic compression stockings (physical object)|	0	303476002|Cardiovascular equipment (physical object)|	0
3.	441031000124122	301913002|Lesion of eyelid (finding)|	0	404684003|Clinical finding (finding)|	0
4.	441731000124124	59008008|Leukoencephalomyelopathy of Rottweilers (disorder)|	0	127326005|Non-human disorder (disorder)|	0
5.	441651000124122	246829005|Lid adhesions (finding)|	0	301913002|Lesion of eyelid (finding)|	0
6.	441311000124125	258560004|Oral saliva sample (specimen)|	0	441620008|Oral fluid specimen (specimen)|	0
7.	441341000124126	119342007|Saliva specimen (specimen)|	0	309051001|Body fluid sample (specimen)|	0

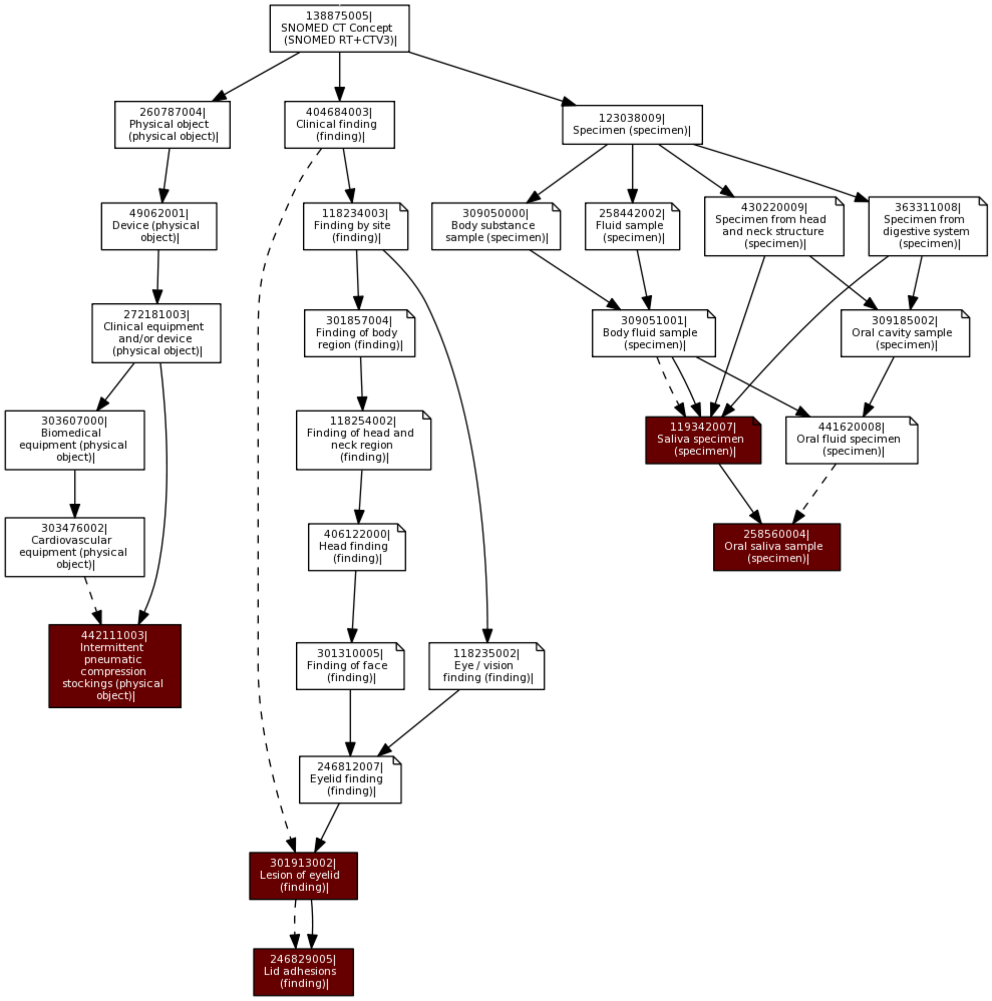


Figure 6-2. Example of core to core "is a" relationships in the US extension (dotted lines).

6.4.1.1.2 Core Concept to Extension Concept

There were five instances where core concepts became subtype of extension concepts (refer to **Table 6-6** and **Figure 6-3**). Two were from the **71388002|Procedure (procedure)|** hierarchy while the other three were from the **123037004|Body structure (body structure)|**, **404684003|Clinical finding (finding)|** and **105590001|Substance (substance)|** hierarchies.

Table 6-6. Core to extension concepts United States extension.

No	Relationship Id	Child Concept	Concept Status	Parent Concept	Concept Status
1.	7321000124125	58332002|Allergy education (procedure)|	0	414011000124106|Education about immune disorder (procedure)|	0
2.	441681000124125	395072006|Counseling for postnatal depression (procedure)|	0	428201000124103|Counseling for depression (procedure)|	0
3.	441671000124128	195226001|Dissection of cerebral arteries, nonruptured (disorder)|	0	429221000124104|Dissectiing aneurysm of cerebral artery (disorder)|	0
4.	9671000124129	264194004|Entire pituitary fossa (body structure)|	0	5311000124101|Anatomic landmark of head (body structure)|	0

No	Relationship Id	Child Concept	Concept Status	Parent Concept	Concept Status
5.	7721000124126	409591007|Sheep blood agar (substance)|	0	5011000124103|Blood agar (substance)|	0

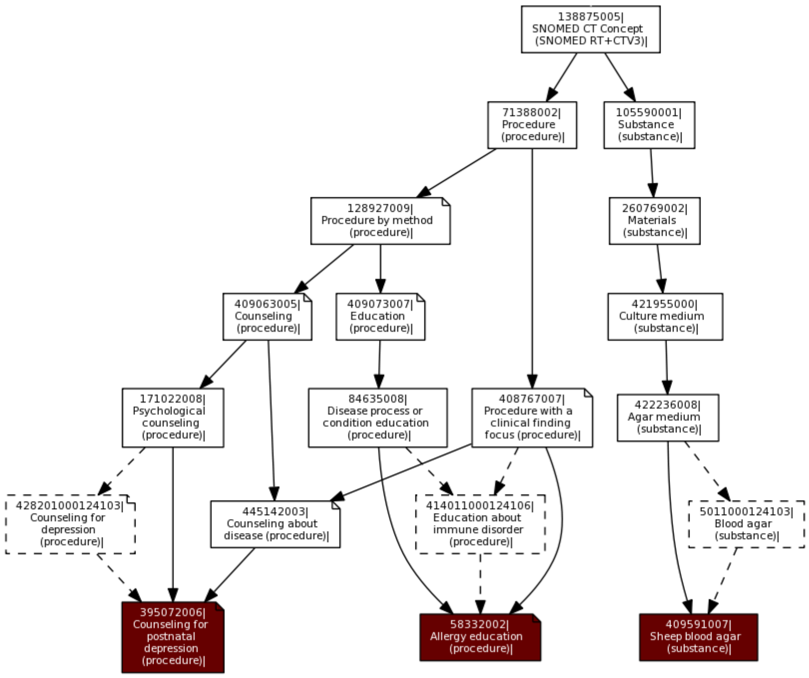


Figure 6-3. Example of core to extension "is a" relationships in the US extension (dotted lines).

6.4.1.2 Non "Is A" Relationships

The results of the non "is a" relationships are shown in **Table 6-7**. The core concepts were redefined in the US extension with other core concepts (n=10) and extension concepts (n=1). Although the results of the CA extension show that the core concepts were redefined with extension concepts (n=32), that is not exactly correct because the incorrect **CharacteristicType** was used in the defining of historical relationships. In this case, **CharacteristicType=0** (defining) was used instead of **CharacteristicType=2** (historical). The examples in the rest of this non "is a" relationships section will focus on the US Extension as there are examples of the four types of non "is a" relationships.

Table 6-7. Results of non "is a" relationships in the CA, US and UK extensions.

No	Defining Non Is A Relationships (Total, Not By Concept)	CA		US		UK	
		Total	%	Total	%	Total	%
1.	Core redefined with core			10	1.3%		
2.	Core redefined with extension	32	1.7%	1	0.1%		
3.	Extension defined with core	1,862	98.3%	718	96.9%	24,561	99.7%
4.	Extension defined with extension			12	1.6%	70	0.3%
	Total	**1,894**	**100.0%**	**741**	**100.0%**	**24,631**	**100.0%**

6.4.1.2.1 Core Redefined with Core

The 10 redefinitions of core concepts represent eight unique concepts (refer to **Table 6-8**). The new definitions were a result of refining current definitions as well as adding new definitions. In Example 2, **49755003|Morphologically abnormal structure (morphologic abnormality)|** has been refined to **42685002|Adhesion (morphologic abnormality)|**. In Example 3, a new attribute **118170007|Specimen source identity (attribute)|** was added in addition to the current attribute **118171006|Specimen procedure (attribute)|**.

Table 6-8. Concepts in the US extension that have been redefined with non "is a" relationships.

No	Concept	Definition From Core	Definition Added in Extension
1.	57684003|Parenchymal renal hypertension (disorder)|	363698007|Finding site (attribute)|= 51840005|Systemic circulatory system structure (body structure)|, 363698007|Finding site (attribute)|= 64033007|Kidney structure (body structure)|, 363705008|Has definitional manifestation (attribute)|= 24184005|Finding of increased blood pressure (finding)|, 47429007|Associated with (attribute)|= 90708001|Kidney disease (disorder)|	363698007|Finding site (attribute)|= 29704000|Structure of parenchyma of kidney (body structure)|
2.	246829005|Lid adhesions (finding)|	{116676008|Associated morphology (attribute)|= 49755003|Morphologically abnormal structure (morphologic abnormality)|, 363698007|Finding site (attribute)|= 80243003|Eyelid structure (body structure)|}	{116676008|Associated morphology (attribute)|= 42685002|Adhesion (morphologic abnormality)|, 363698007|Finding site (attribute)|= 80243003|Eyelid structure (body structure)|}
3.	258537007|Surface swab (specimen)|	118171006|Specimen procedure (attribute)|= 285570007|Taking of swab (procedure)|	118170007|Specimen source identity (attribute)|= 276339004|Environment (environment)|
4.	258560004|Oral saliva sample (specimen)|	{118169006|Specimen source topography (attribute)|= 75535001|Salivary structure (body structure)|, 370133003|Specimen substance (attribute)|= 256897009|Saliva (substance)|}	118169006|Specimen source topography (attribute)|= 74262004|Oral cavity structure (body structure)|
5.	300920004|Carotid atherosclerosis (disorder)|	{116676008|Associated morphology (attribute)|= 28960008|Arteriosclerosis (morphologic abnormality)|, 363698007|Finding site (attribute)|= 69105007|Carotid artery structure (body structure)|}	{116676008|Associated morphology (attribute)|= 38716007|Atherosclerosis (morphologic abnormality)|}
6.	371040005|Thrombotic stroke (disorder)|	363698007|Finding site (attribute)|= 28661005|Cerebrovascular system structure (body structure)|, 42752001|Due to (attribute)|= 65198009|Arterial thrombosis (disorder)|	{116676008|Associated morphology (attribute)|= 396339007|Thrombus (morphologic abnormality)|, 363698007|Finding site (attribute)|= 47990007|Structure of artery of head, neck AND/OR brain (body structure)|}
7.	401174001|Depression management program (regime/therapy)|		363702006|Has focus (attribute)|= 35489007|Depressive disorder (disorder)|
8.	427419006|Transformed migraine (disorder)|	263502005|Clinical course (attribute)|= 90734009|Chronic (qualifier value)|, 363698007|Finding site (attribute)|= 281231009|Vascular structure of head (body structure)|, 363705008|Has definitional manifestation	

No	Concept	Definition From Core	Definition Added in Extension
		(attribute)|= 25064002|Headache (finding)|	

6.4.1.2.2 *Core Redefined with Extension*

The US extension added an additional defining attribute to a core concept using an extension concept. **392163005|Neuroretinal rim finding (finding)|** is defined as **363698007|Finding site (attribute)|=81016008|Optic disc structure (body structure)|** in the core but the US extension added an additional definition of **363698007|Finding site (attribute)|=431321000124109|Structure of neuroretinal rim of optic disc (body structure)|**. It should be noted that **81016008|Optic disc structure (body structure)|** is a supertype of **431321000124109|Structure of neuroretinal rim of optic disc (body structure)|**.

6.4.2 Auditing Results

This section reports on the auditing results by the six categories. A summary of the auditing results are shown in **Table 6-9**. The CA extension had errors in three categories while the US and UK extensions had errors in five categories each. As each category contains verification rules of varying levels of importance, the summary table should not be used to judge the overall quality of the extension. Instead, the summary table should be used as a table of contents to hone into the categories where errors occurred and to review the detailed results. It should also be noted that an error can be counted in different categories. For example, if a description with a fully specified name has been defined using the language code **en-US**, it will fail the intra-dependency verification rule of not having all descriptions that are fully specified name with an English language code and fail the occurrence verification rule of requiring that each concept have a fully specified name with an English language code (**en**). The rest of this section describes the errors identified in detail.

Table 6-9. Summary of auditing results.

Category		Concepts			Descriptions			Relationships		
		CA	US	UK	CA	US	UK	CA	US	UK
Identifiers			1			265				
Value Sets				677						
Dependencies	**Intra**					1,268	2,115	950		337
	Inter									2,165
Machine Readable Concept Model								64	10	11
Inferred Relationships									596	39
Occurrences		41	1,327	11,729	41		1	34	45	

6.4.2.1 SNOMED CT Identifiers

This section reports on the results of the SNOMED CT identifiers auditing in four areas: namespace, partition identifier, check-digit and identifier duplication.

6.4.2.1.1 Namespace

All component identifiers were found to have used the correct namespace.

6.4.2.1.2 Partition Identifiers

All extension concepts and relationships were found to have used the correct partition identifiers. However, there were 264 descriptions in the US extension that used the partition identifier of **01** instead of **11** and they were used for core concepts as opposed to extension concepts. These descriptions centred on "Escherichia coli serogroup" concepts.

6.4.2.1.3 Check Digit

All the component identifiers were found to have used the correct check digit.

6.4.2.1.4 Identifier Duplication

The US extension contained one duplicate extension concept and 265 duplicate extension descriptions. The extension concept was **1111000119100|Sebaceous nevus (disorder)|**. It differed from the record contained in the core in that it was a current (**ConceptStatus=0**) and fully defined (**IsPrimitive=0**) concept while in the core it was a duplicate (**ConceptStatus=2**) and primitive (**IsPrimitive=1**) concept. Of the 265 duplicate descriptions, one was for the fully specified name for "sebaceous nevus," while the other 264 were descriptions for "Escherichia coli serogroup" concepts. It should be noted that the July 2012 release version of SNOMED CT includes 58 concepts and 58 descriptions from the US extension (namespace: **1000119**).

6.4.2.2 Value Sets

All the data elements that used value sets contained the appropriate values.

6.4.2.3 Dependencies

This section reports on the results of the intra and inter component dependency auditing.

6.4.2.3.1 Intra Component Dependencies

6.4.2.3.1.1 Concept Dependencies

All inactive concepts in the CA and US extensions were primitive concepts. There were however, 677 inactive concepts in the UK extension that were fully defined. Of the 677 concepts, 653 were **Erroneous** (e.g., **328551000000101|Action research - National Public Health Classification (administrative concept)|**), 112 were **Duplicate** (e.g., **509591000000107|Difficulty managing medication|**) while two were **Ambiguous** (e.g., **509811000000100|Anterior uveitis|**). It should be noted that in the July 2012 release version of SNOMED CT, all 99,872 inactive concepts were primitive.

6.4.2.3.1.2 Description Dependencies

All active (**DescriptionStatus=0**) and pending move (**DescriptionStatus=11**) CA fully specified name (**DescriptionType=3**) descriptions used the language code of English (**en**). There were 1,268 and 2,115 active fully specified name descriptions from the US and UK extensions that used the language code dialect of United States (**en-US**) and Great Britain (**en-GB**). The US and UK extensions do include 44 and 21,875 active and pending move fully specified name descriptions in English (en).

6.4.2.3.1.3 Relationship Dependencies

The results of the intra relationship dependency auditing are shown in **Table 6-10**.

Table 6-10. Summary of intra relationship dependency errors.

No	Description	CA	US	UK
1.	"Is a" attribute must have a defining characteristic type, not refinable refinability and must be in relationship group 0	0	0	0
2.	Concept model attributes (with the exception of "associated finding," "associated procedure," "is a" and "part of") not linked to qualifier value ranges must have a defining characteristic type, and optional refinability	886	0	337
3.	Concept model attributes linked the qualifier value ranges (plus "associated finding" and "associated procedure", minus "surgical approach" and "using device") must have a defining (with optional refinability) or qualifying (with mandatory refinability) characteristic type	0	0	0
4.	"Surgical approach" attribute must have defining or qualifying characteristic type and refinability must be optional	0	0	0
5.	"Using device" attribute must have defining (with optional refinability) or qualifying (with either non refinable or optional refinability)	0	0	0
6.	"Laterality" attribute must have a defining (with refinability optional) or qualifying (with refinability mandatory) characteristic type and relationship group must be 0	0	0	0
7.	Concept history attributes must have a historical characteristic type, refinability is not refinable and relationship group must be 0	64	0	0
8.	"Part of" attribute must have an additional characteristic type, refinability is not refinable and relationship group must be 0	0	0	0

There were 886 and 337 relationships in the CA and UK extensions that did not have the appropriate characteristic type or refinability. All 886 of the CA extension and 282 in the UK extension relationships used not refinable (**Refinability=0**) instead of optional (**Refinability=1**). The defining attributes for CA extension concept **1311000087107|Ultrasonography of right wrist (procedure)|** are shown in **Figure 6-4**. According to the relationship **14741000087120**, the **405813007|Procedure site - Direct (attribute)|** of **74670003|Wrist joint structure (body structure)|** is not refinable (**Refinability=0**). A review of all core relationships in the July 2012 release version of SNOMED CT indicates that the refinability used in conjunction with **405813007|Procedure site - Direct (attribute)|** is always optional (**Refinability=1**).

```
1311000087107|Ultrasonography of right wrist (procedure)|:
    {260686004|Method (attribute)|=
        278292003|Ultrasound imaging - action (qualifier value)|,
    405813007|Procedure site - Direct (attribute)|=
        74670003|Wrist joint structure (body structure)|}
```

Figure 6-4. Defining attributes of CA extension concept "1311000087107|Ultrasonography of right wrist (procedure)|".

In the UK extension, an additional 55 used a qualifying characteristic type (**CharacteristicType=1**) and mandatory refinability (**Refinability=2**). The error in this case is that the relationship types (**47429007|Associated with (attribute)|**, **118168003|Specimen source morphology (attribute)|**, **118169006|Specimen source topography (attribute)|**, **363698007|Finding site (attribute)|**) used are not linked to qualifier ranges. It is likely that characteristic type should have been defining (**CharacteristicType=0**) instead of qualifying (**CharacteristicType=1**). While the core does include 343,686 relationships that have a mandatory refinability (**Refinability=2**), they refer to only qualifying (**CharacteristicType=1**) relationships. The refinability was set to mandatory (**Refinability=2**) probably because the concepts are not well defined and the range concepts are very high level (e.g., **49062001|Device (physical object)|**, **280115004|Acquired body structure (body structure)|**).

There were 64 historical relationships from the CA extension that used the incorrect **CharacteristicType**. There were 32 concepts that had the relationship type of **370125004|MOVED TO (attribute)|** and **384598002|MOVED FROM (attribute)|** each, which should have been used in conjunction with **CharacteristicType=2** but instead **CharacteristicType=0** was used.

6.4.2.3.2 Inter Component Dependencies

6.4.2.3.2.1 SNOMED CT Identifiers

All extension descriptions and relationships were found to contain valid SNOMED CT identifiers for the fields **ConceptId**, **ConceptId1**, **RelationshipType** and **ConceptId2**. However, the UK extension was found to have used a new **RelationshipType**, **84971000000100|PBCL flag true (attribute)|**, for 3,088 concepts that were linked to 44 unique extension concepts (refer to Figure 6-5). While **84971000000100|PBCL flag true (attribute)|** is a subtype of **246061005|Attribute (attribute)|**, it was not a subtype of **410662002|Concept model attribute (attribute)|**, which is where Concept Model attributes should be located. The concept **123351000000105|Pathology Bounded Code List flag setting (qualifier value)|** has 46 subtypes, 44 of which were used as defining attributes.

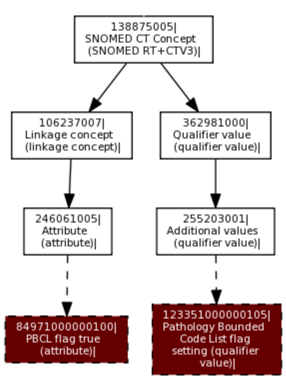


Figure 6-5. Location of new UK extension concepts 84971000000100|PBCL flag true (attribute)| and 123351000000105|Pathology Bounded Code List flag setting (qualifier value)|.

6.4.2.3.3 Concept Status and Description Status

All extension descriptions were found to have contained the appropriate corresponding concept status for each description status.

6.4.2.3.4 Concept Status and Relationship Type

The results of the inter component dependency of concept status and relationship type are shown in **Table 6-11**. The concept history attribute **168666000|SAME AS (attribute)|** was the only issue detected and it was all within the UK extension. It was narrowed down to 2,165 duplicate concepts (**ConceptStatus=2**) being linked to pending move (**ConceptStatus=11**). The international release version of SNOMED CT does not include pending move concepts. Since the UK extension is an extension concept, it may be permissible to link duplicate concepts to pending move concepts.

Table 6-11. Summary of concept status and relationship type errors.

No	Definition	CA	US	UK
1.	370124000|REPLACED BY (attribute)| should be used to link retired (1), outdated (3) or erroneous (5) to current (0) concepts.	0	0	0
2.	168666000|SAME AS (attribute)| should be used to link duplicate (2) concepts to current (0) concepts, or limited (6) concepts to other limited (6) concepts	0	0	2,165
3.	149016008|MAY BE A (attribute)| should be used to link ambiguous (4) concepts to current (0) concepts	0	0	0
4.	159083000|WAS A (attribute)| should be used to link limited (6) concepts to current concepts (0)	0	0	0
5.	370125004|MOVED TO (attribute)| should be used to link moved elsewhere (10) concepts to current (0) concepts.	0	0	0

6.4.2.4 Machine Readable Concept Model

The results of the MRCM verification are shown in **Table 6-12**. Issues were identified with the "appropriateness of attribute with domain" and "appropriateness of range with attribute" categories.

Table 6-12. Summary of Machine Readable Concept Model errors.

No	Definition	CA	US	UK
1.	Appropriateness of attribute with domain	0	4	3,088
2.	Appropriateness of range with attribute	64	6	11
3.	Concepts should either be defined with (associated finding and finding context) or (associated procedure and procedure context) but not a combination of the two	0	0	0

Domain-attribute errors were identified in the US (n=4) and UK (n=3,088) extensions. The errors in the UK extension can be traced back to the use of a new attribute **84971000000100|PBCL flag true (attribute)|** that is not part of the international MRCM. The errors in the US extension are shown in **Table 6-13**. The main issue was that **272741003|Laterality (attribute)|** was used with the domain **404684003|Clinical finding (finding)|** when it should have been used with the domain **123037004|Body structure (body structure)|**.

Table 6-13. Domain-attribute errors identified in the US extension.

No	RelationshipId	ConceptId1	RelationshipType	ConceptId2
1.	444611000124124	430091000124101|Primary right vesicoureteral reflux (disorder)|	272741003|Laterality (attribute)|	24028007|Right (qualifier value)|
2.	444841000124124	430231000124105|Primary left vesicoureteral reflux (disorder)|	272741003|Laterality (attribute)|	7771000|Left (qualifier value)|
3.	445111000124125	430391000124103|Secondary left vesicoureteral reflux (disorder)|	272741003|Laterality (attribute)|	7771000|Left (qualifier value)|
4.	445131000124120	430401000124101|Secondary right vesicoureteral reflux (disorder)|	272741003|Laterality (attribute)|	24028007|Right (qualifier value)|

The CA extension included 64 attribute-range errors only because the characteristic type of the previously mentioned 64 historical relationships used the incorrect characteristic type. The US and UK extensions contained six and 11 errors as well (refer to **Table 6-14** and **Table 6-15**). The errors in the US extension occurred as a result of the incorrect range used for **246090004|Associated finding (attribute)|**, **246454002|Occurrence (attribute)|** and **370135005|Pathological process (attribute)|**. The errors in the US extension occurred as defining attributes (**CharacteristicType=0**) while the errors in the UK extension occurred as qualifying relationships (**CharacteristicType=1**).

Table 6-14. Attribute-range errors identified in the US extension.

No	RelationshipId	ConceptId1	RelationshipType	ConceptId2
1.	455091000124121	428611000124102|Maternal human T-lymphotrophic virus type I infection (situation)|	246090004|Associated finding (attribute)|	416471007|Family history of clinical finding (situation)|
2.	455001000124125	428561000124100|Maternal substance abuse (situation)|	246090004|Associated finding (attribute)|	66214007|Substance abuse (disorder)|
3.	451811000124120	431751000124104|Cramp in lower limb associated with sleep (finding)|	246454002|Occurrence (attribute)|	309610004|During sleep (qualifier value)|
4.	451041000124126	33401000119102|Inflammatory lesion of eyelid (finding)|	370135005|Pathological process (attribute)|	257552002|Inflammation (qualifier value)|
5.	452251000124126	431941000124103|Idiopathic urethral stricture (disorder)|	370135005|Pathological process (attribute)|	54690008|Unknown (origin) (qualifier value)|
6.	452271000124120	431951000124101|Idiopathic nystagmus (disorder)|	370135005|Pathological process (attribute)|	54690008|Unknown (origin) (qualifier value)|

Table 6-15. Attribute-range errors identified in the UK extension.

No	RelationshipId	ConceptId1	RelationshipType	ConceptId2
1.	618811000000128	119339001|Stool specimen (specimen)|	118169006|Specimen source topography (attribute)|	91722005|Physical anatomical entity (body structure)|
2.	618851000000129	119350003|Calculus specimen (specimen)|	118169006|Specimen source topography (attribute)|	91722005|Physical anatomical entity (body structure)|
3.	618991000000126	127454002|Device specimen (specimen)|	118169006|Specimen source topography (attribute)|	91722005|Physical anatomical entity (body structure)|
4.	619011000000129	257261003|Swab (specimen)|	118169006|Specimen source topography (attribute)|	91722005|Physical anatomical entity (body structure)|
5.	619031000000121	258442002|Fluid sample (specimen)|	118169006|Specimen source topography (attribute)|	91722005|Physical anatomical entity (body structure)|
6.	619211000000120	258661006|Slide (specimen)|	118169006|Specimen source topography (attribute)|	91722005|Physical anatomical entity (body structure)|
7.	693071000000120	285191000000107|Specimen obtained by local excision (specimen)|	118169006|Specimen source topography (attribute)|	91722005|Physical anatomical entity (body structure)|
8.	693101000000127	405902005|Specimen obtained by radical excision (specimen)|	118169006|Specimen source topography (attribute)|	91722005|Physical anatomical entity (body structure)|
9.	1154701000000123	361361000000107|Pilonidal sinus swab (specimen)|	118169006|Specimen source topography (attribute)|	91722005|Physical anatomical entity (body structure)|
10.	1154751000000124	361381000000103|Infusion line swab (specimen)|	118169006|Specimen source topography (attribute)|	91722005|Physical anatomical entity (body structure)|
11.	4188241000000122	803301000000105|Device tip sample (specimen)|	118169006|Specimen source topography (attribute)|	91722005|Physical anatomical entity (body structure)|

6.4.2.5 Inferred Relationships

The CA, US and UK extensions included 1,625, 1,465 and 72,027 **116680003|Is a (attribute)|** relationships respectively. In the US and UK extensions, 596 (36.7%) and 39 (0.1%) of children concepts excluded the inferred relationships. Examples are shown in Table 6-16. In the first example, a new extension concept **6321000124100|On rectal examination of prostate, abnormality detected (finding)|** was added as a child of the core concept **309663007|On examination - prostate (finding)|** but no defining attributes were added for the extension concept. In the second example, a new extension concept **2071000124101|One chronic disease (situation)|** was added as a child of another new extension concept **2061000124108|Chronic disease present (situation)|** but only the latter had defining attributes. In the third example, a new **116680003|Is a (attribute)|** was added that linked two core concepts to each other **258560004|Oral saliva sample (specimen)|** and **441620008|Oral fluid specimen (specimen)|**. In addition, a new ungrouped defining attribute, **118169006|Specimen source topography (attribute)|=74262004|Oral cavity structure (body structure)|**, was added for **258560004|Oral saliva sample (specimen)|**.

Table 6-16. Examples of concepts in the US extension that were missing inferred relationships.

No	RelationshipId	ConceptId1 (SubType)	ConceptId2 (SuperType)
1.	10011000124124	6321000124100|On rectal examination of prostate, abnormality detected (finding)|	309663007|On examination - prostate (finding)|: 363698007|Finding site (attribute)|= 41216001|Prostatic structure (body structure)|, 418775008|Finding method (attribute)|= 5880005|Physical examination procedure (procedure)|, 419066007|Finding informer (attribute)|= 420158005|Performer of method (person)|
2.	2831000124122	2071000124101|One chronic disease (situation)|	2061000124108|Chronic disease present (situation)|: {246090004|Associated finding (attribute)|= 27624003|Chronic disease (disorder)|, 408729009|Finding context (attribute)|= 410515003|Known present (qualifier value)|, 408731000|Temporal context (attribute)|= 410512000|Current or specified time (qualifier value)|, 408732007|Subject relationship context (attribute)|= 410604004|Subject of record (person)|}
3.	441311000124125	258560004|Oral saliva sample (specimen)|: 118169006|Specimen source topography (attribute)|= 74262004|Oral cavity structure (body structure)| {118169006|Specimen source topography (attribute)|= 75535001|Salivary structure (body structure)|, 370133003|Specimen substance (attribute)|= 256897009|Saliva (substance)|}	441620008|Oral fluid specimen (specimen)|: {118169006|Specimen source topography (attribute)|= 74262004|Oral cavity structure (body structure)|, 370133003|Specimen substance (attribute)|= 32457005|Body fluid (substance)|}

6.4.2.6 Occurrences

This section reports on the occurrence or lack of occurrence of components and is organised by descriptions and relationships.

6.4.2.6.1 Descriptions

6.4.2.6.1.1 Fully Specified Name

There results of the fully specified name auditing are shown in **Table 6-17**.

Table 6-17. Summary of fully specified name occurrence errors.

No	Description	CA	US	UK
1.	Each current extension concept (ConceptStatus=0) should have one current (DescriptionStatus=0) fully specified name (DescriptionType=3) in English (LanguageCode=en)	0	1,266	2,111
2.	Each current extension concept (ConceptStatus=0) should have only one current (DescriptionStatus=0) fully specified name (DescriptionType=3) in English (LanguageCode=en)	0	0	0
3.	Each pending move extension concept (ConceptStatus=11) should have one pending move	0	0	0

No	Description	CA	US	UK
	(DescriptionStatus=11) fully specified name (DescriptionType=3) in English (LanguageCode=en)			
4.	Each pending move extension concept (ConceptStatus=11) should have only one pending move (DescriptionStatus=11) fully specified name (DescriptionType=3) in English (LanguageCode=en)	0	0	0
5.	Each inactive extension concept (ConceptStatus!=0 AND ConceptStatus!=6 AND ConceptStatus!=11) should have one concept inactive (DescriptionStatus=8) fully specified name (DescriptionType=3) in English (LanguageCode=en)	41	61	9,348
6.	Each inactive extension concept (ConceptStatus!=0 AND ConceptStatus!=6 AND ConceptStatus!=11) should have only one concept inactive (DescriptionStatus=8) fully specified name (DescriptionType=3) in English (LanguageCode=en)	0	0	23
7.	Each limited extension concept (ConceptStatus=6) should have one pending move (DescriptionStatus=6) fully specified name (DescriptionType=3) in English (LanguageCode=en)	0	0	247
8.	Each limited extension concept (ConceptStatus=6) should have only one pending move (DescriptionStatus=6) fully specified name (DescriptionType=3) in English (LanguageCode=en)	0	0	0

There were 1,266 and 2,111 extension concepts from the US and UK extensions respectively that did not have a current (**DescriptionStatus=0**) fully specified name (**DescriptionType=3**) description in English (**LanguageCode=EN**). One US extension concept had two fully specified names. They were both linked to the concept **428251000124104**, whose fully specified name from the concepts table was **Tetanus, diphtheria and acellular pertussis vaccination (procedure)**. However, in the descriptions table, there were two descriptions designated as fully specified names with the same concept id: **619091000124119|Tetanus, diphtheria and acellular pertussis vaccination (procedure)|** and **618061000124119|Anaphylaxis due to Haemophilus influenzae type b vaccine (disorder)|**. There were inactive concepts that were missing fully specified name descriptions with the appropriate description status.

The UK extension also included four concepts had two fully specified name description extensions, each with a language code of **en** and **en-GB** (refer to Table 6-18).

Table 6-18. Current concepts with multiple current fully specified names.

No	Concept	Fully Specified Name #1	LC #1	Fully Specified Name #2	LC #2
1.	22711000000107|GP82 arrest of dental haemorrhage claim form sent to Health Board (finding)|	1645321000000119GP82 arrest of dental haemorrhage claim form sent to Health Board (finding)|	en	47671000000114|GP82 - sent to Health Board (finding)|	en-GB
2.	295861000000105|Skin structure of center of chest (body structure)|	528231000000113|Skin structure of center of chest (body structure)|	en	521201000000114|Skin structure of centre of chest (body structure)|	en-GB
3.	310871000000100|Tumor hormone receptor status (observable entity)|	560191000000111|Tumor hormone receptor status (observable entity)|	en	560801000000114|Tumour hormone receptor status (observable entity)|	en-GB
4.	334251000000108|Hemodiafiltration route (qualifier value)|	623561000000116|Hemodiafiltration route (qualifier value)|	en	618711000000114|Haemodiafiltration route (qualifier value)|	en-GB

6.4.2.6.1.2 Preferred Term

Of the eight auditing rules for occurrences of preferred terms, only one rule was not adhered to: "Each inactive extension concept (**ConceptStatus!=0** AND **ConceptStatus!=6** AND **ConceptStatus!=11**) should have at least one concept inactive (**DescriptionStatus=8**) preferred term (**DescriptionType=1**)." It occurred 41 times in the CA extension and one time in the UK extension. The CA extension included concept statues of moved elsewhere (n=36), duplicate (n=4) and erroneous (n=1). The main issue with these 41 concepts is that the description status should have been identified as concept inactive (**DescriptionStatus=8**). In the UK extension, the concept was **338411000000102|Creme egg (substance)|** and had a concept status of retired without a stated reason (**ConceptStatus=1**). There was a preferred term **629261000000117|Creme egg|** but it should have used the description status of concept inactive (**DescriptionStatus=8**) instead of retired without a stated reason (**DescriptionStatus=1**).

6.4.2.6.2 Relationships

The results of the occurrence of relationships verification are grouped into six subsections: (1) duplicate relationships; (2) redundant relationships; (3) refined relationships; (4) concepts belonging to multiple hierarchies; (5) inactive concepts with non "is a" defining attributes or non historical relationships; and (6) erroneous historical relationships.

6.4.2.6.2.1 Duplicate Relationships

Only the US extension contained duplicate relationships and they all centred on **116680003|Is a (attribute)|** relationships. The 29 duplicate relationships are shown in Table 6-19. Twenty seven of the 29 duplicate "is a" relationships can be attributed to the duplicate relationships to the concept **112283007|Escherichia coli (organism)|**. The example of **119342007|Saliva specimen (specimen)|** and **309051001|Body fluid sample (specimen)|** is shown in Figure 6-6.

Table 6-19. Duplicate "is a" relationships contained in the US extension.

No	Core Relationship	Extension Relationship	ConceptId1 (SubType)	ConceptId2 (SuperType)
1.	2539551021	441341000124126	119342007|Saliva specimen (specimen)|	309051001|Body fluid sample (specimen)|
2.	65648024	441651000124122	246829005|Lid adhesions (finding)|	301913002|Lesion of eyelid (finding)|
3.	4637168027	5011000124126	407193009|Escherichia coli serogroup O169 (organism)|	112283007|Escherichia coli (organism)|
4.	4637207025	5061000124128	407304004|Escherichia coli serogroup O165 (organism)|	112283007|Escherichia coli (organism)|
5.	4637185025	5081000124122	407244009|Escherichia coli serogroup O164 (organism)|	112283007|Escherichia coli (organism)|
6.	4637166028	5161000124129	407187001|Escherichia coli serogroup O159 (organism)|	112283007|Escherichia coli (organism)|
7.	4637291021	5171000124120	427294000|Escherichia coli serogroup O158 (organism)|	112283007|Escherichia coli (organism)|
8.	4637157027	5181000124123	407166006|Escherichia coli serogroup O157 (organism)|	112283007|Escherichia coli (organism)|
9.	4637184026	5291000124128	407242008|Escherichia coli serogroup O152 (organism)|	112283007|Escherichia coli (organism)|
10.	4637182027	5451000124124	407238005|Escherichia coli serogroup	112283007|Escherichia coli (organism)|

No	Core Relationship	Extension Relationship	ConceptId1 (SubType)	ConceptId2 (SuperType)
			0143 (organism)|	
11.	4637176025	5461000124121	407223009|Escherichia coli serogroup 0142 (organism)|	112283007|Escherichia coli (organism)|
12.	4637204021	5531000124129	407296007|Escherichia coli serogroup 0137 (organism)|	112283007|Escherichia coli (organism)|
13.	4637163020	5601000124122	407181000|Escherichia coli serogroup 0128 (organism)|	112283007|Escherichia coli (organism)|
14.	4637175026	5621000124128	407215007|Escherichia coli serogroup 0127 (organism)|	112283007|Escherichia coli (organism)|
15.	4637174027	5631000124125	407212005|Escherichia coli serogroup 0126 (organism)|	112283007|Escherichia coli (organism)|
16.	4637173022	5651000124121	407210002|Escherichia coli serogroup 0125 (organism)|	112283007|Escherichia coli (organism)|
17.	4637180024	5661000124123	407232006|Escherichia coli serogroup 0124 (organism)|	112283007|Escherichia coli (organism)|
18.	4637172028	5741000124129	407208004|Escherichia coli serogroup 0119 (organism)|	112283007|Escherichia coli (organism)|
19.	4637171024	5881000124125	407205001|Escherichia coli serogroup 0114 (organism)|	112283007|Escherichia coli (organism)|
20.	4637200028	6031000124128	407282001|Escherichia coli serogroup 0104 (organism)|	112283007|Escherichia coli (organism)|
21.	4637197029	6261000124129	407274001|Escherichia coli serogroup 083 (organism)|	112283007|Escherichia coli (organism)|
22.	4637162026	6331000124125	407178005|Escherichia coli serogroup 078 (organism)|	112283007|Escherichia coli (organism)|
23.	4637169024	6561000124127	407195002|Escherichia coli serogroup 055 (organism)|	112283007|Escherichia coli (organism)|
24.	4637193025	6681000124120	407266002|Escherichia coli serogroup 045 (organism)|	112283007|Escherichia coli (organism)|
25.	4635974027	6821000124120	131260002|Escherichia coli serogroup 026 (organism)|	112283007|Escherichia coli (organism)|
26.	4637158021	6851000124127	407167002|Escherichia coli serogroup 025 (organism)|	112283007|Escherichia coli (organism)|
27.	4637192024	6901000124121	407261007|Escherichia coli serogroup 022 (organism)|	112283007|Escherichia coli (organism)|
28.	4637155024	6961000124122	407162008|Escherichia coli serogroup 015 (organism)|	112283007|Escherichia coli (organism)|
29.	4637191028	6971000124126	407259003|Escherichia coli serogroup 014 (organism)|	112283007|Escherichia coli (organism)|

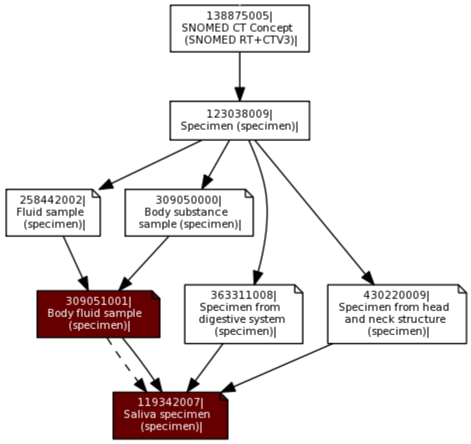


Figure 6-6. Example of duplicate "is a" relationships in the US extension (dotted-lines represent duplicate relationships).

6.4.2.6.2.2 Redundant Relationships

The US extension contained 14 redundant **16680003|Is a (attribute)|** relationships (refer to **Table 6-20** and **Figure 6-7**). An example of a redundant relationship is between **246829005|Lid adhesions (finding)|** and **301913002|Lesion of eyelid (finding)|** (**RelationshipId=65648024**) in which an additional "is a" relationship was added (**RelationshipId=441651000124122**). In addition, a relationship between **301913002|Lesion of eyelid (finding)|** and **404684003|Clinical finding (finding)|** (**RelationshipId=441031000124122**) was added. This is unnecessary as **301913002|Lesion of eyelid (finding)|** is already a subtype of **404684003|Clinical finding (finding)|**.

Table 6-20. Redundant "is a" relationships in the US extension.

No	RelationshipId	ConceptId1 (SubType)	ConceptId2 (SuperType)
1.	3841000124124	133121000119109|Severe recurrent seasonal major depression (disorder)|	58184002|Recurrent disease (disorder)|
2.	441031000124122	301913002|Lesion of eyelid (finding)|	404684003|Clinical finding (finding)|
3.	442711000124123	429261000124105|Dissecting aneurysm of basilar artery (disorder)|	9406001|Dissecting aneurysm of artery (disorder)|
4.	447551000124129	431531000124101|Health literacy assessment (procedure)|	386053000|Evaluation procedure (procedure)|
5.	452941000124123	432081000124100|Counseling for alternatives to driving (procedure)|	409063005|Counseling (procedure)|
6.	452981000124128	432091000124102|Counseling for risks of driving (procedure)|	409063005|Counseling (procedure)|
7.	454401000124122	432441000124100|Intermittent pneumatic compression device (physical object)|	272181003|Clinical equipment and/or device (physical object)|
8.	454431000124125	428421000124107|Intermittent pneumatic compression sleeve (physical object)|	303476002|Cardiovascular equipment (physical object)|
9.	454451000124121	428411000124104|Intermittent pneumatic compression boot (physical object)|	303476002|Cardiovascular equipment (physical object)|
10.	455761000124126	428671000124106|Mother receiving antimetabolite therapy (situation)|	297249002|Family history of procedure (situation)|
11.	455781000124120	428681000124109|Mother receiving antiretroviral medication therapy (situation)|	297249002|Family history of procedure (situation)|

No	RelationshipId	ConceptId1 (SubType)	ConceptId2 (SuperType)
12.	6941000124123	443189004|Enteropathogenic Escherichia coli serogroup O18 (organism)|	112283007|Escherichia coli (organism)|
13.	7051000124121	407156003|Enterotoxigenic Escherichia coli, serotype O8 (organism)|	112283007|Escherichia coli (organism)|
14.	7161000124122	406627005|Enterotoxigenic Escherichia coli serogroup O6 (organism)|	112283007|Escherichia coli (organism)|

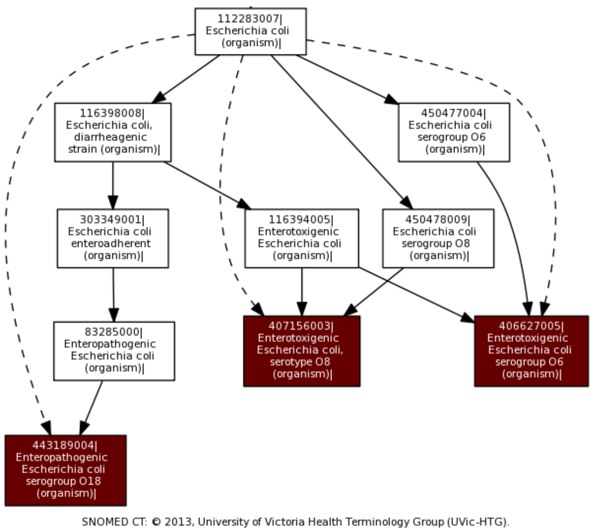


Figure 6-7. Example of redundant "is a" relationships in the US extension.

6.4.2.6.2.3 Refined Relationships

There were three relationships in the US extension that were refinements of current defining attributes (refer to **Table 6-21**). In the first example **300920004|Carotid atherosclerosis (disorder)|**, **116676008|Associated morphology (attribute)|=38716007|Atherosclerosis (morphologic abnormality)|** was added. In this case **38716007|Atherosclerosis (morphologic abnormality)|** is a subtype of **28960008|Arteriosclerosis (morphologic abnormality)|**. In the second example **57684003|Parenchymal renal hypertension (disorder)|**, **363698007|Finding site (attribute)|=29704000|Structure of parenchyma of kidney (body structure)|** was added. In this case **29704000|Structure of parenchyma of kidney (body structure)|** is a subtype of **64033007|Kidney structure (body structure)|**. In the third example, **392163005|Neuroretinal rim finding (finding)|**, **363698007|Finding site (attribute)|=431321000124109|Structure of neuroretinal rim of optic disc (body structure)|** was added. In this case **431321000124109|Structure of neuroretinal rim of optic disc (body structure)|** is a new extension concept and is a subtype **of 81016008|Optic disc structure (body structure)|**. The core relationship concepts and the extension refined relationship concepts are shown in **Figure 6-8**.

Table 6-21. Refined relationships in the US extension.

No	Concept and Defining Attributes (Extension in Bold)	Relationship
1.	300920004|Carotid atherosclerosis (disorder)|: {116676008|Associated morphology (attribute)|= 28960008|Arteriosclerosis (morphologic abnormality)|, **116676008|Associated morphology (attribute)|=** **38716007|Atherosclerosis (morphologic abnormality)|**, 363698007|Finding site (attribute)|= 69105007|Carotid artery structure (body structure)|}	441611000124121
2.	57684003|Parenchymal renal hypertension (disorder)|: **363698007|Finding site (attribute)|=** **29704000|Structure of parenchyma of kidney (body structure)|**, 363698007|Finding site (attribute)|= 51840005|Systemic circulatory system structure (body structure)|, 363698007|Finding site (attribute)|= 64033007|Kidney structure (body structure)|, 363705008|Has definitional manifestation (attribute)|= 24184005|Finding of increased blood pressure (finding)|, 47429007|Associated with (attribute)|= 90708001|Kidney disease (disorder)|	441591000124126
3.	392163005|Neuroretinal rim finding (finding)|: **363698007|Finding site (attribute)|=** **431321000124109|Structure of neuroretinal rim of optic disc (body structure)|**, 363698007|Finding site (attribute)|= 81016008|Optic disc structure (body structure)|	441601000124123

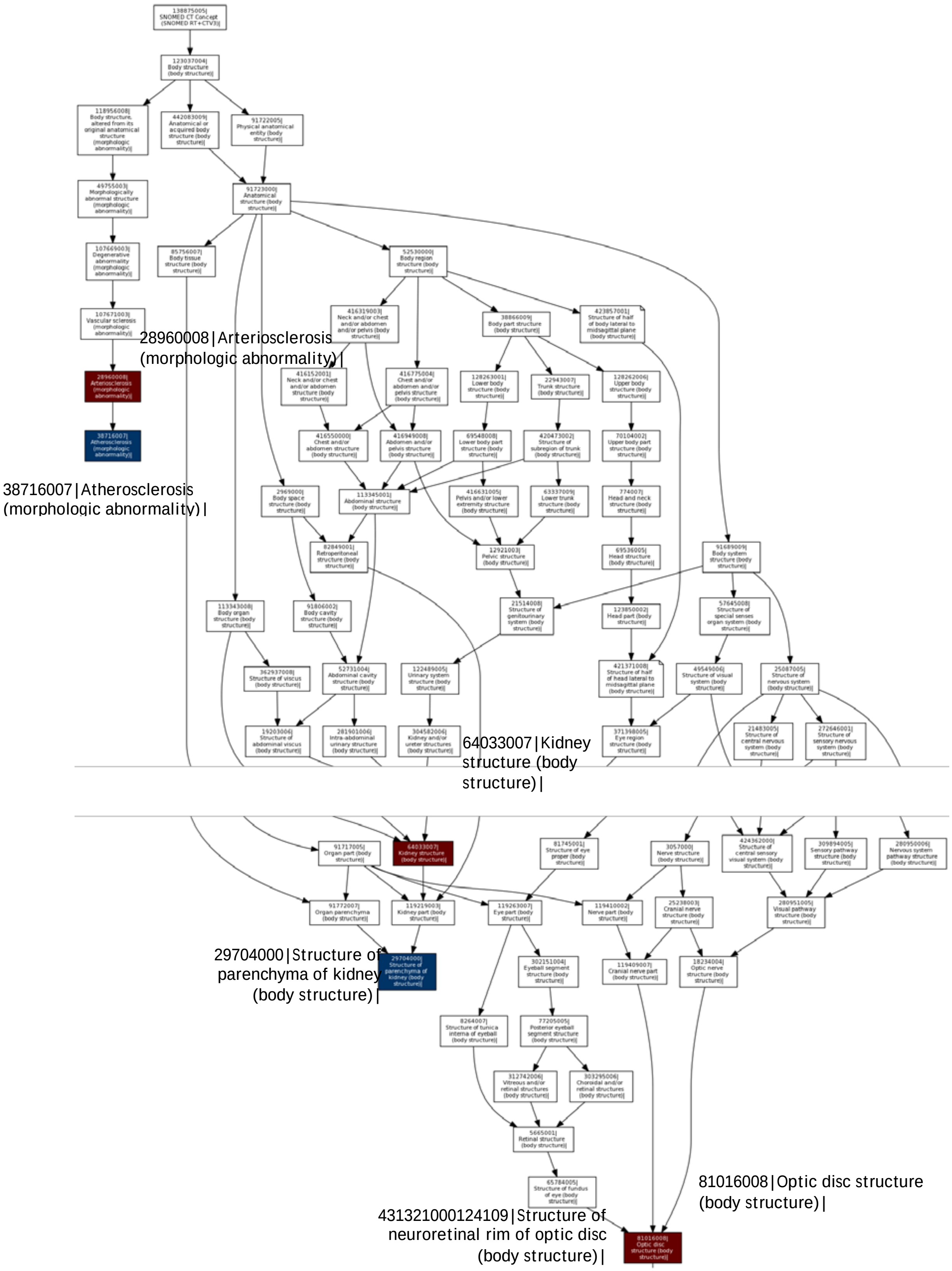


Figure 6-8. Location of concept used in the refining of core concepts (red shows the original definitions; blue shows the new refined definitions).

6.4.2.6.2.4 Concept Belongs to Multiple Hierarchies

Two concepts from the US extension were found to belong to two hierarchies (refer to **Figure 6-9**). The first, **428251000124104|Tetanus, diphtheria and acellular pertussis vaccination (procedure)|**, should only belong to the **71388002|Procedure (procedure)|** rather than the **404684003|Clinical finding (finding)|** hierarchy as the suffix indicates "procedure." The second, **1111000119100|Sebaceous nevus (disorder)|**, is a current (**ConceptStatus=0**) extension concept but has a relationship, **451961000124126**, that links it using **116680003|Is a (attribute)|** to a current concept **239107007|Epidermal nevus (disorder)|**. The complication is that **1111000119100|Sebaceous nevus (disorder)|** is also a core concept but has a concept status of duplicate (**ConceptStatus=2**). As a duplicate concept, there is a historical relationship (**168666000|SAME AS (attribute)|**) that links it to an active concept (**52298009|Linear sebaceous nevus sequence (disorder)|**).

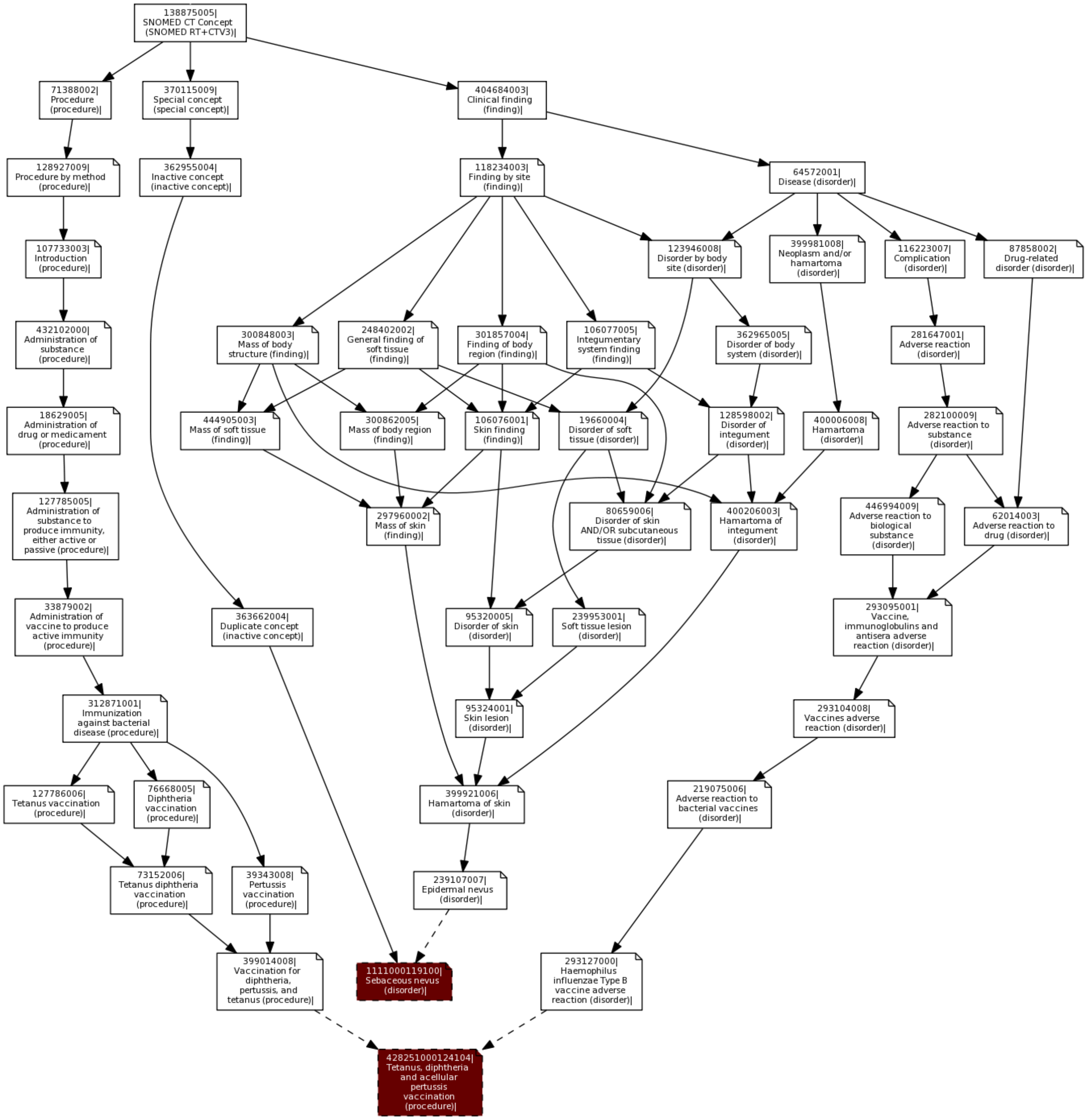


Figure 6-9. Location of "428251000124104|Tetanus, diphtheria and acellular pertussis vaccination (procedure)|" and "1111000119100|Sebaceous nevus (disorder)|" in the hierarchy.

6.4.2.6.2.5 Inactive Concepts with Non "Is A" Defining Attributes or Non Historical Relationships

There were 33 concepts from the CA extension that had non"is a" defining attributes. However, 32 of them were due to the incorrect **CharacteristicType** being used therefore only inactive concept had defining attributes (refer to **Figure 6-10**). The concept **935907381000087100|Congenital rubella syndrome (inactive concept)|** has a concept status of moved elsewhere (**ConceptStatus=10**) but it has four sets of defining attributes. It should also be noted that the suffix **inactive concept** is invalid.

```
935907381000087100|Congenital rubella syndrome (inactive concept)|:
        246075003|Causative agent (attribute)|=
                5210005|Rubella virus (organism)|,
        246454002|Occurrence (attribute)|=
                255399007|Congenital (qualifier value)|,
        363705008|Has definitional manifestation (attribute)|=
                238812004|Exanthematous disorder (disorder)|,
        370135005|Pathological process (attribute)|=
                441862004|Infectious process (qualifier value)|
```

Figure 6-10. Inactive concept in the Canada extension that had defining attributes.

6.4.2.6.2.6 Erroneous Historical Relationships

Of the historical relationships auditing rules, only one was not followed. The CA extension **935907381000087100|Congenital rubella syndrome (inactive concept)|** has a concept status of moved elsewhere (**ConceptStatus=10**). Instead of being a child concept of **370126003|Moved elsewhere (inactive concept)|**, it was a child concept of **733956191000087102|Congenital rubella (disorder)|**, a current concept (refer to **Figure 6-11**).

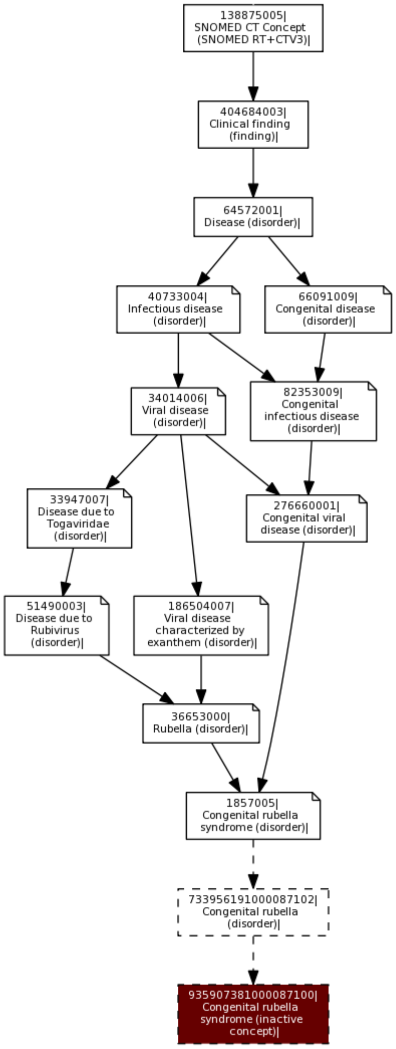


Figure 6-11. Location of "935907381000087100|Congenital rubella syndrome (inactive concept)|" in the hierarchy.

6.5 Discussion

In this study, an auditing method was developed to verify that extensions conform to the guidelines specified in the UG and TIG. A total of 89 verification rules were identified, organised according to the SNOMED CT Extensions Auditing Framework, represented in a machine-readable format and applied to three country-level extensions. The discussion section is divided into three parts. First, observations of the development of the auditing method. Second, a discussion on the errors contained in the extensions. Third, implications and challenges of using extensions. The limitations of the auditing method and future work are described in **Chapter Ten**.

6.5.1 Auditing Method

Compiling the list of verification guidelines that formed the auditing method was not overly difficult but was a tedious process. For example, verifying that retired (**ConceptStatus=1**), outdated (**ConceptStatus=3**) or erroneous (**ConceptStatus=5**) concepts should have a concept history attribute of **370124000|REPLACED BY (attribute)|** that was linked to a current concept required a SQL statement that included three "left joins" and multiple "or" conditions. Another example was verifying the presence of a fully specified name and preferred term for inactive concepts where pending move, limited and other inactive concepts required different description statuses. Additional unintended identification of errors was discovered when generating routine frequency count reports. For example, there wasn't an intention to check whether inactive concepts were fully defined. But when the analysis results of the **IsPrimitive** status showed that the number of active fully defined concepts did not match the total number of fully defined concepts, an inter dependency verification rule was created to identify this type of error.

As demonstrated in this study, auditing extensions for consistency is not difficult and for the most part can be carried out using SQL. As not all health care organisations have the resources to develop these guidelines, the IHTSDO should make a similar set of these reusable guidelines available. It is evident that health care organisations are not running these verification guidelines as demonstrated by the errors identified. By auditing the extensions before using them within their own health care organisation or releasing them as country-wide extensions, it will ensure that distributed extensions will not contain inconsistencies.

6.5.2 Errors in Extensions

All three SNOMED CT extension reviewed contained errors with varying consequences. Non-critical errors such as inactive concepts that are considered fully defined concepts, inactive concepts with non **16680003|Is a (attribute)|** defining attributes or duplicate **16680003|Is a (attribute)|** relationships should have little impact on implementation. On the other hand, critical errors such as missing inferred relationships, incorrect domain-attribute and attribute-range definition use, incorrect characteristic type and missing historical relationships will cause problems when testing for equivalency and subsumption. Going forward, descriptions that use the incorrect language code for the fully specified name and preferred term will have less of an impact as preferred terms have been moved to reference sets and the language codes in the descriptions have all been set to English (**en**) in RF2.

6.5.3 Implications and Challenges

Semantic Interoperability. One of the reasons for using standardised clinical reference terminologies is to ensure interoperability as every user of the terminology will view exactly the same content in a release version. With extensions, there is a likelihood of diminishing semantic interoperability. As demonstrated in the analysis, the US extensions are changing the core by adding additional "is a" relationships between core concepts and re-defining the defining attributes of core concepts. When adding new "is a" relationships to SNOMED CT, five new "types" of "is a" relationships may arise (refer to Figure 6-12). First, new "is a" relationships between core concepts may be created. Second, an extension concept may be a child concept of a core concept. Third, a core concept may be a child of an extension concept. Fourth, an extension concept may be a child of an extension concept. Fifth, "is a" relationships between core concepts may be deleted. This is possible only through the RF2 as that format includes an active field that can deactivate core relationships. Since this analysis was based on RF1 extensions, it was out of scope in this study. The addition of "is a" relationships will yield different results for subsumption queries.

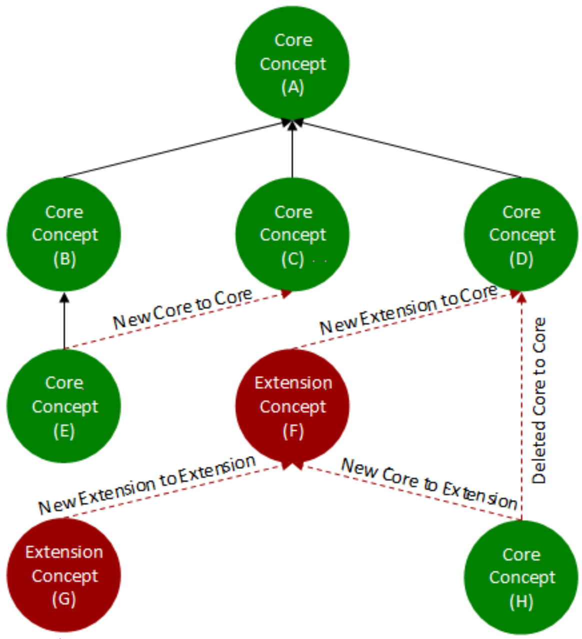


Figure 6-12. Types of "is a" relationships that result from new extension concepts.

The question arises on how to ensure semantic interoperability if extension concepts are transmitted to external organisations. If the extension concept is fully defined and does not include any extension concepts as a proximal primitive or as part of the defining attributes, sending the normal form that comprises only of core concepts would be an alternative. For example, **1171000119108|Recurrent inguinal hernia (disorder)|** is a fully defined concept in the US extension with the parent concept of **396232000|Inguinal hernia (disorder)|**. **1171000119108|Recurrent inguinal hernia (disorder)|** is differentiated from **396232000|Inguinal hernia (disorder)|** with the defining attributes of **263502005|Clinical course (attribute)|=255227004|Recurrent (qualifier value)|**. Therefore, instead of sending **1171000119108|Recurrent inguinal hernia (disorder)|** to an external organisation, it would be better to use **396232000|Inguinal hernia (disorder)|:263502005|Clinical course (attribute)|=255227004|Recurrent (qualifier value)|**. If the extension concept is a primitive concept, or the proximal primitives or defining attributes contain extension concepts, an alternative solution would be needed. In a survey of SNOMED CT implementations,[12] the respondents indicated that the parent core concept of an extension would be used when sending a SNOMED CT coded document beyond their health care organisation. While this does reduce the granularity of what was recorded, it is the only known feasible solution. The impact of using a new attribute, such as in the UK extension (**84971000000100|PBCL flag true (attribute)|**), remains to be seen.

Improved Definitions. The proportion of active primitive concepts in the CA (1,333/1,337=99.7%) and UK (23,940/23,986=99.8%) extensions were significantly higher than the core (229237/296237=77.3%) while the US extension was slightly better than the core (1,030/1,266=81.4%). While primitive concepts are unavoidable in hierarchies such as **410607006|Organism (organism)|**, **362981000|Qualifier value (qualifier value)|** and **48176007|Social context (social concept)|** because the MRCM does not allow these concepts to be defined, concepts in hierarchies such as **404684003|Clinical finding (finding)|** and **71388002|Procedure (procedure)|** should be defined as fully as possible. For example, the primitive CA extension concept **11000087100|Fluoroscopy of left hip (procedure)|** has been defined as a child of **418880005|Fluoroscopy of hip (procedure)|** with the defining attributes of **260686004|Method (attribute)|=312275004|Fluoroscopic imaging - action (qualifier value)|** and **405813007|Procedure site - Direct (attribute)|=29836001|Hip region structure (body structure)|**. Instead of using **29836001|Hip region structure (body structure)|**, the subtype **287679003|Left hip region structure (body structure)|** should have been used, which would allow the concept to be considered fully defined.

6.6 References

[1] Lee, D., de Keizer, N., Lau, F., & Cornet, R. (2013). Literature review of SNOMED CT use. Journal of the American Medical Informatics Association. http://www.ncbi.nlm.nih.gov/pubmed/23828173.

[2] Lee, D. H., Lau, F. Y., & Quan, H. (2010). A method for encoding clinical datasets with SNOMED CT. BMC medical informatics and decision making, 10(1), 53. http://www.ncbi.nlm.nih.gov/pubmed/20849611.

[3] Green, J. M., Wilcke, J. R., Abbott, J., & Rees, L. P. (2006). Development and evaluation of methods for structured recording of heart murmur findings using SNOMED-CT® post-coordination. Journal of the American Medical Informatics Association, 13(3), 321-333. http://www.ncbi.nlm.nih.gov/pubmed/16501179.

[4] SNOMED CT Namespace Registry. http://www.ihtsdo.org/fileadmin/user_upload/Docs_01/SNOMED_CT/Namespaces/SNOMED_CT_Namespace_Registry_-_OFFICIAL_20130611.pdf.

[5] Dolin, R. H., Mattison, J. E., Cohn, S., Campbell, K. E., Wiesenthal, A. M., Hochhalter, B., ... & Zingo, C. (2004). Kaiser Permanente's convergent medical terminology. Medinfo, 11(Pt 1), 346-50. http://www.ncbi.nlm.nih.gov/pubmed/15360832.

[6] Ceusters, W., Smith, B., Kumar, A., & Dhaen, C. (2004). Mistakes in medical ontologies: where do they come from and how can they be detected? Studies in Health Technolology and Informatics. 102:145-63. http://www.ncbi.nlm.nih.gov/pubmed/15853269.

[7] Wang, Y., Halper, M., Min, H., Perl, Y., Chen, Y., & Spackman, K. A. (2007). Structural methodologies for auditing SNOMED. Journal of Biomedical Informatics, 40(5), 561-581. http://www.ncbi.nlm.nih.gov/pubmed/17276736.

[8] Min, H., Perl, Y., Chen, Y., Halper, M., Geller, J., & Wang, Y. (2006). Auditing as part of the terminology design life cycle. Journal of the American Medical Informatics Association, 13(6), 676-690. http://www.ncbi.nlm.nih.gov/pubmed/16929044.

[9] Cornet, R., & Abu-Hanna, A. (2005). Description logic-based methods for auditing frame-based medical terminological systems. Artificial Intelligence in Medicine, 34(3), 201-217. http://www.ncbi.nlm.nih.gov/pubmed/15994071.

[10] International Health Terminology Standards Development Organisation. User Guide. http://www.ihtsdo.org/fileadmin/user_upload/doc/ug/index.html.

[11] International Health Terminology Standards Development Organisation. Technical Implementation Guide. http://www.ihtsdo.org/fileadmin/user_upload/doc/tig/index.html.

[12] Lee, D., Cornet, R., Lau, F., & De Keizer, N. (2012). A survey of SNOMED CT implementations. Journal of biomedical informatics, 46(1), 87-96. http://www.ncbi.nlm.nih.gov/pubmed/23041717.

7. Encoding Method & Results

7.1 Introduction

The majority of published papers on the Systematised Nomenclature of Medicine Clinical Terms (SNOMED CT) have focused on encoding local terms with SNOMED CT and comparing SNOMED CT to other terminology systems.[1] When pre-coordinated concepts are inadequate to represent the full meaning of a local term, the compositional feature of SNOMED CT allows for the combining of multiple concepts together to create a post-coordinated expression that can convey the full meaning of the local term. This compositional feature of SNOMED CT is important because it is not feasible to create pre-coordinated concepts for every single clinical term because it will vastly increase the size of SNOMED CT, make the maintenance of SNOMED CT more difficult and make it harder for health care organisations to implement an even larger terminology. There are four main types of post-coordination: Concept Model, refinement, qualification and combination. The Concept Model is a broad set of guidelines that define how different domains and ranges can be linked together using Concept Model attributes. Refinement is where a defining attribute of a concept is refined to a more granular concept. Qualification is where qualifiers are used to modify a concept. Combination is where two concepts are joined together using a plus (+) sign.

As over a hundred papers have described how local terms have been mapped to SNOMED CT using both pre-coordinated concepts and post-coordinated expressions, the question arises, why is there a need for developing an encoding method? The main reason is that there are inadequate guidelines on creating post-coordinated expressions and the literature includes examples of errors in post-coordinated expressions.[2, 3] Post-coordinated expressions that contain errors can reduce the accuracy of the clinical statement they are supposed to represent as well as reduce the accuracy of data retrieval when testing for equivalency and subsumption.

In a previous study,[4] the need for additional guidelines for crafting post-coordinated expressions was discussed as a next step along with the challenge of abbreviations and acronyms when using lexical matching. While the ultimate goal is to develop an *algorithm* to enable automated post-coordination, that was not realistic given the timeframe of this PhD research and the complexity of post-coordination. Therefore the objective of this study was to develop a *method* for automatically creating post-coordination expressions from free text clinical statements that conform to the SNOMED CT Concept Model. It should be noted that all automatically encoded clinical statements still should be manually reviewed and that the encoding method is designed to suggest the best possible match. The method developed adds to the current state of knowledge by defining how multiple exact matches should be handled, what concepts should be excluded in lexical matching, a proposed Extended Concept Model to help improve the accuracy of post-coordinated expressions, the proposed use of encoding templates and the nine scenarios of modeling concepts.

The rest of this section describes some of the challenges of acronyms, abbreviations and synonyms in SNOMED CT, similar descriptions and exact matches, and errors in post-coordinated expressions found in the literature, which supports the case for needing a post-coordinated encoding method. The rest of this chapter is organised into the materials used, the method developed, the results of using the encoding method to encode problem lists and encounter diagnoses from a primary care dataset (Ethics Waiver: Protocol Number 10-143) and 10 clinical statements from the literature, a discussion of the encoding method and results, and finally the conclusion. For brevity, detailed examples and explanations have been moved into **Appendix E**.

7.1.1 Inconsistent or Incomplete Acronyms, Abbreviations and Synonyms

Abbreviations and acronyms in SNOMED CT are inconsistently applied to descriptions. For example, the concept **13645005|Chronic obstructive lung disease (disorder)|** is represented with seven descriptions using a mixture of acronyms and expanded terms. The challenge arises when a term "COAD" needs to be encoded and there is no exact match since the description that includes "COAD" also includes the expanded term "chronic obstructive airway disease." Refer to **Table 12-33** in **Appendix E** for all seven descriptions. Moreover, there is no indication of whether a description refers to an acronym. For example, **23289016|COLD|** refers to "chronic obstructive lung disease" but from the description itself, it is not possible to know that. The **InitialCapitalStatus** data element indicates that the first letter should be capitalised, but there is little other information and there is the possibility of confusing it with descriptions such as **139573015|Cold|**, which refers to **84162001|Cold sensation quality (qualifier value)|** or **504996015|Cold|**, which refers to **82272006|Common cold (disorder)|**.

A second example is the way in which cancer is represented. A review of concepts with a fully specified name that starts with "malignant tumor" revealed that there are at least six different descriptions of representing cancer (refer to **Table 12-34** in **Appendix E**) and the synonyms are incomplete. From that table, it can be seen there are descriptions lacking for "cancer of breast," "lung cancer," "cancer of pelvis" and "pelvic cancer." The challenge arises when these terms need to be encoded and there are no lexical matches.

7.1.2 Similar or Same Descriptions

There are over 13,000 descriptions in SNOMED CT that are very similar or exactly the same, both within the same hierarchy and across different hierarchies. An example of concepts that are similar are **229738006|Word finding difficulty (disorder)|** and **286384007|Difficulty finding words (finding)|**. An example of concepts that have the same description are **263208005|Fracture of distal end of radius and ulna (disorder)|** and **82065001|Fracture of carpal bone (disorder)|**, which both have a description of "fracture of wrist," (i.e., **391365014|Fracture of wrist|** and **136124017|Fracture of wrist|**) but they are unrelated concepts. While exact matches are usually the gold standard,[5, 6, 7] in some cases, an exact match is not always the appropriate concept. For example, "warts" is exactly matched to **50689014|Warts|**, which belongs to the concept **30285000|Verruca (morphologic abnormality)|**. The concept desired is actually **57019003|Verruca vulgaris (disorder)|** but it only has a synonym of "wart" and not "warts." These subtle differences can cause a clinician to use an incorrect term.

7.1.3 Errors in Post-coordination in Literature

A review of some of the examples of post-coordinated expressions included in the literature were found to contain errors in the form of syntax errors,[5] incorrect type of post-coordination,[8, 9, 10, 11] incorrect domains,[12] incorrect attributes[13] and incorrect ranges.[14] For example **3457005|Patient referral (procedure)|:370131001|Recipient category (attribute)|=66862007|Radiologist (occupation)|**[13] is incorrect because **66862007|Radiologist (occupation)|** is not a valid range for **370131001|Recipient category (attribute)|**. Additional examples of the errors are shown in **Appendix E**.

7.2 Materials

An anonymised dataset was extracted from a primary care electronic medical record (EMR) system. The dataset contained 13,013 patient records and included appointments, problem lists, encounter diagnoses, laboratory results, billing diagnostic codes, medications and examinations. The data ranged from June 27, 2001 to December 20, 2012. The EMR system contained a master problem list that was used to populate the problem lists and encounter diagnoses and was partially mapped to the International Statistical Classification of Diseases and Related Health Problems, Ninth Revision, Clinical Modification (ICD-9-CM). When clinicians could not find a term they wanted in the master problem list, they would enter the free text into the problem list or encounter diagnosis. In this study, the free text of the problem lists (n=20,141) and encounter diagnoses (n=266,032) were extracted for encoding with SNOMED CT. The release version of SNOMED CT used to develop the method was the July 31, 2012 international release version in the release format 1 (RF1).

Ten clinical statements were extracted in from the literature (which were also discussed in **Section 7.1.3** and **Appendix E**) to determine if the encoding method could encode the clinical statements correctly (refer to **Table 7-1**).

Table 7-1. Ten clinical statements from the literature and the references.

No	Clinical Statement	Reference
1.	Multiple chemical sensitivity	Sampalli, T., Shepherd, M., Duffy, J., & Fox, R. (2010). An evaluation of SNOMED CT® in the domain of complex chronic conditions. International Journal of Integrated Care, 10.[5]
2.	Heightened sensitivity to environment	Sampalli, T., Shepherd, M., Duffy, J., & Fox, R. (2010). An evaluation of SNOMED CT® in the domain of complex chronic conditions. International Journal of Integrated Care, 10.[5]
3.	Cellulitis of leg	Grain, H. (2010). Clinical terminology. Studies in Health Technology and Informatics, 151, 70-83.[10]
4.	Spirometry without bronchodilation	Randorff, H. A., & Rosenbeck, G. K. (2012). SNOMED CT Implementation. Mapping Guidelines Facilitating Reuse of Data. Methods of Information in Medicine, 51(6).[9]
5.	Effective parenting	Park, H. A., Lundberg, C. B., Coenen, A., & Konicek, D. J. (2009). Evaluation of the content coverage of SNOMED-CT to represent ICNP Version 1 catalogues. Studies in Health Technology and Informatics, 146, 303.[12]
6.	Pain; mouth	Wang Y, Patrick J, Miller G, O'Hallaran J. A computational linguistics motivated mapping of ICPC-2 PLUS to SNOMED CT. BMC Medical Informatics and Decision Making. 2008 Oct 27;8 Suppl 1:S5.[13]

No	Clinical Statement	Reference
7.	Referral; radiologist	Wang Y, Patrick J, Miller G, O'Hallaran J. A computational linguistics motivated mapping of ICPC-2 PLUS to SNOMED CT. BMC Medical Informatics and Decision Making. 2008 Oct 27;8 Suppl 1:S5.[13]
8.	Abuse; verbal; relative	Wang Y, Patrick J, Miller G, O'Hallaran J. A computational linguistics motivated mapping of ICPC-2 PLUS to SNOMED CT. BMC Medical Informatics and Decision Making. 2008 Oct 27;8 Suppl 1:S5.[13]
9.	Dislocation; knee; simple	Wang Y, Patrick J, Miller G, O'Hallaran J. A computational linguistics motivated mapping of ICPC-2 PLUS to SNOMED CT. BMC Medical Informatics and Decision Making. 2008 Oct 27;8 Suppl 1:S5.[13]
10.	Age of first statin	Pathak, J., Wang, J., Kashyap, S., Basford, M., Li, R., Masys, D. R., & Chute, C. G. (2011). Mapping clinical phenotype data elements to standardized metadata repositories and controlled terminologies: the eMERGE Network experience. Journal of the American Medical Informatics Association, 18(4), 376-386.[14]

7.3 Methods

In this study, the encoding method was developed and applied against the problem lists and encounter diagnoses of the primary care dataset. The post-coordinated expressions that were generated that had a frequency of ≥10 were manually reviewed for accuracy as well as a random selection of one percent of all post-coordinated expressions. In addition, 10 clinical statements from the literature where encoded using the encoding method and were manually reviewed for accuracy. The manual review was conducted by me and the correct encoding shown in the results were based on my judgment of how the problem, encounter or clinical statement from the literature should have been encoded.

The encoding method developed in this study consists of three parts. First, improving the quality of search results. This is achieved by improving the content coverage by generating additional acronyms and synonyms, excluding undesired concepts such as unsanctioned qualifier values and groups of anatomical entity concepts that should not be used, and sorting the results to ensure the most appropriate concept is retrieved first. Second, a nine-scenario process of modeling two or more expressions. These scenarios include checking whether the most appropriate concepts are used when modeling and identifying the appropriate Concept Model attribute. Third, cleaning up the post-coordinated expression. This includes inferring context to ensure the appropriate context is used in clinical statements that contain multiple conditions, the addition of required attributes to ensure the post-coordinated expression is complete and to simplify the expression by locating pre-coordinated concepts for portions of or the whole post-coordinated expression where applicable. The development of the encoding method was iterative and lessons learned from reviewing the results of earlier rounds of encoding were incorporated into the method that was used in subsequent rounds. A summary of the encoding method is shown in **Table 7-2**. The rest of this methods section is devoted to explaining each part of the encoding method.

Table 7-2. Summary of encoding method.

Part 1: Improving the quality of the search results	Part 2: Modeling two or more expressions together	Part 3: Cleaning up the post-coordinated expression
▪ Improving content coverage - Generating acronyms - Generating synonyms from	▪ Candidate is laterality and predicate includes body structure ▪ Candidate is a predicate's defining	▪ Inferring context ▪ Additional post-coordination ▪ Compacting post-coordinated expression

Part 1: Improving the quality of the search results	Part 2: Modeling two or more expressions together	Part 3: Cleaning up the post-coordinated expression
eponymously named diseases - Generating synonyms from implicit clinical findings ▪ Excluding undesired concepts and descriptions - Unsanctioned quality value - Groups of anatomical entities - Concept clusters - Exclude descriptions ▪ Sorting search results - Concepts from different hierarchies - Concepts with an "is a" relationship - A description is the preferred term - A description is more similar to a fully specified name - Manual review	attribute ▪ Candidate is a refinement of predicate's defining attribute ▪ Candidate uses same attribute of predicate's defining attribute ▪ Candidate is a body structure and predicate is a finding or procedure ▪ Encoding template exist for predicate and candidate ▪ Concept model contains appropriate attribute ▪ Candidate is contextual qualifier and predicate is a finding or procedure ▪ Candidate is explicit Concept Model attribute	- General concepts - Contextual qualifiers

7.3.1 Improving the Quality of Search Results

The quality of the search results can be improved by generating additional acronyms and synonyms so as to improve the content coverage, excluding concepts that should not be used in post-coordination and sorting the search results to ensure the most relevant results are retrieved first.

7.3.1.1 Improving Content Coverage

Cleaning candidate terms and developing advanced encoding algorithms can improve lexical matching results. At the same time, having a more complete target set of terms will also improve the lexical matching results. This section describes the process of generating acronyms and their expanded forms, synonyms from eponymously named diseases and implicit clinical findings from within SNOMED CT. Synonyms may also be generated from an external source such as the Unified Medical Language System (UMLS) and other terminologies that have been mapped to SNOMED CT but that is beyond the scope of this method. Structured query language (SQL) statements are provided where applicable and the results should be manually reviewed when creating lists of acronyms and synonyms as these SQL statements are meant to retrieve only potential concepts.

7.3.1.1.1 Generating Acronyms

SNOMED CT descriptions contain a mixture of acronyms and fully expanded descriptions, which can cause difficulties in lexical matching because not all acronyms are available as a description on its own (refer to the example in **Section 7.1.1**). An option would be to generate a list of all acronyms available in SNOMED CT. The first step is to retrieve all descriptions that start with a capital letter and include a dash (-). The general format of descriptions that include both an acronym and fully expanded description is that they include a dash in between them (e.g., **475427019|COPD - Chronic obstructive pulmonary disease|**). Without a manual review of all descriptions it is not possible to know the exact number. Descriptions starting with "O/E" (on examination), and "C/O" (complaint of) and "FH" (family history) should be omitted. An SQL statement is provided in **Figure 7-1**. The second step is to split the description into two parts (i.e., before and after the dash). The first part of the description should be checked if it contains all capital letters as some of the descriptions retrieved may not be an acronym. The third step is to check if the letters contained in the acronym match the second part of the description. This may inevitably omit other acronyms that do not use this structure but it should catch the majority of acronyms available. Note that there may be some acronyms that refer to different concepts so caution should be exercised even when there is an exact match to an acronym.

```
SELECT * FROM SCT_Descriptions WHERE DescriptionStatus=0 AND InitialCapitalStatus=1 AND Term NOT LIKE
'O/E%' AND Term NOT LIKE 'C/O%' AND Term NOT LIKE 'FH%' AND Term LIKE '% - %'
```

Figure 7-1. SQL statement for retrieving potential acronyms.

7.3.1.1.2 Generating Synonyms from Eponymously Named Diseases

A review of historical patient records showed that eponymously name diseases often did not include the word "disease" or "syndrome." For example, "Parkinson's disease" and "Alzheimer's disease" were often referred to as just "Parkinson's" or "Alzheimer's" respectively. The challenge is that these will result in partial matches instead of exact matches. Although there are issues with exact matches, they are still seen as better quality matches than partial matches. One option would be to generate a list of diseases that were named after individuals and this was carried using an SQL statement (refer to **Figure 7-2**). The criteria were that the description had to have an apostrophe followed by the letter s, followed by "disease," "disorder" "or syndrome." In addition, body structures should not be included. For example, a synonym of "Crohn's" was created for **34000006|Crohn's disease (disorder)|** but not for **50440006|Crohn's disease of colon (disorder)|**.

```
SELECT C.ConceptId, C.FullySpecifiedName, C.IsPrimitive, D.DescriptionId, D.Term FROM SCT_Concepts C,
SCT_Descriptions D WHERE C.ConceptId=D.ConceptId AND C.ConceptStatus=0 AND D.DescriptionStatus=0 AND
D.DescriptionType!=3 AND (D.Term LIKE "%'s dis%" OR D.Term LIKE "%'s syndrome%") ORDER BY
C.FullySpecifiedName, D.Term
```

Figure 7-2. SQL statement for retrieving potential eponymously named disease.

7.3.1.1.3 Generating Synonyms from Implicit Clinical Findings

There are clinical terms that implicitly refer to a body structure without explicitly referring to the body structure itself. For example, "basal cell carcinoma" implicitly refers to the skin, but the concept **254701007|Basal cell carcinoma of skin (disorder)|** does not include a synonym for just "basal cell carcinoma." Refer to **Table 12-36** in **Appendix E** for additional examples. From a review of historical patient records, implicit body structures are not used. However, without the synonym that excludes an implicit body structure, it is less likely an exact match will be found. One possibility would be to retrieve all **49755003|Morphologically abnormal structure (morphologic abnormality)|** concepts and retrieve **404684003|Clinical finding (finding)|** that are defined using those concepts. For each of the concepts retrieved, filter out the subtype concepts. The descriptions of the **49755003|Morphologically**

abnormal structure (morphologic abnormality)| may be checked to see if any are applicable to be added a synonyms of the **404684003|Clinical finding (finding)|** concept.

7.3.1.2 Excluding Undesired Concepts and Descriptions

In order to include only relevant concepts, there is a need to exclude descriptions and concepts that are ambiguous or should not be used. Thus far, two sets of active concepts have been identified that should be excluded from use in post-coordination. The first are unsanctioned qualifier values while the second are groups of anatomical concepts. There are also clusters of concepts that have been defined incorrectly or are ambiguous and should be avoided. The reasons for excluding them are detailed in the following sub-sections.

7.3.1.2.1 Unsanctioned Qualifier Values

There are 8,982 concepts from the **362981000|Qualifier value (qualifier value)|** hierarchy but only 1,884 (21.0%) of them can be accessed via the Concept Model. There are 21 Concept Model attributes (e.g., **272741003|Laterality (attribute)|**, **263502005|Clinical course (attribute)|** and **246456000|Episodicity (attribute)|**) that are linked to 30 ranges of qualifier values (e.g., **182353008|Side (qualifier value)|**, **288524001|Courses (qualifier value)|** and **288526004|Episodicities (qualifier value)|**). The main challenge of using unsanctioned qualifier values is that they cannot be combined with other concepts to create post-coordinated expressions while adhering to the constraints of the Concept Model. For example, encoding "pain controlled" using **22253000|Pain (finding)|** and **31509003|Controlled (qualifier value)|** is different from **225782006|Pain control (procedure)|** as the latter refers to the management of pain while the former indicates pain but the context of "controlled" is unknown. The list of exclusion concepts can be generated by excluding all **362981000|Qualifier value (qualifier value)|** concepts that are not accessible via the Concept Model (refer to **Figure 7-3**).

```
SELECT SubTypeId FROM SCT_TransitiveClosure WHERE SuperTypeId=362981000 AND SubTypeId
NOT IN (SELECT DISTINCT(TC.SubTypeId) FROM SCT_ConceptModel CM, SCT_TransitiveClosure
HR, SCT_TransitiveClosure TC WHERE CM.RangeId=HR.SubTypeId AND HR.SuperTypeId=362981000
AND CM.RangeId=TC.SuperTypeId)
```

Figure 7-3. SQL statement for excluding concepts from the "362981000|Qualifier value (qualifier value)|" hierarchy that cannot be accessed via the Concept Model.

7.3.1.2.2 Groups of Anatomical Entities

There is a section in the hierarchy that contain concepts which include multiple body structures. The supertype concept is **420864000|Group of anatomical entities (body structure)|** and there are 370 subtype concepts. Examples include **69948000|Both ankles (body structure)|**, **110736001|Bronchus and lung (combined site) (body structure)|** and **110536004|Tibia and fibula (combined site) (body structure)|**. The reason for excluding these body structures when encoding is that none of these concepts are used in any defining attributes. Any **404684003|Clinical finding (finding)|** or **71388002|Procedure (procedure)|** concepts that refer to multiple body structures use separate body structure concepts so the use of the combined sites will impair the ability to test for equivalency and subsumption when compared against other pre-coordinated concepts. Refer to **Figure 12-12** in **Appendix E** for some examples.

7.3.1.2.3 Concept Clusters

In this context, concept clusters refer to a set of concepts that are poorly defined that can impact the testing for equivalency and subsumption. One such cluster of concepts has to do with **44077006|Numbness (finding)|** (refer to **Figure 7-4**). It should be noted that only the concepts discussed below are shown in the figure in red as plotting out the entire subtype hierarchy of **44077006|Numbness (finding)|** is not viewable on this page. There are multiple issues with the definitions and hierarchical arrangement of these concepts, a few of which are listed here. First, **309537005|Numbness of lower limb (finding)|** and **298753001|Numbness of upper limb (finding)|** should be a subtype of **310501001|Numbness of limbs (finding)|**. Second, **310501001|Numbness of limbs (finding)|** should be a subtype of **102603008|Numbness of skin (finding)|**. Third, **309557009|Numbness of face (finding)|** should be a subtype of **102603008|Numbness of skin (finding)|**. Since numbness generally has to do with the **39937001|Skin structure (body structure)|**, **102603008|Numbness of skin (finding)|** should be used and refined rather than using the other pre-coordinated concepts unless numbness does not deal with a skin structure. The reason is that when querying for **310501001|Numbness of limbs (finding)|** and its subtypes, concepts such as **309537005|Numbness of lower limb (finding)|** and **298753001|Numbness of upper limb (finding)|** will be missed due to the inaccurate hierarchy. Therefore instead of using the pre-coordinated concept **309537005|Numbness of lower limb (finding)|**, an alternative would be the use post-coordination (i.e., **44077006|Numbness (finding)|:363698007|Finding site (attribute)|=371304004|skin structure of lower extremity (body structure)|**). While this may be unwieldy, it is an interim solution until the hierarchy has been fixed.

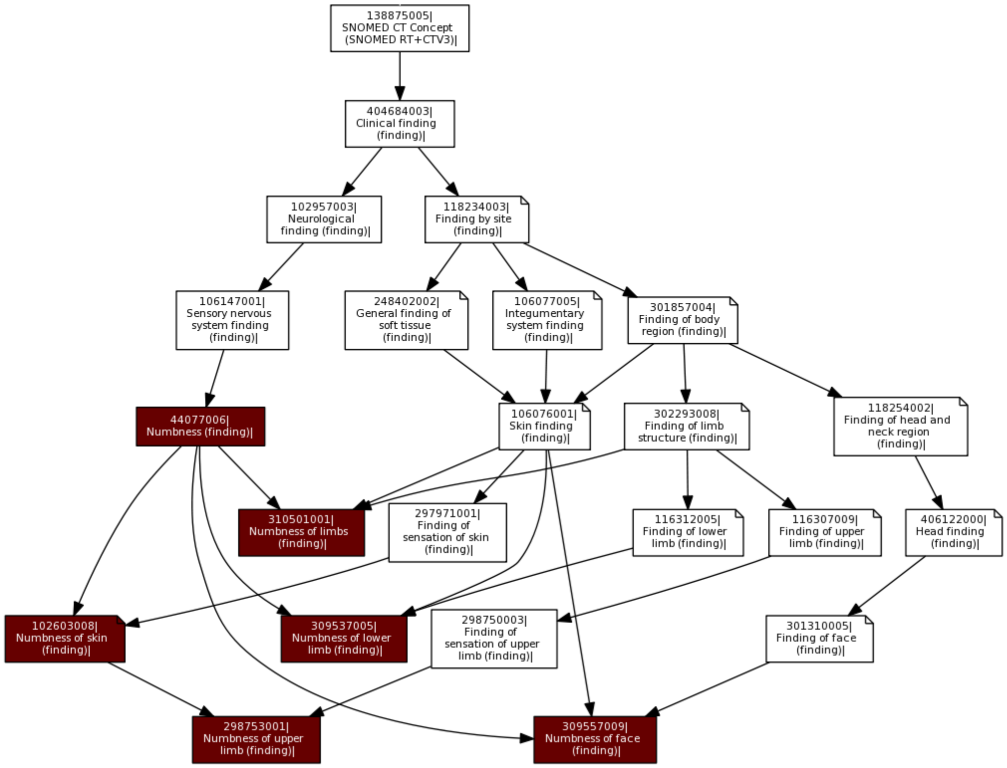


Figure 7-4. Location of "44077006|Numbness (finding)|" and selected subtypes in the hierarchy.

A second set of concepts that have been identified have to do with sanctioned qualifier values. While a **362981000|Qualifier value (qualifier value)|** concept may be sanctioned for use in the Concept Model, it does not necessarily mean it should be used but caution should be exercised when using them. One such particular range of qualifiers is **404684003|Clinical finding (finding)|:363713009|Has interpretation (attribute)|=260245000|Findings values (qualifier value)|**. The concept **260245000|Findings values (qualifier value)|** has 185 subtype concepts such as **373067005|No (qualifier value)|** and **2667000|Absent (qualifier value)|**. However, when encoding a finding such as "no pain" or "pain absent," it should not be encoded as **22253000|Pain (finding)|:363713009|Has interpretation (attribute)|=373067005|No (qualifier value)|** or **22253000|Pain (finding)|:363713009|Has interpretation (attribute)|=2667000|Absent (qualifier value)|**. The defining attributes of the pre-coordinated concept **81765008|No pain (situation)|** shows that it should be encoded using **408729009|Finding context (attribute)|=410516002|Known absent (qualifier value)|** instead. The significance is that post-coordinating **22253000|Pain (finding)|** with the three different qualifiers (i.e., **373067005|No (qualifier value)|**, **2667000|Absent (qualifier value)|** and **410516002|Known absent (qualifier value)|**) will yield unrelated expressions. While **373067005|No (qualifier value)|** is not used in any defining attributes, **2667000|Absent (qualifier value)|** is used 264 times and is generally used in a grouped relationship with **363714003|Interprets (attribute)|**.

7.3.1.2.4 Exclude Descriptions

SNOMED CT also includes descriptions that are general and may be interpreted incorrectly. For example, **161712005|Menopause, function (observable entity)|** has a description of **251994015|The change|**, which may be incorrectly used to encode a phrase such as "the change of dressing was done." Another example is **248103001|Flashing (disorder)|** with a description of **370393013|Flashing|**, which may be incorrectly used for encoding a phrase such as "sees flashing lights" when **248103001|Flashing (disorder)|** actually is a subtype of **58349009|Exhibitionism (disorder)|**. There is no streamlined process for compiling a list of descriptions that should be excluded when encoding but concepts used in encoding a dataset should be manually reviewed to detect any forms of irregularities.

7.3.1.3 Sorting Search Results

While exact matching is considered the gold standard when locating SNOMED CT concepts, there are cases whereby there may be more than one exact match. For example, the word "patch" is a description used in six different concepts (refer to **Table 7-3**) across four hierarchies. This presents a challenge when trying to locate the appropriate concept if only the description is displayed without the fully specified name or defining attributes. Therefore what is needed is a method to sort the search results to ensure the relevant results appear first.

Table 7-3. Concepts that have a description of "patch."

No	Concept	Fully Specified Name	Preferred Term	Synonym
1.	413676000	2530860017|Body tissue patch material (substance)|	2534326013|Body tissue patch material|	2537315013|Patch|
2.	385114002	2660372013|Transdermal patch (qualifier value)|	1479376011|Transdermal patch|	2621574015|Patch|
3.	262302002	1472138010|Human patch material (substance)|	1489879018|Human patch material|	390233017|Patch|
4.	1522000	741981018|Plaque (morphologic abnormality)|	3647018|Plaque|	3650015|Patch|
5.	386028003	1460207010|Surgical patch (physical object)|	1480215018|Surgical patch|	1491437015|Patch|
6.	419702001	2613873010|Patch - unit of product usage (qualifier value)|	2622094015|Patch - unit of product usage|	2577161011|Patch|

There are 13,867 descriptions with literally the same string with the majority occurring within the **123037004|Body structure (body structure)|** hierarchy (n=13,203), between the **105590001|Substance (substance)|** and **373873005|Pharmaceutical / biologic product (product)|** hierarchies (n=3,932), within the **404684003|Clinical finding (finding)|** hierarchy (n=1,506) and between the **123037004|Body structure (body structure)|** and **404684003|Clinical finding (finding)|** hierarchies (n=999). Refer to **Table 12-35** in **Appendix E** for the full frequency counts.

A five step process has been proposed to deal with multiple exact matches (refer to **Figure 7-5**). The first four steps may be done automatically while the last step requires a manual review. The following sub-sections describe each of the steps in more detail.

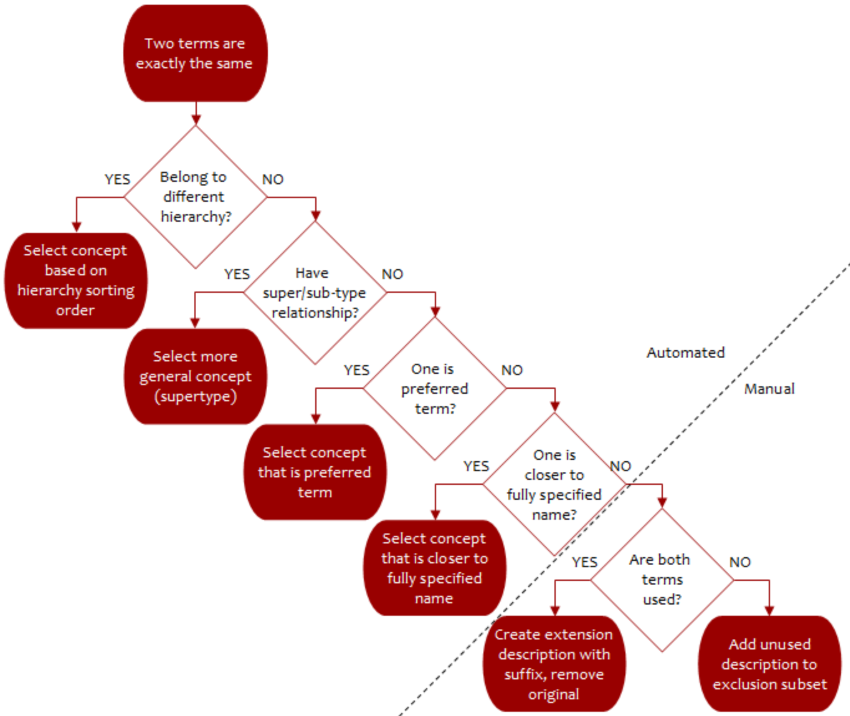


Figure 7-5. How to select the appropriate concept with multiple exact matches.

7.3.1.3.1 Concept from Different Hierarchies

For concepts that have the same description from different hierarchies, selecting the appropriate concept is a straightforward process based on the hierarchy and the context that it is being used. For example, "abrasion" when used in the context of a problem list or encounter diagnosis, would more likely refer to **1767883012|Abrasion (disorder)|** rather than **825994018|Abrasion (procedure)|**. Refer to **Table 7-4** and **Figure 7-6** for more details. In most cases, if two concepts from the **404684003|Clinical finding (finding)|** and **49755003|Morphologically abnormal structure (morphologic abnormality)|** supertype concepts have the same description, it usually is such that the **404684003|Clinical finding (finding)|** is defined by **116676008|Associated morphology (attribute)|=49755003|Morphologically abnormal structure (morphologic abnormality)|**. For example, **399963005|Abrasion (disorder)|** is defined with **116676008|Associated morphology (attribute)|= 400061001|Abrasion (morphologic abnormality)|**. The same goes for concepts from the **373873005|Pharmaceutical / biologic product (product)|** and **105590001|Substance (substance)|** hierarchies where **373873005|Pharmaceutical / biologic product (product)|** is usually defined by **127489000|Has active ingredient (attribute)|=105590001|Substance (substance)|**.

Table 7-4. Concepts that have a description of "abrasion."

No	Concept	Fully Specified Name	Preferred Term	Synonym
1.	400061001	1767981013|Abrasion (morphologic abnormality)|	1779330015|Abrasion|	
2.	399963005	1767883012|Abrasion (disorder)|	1779244012|Abrasion|	
3.	8420001	825994018|Abrasion (procedure)|	1215797018|Abrasion procedure|	195131013|Abrasion|

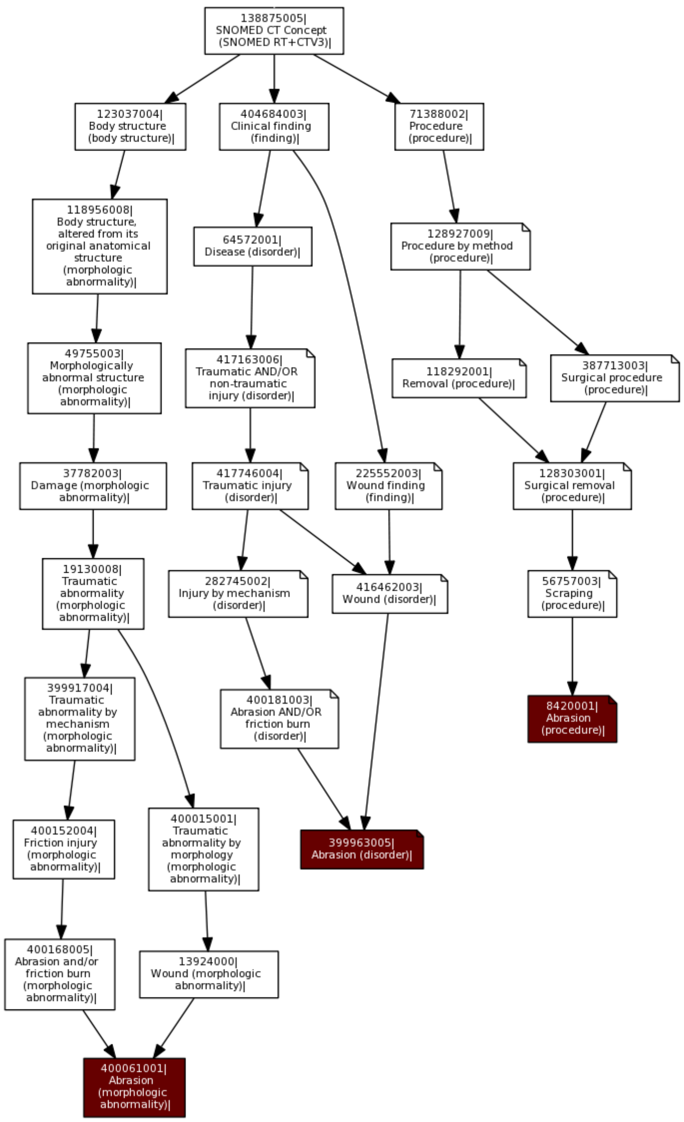


Figure 7-6. Location of concepts "400061001|Abrasion (morphologic abnormality)|", "399963005|Abrasion (disorder)|" and "8420001|Abrasion (procedure)|" in the hierarchy.

The order of hierarchical sorting will depend on the context of dataset being encoded. An example of the order of a problem list sorting order is shown in **Table 7-5**. This sorting order was derived as a result of reviewing concepts from hierarchies that have the same descriptions.

Table 7-5. Hierarchy sorting order for problem list.

No	Hierarchy
1.	404684003|Clinical finding (finding)|
2.	71388002|Procedure (procedure)|
3.	243796009|Situation with explicit context (situation)|
4.	272379006|Event (event)|
5.	363787002|Observable entity (observable entity)|
6.	308916002|Environment or geographical location (environment / location)|
7.	123037004|Body structure (body structure)|
8.	48176007|Social context (social concept)|
9.	410607006|Organism (organism)|
10.	373873005|Pharmaceutical / biologic product (product)|
11.	105590001|Substance (substance)|
12.	362981000|Qualifier value (qualifier value)|
13.	78621006|Physical force (physical force)|
14.	260787004|Physical object (physical object)|
15.	123038009|Specimen (specimen)|
16.	254291000|Staging and scales (staging scale)|
17.	106237007|Linkage concept (linkage concept)|
18.	370115009|Special concept (special concept)|

7.3.1.3.2 Concepts with an "Is a" Relationship

The challenge arises when there are descriptions that are the same from different concepts within the same hierarchy. In this case, the "is a" hierarchy may be used to locate the more general concept. Refer to **Table 12-37** in **Appendix E** for the frequency count of concepts with the same descriptions by hierarchy. For concepts within the same hierarchy that have the same descriptions but also have an "is a" relationship, locating the most general concept can be done by querying the transitive closure table. For example, the term "arm" is exactly matched to four concepts (refer to **Table 7-6**). Querying the transitive closure table indicates that **53120007|Upper limb structure (body structure)|** is the most general concept (refer to **Figure 7-7**). The rationale for selecting the more general (supertype) concept is that a subtype concept may be too specific and make inferences that may not be accurate. Returning to the example of "arm," it can refer to "upper limb" or "upper arm." Since we cannot be certain that "arm" refers to the "upper arm" and thus excludes other parts of the arm such as the "hand" or "forearm," the "upper limb" is a safer choice although it may be less specific. In most cases involving concepts that have an "is a" relationship from the **123037004|Body structure (body structure)|** hierarchy, it usually involves the subtype concept including the term "entire." A review of the defining attributes shows that the "entire" concepts are rarely used. For example, **102556003|Pain in upper limb (finding)|** is defined with **53120007|Upper limb structure (body structure)|** and not **182245002|Entire upper limb (body structure)|**.

Table 7-6. Concepts that have a description of "arm."

No	Fully Specified Name	Preferred Term	Synonym
1.	302538001|Entire upper arm (body structure)|	444276014|Entire upper arm|	2551347016|Arm|
2.	182245002|Entire upper limb (body structure)|	281752018|Entire upper limb|	2550843019|Arm|
3.	53120007|Upper limb structure (body structure)|	496560017|Upper limb structure|	2552021018|Arm|
4.	40983000|Upper arm structure (body structure)|	492856010|Upper arm structure|	68363017|Arm|

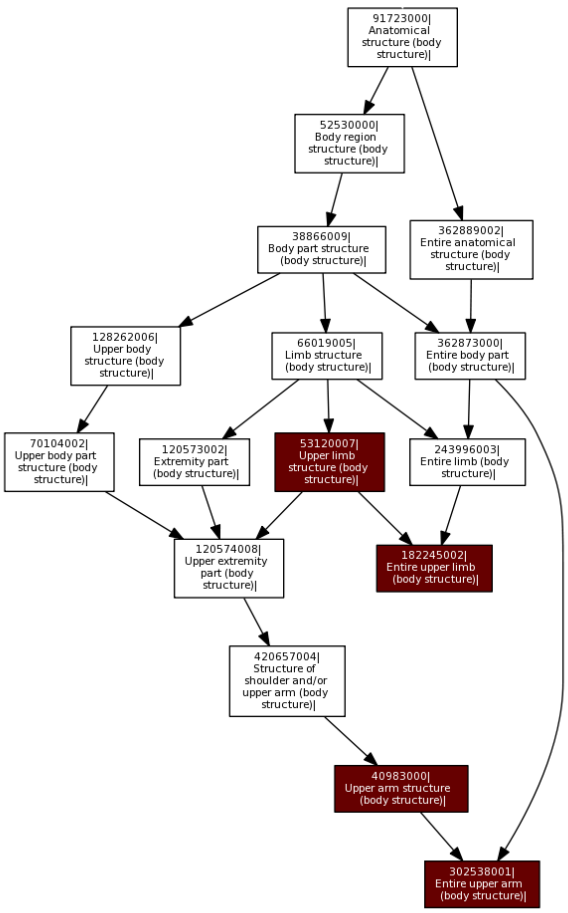


Figure 7-7. Location of concepts "302538001|Entire upper arm (body structure)|", "182245002|Entire upper limb (body structure)|", "53120007|Upper limb structure (body structure)|" and "40983000|Upper arm structure (body structure)|" in the hierarchy.

7.3.1.3.3 A Description is the Preferred Term

The next step would be to check if one of the terms that are the same is a preferred term while another is a synonym. For example, "sunstroke" is a description for both **18615009|Sunstroke (disorder)|** and **52072009|Heat stroke (disorder)|** but **31390012|Sunstroke|** is a preferred term while **496199018|Sunstroke|** is a synonym (refer to

Table 7-7 and **Figure 7-8**). In this case, more weight is given to a description that is a preferred term over a description that is over a synonym.

Table 7-7. Concepts that have a description of "sunstroke."

No	Concept	Fully Specified Name	Preferred Term	Synonym
1.	18615009	Sunstroke (disorder)	31390012|Sunstroke|	▪ 31393014|Siriasis| ▪ 31392016|Ictus solaris|
2.	52072009	Heat stroke (disorder)	86678016|Heat stroke|	▪ 496199018|Sunstroke| ▪ 86681014|Heat apoplexy| ▪ 496197016|Heat exhaustion| ▪ 496198014|Heat hyperpyrexia| ▪ 496199018|Sunstroke| ▪ 86680010|Thermoplegia|

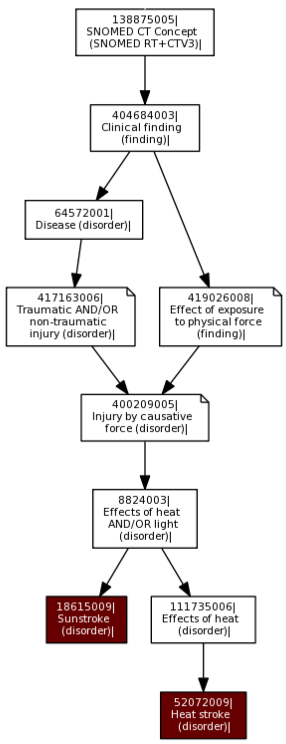


Figure 7-8. Location of concepts "18615009|Sunstroke (disorder)" and "52072009|Heat stroke (disorder)|" in the hierarchy.

7.3.1.3.4 A Description is More Similar to a Fully Specified Name

The following step would be to conduct a comparison between the description and fully specified name to determine if one synonym "leans" more towards one concept. For example, the phrase "abdomen feels bloated" is matched to two different concepts (refer to **Table 7-8**).

Table 7-8. Concepts that have a description of "abeomdn feels boated."

No	Concept	Fully Specified Name	Preferred Term	Synonym
1.	116289008	676919019|Abdominal bloating (finding)|	179244013|Abdominal bloating|	▪ 2475423012|Abdomen feels bloated|

No	Concept	Fully Specified Name	Preferred Term	Synonym
2.	248490000	638799011|Bloating symptom (finding)|	370924017|Bloating symptom|	▪ 2475942016|Abdomen feels bloated|

In this case, "abdomen feels bloated" may be better suited to **676919019|Abdominal bloating (finding)|** as it explicitly refers to "abdominal" while **638799011|Bloating symptom (finding)|** only implicitly refers to "abdominal." It should be noted that both concepts have a defining attribute of **363698007|Finding site (attribute)|=113345001|Abdominal structure (body structure)|** (refer to **Figure 7-9**). It should also be noted that these two concepts are sibling concepts (refer to **Figure 7-10**).

116289008|Abdominal bloating (finding)|:
{363698007|Finding site (attribute)|=
113345001|Abdominal structure (body structure)|}

248490000|Bloating symptom (finding)|:
{363698007|Finding site (attribute)|=
113345001|Abdominal structure (body structure)|,
419066007|Finding informer (attribute)|=
419358007|Subject of record or other provider of history (person)|}

Figure 7-9. Defining attributes of "676919019|Abdominal bloating (finding)|" and "638799011|Bloating symptom (finding)|".

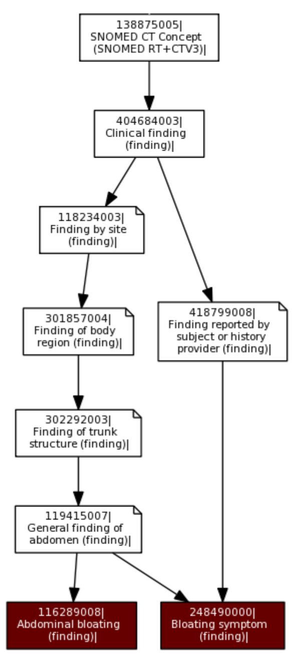


Figure 7-10. Location of concepts "676919019|Abdominal bloating (finding)|" and "638799011|Bloating symptom (finding)|" in the hierarchy.

7.3.1.3.5 Manual Review

At this stage, a manual review is probably necessary to distinguish between the two concepts. If the concepts are within the **404684003|Clinical finding (finding)|** hierarchy, one option may be to select a disorder (**64572001|Disease (disorder)|**) over a non-disorder or vice versa (refer to **Table 7-9**).

Table 7-9. Examples of concepts that have the same fully specified name sans the suffix from the 404684003|Clinical finding (finding)| hierarchy.

No	Term	Disorder	Finding
1.	Anterior capsule opacification	410568009|Anterior capsule opacification (disorder)|	370948005|Anterior capsule opacification (finding)|
2.	Azoospermia	425558002|Azoospermia (disorder)|	48188009|Azoospermia (finding)|
3.	Food intolerance	235719002|Food intolerance (disorder)|	75051000|Food intolerance (finding)|
4.	May-Hegglin anomaly	234484005|May-Hegglin anomaly (disorder)|	250279001|May-Hegglin anomaly (finding)|
5.	Neutrophilia	414850009|Neutrophilia (disorder)|	165518003|Neutrophilia (finding)|
6.	Nutritional deficiency	70241007|Nutritional deficiency (disorder)|	47563007|Nutritional deficiency (finding)|

In some situations, it may not be possible to determine which concept a description may apply to just from the fully specified name or preferred term. For example, "fracture of wrist" may apply to **263208005|Fracture of distal end of radius and ulna (disorder)|** or **82065001|Fracture of carpal bone (disorder)|** and there is nothing within the fully specified name or preferred term that gives an indication of which may be more appropriate (refer to **Table 7-10**).

Table 7-10. Descriptions of "263208005|Fracture of distal end of radius and ulna (disorder)|" and "82065001|Fracture of carpal bone (disorder)|".

No	Concept	Preferred Term	Synonym
1.	263208005|Fracture of distal end of radius and ulna (disorder)|	391364013|Fracture of distal end of radius and ulna|	▪ 391365014|Fracture of wrist| ▪ 391366010|Fracture of lower end of both ulna and radius|
2.	82065001|Fracture of carpal bone (disorder)|	136121013|Fracture of carpal bone|	▪ 1234675010|Fracture of carpus| ▪ 136124017|Fracture of wrist| ▪ 1234676011|Hand fracture - carpal bone|

Plotting the concepts (refer to **Figure 7-11**) shows that **82065001|Fracture of carpal bone (disorder)|** is a subtype of **208388003|Fracture at wrist and/or hand level (disorder)|** and **125598003|Injury of wrist (disorder)|** (not pictured) so it is possible that "fracture of wrist" is better suited to **82065001|Fracture of carpal bone (disorder)|** than **263208005|Fracture of distal end of radius and ulna (disorder)|**. If "fracture of wrist" is still a term that is required, creating extension such as "fracture of wrist (carpal bone)" and "fracture of wrist (ulna and radius)" may be one way to help distinguish between the two. Also, the descriptions **391365014|Fracture of wrist|** and **136124017|Fracture of wrist|** should be added to a list of exclusion descriptions.

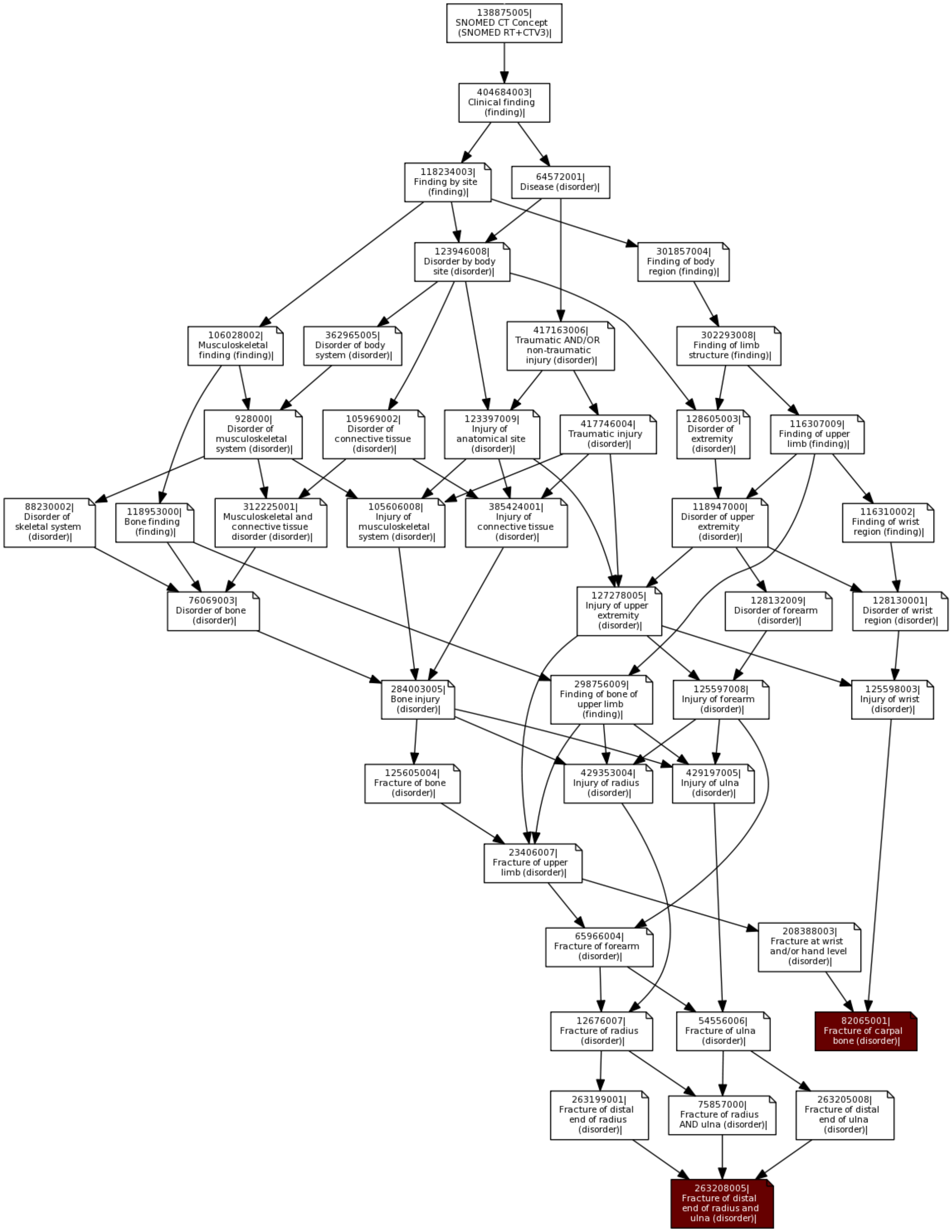


Figure 7-11. Location of concepts “263208005|Fracture of distal end of radius and ulna (disorder)|” and “82065001|Fracture of carpal bone (disorder)|” in the hierarchy. Both concepts have synonyms of “fracture of wrist” (which is not shown in the figure).

7.3.2 Modeling Scenarios

This section describes how to combine two or more concepts together to form a post-coordinated expression and builds on the work by Wang, et al.,[13] Sampalli, et al.,[5] and Richesson, et al.,[3] as well as the SNOMED CT User Guide (UG).[15] In this context, the candidate refers to the concept or expression that is being merged into another concept while predicate refers to the concept or expression that is being merged with. While it may be possible to merge two concepts together, it does not necessarily mean that they should be merged. For example, a clinical statement such as "stroke/dementia" should be encoded separately as **230690007|Cerebrovascular accident (disorder)|** and **52448006|Dementia (disorder)|** rather than a complex post-coordinated expression such as **230690007|Cerebrovascular accident (disorder)|:47429007|Associated with (attribute)|=52448006|Dementia (disorder)|**. The main reason is that linking **404684003|Clinical finding (finding)|** and **404684003|Clinical finding (finding)|** without an explicit relationship can create erroneous inferences as well as complex post-coordinated expressions that may not be retrieved correctly. For example, querying for **52448006|Dementia (disorder)|** and its subtypes will not yield **230690007|Cerebrovascular accident (disorder)|:47429007|Associated with (attribute)|=52448006|Dementia (disorder)|** as the focus concept **230690007|Cerebrovascular accident (disorder)|** is not a subtype of **52448006|Dementia (disorder)|**.

The modeling of concepts between the **404684003|Clinical finding (finding)|** and **71388002|Procedure (procedure)|** hierarchies were also disabled. For example, a clinical statement such as "Bartholin Gland Cyst Drainage" can be encoded with two concepts: **57044006|Cyst of Bartholin's gland duct (disorder)|** and **22462000|Drainage procedure (procedure)|**. There are four different Concept Model attributes that can be used to link **404684003|Clinical finding (finding)|** and **71388002|Procedure (procedure)|** (i.e., **255234002|After (attribute)|**, **47429007|Associated with (attribute)|**, **42752001|Due to (attribute)|** and **363705008|Has definitional manifestation (attribute)|**). However, it may be incorrect to infer, for example, that **57044006|Cyst of Bartholin's gland duct (disorder)|** is **42752001|Due to (attribute)|** the **22462000|Drainage procedure (procedure)|**. It is plausible that the two concepts should be linked via **363702006|Has focus (attribute)|**, but due to the number of potential Concept Model attributes and the possibility of making inappropriate inferences, rules were created to disable the modeling of **404684003|Clinical finding (finding)|** and **71388002|Procedure (procedure)|** if there are no explicit Concept Model attributes to link the two concepts together (refer to **Table 7-11**).

Table 7-11. Concept Model attributes that link "404684003|Clinical finding (finding)|" to "71388002|Procedure (procedure)|".

Concept Model	404684003|Clinical finding (finding)|	71388002|Procedure (procedure)|
404684003|Clinical finding (finding)|	▪ 255234002|After (attribute)| ▪ 47429007|Associated with (attribute)| ▪ 42752001|Due to (attribute)| ▪ 363705008|Has definitional manifestation (attribute)|	▪ 255234002|After (attribute)| ▪ 47429007|Associated with (attribute)| ▪ 418775008|Finding method (attribute)|
71388002|Procedure (procedure)|	▪ 363702006|Has focus (attribute)|	▪ 363702006|Has focus (attribute)|

The modeling method consists of nine different scenarios of how the predicate and candidate may be linked and how to identify an appropriate Concept Model attribute. A summary of the modeling method and examples are shown in **Figure 7-12** followed by a brief description of each scenario. Additional details for each scenario are also available in **Section 12.6.7** of **Appendix E**.

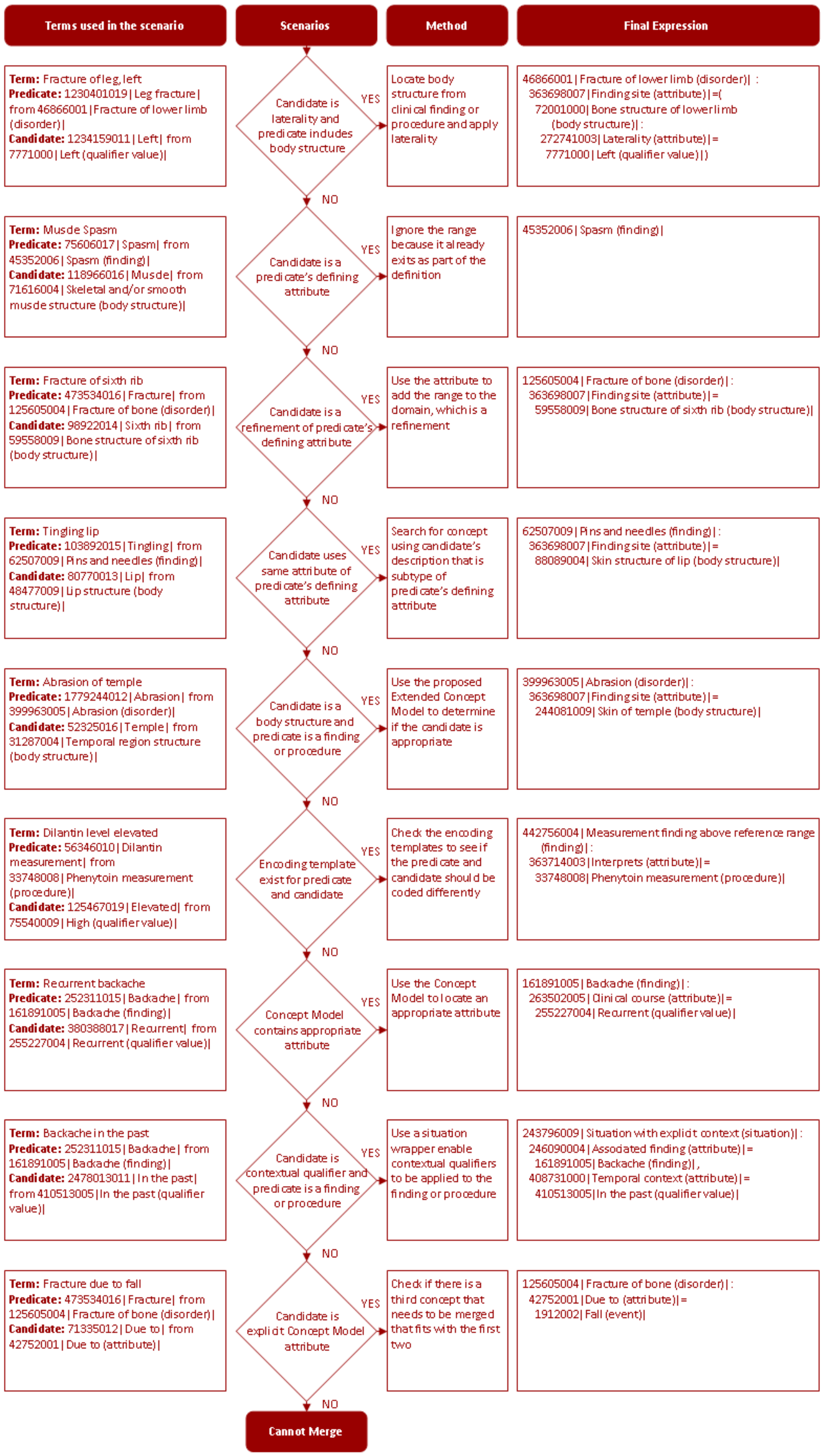


Figure 7-12. Summary of the modeling method.

7.3.2.1 Candidate Is Laterality and Predicate Includes Body Structure

This is the first scenario because a review of the encoding results revealed that laterality was the most commonly used form of post-coordination. The predicate in this scenario must either be a **123037004|Body structure (body structure)|**, **404684003|Clinical finding (finding)|**, **71388002|Procedure (procedure)|** or **243796009|Situation with explicit context (situation)|** that includes **123037004|Body structure (body structure)|** as part of the defining attributes (implicit) or already post-coordinated expression (explicit). Additional details are available in **Section 12.6.7.1** of **Appendix E**.

7.3.2.2 Candidate is a Predicate's Defining Attribute

In this scenario, a query is conducted to determine if the candidate is part of the predicate's defining attributes. If the candidate is a defining attribute of the predicate, it is not necessary to combine the concepts and the candidate concept can be discarded as it is already represented in the predicate. Additional details are available in **Section 12.6.7.2** of **Appendix E**.

7.3.2.3 Candidate is a Refinement of Predicate's Defining Attribute

In this scenario, a query is conducted to determine if the candidate is a refinement of one of the predicate's defining attributes. If the candidate is a refinement of a defining attribute of the predicate, then we are creating a refinement post-coordinated expression. If it is not a subtype, additional checks need to be conducted in the scenarios to follow to ensure the candidate is the appropriate concept. Additional details are available in **Section 12.6.7.3** of **Appendix E**.

7.3.2.4 Candidate Uses Same Attribute of Predicate's Defining Attribute

In this scenario, a query is conducted to determine if the candidate should be combined to the predicate with an attribute that already exists but unlike the previous scenario, the candidate is not equivalent to and is not a subtype of the range of the attribute. If the attribute does not exist, it can be added. If the attribute does exist but the candidate is not a subtype of the range, additional checks need to be conducted in the scenarios to follow to ensure the candidate is the appropriate concept. Additional details are available in **Section 12.6.7.4** of **Appendix E**.

7.3.2.5 Candidate is a Body Structure and Predicate is a Finding or Procedure

While qualification and refinement have very strict and specific rules as to which concepts can be used, the use of sanctioned Concept Model attributes opens the possibility of creating verified but nonsensical post-coordinated expressions. In this case "verified" means the expression is technically correct as it conforms to the Concept Model. For example, consider the clinical statements "ankle abrasion," "ankle fracture" and "ankle sprain." Although all three refer to "ankle," different body structure concepts are used (refer to **Figure 7-11**). Representing "abrasion of ankle" using **67211007|Structure of ligament of ankle joint (body structure)|** would conform to the Concept Model but would be semantically incorrect.

Table 7-12. Example of type of "ankle" concepts.

No	Concept	Associated morphology (attribute)	Finding site (attribute)
1.	44465007|Sprain of ankle (disorder)|	384709000|Sprain (morphologic abnormality)|	67211007|Structure of ligament of ankle joint (body structure)|
2.	211334007|Abrasion, ankle (disorder)|	400061001|Abrasion (morphologic abnormality)|	67269001|Skin structure of ankle (body structure)|
3.	16114001|Fracture of ankle (disorder)|	72704001|Fracture (morphologic abnormality)|	33696004|Bone structure of ankle (body structure)|

While it may not be possible to completely remove the possibility of creating nonsensical statements, it is possible to improve the quality of post-coordinated expressions by ensuring the appropriate body structure concept is used. An analysis of the defining attributes of concepts reveals that there are four attributes that are linked to **123037004|Body structure (body structure)|** and another four that are linked to **49755003|Morphologically abnormal structure (morphologic abnormality)|** (refer to **Table 7-13**). The number of target concepts is shown in **Table 7-14**.

Table 7-13. Sites and morphology in SNOMED CT.

	404684003|Clinical finding (finding)|	71388002|Procedure (procedure)|
Body Structures	363698007|Finding site (attribute)|	▪ 363704007|Procedure site (attribute)| ▪ 405813007|Procedure site - Direct (attribute)| ▪ 405814001|Procedure site - Indirect (attribute)|
Morphology Abnormalities	116676008|Associated morphology (attribute)|	▪ 405816004|Procedure morphology (attribute)| ▪ 363700003|Direct morphology (attribute)| ▪ 363709002|Indirect morphology (attribute)|

Table 7-14. Number of concepts, unique and total body structures and morphologic abnormalities used in defining attributes.

Concept Model Attributes	Concepts	Unique Targets	Total Targets
Body Structures			
363698007|Finding site (attribute)|	59,009	3,734	70,576
363704007|Procedure site (attribute)|	3,349	585	3,448
405813007|Procedure site - Direct (attribute)|	24,801	2,565	30,932
405814001|Procedure site - Indirect (attribute)|	7,249	1,032	7,897
All Procedure Sites	32,614	2,791	42,277
All Sites	**91,623**	**4,772**	**112,853**
Morphologic Abnormalities			
116676008|Associated morphology (attribute)|	41,593	1,765	50,950
405816004|Procedure morphology (attribute)|	174	35	179
363700003|Direct morphology (attribute)|	6,597	308	7,589
363709002|Indirect morphology (attribute)|	514	66	534
All Procedure Morphology	7,122	317	8,302
All Morphologies	**48,715**	**1,834**	**59,252**

An analysis of how **49755003|Morphologically abnormal structure (morphologic abnormality)|** and **123037004|Body structure (body structure)|** concepts are used in defining attributes revealed there are 1,586 **116676008|Associated morphology (attribute)|<<49755003|Morphologically abnormal structure (morphologic abnormality)|** attribute-ranges that are used in conjunction with 3,365 unique **363698007|Finding site (attribute)|<<123037004|Body structure (body structure)|**. The matching was done by relationship group as opposed to all the attributes in a concept. With these 1,586 combinations of **49755003|Morphologically abnormal structure (morphologic abnormality)|** and **123037004|Body structure (body structure)|** concepts, we now have a proposed Extended Concept Model that can assist in locating the appropriate **123037004|Body structure (body structure)|** concept for use with an expression with **49755003|Morphologically abnormal structure (morphologic abnormality)|**, which constrains the permissible values to a smaller subset than the broad guidelines of the Concept Model. Additional details are available in **Section 12.6.7.5** of **Appendix E** on how to practically utilise the Extended Concept Model.

7.3.2.6 Encoding Template Exists for Predicate and Candidate

While search algorithms, word equivalency and addition of acronyms and synonyms can help to improve the likelihood of finding a suitable match, there are still deficiencies as these methods do not take into account the intricacies of how SNOMED CT concepts are defined. A proposed solution is to create encoding templates that help to construct post-coordinated expressions. These templates are based on "trigger" words and how these words can be combined with other concepts. Sample templates are available in **Section 12.6.7.6** in **Appendix E**. They are meant to demonstrate some of the different types of encoding templates that should be created and should not be considered definitive as templates may need to be customised based on the dataset being encoded.

7.3.2.7 Concept Model Contains Appropriate Attribute

If two concepts are to be merged and no Concept Model attributes have been explicitly specified to link the two concepts together, the Concept Model can be queried to determine the most appropriate attribute (refer to **Figure 7-13**).

```
SELECT CM.AttributeId FROM SCT_TransitiveClosure TC1, SCT_TransitiveClosure TC2,
SCT_ConceptModel CM WHERE TC1.SuperTypeId=$DomainId AND TC1.SubTypeId=CM.DomainId AND
TC2.SuperTypeId=$RangeId AND TC2.SubTypeId=CM.RangeId GROUP BY CM.AttributeId
```

Figure 7-13. SQL statement for checking which Concept Model attribute links two concepts together where $DomainId refers to the first concept and $RangeId refers to the second concept.

The Concept Model is comprised of 125 combinations of domain-attribute-ranges (e.g., **404684003|Clinical finding (finding)|:116676008|Associated morphology (attribute)|<<49755003|Morphologically abnormal structure (morphologic abnormality)|**) that are defined in the SNOMED CT UG. While 68 domains and ranges can only be linked by a single Concept Model attribute, there are 24 domains and ranges in which multiple Concept Model attributes may be used. Refer to **Table 7-15** for some examples. The full table is available in **Table 12-43** in **Appendix E**. For example, if two concepts that belong to the **404684003|Clinical finding (finding)|** and

410607006|Organism (organism)| hierarchy need to be combined, there are two Concept Model attributes that may be used: **47429007|Associated with (attribute)|** and **246075003|Causative agent (attribute)|**. A check of the relationships in SNOMED CT shows that **47429007|Associated with (attribute)|** is used only once while **246075003|Causative agent (attribute)|** is used 5,942 times. Therefore in this case, **246075003|Causative agent (attribute)|** would be the better choice. If both concepts are from the **404684003|Clinical finding (finding)|** hierarchy, there are four possibilities. Even though **363705008|Has definitional manifestation (attribute)|** is used in 60% of the cases, we cannot make the assumption that that is the appropriate Concept Model attribute and therefore should not combine the two concepts if no explicit Concept Model attribute is used.

Table 7-15. Sample domains and ranges that have multiple attributes that link the two together.

No	Domain	Attributes	Range	Frequency	Percent	Default
1.	404684003|Clinical finding (finding)|	363705008|Has definitional manifestation (attribute)|	404684003|Clinical finding (finding)|	5,287	59.9%	N/A
		42752001|Due to (attribute)|		1,661	18.8%	
		47429007|Associated with (attribute)|		1,141	12.9%	
		255234002|After (attribute)|		731	8.3%	
2.	404684003|Clinical finding (finding)|	42752001|Due to (attribute)|	272379006|Event (event)|	120	92.3%	Y
		47429007|Associated with (attribute)|		10	7.7%	
3.	404684003|Clinical finding (finding)|	246075003|Causative agent (attribute)|	410607006|Organism (organism)|	5,942	99.98%	Y
		47429007|Associated with (attribute)|		1	0.02%	
4.	404684003|Clinical finding (finding)|	246075003|Causative agent (attribute)|	78621006|Physical force (physical force)|	629	100.0%	Y
		47429007|Associated with (attribute)|		0	0.0%	
5.	404684003|Clinical finding (finding)|	246075003|Causative agent (attribute)|	260787004|Physical object (physical object)|	202	93.5%	Y
		47429007|Associated with (attribute)|		14	6.5%	
6.	404684003|Clinical finding (finding)|	363714003|Interprets (attribute)|	71388002|Procedure (procedure)|	5,578	50.4%	N/A
		418775008|Finding method (attribute)|		3,843	34.7%	
		255234002|After (attribute)|		942	8.5%	
		47429007|Associated with (attribute)|		703	6.4%	
7.	404684003|Clinical finding (finding)|	246075003|Causative agent (attribute)|	105590001|Substance (substance)|	10,118	99.3%	Y
		47429007|Associated with (attribute)|		74	0.7%	
		363589002|Associated procedure (attribute)|		0	0.0%	

7.3.2.8 Candidate is Contextual Qualifier and Predicate is a Finding or Procedure

There are four types of contextual qualifiers: **408729009|Finding context (attribute)|**, **408730004|Procedure context (attribute)|**, **408732007|Subject relationship context (attribute)|** and **408731000|Temporal context (attribute)|**. If a predicate contains a range of a contextual qualifier and the predicate is a **404684003|Clinical finding (finding)|** or **71388002|Procedure (procedure)|**, a situation wrapper in the form of **243796009|Situation with explicit context (situation)|** needs to be added in order for the predicate and candidate to be merged into a single expression.

404684003|Clinical finding (finding)| and **71388002|Procedure (procedure)|** concepts will be linked to the **243796009|Situation with explicit context (situation)|** using **246090004|Associated finding (attribute)|** and **363589002|Associated procedure (attribute)|** respectively while an appropriate attribute for the contextual qualifier can be located using the scenario in **Section 7.3.2.7**. Additional details are available in **Section 12.6.7.7** of **Appendix E**.

7.3.2.9 Candidate is Explicit Concept Model Attribute

An explicit Concept Model attribute may be used in a phrase such as "fracture after fall," in which the individual words are encoded as **125605004|Fracture of bone (disorder)|**, **255234002|After (attribute)|** and **1912002|Fall (event)|**. In this case, it is permissible to combine the three concepts together as a post-coordinated expression. However, it should be noted that just because a phrase uses exactly the same description as a Concept Model attribute does not mean that is the correct one. This is evident by analysing the similarities and differences between SNOMED CT's textual descriptions and concept definitions. A textual description refers to SNOMED CT descriptions while concept definitions refer to SNOMED CT defining attributes. There are basically three possibilities of textual descriptions and concept definitions (refer to **Figure 7-14**).

Combination of Concept Definition and Textual Description		Concept Definition	
		Yes	No
Textual Description	Yes	Ideal usage as textual description reflects the concept definition	It is possible that the concept has not been adequately defined to reflect the textual description of a concept
	No	Textual description is implicit	N/A

Figure 7-14. SNOMED CT concept definitions and textual descriptions.

Examples of "associated with," which demonstrate the three possibilities are shown in **Table 7-16**. In the first example, **421154002|Fever associated with AIDS (disorder)|** contains the phrase "associated with" and is defined with **47429007|Associated with (attribute)|=62479008|Acquired immune deficiency syndrome (AIDS) (disorder)|**. In the second example, **247855001|Fear associated with illness and body function (finding)|** also contains the phrase "associated with" but does not have a defining attribute of **47429007|Associated with (attribute)|**. In the third example, **235076008|Disorder of dental prosthesis (disorder)|** does not contain the phrase "associated with" but does have a defining attribute **47429007|Associated with (attribute)|=71388002|Procedure (procedure)|**.

Table 7-16. Examples of "associated with" of textual description and concept definition.

No	Textual Description / Concept Definition	Concept	Defining Attributes
1.	Yes – Textual Description Yes – Concept Definition	421154002|Fever **associated with** AIDS (disorder)|	421154002|Fever associated with AIDS (disorder)|: 363714003|Interprets (attribute)|= 386725007|Body temperature (observable entity)|, 47429007|**Associated with** (attribute)|= 62479008|Acquired immune deficiency syndrome (AIDS) (disorder)|

No	Textual Description / Concept Definition	Concept	Defining Attributes
2.	Yes – Textual Description No – Concept Definition	247855001|Fear **associated with** illness and body function (finding)|	247855001|Fear associated with illness and body function (finding)|: 363714003|Interprets (attribute)|= 285854004|Emotion (observable entity)|
3.	No – Textual Description Yes – Concept Definition	235076008|Disorder of dental prosthesis (disorder)|	235076008|Disorder of dental prosthesis (disorder)|: 47429007|**Associated with** (attribute)|= 71388002|Procedure (procedure)|

The Concept Model attributes **1220306019|Morphology|**, **367805018|Component|** and **388315014|Method|** appear frequently in SNOMED CT descriptions but under closer inspection, these terms can be taken out of context and are not useful in linking clinical statements. Concept Model attributes such as **1490773016|Intent|**, **388521012|Priority|** and **2643091013|Using device|** appear only once in SNOMED CT descriptions and may be omitted. In the end, four Concept Model attribute descriptions proved to be feasible: **380400010|After|**, **79074014|Associated with|**, **71336013|Caused by|** and **71335012|Due to|**. Incidentally, **47429007|Associated with (attribute)|** is a supertype of **255234002|After (attribute)|** and **42752001|Due to (attribute)|**. **246075003|Causative agent (attribute)|** is also a subtype of **79074014|Associated with|** but the phrase "causative agent" is not used in any of the textual descriptions except for itself and one other concept, **362944004|Causative_agents (qualifier value)|**. It should also be noted that **71335012|Due to|** is used with both **404684003|Clinical finding (finding)|** and **272379006|Event (event)|** although the description is used with far greater frequency in the former than to the latter.

A close investigation of the three Concept Model attributes that are feasible shows that the textual description and concept definition do not always match. For example, even though the phrase "causative agent" does not appear in SNOMED CT descriptions except in the two cases described previously, phrases such as "after," "associated with" and "due to" are used in conjunction with the concept definition of **246075003|Causative agent (attribute)|** (refer to **Table 7-17**). Therefore we cannot simply assume the appropriate Concept Model attribute is correct just because there is a lexical match. A review of the original clinical statement is necessary to determine its intended meaning.

Table 7-17. Examples of concepts with a concept definition of "246075003|Causative agent (attribute)|".

No	Textual Description	Concept	Defining Attributes
1.	After	439859001|Adverse effect of correct biological substance **after** proper administration (disorder)|	439859001|Adverse effect of correct biological substance after proper administration (disorder)|: **246075003|Causative agent (attribute)|**= 115668003|Biological substance (substance)|
2.	Associated with	428885001|Bacteremia **associated with** intravascular line (finding)|	428885001|Bacteremia associated with intravascular line (finding)|: **246075003|Causative agent (attribute)|**= 409822003|Superkingdom Bacteria (organism)|, 363698007|Finding site (attribute)|= 442083009|Anatomical or acquired body structure (body structure)|}
3.	Due to	402153000|Dermatosis **due to** scorpion (disorder)|	402153000|Dermatosis due to scorpion (disorder)|: **246075003|Causative agent (attribute)|**= 76222001|Class Arachnida (organism)|, 363698007|Finding site (attribute)|= 39937001|Skin structure (body structure)|

A review of the use of the textual description "due to" shows that is most often used in conjunction with the attribute **246075003|Causative agent (attribute)|** (refer to Table 7-18). Therefore if a lexical match is done for "due to", it is more likely that **246075003|Causative agent (attribute)|**.

Table 7-18. Comparing the textual descriptions with the attributes 255234002|After (attribute)|, 47429007|Associated with (attribute)|, 246075003|Causative agent (attribute)| and 42752001|Due to (attribute)|.

Textual Description (Columns) / Concept Definition (Rows)	"After"	"Associated with"	"Due to"	"Causative agent"	None	Total
255234002|After (attribute)|	45	3	37		1,239	**1,324**
47429007|Associated with (attribute)|		282	71		1,366	**1,719**
246075003|Causative agent (attribute)|	5	14	1,201		14,710	**15,925**
42752001|Due to (attribute)|	2	24	358		996	**1,380**
255234002|After (attribute)| 47429007|Associated with (attribute)|			3		84	**87**
255234002|After (attribute)| 246075003|Causative agent (attribute)|	2		13		87	**102**
255234002|After (attribute)| 42752001|Due to (attribute)|	1		3		75	**79**
255234002|After (attribute)| 47429007|Associated with (attribute)| 246075003|Causative agent (attribute)|		1			3	**4**
47429007|Associated with (attribute)| 246075003|Causative agent (attribute)|		40	15		77	**132**
47429007|Associated with (attribute)| 42752001|Due to (attribute)|		8	8		36	**52**
246075003|Causative agent (attribute)| 42752001|Due to (attribute)|			103		135	**238**
47429007|Associated with (attribute)| 246075003|Causative agent (attribute)| 42752001|Due to (attribute)|		6			5	**11**

7.3.3 Cleaning Up the Post-coordinated Expression

The third part of the encoding method focuses on cleaning up the post-coordinated expression and consists of inferring context to ensure the appropriate context was used throughout the clinical statement, additional post-coordination to ensure the expression is complete and compacting expressions to the closer-to-user form where applicable.

7.3.3.1 Inferring Context

When encoding free text, as often the case, multiple clinical conditions may exist (e.g., "cancer of lung, colon and liver"). As encoding the entire sentence will not yield an exact match, sentences are broken down into phrases by punctuations and conjunctions and encoded individually. In this case, "cancer of lung, colon and liver" can be broken down into "cancer of lung," "colon" and "liver." When individually encoded, the result would be **363358000|Malignant tumor of lung (disorder)|**, **10200004|Liver structure (body structure)|** and **71854001|Colon**

structure (body structure)|. A general rule would be not to leave any concepts as a **123037004|Body structure (body structure)|**, especially if the context is a problem or encounter diagnosis because on their own **123037004|Body structure (body structure)|** concepts do not indicate a problem. One possibility would be to put the **123037004|Body structure (body structure)|** concept into a **404684003|Clinical finding (finding)|** wrapper (e.g., **404684003|Clinical finding (finding)|:363698007|Finding site (attribute)|=10200004|Liver structure (body structure)|**) but that adds little clinical value. Another possibility would be to try and infer the context from other concepts within the free text that have been encoded as a **404684003|Clinical finding (finding)|** or **71388002|Procedure (procedure)|**. Note that this method works best if a concept is fully defined. There is only one attribute that links a **404684003|Clinical finding (finding)|** to a **123037004|Body structure (body structure)|** (i.e., **363698007|Finding site (attribute)|**) while there are three attributes that link a **71388002|Procedure (procedure)|** to a **123037004|Body structure (body structure)|** (i.e., **363704007|Procedure site (attribute)|**, which is also the supertype of **405813007|Procedure site - Direct (attribute)|** and **405814001|Procedure site - Indirect (attribute)|**).

The first step would be to strip the **123037004|Body structure (body structure)|** concept from the **404684003|Clinical finding (finding)|** or **71388002|Procedure (procedure)|** concept. The defining attributes of **363358000|Malignant tumor of lung (disorder)|** are shown in **Figure 7-15**. In this case, we end up with **363358000|Malignant tumor of lung (disorder)|:116676008|Associated morphology (attribute)|=367651003|Malignant neoplasm of primary, secondary, or uncertain origin (morphologic abnormality)|**.

```
363358000|Malignant tumor of lung (disorder)|:
    {116676008|Associated morphology (attribute)|=
        367651003|Malignant neoplasm of primary, secondary, or uncertain origin (morphologic abnormality)|,
    363698007|Finding site (attribute)|=
        39607008|Lung structure (body structure)|}
```

Figure 7-15. Defining attributes for "363358000|Malignant tumor of lung (disorder)|".

The next step would be to locate a pre-coordinated concept for the expression that was extracted. The expression extracted should be converted to the long normal form (**64572001|Disease (disorder)|:116676008|Associated morphology (attribute)|=367651003|Malignant neoplasm of primary, secondary, or uncertain origin (morphologic abnormality)|**) and tested against the pre-coordinated concepts in SNOMED CT. In this case, **363346000|Malignant neoplastic disease (disorder)|** is an equivalent pre-coordinated concept. It may not always be possible to locate a pre-coordinated concept. In that case, the certainty of the encoding (discussed later) will be reduced. Cases where there are multiple proximal primitive concepts can also reduce the possibility of locating an equivalent pre-coordinated concept. While it is possible to continue with the extracted expression, the certainty of the encoding should be lower.

Finally, the newly inferred context (i.e., **363346000|Malignant neoplastic disease (disorder)|**) can be applied to the "liver" and "colon" concepts (i.e., **363346000|Malignant neoplastic disease (disorder)|:363698007|Finding site (attribute)|=10200004|Liver structure (body structure)|** and **363346000|Malignant neoplastic disease**

(disorder)|:363698007|Finding site (attribute)|=71854001|Colon structure (body structure)|). Additional equivalency testing can be done to determine if a pre-coordinated concept exists for the two concepts (discussed further in Section 7.3.3.3).

This method can be used to infer other context such as **408729009|Finding context (attribute)|**, **408731000|Temporal context (attribute)|** and **408732007|Subject relationship context (attribute)|** although some medical language processing tools may also be able to identify these contexts.

7.3.3.2 Additional Post-coordination

In some cases, additional post-coordination may be needed to accurately represent all the necessary semantics. For example, representing "acute pneumonia" as **233604007|Pneumonia (disorder)|:263502005|Clinical course (attribute)|=424124008|Sudden onset AND/OR short duration (qualifier value)|** may appear to be logical and complete but it is also necessary to refine **23583003|Inflammation (morphologic abnormality)|** to **4532008|Acute inflammation (morphologic abnormality)|**. It is unclear how many of these types of additional post-coordination exist. The example shown was discovered by accident but does demonstrate that such cases exist. Additional information on encoding "acute" finding is available in Section 12.6.8 in Appendix E.

7.3.3.3 Compacting Post-coordinated Expressions

As a candidate term that is being encoded with SNOMED CT may be a sentence that contains multiple clinical conditions that require inferences, equivalency testing should be conducted so that the final expressions will be closer-to-user-form. Two types of compacting have been identified here: general concepts and contextual qualifiers.

7.3.3.3.1 General Concepts

Returning to the example of "Cancer of lung, colon and liver," after inferring that "colon" and "liver" should not be left as concepts from the **123037004|Body structure (body structure)|** hierarchy, we end up the expressions shown in Table 7-19. For "colon" and "liver," these expressions should be converted to the long normal form and a search for a pre-coordinated concept should be carried out, which would yield the concepts **363406005|Malignant tumor of colon (disorder)|** and **93870000|Malignant neoplasm of liver (disorder)|**.

Table 7-19. Preliminary encoding of "cancer of lung, colon and liver."

No	Free Text	SNOMED CT Encoding
1.	Cancer of lung	363358000|Malignant tumor of lung (disorder)|
2.	Colon	363346000|Malignant neoplastic disease (disorder)|: 363698007|Finding site (attribute)|= 71854001|Colon structure (body structure)|
3.	Liver	363346000|Malignant neoplastic disease (disorder)|: 363698007|Finding site (attribute)|= 10200004|Liver structure (body structure)|

7.3.3.3.2 Contextual Qualifiers

According to the SNOMED CT UG, page 55, when a SNOMED CT concept does not have any explicit contextual qualifiers, the concept is referred to as having a default context. That means that a concept from the **404684003|Clinical finding (finding)|** hierarchy is implied to have **408729009|Finding context (attribute)|= 410515003|Known present (qualifier value)|** while a concept from the **71388002|Procedure (procedure)|** hierarchy is implied to have **408730004|Procedure context (attribute)|=385658003|Done (qualifier value)|**. The additional default contextual qualifiers for both concepts are **408731000|Temporal context (attribute)|=410512000|Current or specified time (qualifier value)|** and **408732007|Subject relationship context (attribute)|=410604004|Subject of record (person)|**.

When modeling contextual qualifier values, it is necessary to apply the values to **243796009|Situation with explicit context (situation)|** and then apply either the **246090004|Associated finding (attribute)|** or **363589002|Associated procedure (attribute)|** attributes. There are, however, some pre-coordinated concepts that already have the contextual values applied (refer to **Table 7-20**, **Table 7-21** and **Table 7-22**). If only one contextual value needs to be merged, then using these pre-coordinated concepts may be easier. The full list is shown in **Appendix E**. By using these pre-defined concepts, the post-coordinated expressions generated will be easier to read and recorded in the closer-to-user form. For example, instead of encoding the term "history of breast cancer" as **243796009|Situation with explicit context (situation)|:246090004|Associated finding (attribute)|=254837009|Malignant tumor of breast (disorder)|,408731000|Temporal context (attribute)|=410513005|In the past (qualifier value)|**, the expression can be simplified to **417662000|History of clinical finding in subject (situation)|:246090004|Associated finding (attribute)|=254837009|Malignant tumor of breast (disorder)|**. These default contexts can also be extended to expressions like family history or no family history.

Table 7-20. 410510008|Temporal context value (qualifier value)|=408731000|Temporal context (attribute)|

No	408731000|Temporal context (attribute)|	Pre-coordinated Concept
1.	410513005|In the past (qualifier value)|	417662000|History of clinical finding in subject (situation)|

Table 7-21. 408729009|Finding context (attribute)|=410514004|Finding context value (qualifier value)|

No	Finding Context Value	Pre-coordinated Concept with Finding Context
1.	410605003|Confirmed present (qualifier value)|	395098000|Disorder confirmed (situation)|
2.	410516002|Known absent (qualifier value)|	373572006|Clinical finding absent (situation)|
3.	410590009|Known possible (qualifier value)|	443859009|Possible clinical finding (situation)|
4.	410515003|Known present (qualifier value)|	373573001|Clinical finding present (situation)|
5.	428263003|NOT suspected (qualifier value)|	444436002|Clinical finding not suspected (situation)|
6.	415684004|Suspected (qualifier value)|	41769001|Disease suspected (situation)|

Table 7-22. 408730004|Procedure context (attribute)|=288532009|Context values for actions (qualifier value)|

No	Context Values for Actions	Pre-coordinated Concept
1.	385657008|Abandoned (qualifier value)|	180863000|Operation abandoned (situation)|
2.	410537005|Action status unknown (qualifier value)|	399714002|Procedure status unknown (situation)|
3.	410536001|Contraindicated (qualifier value)|	183932001|Procedure contraindicated (situation)|
4.	410546004|Discontinued (qualifier value)|	416406003|Procedure discontinued (situation)|
5.	385658003|Done (qualifier value)|	443938003|Procedure carried out on subject (situation)|
6.	410525008|Needed (qualifier value)|	417451006|Procedure needed (situation)|
7.	385660001|Not done (qualifier value)|	416237000|Procedure not done (situation)|
8.	410534003|Not indicated (qualifier value)|	428119001|Procedure not indicated (situation)|
9.	410530007|Not offered (qualifier value)|	416064006|Procedure not offered (situation)|
10.	410528005|Not wanted (qualifier value)|	416432009|Procedure not wanted (situation)|
11.	385650005|Organized (qualifier value)|	416662009|Procedure organized (situation)|
12.	397943006|Planned (qualifier value)|	183976008|Operative procedure planned (situation)|
13.	443390004|Refused (qualifier value)|	183944003|Procedure refused (situation)|
14.	385644000|Requested (qualifier value)|	400999005|Procedure requested (situation)|
15.	443942000|Requested by recipient (qualifier value)|	183995001|Patient requested procedure (situation)|
16.	385652002|Started (qualifier value)|	394906002|Procedure started (situation)|
17.	410545000|Stopped before completion (qualifier value)|	394908001|Procedure stopped (situation)|

7.4 Results

Two sets of results are presented here. First, the results of encoding the problem lists and encounter diagnoses. Second, the results of encoding 10 clinical statements from the literature.

7.4.1 Encoding Results

The results of encoding the problem lists and encounter diagnoses that were extracted from the primary care EMR system are shown in Table 7-23. The results are broken down into three categories: by record, by phrase, and by pre- or post-coordination. Fully encoded records mean the entire meaning was captured with the encoding, either via pre-coordinated concepts or post-coordinated expressions. For example, "backache" and "alcohol dependence chronic" were encoded as **161891005|Backache (finding)|** and **66590003|Alcohol dependence (disorder)|:263502005|Clinical course (attribute)|=90734009|Chronic (qualifier value)|** respectively. It should be noted that even though the entire record was encoded, it does not necessarily mean the encodings are correct. Partially, or incompletely, encoded records mean that only part of the record could be assigned SNOMED CT concepts. For example, "family situation stressful" was only encoded as **35359004|Family (social concept)|** as there were no matches to "situation" and "stressful." Partially encoded records are likely to be incorrect since there are deficiencies with the encoding. Fully unencoded records mean that the entire record could not be assigned any SNOMED CT concepts. For example, "compliant" was left unencoded because there are exact matches for "compliant," which in of itself ambiguous. Results by phrases break down the records into individual phrases by punctuations and conjunctions. Results by pre- and post-coordination are self-explanatory.

Table 7-23. Summary encoding results of problem list and encounter diagnoses.

No	Results	Problem List		Encounter Diagnoses	
		Total	Unique	Total	Unique
1.	Number of Records	20,141	2,881	266,032	15,123
	Fully Encoded Record	18,298	2,246	231,963	8,541
	Partially Encoded Record	1,666	582	27,198	5,794
	Fully Unencoded Record	177	53	6,868	787
	Nothing to Encode	0	0	3	1
2.	By Phrases				
	Encoded Phrases	23,382	2,239	299,300	7,357
	Unencoded Phrases	1,930	500	38,048	5,049
3.	Pre vs Post-coordination				
	Pre-coordination	21,114	1,689	270,937	4,541
	Post-coordination	2,268	561	28,363	2,895

The results of encoding the problem lists and encounter diagnoses by the top-level hierarchy are shown in **Table 7-24**. The hierarchies are based on pre-coordinated concepts and the focus concepts of post-coordinated expressions. For the problem list, the most common hierarchies were the **404684003|Clinical finding (finding)|**, **362981000|Qualifier value (qualifier value)|** and **105590001|Substance (substance)|**. For the encounter diagnoses, the most common hierarchies were the **404684003|Clinical finding (finding)|**, **71388002|Procedure (procedure)|** and **363787002|Observable entity (observable entity)|**. Examples of encounter diagnoses for each of the top-level hierarchies are shown in **Table 7-25**. The examples were selected to highlight problematic encodings.

Table 7-24. Encoding results of problem list and encounter diagnoses by top-level hierarchy.

No	Results	Problem List		Encounter Diagnoses	
		Total	Unique	Total	Unique
1.	123037004|Body structure (body structure)|	145	52	2,007	243
2.	404684003|Clinical finding (finding)|	19,990	1,826	197,248	4,905
3.	308916002|Environment or geographical location (environment / location)|	30	6	688	52
4.	272379006|Event (event)|	56	11	747	38
5.	106237007|Linkage concept (linkage concept)|	34	3	476	10
6.	363787002|Observable entity (observable entity)|	70	22	8,004	152
7.	410607006|Organism (organism)|	292	15	1,838	72
8.	373873005|Pharmaceutical / biologic product (product)|	64	6	1,370	18
9.	78621006|Physical force (physical force)|	1	1	19	6

No	Results	Problem List		Encounter Diagnoses	
		Total	Unique	Total	Unique
10.	260787004|Physical object (physical object)|	18	10	648	92
11.	71388002|Procedure (procedure)|	720	108	66,643	565
12.	362981000|Qualifier value (qualifier value)|	837	42	6,540	177
13.	419891008|Record artifact (record artifact)|			59	3
14.	243796009|Situation with explicit context (situation)|	345	62	4,620	698
15.	48176007|Social context (social concept)|	34	14	947	63
16.	370115009|Special concept (special concept)|	5	1	168	9
17.	123038009|Specimen (specimen)|			10	6
18.	105590001|Substance (substance)|	741	60	7,262	247

Table 7-25. Examples of encodings for the encounter diagnoses from each top-level hierarchy.

No	Hierarchy	Encounter Diagnoses	Encodings
1.	123037004|Body structure (body structure)|	Abdomen Gas	▪ 113345001|Abdominal structure (body structure)| ▪ 74947009|Gaseous substance (substance)|
2.	404684003|Clinical finding (finding)|	Backache	▪ 161891005|Backache (finding)|
3.	308916002|Environment or geographical location (environment / location)|	phone call from hospital	▪ 386472008|Telephone consultation (procedure)| ▪ 22232009|Hospital (environment)|
4.	272379006|Event (event)|	Assault	▪ 52684005|Assault (event)|
5.	106237007|Linkage concept (linkage concept)|	Anxiety Depressive Component	▪ 48694002|Anxiety (finding)| ▪ 35489007|Depressive disorder (disorder)| ▪ 246093002|Component (attribute)|
6.	363787002|Observable entity (observable entity)|	Nutrition Consult	▪ 384759009|Nutrition, function (observable entity)| ▪ 223475005|Consulting with (procedure)|
7.	410607006|Organism (organism)|	Trichomonas	▪ 3764006|Trichomonas (organism)|
8.	373873005|Pharmaceutical / biologic product (product)|	Hepatitis B Vaccine	▪ 34689006|Hepatitis B virus vaccine (product)|
9.	78621006|Physical force (physical force)|	Vaginal Spotting Light	▪ 9126005|Menstrual spotting (finding)| ▪ 56242006|Light, electromagnetic radiation (physical force)|
10.	260787004|Physical object (physical object)|	needle exchange	▪ 79068005|Needle, device (physical object)|
11.	71388002|Procedure (procedure)|	Colonoscopy	▪ 73761001|Colonoscopy (procedure)|
12.	362981000|Qualifier value (qualifier value)|	Benzodiazepine Use Chronic	▪ 372664007|Benzodiazepine (substance)| ▪ 419385000|Use - dosing instruction imperative (qualifier value)| ▪ 90734009|Chronic (qualifier value)|
13.	419891008|Record artifact (record artifact)|	death certificate	▪ 307930005|Death certificate (record artifact)|

No	Hierarchy	Encounter Diagnoses	Encodings
14.	243796009|Situation with explicit context (situation)|	Colon cancer family history	▪ 57177007|Family history with explicit context (situation)|
15.	48176007|Social context (social concept)|	Family Situation Stressful	▪ 35359004|Family (social concept)|
16.	370115009|Special concept (special concept)|	Abscess Soft Tissue	▪ 128477000|Abscess (disorder)| ▪ 181607009|Soft tissue (navigational concept)|
17.	123038009|Specimen (specimen)|	Stool Sample bottle given	▪ 119339001|Stool specimen (specimen)| ▪ 68276009|Bottle, device (physical object)|
18.	105590001|Substance (substance)|	Marijuana smoking	▪ 398705004|Cannabis (substance)| ▪ 365981007|Finding of tobacco smoking behavior (finding)|

The most commonly used attributes and most commonly used post-coordinated expressions (excluding duplicate types of expressions such as the use of **263502005|Clinical course (attribute)|=90734009|Chronic (qualifier value)|**) are shown in **Table 7-26** while the commonly used post-coordinated expressions are shown in **Table 7-27**.

Table 7-26. Most commonly used Concept Model attributes in the post-coordination expressions.

No	Attributes (Including Combinations of Attributes)	Frequency
1.	▪ 363698007|Finding site (attribute)|	1,206
2.	▪ 272741003|Laterality (attribute)|	406
3.	▪ 263502005|Clinical course (attribute)|	327
4.	▪ 246090004|Associated finding (attribute)| ▪ 408729009|Finding context (attribute)|	288
5.	▪ 363698007|Finding site (attribute)| ▪ 272741003|Laterality (attribute)|	141
6.	▪ 246090004|Associated finding (attribute)|	138
7.	▪ 363713009|Has interpretation (attribute)|	97
8.	▪ 363714003|Interprets (attribute)|	93
9.	▪ 363701004|Direct substance (attribute)|	67
10.	▪ 363589002|Associated procedure (attribute)| ▪ 408730004|Procedure context (attribute)|	64

Table 7-27. Most commonly used post-coordinated expressions in the encounter diagnoses.

No	Post-coordinated Expression	Original Phrase (Frequency)	Frequency
1.	103742009|Renewal of prescription (procedure)|: 405815000|Procedure device (attribute)|= 359993007|Telephone (physical object)|	▪ prescription refilled by telephone (6,424) ▪ Medication refill - telephone (1) ▪ Medication refill by telephone. (1) ▪ medication renewal by telephone. (1) ▪ prescription renewal by telephone. (1)	6,428

No	Post-coordinated Expression	Original Phrase (Frequency)	Frequency
2.	161891005|Backache (finding)|: {263502005|Clinical course (attribute)|= 90734009|Chronic (qualifier value)|, 363698007|Finding site (attribute)|= 122496007|Lumbar spine structure (body structure)|}	▪ Back pain lumbar chronic (712)	712
3.	308728002|Cervical smear biopsy taken (procedure)|: 363703001|Has intent (attribute)|= 360156006|Screening - procedure intent (qualifier value)|	▪ papanicolaou smear screening (652)	652
4.	243796009|Situation with explicit context (situation)|: 246090004|Associated finding (attribute)|= 35489007|Depressive disorder (disorder)|, 408729009|Finding context (attribute)|= 410590009|Known possible (qualifier value)|	▪ Depression - query (548) ▪ Depression query (1) ▪ depression query (1) ▪ Query depressive disorder (1)	551
5.	233604007|Pneumonia (disorder)|: 263502005|Clinical course (attribute)|= 373933003|Acute onset (qualifier value)|	▪ Acute pneumonia (428) ▪ Pneumonia acute (1)	429
6.	243796009|Situation with explicit context (situation)|: 246090004|Associated finding (attribute)|= 165816005|Human immunodeficiency virus positive (finding)|, 408729009|Finding context (attribute)|= 410519009|At risk context (qualifier value)|	▪ HIV risk (373)	373
7.	66590003|Alcohol dependence (disorder)|: 263502005|Clinical course (attribute)|= 90734009|Chronic (qualifier value)|	▪ Alcohol dependence chronic (337)	337
8.	59108006|Injection (procedure)|: 363701004|Direct substance (attribute)|= 387381009|Methotrexate (substance)|	▪ methotrexate injection (253) ▪ Methotrexate Injection (1)	254
9.	21522001|Abdominal pain (finding)|	▪ Abdominal pain (516) ▪ Abdomen pain (149) ▪ pain (94) ▪ abdominal pain (8) ▪ abdominal (2) ▪ .abdominal pain. (1) ▪ abdomen pain (1) ▪ Abdomen- (1) ▪ Abdominal pain - (1) ▪ abdominal pain. (1)	247
10.	45326000|Shoulder pain (finding)|: 363698007|Finding site (attribute)|= 91774008|Structure of right shoulder region (body structure)|	▪ Shoulder pain - right (227) ▪ pain shoulder . right (1) ▪ Right shoulder pain (1) ▪ right shoulder pain (1) ▪ right shoulder pain. (1)	231
11.	73595000|Stress (finding)|: 363713009|Has interpretation (attribute)|= 75540009|High (qualifier value)|	▪ high stress (185)	185
12.	267073005|Suicidal (finding)|: 363714003|Interprets (attribute)|= 89090003|Ideation, function (observable entity)|	▪ Suicidal ideation (130) ▪ for suicidal ideation (1) ▪ suicidal ideation (1) ▪ suicidal ideations. (1)	133
13.	300950007|Infected hand (disorder)|	▪ Hand infection (116)	116

No	Post-coordinated Expression	Original Phrase (Frequency)	Frequency
14.	56265001|Heart disease (disorder)|: 363714003|Interprets (attribute)|= 80943009|Risk factor (observable entity)|	▪ Cardiac disease risk factor (71)	71
15.	95324001|Skin lesion (disorder)|: 363698007|Finding site (attribute)|= 73897004|Skin structure of face (body structure)|	▪ Face skin lesion (50) ▪ Face lesion (10) ▪ Face lesions (1) ▪ face lesions (1)	62

7.4.2 Review of Post-coordinated Expressions

Two sets of post-coordinated expressions were selected for manual review. First, the most frequent post-coordinated expressions. Second, one percent of all post-coordinated expressions were randomly selected.

7.4.2.1 Most Frequent Post-coordinated Expressions

There were 2,895 unique post-coordinated expressions used in the encoding of the encounter diagnoses. Of the 2,895 expressions, 366 occurred at least 10 times. A preliminary review of the expressions found that 300 (82.0%) were accurate while 66 were incorrect or I was uncertain of the encoding. Examples of the incorrect or uncertain encodings are shown in **Table 7-28** along with a brief explanation on why it was encoded incorrectly. In all cases it was because there was no synonym so the term was split into smaller chunks, encoded individually and merged together.

Table 7-28. Examples of incorrect post-coordinated expressions from the most frequently used post-coordinated expressions and reason for the error.

No	Original Term	Encoding	Reason
1.	Drug counselling	*Post-coordinated Expression* 409063005|Counseling (procedure)|: 363701004|Direct substance (attribute)|= 410942007|Drug or medicament (substance)| *Correct Encoding* 313071005|Counseling for substance abuse (procedure)|	There is no synonym for "drug counselling." The subtype of 313071005|Counseling for substance abuse (procedure)| is 60112009|Drug addiction counseling (procedure)| but it may be too specific.
2.	Maternal anxiety	*Post-coordinated Expression* 243796009|Situation with explicit context (situation)|: 246090004|Associated finding (attribute)|= 48694002|Anxiety (finding)|, 408732007|Subject relationship context (attribute)|= 72705000|Mother (person)| *Correct Encoding* 270472006|Maternal concern (finding)|	There is no synonym for "maternal anxiety." A new synonym should be added for "maternal anxiety" or "anxiety" should be an equivalent term for "concern."
3.	Prostate hypertrophy	*Post-coordinated Expression* 64572001|Disease (disorder)|: {116676008|Associated morphology (attribute)|= 56246009|Hypertrophy (morphologic abnormality)|, 363698007|Finding site (attribute)|= 41216001|Prostatic structure (body structure)|} *Correct Encoding* 266569009|Benign prostatic hyperplasia (disorder)|	"Prostate" was not equated with "prostatic" so the match did not occur with one of the synonyms.

No	Original Term	Encoding	Reason
4.	Mental retardation high functioning	*Post-coordinated Expression* 91138005|Mental retardation (disorder)|: {363713009|Has interpretation (attribute)|= 75540009|High (qualifier value)|, 363714003|Interprets (attribute)|= 246464006|Function (observable entity)|}	No term was found
5.	Eye redness	*Post-coordinated Expression* 386713009|Red color (finding)|: 363698007|Finding site (attribute)|= 81745001|Structure of eye proper (body structure)| *Correct Encoding* 75705005|Red eye (disorder)|	There is no synonym for "eye redness."
6.	Eye problem	*Post-coordinated Expression* 55607006|Problem (finding)|: 363698007|Finding site (attribute)|= 81745001|Structure of eye proper (body structure)|	There is no synonym for "eye problem." 371405004|Disorder of eye proper (disorder)| may be most appropriate?
7.	Annual physical examination	*Post-coordinated Expression* 5880005|Physical examination procedure (procedure)|: 370134009|Time aspect (attribute)|= 53281000|Annual (qualifier value)| *Correct Encoding* 78318003|History and physical examination, annual for health maintenance (procedure)|	There is no shorter synonym for just "annual physical examination."
8.	Inguinal hernia reducible	*Post-coordinated Expression* 396232000|Inguinal hernia (disorder)|: 363713009|Has interpretation (attribute)|= 255359009|Reducible (qualifier value)|	It is uncertain if "reducible" is encoded correctly. While it is technically correct, it may not be semantically correct.
9.	Adult failure to thrive	*Post-coordinated Expression* 243796009|Situation with explicit context (situation)|: 408732007|Subject relationship context (attribute)|= 133936004|Adult (person)|, 246090004|Associated finding (attribute)|= 433476000|Failure to thrive in infant (disorder)| *Correct Encoding* 129588001|Adult failure to thrive syndrome (disorder)|	The concept 129588001|Adult failure to thrive syndrome (disorder)| does not have a synonym without "syndrome." Since an exact match did not occur, the term was split and encoded individually. "Failure to thrive" as a synonym for 433476000|Failure to thrive in infant (disorder)| may be too ambiguous.

7.4.2.2 Randomly Selected

One percent (284) of all post-coordinated expressions (28,363) were randomly selected and manually reviewed for accuracy. A preliminary review of the expressions found that 241 (84.9%) were accurate while 47 were incorrect or I was uncertain of the encoding. Examples of the incorrect or uncertain encodings are shown in **Table 7-29** along with a brief explanation on why it was encoded incorrectly.

Table 7-29. Examples of incorrect post-coordinated expressions from the random sample and reason for the error.

No	Original Term	Post-coordinated Expression	Reason
1.	Toe callus	*Post-coordinated Expression* 201040000|Callosity (disorder)|: {363698007|Finding site (attribute)|= 52034004|Skin structure of toe (body structure)|}	The correct concept has synonyms such as "callosity on toe" and "callus plantar digital" but not "toe callus."

No	Original Term	Post-coordinated Expression	Reason
		Correct Encoding 201041001|Callosity on toe (disorder)|	While the expression created is technically correct, it lacks the definition of "246075003|Causative agent (attribute)|= 263762005|Friction (physical force)|", which is how encoding of the correct concept is defined.
2.	Plantar fasciitis - right foot	*Post-coordinated Expression* 202882003|Plantar fasciitis (disorder)|: {363698007|Finding site (attribute)|= 56459004|Foot structure (body structure)|} {363698007|Finding site (attribute)|=(64799002|Plantar aponeurosis structure (body structure)|: 272741003|Laterality (attribute)|= 24028007|Right (qualifier value)|)} *Correct Encoding* 202882003|Plantar fasciitis (disorder)|: {363698007|Finding site (attribute)|=(64799002|Plantar aponeurosis structure (body structure)|: 272741003|Laterality (attribute)|= 24028007|Right (qualifier value)|)}	The main issue with the expression is that "56459004|Foot structure (body structure)|" is not necessary. "Plantar fasciitis" implicit refers to the foot.
3.	Acute pneumonia	*Post-coordinated Expression* 233604007|Pneumonia (disorder)|: 63502005|Clinical course (attribute)|= 373933003|Acute onset (qualifier value)| *Correct Encoding* 233604007|Pneumonia (disorder)|: 63502005|Clinical course (attribute)|= 424124008|Sudden onset AND/OR short duration (qualifier value)| 116676008|Associated morphology (attribute)|= 4532008|Acute inflammation (morphologic abnormality)|	"123592004|Acute mucous pneumonia (disorder)|" is also a potential match as it has the same defining attributes as the correct encoding. As described in the section on additional post-coordination, there are cases where there is a need to add additional post-coordination in order to capture all the necessary semantics.
4.	Suicidal ideation	*Post-coordinated Expression* 267073005|Suicidal (finding)|: 363714003|Interprets (attribute)|= 89090003|Ideation, function (observable entity)| *Correct Encoding* 6471006|Suicidal thoughts (finding)|	While not technically incorrect, there is a pre-coordinated concept that should be used. 89090003|Ideation, function (observable entity)| is also probably the incorrect concept.
5.	Abuse physical neglect	*Post-coordinated Expression* 418506006|Physical abuse (event)|: 47429007|Associated with (attribute)|= 95930005|Victim of neglect (finding)|	It is uncertain if the post-coordinated expression is appropriate.
6.	tuberculin skin test	*Post-coordinated Expression* 53309004|Skin test (procedure)|: 363701004|Direct substance (attribute)|= 423321003|Tuberculin purified protein derivative (product)|	It is uncertain what the correct encoding should be but the expression is probably incorrect.
7.	tuberculosis read	*Post-coordinated Expression* 56717001|Tuberculosis (disorder)|: 363714003|Interprets (attribute)|= 50360004|Reading (observable entity)|	It is uncertain what the correct encoding should be but the expression is probably incorrect.

No	Original Term	Post-coordinated Expression	Reason
8.	Injection drug use	*Post-coordinated Expression* 59108006|Injection (procedure)|: {260686004|Method (attribute)|= 419385000|Use - dosing instruction imperative (qualifier value)|, 363701004|Direct substance (attribute)|= 410942007|Drug or medicament (substance)|} *Correct Encoding* 228388006|Intravenous drug user (finding)|	"Use" was incorrectly encoded as 419385000|Use - dosing instruction imperative (qualifier value)|. In addition, there is a pre-coordinated concept that should be used.
9.	Arm cast	*Post-coordinated Expression* 64572001|Disease (disorder)|: {116676008|Associated morphology (attribute)|= 21718009|Cast (morphologic abnormality)|, 363698007|Finding site (attribute)|= 53120007|Upper limb structure (body structure)|} *Correct Encoding* 438409006|Application of short arm cast (procedure)| 440490004|Application of long arm cast (procedure)|	"Cast" was inappropriately encoded. The possible correct encodings do not have synonyms for "arm cast" and would have result in partial matches.

7.4.3 Comparison with Other Methods

Ten clinical statements were extracted from the literature in which errors were identified in **Section 7.1.3** to determine if the encoding method could accurately encode them and were manually reviewed. The results are shown in **Table 7-30**. Of the 10 clinical statements, three were deemed to be correct, three were incorrect, three were partially correct and one was not encoded.

In the first two statements, **35209006|Sensitivity (finding)|** is incorrect because it refers to **106131003|Mood finding (finding)|**. A manual lookup did not yield any concepts that could be used. In the second example, **260365005|Heightened (qualifier value)|**, while technically correct, is probably semantically incorrect. It should be noted **that 260365005|Heightened (qualifier value)|** is not used in any of the defining attributes so it is unclear how it should be used.

In the third statement, there is a concept pre-coordinated concept **238402004|Cellulitis of leg (disorder)|**. To ensure the encoding method would create a post-coordinated expression, **238402004|Cellulitis of leg (disorder)|** was added to the list of exclude concepts. The resulting encoding **128045006|Cellulitis (disorder)|:363698007|Finding site (attribute)|=61685007|Lower limb structure (body structure)|** was correct.

In the fourth statement, there was no exact match for the word "bronchodilation" and therefore only "spirometry" was encoded. The concept that was used in the literature, **255274008|Pre-bronchodilation (qualifier value)|**, cannot be accessed via the Concept Model. In the fifth statement, there was no exact match for the word "parenting." There was an exact match for effective (**254648000|Effective (qualifier value)|**) but it cannot be accessed via the Concept Model.

In the sixth statement, the coding in the literature used **21082005|Entire mouth region (body structure)|** while the encoding method used **123851003|Mouth region structure (body structure)|**. The algorithm to sort the search results indicated that the more general concept should be used (refer to Section 7.3.1.3.2) so this encoding was deemed correct.

In the seventh statement, a pre-coordinated concept **305997006|Referral by radiologist (procedure)|** was identified. That concept has been active since 2002 and the authors were using SNOMED CT so it is unclear why they did not use that concept. In the eighth statement, "abuse verbal relative" is ambiguous as to whether an individual was verbally abused by a relative or a relative was verbally abused. The encoding method encoded the statement as the former but should have also encoded it as **225825002|Victim of verbal abuse (finding)|**.

In the ninth statement, "dislocation knee simple" was encoded as **58320001|Traumatic dislocation of knee joint (disorder)|** with "simple" not being captured. There is a concept **50399007|Closed dislocation (morphologic abnormality)|** that could have been used but the algorithm encoded "dislocation knee" first. "Dislocation" and "simple" were also separated by the word "knee," which would not have made it possible to encode as one concept. Incidentally, if "dislocation simple knee" was used, it would have been encoded as **64572001|Disease (disorder)|:116676008|Associated morphology (attribute)|=50399007|Closed dislocation (morphologic abnormality)|,363698007|Finding site (attribute)|=72696002|Knee region structure (body structure)|**. **58320001|Traumatic dislocation of knee joint (disorder)|** has a **363698007|Finding site (attribute)|** of **49076000|Knee joint structure (body structure)|**, which is a subtype of **72696002|Knee region structure (body structure)|**.

In the last example, the encoding method identified "statin" correctly but incorrectly encoded "age" as **248280005|Aging (finding)|**. There does not appear to be a concept that can represent "age of first statin."

Table 7-30. Results of encoding 10 clinical statements from the literature.

No	Clinical Statement	Literature	Encoding Method	Result
1.	Multiple chemical sensitivity	281867008|Multisystem Disorder|: {47429007|associated with|= 35209006|sensitivity|: 410658008|triggered by|= 441900009|chemical| 276339004|environment|}	441900009|Chemical (substance)| 35209006|Sensitivity (finding)|	Incorrect
2.	Heightened sensitivity to environment	35209006|sensitivity|: {37135001|tolerance related finding|= 260365005|heightened|: 44190009|triggered by|= 276339004|environment|}	35209006|Sensitivity (finding)|: 363713009|Has interpretation (attribute)|= 260365005|Heightened (qualifier value)| 276339004|Environment (environment)|	Incorrect
3.	Cellulitis of leg	128045008 Cellulitis (disorder) + 61685007 lower limb structure (body structure)	128045006|Cellulitis (disorder)|: 363698007|Finding site (attribute)|= 61685007|Lower limb structure (body structure)|	Correct
4.	Spirometry without bronchodilation	127783003|Spirometry (procedure)|+ 255274008|Pre-bronchodilation (qualifier value)|	127783003|Spirometry (procedure)|	Partial
5.	Effective parenting	405175009|Parenting behavior (observable entity)|+ 254648000|Effective (qualifier value)|	N/A	

No	Clinical Statement	Literature	Encoding Method	Result
6.	Pain; mouth	22253000 pain (clinical finding) + 21082005 entire mouth region (body structure): relationship type = 363698007 finding site (attribute)	22253000|Pain (finding)|: 363698007|Finding site (attribute)|= 123851003|Mouth region structure (body structure)|	Correct
7.	Referral; radiologist	3457005 patient referral (procedure) + 66862007 radiologist (occupation): relationship type = 370131001 recipient category (attribute)	305997006|Referral by radiologist (procedure)|	Correct
8.	Abuse; verbal; relative	125677006 Relative (person) + 225825002 Victim of verbal abuse (clinical finding): relationship type = indeterminate	243796009|Situation with explicit context (situation)|: 246090004|Associated finding (attribute)|= 225478005|Verbally abusive behavior (finding)|, 408732007|Subject relationship context (attribute)|= 125677006|Relative (person)|	Partial
9.	Dislocation; knee; simple	13673007 Simple (qualifier value) + 129156001 Traumatic dislocation of knee joint (clinical finding): relationship type = 246100006 onset (attribute)	58320001|Traumatic dislocation of knee joint (disorder)|	Partial
10.	Age of first statin	363819003|drug therapy observable|: {24645002|occurrence|=255216001|first|}, {127489000|has active ingredient|= 6302009|HMG-CoA reductase inhibitor|}	248280005|Aging (finding)| 372912004|3-Hydroxy-3-methylglutaryl-coenzyme A (HMG-CoA) reductase inhibitor (substance)|	Incorrect

7.5 Discussion

In this study, an encoding method was developed and was used to encode problem lists and encounter diagnoses from a primary care dataset as well as 10 clinical statement extracted from the literature. The encoding method consisted of three parts: improving the quality of search results, a nine-scenario process of modeling two or more expressions together, and a cleaning up of the post-coordinated expression. The discussion is divided into four parts. First, general observations of the encoding results. Second, challenges encountered in the development of the encoding method. Third, a comparison with other methods. Fourth, the certainty of encoding. The limitations of the encoding method and future work are described in **Chapter Ten**.

7.5.1 General Observations of the Encoding Results

A review of the post-coordinated expressions that were used at least 10 times in the encounter diagnoses showed that they were correct 82% of the time while a review of a random selection of one percent of all post-coordinated expressions showed that they were correct 85% of the time. The encoding method performed well for commonly used post-coordination such as chronicity (e.g., "alcohol dependence chronic" **66590003|Alcohol dependence (disorder)|:263502005|Clinical course (attribute)|=90734009|Chronic (qualifier value)|**) and laterality (e.g., "shoulder pain right" **45326000|Shoulder pain (finding)|:363698007|Finding site (attribute)|=91774008|Structure of right shoulder region (body structure)|**). The compacting function also worked well to identify pre-coordinated concepts from post-coordinated expressions. For example, it identified that

16982005|Shoulder region structure (body structure)|:272741003|Laterality (attribute)|=24028007|Right (qualifier value)| was equivalent to **91774008|Structure of right shoulder region (body structure)|**. It was also able to identify the correct body structure using the proposed Extended Concept Model for the examples reviewed. For example, for the term "face skin lesion," the concept **73897004|Skin structure of face (body structure)|** was identified instead of **89545001|Face structure (body structure)|**. In these cases, the automatically generated post-coordinated expressions were able to represent the original terms accurately and would have been more efficient than manually crafting them.

While the encoding method was not limited to specific hierarchies, a review of the results by hierarchy can help to identify which SNOMED CT expressions are suitable for a problem list or encounter diagnoses. For example, problem lists should ideally contain expressions (pre-coordinated concepts or focus concepts of post-coordinated expressions) from the **404684003|Clinical finding (finding)|**, **272379006|Event (event)|**, **71388002|Procedure (procedure)|** and **243796009|Situation with explicit context (situation)|** hierarchies.[16] Therefore expressions that do not fall into these hierarchies should be reviewed. A review of examples of expressions from outside the four hierarchies identified indicated that the encoding was incomplete and the expressions require additional modelling.

7.5.1.1 Uncertainty of Accuracy

There remains uncertainty of the accuracy of some of the encoding results. For example, "HIV risk" was encoded as a **243796009|Situation with explicit context (situation)|**, but there are a set of concepts that deal with "at risk" findings (i.e., **281694009|Finding of at risk (finding)|**). A challenge is that the subtype concepts are primitive and are not linked to the actual diseases. For example, **161639008|At risk of heart disease (finding)|** does not have **56265001|Heart disease (disorder)|** as a defining attribute. In the post-coordinated expression used, **165816005|Human immunodeficiency virus positive (finding)|** (or possibly **86406008|Human immunodeficiency virus infection (disorder)|**) is part of the expression and **410519009|At risk context (qualifier value)|** is used to indicate the "at risk" portion. It is unclear if this is the best way to encode "at risk" conditions but the advantage over using concepts that are subtypes of **281694009|Finding of at risk (finding)|** is that the concept is part of the definition. There are only three concepts that use **410519009|At risk context (qualifier value)|** as part of the defining attributes: **443999008|Risk of exposure to communicable disease (situation)|**, **398819009|Diabetic foot at risk (situation)|** and **444237009|Risk of exposure to Leptospira (situation)|**. There needs to be a forum where uncertain encodings can be discussed.

7.5.1.2 Identification of Synonyms

An unintended consequence of using this encoding method is that it can help to identify synonyms for pre-coordinated concepts. In the sample results in Table 7-27, there were two pre-coordinated concepts even though it was identified as post-coordination. The reason is that the post-coordinated expressions were "compacted," and an equivalent pre-coordinated concept was found. In the first example, "abdomen pain" was found to be a synonym for "abdominal pain." As there was no exact match for "abdomen pain," the term was encoded individually as **113345001|Abdominal structure (body structure)|** using the description **187547015|Abdomen|** and **22253000|Pain (finding)|** using the description **37361011|Pain|**. The two concepts were merged together as **22253000|Pain (finding)|:363698007|Finding site (attribute)|=113345001|Abdominal structure (body structure)|** and during the compacting phase, it was found to be equivalent to **113345001|Abdominal structure (body structure)|**. In the second example, "hand infection" was found to be a synonym for "infected hand." "Hand infection" was encoded separately as **85562004|Hand structure (body structure)|** using description **141819019|Hand|** and **40733004|Infectious disease (disorder)|** using description **67942010|Infection|**. The two concepts were merged as **40733004|Infectious disease (disorder)|:363698007|Finding site (attribute)|= 85562004|Hand structure (body structure)|** and during the compacting phase, it was found to be equivalent to **300950007|Infected hand (disorder)|**. When found, these synonyms should be added as extension descriptions or to the core to improve the content coverage.

7.5.2 Challenges

7.5.2.1 Incorrect Encoding

Creating incorrect post-coordinated expression continues to be a challenge. A review of the post-coordinated expressions that occurred at least 10 times in the encounter diagnoses showed that there were still semantic errors in the encodings. Three main reasons were identified during a review of the encoding results.

First, the original term was ambiguous; therefore there was only so much the encoding method could do. For example, the encounter diagnosis "Trichomonas" was encoded as **3764006|Trichomonas (organism)|** when it probably should have been encoded as **56335008|Infection by Trichomonas (disorder)|**. A second example is when the encounter diagnosis "Hepatitis B Vaccine" was encoded as **34689006|Hepatitis B virus vaccine (product)|** when it probably should have been encoded as **16584000|Hepatitis B vaccination (procedure)|**. In both cases the encoding was not completely off as **56335008|Infection by Trichomonas (disorder)|** is defined with **246075003246075003|Causative agent (attribute)|=3764006|Trichomonas (organism)|** and **16584000|Hepatitis B vaccination (procedure)|** is defined with **363701004|Direct substance (attribute)|=396424005|Hepatitis B virus vaccine (substance)|**. Inference could be made to locate a **404684003|Clinical finding (finding)|** concept but in some cases more than one concept may be applicable and automatically associating it with a concept may be presumptuous.

Second, the lack of synonyms available to locate a match. For example, "abdomen gas" was encoded separately as **113345001|Abdominal structure (body structure)|** and **74947009|Gaseous substance (substance)|**. The correct encoding may be **271835004|Abdominal distension, gaseous (finding)|** but it does not have a synonym for "abdomen gas"; therefore "abdomen gas" was encoded as **74947009|Gaseous substance (substance)|**. In this case, "gas" should probably have been encoded as **392610001|Intestinal gas (substance)|**. In another example, the post-coordinated expression used to represent "suicidal ideation" is technically correct, but a more appropriate concept is probably **6471006|Suicidal thoughts (finding)|**. There are no concepts that are defined with **89090003|Ideation,**

function (observable entity)|. Only a partial match would have identified **6471006|Suicidal thoughts (finding)|**. Additional work needs to be done to add these synonyms to SNOMED CT, either as extension descriptions at an organisational level or at the international level if these synonyms are used frequently.

Third, the further need to extend the Concept Model. In the example "high stress," the expression was probably encoded incorrectly as there are pre-coordinated concepts such as **23085004|Increased stress (finding)|** and **424582000|Stress overload (finding)|**. **75540009|High (qualifier value)|** is used in defining 10 concepts such as **44393007|Increased granulocyte destruction (finding)|** and **79738005|Increased platelet production (finding)|**. There is therefore a need to constrain how general qualifier concepts such as **75540009|High (qualifier value)|** are used and to ensure concepts are used in the appropriate context.

7.5.2.2 Incomplete Encoding

"Acute pneumonia" was encoded as **233604007|Pneumonia (disorder)|:263502005|Clinical course (attribute)|=373933003|Acute onset (qualifier value)|**, when it actually should have been encoded as **233604007|Pneumonia (disorder)|:4532008|Acute inflammation (morphologic abnormality)|,263502005|Clinical course (attribute)|=424124008|Sudden onset AND/OR short duration (qualifier value)|**. Section 7.3.3.1 was written to help resolve this incompleteness encoding challenge. It is uncertain how many other types of expressions require multiple qualifications and refinements in order to accurately represent the clinical condition.

7.5.2.3 Equivalency

The encoding of "prescription refilled by telephone" appears to be correct although **405815000|Procedure device (attribute)|** could have been **424226004|Using device (attribute)|** as well. **363699004|Direct device (attribute)|** and **363710007|Indirect device (attribute)|** are subtypes of **405815000|Procedure device (attribute)|** while **425391005|Using access device (attribute)|** is a subtype of **424226004|Using device (attribute)|**. Although there are concepts such as **185317003|Telephone encounter (procedure)|** and **401267002|Telephone triage encounter (procedure)|**, those concepts do not use **359993007|Telephone (physical object)|** as part of the defining attributes. There is only one concept, **312635000|Provision of telephone (procedure)|** that uses **359993007|Telephone (physical object)|** as part of the defining attributes and that is through the attribute **363699004|Direct device (attribute)|**. In the future if a pre-coordinated concept for "prescription refilled by telephone" is released, it is uncertain if that concept will be equivalent to the post-coordinated expression created here seeing that concepts that refer to "telephone" do not use it in the defining attributes.

In a second example, "Back pain lumbar chronic" was an interesting candidate term for encoding because there are many ways in which it can be represented. There are pre-coordinated concepts such as **134407002|Chronic back pain (finding)|** and **279039007|Low back pain (finding)|**, with the latter concept having a synonym of **416144015|Lumbar pain|**. If starting with the first concept, and using the **363698007|Finding site (attribute)|=52612000|Lumbar region back structure (body structure)|** that was in **279039007|Low back pain**

(finding)|, we would have the expression **134407002|Chronic back pain (finding)|:363698007|Finding site (attribute)|=52612000|Lumbar region back structure (body structure)|**. If starting with the second concept, and using **263502005|Clinical course (attribute)|=90734009|Chronic (qualifier value)|** to represent "chronic," we would end up with **279039007|Low back pain (finding)|:263502005|Clinical course (attribute)|=90734009|Chronic (qualifier value)|**. There is also a concept **278860009|Chronic low back pain (finding)|**, which is a subtype of **279039007|Low back pain (finding)|**, has a preferred term of **216213018|Chronic low back pain|** and synonym of **415888015|CLBP - Chronic low back pain|** but not "chronic lumbar pain," and also uses **363698007|Finding site (attribute)|=52612000|Lumbar region back structure (body structure)|** but does include **279039007|Low back pain (finding)|:263502005|Clinical course (attribute)|=90734009|Chronic (qualifier value)|**. The four expressions are shown in **Table 7-31**. Herein lies a challenge where there are multiple ways that could be used to represent "back pain lumbar chronic," but they are not equivalent. **216213018|Chronic low back pain|** should probably be used, however, as it does not have "lumbar" as one of its synonyms, therefore the lexical match failed.

Table 7-31. Different ways of representing "back pain lumbar chronic."

No	Closer-to-User Form	Frequency
1.	161891005|Backache (finding)|: {263502005|Clinical course (attribute)|= 90734009|Chronic (qualifier value)|, 363698007|Finding site (attribute)|= 122496007|Lumbar spine structure (body structure)|}	Subsumes 3rd expression, unrelated to 2nd and 4th expressions
2.	134407002|Chronic back pain (finding)|: 363698007|Finding site (attribute)|= 52612000|Lumbar region back structure (body structure)|	Unrelated to all three other expressions.
3.	279039007|Low back pain (finding)|: 263502005|Clinical course (attribute)|= 90734009|Chronic (qualifier value)|	Subsumed by 1st expression, unrelated to 2nd and 4th expressions
4.	278860009|Chronic low back pain (finding)|	Unrelated to all three other expressions but probably the accurate concept

7.5.3 Comparison with Other Methods

Although encoding the 10 clinical statements extracted from the literature did not fare well, there are a number of valid reasons for that. First, the concepts do not exist in SNOMED CT. For example, there is no known concept to represent "chemical sensitivity," "sensitivity to environment" or "effective parenting." Interestingly there are concepts to represent impaired, deficient and abusive parenting. Second, the descriptions do not exist in SNOMED CT. For example, there was no exact match to "bronchodilation" and it is difficult to equate "bronchodilation" with "pre-bronchodilation." There is nothing that can be done to locate discrete concepts if they do not exist. Third, the original clinical statements were ambiguous. The "relative" in the "abuse verbal relative" was ambiguous. Nevertheless, the pre-coordinated expressions that were deemed correct are promising as they were between findings and body structures, which are typically the most common types of post-coordination (refer to **Table 7-26**).

7.5.4 Certainty of Encoding

As the encoding method consists of many components, there is a need to measure how confident we are with the results of the encoding. As thousands of post-coordinated expressions were created, it is not possible to manually review all of them. A review of the most common post-coordinated expressions is also not very helpful as the most common post-coordinated expression are also generally the simplest, such as chronicity and laterality, that leaves little room for error. The certainty of encoding need not necessarily be a statistical indicator but should give a general sense on the encoding accuracy. Below are some factors that should be taken into account when developing a certainty of encoding score.

Type of Lexical Match. In this encoding method, only exact matches, both against the original and normalised SNOMED CT descriptions, and against the list of acronyms generated are used. Of the three, an exact match against the original description has the highest chance of being accurate. Normalised descriptions are sorted alphabetically so there is a chance of misinterpretation. Acronyms are generally problematic because of the wide range of conditions they may refer to. Certainly acronyms like COPD are unambiguous but that would require ranking the acronyms as well.

Type of Post-coordination. The encoding method expanded the type of post-coordination to include template encoding and the extended Concept Model. Refinement is probably the most accurate post-coordination method as we are refining what already exists in a concept's definition. Qualification is also reasonably accurate, although caution should be exercised when using **363713009|Has interpretation (attribute)|** and when modelling clinical statements asserting acute clinical course. Using an encoding template and the extended Concept Model will probably produce more accurate results than just using the Concept Model.

Concept Model Attributes and Ranges. The attributes used, or the combination of attributes used, in constructing the post-coordinated expressions can also be used to help determine the certainty of encoding. For example, **370134009|Time aspect (attribute)|** was used only eight times and closer attention should be paid to these attributes. Similarly, the range used in the post-coordinated expression also merits review. As qualifier values such as **75540009|High (qualifier value)|** can be ambiguous, certain ranges should be reviewed.

What Was Encoded. The encoding method developed was geared towards clinical statements that should be in phrases rather than embedded in paragraphs or sentences. If the entire candidate term could be encoded, there is more confidence in the encoding than if only parts of the candidate term were encoded.

Inferences. Encoding candidate terms with explicit context should yield better encoding results than candidate terms that contain multiple clinical terms that need to be parsed and encoded. Although inferring a clinical condition that should be applied to a body structure is a novel method of not leaving a concept as a body structure, it does open up the possibility of incorrect inferences.

7.6 References

[1] Lee, D., de Keizer, N., Lau, F., & Cornet, R. (2013). Literature review of SNOMED CT use. Journal of the American Medical Informatics Association. http://www.ncbi.nlm.nih.gov/pubmed/23828173.

[2] Navas, H., Lopez, O. A., Gambarte, L., Elías, L. G., Wasserman, S., Orrego, N., ... & de Quirós, F. G. (2010). Implementing rules to improve the quality of concept post-coordination with SNOMED CT. Studies in Health Technology and Informatics, 160(Pt 2), 1045. http://www.ncbi.nlm.nih.gov/pubmed/20841843.

[3] Richesson, R. L., Andrews, J. E., & Krischer, J. P. (2006). Use of SNOMED CT to represent clinical research data: a semantic characterization of data items on case report forms in vasculitis research. Journal of the American Medical Informatics Association, 13(5), 536-546. http://www.ncbi.nlm.nih.gov/pubmed/16799121.

[4] Lee, D. H., Lau, F. Y., & Quan, H. (2010). A method for encoding clinical datasets with SNOMED CT. BMC Medical Informatics and Decision Making, 10(1), 53. http://www.ncbi.nlm.nih.gov/pubmed/20849611.

[5] Sampalli, T., Shepherd, M., Duffy, J., & Fox, R. (2010). An evaluation of SNOMED CT® in the domain of complex chronic conditions. International Journal of Integrated Care, 10. http://www.ncbi.nlm.nih.gov/pubmed/20422022.

[6] Penz, J. F., Brown, S. H., Carter, J. S., Elkin, P. L., Nguyen, V. N., Sims, S. A., & Lincoln, M. J. (2004). Evaluation of SNOMED coverage of Veterans Health Administration terms. Medinfo, 11(Pt 1), 540-4. http://www.ncbi.nlm.nih.gov/pubmed/15360871.

[7] Fung, K. W., Xu, J., Rosenbloom, S. T., Mohr, D., Maram, N., & Suther, T. (2011). Testing Three Problem List Terminologies in a Simulated Data Entry Environment. In AMIA Annual Symposium Proceedings (Vol. 2011, p. 445). American Medical Informatics Association. http://www.ncbi.nlm.nih.gov/pubmed/22195098.

[8] Andrews, J. E., Richesson, R. L., & Krischer, J. (2007). Variation of SNOMED CT coding of clinical research concepts among coding experts. Journal of the American Medical Informatics Association, 14(4), 497-506. http://www.ncbi.nlm.nih.gov/pubmed/17460128.

[9] Randorff, H. A., & Rosenbeck, G. K. (2012). SNOMED CT Implementation. Mapping Guidelines Facilitating Reuse of Data. Methods of Information in Medicine, 51(6). http://www.ncbi.nlm.nih.gov/pubmed/23038162.

[10] Grain, H. (2010). Clinical terminology. Studies in Health Technology and Informatics, 151, 70-83. http://www.ncbi.nlm.nih.gov/pubmed/20407153.

[11] Bakhshi-Raiez, F., Ahmadian, L., Cornet, R., de Jonge, E., & de Keizer, N. F. (2010). Construction of an interface terminology on SNOMED CT: Generic approach and its application in intensive care. Methods of Information in Medicine, 49(4), 349-59. http://www.ncbi.nlm.nih.gov/pubmed/20582384.

[12] Park, H. A., Lundberg, C. B., Coenen, A., & Konicek, D. J. (2009). Evaluation of the content coverage of SNOMED-CT to represent ICNP Version 1 catalogues. Studies in Health Technology and Informatics, 146, 303-7. http://www.ncbi.nlm.nih.gov/pubmed/19592854.

[13] Wang Y, Patrick J, Miller G, & O'Hallaran J. A computational linguistics motivated mapping of ICPC-2 PLUS to SNOMED CT. BMC Medical Informatics and Decision Making. 2008 Oct 27;8 Suppl 1:S5. http://www.ncbi.nlm.nih.gov/pubmed/19007442.

[14] Pathak, J., Wang, J., Kashyap, S., Basford, M., Li, R., Masys, D. R., & Chute, C. G. (2011). Mapping clinical phenotype data elements to standardized metadata repositories and controlled terminologies: the eMERGE Network experience. Journal of the American Medical Informatics Association, 18(4), 376-386. http://www.ncbi.nlm.nih.gov/pubmed/21597104.

[15] International Health Terminology Standards Development Organisation. User Guide. http://www.ihtsdo.org/fileadmin/user_upload/doc/ug/index.html.

[16] Fung K.W., McDonald C., & Srinivasan S. The UMLS-CORE project: a study of the problem list terminologies used in large healthcare institutions. Journal of the Americal Medical Informatics Association. 2010 Nov-Dec;17(6):675-80. http://www.ncbi.nlm.nih.gov/pubmed/20962130.

8. RETRIEVAL METHOD & RESULTS

8.1 Introduction

One of the benefits of the Systematised Nomenclature of Medicine Clinical Terms (SNOMED CT) is the arrangement of concepts in the subsumption, or "is a" hierarchy, whereby more general concepts occur higher up the hierarchy and more specialised concepts occur lower down the hierarchy. This can simplify the retrieval of concepts whereby only a higher level concept needs to be specified in order to retrieve all relevant subtype concepts. For example, querying for patients who have lung cancer using the concept **363358000|Malignant tumor of lung (disorder)|** and its subtypes will retrieve concepts such as **254626006|Adenocarcinoma of lung (disorder)|**, **254634000|Squamous cell carcinoma of lung (disorder)|** and **94391008|Secondary malignant neoplasm of lung (disorder)|**. Another benefit is the defining attributes, whereby concepts are defined with other concepts. For example, **371087003|Diabetic foot ulcer (disorder)|** is defined with three sets of attributes (i.e., **47429007|Associated with (attribute)|=73211009|Diabetes mellitus (disorder)|**, **116676008|Associated morphology (attribute)|=56208002|Ulcer (morphologic abnormality)|** and **363698007|Finding site (attribute)|=56459004|Foot structure (body structure)|**). These attributes can help to facilitate the retrieval of patients who have been diagnosed with diabetes (**73211009|Diabetes mellitus (disorder)|**), patients who have problems with their feet (**56459004|Foot structure (body structure)|**), or patients who suffer from ulcers (**56208002|Ulcer (morphologic abnormality)|**).

A literature review of SNOMED CT use[1] revealed there are few studies that use SNOMED CT to "retrieve or analyse patient data." Even fewer are the studies that used the semantic relationships in SNOMED CT. Three studies used SNOMED CT descriptions to search a clinical dataset,[2, 3, 4] three studies used SNOMED CT descriptions to annotate free text summaries and reports[5, 6] and another used SNOMED CT descriptions to train natural language processing tools.[7] The papers that did touch on the semantic relationships of SNOMED CT reported that concept searches performed better than keyword searches.[8]

While using a description logic classifier is the ideal method of testing whether two expressions are related to each other,[9] the reality is that at the time of this study, there are no known tools that have been tailored specifically for SNOMED CT that are widely available and easy to use. In addition, most technical analysts in health care organisations are familiar with relational databases and do not have the expertise of how to use description logic classifiers. Therefore, while description logic tools are being developed for production applications, an interim solution is to use structural subsumption. However, computing structural subsumption can be slow. On a conservative estimate of 0.01 seconds it takes to compare two expressions for equivalency using the long normal form method,[10] it will take over a minute to test each expression in a dataset that contains 7,000 unique expressions. The long normal form is a view of an expression that is generated by applying a set of logical transformation rules[11] that decomposes an expression into its primitive defining supertypes.[10]

This estimate was based on how long it takes to execute a structured query language (SQL) statement in a transitive closure table for a single pair of concepts (e.g., **SELECT * FROM SCT_TransitiveClosure WHERE SuperTypeId=X and SubTypeId=Y**, where **X** and **Y** refer to SNOMED CT concept identifiers) in MySQL, a database management program. Using the long normal form method will generally require at least four comparisons (i.e., the focus concept and three contextual qualifiers: finding/procedure, temporal and subject relationship) depending on what types of expressions are being compared (i.e., pre-coordination vs post-coordination). If it takes 0.05 seconds to compare each pair of expressions, that would be well over five minutes to compare a dataset that has 7,000 expressions. There also have not been any studies on the optimal method of storing post-coordinated expressions. While the SNOMED CT Technical Implementation Guide (TIG) published by the International Health Terminology Standards Development Organisation (IHTSDO) suggests four methods, no studies have been published on which method is ideal.

There are three objectives of this study. First, to review the suggested methods of storing post-coordinated expressions. Second, to develop an optimised structural subsumption retrieval method. Third, to explore the use of SNOMED CT queries to identify chronic diseases[12] in an anonymised primary care dataset. The queries are intended to explore how SNOMED CT expressions that can be retrieved as opposed to identifying patient counts. While the results of the queries are important, it is also important to note that the retrieval method developed is also a type of result. The study adds to the current state of knowledge by defining a data storage schema for storing post-coordinated expressions that is suited for structural subsumption and a method for improving the efficiency of executing structural subsumption queries while maintaining its effectiveness.

The rest of the paper is organised into the materials used, the retrieval method that was developed, the results of applying the retrieval method to the encounter diagnoses from a primary care dataset that was encoded with SNOMED CT, a discussion of the retrieval method and results. The appendix also contains a section that reviews the fundamentals of SNOMED CT queries in **Section 12.7.1** of **Appendix F**.

8.2 Materials

There were four main materials used in this study. First, the relevant SNOMED CT tables. Second, an anonymised primary care dataset that was encoded with SNOMED CT. Third, the data storage methods suggested by the IHTSDO in the TIG. Fourth, chronic diseases listed on the Public Health Agency of Canada's website.

8.2.1 SNOMED CT

The pre-requisite for computing structural subsumption are the concepts, relationships, canonical, transitive closure and Concept Model tables. The first three are distributed as part of the semi-annual release of SNOMED CT. The transitive closure can be generated using the relationships table and the TIG includes instructions on how this can be done. The Concept Model table can be obtained from the IHTSDO or can be copied manually from the TIG.

8.2.2 *Primary Care Dataset*

An anonymised dataset was extracted from a primary care clinic. It consisted of 13,013 patient records, which included appointments, problems, encounter diagnoses, laboratory results, billing diagnostic codes, medications and physical examinations. The problems, encounter diagnoses and medications were originally in free text and were encoded in SNOMED CT. The process in which the dataset was encoded is found in **Chapter Seven**. Only the encounter diagnoses were used in the queries and it contained 4,595 pre-coordinated concepts and 2,761 post-coordinated expressions. Only the encounter diagnoses were used in this study as the encounter diagnoses in the dataset was 10 fold of the problems.

8.2.3 *Data Storage Methods*

The TIG suggests four methods of storing post-coordinated expressions (refer to **Table 8-1**). These four data storage methods were reviewed to determine the support for storing post-coordinated expressions, the level of difficulty of implementation and the support for data retrieval.

Table 8-1. Summary of analysis of data storage methods suggested by the SNOMED CT Technical Implementation Guide.

No	Data Storage Method	Description
1.	Parsable Text Representation	Each expression is stored in a single row as a string
2.	Unrestricted Relational Representation	Each expression is stored as multiple rows as attribute/value pairs
3.	Restricted Relational Representation	Each expression is stored in a single row with multiple attribute/value pairs
4.	XML Representations	Each expression is stored using XML

8.2.4 *Chronic Diseases*

Chronic diseases were selected from the Public Health Agency of Canada's website,[a] which included the category of chronic disease and examples of diseases in each category (refer to **Table 8-2**).

Table 8-2. Chronic diseases from the Public Health Agency of Canada's website.

No	Category	Disease Examples
1.	Cancer	Prostate cancer, lung cancer, thyroid cancer, bone cancer, breast cancer, cervical cancer, colorectal cancer, melanoma skin cancer, non melanoma skin cancer
2.	Heart (Cardiovascular) Disease	Ischaemic heart disease, cerebrovascular disease, peripheral vascular disease, heart failure, rheumatic heart disease, congenital heart disease
3.	Chronic Respiratory Disease	Asthma, chronic obstructive pulmonary disease, sleep apnoea
4.	Diabetes	Type 1, type 2, gestational
5.	Mental Illness	Mood disorder, major depression, bipolar disorder, schizophrenia, anxiety disorder, personality disorder, eating disorder, problem gambling, substance dependency
6.	Neurological Conditions	Paralysis, weakened muscles, low coordination, loss of sensation, seizures, confusion, pain, altered cognitive functions
7.	Musculoskeletal Diseases	Arthritis, osteoporosis

[a] http://www.phac-aspc.gc.ca/cd-mc/index-eng.php

8.3 Methods

Four methods were used in this study. First, structural subsumption was reviewed and an optimised structural subsumption method was developed. Second, the four methods for storing post-coordinated expressions were reviewed and two post-coordinated expressions were used to see if they could be stored in each of the storage methods. Third, the chronic disease categories and example of diseases from the Public Health Agency of Canada's website were manually encoded with SNOMED CT. Fourth, queries were executed against the primary care dataset to explore how SNOMED CT queries could be used and the efficiency of the retrieval method.

8.3.1 Retrieval Method

This retrieval method consists of three steps. First, structure the predicate expression. Second, identify candidate expressions. Third, compare the expressions. An overview of the steps is shown in **Figure 8-1**. The rest of this section describes each step in detail.

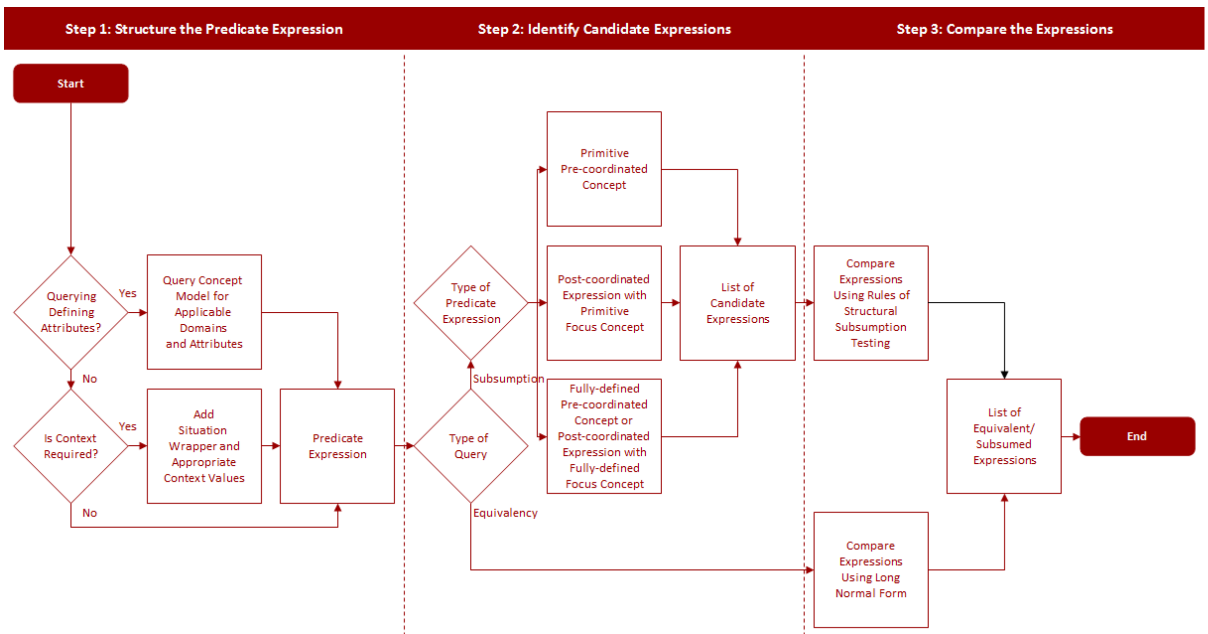


Figure 8-1. Overview of the retrieval method.

8.3.1.1 Step 1: Structure the Predicate Expression

The first step is to create the predicate expression that will be used in the query. There are basically three types of predicates that need to be structured: (1) defining attributes; (2) with context; and (3) no context or default context.

8.3.1.1.1 Defining Attributes

If the predicate is needed to retrieve concepts or expressions that use a particular concept as part of the defining attributes or post-coordinated expression, the first step is to determine the domain of the concept of interest. For example, if we are interested in finding concepts that use **73211009|Diabetes mellitus (disorder)|**, we have to determine which Concept Model constraints allow **73211009|Diabetes mellitus (disorder)|** to be defined as a range value. Assuming the Concept Model is stored in a relational database and the table has columns of **DomainId**, **AttributeId** and **RangeId**, a possibility would be to use the SQL statement as shown in **Figure 8-2**.

```
SELECT
    TC.DomainId,
    TC.AttributeId
FROM
    SCT_TransitiveClosure TC,
    SCT_ConceptModel CM
WHERE
    CM.RangeId=TC.SuperTypeId AND
    TC.SubTypeId=73211009
```

Figure 8-2. SQL Statement to retrieve the domain(s) and attribute(s) from a range value.

The SQL statement will generate 11 results (refer to **Table 8-3**). The fully-specified name was added for clarity.

Table 8-3. Results of retrieving the domain and attributes.

No	Domain	Attribute
1.	404684003|Clinical finding (finding)|	255234002|After (attribute)|
2.	404684003|Clinical finding (finding)|	47429007|Associated with (attribute)|
3.	404684003|Clinical finding (finding)|	246075003|Causative agent (attribute)|
4.	404684003|Clinical finding (finding)|	42752001|Due to (attribute)|
5.	404684003|Clinical finding (finding)|	363705008|Has definitional manifestation (attribute)|
6.	272379006|Event (event)|	255234002|After (attribute)|
7.	272379006|Event (event)|	47429007|Associated with (attribute)|
8.	272379006|Event (event)|	246075003|Causative agent (attribute)|
9.	272379006|Event (event)|	42752001|Due to (attribute)|
10.	71388002|Procedure (procedure)|	363702006|Has focus (attribute)|
11.	243796009|Situation with explicit context (situation)|	246090004|Associated finding (attribute)|

As some of the Concept Model attributes have a supertype/subtype relationship (refer to **Figure 8-3**) it is possible to remove the subtype attributes as they are redundant, which will result in five predicate expressions (refer to **Table 8-4**). Essentially what we have created are sets of predicate expressions that can subsume candidate concepts with the related defining attributes.

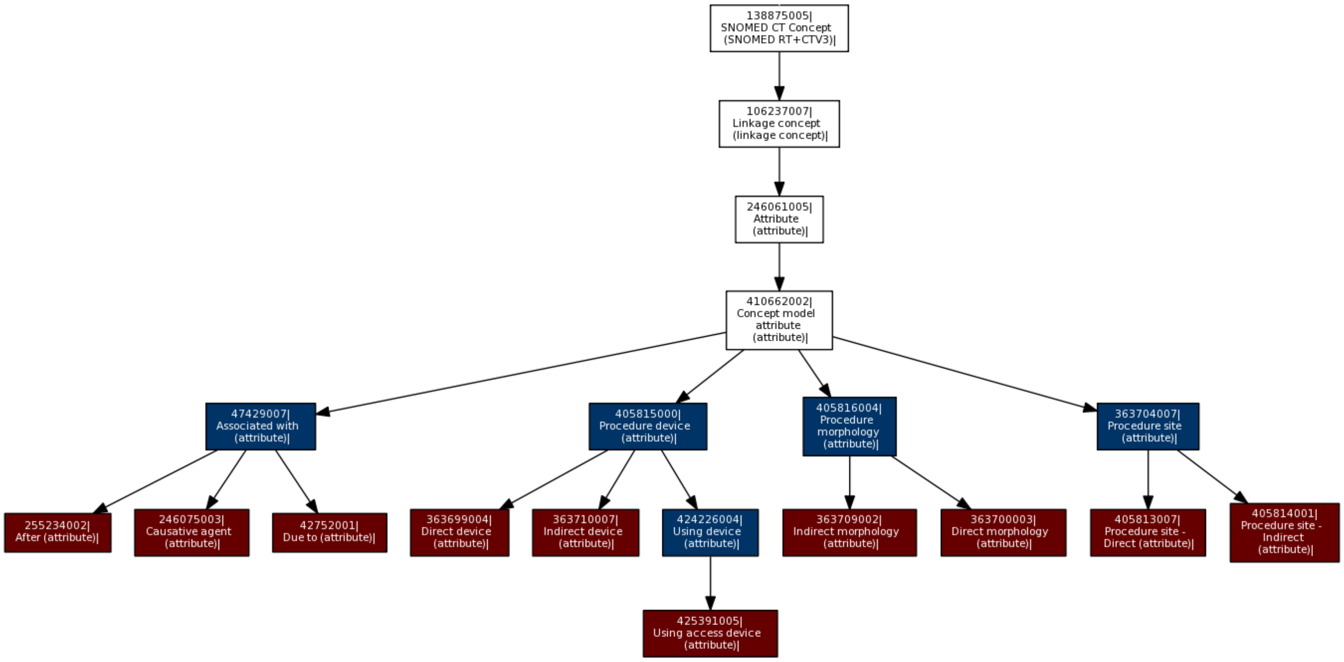


Figure 8-3. Concept Model attributes (blue) that have subtype concepts.

Table 8-4. Predicate expressions for retrieving concepts related to "73211009|Diabetes mellitus (disorder)|".

No	Predicate Expressions
1.	404684003|Clinical finding (finding)|: 47429007|Associated with (attribute)|= 73211009|Diabetes mellitus (disorder)|
2.	404684003|Clinical finding (finding)|: 363705008|Has definitional manifestation (attribute)|= 73211009|Diabetes mellitus (disorder)|
3.	272379006|Event (event)|: 47429007|Associated with (attribute)|= 73211009|Diabetes mellitus (disorder)|
4.	71388002|Procedure (procedure)|: 363702006|Has focus (attribute)|= 73211009|Diabetes mellitus (disorder)|
5.	243796009|Situation with explicit context (situation)|: 246090004|Associated finding (attribute)|= 73211009|Diabetes mellitus (disorder)|

Even though it is possible to query the relationships table to determine which concepts have a defining attribute that refers to **73211009|Diabetes mellitus (disorder)|** (e.g., **SELECT * FROM SCT_Relationships WHERE CharacteristicType=0 AND RelationshipType=116680003 AND ConceptId=73211009**), if a refinement of **73211009|Diabetes mellitus (disorder)|** was used, those concepts would have been missed. For example, when reviewing the type of pre-coordinated concepts that have been defined with **73211009|Diabetes mellitus (disorder)|** or one of its subtype, we find there are four concepts (refer to **Table 8-5**). If the above SQL statement were used, 39 concepts would have been omitted.

Table 8-5. Concepts that are defined using "73211009|Diabetes mellitus (disorder)|" or one of its subtypes.

No	Predicate Expressions	Total
1.	73211009|Diabetes mellitus (disorder)|	106
2.	46635009|Diabetes mellitus type 1 (disorder)|	18
3.	44054006|Diabetes mellitus type 2 (disorder)|	18
4.	75524006|Malnutrition related diabetes mellitus (disorder)|	3

It is possible to incorporate a transitive closure table to handle the subtypes (e.g., **SELECT R.ConceptId1 FROM SCT_Relationships R, SCT_TransitiveClosure TC WHERE R.RelationshipType=47429007 AND R.ConceptId2=TC.SubTypeId AND TC.SuperTypeId=73211009**) but that assumes that the candidate expressions are either pre-coordinated concepts or post-coordinated expressions in the close-to-user form.

8.3.1.1.2 With Context

When contexts are involved, there is a need to create a post-coordinated predicate expression. In order to specify context, the **243796009|Situation with explicit context (situation)|** wrapper should be used. The **246090004|Associated finding (attribute)|** and **363589002|Associated procedure (attribute)|** attributes will be linked to the appropriate **404684003|Clinical finding (finding)|** and **71388002|Procedure (procedure)|** concepts while the rest of the context values may be manipulated for the desired result (refer to **Figure 8-5**). Only contexts that are needed and that are not the default context values should be added to the predicate.

In order to retrieve all types of expressions regardless of the three types of context (i.e., finding/procedure, temporal and subject relationship), the topmost supertype concept in the range may be specified (refer to **Table 8-6** and **Figure 8-4**).

Table 8-6. SNOMED CT contexts, range and default values.

No	Concept Model Attribute	Supertype Value	Default Value
1.	408729009|Finding context (attribute)|	410514004|Finding context value (qualifier value)|	410515003|Known present (qualifier value)|
2.	408730004|Procedure context (attribute)|	288532009|Context values for actions (qualifier value)|	385658003|Done (qualifier value)|
3.	408731000|Temporal context (attribute)|	410510008|Temporal context value (qualifier value)|	410512000|Current or specified time (qualifier value)|
4.	408732007|Subject relationship context (attribute)|	125676002|Person (person)|	410604004|Subject of record (person)|

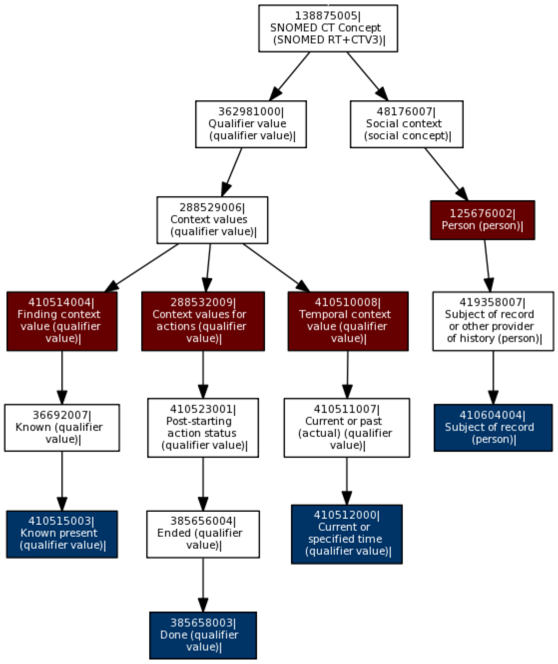


Figure 8-4. Plotting the range and default value of SNOMED CT contexts (red refers to the supertype value; blue refers to the default value).

The predicate expressions are shown in **Figure 8-5**.

404684003|Clinical finding (finding)|	**71388002|Procedure (procedure)|**
243796009|Situation with explicit context (situation)|: {246090004|Associated finding (attribute)|= 404684003|Clinical finding (finding)|, 408729009|Finding context (attribute)|= 410514004|Finding context value (qualifier value)|, 408731000|Temporal context (attribute)|= 410510008|Temporal context value (qualifier value)|, 408732007|Subject relationship context (attribute)|= 125676002|Person (person)|}	243796009|Situation with explicit context (situation)|: {363589002|Associated procedure (attribute)|= 71388002|Procedure (procedure)|, 408730004|Procedure context (attribute)|= 385658003|Done (qualifier value)|, 408731000|Temporal context (attribute)|= 410510008|Temporal context value (qualifier value)|, 408732007|Subject relationship context (attribute)|= 125676002|Person (person)|}

Figure 8-5. Predicate expression to retrieve all expressions for "404684003|Clinical finding (finding)|" and "71388002|Procedure (procedure)|" regardless of context.

8.3.1.1.3 No Context or Default Context

If no contexts are required in the predicate, the pre-coordinated concept or post-coordinated expression can be used as is.

8.3.1.2 Step 2: Retrieve Potential Candidates

This step basically identifies potential candidate expressions that have a possibility of being subsumed by the predicate. There are two main ways of doing this. First, include only candidates that have the same or refined context. For example, if the predicate is **266894000|Family history: Cardiovascular disease (situation)|**, the three contextual values need to be identified. When the predicate is normalised, we can determine that **408729009|Finding context (attribute)|=410515003|Known present (qualifier value)|**, **408731000|Temporal context (attribute)|=410589000|All times past (qualifier value)|** and **408732007|Subject relationship context**

(attribute)|=444148008|Person in family of subject (person)|. For each of the three contextual values, we need to retrieve the subtypes. **410515003|Known present (qualifier value)|** has three subtype concepts: **410605003|Confirmed present (qualifier value)|**, **410591008|Definitely present (qualifier value)|**, **410592001|Probably present (qualifier value)|**; **410589000|All times past (qualifier value)|** does not have any subtypes; while **444148008|Person in family of subject (person)|** has 20 subtype concepts. Therefore when retrieving potential candidates, it must be ensured that the candidate matches one of the four **408729009|Finding context (attribute)|** range values, it matches exactly to the single **408731000|Temporal context (attribute)|** range value, and it matches one of the 21 **408732007|Subject relationship context (attribute)|** range values.

The second part is to retrieve candidates that have a proximal primitive that is equivalent or a subtype of the proximal primitive of the predicate and also includes the concept model attributes in the long normal form of the predicate. To recap, a concept that is fully defined means that it can be sufficiently differentiated from its parent and sibling concepts via its defining relationships. On the other hand, the defining attributes of a primitive concept do not fully express its meaning. The proximal primitive of a primitive concept refers to itself, while the proximal primitive(s) of a fully defined concept refers to the closest primitive supertype concepts.

Referring back to the long normal form of the predicate **266894000|Family history: Cardiovascular disease (situation)|**, the focus concept refers to **49601007|Disorder of cardiovascular system (disorder)|**, which has a proximal primitive of **64572001|Disease (disorder)|** and has one defining attribute **363698007|Finding site (attribute)|**. Since **49601007|Disorder of cardiovascular system (disorder)|** is a fully-defined concept, we have to retrieve all subtypes of its proximal primitive **64572001|Disease (disorder)|**. Even by filtering potential candidate expression by the proximal primitive, we may still end up with a fairly large number of potential candidate expressions. For example, the concept **49601007|Disorder of cardiovascular system (disorder)|**, has a proximal primitive concept of **64572001|Disease (disorder)|**, which is near the very top of the **404684003|Clinical finding (finding)|** hierarchy. Looking at the entire content of SNOMED CT as candidates, **64572001|Disease (disorder)|** has less subtype concepts (65,396) than **404684003|Clinical finding (finding)|** (98,568), therefore by checking to ensure that candidate concepts or proximal primitives of **64572001|Disease (disorder)|**, the number of potential concepts has been reduced by over 33,000 concepts. To further reduce the number of potential concepts, we can look at the defining attributes that are contained in the long normal form of the predicate expression. In this case, **49601007|Disorder of cardiovascular system (disorder)|** only has one defining attribute **363698007|Finding site (attribute)|**. By ensuring that the candidate is a subtype of **64572001|Disease (disorder)|** and has **363698007|Finding site (attribute)|** in its long normal form, we have now reduced the number of potential candidates to 47,592. That might seem a lot, but we have more than halved the original number of potential concepts. With predicate focus concepts that are primitive, the number of candidates can usually be further reduced as only subtypes are needed and not a different proximal primitive (since the proximal primitive of primitive concepts refers to itself) which may lie higher up in the hierarchy.

This can be shown in the SQL statement in **Figure 8-6**. These potential candidates are not necessarily going to be subsumed by the predicate as the target of the **363698007|Finding site (attribute)|** may not be subsumed by the predicate, but it narrows down the number of potential candidates to be tested. It is possible to query for the three contextual values against the long normal form, but querying against a full-text index is generally slower than an exact match to an index.

```
SELECT
    *
FROM
    ExpressionsTable ET,
    SCT_TransitiveClosure TC
WHERE
    ET.FindingContext IN (410515003, 410605003, 410591008, 410592001) AND
    ET.TemporalContext IN (410589000) AND
    ET.SubjectRelationshipContext IN (444191003, 444303004, 444192005, 444194006, 444295003,
            444193000, 444242001, 444053001, 444243006, 444244000, 444301002, 444294004, 444292000,
            444293005, 444148008, 444302009, 444304005, 444241008, 444052006, 444055008, 444054007)
            AND
    ET.ProximalPrimitive=TC.SubTypeId AND
    TC.SuperTypeId=64572001 AND
    MATCH(ET.LongNormalForm) AGAINST (363698007)
```

Figure 8-6. SQL Statement to retrieve potential candidates.

The most common concepts that are proximal primitives of other fully-defined concepts in the January 31, 2012 release version of SNOMED CT and the number of subtype concepts they have are shown in **Table 8-7**.

Table 8-7. Most common proximal primitives from the "404684003|Clinical finding (finding)|" hierarchy.

No	Proximal Primitive	Used as Proximal Primitive	Number of Subtype Concepts
1.	64572001|Disease (disorder)|	28,446	65,396
2.	404684003|Clinical finding (finding)|	5,500	98,568
3.	281647001|Adverse reaction (disorder)|	1,626	1,838
4.	118228005|Functional finding (finding)|	1,577	8,094
5.	106190000|Allergic state (disorder)|	1,520	1,619
6.	248536006|Finding of functional performance and activity (finding)|	1,407	1,555
7.	75478009|Poisoning (disorder)|	1,192	3,730
8.	72431002|Accidental poisoning (disorder)|	703	761
9.	298325004|Finding of movement (finding)|	570	843
10.	55680006|Drug overdose (disorder)|	536	1,965
11.	410061008|Intentional poisoning (disorder)|	519	523
12.	269736006|Poisoning of undetermined intent (disorder)|	517	518
13.	129456006|Specific enzyme deficiency (disorder)|	454	616
14.	59369008|Accidental drug overdose (disorder)|	453	454
15.	59274003|Intentional drug overdose (disorder)|	452	453
16.	371341003|Drug overdose of undetermined intent (disorder)|	451	452
17.	127326005|Non-human disorder (disorder)|	348	492
18.	106146005|Reflex finding (finding)|	335	371

No	Proximal Primitive	Used as Proximal Primitive	Number of Subtype Concepts
19.	118231006|Communication finding (finding)|	332	469
20.	22253000|Pain (finding)|	290	788

In subsumption testing, there are basically three different types of predicate expressions: (1) predicate is a pre-coordinated concept that is primitive; (2) predicate is a post-coordinated expression with a primitive focus concept; and (3) predicate is pre-coordinated concept that is fully-defined or post-coordinated expression with a fully-defined focus concept. Different algorithms should be used for the different types of predicate to improve the efficiency of retrieval.

8.3.1.2.1 Predicate is a Pre-coordinated Concept that is Primitive

In this scenario, only candidates that have a proximal primitive concept that is a subtype of the predicate concept should be retrieved. Since the predicate is a primitive pre-coordinated concept, the proximal primitive concept refers to itself and it is not possible to create an expression that will be subsumed by predicate unless the focus concept of the candidate is a subtype of the predicate. For example, the concept **254628007|Carcinoma of lung parenchyma (disorder)|** is a primitive concept. It has the defining attributes of **116676008|Associated morphology (attribute)|=86049000|Malignant neoplasm, primary (morphologic abnormality)|** and **363698007|Finding site (attribute)|=113255004|Structure of parenchyma of lung (body structure)|**. Since **254628007|Carcinoma of lung parenchyma (disorder)|** is a primitive concept, its proximal primitive refers to itself. Creating a post-coordinated expression with a focus concept of **64572001|Disease (disorder)|** with the same refinements as the defining attributes of **254628007|Carcinoma of lung parenchyma (disorder)|** will not make this concept subsume it even though the defining attributes can be subsumed as the focus concept **64572001|Disease (disorder)|** is not a subtype of **254628007|Carcinoma of lung parenchyma (disorder)|**. It is also unnecessary to further test if a candidate post-coordinated expression with a focus concept that is a subtype of the predicate concept because any additional post-coordination in the candidate will not affect subsumption as the defining attributes in the candidate are the same or specialisation of the defining attributes of the predicate concept. Therefore if a candidate **254628007|Carcinoma of lung parenchyma (disorder)|** was qualified with **246456000|Episodicity (attribute)|= 272131007|Old episode (qualifier value)|**, it will not make any difference and be still subsumed by the predicate.

8.3.1.2.2 Predicate is a Post-coordinated Expression with a Primitive Focus Concept

Even though the predicate is a post-coordinated expression, it is still necessary to retrieve pre-coordinated candidate concepts because it is possible that the additional refinements or qualifications that were added to the predicate expression are subsumed by a pre-coordinated concept. For example, "diabetes mellitus during infancy" could be encoded as **73211009|Diabetes mellitus (disorder)|:246454002|Occurrence (attribute)|=3658006|Infancy (qualifier value)|**. This post-coordinated expression will subsume the concept **49817004|Neonatal diabetes mellitus**

(disorder)| as one of the defining attributes of **49817004|Neonatal diabetes mellitus (disorder)|** is **246454002|Occurrence (attribute)|= 255407002|Neonatal (qualifier value)|**. It is not a given that a pre-coordinated candidate concept will be subsumed by a post-coordinated predicate concept even though the candidate concept is a subtype of a focus concept of the predicate expression. For example, "diabetes mellitus during childhood" could be encoded as **73211009|Diabetes mellitus (disorder)|:246454002|Occurrence (attribute)|=255398004|Childhood (qualifier value)|** but this expression will not subsume because **255407002|Neonatal (qualifier value)|** is not a subtype of **255398004|Childhood (qualifier value)|**. The concepts just discussed are shown in **Figure 8-7**.

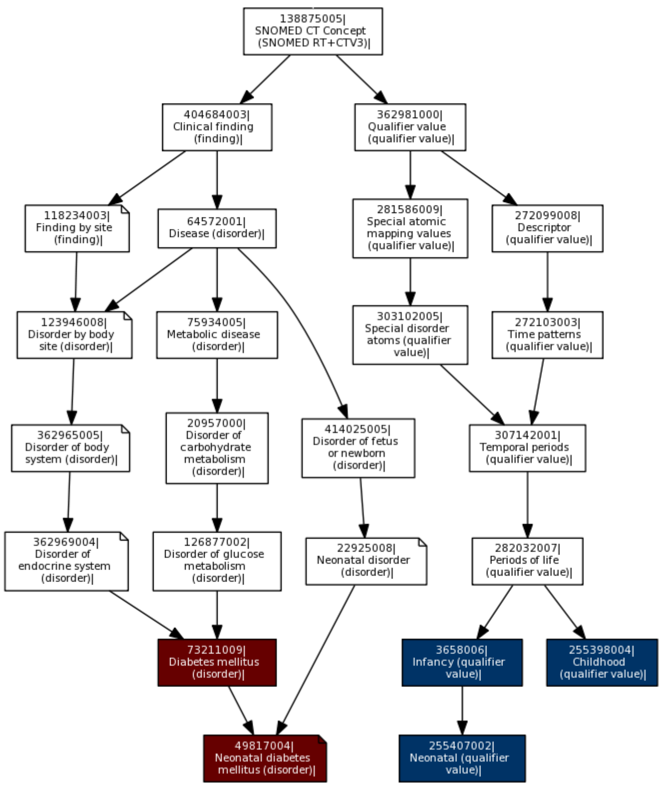


Figure 8-7. Location of concepts "73211009|Diabetes mellitus (disorder)|", "49817004|Neonatal diabetes mellitus (disorder)|", "3658006|Infancy (qualifier value)|", "255407002|Neonatal (qualifier value)|" and "255398004|Childhood (qualifier value)|" in the hierarchy.

8.3.1.2.3 Predicate is a Pre-coordinated Concept that is Fully-defined or Post-coordinated Expression with a Fully-defined Focus Concept

If the predicate is a fully-defined concept, it is important not to exclude any candidates with a focus concept that is a supertype of the focus concept of the predicate. For example, the concept **21522001|Abdominal pain (finding)|** is fully-defined and has a proximal primitive concept of **22253000|Pain (finding)|**. Any candidate expression with a focus concept of **22253000|Pain (finding)|** when post-coordinated with the appropriate attributes can be subsumed by **21522001|Abdominal pain (finding)|**. When encoding a phrase such as "pain in right lower quadrant of abdomen," it may be represented with an expression such as **22253000|Pain (finding)|:363698007|Finding site (attribute)|=48544008|Structure of right lower quadrant of abdomen (body**

structure)| as opposed to **21522001|Abdominal pain (finding)|:363698007|Finding site (attribute)|=48544008|Structure of right lower quadrant of abdomen (body structure)|**. When an exact match is not found for a phrase such as "pain in right lower quadrant of abdomen," it is likely that the phrase will be broken into smaller chunks and searches will be performed on these smaller chunks. In this example, "pain" will be matched to **22253000|Pain (finding)|** while "right lower quadrant of abdomen" will be matched to **48544008|Structure of right lower quadrant of abdomen (body structure)|** and the two concepts be joined by **363698007|Finding site (attribute)|**. Depending on how the expression is constructed, it may be less likely that **21522001|Abdominal pain (finding)|** will be the focus concept followed by a refinement of the finding site because in this phrase, "pain" is the first word while "abdomen" is the last word and it is unlikely that an encoding algorithm will ignore the words in between and encode the first and last words with a single concept. In this case, the post-coordinated expression **22253000|Pain (finding)|:363698007|Finding site (attribute)|=48544008|Structure of right lower quadrant of abdomen (body structure)|** is subsumed by **21522001|Abdominal pain (finding)|**.

In the second example, there are concepts for primary and secondary jaw cancer (**93845000|Primary malignant neoplasm of jaw (disorder)|** and **94356000|Secondary malignant neoplasm of jaw (disorder)|**) but there does not exist a pre-coordinate concept for just "jaw cancer." A refinement post-coordination is less likely to be done from a refinement of concepts such as **126634001|Neoplasm of jaw (disorder)|** or **363501002|Malignant tumor of face (disorder)|** although they are one of the parent concepts of the primary and secondary jaw concepts and the post-coordination could be technically correct. For example, in the former concept, it could be refined with **116676008|Associated morphology (attribute)|=367651003|malignant neoplasm of primary, secondary, or uncertain origin (morphologic abnormality)|** while the former could be refined with **363698007|Finding site (attribute)|=661005|Jaw region structure (body structure)|**. However, it is more likely that "jaw" will get matched to **661005|Jaw region structure (body structure)|** and "cancer" will be matched to **363346000|Malignant neoplastic disease (disorder)|** and the two concepts will be linked using **363698007|Finding site (attribute)|** thereby forming the post-coordinated expression **363346000|Malignant neoplastic disease (disorder)|:363698007|Finding site (attribute)|=661005|Jaw region structure (body structure)|**. Therefore, when querying for all types of facial cancer (**363501002|Malignant tumor of face (disorder)|**), we cannot discount retrieving post-coordinated expression that have a focus concept of **363346000|Malignant neoplastic disease (disorder)|** even though it is a supertype concept of **363501002|Malignant tumor of face (disorder)|**.

8.3.1.3 Step 3: Compare Potential Candidates

The last step is to compare each potential candidate against the predicate. The rules for comparing two expressions for equivalency and subsumption have been enumerated in the TIG and therefore are not repeated here. By this stage, most unrelated candidates should have been omitted. While querying for **73211009|Diabetes mellitus (disorder)|** and its subtypes such as **46635009|Diabetes mellitus type 1 (disorder)|** and **44054006|Diabetes mellitus**

type 2 (disorder)|, there may be cases where pregnancy-related diabetes concepts such as **359939009|Maternal diabetes mellitus (disorder)|** and **199223000|Diabetes mellitus during pregnancy, childbirth and the puerperium (disorder)|** are not wanted (refer to Table 8-8). Therefore queries must be designed in such a way as to include and exclude expressions. Multiple predicate expressions consisting of inclusion and exclusion criteria may be necessary in order to retrieve a relevant set of concepts. When evaluating the inclusion and exclusion criteria, the exclusion criteria should have priority over the inclusion criteria.

Table 8-8. Inclusion and exclusion criteria.

No	Predicate	Equivalency	Include / Exclude
1.	73211009|Diabetes mellitus (disorder)|	Equivalency and subsumption	Include
2.	359939009|Maternal diabetes mellitus (disorder)|	Equivalency and subsumption	Exclude
3.	199223000|Diabetes mellitus during pregnancy, childbirth and the puerperium (disorder)|	Equivalency and subsumption	Exclude

8.3.2 Method for Evaluating Data Storage Methods

In order to determine which data storage method is best, we must first enumerate what we want to do with the expressions stored within the table. A review of the types of data retrieval methods previously listed indicates that there are two main types of retrieval: (1) equivalency and (2) subsumption. As mentioned earlier, the querying for defining attributes may also be defined as a type of subsumption query. In order to test expressions for equivalency, we need to store the long normal form. In order to test post-coordinated expressions for subsumption, we need the long normal form for the predicate expression and the short normal form for the candidate expression. However, as the long normal form can also be used for the candidate expression and the short normal form is a subset of the long normal form, it is permissible to use the long normal form for both the predicate and candidate expressions. The main reason for storing the pre-calculated long normal form rather than calculating it in real-time is to improve the efficiency of the query.

The long normal form of concepts can be broken up into four main parts. First, the **246090004|Associated finding (attribute)|** or **363589002|Associated procedure (attribute)|**. The value of these attributes contains the main clinical statement while the next three refer to different types of context. Second, the **408729009|Finding context (attribute)|** or **408730004|Procedure context (attribute)|**. The value of these attributes indicate the context of the **246090004|Associated finding (attribute)|** or **363589002|Associated procedure (attribute)|**. Third, the **408731000|Temporal context (attribute)|**. The value of this attribute indicates the timeframe of the clinical statement. Fourth, the **408732007|Subject relationship context (attribute)|**. The value of this attribute indicates who the clinical statement applies to. As context can only be applied to concepts from the **404684003|Clinical finding (finding)|**, **71388002|Procedure (procedure)|** and **243796009|Situation with explicit context (situation)|** hierarchies, contextual information for expressions from other hierarchies will not need to be stored.

When attempting to retrieve clinical statements that match a specified criterion, these three contexts should always be evaluated. For example, when searching for patients who currently have lung cancer, we want to ensure that the finding context refers to the present (**410515003|Known present (qualifier value)|**) rather than the absent (**410516002|Known absent (qualifier value)|**), that the temporal context refers to the present (or current) time (**410512000|Current or specified time (qualifier value)|**) rather than the past (**410513005|In the past (qualifier value)|**), and that it is the patient (**410604004|Subject of record (person)|**) who has the lung cancer and not a family member (**444148008|Person in family of subject (person)|**).

While it is possible to derive the context by parsing each long normal form, doing so would be less efficient. Ideally these contexts should be stored as discrete data elements. It should also be noted that the range of values that these context contain cannot be further refined or qualified thereby making it easier to store. As for the main non-contextual portion of the clinical statement, the entire statement will need to be parsed as there are many ways of representing the same clinical condition. However, we ideally would like to retrieve a set of potential expressions from all candidate expressions. For example, if we want to retrieve expressions that are subsumed by **363358000|Malignant tumor of lung (disorder)|** from a subset that contains 3,000 **404684003|Clinical finding (finding)|**, 1,000 **71388002|Procedure (procedure)|** and **500 272379006|Event (event)|**, we should ideally exclude the 1,500 concepts that are **71388002|Procedure (procedure)|** and **272379006|Event (event)|**. In order to do so efficiently, we can conduct tests to determine if the focus concept is a supertype or subtype of **363358000|Malignant tumor of lung (disorder)|**. Any focus concept that is a subtype of **363358000|Malignant tumor of lung (disorder)|** and has the appropriate contexts should be relevant. However, as mentioned earlier, the focus concepts that are supertypes of **363358000|Malignant tumor of lung (disorder)|** should also be retrieved as they potentially could be subsumed by **363358000|Malignant tumor of lung (disorder)|**. So rather than retrieving all 3,000 expressions whose focus concept is from the **404684003|Clinical finding (finding)|** hierarchy, we can start to filter the relevant supertype and subtype concepts. All subtype concepts of **363358000|Malignant tumor of lung (disorder)|** will be subsumed while some supertype concepts may be subsumed.

If the predicate focus concept is its primitive concept, then there is no need to retrieve the supertype concepts because it is not possible for a candidate expression whose focus concept is a supertype of a primitive predicate expression to be subsumed by the predicate concept. The need to retrieve the supertype concepts also depends on how the subset was crafted. If the subset was crafted with rules that govern that pre-coordinated concepts should be used first and that the closest concept should be used instead of post-coordination, then these rules may apply differently.

8.4 Results

Two types of results are presented here. First, the results of the analysis of the data storage methods. Second, the results of encoding the chronic disease categories and examples. Third, the results of querying the primary care dataset for the seven chronic diseases.

8.4.1 Results of Analysis of Data Storage Methods

A summary of the results of the analysis of the data storage methods suggested by the TIG is shown in **Table 8-9**.

Table 8-9. Summary of analysis of data storage methods suggested by the SNOMED CT Technical Implementation Guide.

No	Data Storage Method	Description	Support for Post-coordination	Implementation Difficulty	Support for Data Retrieval	Recommendation
1.	Parsable Text Representation	Each expression is stored in a single row as a string	Unlimited	Easy	Less efficient	Recommended if only pre-coordination
2.	Unrestricted Relational Representation	Each expression is stored as multiple rows as attribute/value pairs	Unlimited	Complex	More efficient	Best method of the four but still has limitations
3.	Restricted Relational Representation	Each expression is stored in a single row with multiple attribute/value pairs	Limited	Easy	More efficient	Not recommended as limited support for post-coordination
4.	XML Representations	Each expression is stored using XML	Unlimited	Moderate	Less efficient	Recommended only if implementers are familiar with XML and use an XML database

As the TIG does not provide the schemas for the tables the schema representations used here may not be exactly what the authors of the TIG had in mind. Without the details of the schemas, it is up to implementers to decide how they should be designed. Two expressions were used to test the data storage methods. The expressions represent the close-to-user form and long normal form of how "cancer of left lung" could be represented (refer to **Figure 8-8**).

<table>
<tr><td>Close-to-User Form</td><td>363358000|Malignant tumor of lung (disorder)|:{
363698007|Finding site (attribute)|=
44029006|Left lung structure (body structure)|}</td></tr>
<tr><td>Long Normal Form</td><td>243796009|Situation with explicit context (situation)|:
{246090004|Associated finding (attribute)|=(
64572001|Disease (disorder)|:
{116676008|Associated morphology (attribute)|=
367651003|Malignant neoplasm of primary, secondary, or uncertain origin (morphologic abnormality)|,
363698007|Finding site (attribute)|=(
39607008|Lung structure (body structure)|:
272741003|Laterality (attribute)|=
7771000|Left (qualifier value)|
)}
),
408729009|Finding context (attribute)|=
410515003|Known present (qualifier value)|,
408731000|Temporal context (attribute)|=
410512000|Current or specified time (qualifier value)|,
408732007|Subject relationship context (attribute)|=
410604004|Subject of record (person)|}</td></tr>
</table>

Figure 8-8. The close-to-user and long normal form of "cancer of left lung."

8.4.1.1 Parsable Text Representation

In the parsable text representation, the expressions are stored in a single data element, **Expression**, in a table (refer to **Table 8-10**). The advantage of this method is that any expression can be stored and requires no parsing in order to be stored. The disadvantage is that the entire expression needs to parsed before we can determine if it is useful for a particular query.

Table 8-10. Storing post-coordinated expressions using the parsable text representation.

No	Expression
1.	363358000:{363698007=44029006}
2.	243796009:{246090004=(64572001:{116676008=367651003,363698007=(39607008:272741003=7771000)}),408729009=410515003,408731000=410512000,408732007=410604004}

8.4.1.2 Unrestricted Relational Representation

The unrestricted relational representation is arguably the most complex of the four suggested representations as each expression requires multiple records in a table. While the TIG refers to the storing of "attribute-value pairs," there are few other details. Ideally there should be two tables, the first to record the human readable description of the expression and the second to record the expression itself. Each portion of an expression would require an internal identifier (**InternalId**). In order to accommodate nested expressions, there must be an identifier to indicate a parent/child relationship (**InternalParentId**). To accommodate defining attributes that can be grouped and ungrouped, one more identifier is needed (**InternalGroupId**). These three ids, **InternalId**, **InternalParentId** and **InternalGroupId** do not contain any SNOMED CT semantics but are required to restructure an expression that has been broken down into sections for storage in an unrestricted relational representation. The data elements are listed in **Table 8-11**.

Table 8-11. Data elements in the unrestricted relational representation.

No	Data Element	Description
1.	ExpressionId	External identifier of an expression.
2.	InternalId	Internal identifier of a portion of an expression.
3.	InternalParentId	Internal identifier of the parent portion of an expression.
4.	InternalGroupId	Internal identifier of the group of an expression. Ungrouped portions are assigned the value 0, while grouped are assigned a value of ≥1.
5.	Attribute	The attribute concept from the Concept Model.
6.	Value	The range concept in a refinement/qualification.

One way to understand the need for these three ids is to represent an expression using a multi-dimensional array. Each attribute-value pair must have an **InternalParentId**, **InternalGroupId** and **InternalId**. The first focus concept in the expression will always have an **InternalParentId** of **0**. Ungrouped attributes are assigned an **InternalGroupId** of **0** while grouped attributes are assigned an **InternalGroupId** starting from **101**. The number starts from **101** as opposed to **1** so as to avoid confusion with defining attributes. The **InternalGroupId** for each nesting level can restart in each level. A post-coordinated expression can be broken down to smaller portions by each attribute-value pair. Only the very first focus concept will not have an attribute.

In the example of "cancer of left lung," the focus concept is **363358000|Malignant tumor of lung (disorder)|**. Since it is the first focus concept, the **InternalParentId** is **0** (refer to Figure 8-9, line 3). The first focus concept is never grouped, therefore the **InternalGroupId** is **0** (line 5). Subsequent focus concepts may be grouped, depending on how an expression is constructed. The **InternalId** is an auto-number for each expression and since this is the first focus concept, the **InternalId** is **1** (line 7). As this is the first focus concept, there is no **Attribute** (line 9) while the **Value** is **363358000** (line 10). The next portion of the expression is the refinement of **363358000|Malignant tumor of lung (disorder)|**. Since the refinement, **{363698007|Finding site (attribute)|=44029006|Left lung structure (body structure)|}**, must be linked to the focus concept, the **InternalParentId** (line 14) must refer to the **InternalId** of **363358000|Malignant tumor of lung (disorder)|**, which is **1** (line 7). Since the refinement is grouped, i.e., **{}**, the **InternalGroupId** (line 16) is assigned a value of **101**. The **ExpressionId** of this refinement is therefore **2** (line 18), since it is the next sequential number after **1**.

```
Line                          Array                                Description

   1  Array
   2  (
   3      [0] => Array                                             ExpressionParentId
   4          (
   5              [0] => Array                                     ExpressionGroupId
   6                  (
   7                      [1] => Array                             ExpressionId
   8                          (
   9                              [Attribute] =>
  10                              [Value] => 363358000
  11                          )
  12                  )
  13          )
  14      [1] => Array                                             ExpressionParentId
  15          (
  16              [101] => Array                                   ExpressionGroupId
  17                  (
  18                      [2] => Array                             ExpressionId
  19                          (
  20                              [Attribute] => 363698007
  21                              [Value] => 44029006
  22                          )
  23                  )
  24          )
  25  )
```

Figure 8-9. Representing a post-coordinated expression in a multi-dimensional array.

The full representation of "cancer of left lung" is shown in Table 8-12 with **ExpressionId=1** while the long normal form is represented with **ExpressionId=2**.

Table 8-12. Unrestricted relational representation of SNOMED CT expressions.

Expression Id	Internal Parent Id	Internal Group Id	Internal Id	Attribute	Concept
1	0	0	1		363358000
1	1	101	2	363698007	44029006
2	0	0	1		243796009
2	1	101	2	246090004	64572001

Expression Id	Internal Parent Id	Internal Group Id	Internal Id	Attribute	Concept
2	1	101	6	408729009	410515003
2	1	101	7	408731000	410512000
2	1	101	8	408732007	410604004
2	2	101	3	116676008	367651003
2	2	101	4	363698007	39607008
2	4	0	5	272741003	7771000

8.4.1.3 Restricted Relational Representation

There are three options for storing the restricted relational representation.

8.4.1.3.1 Option 1

According to the TIG, page 643, "Each qualifier is represented by two ConceptId fields (one for the attribute and one for the value) and an optional field for Relationship group field. With this option the only restriction is the total number of qualifiers or modifiers that can be stored for each Concept." In this option, refinements are stored as attribute-value pairs in additional data elements within the same row. The limitation of this option is that nested refinements cannot be represented. The long normal form of "cancer of left lung" contains eight portions and although it is possible to have eight columns of attribute-value pairs, plus relationship group ids, it still will not be able to accommodate the nested refinements (refer to **Table 8-13**). Moreover, storing and retrieving expressions in this method is inefficient as the entire expression would have to be retrieved and parsed each time.

Table 8-13. Restricted relational representation, option 1.

No	Focus Concept	Attribute1	Value1	Relationship Group1	Attribute2	Value2	Relationship Group2
1.	363358000	363698007	44029006	101			
2.	243796009	24609000	6457200	0			

8.4.1.3.2 Option 2

According to the TIG, page 643, "Each qualifier is represented as a single ConceptId and carries the value of a qualifier attribute specific to that field. This restricts the usable qualifiers in the database schema." The main disadvantage of this storage method is that it would require numerous columns in order to accommodate all possible number of refinements or qualifiers, and even then, it still would not support nested post-coordinated expressions (refer to **Table 8-14**).

Table 8-14. Restricted relational representation, option 2.

No	Focus Concept	Associated Finding	Associated Procedure	Finding context	Procedure context	Temporal context	Subject Relationship Context
1.	363358000						
2.	243796009	64572001		410515003		410512000	410604004

8.4.1.3.3 Option 3

According to the TIG, page 643, "Similar to the above, but with different sets of qualifying attributes available according to the semantic type of the primary Concept in the statement. There are various ways of implementing this approach to ensure the appropriate interpretation is applied to each row of the table." In another words, this option uses a different set of attributes in the columns. For example, when dealing with a **404684003|Clinical finding (finding)|**, the columns will refer to **246090004|Associated finding (attribute)|** and **408729009|Finding context (attribute)|**, but when referring to a **71388002|Procedure (procedure)|**, the columns will refer to **363589002|Associated procedure (attribute)|** and **408730004|Procedure context (attribute)|**. This will require a separate table for each domain with each table having a variable number of columns but it still will not be able to accommodate nested post-coordinated expressions.

8.4.1.4 XML Representation

The XML representation is similar to the unrestricted relational representation except that each attribute-value pair is stored in XML (refer to **Figure 8-10**) while in the unrestricted relational representation each attribute-value pair is stored as a new row in a table. In addition to XML, this representation can also be represented in other formats such as the JavaScript Object Notation (JSON).

Close-to-User Form	`<code xsi:type="CD" code="363358000" displayName="Malignant tumor of lung" codeSystem="2.16.840.1.113883.6.96" codeSystemName="SNOMED CT"> <qualifier> <name xsi:type="CD" code="363698007" displayName="Finding site" codeSystem="2.16.840.1.113883.6.96" codeSystemName="SNOMED CT"/> <value xsi:type="CD" code="44029006" displayName="Left lung structure" codeSystem="2.16.840.1.113883.6.96" codeSystemName="SNOMED CT"/> </qualifier> </code>`
Long Normal Form	see below

```
<code xsi:type="CD" code="243796009" displayName="Situation with explicit context"
        codeSystem="2.16.840.1.113883.6.96" codeSystemName="SNOMED CT">
    <qualifier>
        <name xsi:type="CD" code="246090004" displayName="Associated finding"
                codeSystem="2.16.840.1.113883.6.96" codeSystemName="SNOMED CT"/>
        <value xsi:type="CD" code="64572001" displayName="Disease"
                codeSystem="2.16.840.1.113883.6.96" codeSystemName="SNOMED CT">
            <qualifier>
                <name xsi:type="CD" code="116676008" displayName="Associated
                        morphology" codeSystem="2.16.840.1.113883.6.96"
                        codeSystemName="SNOMED CT"/>
                <value xsi:type="CD" code="367651003" displayName="Malignant
                        Neoplasm (Morphology)" codeSystem="2.16.840.1.113883.6.96"
                        codeSystemName="SNOMED CT"/>
                <name xsi:type="CD" code="363698007" displayName="Finding site"
                        codeSystem="2.16.840.1.113883.6.96" codeSystemName="SNOMED
                        CT"/>
                <value xsi:type="CD" code="39607008" displayName="Lung structure"
                        codeSystem="2.16.840.1.113883.6.96" codeSystemName="SNOMED
                        CT">
                    <qualifier>
                        <name xsi:type="CD" code="272741003"
                                displayName="Laterality"
                                codeSystem="2.16.840.1.113883.6.96"
                                codeSystemName="SNOMED CT"/>
                        <value xsi:type="CD" code="7771000" displayName="Left"
                                codeSystem="2.16.840.1.113883.6.96"
                                codeSystemName="SNOMED CT"/>
                    </qualifier>
                </value>
            </qualifier>
        </value>
        <name xsi:type="CD" code="408729009" displayName="Finding context"
                codeSystem="2.16.840.1.113883.6.96" codeSystemName="SNOMED CT"/>
        <value xsi:type="CD" code="410515003" displayName="Known present"
                codeSystem="2.16.840.1.113883.6.96" codeSystemName="SNOMED CT"/>
        <name xsi:type="CD" code="408731000" displayName="Temporal context"
                codeSystem="2.16.840.1.113883.6.96" codeSystemName="SNOMED CT"/>
        <value xsi:type="CD" code="410512000" displayName="Current or specified"
                codeSystem="2.16.840.1.113883.6.96" codeSystemName="SNOMED CT"/>
        <name xsi:type="CD" code="408732007" displayName="Subject relationship
                context" codeSystem="2.16.840.1.113883.6.96"
                codeSystemName="SNOMED CT"/>
        <value xsi:type="CD" code="410604004" displayName="Subject of record"
                codeSystem="2.16.840.1.113883.6.96" codeSystemName="SNOMED CT"/>
    </qualifier>
</code>
```

Figure 8-10. Storing a post-coordinated expression in an XML representation.

8.4.1.5 A Proposed Data Storage Method: Enhanced Parsable Representation

While the above representations may be adequate to store post-coordinated expressions of variable complexities, they are not necessarily optimised for data retrieval. This section discusses how post-coordinated expressions can be stored to ensure efficient retrieval. As previously mentioned, there are three main types of context in SNOMED CT. The range of these contexts are **362981000|Qualifier value (qualifier value)|** and **125676002|Person (person)|** concepts (refer to **Table 8-15**), which cannot further be post-coordinated. Therefore to speed up data retrieval, these contexts can be stored as discrete values.

Table 8-15. Contexts used in post-coordinated expressions.

No	Attribute	Range	No of Subtype	Default Value	Subtypes of Default Value
1.	408729009|Finding context (attribute)|	410514004|Finding context value (qualifier value)|	18	410515003|Known present (qualifier value)|	▪ 410605003|Confirmed present (qualifier value)| ▪ 410591008|Definitely present (qualifier value)| ▪ 410515003|Known present (qualifier value)| ▪ 410592001|Probably present (qualifier value)|
2.	408730004|Procedure context (attribute)|	288532009|Context values for actions (qualifier value)|	50	385658003|Done (qualifier value)|	▪ 410542002|Attended (qualifier value)| ▪ 398166005|Performed (qualifier value)|
3.	408731000|Temporal context (attribute)|	410510008|Temporal context value (qualifier value)|	11	410512000|Current or specified time (qualifier value)|	▪ 410589000|All times past (qualifier value)| ▪ 15240007|Current (qualifier value)| ▪ 410584005|Current - time specified (qualifier value)| ▪ 410585006|Current - time unspecified (qualifier value)| ▪ 410587003|Past - time specified (qualifier value)| ▪ 410586007|Specified time (qualifier value)|
4.	408732007|Subject relationship context (attribute)|	125676002|Person (person)|	425	410604004|Subject of record (person)|	

As the long normal form is used to retrieve equivalent expressions, it should be stored pre-computed. The proximal primitive can be used as an indicator of the concept highest up in the hierarchy that could be relevant. Therefore the proximal primitive concept should be stored in a separate field. Multiple proximal primitives may be stored in this field. **Table 8-16** lists all the data fields of the enhanced parsable representation along with a description, rationale and example for each field.

Table 8-16. Enhanced parsable representation (* indicates human readable descriptions added for clarity).

No	Data Element	Description	Rationale	Example
1.	HumanReadable	Human readable description of the expression		Family history of lung cancer
2.	CloserToUserForm*	The close-to-user form of the expression		429011007|Family history of malignant neoplasm of lung (situation)|
3.	LongNormalForm	The long normal form of the expression.	To enable testing for equivalency quickly.	243796009:{246090004=(64572001:{116676008=367651003,363698007=39607008}),408729009=410515003,408731000=410511007,408732007=444148008}
4.	FocusConcept*	The first concept before any post-coordination	Facilitates query efficiency	363358000|Malignant tumor of lung (disorder)|
5.	Proximal Primitive*	Proximal primitive concepts of the focus concept	Facilitates query efficiency	64572001|Disease (disorder)|

No	Data Element	Description	Rationale	Example
6.	AssociatedFindingProcedure*	Associated finding or procedure	Facilitates query efficiency	363358000|Malignant tumor of lung (disorder)|
7.	FindingProcedureContext*	Finding or procedure context	Facilitates query efficiency	410515003|Known present (qualifier value)|
8.	TemporalContext*	Temporal context	Facilitates query efficiency	410511007|Current or past (actual) (qualifier value)|
9.	SubjectRelationshipContext*	Subject relationship context	Facilitates query efficiency	444148008|Person in family of subject (person)|
10.	ReleaseVersion	Release version used in the expression	Facilitate version control	20120131

In theory this table structure should be able to accommodate any expression. However, there are 23 expressions from the **243796009|Situation with explicit context (situation)|** hierarchy (refer to **Table 8-17**) that have multiple **246090004|Associated finding (attribute)|**, **363589002|Associated procedure (attribute)|**, **408729009|Finding context (attribute)|**, **408730004|Procedure context (attribute)|**, **408731000|Temporal context (attribute)|** and **408732007|Subject relationship context (attribute)|** values. The main reason there are multiple contextual values is because there are multiple procedures or findings. The multiple contextual values can refer to different target values. For example, **433807000|History of occlusion of cerebral artery without cerebral infarction (situation)|** has two different **246090004|Associated finding (attribute)|**; one is refers to a **408729009|Finding context (attribute)|** of **410516002|Known absent (qualifier value)|** while the other **410515003|Known present (qualifier value)|**. In another example, **161079000|Both parents smoke (situation)|** has two of the same **246090004|Associated finding (attribute)|**; however, one refers to a **408732007|Subject relationship context (attribute)|** of **444295003|Father of subject (person)|** while the other refers to **444301002|Mother of subject (person)|**. It is uncertain if these concepts with multiple contextual values can be modelled in a different way or what the implications are. These concepts have been identified as they cannot be accommodated by the enhanced parsable representation but the reason for the modelling is beyond the scope of this research.

Table 8-17. Concepts from the "243796009|Situation with explicit context (situation)|" hierarchy that have multiple contextual values.

No	Concept	246090004|Associated finding (attribute)|	363589002|Associated procedure (attribute)|	408729009|Finding context (attribute)|	408730004|Procedure context (attribute)|	408731000|Temporal context (attribute)|	408732007|Subject relationship context (attribute)|
1.	397680002|Absence of signs and symptoms of infection (situation)|	2		2		1	2
2.	442322007|Alexia and agraphia present (situation)|	2		2		2	2
3.	439138006|Both parents misuse drugs (situation)|	2		2		2	2
4.	161079000|Both parents smoke (situation)|	2		2		2	2

No	Concept	246090004|Associated finding (attribute)|	363589002|Associated procedure (attribute)|	408729009|Finding context (attribute)|	408730004|Procedure context (attribute)|	408731000|Temporal context (attribute)|	408732007|Subject relationship context (attribute)|
5.	183977004|Ear, nose and throat operation planned (situation)|		2		2	1	1
6.	371622005|Elevated blood pressure reading without diagnosis of hypertension (situation)|	2		2		2	2
7.	433491003|Family history of bilateral hip replacements (situation)|		2		2	2	2
8.	412784002|Family history: Enuresis (situation)|	2		2		2	2
9.	440706004|Full renal function recovered (situation)|	2		2		2	2
10.	161680001|History of - artificial limb (situation)|		2		2	2	2
11.	161667004|History of - heart valve recipient (situation)|		2		2	2	2
12.	308067002|History of - Stroke in last year (situation)|	2		2		2	2
13.	161663000|History of - tissue/organ recipient (situation)|		2		2	2	2
14.	428043001|History of arthroscopic procedure on shoulder (situation)|		2		2	2	2
15.	428265005|History of arthroscopy of elbow (situation)|		2		2	2	2
16.	429677001|History of arthroscopy of knee joint (situation)|		2		2	2	2
17.	433807000|History of occlusion of cerebral artery without cerebral infarction (situation)|	2		2		2	2
18.	390855002|On examination - diabetic maculopathy absent both eyes (situation)|	2		2		2	2
19.	390854003|On examination - diabetic maculopathy present both eyes (situation)|	2		2		2	2
20.	394885002|Parents do not smoke (situation)|	2		2		2	2
21.	253987006|Unilateral incomplete cleft lip and/or alveolus (situation)|	2		2		2	2
22.	204613000|Unilateral incomplete cleft palate with cleft lip (situation)|	2		2		2	2
23.	71101001|Unilateral modified radical dissection of neck (situation)|		2		2	2	2

8.4.2 Result of Encoding Chronic Conditions

The results of encoding the chronic disease are shown in **Table 8-18**. All seven categories and the 40 examples of diseases were encoded. One disease, non melanoma skin cancer was encoded using two concepts: **402815007|Squamous cell carcinoma (disorder)|** and **254701007|Basal cell carcinoma of skin (disorder)|**.

Table 8-18. Results of encoding the chronic diseases (disease categories in bold).

No	Chronic Disease Category and Example	SNOMED CT Encoding
1.	**Cancer**	363346000|Malignant neoplastic disease (disorder)|
2.	Prostate cancer	399068003|Malignant tumor of prostate (disorder)|
3.	Lung cancer	363358000|Malignant tumor of lung (disorder)|
4.	Thyroid cancer	363478007|Malignant tumor of thyroid gland (disorder)|
5.	Bone cancer	428281000|Malignant neoplasm of bone (disorder)|
6.	Breast cancer	254837009|Malignant tumor of breast (disorder)|
7.	Cervical cancer	363354003|Malignant tumor of cervix (disorder)|
8.	Colorectal cancer	363510005|Malignant tumor of large intestine (disorder)|
9.	Melanoma skin cancer	372130007|Malignant neoplasm of skin (disorder)|
10.	Non melanoma skin cancer	402815007|Squamous cell carcinoma (disorder)| 254701007|Basal cell carcinoma of skin (disorder)|
11.	**Heart (cardiovascular) disease**	49601007|Disorder of cardiovascular system (disorder)|
12.	Ischaemic heart disease	414545008|Ischemic heart disease (disorder)|
13.	Cerebrovascular disease	62914000|Cerebrovascular disease (disorder)|
14.	Peripheral vascular disease	400047006|Peripheral vascular disease (disorder)|
15.	Heart failure	84114007|Heart failure (disorder)|
16.	Rheumatic heart disease	23685000|Rheumatic heart disease (disorder)|
17.	Congenital heart disease	13213009|Congenital heart disease (disorder)|
18.	**Chronic respiratory disease**	17097001|Chronic disease of respiratory system (disorder)|
19.	Asthma	195967001|Asthma (disorder)|
20.	Chronic obstructive pulmonary disease	13645005|Chronic obstructive lung disease (disorder)|
21.	Sleep apnoea	73430006|Sleep apnea (disorder)|
22.	**Diabetes**	73211009|Diabetes mellitus (disorder)|
23.	Diabetes type 1	46635009|Diabetes mellitus type 1 (disorder)|
24.	Diabetes type 2	44054006|Diabetes mellitus type 2 (disorder)|
25.	Gestational diabetes	11687002|Gestational diabetes mellitus (disorder)|
26.	**Mental Illness**	74732009|Mental disorder (disorder)|
27.	Mood disorder	46206005|Mood disorder (disorder)|
28.	Major depression	370143000|Major depressive disorder (disorder)|
29.	Bipolar disorder	13746004|Bipolar disorder (disorder)|
30.	Schizophrenia	58214004|Schizophrenia (disorder)|
31.	Anxiety disorder	197480006|Anxiety disorder (disorder)|
32.	Personality disorder	33449004|Personality disorder (disorder)|
33.	Eating disorder	72366004|Eating disorder (disorder)|
34.	Problem gambling	18085000|Compulsive gambling (disorder)|
35.	Substance dependency	2403008|Psychoactive substance dependence (disorder)|

No	Chronic Disease Category and Example	SNOMED CT Encoding
36.	**Neurological conditions**	102957003|Neurological finding (finding)|
37.	Paralysis	44695005|Paralysis (finding)|
38.	Weakened muscles	26544005|Muscle weakness (finding)|
39.	Low coordination	302289002|Coordination problem (finding)|
40.	Loss of sensation	44077006|Numbness (finding)|
41.	Seizures	91175000|Seizure (finding)|
42.	Confusion	40917007|Clouded consciousness (finding)|
43.	Pain	22253000|Pain (finding)|
44.	Altered cognitive functions	386806002|Impaired cognition (finding)|
45.	**Musculoskeletal diseases**	928000|Disorder of musculoskeletal system (disorder)|
46.	Arthritis	3723001|Arthritis (disorder)|
47.	Osteoporosis	64859006|Osteoporosis (disorder)|

8.4.3 Result of Querying Chronic Conditions

The seven chronic conditions were queried for equivalency and subsumption, defining attributes, and context free. The results are shown in **Table 8-19**.

Table 8-19. Results of querying the chronic conditions.

No	Category	SNOMED CT	Equivalency and Subsumption	Defining Attributes	Context Free
1.	Cancer	363346000|Malignant neoplastic disease (disorder)|	111	1	12
2.	Heart (cardiovascular) disease	49601007|Disorder of cardiovascular system (disorder)|	249	13	25
3.	Chronic respiratory disease	17097001|Chronic disease of respiratory system (disorder)|	25	0	3
4.	Diabetes	73211009|Diabetes mellitus (disorder)|	4	10	1
5.	Mental Illness	74732009|Mental disorder (disorder)|	152	1	40
6.	Neurological conditions	102957003|Neurological finding (finding)|	359	52	20
7.	Musculoskeletal diseases	928000|Disorder of musculoskeletal system (disorder)|	538	5	34

For the equivalency and subsumption queries, an SQL statement (**SELECT * FROM ExpressionsTable ET, SCT_TransitiveClosure TC WHERE ET.ProximalPrimitive=TC.SubTypeId AND TC.SuperTypeId=Id** where Id refers to the proximal primitive of the predicate expression) was used to retrieve potential candidate expressions from the 7,356 unique expressions used to represent the encounter diagnoses using the method described in **Section 8.3.1.2**. The results are reported in **Table 8-20**. The numbers of candidate expressions retrieved are shown in the **Candidates** column while the number of expressions, after testing for equivalency and subsumption using the normal form, are shown in the **Final** column. The purpose is to demonstrate how the initial SQL statement can be used to identify a set of candidate terms rather than testing all 7,356 expressions.

Table 8-20. Seven categories of disease, the SNOMED CT encoding used, and the candidate and final number of expressions retrieved.

No	Category	SNOMED CT	Is Primitive	Proximal Primitive	Defining Attributes	Candidates	Final
1.	Cancer	363346000|Malignant neoplastic disease (disorder)|	Fully defined	64572001|Disease (disorder)|	116676008|Associated morphology (attribute)|	2,124	111
2.	Heart (cardiovascular) disease	49601007|Disorder of cardiovascular system (disorder)|	Fully defined	64572001|Disease (disorder)|	363698007|Finding site (attribute)|	2,730	249
3.	Chronic respiratory disease	17097001|Chronic disease of respiratory system (disorder)|	Fully defined	64572001|Disease (disorder)|	263502005|Clinical course (attribute)| 363698007|Finding site (attribute)|	239	25
4.	Diabetes	73211009|Diabetes mellitus (disorder)|	Primitive	73211009|Diabetes mellitus (disorder)|	363698007|Finding site (attribute)|	4	4
5.	Mental Illness	74732009|Mental disorder (disorder)|	Primitive	74732009|Mental disorder (disorder)|		152	152
6.	Neurological conditions	102957003|Neurological finding (finding)|	Primitive	102957003|Neurological finding (finding)|		363	363
7.	Musculoskeletal diseases	928000|Disorder of musculoskeletal system (disorder)|	Fully defined	64572001|Disease (disorder)|	363698007|Finding site (attribute)|	2,730	538

The next seven sub-sections describe the results of each of the seven chronic diseases and provide examples of each type of query.

8.4.3.1 Cancer

The nine examples of cancer were all subtypes of the cancer category concept (refer to **Figure 8-11**). What this means is that a subsumption query for just **363346000|Malignant neoplastic disease (disorder)|** will retrieve the nine example concepts.

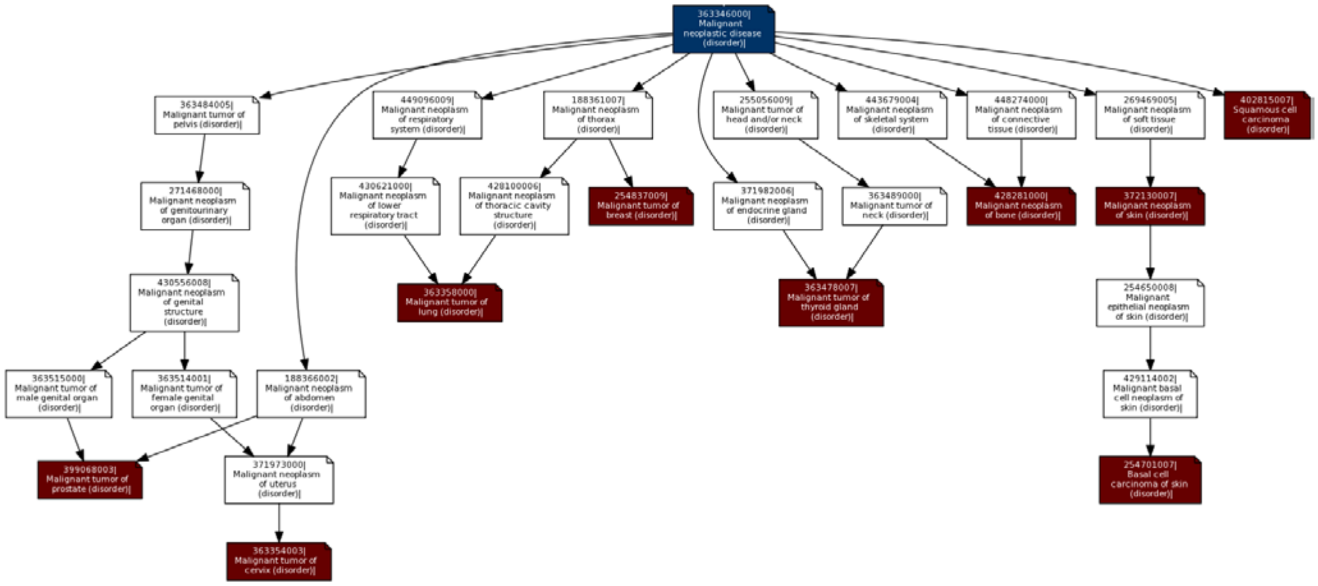


Figure 8-11. Location of cancer category concept and examples of cancer concepts in the hierarchy.

Examples of the subsumption query found in the primary care dataset included **109841003|Liver cell carcinoma (disorder)|**, **118599009|Hodgkin's disease (disorder)|** and **93143009|Leukemia, disease (disorder)|**. The defining attributes query retrieved one concept, **15886004|Screening for cancer (procedure)|**, which has a defining attribute of **363702006|Has focus (attribute)|=363346000|Malignant neoplastic disease (disorder)|**. Examples of the context free query are shown in **Table 8-21**.

Table 8-21. Example of context free query for "363346000|Malignant neoplastic disease (disorder)|".

No	Encounter Diagnosis	SNOMED CT
1.	Possible of basal cell carcinoma	243796009|Situation with explicit context (situation)|: 246090004|Associated finding (attribute)|= **254701007|Basal cell carcinoma of skin (disorder)|**, 408729009|Finding context (attribute)|= 410590009|Known possible (qualifier value)|
2.	No lung cancer	373572006|Clinical finding absent (situation)|: 246090004|Associated finding (attribute)|= **363358000|Malignant tumor of lung (disorder)|**
3.	Family history of colon cancer	57177007|Family history with explicit context (situation)|: 246090004|Associated finding (attribute)|= **363406005|Malignant tumor of colon (disorder)|**

8.4.3.2 Heart (Cardiovascular) Disease

The six examples of cardiovascular disease were all subtypes of the cardiovascular disease category concept (refer to **Figure 8-12**). What this means is that a subsumption query for just **49601007|Disorder of cardiovascular system (disorder)|** will retrieve the six other concepts.

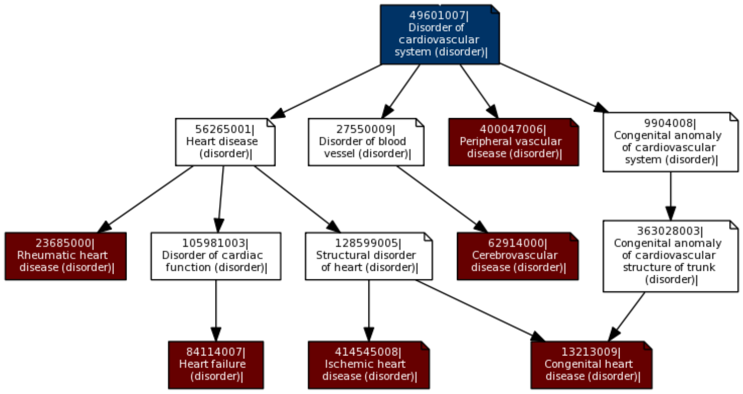


Figure 8-12. Location of heart (cardiovascular) disease category concept and examples of heart (cardiovascular) disease concepts in the hierarchy.

Examples of the subsumption query found in the primary care dataset included **10633002|Acute congestive heart failure (disorder)|**, **128060009|Venous varices (disorder)|** and **2929001|Occlusion of artery (disorder)|**. Examples of defining attributes query are shown in **Table 8-22** while examples of the context free query are shown in **Table 8-23**.

Table 8-22. Example of defining attributes query for "49601007|Disorder of cardiovascular system (disorder)|".

No	Encounter Diagnosis	SNOMED CT
1.	Hypertensive retinopathy	6962006|Hypertensive retinopathy (disorder)|: 363698007|Finding site (attribute)|= 5665001|Retinal structure (body structure)|, 47429007|Associated with (attribute)|= **38341003|Hypertensive disorder, systemic arterial (disorder)|**
2.	Rubella (disorder)	36653000|Rubella (disorder)|: 246075003|Causative agent (attribute)|= 5210005|Rubella virus (organism)|, 363705008|Has definitional manifestation (attribute)|= **238812004|Exanthematous disorder (disorder)|**, 370135005|Pathological process (attribute)|= 441862004|Infectious process (qualifier value)|
3.	Cardioversion	250980009|Cardioversion (procedure)|: 363702006|Has focus (attribute)|= **44808001|Conduction disorder of the heart (disorder)|**, 363704007|Procedure site (attribute)|= 80891009|Heart structure (body structure)|

Table 8-23. Example of context free query for "49601007|Disorder of cardiovascular system (disorder)|".

No	Encounter Diagnosis	SNOMED CT
1.	Possible angina	243796009|Situation with explicit context (situation)|: 246090004|Associated finding (attribute)|= **194828000|Angina (disorder)|**, 408729009|Finding context (attribute)|= 410590009|Known possible (qualifier value)|),
2.	Recent stroke	243796009|Situation with explicit context (situation)|: 246090004|Associated finding (attribute)|= **230690007|Cerebrovascular accident (disorder)|**, 408731000|Temporal context (attribute)|= 6493001|Recent (qualifier value)|)

No	Encounter Diagnosis	SNOMED CT
3.	Risk of coronary heart disease	243796009|Situation with explicit context (situation)|: 246090004|Associated finding (attribute)|= 53741008|Coronary arteriosclerosis (disorder)|, 408729009|Finding context (attribute)|= 410519009|At risk context (qualifier value)|).

8.4.3.3 Chronic Respiratory Disease

Of the three examples of chronic respiratory disease concepts, only one was retrieved using the concept **17097001|Chronic disease of respiratory system (disorder)|** (refer to **Figure 8-13**). The closest concept that covers all three examples of chronic respiratory diseases is **64572001|Disease (disorder)|**.

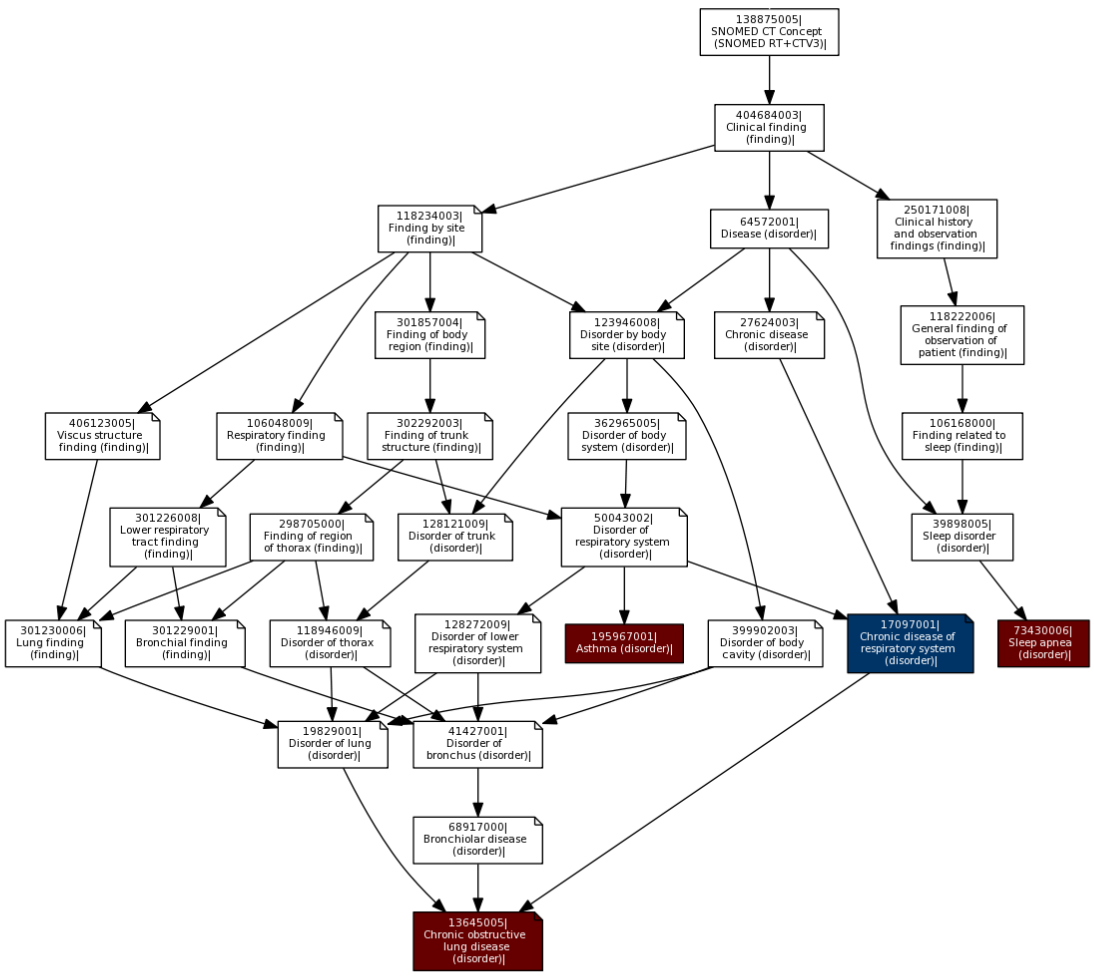


Figure 8-13. Location of chronic respiratory disease category concept and examples of chronic respiratory disease concepts in the hierarchy.

Examples of the subsumption query found in the primary care dataset included **140004|Chronic pharyngitis (disorder)|**, **63480004|Chronic bronchitis (disorder)|** and **425748003|Acute exacerbation of chronic bronchitis (disorder)|**. No defining attributes query results were retrieved. Examples of the context free query are shown in **Table 8-24**.

Table 8-24. Example of context free query for “17097001|Chronic disease of respiratory system (disorder)|”.

No	Encounter Diagnosis	SNOMED CT
1.	Possible intermittent asthma	243796009|Situation with explicit context (situation)|: 246090004|Associated finding (attribute)|=(**195967001|Asthma (disorder)|**: 263502005|Clinical course (attribute)|= 7087005|Intermittent (qualifier value)|), 408729009|Finding context (attribute)|= 410590009|Known possible (qualifier value)|
2.	Possible COPD	243796009|Situation with explicit context (situation)|: 246090004|Associated finding (attribute)|= **13645005|Chronic obstructive lung disease (disorder)|**, 408729009|Finding context (attribute)|= 410590009|Known possible (qualifier value)|
3.	No COPD	373572006|Clinical finding absent (situation)|: 246090004|Associated finding (attribute)|= **13645005|Chronic obstructive lung disease (disorder)|**

8.4.3.4 Diabetes

The three examples of diabetes were all subtypes of the diabetes category concept (refer to **Figure 8-14**). What this means is that a subsumption query for just **73211009|Diabetes mellitus (disorder)|** will retrieve the three other concepts.

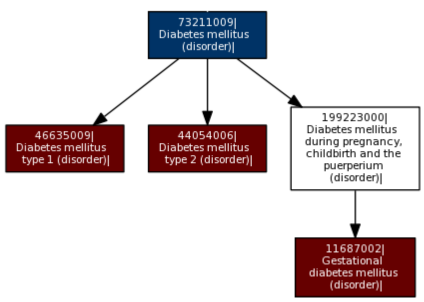


Figure 8-14. Location of diabetes category concept and examples of diabetes concepts in the hierarchy.

Other than the three examples of diabetes and the diabetes category concept, no other expressions were retrieved using the equivalency and subsumption query. Examples of the context free query are shown in **Table 8-25**. One expression was retrieved using the context free query, risk of diabetes (i.e., **243796009|Situation with explicit context (situation)|:246090004|Associated finding (attribute)|=73211009|Diabetes mellitus (disorder)|, 408729009|Finding context (attribute)|= 410519009|At risk context (qualifier value)|**).

Table 8-25. Example of context free query for “73211009|Diabetes mellitus (disorder)|”.

No	Encounter Diagnosis	SNOMED CT
1.	371087003|Diabetic foot ulcer (disorder)|	371087003|Diabetic foot ulcer (disorder)|: 47429007|Associated with (attribute)|= 73211009|Diabetes mellitus (disorder)| {116676008|Associated morphology (attribute)|= 56208002|Ulcer (morphologic abnormality)|, 363698007|Finding site (attribute)|= 56459004|Foot structure (body structure)|}

No	Encounter Diagnosis	SNOMED CT
2.	4855003|Diabetic retinopathy (disorder)|	4855003|Diabetic retinopathy (disorder)|: 47429007|Associated with (attribute)|= 73211009|Diabetes mellitus (disorder)| {116676008|Associated morphology (attribute)|= 49755003|Morphologically abnormal structure (morphologic abnormality)|, 363698007|Finding site (attribute)|= 5665001|Retinal structure (body structure)|}
3.	127013003|Diabetic renal disease (disorder)|	127013003|Diabetic renal disease (disorder)|: 363698007|Finding site (attribute)|= 64033007|Kidney structure (body structure)|, 47429007|Associated with (attribute)|= 56019007|Systemic disease (finding)|, 47429007|Associated with (attribute)|= 73211009|Diabetes mellitus (disorder)|

8.4.3.5 *Mental Illness*

The nine examples of mental illness were all subtypes of the mental illness category concept (refer to **Figure 8-15**). What this means is that a subsumption query for just **74732009|Mental disorder (disorder)|** will retrieve the nine other concepts.

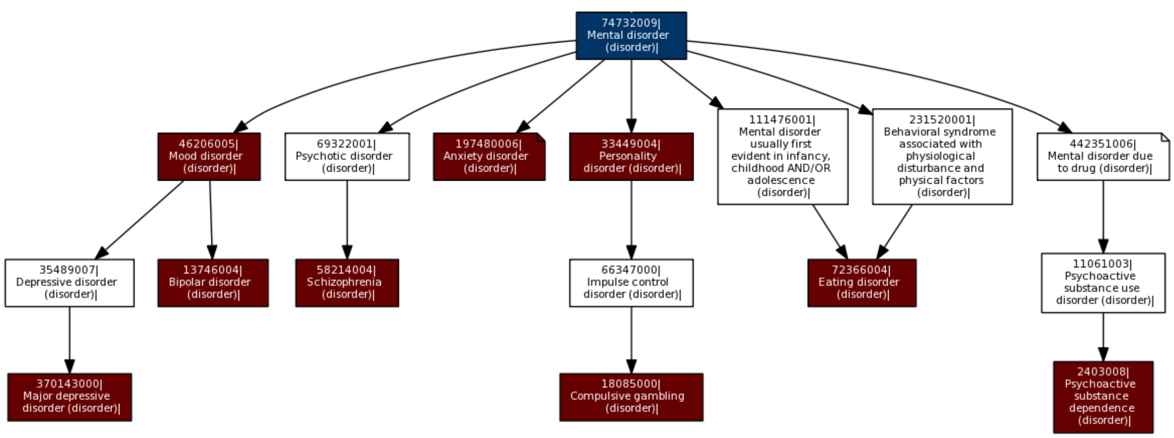


Figure 8-15. Location of mental illness category concept and examples of mental illness concepts in the hierarchy.

Examples of the subsumption query found in the primary care dataset included **1376001|Obsessive compulsive personality disorder (disorder)|**, **2776000|Delirium (disorder)|** and **13601005|Paranoid personality disorder (disorder)|**. There was one expression retrieved via the defining attributes, and that was through blindness associated with mental disorder (i.e., **105597003|Blindness AND/OR vision impairment level (disorder)|:47429007|Associated with (attribute)|=74732009|Mental disorder (disorder)|**). Examples of the context free query are shown in **Table 8-26**.

Table 8-26. Example of context free query for "74732009|Mental disorder (disorder)|".

No	Encounter Diagnosis	SNOMED CT
1.	No bipolar disorder	373572006|Clinical finding absent (situation)|: 246090004|Associated finding (attribute)|= **13746004|Bipolar disorder (disorder)|**

No	Encounter Diagnosis	SNOMED CT
2.	History of attention deficit hyperactivity disorder	417662000|History of clinical finding in subject (situation)|: 246090004|Associated finding (attribute)|= **406506008|Attention deficit hyperactivity disorder (disorder)|**
3.	Family history of psychoactive substance dependence	57177007|Family history with explicit context (situation)|: 246090004|Associated finding (attribute)|= **2403008|Psychoactive substance dependence (disorder)|**

8.4.3.6 Neurological Conditions

Of the eight examples of neurological conditions concepts, only three were retrieved using the concept **102957003|Neurological finding (finding)|** (refer to **Figure 8-16**). The closest concept that covers all eight examples of neurological conditions is **404684003|Clinical finding (finding)|**.

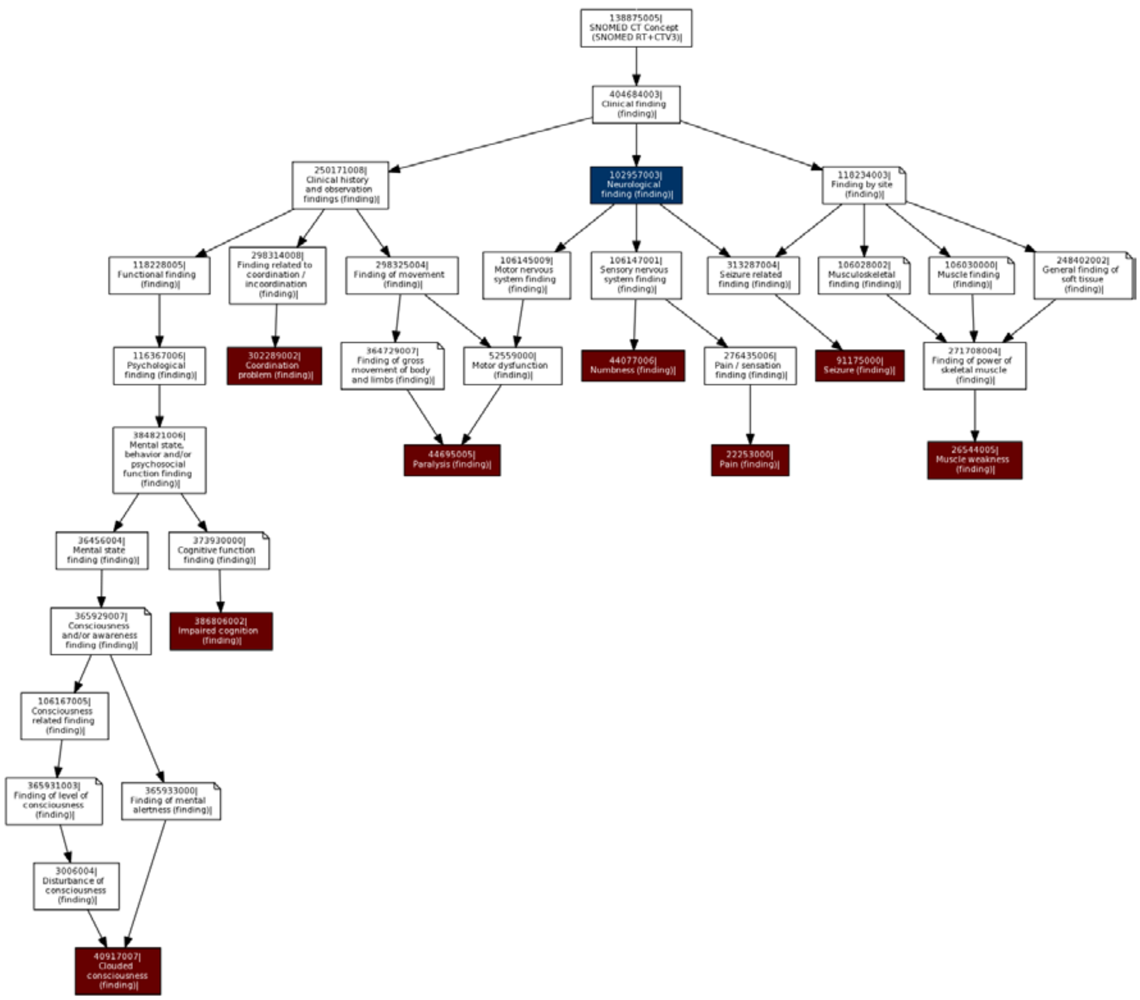


Figure 8-16. Location of neurological conditions category concept and examples of neurological conditions concepts in the hierarchy.

Examples of the subsumption query found in the primary care dataset included **21522001|Abdominal pain (finding)|**, **44169009|Loss of sense of smell (finding)|** and **59770006|Dyslexia (finding)|**. Examples of defining attributes query are shown in **Table 8-27**, while examples of the context free query are shown in **Table 8-28**.

Table 8-27. Example of defining attributes query for "102957003|Neurological finding (finding)|".

No	Encounter Diagnosis	SNOMED CT
1.	Pruritic disorder	279333002|Pruritic disorders (disorder)|: 363698007|Finding site (attribute)|= 39937001|Skin structure (body structure)|, 363705008|Has definitional manifestation (attribute)|= **418363000|Itching of skin (finding)|**
2.	Seizure disorder	128613002|Seizure disorder (disorder)|: 363698007|Finding site (attribute)|= 12738006|Brain structure (body structure)|, 363705008|Has definitional manifestation (attribute)|= **91175000|Seizure (finding)|**
3.	Angina	194828000|Angina (disorder)|: 363705008|Has definitional manifestation (attribute)|= **22253000|Pain (finding)|** {116676008|Associated morphology (attribute)|= 449735000|Structural change due to ischemia (morphologic abnormality)|, 363698007|Finding site (attribute)|= 80891009|Heart structure (body structure)|}

Table 8-28. Example of context free query for "102957003|Neurological finding (finding)|".

No	Encounter Diagnosis	SNOMED CT
1.	Recent paresthesia of foot	243796009|Situation with explicit context (situation)|: 246090004|Associated finding (attribute)|= 309087008|Paresthesia of foot (finding)|, 408731000|Temporal context (attribute)|= 6493001|Recent (qualifier value)|
2.	No history of ankle pain	443508001|No history of clinical finding in subject (situation)|: 246090004|Associated finding (attribute)|= 247373008|Ankle pain (finding)|
3.	No seizure	373572006|Clinical finding absent (situation)|: 246090004|Associated finding (attribute)|= 91175000|Seizure (finding)|

8.4.3.7 Musculoskeletal Disease

The two examples of musculoskeletal disease were both subtypes of the musculoskeletal disease category concept (refer to **Figure 8-17**). What this means is that a subsumption query for just **928000|Disorder of musculoskeletal system (disorder)|** will retrieve the two other concepts.

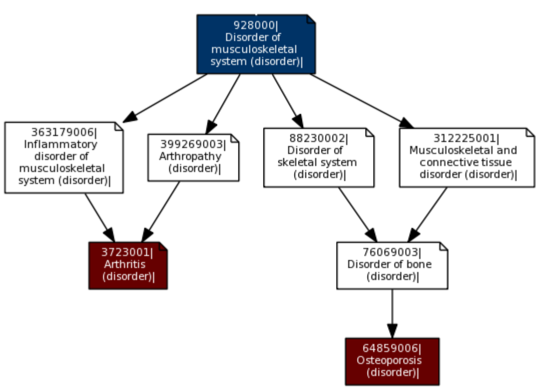


Figure 8-17. Location of musculoskeletal disease category concept and examples of musculoskeletal disease concepts in the hierarchy.

Examples of the subsumption query found in the primary care dataset included **16114001|Fracture of ankle (disorder)|**, **22193007|Degenerative joint disease of hand (disorder)|** and **239873007|Osteoarthritis of knee (disorder)|**. Examples of defining attributes query are shown in **Table 8-29**, which were only from post-coordinated expressions, while examples of the context free query are shown in **Table 8-30**.

Table 8-29. Example of defining attributes query for "928000|Disorder of musculoskeletal system (disorder)|".

No	Encounter Diagnosis	SNOMED CT
1.	Neuropathy due to degenerative disorder	386033004|Neuropathy (disorder)|: 47429007|Associated with (attribute)|= **362975008|Degenerative disorder (disorder)|**
2.	Injury due to discitis	417163006|Traumatic AND/OR non-traumatic injury (disorder)|: 42752001|Due to (attribute)|= **2304001|Discitis (disorder)|**
3.	Neck pain due to arthropathy of articular surface	81680005|Neck pain (finding)|: 42752001|Due to (attribute)|=(**399269003|Arthropathy (disorder)|**: 363698007|Finding site (attribute)|= 70746003|Structure of articular surface of bone (body structure)|)

Table 8-30. Example of context free query for "928000|Disorder of musculoskeletal system (disorder)|".

No	Encounter Diagnosis	SNOMED CT
1.	Possible in left metatarsal bone	243796009|Situation with explicit context (situation)|: 246090004|Associated finding (attribute)|=(**125605004|Fracture of bone (disorder)|**: {363698007|Finding site (attribute)|=(53884002|Metatarsal bone structure (body structure)|: 272741003|Laterality (attribute)|= 7771000|Left (qualifier value)|)}), 408729009|Finding context (attribute)|= 410590009|Known possible (qualifier value)|
2.	No fracture	373572006|Clinical finding absent (situation)|: 246090004|Associated finding (attribute)|= **125605004|Fracture of bone (disorder)|**
3.	Suspected psoriasis with arthropathy	444433005|Suspected clinical finding (situation)|: 246090004|Associated finding (attribute)|= **33339001|Psoriasis with arthropathy (disorder)|**

8.5 Discussion

In this study, the fundamentals of SNOMED CT queries were reviewed (refer to **Appendix F**), an analysis of the data storage methods suggested by the IHTSDO was conducted, a retrieval method to optimise computing structural subsumption was developed and the retrieval method was used to identify chronic diseases in a primary care dataset. The discussion section is divided into four parts. First, the data storage methods. Second, the benefits of the retrieval method developed. Third, the benefits of SNOMED CT retrieval. Fourth, the challenges of SNOMED CT retrieval. The limitations and future work of this chapter are described in **Chapter Ten**.

8.5.1 Data Storage Methods

The TIG suggested three methods (excluding the XML representation) for storing post-coordinated expressions in a relational database format. The "restricted relational representation" should be ruled out for use because it can only accommodate a limited set of post-coordination. The "parasable text representation," although easy to implement and supports unlimited post-coordination is less efficient for data retrieval as the entire expression needs to be parsed before it can be determined if it is useful. On the other hand, the "unrestricted relational representation" is more efficient for data retrieval but is more complex to implement. A fourth relational database format, "enhanced parasable representation," was proposed that attempts to optimise data retrieval by defining discrete data elements for the focus concept, contextual values, and pre-computed long normal form. The pre-computed long normal form allows for quick testing for equivalency, while the discrete contextual values allow for efficient retrieval of expressions within the required context. This enabled the optimised structural subsumption method to retrieve a set of candidate expressions to be tested using the structural subsumption method instead of testing all expressions contained in a dataset.

8.5.2 Retrieval Method

The retrieval method developed contributed to two main areas. First, how to structure predicate expressions to ensure all relevant candidate expressions will be retrieved. Second, how to optimise structural subsumption by classifying the predicate expression into one of three categories (i.e., pre-coordinated concept that is primitive; post-coordinated expression with a primitive focus concept; pre-coordinated concept that is fully-defined or post-coordinated expression with a fully-defined focus concept). The structuring predicate expression steps demonstrated how expressions should be created for querying defining attributes and how to conduct context free queries. This optimised method was demonstrated by conducting queries against a set of clinical care guidelines. The use of an SQL statement to retrieve a set of candidate expressions was especially useful when the focus concept of the predicate expression was a primitive concept. The time it took to execute the query to retrieve the initial set of candidate expressions took the same time as querying a transitive closure table, much like how two pre-coordinated concepts are tested to see if they have a supertype/subtype relationship. For example, for **73211009|Diabetes mellitus (disorder)|**, **74732009|Mental disorder (disorder)|** and **102957003|Neurological finding (finding)|**, the SQL statement was able to narrow down the number of candidate expressions to 4, 152 and 363 respectively. For the fully-defined focus concepts, the number ranged from 239 to 2,730. This means that the initial SQL statement was able to improve the efficiency by reducing the number of expression to test by between 62.9% (7,356-2,730/7,356) and 99.9% (7,356-4/7,356). In all four the cases, the proximal primitive was **64572001|Disease (disorder)|**, which is close to the top of the hierarchy. The range of 239 to 2,730 was due to the number of attributes that were queried, including **116676008|Associated morphology (attribute)|**, **363698007|Finding site (attribute)|** and **263502005|Clinical course (attribute)|**. For the seven queries, a comparison was conducted against all 7,356 expressions and the results showed that the initial SQL statement to retrieve the set of candidate expressions did not miss any of the final expressions that were identified, which demonstrated that this method was just as effective.

8.5.3 Benefits of SNOMED CT Retrieval

Subsumption Queries. For the seven categories of chronic diseases, the concept assigned to the category was able to retrieve all the examples of the disease except for neurological conditions and chronic respiratory disease. The concept **102957003|Neurological finding (finding)|** was only able to retrieve four of the eight concepts while the concept **17097001|Chronic disease of respiratory system (disorder)|** was only able to retrieve one of the three concepts. In all seven categories, the number of unique expressions that were retrieved in the primary care dataset ranged from 4 for **73211009|Diabetes mellitus (disorder)|** to 538 for **928000|Disorder of musculoskeletal system (disorder)|**. **74732009|Mental disorder (disorder)|**, which contained 152 subtype expressions, included concepts such as **1376001|Obsessive compulsive personality disorder (disorder)|**, **2776000|Delirium (disorder)|** and **13601005|Paranoid personality disorder (disorder)|**. These concepts are probably relevant to identifying mental illnesses but were not confirmed by a clinician. Caution should be exercised when interpreting the subsumption results as not all the subtype concepts may apply. For example, the category musculoskeletal diseases (i.e., with **928000|Disorder of musculoskeletal system (disorder)|**) contained the examples **3723001|Arthritis (disorder)|** and **64859006|Osteoporosis (disorder)|**. In addition to retrieving these two concepts, the subsumption query also retrieved concepts such as **16114001|Fracture of ankle (disorder)|**, **428016006|Contusion of rib (disorder)|** and **427782005|Injury of mandible (disorder)|**. These concepts are probably out of the scope of what should be considered chronic diseases.

Defining Attribute Queries. The defining attribute queries were useful for retrieving related concepts. For example, the related concepts for **73211009|Diabetes mellitus (disorder)|** that were retrieved included **371087003|Diabetic foot ulcer (disorder)|**, **4855003|Diabetic retinopathy (disorder)|** and **127013003|Diabetic renal disease (disorder)|**. These concepts are not subtypes of **73211009|Diabetes mellitus (disorder)|** but are implicit that **73211009|Diabetes mellitus (disorder)|** is present. Caution should be exercised when interpreting the results based on defining attributes because they could have a different meaning. For example, **15886004|Screening for cancer (procedure)|** does not mean that cancer is present even though the concept has a defining attribute of **363702006|Has focus (attribute)|=363346000|Malignant neoplastic disease (disorder)|**. In another example, using the defining attributes to locate concepts related to **102957003|Neurological finding (finding)|**, the concepts retrieved included **21522001|Abdominal pain (finding)|**. Although **21522001|Abdominal pain (finding)|** is a subtype of **102957003|Neurological finding (finding)|**, it probably does not apply in this context as a chronic disease.

Context Free Queries. The context free queries allowed expressions that were used in different contexts such as "at risk," "known possible," "suspected," "known absent" to be retrieved. The context free queries were used for exploratory purposes in order to determine what other context the concept is used in.

8.5.4 *Challenges of SNOMED CT Retrieval*

There are a number of challenges that can affect the testing for equivalency and subsumption. A summary of these challenges are listed in **Table 8-31**. These challenges are the result of using structural subsumption as well as how concepts in SNOMED CT are modelled and arranged in the subsumption hierarchy.

Table 8-31. Summary of challenges when testing for equivalency and subsumption.

No	Challenge
1.	Relationship Groups
2.	Primitive Concepts
3.	Modelling of Concepts
4.	Errors in Modelling of Concepts
5.	Errors in Hierarchical Structure of Concepts
6.	Inactive Concepts
7.	Incomplete Modelling of Concepts
8.	Different Encoding Methods

8.5.4.1 Relationship Groups

Relationship groups pose a challenge for structural subsumption because the groups are applied differently in defining attributes and when post-coordinating. "Abrasion of ankle" can be represented with the pre-coordinated concept **211334007|Abrasion, ankle (disorder)|** or the post-coordinated expression **399963005|Abrasion (disorder)|:363698007|Finding site (attribute)|=67269001|Skin structure of ankle (body structure)|**. However, when both concepts are converted to their long normal forms, the pre-coordinated concept has the inherent relationship group while the post-coordinated expression does not (refer to **Figure 8-18**). In this case, the result will be subsumption rather than equivalence. A workaround would be to apply the curly brackets {} when generating to the long normal form to both grouped and ungrouped attributes.

211334007|Abrasion, ankle (disorder)|	**399963005|Abrasion (disorder)|:** **363698007|Finding site (attribute)|=** **67269001|Skin structure of ankle (body structure)|**
64572001|Disease (disorder)|: **{**116676008|Associated morphology (attribute)|= 400061001|Abrasion (morphologic abnormality)|, 363698007|Finding site (attribute)|= 67269001|Skin structure of ankle (body structure)|**}**	64572001|Disease (disorder)|: 116676008|Associated morphology (attribute)|= 400061001|Abrasion (morphologic abnormality)|, 363698007|Finding site (attribute)|= 67269001|Skin structure of ankle (body structure)|

Figure 8-18. Comparing the long normal form of the pre-coordinated concept and post-coordinated expressions for representing "abrasion of ankle".

8.5.4.2 Primitive Concepts

Primitive concepts can cause difficulty when testing for equivalency and subsumption when they are underspecified. As the proximal primitive concept refers to itself, there is no possibility of creating a post-coordinated expression that will be computed to be equivalent to the pre-coordinated concept. For example, there is a pre-coordinated concept for "hand pain," **53057004|Hand pain (finding)|**. A post-coordinated expression that can fully represent the same concept is **22253000|Pain (finding)|:363698007|Finding site (attribute)|=85562004|Hand structure (body structure)|**. However, since **53057004|Hand pain (finding)|** is a primitive concept, even though our post-coordinated expression fully expresses the full semantics of "hand pain," it is not deemed equivalent to the pre-coordinated concept (refer to Figure 8-19). This is in contrasted to the previous example where **47933007|Foot pain (finding)|**, which as a fully-defined concept is computed to be equivalent to **22253000|Pain (finding)|:363698007|Finding site (attribute)|=56459004|Foot structure (body structure)|**.

53057004|Hand pain (finding)|	**22253000|Pain (finding)|: 363698007|Finding site (attribute)|= 85562004|Hand structure (body structure)|**
53057004|Hand pain (finding)|: 363698007|Finding site (attribute)|= 85562004|Hand structure (body structure)|	22253000|Pain (finding)|: 363698007|Finding site (attribute)|= 85562004|Hand structure (body structure)|

Figure 8-19. Comparing the long normal forms of "53057004|Hand pain (finding)|" and "22253000|Pain (finding)|:363698007|Finding site (attribute)|=85562004|Hand structure (body structure)|".

8.5.4.3 Modelling of Concepts

The concepts used to form post-coordinated expressions are not always the same concepts used in the definition of concepts. For example, representing "acute pain" in a post-coordination expression could be represented with **22253000|Pain (finding)|:263502005|Clinical course (attribute)|=373933003|Acute onset (qualifier value)|**. However, this is different from the defining attributes of **274663001|Acute pain (finding)|** (refer to Figure 8-20).

274663001|Acute pain (finding)|	**22253000|Pain (finding)|: 263502005|Clinical course (attribute)|= 373933003|Acute onset (qualifier value)|**
301369003|Finding of pattern of pain (finding)|: 263502005|Clinical course (attribute)|= 424124008|Sudden onset AND/OR short duration (qualifier value)|	22253000|Pain (finding)|: 263502005|Clinical course (attribute)|= 373933003|Acute onset (qualifier value)|

Figure 8-20. Comparing the long normal forms of "274663001|Acute pain (finding)|" and "22253000|Pain (finding)|:263502005|Clinical course (attribute)|=373933003|Acute onset (qualifier value)|".

When comparing the pre-coordinated concept and post-coordinated expression, the result is that they are unrelated. This is because **301369003|Finding of pattern of pain (finding)|** is a subtype of **22253000|Pain (finding)|** while **373933003|Acute onset (qualifier value)|** is a subtype of **424124008|Sudden onset AND/OR short duration (qualifier value)|** (refer to Figure 8-21).

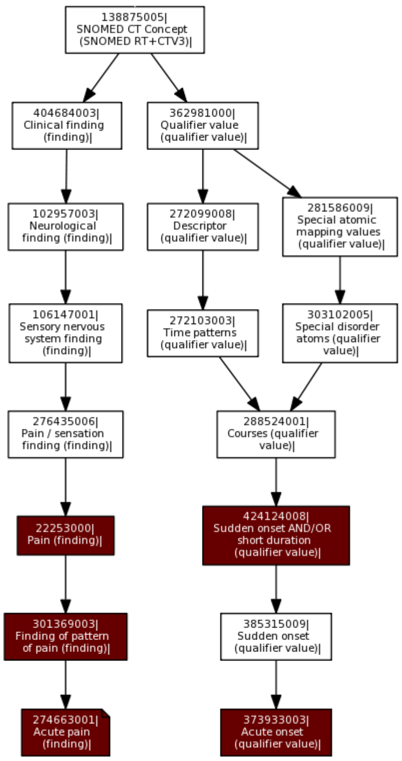


Figure 8-21. Location of concepts used to represent "acute pain" in the hierarchy.

8.5.4.4 Errors in Modelling of Concepts

Errors in the modelling of concepts can cause expressions that are supposed to be equivalent to be computed as subsumption or unrelated. **110168002|Abrasion of chin (disorder)|** is modelled incorrectly using the **363698007|Finding site (attribute)|** of **30291003|Chin structure (body structure)|** and **73897004|Skin structure of face (body structure)|** instead of **23747009|Skin structure of chin (body structure)|** (refer to **Figure 8-22**). If it were to be compared to a correctly post-coordinated expression, **399963005|Abrasion (disorder)|:363698007|Finding site (attribute)|=23747009|Skin structure of chin (body structure)|**, the results will be computed as subsumption rather than equivalent.

110168002|Abrasion of chin (disorder)|	399963005|Abrasion (disorder)|: 363698007|Finding site (attribute)|= 23747009|Skin structure of chin (body structure)|
64572001|Disease (disorder)|: {116676008|Associated morphology (attribute)|= 400061001|Abrasion (morphologic abnormality)|, 363698007|Finding site (attribute)|= 30291003|Chin structure (body structure)|} {116676008|Associated morphology (attribute)|= 400061001|Abrasion (morphologic abnormality)|, 363698007|Finding site (attribute)|= 73897004|Skin structure of face (body structure)|}	64572001|Disease (disorder)|: {116676008|Associated morphology (attribute)|= 400061001|Abrasion (morphologic abnormality)|, 363698007|Finding site (attribute)|= 23747009|Skin structure of chin (body structure)|}

Figure 8-22. Comparing the long normal of forms of "110168002|Abrasion of chin (disorder)|" and "399963005|Abrasion (disorder)|:363698007|Finding site (attribute)|=23747009|Skin structure of chin (body structure)|".

8.5.4.5 Incomplete Modelling of Concepts

While over 93% (16,016 out of 17,140) of the concepts in the **373873005|Pharmaceutical / biologic product (product)|** hierarchy are defined with a concept from the **105590001|Substance (substance)|** hierarchy using the Concept Model attribute **127489000|Has active ingredient (attribute)|**, there are still concepts that are incompletely defined. An example that was identified in this study is that **1182007|Hypotensive agent (product)|** is not defined with **372586001|Hypotensive agent (substance)|**, which was the reason why the results of querying for subtypes of **372586001|Hypotensive agent (substance)|** and **373873005|Pharmaceutical / biologic product (product)|:127489000|Has active ingredient (attribute)|=372586001|Hypotensive agent (substance)|** resulted in a different number of concepts retrieved. Therefore additional work needs to be done to ensure concepts are modelled completely as incompletely modelled concepts may not be retrieved with queries such as the defining attributes query.

8.5.4.6 Errors in Hierarchical Structure of Concepts

Errors in the hierarchy are another hindrance to testing for equivalency and subsumption. For example, the structure of the concept **44077006|Numbness (finding)|** and its subtypes has many issues, the most obvious of which is that **298753001|Numbness of upper limb (finding)|** and **309537005|Numbness of lower limb (finding)|** are not subtype concepts of **310501001|Numbness of limbs (finding)|** (refer to **Figure 8-23**). Therefore, querying the subtypes of **310501001|Numbness of limbs (finding)|** will not yield concepts such as **309521004|Numbness of hand (finding)|**.

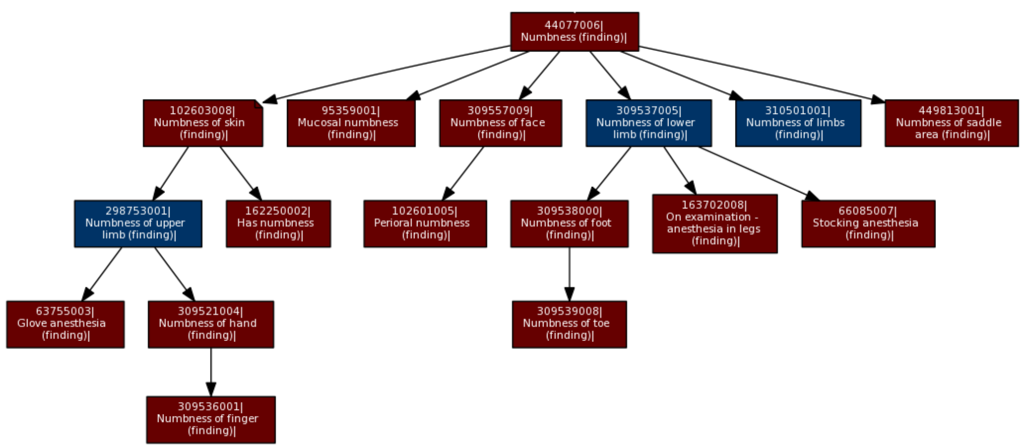


Figure 8-23. Plotting the concept "44077006|Numbness (finding)|" and its subtype concepts.

A closer look at the defining attributes indicates that all the **363698007|Finding site (attribute)|** refers to one of the subtypes of **400199006|Structure of skin and/or surface epithelium (body structure)|**. Therefore by right there should be subsumption between **310501001|Numbness of limbs (finding)|** and the concepts **298753001|Numbness of upper limb (finding)|** and **309537005|Numbness of lower limb (finding)|** (refer to **Table 8-32** and **Figure 8-24**).

Table 8-32. Finding site of concepts "102603008|Numbness of skin (finding)|", "309557009|Numbness of face (finding)|", "310501001|Numbness of limbs (finding)|", "298753001|Numbness of upper limb (finding)|" and "309537005|Numbness of lower limb (finding)|".

102603008|Numbness of skin (finding)|	309557009|Numbness of face (finding)|	310501001|Numbness of limbs (finding)|	298753001|Numbness of upper limb (finding)|	309537005|Numbness of lower limb (finding)|
400199006|Structure of skin and/or surface epithelium (body structure)|	73897004|Skin structure of face (body structure)|	116370005|Skin structure of extremity (body structure)|	371311000|Skin structure of upper extremity (body structure)|	371304004|Skin structure of lower extremity (body structure)|

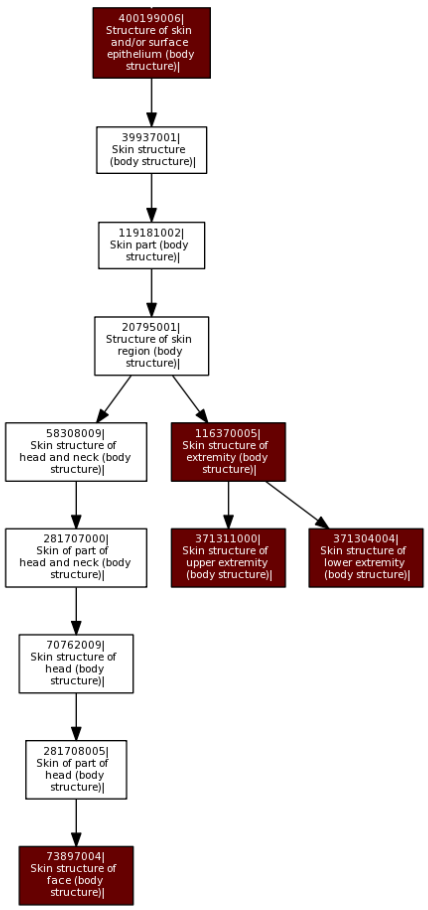


Figure 8-24. Location of concepts "400199006|Structure of skin and/or surface epithelium (body structure)|", "73897004|Skin structure of face (body structure)|", "116370005|Skin structure of extremity (body structure)|", "371311000|Skin structure of upper extremity (body structure)|" and "371304004|Skin structure of lower extremity (body structure)|" in the hierarchy.

8.5.4.7 Inactive Concepts

With each release of SNOMED CT, hundreds to thousands of concepts are inactivated (refer to **Table 8-33**). Therefore when constructing queries, they have to be sophisticated enough to be able to retrieve historical concepts.

Table 8-33. Number of inactive concepts from the past 10 SNOMED CT release versions.

Release Version	Jul 2008	Jan 2009	Jul 2009	Jan 2010	Jul 2010	Jan 2011	Jul 2011	Jan 2012	Jul 2012	Jan 2013
Inactive Concepts	94,201	94,861	98,391	98,817	99,097	99,304	99,491	99,593	99,872	99,985
Difference (+/-)	867	660	3,530	426	280	207	187	102	279	113

The historical relationships (subtype concepts of **410663007|Concept history attribute (attribute)|** used in conjunction with **CharacteristicType=0**) can be used to retrieve active concepts from inactive concepts. **Table 8-34** lists the inactive concepts that are linked to the active concept **73211009|Diabetes mellitus (disorder)|**.

Table 8-34. Inactive concepts linked to 73211009|Diabetes mellitus (disorder)|.

No	Inactive Concept	Concept Status	Historical Relationship
1.	267467004|Diabetes mellitus (& [ketoacidosis]) (disorder)|	Ambiguous	149016008|MAY BE A (attribute)|
2.	154671004|Diabetes mellitus (& [ketoacidosis]) (disorder)|	Ambiguous	149016008|MAY BE A (attribute)|
3.	191044006|Diabetes mellitus (disorder)|	Duplicate	168666000|SAME AS (attribute)|
4.	111553002|Diabetes mellitus associated with unlisted condition (disorder)|	Retired without a stated reason	370124000|REPLACED BY (attribute)|
5.	190324002|Diabetes mellitus NOS with no mention of complication (disorder)|	Limited	159083000|WAS A (attribute)|
6.	190421006|Diabetes mellitus NOS with other specified manifestation (disorder)|	Limited	159083000|WAS A (attribute)|
7.	154674007|Diabetes mellitus with no mention of complication (disorder)|	Limited	159083000|WAS A (attribute)|
8.	190321005|Diabetes mellitus with no mention of complication (disorder)|	Limited	159083000|WAS A (attribute)|
9.	190417004|Diabetes mellitus with other specified manifestation (disorder)|	Limited	159083000|WAS A (attribute)|
10.	267380002|Diabetes mellitus, adult onset, with no mention of complication (disorder)|	Limited	159083000|WAS A (attribute)|
11.	190419001|Diabetes mellitus, adult onset, with other specified manifestation (disorder)|	Limited	159083000|WAS A (attribute)|
12.	267379000|Diabetes mellitus, juvenile type, with no mention of complication (disorder)|	Limited	159083000|WAS A (attribute)|
13.	190418009|Diabetes mellitus, juvenile type, with other specified manifestation (disorder)|	Limited	159083000|WAS A (attribute)|
14.	190323008|Diabetes mellitus: [adult onset, with no mention of complication] or [maturity onset] or [non-insulin dependent] (disorder)|	Ambiguous	149016008|MAY BE A (attribute)|
15.	190322003|Diabetes mellitus: [juvenile type, with no mention of complication] or [insulin dependent] (disorder)|	Ambiguous	149016008|MAY BE A (attribute)|
16.	190336008|Other specified diabetes mellitus with coma (disorder)|	Limited	159083000|WAS A (attribute)|
17.	190382000|Other specified diabetes mellitus with multiple complications (disorder)|	Limited	159083000|WAS A (attribute)|
18.	190420007|Other specified diabetes mellitus with other specified complications (disorder)|	Limited	159083000|WAS A (attribute)|
19.	190383005|Unspecified diabetes mellitus with multiple complications (disorder)|	Limited	159083000|WAS A (attribute)|

8.5.4.8 Different Encoding Methods

Different encoding methods can also cause challenges when retrieving concepts. For example, "childhood asthma" can be represented by at least two pre-coordinated concepts: **161527007|History of - asthma (situation)|** or

233678006|Childhood asthma (disorder)|. The defining attributes of both concepts are shown in Figure 8-25. Aside from the difference in which the primary condition "asthma" is defined, **161527007|History of - asthma (situation)|** uses the defining attributes of **408731000|Temporal context (attribute)|=410513005|In the past (qualifier value)|** to indicate the historical context while **233678006|Childhood asthma (disorder)|** uses the defining attributes of **246454002|Occurrence (attribute)|=255398004|Childhood (qualifier value)|** to indicate the childhood occurrence. While the use of either concept may be acceptable, the two concepts are unrelated. Therefore retrieving one will not retrieve the other. Another challenge with using **255398004|Childhood (qualifier value)|** is that it can only be applied to **404684003|Clinical finding (finding)|** and not to **71388002|Procedure (procedure)|**. Therefore if we wish to encode a phrase such as "tonsillectomy during childhood," we will be unable to post-coordinate the expression.

161527007|History of - asthma (situation)|	233678006|Childhood asthma (disorder)|
243796009|Situation with explicit context (situation)|: {246090004|Associated finding (attribute)|=(195967001|Asthma (disorder)|: {363698007|Finding site (attribute)|= 20139000|Structure of respiratory system (body structure)|}), 408729009|Finding context (attribute)|= 410515003|Known present (qualifier value)|, 408731000|Temporal context (attribute)|= **410513005|In the past (qualifier value)|**, 408732007|Subject relationship context (attribute)|= 410604004|Subject of record (person)|}	243796009|Situation with explicit context (situation)|: {246090004|Associated finding (attribute)|=(233678006|Childhood asthma (disorder)|: 246454002|Occurrence (attribute)|= 255398004|Childhood (qualifier value)| {116676008|Associated morphology (attribute)|= 26036001|Obstruction (morphologic abnormality)|, 363698007|Finding site (attribute)|= 955009|Bronchial structure (body structure)|}), 408729009|Finding context (attribute)|= 410515003|Known present (qualifier value)|, 408731000|Temporal context (attribute)|= 410512000|Current or specified time (qualifier value)|, 408732007|Subject relationship context (attribute)|= 410604004|Subject of record (person)|}

Figure 8-25. Comparing the long normal forms of "161527007|History of - asthma (situation)|" and "233678006|Childhood asthma (disorder)|".

A second example is encoding "at risks." For example, "at risk of diabetes mellitus" can be represented using a pre-coordinated concept **161641009|At risk of diabetes mellitus (finding)|** and a post-coordinated expression **243796009|Situation with explicit context (situation)|:246090004|Associated finding (attribute)|=73211009|Diabetes mellitus (disorder)|,408729009|Finding context (attribute)|=410519009|At risk context (qualifier value)|**. The challenge of using the pre-coordinated concept is that it does not include any defining attributes that links it back to **73211009|Diabetes mellitus (disorder)|**. All subtype concepts of **281694009|Finding of at risk (finding)|** are primitive and do not have any defining attributes aside from **363714003|Interprets (attribute)|=278844005|General clinical state (observable entity)|**. On the other hand, the post-coordinated expression links back to **73211009|Diabetes mellitus (disorder)|** using **246090004|Associated finding (attribute)|=73211009|Diabetes mellitus (disorder)|** and uses the **408729009|Finding context (attribute)|** of **410519009|At risk context (qualifier value)|**. Only three concepts are defined with **410519009|At risk context (qualifier value)|**: **398819009|Diabetic foot at risk (situation)|**, **443999008|Risk of exposure to communicable disease (situation)|** and **444237009|Risk of exposure to Leptospira (situation)|**. There is no general supertype concept for these three "at risk" situation concepts, but they are all subtypes of **413350009|Finding with explicit context (situation)|**. However, any concept from the **404684003|Clinical finding (finding)|** can be qualified using **410519009|At risk context (qualifier value)|**. On the other hand, **281694009|Finding of at risk (finding)|** contains 162 subtype concepts but there is no precedent on how to encode other at risk diseases as none of the subtype concepts have defining attributes that link back to the actual finding or disorder that is at risk.

While it is interesting that there is a concept for **398819009|Diabetic foot at risk (situation)|** that is a situation and uses **410519009|At risk context (qualifier value)|** as a defining attribute, it is also noteworthy that there are eight other concepts that define other diabetic foot at risk from the **404684003|Clinical finding (finding)|** hierarchy, and they are not always related. As shown in Figure 8-26, the concepts **308106006|On examination - Left diabetic foot at risk (disorder)|** and **308105005|On examination - Right diabetic foot at risk (disorder)|** and unrelated to **398819009|Diabetic foot at risk (situation)|**. In addition, they are not supertypes of concepts such as **394675000|On examination - Left diabetic foot at low risk (finding)|** and **394672002|On examination - Right diabetic foot at high risk (finding)|**.

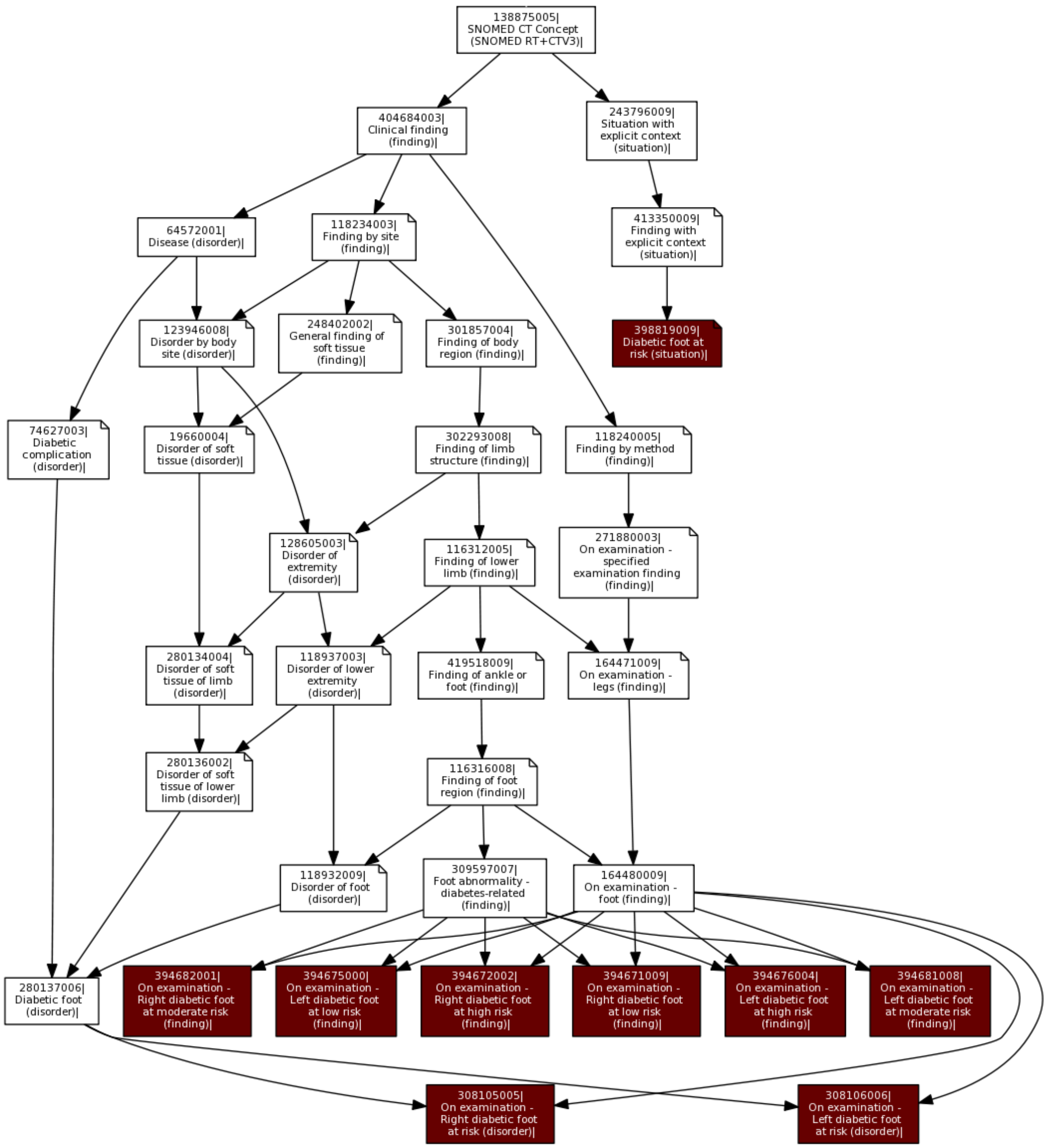


Figure 8-26. Plotting out the concepts that refer to diabetic foot at risk.

8.6 References

[1] Lee, D., de Keizer, N., Lau, F., & Cornet, R. (2013). Literature review of SNOMED CT use. Journal of the American Medical Informatics Association. http://www.ncbi.nlm.nih.gov/pubmed/23828173.

[2] Arnot‐Smith, J., & Smith, A. F. (2010). Patient safety incidents involving neuromuscular blockade: analysis of the UK National Reporting and Learning System data from 2006 to 2008. Anaesthesia, 65(11), 1106-1113. http://www.ncbi.nlm.nih.gov/pubmed/20840604.

[3] Hussain, F., Muller, F., Husain, E. (2010). Under-reporting of invasive malignant melanomas in North East of Scotland. British Journal of Dermatology 2010; 163: 67-

[4] Benoit, S. R., L Clifford McDonald, M. D., English, R., & Tokars, J. I. (2011). Automated surveillance of Clostridium difficile infections using BioSense. Infection Control and Hospital Epidemiology, 32(1), 26-33.

[5] Nguyen, A., Moore, J., Zuccon, G., Lawley, M., & Colquist, S. (2012). Classification of pathology reports for cancer registry notifications. Studies in health technology and informatics, 178, 150-6. http://www.ncbi.nlm.nih.gov/pubmed/22797034.

[6] Liaw, S. T., Chen, H. Y., Maneze, D., Taggart, J., Dennis, S., Vagholkar, S., & Bunker, J. (2012). Health reform: Is routinely collected electronic information fit for purpose?. Emergency Medicine Australasia, 24(1), 57-63. http://www.ncbi.nlm.nih.gov/pubmed/22313561.

[7] Matheny, M. E., FitzHenry, F., Speroff, T., Green, J. K., Griffith, M. L., Vasilevskis, E. E., ... & Brown, S. H. (2012). Detection of infectious symptoms from VA emergency department and primary care clinical documentation. International journal of medical informatics, 81(3), 143-156.http://www.ncbi.nlm.nih.gov/pubmed/22244191.

[8] Koopman, B., Bruza, P., Sitbon, L., & Lawley, M. (2012). Towards semantic search and inference in electronic medical records: An approach using concept-based information retrieval. The Australasian medical journal, 5(9), 482-8. http://www.ncbi.nlm.nih.gov/pubmed/23115582.

[9] Rector, A. L., & Brandt, S. (2008). Why do it the hard way? The case for an expressive description logic for SNOMED. Journal of the American Medical Informatics Association, 15(6), 744-751. http://www.ncbi.nlm.nih.gov/pubmed/18755993.

[10] Spackman, K. A. (2001). Normal forms for description logic expressions of clinical concepts in SNOMED RT. In Proceedings of the AMIA Symposium (p. 627-31). American Medical Informatics Association.

[11] International Health Terminology Standards Development Organisation. Technical Implementation Guide. http://www.ihtsdo.org/fileadmin/user_upload/doc/tig/index.html.

[12] Public Health Agency of Canada. Chronic Diseases. http://www.phac-aspc.gc.ca/cd-mc/index-eng.php. Last accessed: August 4, 2013.

This page is intentionally left blank.

9. Towards Demonstrating the Clinical Value of SNOMED CT

9.1 Introduction

The Systematised Nomenclature of Medicine Clinical Terms (SNOMED CT) is a clinical reference terminology that has a goal of improving quality and safety of care.[1] While over 20 countries[2] have selected SNOMED CT as the preferred choice of a clinical reference terminology, there are few studies published to date on how SNOMED CT should be implemented and used so as to demonstrate clinical value in clinical settings. The majority of articles on SNOMED CT have centred on theoretical work such as describing the different components of a reference terminology, comparing SNOMED CT to other standardised terminologies and mapping local codes to SNOMED CT.[3] Benefits, both realised and anticipated, include improved quality of documentation,[4] improved efficiency and consistency of encoding,[5] reduced time and costs for transcribing, post-coding and quality management,[5, 6, 7] ability to conduct biosurveillance monitoring,[8] ability to audit patient records,[9] support patient case queries,[5] support integration with clinical practice guidelines[17] and enable international benchmarking.[35] While there are stated benefits such as improving patient safety[10] and facilitating decision support systems,[11, 12] these benefits have not been quantified.

In this study, the incremental value of using SNOMED CT in primary health care is demonstrated from multiple perspectives. This study contributes to SNOMED CT knowledge by showing how: (1) SNOMED CT concepts can be incorporated into published inference methods to identify patients with specific health conditions; (2) SNOMED CT semantic features such as subsumption and defining relationships can be used to create SNOMED CT-based query tools; and (3) SNOMED CT may be implemented in ways that add tangible value in clinical settings. For this study, the SNOMED CT Clinical Value Framework, which was described in **Chapter Five**, was used to categorise the different uses of SNOMED CT. The framework is shown in **Table 9-1.**

Table 9-1. SNOMED CT Clinical Value Framework with examples of SNOMED CT use.

Data Use / Data Use Target Audience	Patient	Practice	Population
Data Capture	▪ Suggest billing diagnostic code based on encounter diagnoses ▪ Suggest encounter diagnoses based on billing diagnostic code ▪ Suggest adding encounter diagnosis to problem list	▪ Audit (critique) completeness of problem list and encounter diagnoses ▪ Audit (critique) accuracy of billing diagnostic codes	▪ Compare the completeness of problem lists, encounter diagnoses and billing diagnostic codes between practices and by population
Data Retrieval	▪ Alert clinicians to potential drug-allergy interactions ▪ Alert clinicians to out-of-range values ▪ Remind clinicians to order routine tests ▪ Suggest medications based on out-of-range test results	▪ Audit (critique) evidence-based guideline adherence ▪ Patient case queries and patient recall ▪ Generate statistics on key primary health care indicators	▪ Compare clinical outcomes between practices and by population

Data Use / Data Use Target Audience	Patient	Practice	Population
Data Sharing	▪ Generate standardised patient clinical summary in the form of the Continuity of Care Document (CCD) for referrals that are encoded with SNOMED CT	▪ Generate standardised practice quality reports in the form of the Quality Reporting Document Architecture (QDRA) Category II/ III that are encoded with SNOMED CT	▪ Generate standardised practice quality reports in the form of the Quality Reporting Document Architecture (QDRA) Category III that are encoded with SNOMED CT

9.2 Materials

The materials used in this study included: (1) an anonymised dataset; (2) SNOMED CT and cross map; (3) Continuity of Care Document (CCD); (4) diabetes mellitus, the clinical domain of interest; and (5) reference papers.

9.2.1 Anonymised Dataset

The dataset was extracted from a Canadian primary care clinic and contained 3,298 active patients (i.e., patients who had an appointment within the last year). The data ranged from June 27, 2001 to December 20, 2012 and included appointments, problem lists, encounter diagnoses, results, billing diagnostic codes, medications and examinations. A description of each table, the number of records and number of patients are shown in **Table 9-2**. The problem lists and encounter diagnoses were encoded with SNOMED CT as described in **Chapter Seven**.

Table 9-2. Description of tables and records extracted from the primary care EMR.

No	Table	Description	Records	Patients
1.	Patient	Lists the patients in the EMR. Only the patient id was extracted.	13,013	13,013
2.	Appointments	Records the appointments.	353,169	11,675
3.	Encounters Diagnoses	Records the problems specific to an appointment.	266,032	9,486
4.	Problems Lists	Records the problems not specific to an appointment.	20,131	4,894
5.	Medications	Records the medications that were prescribed.	73,904	6,273
6.	Results	Records the laboratory results that are pending and completed.	1,173,449	6,109
7.	Billing	Records the billing codes that were assigned to an appointment.	188168	8,313
8.	Examination	Records the systolic and diastolic blood pressure, height and weight.	25,218	4,425

9.2.2 SNOMED CT and Cross Map

The National Library of Medicine (NLM) currently publishes two cross maps for the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM). The first is from ICD-9-CM to SNOMED CT, which is a map to convert ICD-9-CM diagnostic codes to SNOMED CT and is based on the July 2011 release version of SNOMED CT and 2012 version of ICD-9-CM.[a] The second is from SNOMED CT to ICD-9-CM, which is a rule based map to support reimbursement and is based on the July 2008 release version of SNOMED CT and 2008 version of ICD-9-CM.[b] Although it is not a complete cross map and only includes about 5,000 mappings, it was

[a] http://www.nlm.nih.gov/research/umls/mapping_projects/icd9cm_to_snomedct.html
[b] http://www.nlm.nih.gov/research/umls/mapping_projects/snomedct_to_icd9cm_reimburse.html

sufficient for this study as the SNOMED CT concepts and ICD-9-CM codes for diabetes are covered in the map. This map was used in this study to demonstrate the automatic generation of ICD-9-CM billing codes from encounter diagnoses. The relevant cross map and cross map targets are shown in **Table 9-3**.

Table 9-3. Snapshot of the SNOMED CT to ICD-9-CM cross map for four concepts of diabetes.

No	Map Concept Id	Map Option	Map Priority	Map Advice	Target Id	Target Codes
1.	73211009|Diabetes mellitus (disorder)|	1	1	Otherwise	4562056	250.00 - Diabetes mellitus without mention of complication, type II or unspecified type, not stated as uncontrolled
2.	46635009|Diabetes mellitus type 1 (disorder)|	1	1	Otherwise	4563053	250.01 - Diabetes mellitus without mention of complication, type I [juvenile type], not stated as uncontrolled
3.	44054006|Diabetes mellitus type 2 (disorder)|	1	1	Otherwise	4562056	250.00 - Diabetes mellitus without mention of complication, type II or unspecified type, not stated as uncontrolled
4.	11687002|Gestational diabetes mellitus (disorder)|	1	1	Otherwise	6690053	648.80 - Abnormal glucose tolerance of mother, unspecified as to episode of care or not applicable

9.2.3 Continuity of Care Document

The CCD is a health data standard based on the Health Level Seven (HL7) Clinical Document Architecture (CDA). It is an extensible markup language (XML) specification for creating clinical summaries and contains 16 sections[c] (refer to **Table 9-4**). The encoding results of the problem list and encounter diagnoses indicated that the SNOMED CT concepts used were from multiple hierarchies, some of which were not considered problems including allergies, vaccinations, procedures as well as personal and family history. The CCD sections were used to review the problem list and encounter diagnoses to determine if the information entered was appropriate.

Table 9-4. Sections in the Continuity of Care Document.

No	Continuity of Care Document Section	Description
1.	Summary Purpose	Contains data on the patient's payers and is used to define which entity is the responsible fiduciary for the financial aspects of a patient's care.
2.	Payers	Contains data defining the patient's advance directives and any reference to supporting documentation.
3.	Advance Directive	Contains instructions should the patient no longer be able to make them due to illness or incapacitation.
4.	Functional Status	Describes the patient's status of normal functioning at the time the Care Record was created.
5.	Problems	Lists and describes all relevant clinical problems at the time the summary is generated.
6.	Family History	Contains data defining the patient's genetic relatives in terms of possible or relevant health risk factors that have a potential impact on the patient's healthcare risk profile.
7.	Social History	Contains data defining the patient's occupational, personal social, and environmental history and health risk factors, as well as administrative data such as marital status, race, ethnicity and religious affiliation.
8.	Allergies, Adverse Reactions and Alerts	Lists and describes any allergies, adverse reactions, and alerts that are pertinent to the patient's current or past medical history.

[c] http://www.hl7.org/implement/standards/product_brief.cfm?product_id=6

No	Continuity of Care Document Section	Description
9.	Medications	Defines a patient's current medications and pertinent medication history.
10.	Medical Equipment	Defines a patient's implanted and external medical devices and equipment that their health status depends on, as well as any pertinent equipment or device history.
11.	Immunisations	Defines a patient's current immunisation status and pertinent immunisation history. The primary use case is to enable communication of a patient's immunisation status.
12.	Vital Signs	Contains current and historically relevant vital signs, such as blood pressure, heart rate, respiratory rate, height, weight, body mass index, head circumference, crown-to-rump length, and pulse oximetry.
13.	Results	Contains the results of observations generated by laboratories, imaging procedures, and other procedures.
14.	Procedures	Defines all interventional, surgical, diagnostic, or therapeutic procedures or treatments pertinent to the patient historically at the time the document is generated.
15.	Encounters	List and describes any healthcare encounters pertinent to the patient's current health status or historical health history.
16.	Plan of Care	Contains data defining pending orders, interventions, encounters, services, and procedures for the patient.

9.2.4 Clinical Domain of Interest

Diabetes mellitus was used throughout this study to demonstrate the different uses of SNOMED CT and the value-added aspects. According to the American Diabetes Association, diabetes is "a group of metabolic diseases characterised by hyperglycaemia resulting from defects in insulin secretion, insulin action, or both."[13] Diabetes was selected for four reasons. First, diabetes has a high prevalence. In Canada, approximately 2.4 million individuals (6.8%) have been diagnosed with diabetes.[14] The dataset used in this study contained about 3,300 active patients so a disease with a relatively high prevalence was needed. Second, diabetes is one of the most documented chronic diseases in terms of the disease, risk factors, complications and treatments. Third, relevant concepts, codes and tests already exist in current standardised terminologies. Fourth, diabetes can be identified by laboratory results, medications, problem lists, encounter diagnoses, and billing diagnostic codes, which enable the disease to be analysed from different approaches.

9.2.5 Reference Papers

Two papers which contained diabetes identification algorithms and care guidelines were used as references. The papers were "A method and knowledge base for automated inference of patient problems from structured data in an electronic medical record" by Wright, et al.,[15] and "Diabetes flow sheets use associated with guideline adherence" by Hahn, et al.[16]

Wright's Criteria. In this paper, Wright, et al.,[15] developed a set of rules based on billing codes, laboratory results, medications and physical examinations to assess the completeness of a patient's problem list for 17 clinical conditions. While there are other papers that have focused on the completeness of the problem list, they were usually limited to only medications.[17, 18, 19] In this study, only diabetes was used and the rules are shown in **Table 9-5**. As the medications used in this study were encoded with SNOMED CT, instead of using the enhanced therapeutic class

(ETC), SNOMED CT concepts were used in the medication definition. While Wright, et al.,[15] looked only at whether or not three SNOMED CT concepts occurred in the problem list (i.e., **73211009|Diabetes mellitus (disorder)|**, **9414007|Impaired glucose tolerance (disorder)|** or **21584002|Syndrome of infant of diabetic mother (disorder)|**), this study also included the subtypes of the three concepts. For example, **46635009|Diabetes mellitus type 1 (disorder)|**, which was a subtype of **73211009|Diabetes mellitus (disorder)|** was considered a match. Therefore instead of looking for three concepts that did not occur in the problem list, a total of 113 concepts were checked (**73211009|Diabetes mellitus (disorder)|** (n=92), **9414007|Impaired glucose tolerance (disorder)|** (n=20), **21584002|Syndrome of infant of diabetic mother (disorder)|** (n=1)).

Table 9-5. Criteria used to identify patients with diabetes mellitus.

Criteria
If patients meets at least one of these definitions ▪ Definition 1 (at least one laboratory result) - Hemoglobin A1c/Hemoglobin.total in Blood - (LOINC: 4548-4) ≥7 - Hemoglobin A1c/Hemoglobin.total in Blood by HPLC - (LOINC: 17856-6) ≥7 - Hemoglobin A1c/Hemoglobin.total in Blood by Electrophoresis - (LOINC: 4549-2) ≥7 ▪ Definition 2 (at least two billing codes) - Diabetes mellitus - (ICD-9: 250) or one of its subtypes ▪ Definition 3 (at least one medication) - 384953001|Antidiabetic preparation (product)| - 373245004|Antidiabetic agent (substance)|.
And do not have any of these problems on their problem list ▪ 21584002|Syndrome of infant of diabetic mother (disorder)| ▪ 9414007|Impaired glucose tolerance (disorder)| ▪ 73211009|Diabetes mellitus (disorder)|
Then suggest adding these problems: ▪ Default problem - 73211009|Diabetes mellitus (disorder)| ▪ Additional options as related terms - 44054006|Diabetes mellitus type 2 (disorder)| - 46635009|Diabetes mellitus type 1 (disorder)| - 9414007|Impaired glucose tolerance (disorder)| - 420422005|Ketoacidosis in diabetes mellitus (disorder)| - 68256003|Increased glucose level (finding)| - 11687002|Gestational diabetes mellitus (disorder)|

Hahn's Criteria. There are many flow sheets available that list the process of caring for patients diagnosed with diabetes. These include identifying risk factors, laboratory tests, medications and screening guidelines. As these flow sheets can include over 30 data elements, it is neither feasible nor practical to report on the adherence of individual data elements. The paper by Hahn, et al.,[16] succinctly compiled a list of the most common and relevant care guidelines[20, 21] and grouped them into three sets (i.e., assessment scores; treatment scores; and target attainment scores). The assessment scores assess the documentation of five types of care guidelines. The treatment scores assess the results of four laboratory tests, examinations and medications. The target attainment scores assess the attainment of two different laboratory test results and blood pressure goals. Refer to **Table 9-6** for the criteria in each set. As the low-density lipoprotein and urine microalbumin test results were recorded as mmol/L and mg/mmol respectively in the primary care dataset, it was simpler to change the care guidelines criteria to those units rather than vice versa. The low-density lipoprotein result was converted from mg/dL to mmol/L by dividing the target result of 100 by 38.66976, which was 2.59.[d] The urine microalbumin test was converted from mg/g to mg/mmol by multiplying the target result of 30 by the factor of 0.113,[e] which was 3.39.

The representation of clinical care guidelines with SNOMED CT to provide decision support is not new.[22, 23, 24] In this study, SNOMED CT was used in three areas. First, to identify patients who had diabetes in either the problem list or encounter diagnoses at least one year prior (i.e., December 20, 2011) to the last date in the dataset. The one year duration was necessary as the maximum length for two of the guidelines was one year (refer to **Table 9-6**). Second, SNOMED CT was used to identify the smoking status of patients as part of the assessment score. The SNOMED CT concepts used as were **365981007|Finding of tobacco smoking behavior (finding)|** and **308512009|Smoking monitoring status (finding)|**. Third, SNOMED CT was used to capture the medications prescribed as part of the treatment scores. Only the active ingredient in the medication was captured and not the dosage.

Table 9-6. Assessment of adherence to care guidelines by Hahn KA.

No	Assessment of Adherence to Care Guidelines
1.	Diabetes **assessment scores** were based on documentation in the medical record of five assessments (20 points each): ▪ Haemoglobin A1c level tested in the past six months ▪ Urine microalbumin level tested in the past year ▪ Low-density lipoprotein (LDL-cholesterol) level tested in the past year ▪ Smoking status ▪ Blood pressure
2.	Diabetes **treatment scores** consisted of documentation in the medical record of 4 measures (25 points each): ▪ LDL-cholesterol below 100 mg/dL (**2.59 mmol/L**) or use of lipid-lowering agent ▪ Haemoglobin A1c level at or below 8% or use of a hypoglycemic agent ▪ Blood pressure at or below 130/85 mm Hg or use of an antihypertensive agent ▪ Urine microalbumin level below 30 mg/g (**3.39 mg/mmol**) creatine or use of an angiotensin-converting enzyme (ACE) inhibitor or angiotensin II receptor blocker (ARB)
3.	Diabetes **target attainment scores** consisted of documentation in the medical record of attainment of 3 targets (33.3 points each): ▪ Haemoglobin A1c level less than 8% ▪ LDL-cholesterol level less than or equal to 100 mg/dL (**2.59 mmol/L**) ▪ Blood pressure less than or equal to 130/85 mm Hg

9.3 Methods

The methods used in this study can be grouped into four categories. First, the methods used to import the primary care dataset into a data repository. Second, the guidelines that were used to demonstrate the clinical value of SNOMED CT. Third, the prototyping method that was used to develop an EMR system prototype. Fourth, the method that was used to solicit feedback on how the guidelines were used and the demonstration of the prototype.

9.3.1 Data Repository

Before importing the data into the repository, four data standardisation, data cleaning and encoding processes had to take place. The database schema for the data repository is shown in **Figure 9-1**.

[d] http://www.onlineconversion.com/cholesterol.htm
[e] http://www.aruplab.com/guides/ug/tests/0050203.jsp

First, the patient id that was used in the primary care EMR (EMR internal patient id, not the government issued primary health number (PHN)) was hashed using a one-way hashing algorithm, SHA-256 (secure hash algorithm), and a passphrase to ensure privacy.

Second, the examination table was standardised. The original examination table recorded systolic and diastolic blood pressure records using an **Exam** field to record the name of the examination and a **TextResult** field to record the examination finding. These were transformed into an examination table, which had discrete data elements for systolic and diastolic pressures. This was done to simplify the data retrieval. In some cases the systolic field contain both the systolic and diastolic values (e.g., 90/60) and had to be separated into discrete data elements.

Third, the billing diagnostic codes were normalised. The original billing diagnostic table contained three data elements per row that could store ICD-9-CM billing codes (i.e., **ICD9Code**, **ICD9Code2**, **ICD9Code3**). These were normalised into individual records. This was done to simplify the data retrieval. In addition, the decimal place for the ICD-9-CM codes were added back where applicable (e.g., from **25001** to **250.01**).

Fourth, the laboratory results were standardised and consisted of three sub-steps. First, the coding systems used to record test names were identified. Even though the EMR system extracted the results from HL7 messages and imported it into a relational database format, there was no indication of the code system identifier in the table. It was eventually determined that two main coding systems were used: the Logical Observation Identifiers Names and Codes (LOINC)[f] and the Current Procedural Terminology (CPT).[g] Second, the relevant laboratory tests that were used in this study were identified by searching the description of the test name data element and the coding systems. The two other coding systems were then mapped to LOINC. Only the relevant laboratory test results were standardised and normalised. Third, test results were checked to ensure the units and data types were the same. Numeric data stored using the HL7 data type of ST (string) was converted to NM (numeric). As the test results data element could contain an interpretation or commentary of the results, it was necessary to parse the result data element and split the result value from the narrative. For example, for the LOINC test **39469-2**, the result value was "3.04
The LDL therapeutic targets for low, moderate and high risk individuals are less than 5.0, 3.5 and 2.0 mmol/L respectively. See Can. J. Cardiol. 2006:22(11): 913-27." In this case, everything after the test value 3.04 was removed and the data type was changed from FT (free text) to NM. As the result units for the same test was not always consistent (e.g., percent vs decimal), they were converted to a standardised unit along with their result values (e.g., Haemoglobin A1c decimal value (0.008) was converted to percent (8%)).

---

[f] http://loinc.org

[g] http://www.ama-assn.org/ama/pub/physician-resources/solutions-managing-your-practice/coding-billing-insurance/cpt.page

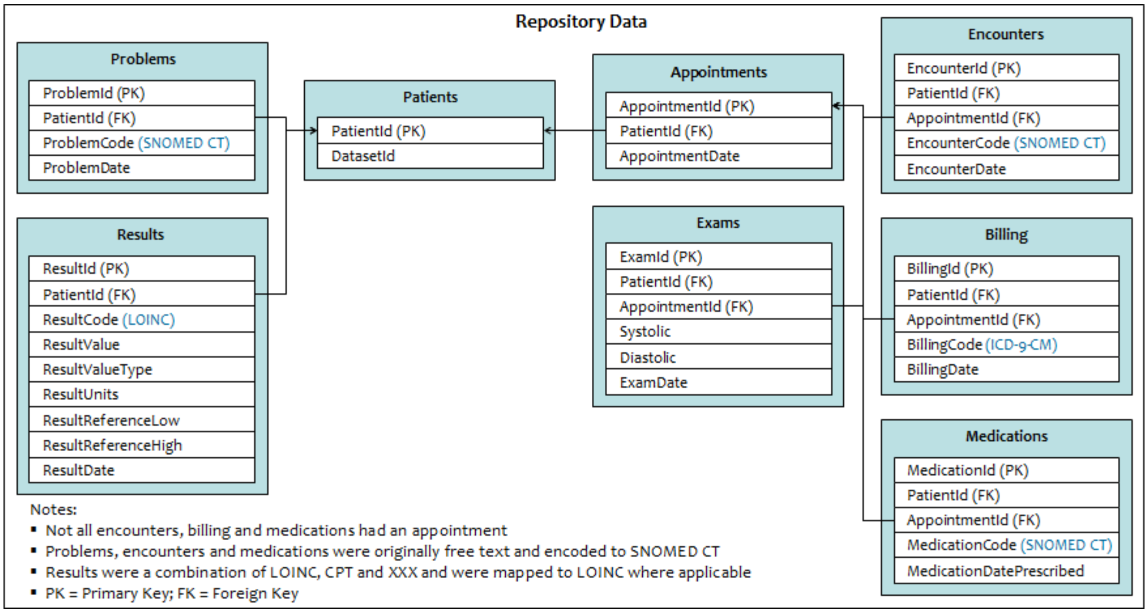


Figure 9-1. Data repository database schema used in this study.

9.3.2 Clinical Value Methods

Four methods were used to demonstrate the value of SNOMED CT. The methods, Clinical Value Targets and relevant data elements are shown in **Table 9-7**.

Table 9-7. Methods of demonstrating the value of SNOMED CT and the data elements used.

No	Method	Clinical Value Level Targets	Data Use	Problem	Encounter	Results	Billing	Medication	Examination
1.	Completeness of problem list	▪ Enter problem list	▪ Data Capture	✓		✓	✓	✓	
2.	Completeness of encounter diagnoses and billing diagnostic codes	▪ Enter encounter notes	▪ Data Capture		✓		✓		
3.	Completeness of clinical summary	▪ Enter allergies ▪ Enter immunisations ▪ Access hospital clinical reports and encounter summaries from the EMR	▪ Data Capture ▪ Data Sharing ▪ Data Retrieval	✓	✓				
4.	Adherence of care guidelines	▪ Generate automated alerts from within the EMR ▪ Generate automated reminders from within the EMR	▪ Data Retrieval	✓	✓	✓		✓	✓

The first and fourth methods were based on two studies that demonstrated how clinical conditions could be inferred and how adherence to care guidelines using diabetes flow sheets could be determined. The second method compared the encounter diagnoses and billing diagnostic codes for completeness using the SNOMED CT to ICD-9-CM cross map. The third method used the CCD to determine if information entered into the problem list and encounter diagnosis could be better stored elsewhere in a clinical record. One or more SNOMED CT predicate expressions were assigned for inclusion and/or exclusion to each of the CCD sections where applicable and queries were executed against the problem list and encounter diagnoses.

9.3.3 Prototyping

Prototyping was used to develop an EMR system prototype to demonstrate the different clinical value of using SNOMED CT. The prototype was a web-based application that was developed using PHP, a server-side scripting programming language, JavaScript, a client-side scripting programming language, and MySQL, a relational database management system. While it could be argued that PHP and JavaScript are not typical prototyping languages, they were sufficient to develop the prototype quickly. As more clinical value aspects of using SNOMED CT were identified, functions were programmed into the prototype.

9.3.4 Group Interview

Two types of expert feedback sessions were conducted in the form of group interviews. First, a clinical feedback session that focused on obtaining feedback from the perspective of a clinician. Second, a technical feedback session that focused on obtaining feedback from technical users (including data managers, programmers and analysts). Ten individuals (including six clinicians) were invited to the clinical feedback session. Five individuals attended the clinical feedback session and included two physicians, an information and technology manager, a data manager and a terminology analyst. Eight individuals were invited to the technical feedback session. It was attended by the same information and technology manager and data manager, along with another information management consultant. For the feedback sessions, three main open-ended questions (i.e., "does it make sense?", "does it add value?" and "what are the potential implementation issues?") were asked for each of the ways SNOMED CT was used. Notes were taken of the feedback provided and a thematic analysis of the discussion was conducted.

9.4 Results

The results of this study are presented in five parts. First, the results of the encodings and mappings. Second, the results of the incremental value of using SNOMED CT. Third, the description, screenshots and scenarios that were built into the prototype. Fourth, the results of the prototype. Fifth, the results of the feedback sessions with the clinicians and terminology subject matter experts. The limitations and future work of this chapter are described in **Chapter Ten**.

9.4.1 Mapping and Encoding Results

Problem List and Encounter Diagnoses. Of the 20,141 problems, 18,298 (90.8%) were fully encoded, 1,666 (8.3%) were partially encoded, and 177 (0.9%) were unencodeable. Of the 265,310 encounter diagnoses, 231,963 (87.2%) were fully encoded, 27,198 (10.2%) were partially encoded, and 6,149 (2.3%) where unencodeable. Refer to **Table 9-8**. Fully encoded means that the SNOMED CT expression used could fully represent the meaning of the term being encoded. Partially encoded means that SNOMED CT could not adequately represent the full meaning of the term being encoded. For example, only "hepatic failure" from the term "hepatic failure end stage" could be encoded. Additional details about the results can be found in **Chapter Seven**.

Table 9-8. Results of encoding the problem list and encounter diagnoses with SNOMED CT.

No	Encoding	Problem List		Encounter Diagnoses	
		Total	Unique	Total	Unique
1.	**By Records**	20,141	2,881	266,032	15,123
	Fully encoded record	18,298	2,246	231,963	8,541
	Partially encoded record	1,666	582	27,198	5,794
	Fully unencoded record	177	43	6,149	779
	Nothing to encode	1	1	3	1
2.	**By Phrases**				
	Encoded phrases	23,382	2,239	299,300	7,357
	Unencoded phrases	1,930	500	38,048	5,049
3.	**Pre- vs Post-coordination**				
	Pre-coordination	21,114	1,689	270,937	2,895
	Post-coordination	2,268	561	28,363	2,838

Medications. The encoding of medications was two-fold. First, the guidelines for diabetic care that referred to medications were encoded. Second, the medications prescribed to patients were encoded. The results of the encodings used by Wright, et al.,[15] and Hahn, et al.,[16] are shown in **Table 9-9** and **Table 9-10**. The concepts have been plotted out in **Figure 9-2**.

Table 9-9. Medications used to in Wright, et al.,[15] paper.

No	Enhanced Therapeutic Class	SNOMED CT Product	SNOMED CT Substance
1.	Injectable Antidiabetic Agents (ETC: 5886)	384953001|Antidiabetic preparation (product)|	373245004|Antidiabetic agent (substance)|
2.	Oral Antidiabetic Agents (ETC: 154)	384953001|Antidiabetic preparation (product)|	373245004|Antidiabetic agent (substance)|

Table 9-10. Medications used in Hahn, et al.,[16] paper

No	Medication	SNOMED CT Product	SNOMED CT Substance
1.	Lipid-lowering agent	57952007|Antilipemic agent (product)|	373267003|Antilipemic agent (substance)|
2.	Hypoglycemic agent	312064005|Hypoglycemic agent (product)|	373299009|Hypoglycemic agent (substance)|
3.	Antihypertensive agent	1182007|Hypotensive agent (product)|	372586001|Hypotensive agent (substance)|
4.	Angiotensin-converting enzyme (ACE) inhibitor	41549009|Angiotensin-converting enzyme inhibitor agent (product)|	372733002|Angiotensin-converting enzyme inhibitor agent (substance)|
5.	Angiotensin II receptor blocker (ARB)	96308008|Angiotensin II receptor antagonist (product)|	372913009|Angiotensin II receptor antagonist (substance)|

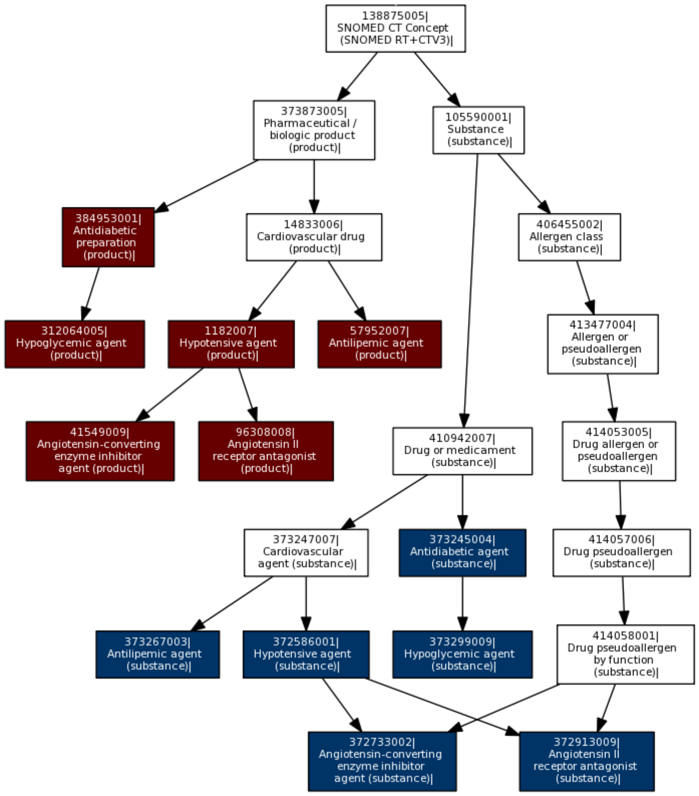


Figure 9-2. Location of the concepts for medications used in this study.

There were 73,904 medication records, 848 (1.1%) of which contain null (blank) records. It is unclear why these records were blank. Of the remaining 73,056 records, 69,513 (95.2%) were encoded with SNOMED CT (refer to **Table 9-11**). In all, 485 unique SNOMED CT expressions were used, including three post-coordinated expressions for 54 records (refer to **Table 9-12**). Medications with a frequency of ≥10 that were not automatically encoded were manually encoded. The active ingredients for medications that only contained a brand name were used to encode the medications. There were nine medications with a frequency of ≥10 that could not be encoded. Seven referred to "blister pack" with no indication of medication, one referred to "Dossette," a medication storage box, and the last one referred to "AeroChamber," a medication inhaler aid, which technically was not a medication.

Table 9-11. Medication encoding results.

No	Type of SNOMED CT Encoding	Total	Unique
1.	Pre-coordination	69,459	19,561
2.	Post-coordination	54	3
3.	Unencoded	3,543	2,666
4.	Nothing to encode	848	1
	Total	**73,904**	**22,231**

Table 9-12. Medications that required post-coordination.

No	Medication	SNOMED CT Expression
1.	Pantoprazole magnesium	373873005|Pharmaceutical / biologic product (product)|: 127489000|Has active ingredient (attribute)|= 105858004|Magnesium AND/OR magnesium compound (substance)|, 127489000|Has active ingredient (attribute)|= 395821003|Pantoprazole (substance)|
2.	Amoxicillin/clarithromycin/lansoprazole	373873005|Pharmaceutical / biologic product (product)|: 127489000|Has active ingredient (attribute)|= 372687004|Amoxicillin (substance)|, 127489000|Has active ingredient (attribute)|= 386888004|Lansoprazole (substance)|, 127489000|Has active ingredient (attribute)|= 387487009|Clarithromycin (substance)|
3.	Multivitamin with iron	373873005|Pharmaceutical / biologic product (product)|: 127489000|Has active ingredient (attribute)|= 373755001|Iron agent (substance)|, 127489000|Has active ingredient (attribute)|= 412250002|Multivitamin agent (substance)|

Laboratory Results. Only the laboratory tests results that were required in this study were cleaned and mapped to LOINC. The patient results contained three different coding systems (i.e., LOINC, CPT and an unidentified coding system). The paper by Wright, et al.,[15] defined three LOINC tests (i.e., **4548-4 – Hemoglobin A1C**, **17856-6 – Hemoglobin A1c/Hemoglobin.total in Blood by HPLC** and **4549-2 – Hemoglobin A1c/Hemoglobin.total in Blood by Electrophoresis**). Out of the three LOINC codes, only **4548-4** occurred in the patient results. The paper by Hahn, et al.,[16] identified three types of laboratory test results (i.e., haemoglobin A1c, urine microalbumin and low-density lipoprotein). As the LOINC codes were not provided, a search of the original description of the test result and a search of LOINC and CPT were conducted. The results of the mapping from CPT and the unidentified coding system to LOINC are shown in **Table 9-13**. In this study, a distinction was not made between the 24 hour period and non-24 hour period tests. Although there were different codes for Albumin/Creatinine and Microalbumin/Creatinine in LOINC, these codes were used interchangeably in the test descriptions in the patient results. For example, the LOINC code **9318-7** had the official description of **Albumin/Creatinine [Mass Ratio] in Urine** but the test descriptions in the dataset were "Urine ACR (Microalbumin/Creatinine Ratio)" and "Urine Microalbumin/Creatinine Ratio."

Table 9-13. Laboratory test codes.

No	Laboratory Test	LOINC Codes	Mapped Codes
1.	Haemoglobin A1c	4548-4 - Hemoglobin A1C	▪ 83020.1 - HEMOGLOBIN A1C ▪ HbA1C - HgA1C
2.	Urine microalbumin	9318-7 - Albumin/Creatinine [Mass Ratio] in Urine	▪ X001 ▪ 82043.1 - Albumin; urine, microalbumin, quantitative
		13705-9 - Albumin/Creatinine [Mass Ratio] in 24 hour Urine	
		14959-1 - Microalbumin/Creatinine [Mass Ratio] in Urine	
		14958-3 - Microalbumin/Creatinine [Mass Ratio] in 24 hour Urine	

No	Laboratory Test	LOINC Codes	Mapped Codes
		30000-4 - Microalbumin/Creatinine [Ratio] in Urine	
3.	Low-density lipoprotein	22748-8 - Cholesterol in LDL [Moles/volume] in Serum or Plasma	
		39469-2 - Cholesterol in LDL [Moles/volume] in Serum or Plasma by calculation	

9.4.2 Demonstrating the Clinical Value of SNOMED CT

This section reports on the results of the four methods applied to demonstrate the clinical value of SNOMED CT by the three data uses.

9.4.2.1 Data Capture

The first data use dimension was data capture. This section reports on how SNOMED CT can be used to help capture information and ensure the completeness of the information captured. The completeness of data capture was demonstrated using two different ways. First, the completeness of problem list using the criteria by Wright, et al.[15] Second, the completeness of encounter diagnoses and billing diagnostic codes.

9.4.2.1.1 Completeness of Problem List

The three definitions by Wright, et al.,[15] to identify diabetic patients netted 274 active patients, 232 (84.7%) of which had "diabetes" (**73211009|Diabetes mellitus (disorder)|**, **46635009|Diabetes mellitus type 1 (disorder)|**, **44054006|Diabetes mellitus type 2 (disorder)|** and **11687002|Gestational diabetes mellitus (disorder)|**) recorded in the problem list (refer to **Figure 9-3**). The patients were identified by medications (n=235), billing diagnostic codes (n=168) and laboratory results (n=110). Of the 274 patients identified, 86 were identified using all three definitions. The subtypes of the medications used in the third definition that were present in the dataset are shown in **Figure 9-4.**

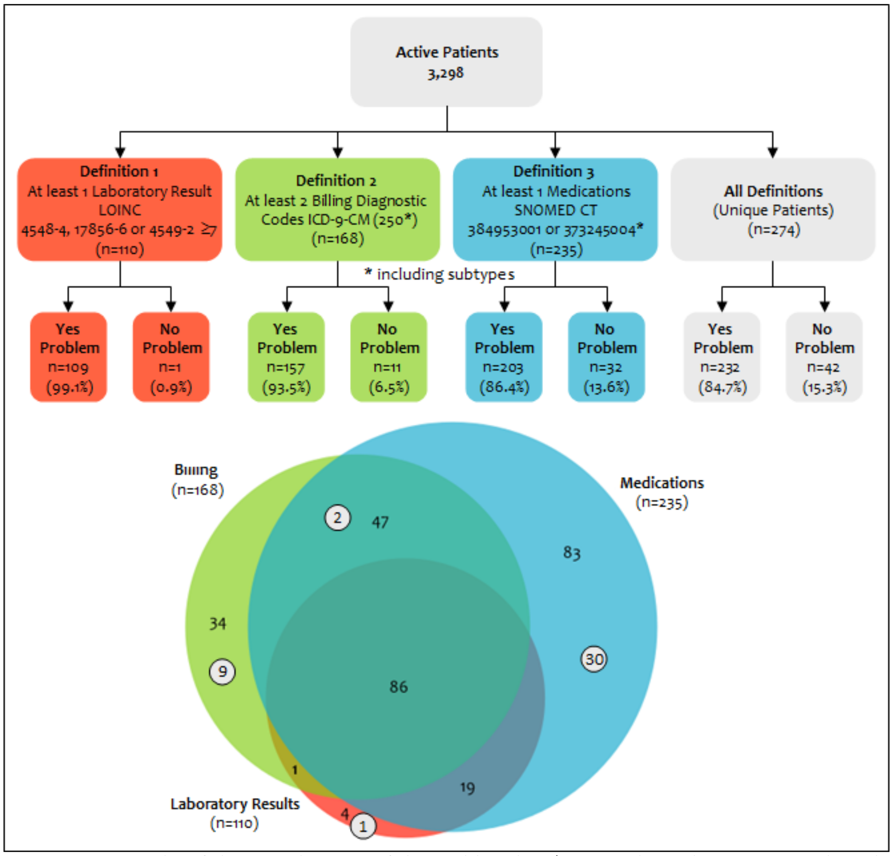


Figure 9-3. Results of the completeness of the problem list (grey circles indicate patients have been identified as having diabetes by one of the three definitions but do not have "diabetes" in the problem list).

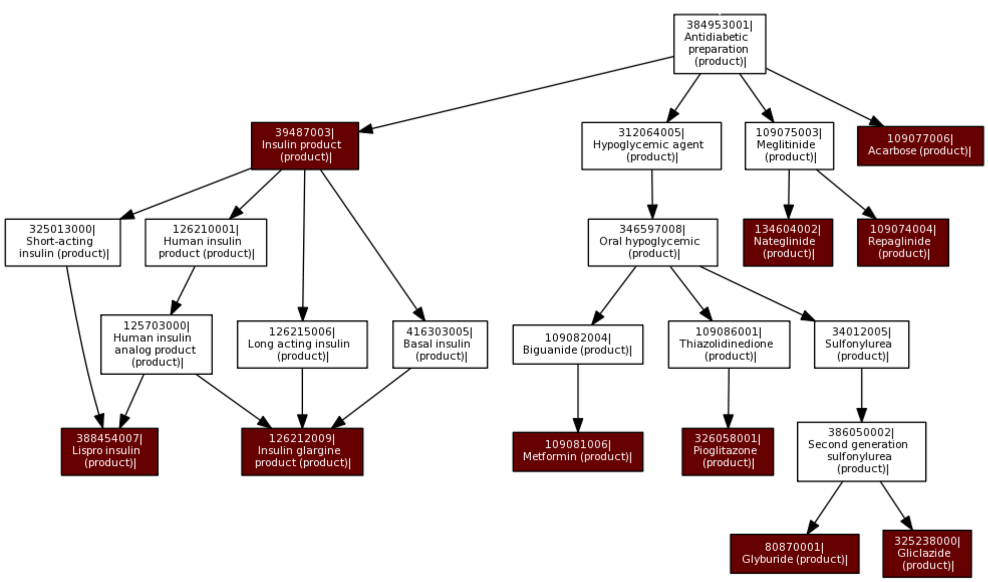


Figure 9-4. Diabetic medications used in the completeness of problem list method.

9.4.2.1.2 Completeness of Encounter Diagnoses and Billing Diagnostic Codes

The concept and subtypes of **73211009|Diabetes mellitus (disorder)|** in the encounter diagnoses were retrieved and compared against the billing diagnostic codes for appointments in the last year. The results are shown in **Table 9-14**. Of the 35,876 appointments that were scheduled in the past year, diabetes was recorded as an encounter diagnosis in 476 appointments. In 140 (29.5%) of the 476 appointments, either the exact ICD-9-CM code as recommended by the cross map was used (n=33, 6.9%) or a more general ICD-9-CM code was used (n=107, 22.5%). For example, the cross map recommends using **250.00** for **73211009|Diabetes mellitus (disorder)|** but **250** was used instead. None of the 35,876 appointments contained the maximum number (three) of billing diagnostic codes. Using the same cross map, the reverse can be done to ensure the completeness of the encounter diagnoses based on the billing diagnostic code recorded. Of the 174 and 48 appointments that included the billing diagnostic codes **250** and **250.01**, 108 (62.1%) and 35 (72.9%) respectively had an entry of **73211009|Diabetes mellitus (disorder)|** or one of its subtype concepts as an encounter diagnosis.

Table 9-14. Comparison of encounter diagnoses with billing diagnostic codes.

SNOMED CT to ICD-9-CM						
No	**SNOMED CT**	**Recommended ICD-9-CM Code**	**Number of Appointments**	**Total Number of Billing Diagnostic Codes**	**Number of Exact Billing Diagnostic Codes**	**Number of Main Billing Diagnostic Codes**
1.	73211009|Diabetes mellitus (disorder)|	**250.00** - Diabetes mellitus without mention of complication, type II or unspecified type, not stated as uncontrolled	125	96	0	14
2.	46635009|Diabetes mellitus type 1 (disorder)|	**250.01** - Diabetes mellitus without mention of complication, type I [juvenile type], not stated as uncontrolled	84	74	33	15
3.	44054006|Diabetes mellitus type 2 (disorder)|	**250.00** - Diabetes mellitus without mention of complication, type II or unspecified type, not stated as uncontrolled	267	217	0	78
4.	11687002|Gestational diabetes mellitus (disorder)|*	**648.80** - Abnormal glucose tolerance of mother, unspecified as to episode of care or not applicable	0	0	0	0
	Totals		**476**	**387**	**33**	**107**

* There were no results for gestational diabetes mellitus because no patients within the last year were seen for this condition.

ICD-9-CM to SNOMED CT				
No	**Billing Diagnostic Code**	**Number of Appointments**	**Number with Matching Diagnoses**	**Number Without Matching Diagnoses**
1.	**250** - Diabetes mellitus	174	108	66
2.	**250.01** - Diabetes mellitus without mention of complication, type I [juvenile type], not stated as uncontrolled	48	35	13
	Totals	**222**	**143**	**79**

9.4.2.2 Data Retrieval

The second data use dimension was retrieval, which was demonstrated in the adherence of care guidelines.

9.4.2.2.1 Adherence of Care Guidelines

A total of 218 patients (6.6% of active patients) were identified as diabetics by having **73211009|Diabetes mellitus (disorder)|** or one of its subtype concepts in the problem list or encounter diagnoses at least one year before and had an appointment at the primary care clinic within the last year. Of these patients, 179 (82.1%) achieved at least one of the five clinical assessment criteria, with blood pressure being the most commonly recorded criterion. Refer to **Table 9-15.** The smoking status of twenty-three patients were identified using three SNOMED CT concepts (**160616005|Trying to give up smoking (finding)|**, **365981007|Finding of tobacco smoking behavior (finding)|** and **65568007|Cigarette smoker (finding)|**). A total of 159 (72.9%) unique patients met at least one of the four treatment criteria with the most common treatment being blood pressure <130/85 or use of an antihypertensive agent (n=114, 52.3%). A total of 159 (72.9%) patients met at least one of the three target attainment criteria, with LDL-C level ≤100 mg/dL (2.59 mmol/L) being the most common target attainment (n=126, 57.8%). Using the scoring system in [16], the assessment, treatment and target attainment scores respectively, the average scores were 43.2/100, 39.8/100 and 40.6/100.

Table 9-15. Results of diabetes assessment scores based on criteria by Hahn, et al.[16]

No	Criteria	Patients	
		Total	Percent
Assessment Scores (20 points per criterion)			
1.	HbA1c level tested in the past six months	86	39.4%
2.	Urine microalbumin level tested in the past year	73	33.5%
3.	LDL-C level tested in the past year	129	59.2%
4.	Smoking Status	23	10.6%
5.	Blood pressure	160	73.4%
	Did not meet any of the criteria above	39	17.9%
	Number of criteria met: five (n=4, 1.8%), four (n=45, 20.6%), three (n=47, 21.6%), two (n=47, 21.6%), one (n=36, 16.5%), none (n=39, 17.9%).		
Treatment Scores (25 points per criterion)			
1.	HbA1c level at or below 8% or use of a hypoglycaemic agent	78	35.8%
2.	Urine microalbumin level below 30 mg/g (3.39mg/mmol) creatine or use of an angiotensin-converting enzyme (ACE) inhibitor or angiotensin II receptor blocker (ARB)	58	26.6%
3.	LDL-C below 100 mg/dL (2.59 mmol/L) or use of lipid-lowering agent	97	44.5%
4.	Blood pressure at or below 130/85 mmHg or use of an antihypertensive agent	114	52.3%
	Did not meet any of the criteria	59	27.1%
	Number of criteria met: four (n=21, 9.6%), three (n=34, 15.6%), two (n=57, 26.1%), one (n=47, 21.6%), none (n=59, 27.1%).		
Target Attainment Scores (33.3 points per criterion)			
1.	HbA1c level less than 8%	67	30.7%
2.	LDL-C level less than or equal to 100 mg/dL (2.59 mmol/L)	90	41.3%

No	Criteria	Patients	
3.	Blood pressure less than or equal to 130/85 mm Hg	109	50.0%
	Did not meet any of the criteria	68	31.2%
	Number of criteria met: three (n=27, 12.4%), two (n=62, 28.4%), one (n=61, 28.0%) and none (n=68, 31.2%).		

9.4.2.3 Data Sharing

The third data use dimension was sharing, which was demonstrated in the completeness of clinical summary.

9.4.2.3.1 Completeness of Clinical Summary

Overall, it was found that 755 and 11,454 terms in the problem list and encounter diagnoses respectively from active patients could be more appropriately recorded in other parts of a clinical record (refer to **Table 9-16**). For example, for the section **Allergies, Adverse Reactions and Alerts**, two predicate expressions were specified: **281647001|Adverse reaction (disorder)|** and **420134006|Propensity to adverse reactions (disorder)|**. A total of 166 (117 patients) adverse reactions and 98 (73 patients) propensities to adverse reactions were recorded in the problem list and encounter diagnoses. Examples included **292055008|Codeine adverse reaction (disorder)|** and **293619005|Ibuprofen allergy (disorder)|**.

Table 9-16. Continuity of Care Document with sections mapped to SNOMED CT concepts and examples of clinical concepts that were found in the problem list and encounter diagnoses that could belong to these sections.

No	Continuity of Care Document Sections	Example of SNOMED CT Predicate Expressions	Problem List		Encounter Diagnoses		Examples
			All Patients	Active Patients	All Patients	Active Patients	
1.	Summary Purpose	▪ **N/A**					
2.	Payers	▪ **N/A**					
3.	Advance Directive	▪ **Include:** 365870005|Finding of resuscitation status (finding)|	16	10	18	13	304253006|Not for resuscitation (finding)|
4.	Functional Status	▪ **Include:** 105501005|Dependence on enabling machine or device (finding)|	0	0	0	0	(no examples in dataset)
5.	Problems	▪ **Include:** 64572001|Disease (disorder)| ▪ **Exclude:** 420134006|Propensity to adverse reactions (disorder)| ▪ **Exclude:** 281647001|Adverse reaction (disorder)|	15,905	8,385	139,335	86,729	73211009|Diabetes mellitus (disorder)|
6.	Family History	▪ **Include:** 7177007|Family history with explicit context (situation)|	21	0	63	3	312824007|Family history of cancer of colon (situation)|

No	Continuity of Care Document Sections	Example of SNOMED CT Predicate Expressions	Problem List		Encounter Diagnoses		Examples
			All Patients	Active Patients	All Patients	Active Patients	
7.	Social History	▪ **Include:** 365949003|Health-related behavior finding (finding)|	436	215	3,881	2,238	65568007|Cigarette smoker (finding)|
8.	Allergies, Adverse Reactions and Alerts	▪ **Include:** 420134006|Propensity to adverse reactions (disorder)| ▪ **Include:** 281647001|Adverse reaction (disorder)|	17	10	386	254	426232007|Environmental allergy (disorder)|
9.	Medications	▪ **Include:** 373873005|Pharmaceutical / biologic product (product)|	64	21	1,381	606	73639000|Prescription drug (product)|
10.	Medical Equipment	▪ **Include:** 303607000|Biomedical equipment (physical object)|	2	1	96	53	87405001|Cane, device (physical object)|
11.	Immunisations	▪ **Include:** 127785005|Administration of substance to produce immunity, either active or passive (procedure)|	2	1	2,402	1,347	86198006|Influenza vaccination (procedure)|
12.	Vital Signs	▪ **Include:** 118227000|Vital signs finding (finding)|	5	3	502	327	61086009|Pulse irregular (finding)|
13.	Results	▪ **Include:** 442618008|Abnormal finding on evaluation procedure (finding)| ▪ **Include:** 441742003|Evaluation finding (finding)|: 363713009|Has interpretation (attribute)|=17621005|Normal (qualifier value)|	14	12	77	61	313172000|Colonoscopy abnormal (finding)|
14.	Procedures	▪ **Include:** 387713003|Surgical procedure (procedure)|	165	94	484	299	18949003|Change of dressing (procedure)|
15.	Encounters	▪ N/A					
16.	Plan of Care	▪ **Include:** 243120004|Regimes and therapies (regime/therapy)|	13	10	2,164	1,756	67516001|Detoxification therapy (regime/therapy)|
	Total		**16,660**	**87,62**	**150,789**	**93,686**	

In addition, it was found that non-communicable diseases such as diabetes were recorded in the encounter diagnoses but not the problem list. The results of the comparison between the problem list and encounter diagnoses for **73211009|Diabetes mellitus (disorder)|**, **46635009|Diabetes mellitus type 1 (disorder)|** and **44054006|Diabetes mellitus type 2 (disorder)|** are shown in **Table 9-17**. There were 271 active patients identified as diabetics from the encounter diagnoses and 155 from the problem list. The total number of unique patients was 275, 151 (54.9%) of which were include in both the encounter diagnoses and problem list. In all three cases, the encounter diagnoses had more diabetes patients than the problem list.

Table 9-17. Comparison between the problem list and encounter diagnoses of patients that have diabetes mellitus.

No	Type of Diabetes	Problems		Encounters		Overlap		
		Patients	Records	Patients	Records	Patients Overlap	Total Patients	Percent
1.	73211009|Diabetes mellitus (disorder)|	66	68	295	1,066	61	300	20.3%
2.	46635009|Diabetes mellitus type 1 (disorder)|	71	74	126	783	67	130	51.5%
3.	44054006|Diabetes mellitus type 2 (disorder)|	168	173	308	2,004	163	313	52.1%
4.	11687002|Gestational diabetes mellitus (disorder)|	4	5	3	8	0	7	0.0%
	Totals (Unique patients; Total frequency)	**282**	**320**	**475**	**3,861**	**271**	**486**	**55.8%**

9.4.3 Results of the Prototype

Six scenarios are presented here to demonstrate the different types of clinical value of using SNOMED CT in an EMR system.

9.4.3.1 Scenario #1: Patient Case Query

Refer to Figure 9-5. In this scenario, a basic search for **73211009|Diabetes mellitus (disorder)|** is conducted. The search automatically conducts a subsumption search and therefore retrieves **73211009|Diabetes mellitus (disorder)|** and three other subtypes. The subsumption search identifies three other subtype concepts (i.e., **46635009|Diabetes mellitus type 1 (disorder)|**, **44054006|Diabetes mellitus type 2 (disorder)|** and **11687002|Gestational diabetes mellitus (disorder)|**). A summary of the search results are presented along with the list of patients. On the right, the hierarchy is plotted out so that end-users will be able to see why the concepts were retrieved.

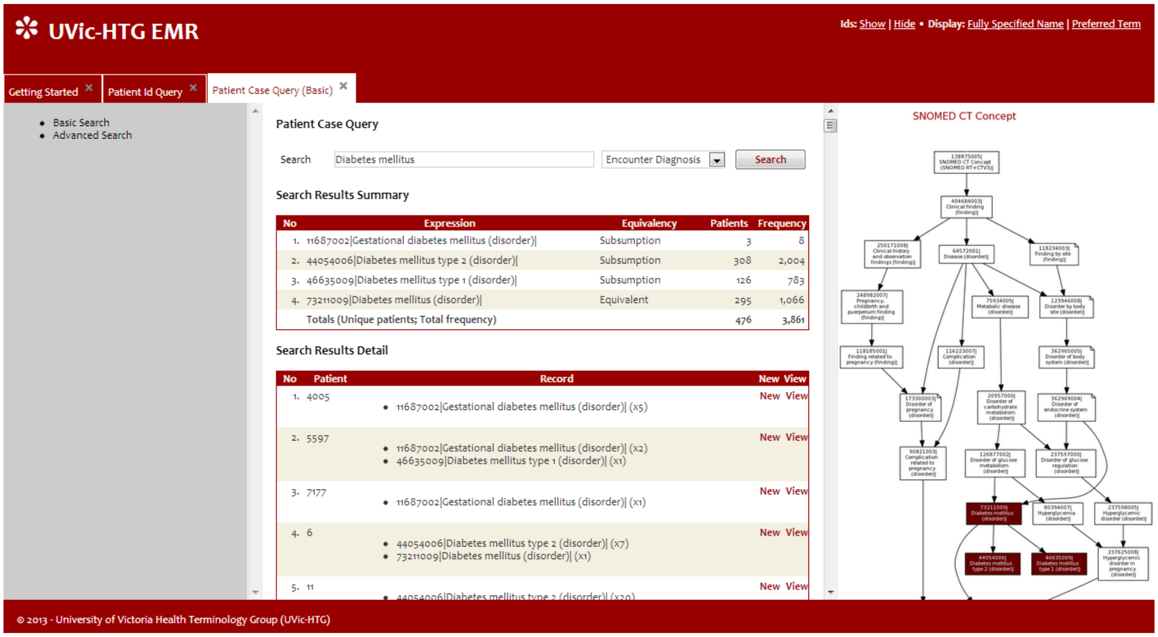


Figure 9-5. EMR prototype screenshot for Scenario #1.

9.4.3.2 *Scenario #2: Suggestions and Reminders*

Refer to **Figure 9-6**. In this scenario, diabetes mellitus is typed into the encounter diagnosis and the auto-complete brings up a list of the most relevant terms.

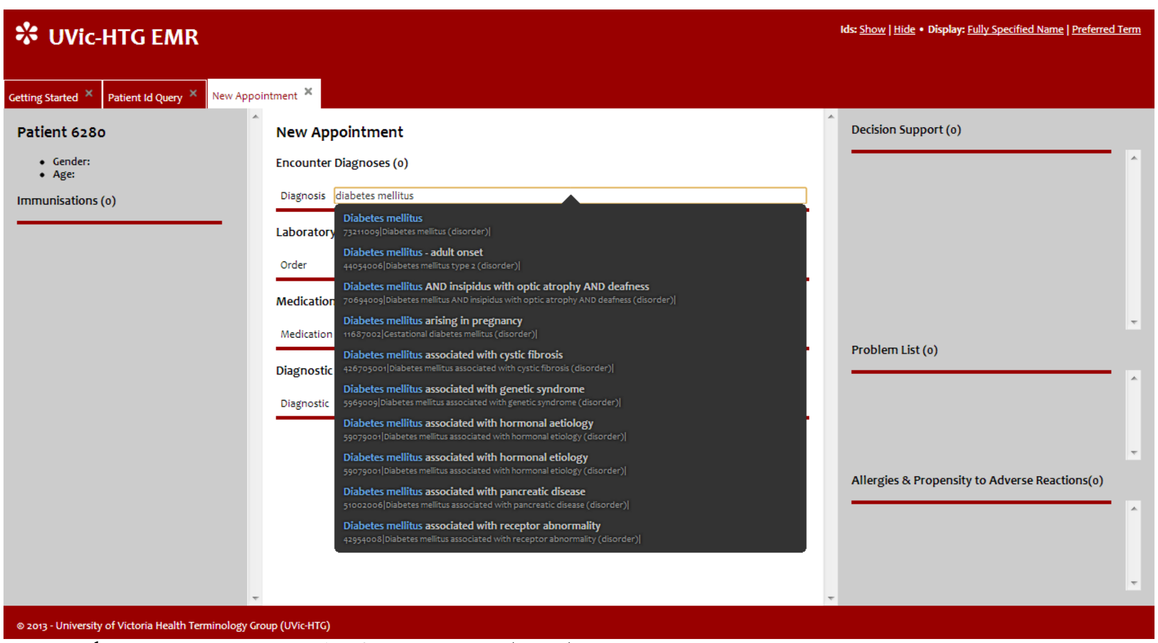


Figure 9-6. EMR prototype screenshot for Scenario #2 (part 1).

Refer to **Figure 9-7**. Once diabetes mellitus has been selected, it is added to the list of encounter diagnoses. The decision support uses the ICD-9-CM cross map to suggest a billing diagnostic code, uses the Hahn, et al.,[16] criteria to remind the clinician to order laboratory tests, and based on **73211009|Diabetes mellitus (disorder)|** being a subtype of **64572001|Disease (disorder)|**, it is suggested that it be added to the problem list. The decision support is shown on the right column. As not all the items in the decision support can be seen, they are listed below:

- Suggestion: CDSS suggests adding "250.00 - Diabetes mellitus without mention of complication, type II or unspecified type, not stated as uncontrolled" to your diagnostic record.
- Reminder: CDSS suggests ordering a laboratory test 4548-4 - Hemoglobin A1c/Hemoglobin.total in Blood.
- Reminder: CDSS suggests ordering a laboratory test 9318-7 - Albumin/Creatinine [Mass Ratio] in Urine
- Reminder: CDSS suggests ordering a laboratory test 22748-8 - Cholesterol in LDL [Moles/volume] in Serum or Plasma.
- Suggestion: CDSS suggests adding 73211009|Diabetes mellitus (disorder)| to the problem list.

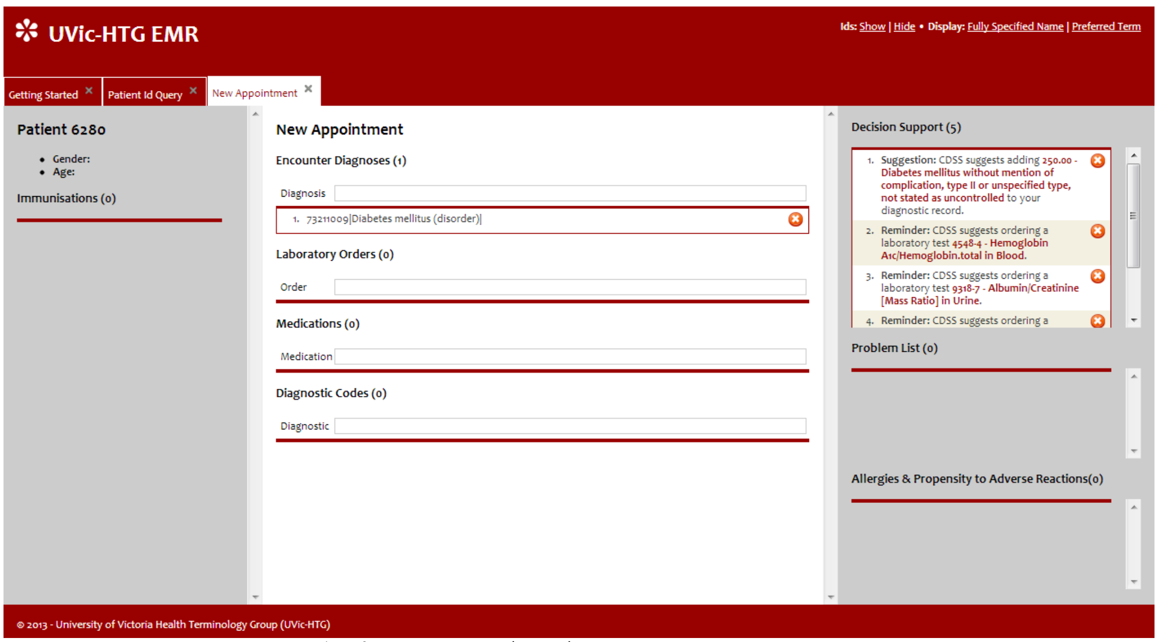


Figure 9-7. EMR prototype screenshot for Scenario #2 (part 2).

Refer to **Figure 9-8**. The items in the decision support can be closed by clicking on the red cross or adhered to by clicking on the words in red. In this case, all five decision support items have been clicked. The three laboratory tests have been added under the section "Laboratory Orders" while the ICD-9-CM code has been added under the section "Diagnostic Codes." Diabetes mellitus has also been added under "Problem list" on the right column.

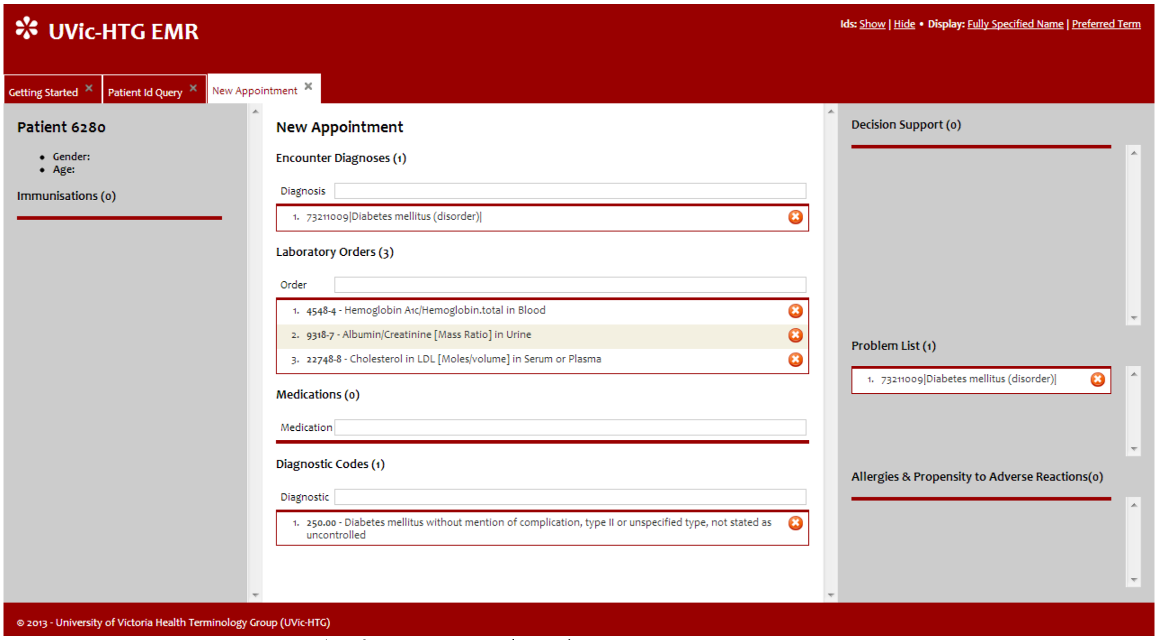


Figure 9-8. EMR prototype screenshot for Scenario #2 (part 3).

9.4.3.3 Scenario #3: Alerts and Suggestions

Refer to **Figure 9-9**. In this scenario, a patient has had a **4548-4 - Hemoglobin A1c/Hemoglobin.total in Blood**, **9318-7 - Albumin/Creatinine [Mass ratio] in Urine** and **39469-2 - Cholesterol in LDL [Moles/volume] in Serum or Plasma by calculation** within the last six months and all of the values, according to Hahn, et al.,[16] fall outside of the normal ranges. The decision support has identified three alerts and two suggestions. Refer to the list below. For the Haemoglobin A1c test, the recommendation should be a **312064005|Hypoglycemic agent (product)|**. But in this case, the decision support has identified that **109081006|Metformin (product)|** and **80870001|Glyburide (product)|** are both subtypes of **312064005|Hypoglycemic agent (product)|** and therefore has not suggested any medications. On the other hand, for the Albumine/Creatinine and Cholesterol in LDL tests, the patient does not have any medications so the decision support suggests **41549009|Angiotensin-converting enzyme inhibitor agent (product)|** and **57952007|Antilipemic agent (product)|**.

- Alert: Laboratory test, 4548-4 - Hemoglobin A1c/Hemoglobin.total in Blood on 2012-12-04 was 8.9, which was above the normal reference range of 8. However, the patient is already prescribed on the following medication(s): 109081006|Metformin (product)|, 80870001|Glyburide (product)|.
- Alert/Suggestion: Laboratory test, 9318-7 - Albumin/Creatinine [Mass ratio] in Urine on 2012-12-04 was 15.40, which was above the normal reference range of 3.39. CDSS suggest 41549009|Angiotensin-converting enzyme inhibitor agent (product)|.
- Alert/Suggestion: Laboratory test, 39469-2 - Cholesterol in LDL [Moles/volume] in Serum or Plasma by calculation on 2012-09-27 was 2.67, which was above the normal reference range of 2.59. CDSS suggest 57952007|Antilipemic agent (product)|.

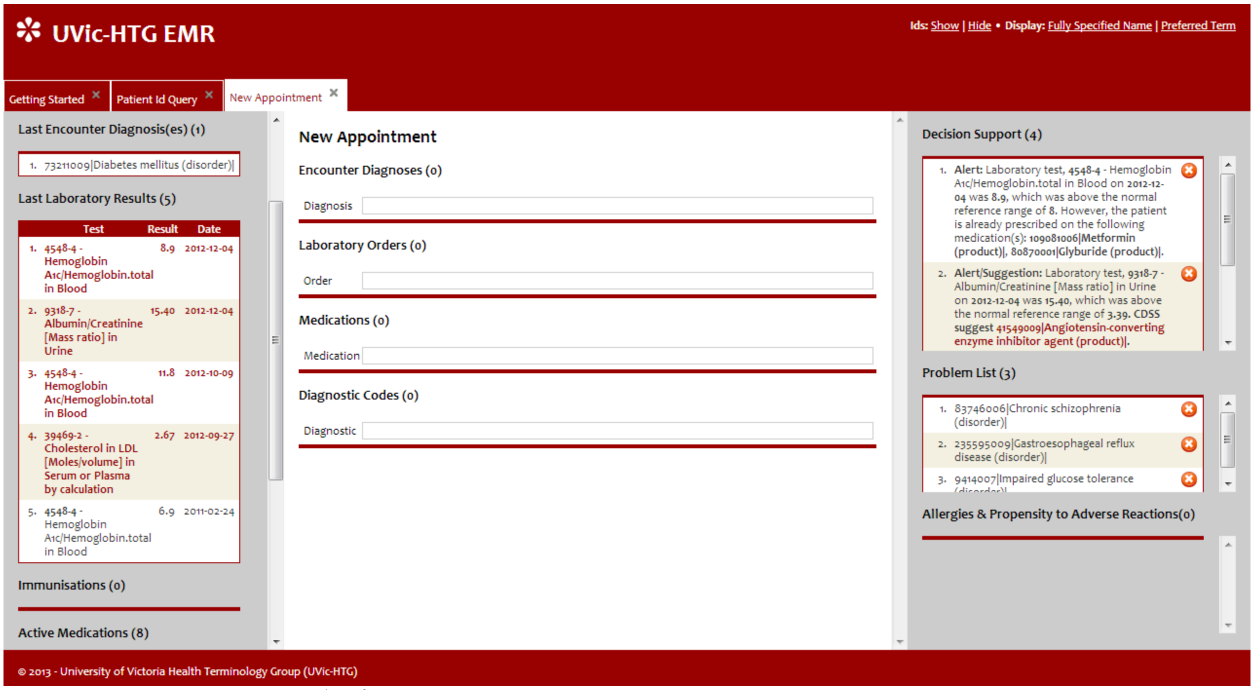


Figure 9-9. EMR prototype screenshot for Scenario #3.

9.4.3.4 Scenario #4: Suggestions

Refer to **Figure 9-10**. In this scenario, **282095007|Allergic reaction to bee sting (disorder)|** has been added as an encounter diagnosis. The decision support retrieves the ICD-9-CM billing diagnostic codes as usual, but also suggests adding **282095007|Allergic reaction to bee sting (disorder)|** and **424213003|Allergy to bee venom (disorder)|** to the list of allergies. **282095007|Allergic reaction to bee sting (disorder)|** was suggested because it is a subtype of **281647001|Adverse reaction (disorder)|**. On the other hand **424213003|Allergy to bee venom (disorder)|** was suggested because the defining attribute of **282095007|Allergic reaction to bee sting (disorder)|** contains **246075003|Causative agent (attribute)|=288328004|Bee venom (substance)|**, which was then post-coordinated as **420134006|propensity to adverse reactions (disorder)|:246075003|Causative agent (attribute)|=288328004|Bee venom (substance)|**, after which it was found to be equivalent to **424213003|Allergy to bee venom (disorder)|**.

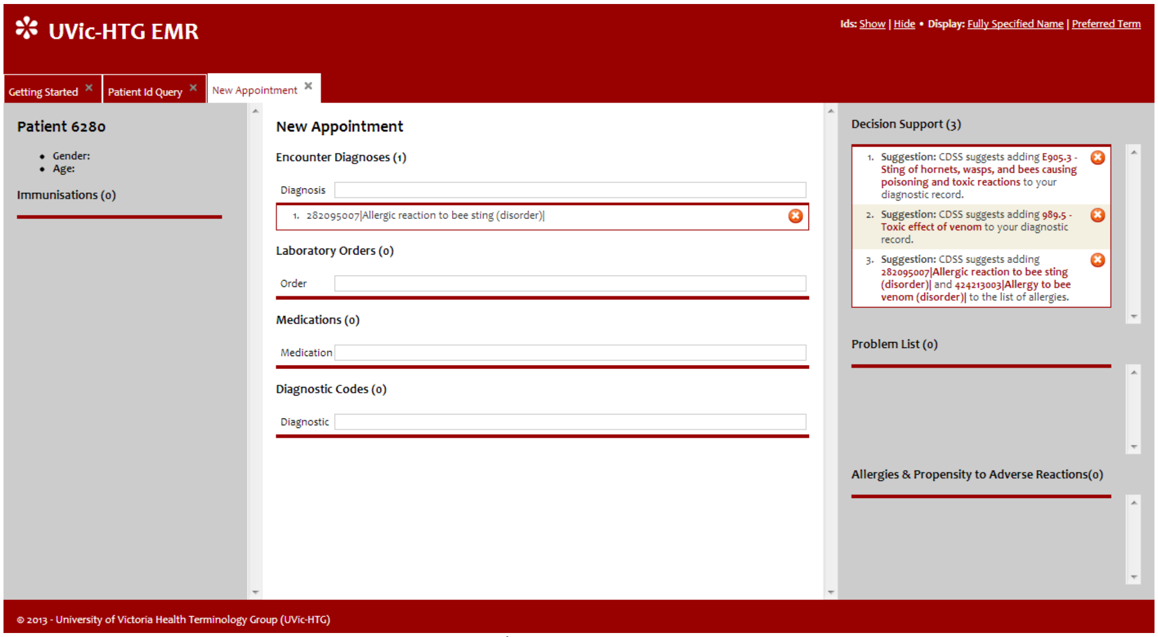


Figure 9-10. EMR prototype screenshot for Scenario #4.

9.4.3.5 Scenario #5: Alerts

Refer to **Figure 9-11**. In this scenario, a patient has an **293619005|Ibuprofen allergy (disorder)|** allergy. When an attempt is made to add **370192006|Ibuprofen 100mg tablet (product)|** to the list of medications, the decision support warns of the allergy to ibuprofen. This is done by extracting the **127489000|Has active ingredient (attribute)|** out of **370192006|Ibuprofen 100mg tablet (product)|**, which is **387207008|Ibuprofen (substance)|**, and compared against the list of allergies of which **293619005|Ibuprofen allergy (disorder)|** has the defining attributes of **246075003|Causative agent (attribute)|=387207008|Ibuprofen (substance)|**.

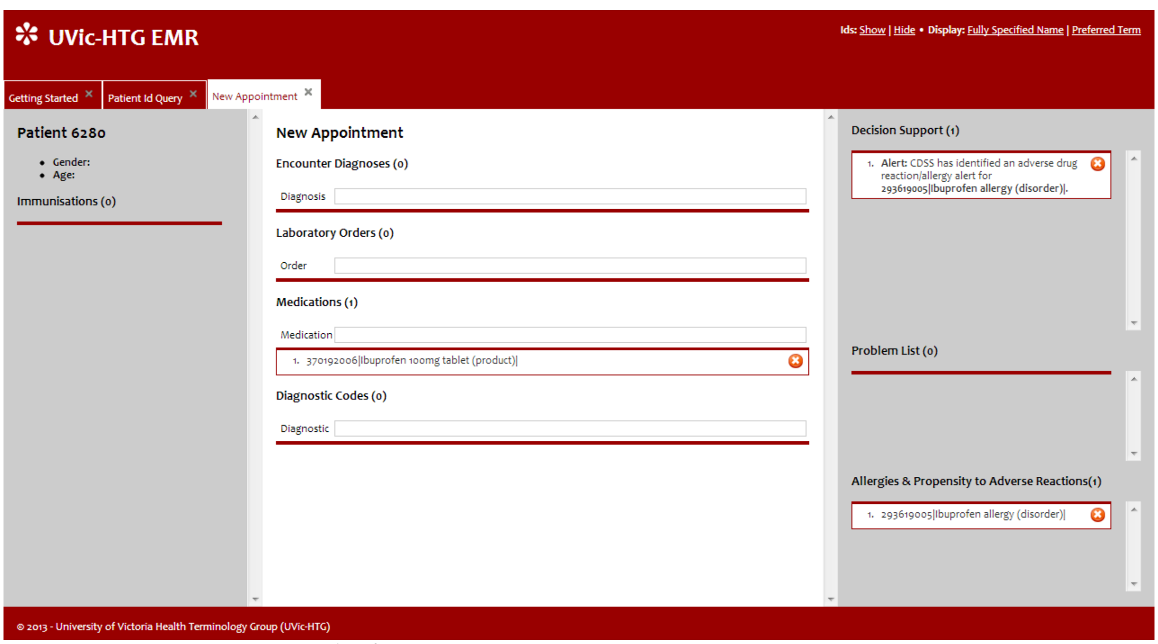


Figure 9-11. EMR prototype screenshot for Scenario #5.

9.4.3.6 *Scenario #6: Suggestions*

Refer to **Figure 9-12**. In this scenario, once a clinical record is opened, the past encounter diagnoses and problem lists are queried in order to determine if some of the clinical conditions captured there may be better recorded in other sections of a clinical record. In this prototype, two main queries are conducted. First, to determine if any condition should be stored under the "Immunisation" section, which uses the predicate concept **127785005|Administration of substance to produce immunity, either active or passive (procedure)|**. Second, to determine if any condition could be stored under the "Allergies & Propensity to Adverse Reactions" section, which uses the predicate concepts **420134006|Propensity to adverse reactions (disorder)|** and **281647001|Adverse reaction (disorder)|**. In this case, four past problem list and encounter diagnoses were suggested to be stored elsewhere in the clinical summary.

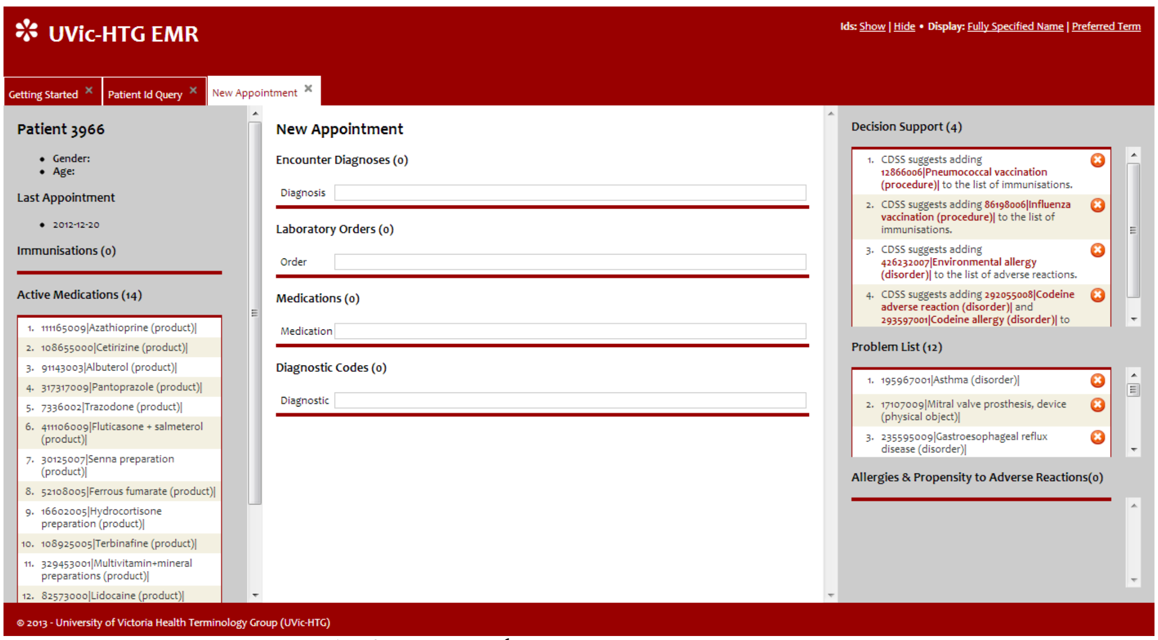


Figure 9-12. EMR prototype screenshot for Scenario #6.

9.4.4 *Feedback Sessions*

This section reports on the feedback of the clinical and technical feedback sessions. The slides used in the feedback sessions are available in **Appendix G**. As there was overlapping feedback between the two sessions, the feedback was combined and reported together. Overall feedback was positive and the work was described as valuable in demonstrating the usefulness of SNOMED CT.

SNOMED CT Clinical Value Framework. The proposed SNOMED CT Clinical Value Framework for demonstrating the value of SNOMED CT resonated with the group, especially the target data users (i.e., patients; practice; and population). It was suggested that the use of SNOMED CT would be most beneficial initially at the practice and population levels where reports are generated and patient case queries are conducted for recall. There is a need to conduct patient recall at a practice level, with standardised criteria and methods being applied as opposed to leaving it up to individual clinicians to make the decision although patient lists generated should still be reviewed by the clinicians. While the suggestions and decision support features at the patient level are useful, it is uncertain whether clinicians would use those features. There were also concerns on whether EMR vendors would be able to provide such functionality.

Clinicians View of Data Quality. One of the main challenges pointed out by a clinician was that despite the added value of SNOMED CT, there are clinicians who simply do not care about data quality or what data is captured. These clinicians feel they know which patients suffer from what conditions and do not need any additional tools. In these cases, the clinician in the interview group stressed the importance of demonstrating the value of SNOMED CT (such as the methods presented) and not just the ability to capture the data but to use the data captured that adds clinical value. The continuity of care, the ability for clinicians to care for another clinician's patient and patient referrals, were seen as potential motivators. It was suggested that incentives, financial ones in particular, may be necessary in order to persuade clinicians to improve the quality of their data as it was expressed that most clinicians would not voluntarily want to improve the quality of their data unless there were incentives or requirements. Routine data quality reports which compare the performance of clinicians within a practice may help by adding some peer pressure on clinicians with poor data quality. The intent is not to embarrass but to enable clinicians to acknowledge the problem.

Clinicians Confidence in Methods and Terminology. There is a need to demonstrate the validity of the methods and terminology to clinicians so that will be amenable to using them. The "confidence in methods," as described by the clinician, stems from being able to demonstrate that the correct patients are retrieved, incorrect patients are not retrieved and that none are omitted. While it may not be possible to completely rule out all false negatives, it is possible to compare the SNOMED CT retrieval methods against current free text searches and demonstrate the accuracy of SNOMED CT queries, which was demonstrated in the technical feedback session but not the clinical feedback session. The "confidence in terminology" is two-fold. First, clinicians who already use coded data want to continue to be able to use the terms they are familiar with; therefore plugging a new subset into an EMR is not feasible as they may not be able to find the terms they want. Instead of using an "off-the-shelf" subset, it is important to review and encode historical records and to create an interface terminology. Second, clinicians do not want their workflow to be interrupted. When using free text, they are able to type without looking at the screen so as not to disrupt eye contact with their patients but that is not possible with pick lists. One solution would be to automatically index the free text and provide the list of codes for clinicians to review. However, medical language processing still has some ways to go before it can automatically identify all the appropriate codes. The clinician who raised these issues stressed the importance of getting other clinicians to understand the importance of using coded structured data.

SNOMED CT Queries. SNOMED CT subsumption queries were seen as useful in retrieving more specific types of diseases as well as the ability to identify diseases using different definitions (e.g., problem list, medications, laboratory results). A question raised was whether a patient who has a family history of breast cancer but does not have breast cancer can be retrieved for further consultation and screening. In this case, it was not possible to demonstrate it because the anonymised primary care dataset only included a few records with family history. The ability of SNOMED CT to record context, both temporal and subject relationship, enables such queries to be executed, provided the data is captured accurately. A second question raised was whether it was possible to differentiate between what was recorded as a diagnostic code (e.g., 250) and the health condition (e.g., screening for diabetes) and to weed out false positives. In this case it would have been possible if screening for diabetes was recorded as **171183004|Diabetes mellitus screening (procedure)|** instead of **73211009|Diabetes mellitus (disorder)|**. The presentation included examples of how SNOMED CT expressions can be structured so as to ensure relevant records are retrieved. There was concern that clinicians would not know how to craft such complex expressions and even terminology analysts may end up with differing expressions. Reporting at the population level may be inconsistent if different criteria were used at the practice levels. Possible solutions include creating pre-defined queries and to develop intuitive interfaces and robust functionality that would enable the necessary predicate expressions to be automatically generated based on a simple interface. Reporting requirements set forth by any organisation should comprise inclusion and exclusion expression criteria. If the reporting requirements are based on standardised terminologies, the queries should be structured in a machine readable format so that the queries will be consistent, unambiguous and easy to use.

Cross Mapping. The automatic generation of diagnostic codes was seen as useful and a time saver. It was noted that the number of diagnostic codes that can be submitted differ by province. BC allows up to three diagnostic codes while Ontario only allows one. Therefore statistical reports generated at a population level that are sourced from clinician submitted diagnostic codes are probably not accurate as it is normal for clinicians to record something completely different in the billing code and diagnostic code. In the prototype demonstration, potential ICD-9-CM codes were automatically generated and displayed for possible inclusion as diagnostic codes. The question was asked why not automatically add them. The answer was that in some cases, more than one diagnostic code could be used. A possible solution would be to review the historical diagnostic codes in a practice and to identify situations in which multiple codes were applicable and to pick the most suitable one. One-to-one maps could be automatically used.

Maintenance. As SNOMED CT is constantly undergoing revision and concepts are inactivated, a question was raised on how to maintain the codes in a subset or EMR. One of the auditing methods for subsets developed in this research (but was not reported in this dissertation) was on versioning, whereby a set of SNOMED CT expressions were compared against two SNOMED CT release versions to identify the changes made. Therefore the solution would be to extract the SNOMED CT expressions used within an EMR every six months and to run it through the subset auditing tool. The tool also suggests active concepts based on the historical relationships for inactive concepts. As historical patient records should not be changed and to prevent an excessive amount of maintenance, the patient records may be stored with interface terminology codes which are mapped to SNOMED CT. Therefore any changes will be reflected in the map as opposed to actual patient records. Only when SNOMED CT queries are required are the mappings looked up. This mapping is hidden from the clinicians.

SNOMED CT Application Programming Interface (API). As the technical feedback session delved into the inner workings of structural subsumption, concerns were expressed on how organisations, particularly vendors, would be able to develop such complex algorithms. Therefore the use of a SNOMED CT Application Programming Interface (API) to help in both encoding and retrieval was seen as beneficial especially to vendors and small health care organisations that do not have time, expertise or resources to develop their own tools. Questions arose as to who should host such a SNOMED CT API or web services. In Canada, Canada Health Infoway would be in a good position to host the services as they are responsible for the distribution of SNOMED CT in Canada, and they have developed and maintain SNOMED CT subsets and extensions. At this moment they probably do not have the technical expertise but could get guidance from the IHTSDO. At the international level, the IHTSDO potentially could host such services. They already have the concept look up service and there have been discussions on finalising the query specification.

Legacy Data. A question was raised on what to do with free text or legacy data. One solution is to run the free text or legacy data through a cleaning and encoding algorithm, as was done with the anonymised dataset in this study, and generate a list of SNOMED CT expressions. As these SNOMED CT expressions will be entered into an EMR and there are clinical and legal implications, it is important that a manual review process take place prior to using the SNOMED CT expressions generated.

Generalisability of Methods. A question of the generalisability of the methods for use in other diseases was raised. The main concern the clinician had was that diabetes has been frequently targeted by informaticians and that the methods developed may not be applicable to other diseases. While the presentation focused on diabetes so as to provide a consistent domain of interest throughout the methods, the methods included a broad range of data sources including problem lists, encounter diagnoses, diagnostic codes, medications, laboratory results and physical examination results. Therefore any disease that contains relevant information from any of the data sources can be used. It should also be noted that all 17 clinical conditions by Wright, et al.,[15] were used in the study but only the results of diabetes were presented.

Value of SNOMED CT. The clinicians felt the true value of SNOMED CT will be through the demonstration of improved patient care, which at the moment is still theoretical or proof of concept via retrospective studies. Clinicians need to be taught on how decision support and standardisation will support front end delivery of care and the continuity of care. While the demonstrator was important to show how SNOMED CT could be implemented and was helpful in moving the focus from the theoretical to the practical, there is still the challenge of how to get vendors to implement it in an operational setting and whether clinicians will use it when it is available.

9.5 Discussion

In this study, SNOMED CT was used to encode the problem list, encounter diagnoses and medications and the value of encoding these three data elements along with laboratory results (LOINC), billing diagnostic codes (ICD-9-CM) and physical examinations were analysed using four methods. The discussion section is divided into three parts. First, how the clinical value of SNOMED CT was demonstrated in this study. Second, the benefits and implications of using SNOMED CT. Third, challenges encountered in this study. The limitations and future work of this study are described in **Chapter Ten**.

9.5.1 Clinical Value of SNOMED CT

The discussion of the clinical value of SNOMED CT is organised according to the proposed SNOMED CT Clinical Value Framework.

9.5.1.1 Data Capture

Completeness of Problem List. The problem list identified 232 patients with diabetes. In contrast, the method described in this study has the potential to improve the completeness of the problem list as 42 additional patients were identified using laboratory results, billing codes and medications. While Wright, et al.,[15] looked only at whether or not three SNOMED CT concepts occurred in the problem list, this study included the subtypes of the three concepts. For example, we looked at all subtypes of **73211009|Diabetes mellitus (disorder)|**, which included three additional concepts in this dataset: **46635009|Diabetes mellitus type 1 (disorder)|**, **44054006|Diabetes mellitus type 2 (disorder)|** and **11687002|Gestational diabetes mellitus (disorder)|**. While this method was applied at a practice level, it can also be applied at a patient level to provide suggestions to clinicians. For example, if a HbA1c test result returned was >7, a check could be done to see if diabetes or one of it subtype concepts was on the patient's problem list and suggest it to the clinician. The suggestion could take place when a clinician opens a patient record or in daily or weekly reports at the practice level and can help fulfil the Clinical Value Target Level 1 of capturing problems in the EMR.

Completeness of Encounter Diagnoses and Billing Diagnostic Codes. A comparison between the encounter diagnoses and billing diagnostic codes showed there was a discrepancy between the two. There are five possible reasons for why a billing diagnostic code for diabetes was recorded but diabetes was not recorded in the encounter diagnoses. First, the incorrect billing diagnostic code was used, which would indicate clinicians might need to be re-trained on recording billing diagnostic codes. Second, diabetes was inadvertently omitted, which would indicate clinicians might need to be re-trained or reminded on recording the encounter diagnosis. Third, it is possible that a patient who came in for the appointment was not there specifically for diabetes but for other problems. Fourth, clinicians document one main encounter diagnosis but also record a secondary billing diagnosis. Fifth, usability issues in the EMR may have contributed to discouraging clinicians from recording multiple encounter diagnoses. The results also showed clinicians were more inclined to use the general **250** code as opposed to more specific codes such as **250.00** for **73211009|Diabetes mellitus (disorder)|** and **250.00** for **44054006|Diabetes mellitus type 2 (disorder)|**, and **250.01** for **46635009|Diabetes mellitus type 1 (disorder)|** although the implications are unclear. It should be noted that the cross map does not have **250** in the list of target codes. If **250** is preferred for any type of diabetes, the cross map would need to be modified and the mappings for other diseases will need to be reviewed before implementation. This method was conducted at the practice level but can also be applied at a patient level where clinicians can be provided suggestions for ICD-9-CM codes in real-time when entering an encounter diagnosis or a suggestion for an encounter diagnosis based on what ICD-9-CM codes were entered. In the United States, having a more complete set of billing codes may mean additional income.[25, 26]

9.5.1.2 Data Retrieval

Adherence of Care Guidelines. From the method of analysis using only the EMR data extracted, the results of the adherence to diabetes care guidelines showed that a small percentage of patients were receiving optimal care that was being recorded in the EMR. There are a few reasons for this. First, the guidelines in this study were based on laboratory results and only those captured electronically were evaluated. Second, patients may visit multiple clinics and therefore all relevant information (e.g., medication prescriptions) may not have been available in the dataset. Third, activities may not be documented in the EMR (e.g. paper prescriptions). Fourth, information may be documented in the wrong places in the EMR. Fifth, results may not be returned to the clinic in a manner that could be accessed through our algorithms (e.g. in a scanned consult letter). It should be noted that although the scores were based on common diabetes guidelines, it still represents a small portion of all available guidelines.

These guidelines were applied at the practice level but can also be applied at the patient level where clinicians are provided reminders for ordering test or suggestions for medications based on alerts for out of range results. This may help to improve the adherence of care guidelines and documentation of care, which may translate into improved outcomes. At a population level, government organisations will be able to compare the levels of care across different jurisdictions.

9.5.1.3 Data Sharing

Completeness of Clinical Summary. As the exchange of machine-processable clinical summaries becomes commonplace, it is vital that information contained in the summaries be accurate and organised appropriately. For example, a clinician who is reviewing a patient's list of allergies may not think to check the problem list when there is a separate section that specifically lists all allergies. In addition, computerised decision support rules in EMRs for allergies may only be triggered by entries contained in the allergies section and not the problem list. It is also likely that only a patient's problems but not encounter diagnoses are included in clinical summaries. In this study, out of the 274 patients that were identified as diabetics in the problem list or encounter diagnoses, there was only a 54.9% overlap, which further illustrates the need to analyse the information recorded in the encounter diagnoses. By specifying predicate expressions for each clinical summary section, it is possible to help improve the completeness of clinical summaries. It also has the potential to simplify clinician data entry as the encounter diagnosis may be used as a single point of entry for a range of data elements, and the EMR is able to direct the information entered into the appropriate sections. The predicate expressions presented in this study are only examples, and further refinement of the inclusion and exclusion criteria are needed.

9.5.2 Benefits and Implications of Using SNOMED CT

Standardisation. One of the main benefits of SNOMED CT is that it is a standardised clinical reference terminology. While this point may sound overly simplistic, it is important to point out the degree of variability a disease such as diabetes may be represented with free text. In the anonymised dataset, diabetes was represented as "Diabetes Mellitus Non Insulin Dependent (Type 2)" (n=1,046), "Diabetes Mellitus Insulin Dependent" (n=670), "Diabetes Type 2" (n=514), "Diabetes Mellitus -Type 2- Non Insulin Dependent" (n=421), "Diabetes Mellitus" (n=404), "Diabetes" (n=355) and many other ways. The data dictionary in the primary care EMR in which the data was extracted was not concept-based and therefore without the SNOMED CT encoding it was difficult to identify all the variations in which diabetes was represetned. By encoding local terms with SNOMED CT, clinicians will still be free to use their preferred terms but any data analysis will be simplified with a concept-based terminology residing in the background.

Simplified Retrieval. The use of SNOMED CT can help to simplify queries as a single expression can be used to retrieve clinical conditions that have multiple synonyms, disease specialisation, implicit conditions and in the appropriate context. For example, it was relatively easy to retrieve all types of diabetes mellitus as well as all related complications of diabetes using a single concept of **73211009|Diabetes mellitus (disorder)|** by specifying the inclusion of subtypes in the retrieval method. It was also able to exclude non-diabetic expressions or diabetes expressions with a different context. **Table 9-18** lists some examples.

Table 9-18. Original terms with the word "diabetes" or "diabetic" that are not related to the presence of diabetes mellitus.

No	Original Term	SNOMED CT	Problems	Encounters
1.	Diabetes mellitus risk	243796009|Situation with explicit context (situation)|: 246090004|Associated finding (attribute)|= 73211009|Diabetes mellitus (disorder)|, 408729009|Finding context (attribute)|= 410519009|At risk context (qualifier value)|	8	68
2.	Diabetes insipidus	15771004|Diabetes insipidus (disorder)|	2	11
3.	Diabetic diarrhoea	235832002|Hollow visceral neuropathy (disorder)|	0	1

There were 68 patients identified as "at risk of diabetes." In SNOMED CT, "at risk" concepts can be encoded either as one of the subtypes of **281694009|Finding of at risk (finding)|** or it can be qualified using the qualifier value concept **410519009|At risk context (qualifier value)|**. The benefit of using the pre-coordinated concept and subtype of

281694009|Finding of at risk (finding)| is that the concept already exists. However, all 162 subtype concepts are primitive and do not have any defining attributes that link it back to the clinical condition that is being defined as at risk. For example, the concept **161639008|At risk of heart disease (finding)|** does not have a defining attribute that is linked to **56265001|Heart disease (disorder)|**, which makes the concept less useful when conducting semantic queries (as opposed to lexical queries). On the other hand, by creating a post-coordinated expression and using the **408729009|Finding context (attribute)|** to define the context, relevant clinical conditions are able to be retrieved and can be differentiated with the context. In this case, 63 patients were retrieved. These patients have not been diagnosed with diabetes but are at risk of diabetes. Executing a string-matching query for "diabetes" would have retrieved these as false positive results. The other example was "diabetes insipidus," which refers to excessive thirst and excretion of large amounts of severely diluted urine. "Diabetic diarrhoea" was also identified using a text string search and does refer to diarrhoea during diabetes but there may be a missing definition that identifies it as a diabetic complication.

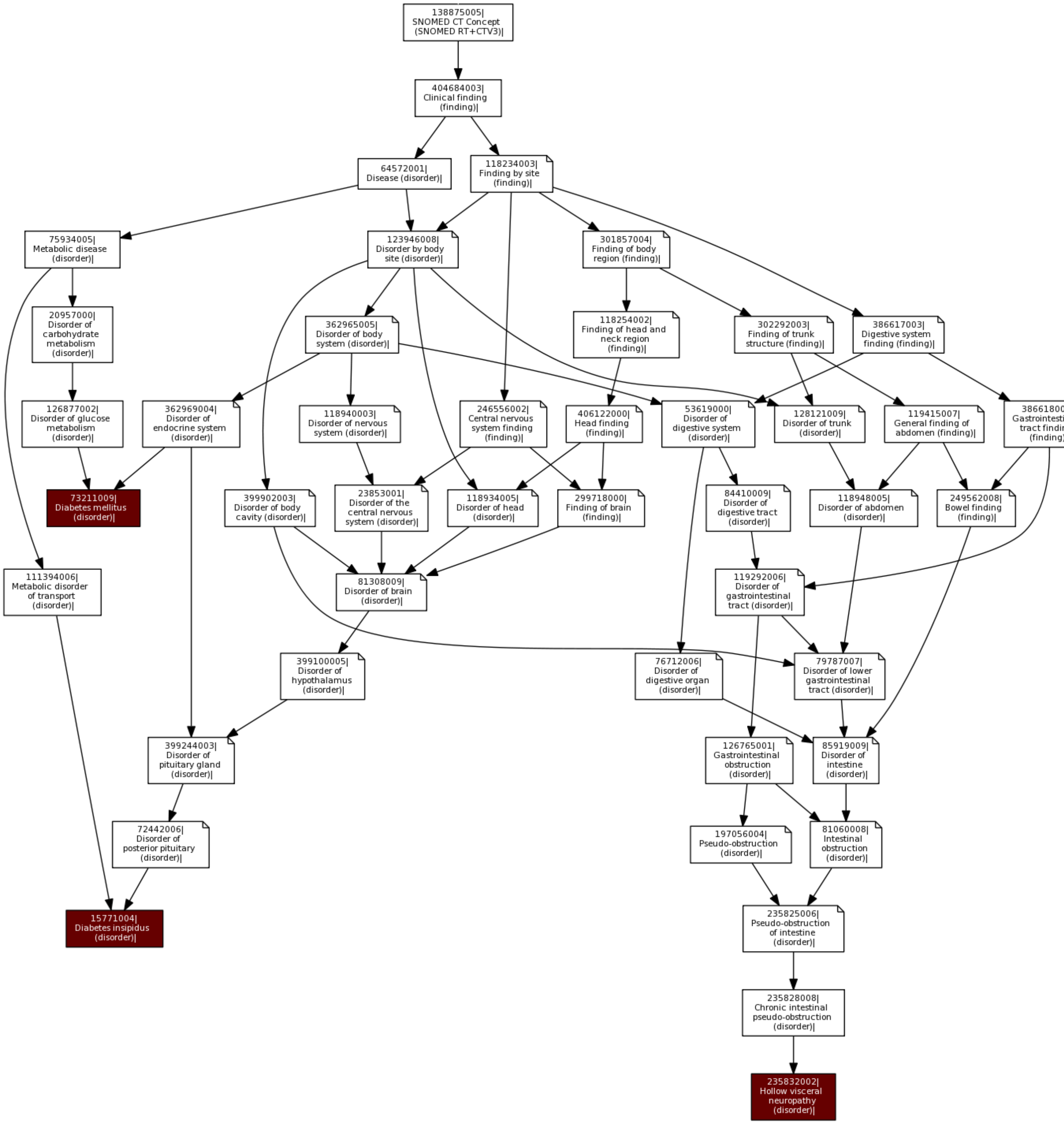


Figure 9-13. Location of concepts that have diabetes in the description but are not related to diabetes mellitus.

Facilitated Decision Support. By capturing and populating clinical data into appropriate sections within a clinical record with SNOMED CT and other standardised terminologies, decision support systems can be easier to implement. Clinical summaries that are exchanged with other providers may be imported seamlessly to ensure critical allergies and conditions are imported into the appropriate sections and can trigger decision support rules as appropriate. Guidelines that are encoded with SNOMED CT and other terminology systems such as LOINC can aid decision support by providing alerts, suggestions and reminders based on evidence-based guidelines. At a practice level, it can also facilitate audits to ensure clinicians are providing the best care for their patients. By creating guidelines that are based on standardised terminologies, these guidelines will be re-usable in different organisations and will not require re-mapping the guidelines to local terminologies.

Improved Completeness. The comparison of problem lists that are SNOMED CT-based can help improve the completeness of the problem list by comparing it with encounter diagnoses, billing codes, medications and laboratory results. Having a complete problem list and encounter diagnoses will help organisations fulfil the Clinical Value Target Level 1 of recording problem lists and encounter diagnoses electronically.

Automated Billing/Classification. The use of cross maps at the point of care can aid data capture by suggesting appropriate classification codes for billing or statistical reporting purposes. Cross maps also ensure that data captured once can be re-used without much effort to recode and that the classification codes are consistent. Cross maps used at a practice level can facilitate audits and ensure clinicians are using the appropriate billing codes.

9.5.3 Challenges

The challenges encountered could be broadly categorised as data quality challenges, which refer to the data quality issues encountered in the primary care dataset, and SNOMED CT-related challenges, which refer to challenges of using SNOMED CT.

9.5.3.1 Data Quality Challenges

There were three main data quality challenges: (1) identifying coding systems; (2) missing dates; (3) different HL7 data types and units; and (4) ambiguous terms.

Identifying Coding Systems. There were at least three different standards used to identify lab test: LOINC, CPT and another unidentified coding system. The code system identifier was not included so the terminology used had to be derived based on the format of the code. Fortunately the three coding systems had distinct patterns. LOINC was NNNNN-N, CPT was NNNNN[.N] and the other was XNNN[NN] where N represents a numeric character and X is a literal representation of the letter X.

Missing Dates. There were missing dates from encounters, problems and laboratory results. As some of the analysis, such as the care guidelines analysis, were time-sensitive, some of the results may have been skewed. The missing laboratory dates for the relevant tests came up to 408. However, they probably were from tests done more than one year ago from a facility that has not provided test results for more than a year, which would not have affected the care guidelines.

Different HL7 Data Types and Units. Several HL7 data types were used to report the results including FT (free text), ST (string) and NM (numeric). Numeric results that were given the ST data type were changed to NM. Results that contained the result and a narrative were split. Results values were presented as both percent (e.g. 11%)

while at other times it was presented as a value (e.g., 0.054). While converting the units and data types was relatively trivial, it still demonstrates that data captured as discrete data elements still need to be cleaned and transformed in order to be used in a decision support system.

Ambiguous Terms. Ambiguous items in the problem list and encounter diagnoses made it difficult to assign an appropriate SNOMED CT concept. Examples were "complaint" (n=150) and "pre operative" (n=64). Without additional information, it was not possible to infer what "compliant" and "pre operative" referred to.

9.5.3.2 SNOMED CT-related Challenges

Two main types SNOMED CT-related challenges were encountered: (1) content coverage; and (2) definitions and structure of the hierarchy.

9.5.3.2.1 Content Coverage

The encoding method in **Chapter Seven** was used to encode the problem list and encounter diagnoses while medications were encoded manually. Content coverage with complete matches was 90.8%, 87.2% respectively and 95.2% respectively. There were the usual data quality challenges associated with free text,[27] included spelling mistakes, abbreviations and ambiguous local terms. The partial matches or no matches could be attributed to (1) lack of concepts; (2) lack of synonyms; (3) lack of brand names and (4) inability to post-coordinate.

Lack of concepts. Examples of concepts that were not represented in SNOMED CT were "charting error" (n=92) and "dental abscess early" (n=23). It was possible that these concepts did exist but under a different synonym, in which case the challenge is then a lack of synonym.

Lack of synonyms. The lack of synonyms resulted in more partial matches. For example, terms such as "unstable housing" (639), "hypertension controlled" (n=101) and "emotional distress" (n=83) were partially and incorrectly encoded to **414418009|Housed (finding)|**, **38341003|Hypertensive disorder, systemic arterial (disorder)|** and **69328002|Distress (finding)|** when they should have been encoded with **81877007|Housing problems (finding)|**, **170577003|Good hypertension control (finding)|** and **442347009|Emotional stress (finding)|**.

Lack of brand names. As medication brand names such as "Pegasys" (n=940) and "Fluanxol" (n=660) were mentioned in the encounter diagnosis and did not exist in SNOMED CT, the active ingredients had to be looked up in medical dictionaries in order to find a suitable match. From the limited use of concepts from the **373873005|Pharmaceutical / biologic product (product)|** hierarchy, it is not feasible to use SNOMED CT primarily as a medications terminology because of the incomplete hierarchy and lack of brand names. Dosages were not used in this study so it is unknown how well the content coverage of medications with dosages is recorded in SNOMED CT.

Inability to post-coordinate. There were encounter diagnoses in which the main concept can be encoded but not the modifiers. For example, it was possible to encode "dental carries" but not the "multiple" in "dental carriers multiple" (n=133). There were pre-coordinated concepts such as **134291007|Multiple fractures (disorder)|** and **248528002|Multiple lumps (finding)|** but they were primitive concepts and the defining attributes did not indicate how “multiple” was represented. There was a qualifier concept **255204007|Multiple (qualifier value)|** but it was not accessible via the Concept Model. Another example is the inability to represent “end stage” diseases. For example, “hepatic failure end stage” (n=65) and “liver disease end stage” (n=33) could only be represented as **59927004|Hepatic failure (disorder)|** and **235856003|Disorder of liver (disorder)|** respectively. There were some “end stage” concepts such as **46177005|End stage renal disease (disorder)|** but it was primitive and did not indicate how “end stage” was represented. There was also the qualifier concept **42796001|End-stage (qualifier value)|** but it too could not be accessed via the Concept Model.

9.5.3.2.2 Definitions and Hierarchy

Three main challenges related to the definitions and hierarchy were identified: (1) incomplete concept definitions; (2) concepts in the hierarchy; and (3) incomplete hierarchy.

Incomplete Concept Definitions. Even though concepts from the **373873005|Pharmaceutical / biologic product (product)|** hierarchy are usually defined with a concept from the **105590001|Substance (substance)|** using the concept model attribute **127489000|Has active ingredient (attribute)|**, it was not the case for these six concepts. For example, **90332006|Acetaminophen (product)|** is defined with **127489000|Has active ingredient (attribute)|= 387517004|acetaminophen (substance)|** but **384953001|Antidiabetic preparation (product)|** was not defined with **373245004|Antidiabetic agent (substance)|**.

Concepts in the Hierarchy. It is important to know how the SNOMED CT hierarchy is structured especially when conducting subsumption queries. For example, when using **73211009|Diabetes mellitus (disorder)|** as the predicate expression, chronic conditions such as **46635009|Diabetes mellitus type 1 (disorder)|** and **44054006|Diabetes mellitus type 2 (disorder)|** may be what is wanted but not acute conditions such as **11687002|Gestational diabetes mellitus (disorder)|**. Therefore more complex predicate expressions with multiple include and exclude criteria may be necessary to attain the desired results.

Incomplete Hierarchy. There is some incompleteness in the SNOMED CT hierarchy. For example, the concepts **82156005|Vitamin A preparation (product)|** and **11563006|Vitamin D preparation (product)|** should be supertypes concepts **of 29987004|Vitamins A and D preparation (product)|** but are not (refer to Figure 9-14).

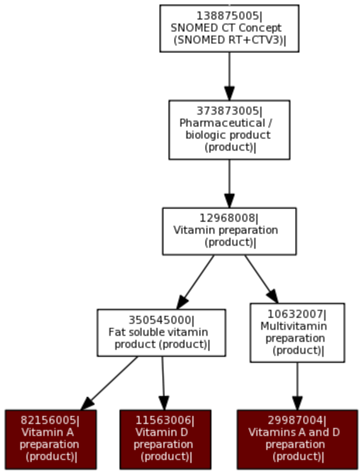


Figure 9-14. Location of "82156005|Vitamin A preparation (product)|", "11563006|Vitamin D preparation (product)|" and "29987004|Vitamins A and D preparation (product)|" in the hierarchy.

9.6 References

[1] SNOMED CT Value Proposition. http://www.ihtsdo.org/snomed-ct/whysnomedct/snomedfeatures. Last accessed: September 18, 2013.

[2] Members of IHTSDO. http://www.ihtsdo.org/members. Last accessed: September 18, 2013.

[3] Lee, D., de Keizer, N., Lau, F., & Cornet, R. (2013). Literature review of SNOMED CT use. Journal of the American Medical Informatics Association. http://www.ncbi.nlm.nih.gov/pubmed/23828173.

[4] Zetterberg, C., Ahlzén, K., Ericsson, E., & Kron, B. (2012). An Example of a Multi-Professional Process-Oriented Structured Documentation Bound to SNOMED CT. Studies in Health Technology and Informatics, 180, 1215-7. http://www.ncbi.nlm.nih.gov/pubmed/22874405.

[5] Patrick, J. D., Ryan, A., & Herkes, R. (2008). Introduction of Enhancement Technologies into the Intensive Care Service, Royal Prince Alfred Hospital, Sydney. Health Information Management Journal, 37(1), 40-5. http://www.ncbi.nlm.nih.gov/pubmed/18245864.

[6] Zetterberg, C., Ahlzén, K., Ericsson, E., & Kron, B. (2012). An Example of a Multi-Professional Process-Oriented Structured Documentation Bound to SNOMED CT. Studies in Health Technology and Informatics, 180, 1215-7. http://www.ncbi.nlm.nih.gov/pubmed/22874405.

[7] Tvede, I., Bredegaard, K., & Andersen, J. S. (2010). Quality improvements based on detailed and precise terminology. Studies in Health Technology and Informatics, 155, 71-7. http://www.ncbi.nlm.nih.gov/pubmed/20543312.

[8] Elkin, P. L., Froehling, D., Wahner-Roedler, D., Trusko, B., Welsh, G., Ma, H., ... & Brown, S. H. (2008). NLP-based identification of pneumonia cases from free-text radiological reports. In AMIA Annual Symposium Proceedings (Vol. 2008, p. 172-6). American Medical Informatics Association. http://www.ncbi.nlm.nih.gov/pubmed/18998791.

[9] Liaw, S. T., Chen, H. Y., Maneze, D., Taggart, J., Dennis, S., Vagholkar, S., & Bunker, J. (2012). Health reform: Is routinely collected electronic information fit for purpose? Emergency Medicine Australasia, 24(1), 57-63. http://www.ncbi.nlm.nih.gov/pubmed/22313561.

[10] Farfán Sedano, F. J., Terron Cuadrado, M., García Rebolledo, E. M., Castellanos Clemente, Y., & Serrano Balazote, P. (2009). Implementation of SNOMED CT to the medicines database of a general hospital. Studies in Health Technology and Informatics, 148, 123-30. http://www.ncbi.nlm.nih.gov/pubmed/19745242.

[11] Kim, H. Y., & Park, H. (2012). Development and evaluation of data entry templates based on the entity-attribute-value model for clinical decision support of pressure ulcer wound management. International journal of medical informatics, 81(7), 485-492. http://www.ncbi.nlm.nih.gov/pubmed/22079242.

[12] Benoit, S. R., L Clifford McDonald, M. D., English, R., & Tokars, J. I. (2011). Automated surveillance of Clostridium difficile infections using BioSense. Infection Control and Hospital Epidemiology, 32(1), 26-33. http://www.ncbi.nlm.nih.gov/pubmed/21128815.

[13] American Diabetes Association. Diagnosis and Classification of Diabetes Mellitus. Diabetic Care. 2001 Jan:33(Supplement 1):S62-S69.

[14] Public Health Agency of Canada. Diabetes in Canada: Facts and Figures from a Public Health Perspective. 2011. http://www.phac-aspc.gc.ca/cd-mc/publications/diabetes-diabete/facts-figures-faits-chiffres-2011/pdf/facts-figures-faits-chiffres-eng.pdf.

[15] Wright, A., Pang, J., Feblowitz, J. C., Maloney, F. L., Wilcox, A. R., Ramelson, H. Z., ... & Bates, D. W. (2011). A method and knowledge base for automated inference of patient problems from structured data in an electronic medical record. Journal of the American Medical Informatics Association, 18(6), 859-867. http://www.ncbi.nlm.nih.gov/pubmed/21613643

[16] Hahn, K. A., Ferrante, J. M., Crosson, J. C., Hudson, S. V., & Crabtree, B. F. (2008). Diabetes flow sheet use associated with guideline adherence. The Annals of Family Medicine, 6(3), 235-238. http://www.ncbi.nlm.nih.gov/pubmed/18474886.

[17] Jao, C., Hier, D., & Galanter, W. (2008, November). Automating the maintenance of problem list documentation using a clinical decision support system. In AMIA... Annual Symposium proceedings/AMIA Symposium. AMIA Symposium (p. 989). http://www.ncbi.nlm.nih.gov/pubmed/18998927.

[18] Galanter, W. L., Hier, D. B., Jao, C., & Sarne, D. (2010). Computerized physician order entry of medications and clinical decision support can improve problem list documentation compliance. International Journal of Medical Informatics, 79(5), 332-338. http://www.ncbi.nlm.nih.gov/pubmed/18599342.

[19] Carpenter, J. D., & Gorman, P. N. (2002). Using medication list--problem list mismatches as markers of potential error. In Proceedings of the AMIA Symposium (p. 106-10). American Medical Informatics Association. http://www.ncbi.nlm.nih.gov/pubmed/12463796.

[20] Adaji, A., Schattner, P., Jones, K. M., Beovich, B., & Piterman, L. (2013). Care planning and adherence to diabetes process guidelines: Medicare data analysis. Australian Health Review, 37(1), 83-87. http://www.ncbi.nlm.nih.gov/pubmed/23157923.

[21] Amed, S., Nuernberger, K., McCrea, P., Reimer, K., Krueger, H., Aydede, S. K., ... & Collet, J. P. (2013). Adherence to Clinical Practice Guidelines in the Management of Children, Youth, and Young Adults with Type 1 Diabetes—A Prospective Population Cohort Study. The Journal of Pediatrics. http://www.ncbi.nlm.nih.gov/pubmed/23523280.

[22] Bernstein, K., & Andersen, U. (2008). Managing Care Pathways combining SNOMED CT, Archetypes and an Electronic Guideline System. Studies in Health Technology and Informatics, 136, 353-8. http://www.ncbi.nlm.nih.gov/pubmed/18487756.

[23] Hrabak, K. M., Campbell, J. R., Tu, S. W., McClure, R., & Weida, R. T. (2007). Creating interoperable guidelines: requirements of vocabulary standards in immunization decision support. Studies in Health Technology and Informatics, 129(Pt 2), 930-4. http://www.ncbi.nlm.nih.gov/pubmed/17911852.

[24] Ahmadian, L., De Keizer, N. F., & Cornet, R. (2009). The use of SNOMED CT for representing concepts used in preoperative guidelines. Studies in Health Technology and Informatics, 150, 658-62. http://www.ncbi.nlm.nih.gov/pubmed/19745393.

[25] Avillach, P., Joubert, M., & Fieschi, M. (2008). Improving the quality of the coding of primary diagnosis in standardized discharge summaries. Health Care Management Science, 11(2), 147-151. http://www.ncbi.nlm.nih.gov/pubmed/18581821.

[26] Bowen M. Understanding, evaluating and enhancing electronic medical record adoption in a primary care setting: A programme to improve electronic medical record data quality and its effect on family practice provision of incentivized and enhanced care for chronic disease patients. MSc Thesis. University of Victoria: Canada. 2013.

[27] Lee, D. H., Lau, F. Y., & Quan, H. (2010). A method for encoding clinical datasets with SNOMED CT. BMC medical informatics and decision making, 10(1), 53. http://www.ncbi.nlm.nih.gov/pubmed/20849611.

10. LESSONS LEARNED

10.1 Introduction

This chapter provides a summary of this PhD research and is divided into five sections. First, a summary of the findings, in which the findings of this research are tied back to the three research questions. Second, the limitations for each of the studies. Third, future work that can be continued from where this research left off. Fourth, how this research has contributed to the current state of knowledge. Fifth, the conclusion.

10.2 Summary of Findings

The purpose of this research was to answer the overarching research question "How can the clinical value of SNOMED CT be demonstrated in the primary health care setting to enhance patient care?" Three specific questions were posed to help answer that question. This section reviews each of the questions and how they have been answered in this research.

The first question was "to what extent is SNOMED CT being adopted?" In **Chapter Three**, the literature review of 488 SNOMED CT related papers from 22 countries that were published in the scientific literature showed that the majority of the papers have centred on comparing SNOMED CT to other terminology systems and prospective content coverage of local terms. This is still encouraging as work is being done to harmonise SNOMED CT with other standardised 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 steadily increased although not many are in operational clinical settings. Data capture using SNOMED CT has been done through check lists, questionnaires, data entry templates, auto-complete fields and indexing via natural language processing algorithms. Data retrieval has mainly been limited to the use of SNOMED CT synonyms to search free text records and natural language processing algorithms. Overall, no studies were found that quantitatively demonstrated how SNOMED CT has improved care in clinical settings. In **Chapter Four**, the interviews with 14 individuals from 13 health care organisations across eight countries examined the implementation of SNOMED CT in healthcare organisations in terms of design, use and maintenance. Data capture is similar to the rest of the literature while data retrieval has included data retrieval subsets, subsumption through the SNOMED CT hierarchy and through the defining attributes. SNOMED CT specific challenges included creation of post-coordinated expressions, compilation of subsets and subsumption queries. Although some health care organisations have successfully adopted SNOMED CT, none have reached the stage of evaluating the benefits of SNOMED CT in terms of quality of care.

The second question was "how can SNOMED CT be used to represent clinical data in primary health care settings in a manner that is accurate, consistent and unambiguous?" The answer is two-fold. First, there must be assurance that the quality of SNOMED CT is consistent. In **Chapter Six**, 89 verification rules were identified, organised into the SNOMED CT auditing framework, represented in a machine-readable format and applied to three country level extensions. Auditing extensions before implementing them is important because they may contain errors, as demonstrated in this study. The auditing rules compiled in this study are relatively easy to use and do not require a high level of technical skills. By consistently applying these verification rules before extensions are distributed and implemented, health care organisations will be able to have a greater confidence in the veracity of the extensions and will not encounter unexpected results when using to test for equivalency and subsumption. Second, an automated encoding algorithm that is based on how SNOMED CT concepts are modelled is needed to ensure clinical data are encoded accurately, consistently and unambiguously. There are many complex components that go into the development of an automated post-coordinated encoding method. In **Chapter Seven**, an encoding method was developed that was comprised of three parts: steps to improve the content coverage and selection of terms, scenarios on how to combine two or more pre-coordinated concepts, and cleaning up the post-coordinated expression. The steps for selecting a term and the scenarios for combining concepts through the proposed Extended Concept Model and encoding templates helped to ensure the most appropriate concept was selected and combined into a post-coordinated expression. A review of the most commonly used post-coordinated expressions and a random sample of the post-coordinated expressions show that they were accurate between 82 and 85% of the time.

The third question was "how can SNOMED CT be used to enhance decision support capabilities in primary health care settings such as identifying high-risk patients and drug-allergy alerts?" In **Chapter Eight**, it was demonstrated that using SNOMED CT queries for seven types of chronic diseases was a straightforward process and that the subsumption queries were more effective than just looking up a set of concept identifiers. The optimised structural subsumption method and storage methods also helped to improve the efficiency of retrieving concepts without sacrificing the effectiveness. In **Chapter Nine**, a knowledge base for inferring clinical conditions based on problem lists, medications, diagnostic billing codes and laboratory results, and a set of clinical care guidelines for diabetes were used to examine a SNOMED CT-encoded primary care dataset. The results showed ways in which clinical documentation could be improved and areas where suggestions and reminders can be incorporated to ensure patients are receiving care according to guidelines. These criteria were programmed into an electronic medical records system prototype to demonstrate how they could be practically implemented as decision support aids and included suggesting billing diagnostic codes, suggesting the prescribing of medications, reminding the ordering of laboratory tests, alerting for out of range values and alerting of potential drug-allergy interactions. While these inference knowledge base and care guidelines are rudimentary, they demonstrated that improving data quality in terms of completeness is not difficult, and that implementing basic decision support features does not need to be complex.

10.3 Limitations

This section describes the limitations for each of the studies that were conducted throughout this research.

10.3.1 Literature Review (Chapter Three)

There were three limitations of the literature review. First, publication bias. Only papers catalogued in PubMed and Embase, and only papers that were published in English or had an English abstract were reviewed and classified. The review of two inventories of SNOMED CT use and the papers included in the study showed that the majority of implementations are not published in the scientific literature or are not captured in PubMed or Embase.

Second, the lack of evaluation of the usage categories and definitions. Although the majority of the usage categories were based on a previous study and the authors agreed upon the definitions that were assigned to each usage category, no evaluation was conducted to determine the effects of refining the definitions.

Third, 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. However, the interrater agreement between the co-authors for the 65 papers was not tracked. Although the initial interrater agreement was not tracked, any differences in the classifying of these 65 papers were reconciled.

10.3.2 Implementation Survey (Chapter Four)

The main limitation of the survey of SNOMED CT implementations was the small sample size. However, there are very few known SNOMED CT implementations in clinical care settings and the 13 interviews covered an adequate proportion of known implementations. Overall, 50 individuals were invited to participate in the study.

10.3.3 Conceptual Framework (Chapter Five)

The SNOMED CT Implementation Conceptual Framework only takes into account three aspects of implementation (i.e., encoding; auditing; and retrieval). It does not include factors such as user interfaces, organisational factors, education, and how it should be implemented in conjunction with an information model. The limitations of the SNOMED CT Extensions Auditing Framework are discussed below in **Section 10.3.4**. The SNOMED CT Clinical Value Framework was developed mainly based on the requirements of the Canada Health Infoway Clinical Value Targets. Although it was sufficient to categories the different ways of demonstrating the clinical value of SNOMED CT, it may not be generalisable to the needs of other health care organisations.

10.3.4 Auditing Method (Chapter Six)

There were two main limitations with the auditing method. First, the verification rules that were used in the extensions auditing method in this study were limited to the verification of identifiers, value sets, dependencies, the Machine Readable Concept Model, inferred relationships and occurrences. The guidelines were also limited to Release Format 1 (RF1), the original format in which SNOMED CT was published, which is gradually replaced by a newer format, Release Format 2 (RF2). The auditing focused on consistency (i.e., "the knowledge should not be self-contradictory") as opposed to completeness (i.e., "it should have the necessary knowledge") or correctness (i.e., "the knowledge should be faithful to the real world"). Second, these verification rules were not externally reviewed and have not been verified.

10.3.5 Encoding Method (Chapter Seven)

There were three main limitations with the encoding method. First, this encoding method is a "brute force" method in the sense that parts of speech are not taken into account and the matching algorithm attempts to encode all words, which explains why the "light" in "vaginal spotting light" would be encoded as **56242006|Light, electromagnetic radiation (physical force)|**. Second, the encoding method was developed based on the problem lists and encounter diagnoses from a single primary care practice and may not be generalisable to all practices or medical domains. Third, neither the encoding method nor the results of the encoding have been externally verified or compared to other encoding methods.

10.3.6 Retrieval Method (Chapter Eight)

There were two main limitations with the retrieval method. First, the retrieval method has not been verified by a clinical terminologist or clinician. Second, the optimised structural subsumption retrieval method was not compared against the results of using a description logic classifier. However, the results were compared against the structural subsumption method, which indicated there was no loss in effectiveness.

10.3.7 Clinical Value (Chapter Nine)

The demonstration of the value of SNOMED CT was built on the encoding and retrieval method previously described and the limitations of those methods are not repeated here. For this study, four limitations were identified.

First, this study was limited to the data extracted, which excluded other clinical documentation, such as subject-object-assessment-plan notes and referrals. About 10% of the problem list and encounter diagnoses were not encoded to SNOMED CT, which may have skewed the results. However, problems, encounter diagnoses and medications that occurred ≥10 times were manually reviewed to ensure they were encoded and encoded correctly. It cannot be expected that all terms be manually reviewed as there were over 266,000 encounter diagnoses, 20,000 problems and 73,000 medications. In addition, documentation limitations in the EMR may have lowered the estimates of clinical quality (e.g., paper prescriptions, laboratory results managed in collaboration with a specialist and not sent to the EMR). The missing dates in the laboratory results may have skewed the results of the analysis of adherence to care guidelines, which was time-sensitive.

Second, the guidelines used in this study differed from each other. The guidelines used in this study for levels of haemoglobin A1c were 7% for the criteria by Wright, et al., and 8% for the criteria by Hahn, et al. Results of this test are also dependent on age and gender. As those data elements were not available, the analysis was based on the raw numbers provided by the papers. The laboratory test results were imported from HL7 messages, which usually included a reference range and abnormal flags. The reference range for haemoglobin A1c was generally between 4.2 to 4.5% and 6.0 to 6.2%. Therefore an individual who had a result of 6.5% would have received an "abnormal" or "high" flag. However, based on the 7% and 8%, it would have been considered normal. In this study, the exact guidelines were used and not the interpretations included.

Third, while the smoking status was determined via the encounter diagnoses, there were medications that had the active ingredient of nicotine (i.e., Nicoderm patches; Prostep nicotine patch 22 mg; Habitrol patch; and Nicotrol NS) that could possibly be used to indicate that the smoking status was being monitored. However, it was beyond the scope of this study but may be looked at in the future.

Fourth, the results were not verified. Although the results were briefly reviewed by a clinician from the primary care clinic in which the data was extracted, the results were not compared to any other reports.

10.4 Future Work

There is still much work to be done. First, as the use of SNOMED CT has shifted from Release Format 1 to Release Format 2, there is a need to update the verification rules for auditing extensions. As the CA and UK extensions were only available in the RF1 format, the verification rules in this study were written to conform to the RF1 specification. Additional work is needed to transform these verification rules to RF2. As the general rules have already been enumerated, converting the rules to RF2, although not trivial, should not be overly difficult. Future work would also include extending these rules to include validation and to incorporate the auditing method as part of the development and publishing processes.

Second, the proposed Extended Concept Model needs to be expanded beyond the scope of **123037004|Body structure (body structure)|** and **49755003|Morphologically abnormal structure (morphologic abnormality)|** and additional encoding templates need to be developed to ensure that the appropriate concepts are used in post-coordinated expressions. As the encoding of individual phrases thus far has been through a "brute force" method, medical language processing should be included to identify the parts of speech and context of clinical statements. It is expected that this encoding method would still be useful as medical language processing tools such as cTakes and MetaMap do not currently output post-coordinated expressions. Additional encoding templates need to be developed to handle different types of clinical statements and the addition of synonyms identified from encoding datasets needs to be assimilated back into SNOMED CT either as extensions or into the core.

Third, while the optimised structural subsumption is an interim solution, there is a need to develop description logic tools that can be used in clinical settings to retrieve SNOMED CT-encoded data. While there have been publications on the use of description logic to assist in quality assurance and mapping, there are no known publications in the scientific literature that deal specifically with how to use a description logic classifier to compare two SNOMED CT post-coordinated expressions for equivalency or subsumption.

Fourth, continued efforts need to be focused on representing clinical guidelines that have been encoded with SNOMED CT and other terminology systems into standardised formats such the openEHR archetypes and Clinical Models.

10.5 Contributions in this Research

There are five main areas in which this research contributes to the current state of knowledge. First, verification rules for auditing extensions. Second, the automated post-coordinated encoding method. Third, the optimised data storage method for storing post-coordinated expressions. Forth, the optimised structural subsumption method for data retrieval. Fifth, the examples of demonstrating the clinical value of SNOMED CT.

Verification Rules for Auditing Extensions. Eighty-nine verification rules for identifying errors in SNOMED CT extensions were identified, organised into the SNOMED CT Auditing Extensions Framework and represented in a machine-readable format that can be re-used to verify extensions before they are published. Although these verification rules are not ground-breaking research, the need for having them was confirmed when errors were identified in three country level extensions that have been publicly published. As more namespaces have been assigned and health care organisations start to develop their own extensions, the verification rules will become increasingly important to ensure the extensions are consistent.

Automated Post-coordinated Encoding Method. The need for an automated post-coordinated encoding method is evident by the errors contained in the examples of post-coordinated expressions in the literature. Post-coordination is not easy, and while the encoding method developed is not perfect, the sorting algorithm to assist in selecting the appropriate concept when there are more than one exact matches, the proposed Extended Concept Model to assist in selecting the appropriate body structures, the encoding templates that help in constructing a post-coordinated expression, and the general scenarios of modelling two concepts have proven to be useful in encoding a primary care dataset.

Optimised Data Storage Method for Post-coordinated Expressions. The Technical Implementation Guide suggested three methods for storing post-coordinated expressions in relational databases but does not indicate a preference for the most suitable method. After an analysis of the three methods, only one was deemed suitable. As it was not optimised for structural subsumption, a fourth method was proposed that includes storing the pre-computed long normal form and the contextual values in discrete data elements and supports the optimised structural method developed.

Optimised Structural Subsumption Method. Even though the use of a description logic classifier is the preferred method of testing expressions for equivalency and subsumption, an interim solution is needed as the use of description logic classifiers is not common place in health care organisations. An optimised structural subsumption method was developed based on three types of predicate expressions to retrieve an initial set of candidate expressions before using the long normal forms. The reduction of the number of expressions that were compared ranged from 62.9% to 99.9% for seven chronic disease categories tested.

Examples of Demonstrating the Clinical Value of SNOMED CT. With the rich dataset that was extracted from a primary care clinic and encoded with SNOMED CT, the clinical value of SNOMED CT was demonstrated though the proposed SNOMED CT Clinical Value Framework. They included suggestions of billing diagnostic codes (using the ICD-9-CM reimbursement cross map), suggestions for adding a problem into the problem list (based on the criteria by Wright, et al.) suggestions for ordering laboratory tests based on the encounter diagnosis (based on the criteria by Hahn, et al.), suggestions for adding non-transitive illnesses (i.e., a disease rather than a symptom or finding) from the encounter diagnosis into the problem list (as a result of the encounter diagnosis being a subtype of a specific concept), suggestions for prescribing medications (sans allergic medications and currently prescribed medications) based on out of range results, alerts for potential drug-allergy interactions (based on the active ingredient of medications and the causative agent in allergies) and suggestions for improving the organisation of a clinical summary (based on the assignment of predicate expressions to sections of the clinical summary).

10.6 Conclusion

In this PhD research, the potential clinical value of SNOMED CT was demonstrated by improving the completeness of clinical records and facilitating decision support features such as alerting clinicians to potential drug-allergy interactions, and reminding clinicians to order routine tests. The realisation of the potential clinical value was based upon the accurate and unambiguous manner in which clinical terms were encoded using the encoding method, the efficient and effective retrieval of relevant concepts using the retrieval method, and to a lesser extent, the ensuring that the concepts used were consistent using the auditing method.

11. Bibliography

[PowerPoint Presentation] Canada Health Infoway. (2011). Health System Use Demonstration Projects Overview.

Adaji, A., Schattner, P., Jones, K. M., Beovich, B., & Piterman, L. (2013). Care planning and adherence to diabetes process guidelines: Medicare data analysis. Australian Health Review, 37(1), 83-87. http://www.ncbi.nlm.nih.gov/pubmed/23157923.

Ahmadian, L., De Keizer, N. F., & Cornet, R. (2009). The use of SNOMED CT for representing concepts used in preoperative guidelines. Studies in Health Technology and Informatics, 150, 658-62. http://www.ncbi.nlm.nih.gov/pubmed/19745393.

Amed, S., Nuernberger, K., McCrea, P., Reimer, K., Krueger, H., Aydede, S. K., ... & Collet, J. P. (2013). Adherence to Clinical Practice Guidelines in the Management of Children, Youth, and Young Adults with Type 1 Diabetes—A Prospective Population Cohort Study. The Journal of Pediatrics. http://www.ncbi.nlm.nih.gov/pubmed/23523280.

American Diabetes Association. Diagnosis and Classification of Diabetes Mellitus. Diabetic Care. 2001 Jan:33(Supplement 1):S62-S69.

Andrews, J. E., Richesson, R. L., & Krischer, J. (2007). Variation of SNOMED CT coding of clinical research concepts among coding experts. Journal of the American Medical Informatics Association, 14(4), 497-506. http://www.ncbi.nlm.nih.gov/pubmed/17460128.

Arnot - Smith, J., & Smith, A. F. (2010). Patient safety incidents involving neuromuscular blockade: analysis of the UK National Reporting and Learning System data from 2006 to 2008. Anaesthesia, 65(11), 1106-1113. http://www.ncbi.nlm.nih.gov/pubmed/20840604.

Avillach, P., Joubert, M., & Fieschi, M. (2008). Improving the quality of the coding of primary diagnosis in standardized discharge summaries. Health Care Management Science, 11(2), 147-151. http://www.ncbi.nlm.nih.gov/pubmed/18581821.

Babbie, E. (2012). The practice of social research. CengageBrain. com.

Bakhshi-Raiez, F., Ahmadian, L., Cornet, R., de Jonge, E., & de Keizer, N. F. (2010). Construction of an interface terminology on SNOMED CT: Generic approach and its application in intensive care. Methods of Information in Medicine, 49(4), 349-59. http://www.ncbi.nlm.nih.gov/pubmed/20582384.

Bakhshi-Raiez, F., de Keizer, N. F., Cornet, R., Dorrepaal, M., Dongelmans, D., & Jaspers, M. W. (2012). A usability evaluation of a SNOMED CT based compositional interface terminology for intensive care. International Journal of Medical Informatics, 81(5), 351-362. http://www.ncbi.nlm.nih.gov/pubmed/22030036.

Bennett, K. J., & Steen, C. (2010). Electronic medical record customization and the impact upon chart completion rates. Family Medicine, 42(5), 338-42. http://www.ncbi.nlm.nih.gov/pubmed/20461565.

Benoit, S. R., L Clifford McDonald, M. D., English, R., & Tokars, J. I. (2011). Automated surveillance of Clostridium difficile infections using BioSense. Infection Control and Hospital Epidemiology, 32(1), 26-33.

Bernstein, K., & Andersen, U. (2008). Managing Care Pathways combining SNOMED CT, Archetypes and an Electronic Guideline System. Studies in Health Technology and Informatics, 136, 353-8. http://www.ncbi.nlm.nih.gov/pubmed/18487756.

Bird, L., Brooks, C., Cheong, Y. C., & Tun, N. N. (2011). A logical approach to semantic interoperability in healthcare. Studies in Health Technology and Informatics. v168, 1-9. http://www.ncbi.nlm.nih.gov/pubmed/21893905.

Bowen M. Understanding, evaluating and enhancing electronic medical record adoption in a primary care setting: A programme to improve electronic medical record data quality and its effect on family practice provision of incentivized and enhanced care for chronic disease patients. MSc Thesis. University of Victoria: Canada. 2013.

Boyatzis, R. (1998). Transforming qualitative information: Thematic analysis and code development. Thousand Oaks, California: Sage Publications, Incorporated.

Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2). pp. 77-101.

Brooks, R. J. (2007). Conceptual modelling: framework, principles, and future research. (Management Science Working Paper Series). Lancaster University: The Department of Management Science.

Bryman, A. (2012). Chapter 15: Interviewing in qualitative research. In Social research methods. Oxford university press.

Canada Health Infoway. SNOMED CT RefSets. https://infocentral.infoway-inforoute.ca/2_Standards/1_pan-Canadian_Standards/Terminology/4_SNOMED_CT_Terminologies.

Canada Health Infoway. SNOMED CT. https://infocentral.infoway-inforoute.ca/2_Standards/1_pan-Canadian_Standards/Terminology/4_SNOMED_CT_Terminologies.

Canadian Institute for Health Information. (2012). Draft Pan-Canadian Primary Health Care Electronic Medical Record Content Standard, Version 2.1 – Implementation Guide. https://secure.cihi.ca/free_products/PHC_EMR_CS_Implementation_Guide.pdf.

Canadian Institute for Health Information. (2012). Pan-Canadian Primary Health Care Indicator Report. https://secure.cihi.ca/free_products/Pan-Canadian_PHC_Indicator_Update_Report_en_web.pdf.

Cao, F., Sun, X., Wang, X., Li, B., Li, J., & Pan, Y. (2011). Ontology-based knowledge management for personalized adverse drug events detection. Studies in health technology and informatics, 169, 699-703. http://www.ncbi.nlm.nih.gov/pubmed/21893837.

Carpenter, J. D., & Gorman, P. N. (2002). Using medication list--problem list mismatches as markers of potential error. In Proceedings of the AMIA Symposium (p. 106-10). American Medical Informatics Association. http://www.ncbi.nlm.nih.gov/pubmed/12463796.

Centre for Health Information Research and Development Implementing SNOMED CT within national electronic record solutions. http://www.chirad.org.uk/paper_one.htm.

Centres for Medicare and Medicaid Services. Eligible Professional Meaningful Use Core Measures Measure 3 of 15. http://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/downloads/3MaintainProblemList.pdf.

Ceusters, W., Smith, B., Kumar, A., & Dhaen, C. (2004). Mistakes in medical ontologies: where do they come from and how can they be detected? Studies in Health Technolology and Informatics. 102:145-63. http://www.ncbi.nlm.nih.gov/pubmed/15853269.

Cimino, J. J. (2006). In defense of the Desiderata. Journal of biomedical informatics, 39(3), 299-306. http://www.ncbi.nlm.nih.gov/pubmed/16386470.

Clinical Value: "Meaningful Use" in Canada. http://infowayconnects.infoway-inforoute.ca/blog/clinicians-and-health-informatics/220-clinical-value-meaningful-use-in-canada.

Conceptual modelling for shared understanding. http://learningforsustainability.net/social_learning/conceptual_modelling.php.

Cornet, R., & Abu-Hanna, A. (2005). Description logic-based methods for auditing frame-based medical terminological systems. Artificial Intelligence in Medicine, 34(3), 201-217. http://www.ncbi.nlm.nih.gov/pubmed/15994071.

Cornet, R., & de Keizer, N. (2008). Forty years of SNOMED: a literature review. BMC medical informatics and decision making, 8(Suppl 1), S2. http://www.ncbi.nlm.nih.gov/pubmed/19007439.

Cuggia, M., Bayat, S., Garcelon, N., Sanders, L., Rouget, F., Coursin, A., & Pladys, P. (2009). A full-text information retrieval system for an epidemiological registry. Studies in health technology and informatics, 160(Pt 1), 491-495. http://www.ncbi.nlm.nih.gov/pubmed/20841735.

Daly, J., Kellehear, A. & Gliksman, M. (1997). The public health researcher: A methodological approach. Melbourne, Australia: Oxford University Press.

de Lusignan, S., Chan, T., & Jones, S. (2011). Large complex terminologies: more coding choice, but harder to find data--reflections on introduction of SNOMED CT (Systematized Nomenclature of Medicine--Clinical Terms) as an NHS standard. Informatics in Primary Care, 19(1), 3. http://www.ncbi.nlm.nih.gov/pubmed/22118330.

De Silva, T. S., MacDonald, D., Paterson, G., Sikdar, K. C., & Cochrane, B. (2011). Systematized nomenclature of medicine clinical terms (SNOMED CT) to represent computed tomography procedures. Computer methods and programs in biomedicine, 101(3), 324-329. http://www.ncbi.nlm.nih.gov/pubmed/21316117.

Design Science Research in Information Systems. http://desrist.org/design-research-in-information-systems. Last accessed: July 24, 2013.

Dolin, R. H., Mattison, J. E., Cohn, S., Campbell, K. E., Wiesenthal, A. M., Hochhalter, B., ... & Zingo, C. (2004). Kaiser Permanente's convergent medical terminology. Medinfo, 11(Pt 1), 346-50. http://www.ncbi.nlm.nih.gov/pubmed/15360832.

Dolin, R. H., Spackman, K. A., & Markwell, D. (2002). Selective retrieval of pre-and post-coordinated SNOMED concepts. In Proceedings of the AMIA Symposium (p. 210). American Medical Informatics Association.

Elhanan, G., Perl, Y., & Geller, J. (2011). A survey of SNOMED CT direct users, 2010: impressions and preferences regarding content and quality. Journal of the American Medical Informatics Association, 18(Suppl 1), i36-i44. http://www.ncbi.nlm.nih.gov/pubmed/21836159.

Elkin, P. L., Froehling, D., Wahner-Roedler, D., Trusko, B., Welsh, G., Ma, H., ... & Brown, S. H. (2008). NLP-based identification of pneumonia cases from free-text radiological reports. In AMIA Annual Symposium Proceedings (Vol. 2008, p. 172-6). American Medical Informatics Association. http://www.ncbi.nlm.nih.gov/pubmed/18998791.

Elkin, P. L., Ruggieri, A. P., Brown, S. H., Buntrock, J., Bauer, B. A., Wahner-Roedler, D., ... & Bergstrom, L. (2001). A randomized controlled trial of the accuracy of clinical record retrieval using SNOMED-RT as compared with ICD9-CM. In Proceedings of the AMIA Symposium (p. 159). American Medical Informatics Association.

Elkin, P. L., Trusko, B. E., Koppel, R., Speroff, T., Mohrer, D., Sakji, S., ... & Brown, S. H. (2010). Secondary use of clinical data. Studies in Health Technology and Informatics, 155, 14-29. http://www.ncbi.nlm.nih.gov/pubmed/20543306.

Farfán Sedano, F. J., Terron Cuadrado, M., García Rebolledo, E. M., Castellanos Clemente, Y., & Serrano Balazote, P. (2009). Implementation of SNOMED CT to the medicines database of a general hospital. Studies in Health Technology and Informatics, 148, 123-30. http://www.ncbi.nlm.nih.gov/pubmed/19745242.

Floyd, C. (1984). A systematic look at prototyping. In Approaches to prototyping (pp. 1-18). Springer Berlin Heidelberg.

Fung K. W., McDonald C. & Srinivasan S. The UMLS-CORE project: a study of the problem list terminologies used in large healthcare institutions. Journal of the Americal Medical Informatics Association. 2010 Nov-Dec;17(6):675-80. http://www.ncbi.nlm.nih.gov/pubmed/20962130.

Fung, K. W., Xu, J., Rosenbloom, S. T., Mohr, D., Maram, N., & Suther, T. (2011). Testing Three Problem List Terminologies in a Simulated Data Entry Environment. In AMIA Annual Symposium Proceedings (Vol. 2011, p. 445). American Medical Informatics Association. http://www.ncbi.nlm.nih.gov/pubmed/22195098.

Galanter, W. L., Hier, D. B., Jao, C., & Sarne, D. (2010). Computerized physician order entry of medications and clinical decision support can improve problem list documentation compliance. International Journal of Medical Informatics, 79(5), 332-338. http://www.ncbi.nlm.nih.gov/pubmed/18599342.

Ghosh, T. P. (2013). Information Systems Control and Audit. Systems Development Life Cycle Methodology. The Institute of Chartered Accountants of India.

Grain, H. (2010). Clinical terminology. Studies in Health Technology and Informatics, 151, 70-83. http://www.ncbi.nlm.nih.gov/pubmed/20407153.

Green, J. M., Wilcke, J. R., Abbott, J., & Rees, L. P. (2006). Development and evaluation of methods for structured recording of heart murmur findings using SNOMED-CT® post-coordination. Journal of the American Medical Informatics Association, 13(3), 321-333. http://www.ncbi.nlm.nih.gov/pubmed/16501179.

Guest, G. (2012). Applied thematic analysis. Thousand Oaks, California: Sage Publications, Incorporated.

Hahn, K. A., Ferrante, J. M., Crosson, J. C., Hudson, S. V., & Crabtree, B. F. (2008). Diabetes flow sheet use associated with guideline adherence. The Annals of Family Medicine, 6(3), 235-238. http://www.ncbi.nlm.nih.gov/pubmed/18474886.

Hardiker, N. R., Casey, A., Coenen, A., & Konicek, D. (2006). Mutual enhancement of diverse terminologies. In AMIA Annual Symposium Proceedings (Vol. 2006, p. 319-23). American Medical Informatics Association. http://www.ncbi.nlm.nih.gov/pubmed/17238355.

HIMSS EMR Adoption Model. http://www.himssanalytics.org/hc_providers/emr_adoption.asp. Last accessed: July 24, 2013.

Holmes, C. (2011). The problem list beyond meaningful use. Part I: The problems with problem lists. Journal of AHIMA/American Health Information Management Association, 82(2), 30-3. http://www.ncbi.nlm.nih.gov/pubmed/21337850.

Hrabak, K. M., Campbell, J. R., Tu, S. W., McClure, R., & Weida, R. T. (2007). Creating interoperable guidelines: requirements of vocabulary standards in immunization decision support. Studies in Health Technology and Informatics, 129(Pt 2), 930-4. http://www.ncbi.nlm.nih.gov/pubmed/17911852.

Hurrell M.J., Monk T.G., Nicol A., Norton A.N., Reich D.L., Walsh J.L.. (2011). Implementation of a standards-based, CDA-compliant anesthesia record. Journal of Clinical Monitoring and Computing, 25, 5-16.

Hussain, F., Muller, F., Husain, E. (2010). Under-reporting of invasive malignant melanomas in North East of Scotland. British Journal of Dermatology 2010; 163: 67-

Implementation experience. http://www.ihtsdo.org/snomed-ct/snomed-present/implementation-experience.

Ingenerf, J., Reiner, J., & Seik, B. (2001). Standardized terminological services enabling semantic interoperability between distributed and heterogeneous systems. International Journal of Medical Informatics, 64(2), 223-240. http://www.ncbi.nlm.nih.gov/pubmed/11734388.

International Health Terminology Standards Development Organisation. Technical Implementation Guide. http://www.ihtsdo.org/fileadmin/user_upload/doc/tig/index.html.

International Health Terminology Standards Development Organisation. User Guide. http://www.ihtsdo.org/fileadmin/user_upload/doc/ug/index.html.

James, A. G., & Spackman, K. A. (2008). Representation of disorders of the newborn infant by SNOMED CT. Studies in Health Technology and Informatics, 136, 833-8. http://www.ncbi.nlm.nih.gov/pubmed/18487835.

Jao, C., Hier, D., & Galanter, W. (2008, November). Automating the maintenance of problem list documentation using a clinical decision support system. In AMIA... Annual Symposium proceedings/AMIA Symposium. AMIA Symposium (p. 989). http://www.ncbi.nlm.nih.gov/pubmed/18998927.

Johansson, I. (2006). Bioinformatics and biological reality. Journal of Biomedical Informatics, 39(3), 274-287. http://www.ncbi.nlm.nih.gov/pubmed/16198638.

Kim, H. Y., & Park, H. (2012). Development and evaluation of data entry templates based on the entity-attribute-value model for clinical decision support of pressure ulcer wound management. International journal of medical informatics, 81(7), 485-492. http://www.ncbi.nlm.nih.gov/pubmed/22079242.

Koopman, B., Bruza, P., Sitbon, L., & Lawley, M. (2012). Towards semantic search and inference in electronic medical records: An approach using concept-based information retrieval. The Australasian medical journal, 5(9), 482-8. http://www.ncbi.nlm.nih.gov/pubmed/23115582.

Kotiadis, K., & Robinson, S. (2008). Conceptual modelling: Knowledge acquisition and model abstraction. In Proceedings of the 40th Conference on Winter Simulation (pp. 951-958). Winter Simulation Conference.

Kvale, S., & Brinkmann, S. (2009). Interviews: Learning the craft of qualitative research interviewing. Sage Publications, Incorporated.

Lee, D. H., Lau, F. Y., & Quan, H. (2010). A method for encoding clinical datasets with SNOMED CT. BMC medical informatics and decision making, 10(1), 53. http://www.ncbi.nlm.nih.gov/pubmed/20849611.

Lee, D., Cornet, R., & Lau, F. (2011). Implications of SNOMED CT versioning. International Journal of Medical Informatics, 80(6), 442-453. http://www.ncbi.nlm.nih.gov/pubmed/21450517.

Lee, D., Cornet, R., Lau, F., & De Keizer, N. (2013). A survey of SNOMED CT implementations. Journal of Biomedical Informatics, 46(1), 87-96. http://www.ncbi.nlm.nih.gov/pubmed/23041717.

Lee, D., de Keizer, N., Lau, F., & Cornet, R. (2013). Literature review of SNOMED CT use. Journal of the American Medical Informatics Association. http://www.ncbi.nlm.nih.gov/pubmed/23828173.

Lee, N. J., & Bakken, S. (2007). Development of a prototype personal digital assistant-decision support system for the management of adult obesity. International Journal of Medical Informatics, 76, S281-S292. http://www.ncbi.nlm.nih.gov/pubmed/17606400.

Lee, S., Tsirbas, A., Goldberg, R. A., & McCann, J. D. (2006). Standardized terminology for aesthetic ophthalmic plastic surgery. Ophthalmic Plastic & Reconstructive Surgery, 22(5), 371-374. http://www.ncbi.nlm.nih.gov/pubmed/16985422.

Liaw, S. T., Chen, H. Y., Maneze, D., Taggart, J., Dennis, S., Vagholkar, S., & Bunker, J. (2012). Health reform: Is routinely collected electronic information fit for purpose? Emergency Medicine Australasia, 24(1), 57-63. http://www.ncbi.nlm.nih.gov/pubmed/22313561.

Lieberman, M. I., & Ricciardi, T. N. (2003). The Use of SNOMED© CT Simplifies Querying of a Clinical Data Warehouse. In AMIA Annual Symposium Proceedings (Vol. 2003, p. 910). American Medical Informatics Association. http://www.ncbi.nlm.nih.gov/pubmed/14728416.

Liu, J., Lane, K., Lo, E., Lam, M., Truong, T., & Veillette, C. (2010). Addressing SNOMED CT implementation challenges through multi-disciplinary collaboration. Studies in Health Technology and Informatics, 160(Pt 2), 981-5. http://www.ncbi.nlm.nih.gov/pubmed/20841830.

Liu, S., Ni, Y., Mei, J., Li, H., Xie, G., Hu, G., ... & Pan, Y. (2009). ismart: Ontology-based semantic query of cda documents. In AMIA Annual Symposium Proceedings (Vol. 2009, p. 375). American Medical Informatics Association.

Long, W. (2005). Extracting diagnoses from discharge summaries. In AMIA annual symposium proceedings (Vol. 2005, p. 470-4). American Medical Informatics Association. http://www.ncbi.nlm.nih.gov/pubmed/16779084.

Luqi, L., & Steigerwald, R. (1992,). Rapid software prototyping. In System Sciences, 1992. Proceedings of the Twenty-Fifth Hawaii International Conference on (Vol. 2, pp. 470-479). IEEE.

Lusky, K. (2005). Pilot points way to speedier cancer surveillance. CAP today/College of American Pathologists, 19(2), 5-6, 8. http://www.ncbi.nlm.nih.gov/pubmed/15787106.

Lussier, Y. A., & Bourque, M. (1997). Comparing SNOMED and ICPC retrieval accuracies using relational database models. In Proceedings of the AMIA Annual Fall Symposium (p. 514). American Medical Informatics Association.

Manitoba e-Health. EMR Adoption Program: Program Guidelines and Detailed Requirements V2.1. http://www.manitoba-ehealth.ca/commPhysicians/files/Prog_Guidelines.pdf.

Matheny, M. E., FitzHenry, F., Speroff, T., Green, J. K., Griffith, M. L., Vasilevskis, E. E., ... & Brown, S. H. (2012). Detection of infectious symptoms from VA emergency department and primary care clinical documentation. International journal of medical informatics, 81(3), 143-156. http://www.ncbi.nlm.nih.gov/pubmed/22244191.

McConnell, S. (1996). Rapid Development: Taming Wild Software Schedules. Redmond: Microsoft Press.

Meaningful Use Quality Performance Measures Benefit from New SNOMED CT "Public Good" Use Policy. http://www.ihtsdo.org/news/article/article/meaningful-use-quality-performance-measures-benefit-from-new-snomed-ct-public-good-use-policy.

Members of IHTSDO. http://www.ihtsdo.org/members.

Min, H., Perl, Y., Chen, Y., Halper, M., Geller, J., & Wang, Y. (2006). Auditing as part of the terminology design life cycle. Journal of the American Medical Informatics Association, 13(6), 676-690. http://www.ncbi.nlm.nih.gov/pubmed/16929044.

Nadkarni, P. M., & Darer, J. A. (2010). Migrating existing clinical content from ICD-9 to SNOMED. Journal of the American Medical Informatics Association, 17(5), 602-607. http://www.ncbi.nlm.nih.gov/pubmed/20819871.

Navas, H., Lopez, O. A., Gambarte, L., Elías, L. G., Wasserman, S., Orrego, N., ... & de Quirós, F. G. (2010). Implementing rules to improve the quality of concept post-coordination with SNOMED CT. Studies in health technology and informatics, 160(Pt 2), 1045. http://www.ncbi.nlm.nih.gov/pubmed/20841843.

Nguyen, A., Moore, J., Zuccon, G., Lawley, M., & Colquist, S. (2012). Classification of pathology reports for cancer registry notifications. Studies in health technology and informatics, 178, 150-6. http://www.ncbi.nlm.nih.gov/pubmed/22797034.

Olivé, A. (2000). An introduction to conceptual modelling of information systems. Advanced database technology and design. Artech House, 25-57.

Osornio, A. L., Luna, D., Gambarte, M. L., Gomez, A., Reynoso, G., & de Quiros, F. G. (2007). Creation of a local interface terminology to SNOMED CT. Studies in health technology and informatics, 129(Pt 1), 765. http://www.ncbi.nlm.nih.gov/pubmed/17911820.

Øvretveit, J., Scott, T., Rundall, T. G., Shortell, S. M., & Brommels, M. (2007). Improving quality through effective implementation of information technology in healthcare. International Journal for Quality in Health Care, 19(5), 259-266. http://www.ncbi.nlm.nih.gov/pubmed/17717038.

Park, H. A., Lundberg, C. B., Coenen, A., & Konicek, D. J. (2009). Evaluation of the content coverage of SNOMED-CT to represent ICNP Version 1 catalogues. Studies in Health Technology and Informatics, 146, 303-7. http://www.ncbi.nlm.nih.gov/pubmed/19592854.

Pathak, J., Wang, J., Kashyap, S., Basford, M., Li, R., Masys, D. R., & Chute, C. G. (2011). Mapping clinical phenotype data elements to standardized metadata repositories and controlled terminologies: the eMERGE Network experience. Journal of the American Medical Informatics Association, 18(4), 376-386. http://www.ncbi.nlm.nih.gov/pubmed/21597104.

Patrick, J. D., Ryan, A., & Herkes, R. (2008). Introduction of Enhancement Technologies into the Intensive Care Service, Royal Prince Alfred Hospital, Sydney. Health Information Management Journal, 37(1), 40-5. http://www.ncbi.nlm.nih.gov/pubmed/18245864.

Patton, M. Q. (1990). Qualitative evaluation and research methods. Sage Publications, Incorporated.

Penz, J. F., Brown, S. H., Carter, J. S., Elkin, P. L., Nguyen, V. N., Sims, S. A., & Lincoln, M. J. (2004). Evaluation of SNOMED coverage of Veterans Health Administration terms. Medinfo, 11(Pt 1), 540-4. http://www.ncbi.nlm.nih.gov/pubmed/15360871.

Pole, D. (2010). Electronic Patient Records in Sri Lankan Hospitals. Sri Lanka Journal of Bio-Medical Informatics, 1(1):43-45.

Public Health Agency of Canada. Chronic Diseases. http://www.phac-aspc.gc.ca/cd-mc/index-eng.php.

Public Health Agency of Canada. Diabetes in Canada: Facts and Figures from a Public Health Perspective. 2011. http://www.phac-aspc.gc.ca/cd-mc/publications/diabetes-diabete/facts-figures-faits-chiffres-2011/pdf/facts-figures-faits-chiffres-eng.pdf.

Rajeev, D., Staes, C. J., Evans, S. R., Mottice, S., Rolfs, R., Samore, M. H., ... & Huff, S. M. (2010). In response to letter to the editor:'Concerning SNOMED-CT content for public health case reports'. Journal of the American Medical Informatics Association, 17(5), 613-614. http://www.ncbi.nlm.nih.gov/pubmed/20842802.

Randorff, H. A., & Rosenbeck, G. K. (2012). SNOMED CT Implementation. Mapping Guidelines Facilitating Reuse of Data. Methods of Information in Medicine, 51(6). http://www.ncbi.nlm.nih.gov/pubmed/23038162.

Rector, A. L., & Brandt, S. (2008). Why do it the hard way? The case for an expressive description logic for SNOMED. Journal of the American Medical Informatics Association, 15(6), 744-751. http://www.ncbi.nlm.nih.gov/pubmed/18755993.

Rector, A. L., Brandt, S., & Schneider, T. (2011). Getting the foot out of the pelvis: modelling problems affecting use of SNOMED CT hierarchies in practical applications. Journal of the American Medical Informatics Association, 18(4), 432-440. http://www.ncbi.nlm.nih.gov/pubmed/21515545.

Results of the Survey to Gather the Use, Benefits and Tools of SNOMED CT. http://www.ihtsdo.org/fileadmin/user_upload/Docs_01/Publications/Implementation/Results_of_Survey_on_Use_of_SNOMED_V1.0._Binder.pdf. January 19, 2011.

Rice, P., & Ezzy, D. (1999). Qualitative research methods: A health focus. Melbourne: Oxford University Press.

Richesson, R. L., Andrews, J. E., & Krischer, J. P. (2006). Use of SNOMED CT to represent clinical research data: a semantic characterization of data items on case report forms in vasculitis research. Journal of the American Medical Informatics Association, 13(5), 536-546. http://www.ncbi.nlm.nih.gov/pubmed/16799121.

Richesson, R., Young, K., Guillette, H., Tuttle, M., Abbondondolo, M., & Krischer, J. (2006). Standard terminology on demand: facilitating distributed and real-time use of SNOMED CT during the clinical research process. In AMIA Annual Symposium Proceedings (Vol. 2006, p. 1076). American Medical Informatics Association. http://www.ncbi.nlm.nih.gov/pubmed/17238695.

Robinson, S. (2004). Simulation: The Practice of Model Development and Use. Chichester: John Wiley & Sons.

Robinson, S. (2010). Conceptual Modelling: Who Needs It? SCS M&S Magazine.

Robinson, T. J., DuVall, S. L., & Wiggins III, R. H. (2011). Creation and storage of standards-based pre-scanning patient questionnaires in PACS as DICOM objects. Journal of Digital Imaging, 24(5), 823-827. http://www.ncbi.nlm.nih.gov/pubmed/20976611.

Rosenbloom, S. T., Miller, R. A., Johnson, K. B., Elkin, P. L., & Brown, S. H. (2006). Interface terminologies facilitating direct entry of clinical data into electronic health record systems. Journal of the American Medical Informatics Association, 13(3), 277-288. http://www.ncbi.nlm.nih.gov/pubmed/16501181.

Saldana, Johnny. (2009). The Coding Manual for Qualitative Researchers. Thousand Oaks, California: Sage Publications, Incorporated.

Sampalli, T., Shepherd, M., Duffy, J., & Fox, R. (2010). An evaluation of SNOMED CT® in the domain of complex chronic conditions. International journal of integrated care, 10. http://www.ncbi.nlm.nih.gov/pubmed/20422022.

Seidman, I. (1998). Interviewing as qualitative research: A guide for researchers in education and the social sciences. New York, NY: Teachers College Press.

Sevenster, M., van Ommering, R., & Qian, Y. (2012). Algorithmic and user study of an autocompletion algorithm on a large medical vocabulary. Journal of Biomedical Informatics, 45(1), 107-119. http://www.ncbi.nlm.nih.gov/pubmed/22019376.

Sherman, S., Shats, O., Fleissner, E., Bascom, G., Yiee, K., Copur, M., ... & Cowan, K. (2011). Multicenter breast cancer collaborative registry. Cancer informatics, 10, 217-26. http://www.ncbi.nlm.nih.gov/pubmed/21918596.

Sherman, S., Shats, O., Ketcham, M. A., Anderson, M. A., Whitcomb, D. C., Lynch, H. T., ... & Brand, R. E. (2011). PCCR: Pancreatic cancer collaborative registry. Cancer informatics, 10, 83-91. http://www.ncbi.nlm.nih.gov/pubmed/21552494.

SNOMED CT Australia. http://www.nehta.gov.au/connecting-australia/terminology-and-information/clinical-terminology/snomed-ct-au.

SNOMED CT Benefits. http://www.ihtsdo.org/snomed-ct/whysnomedct/benefits.

SNOMED CT Canada. https://sl.infoway-inforoute.ca/content/disppage.asp?cw_page=snomedct_e.

SNOMED CT in use. https://www.infoway-inforoute.ca/standards-collaborative/snomed-ctr/snomed-ct-in-use.

SNOMED CT Namespace Registry. http://www.ihtsdo.org/fileadmin/user_upload/Docs_01/SNOMED_CT/Namespaces/SNOMED_CT_Namespace_Registry_-_OFFICIAL_20130611.pdf.

SNOMED CT Value Proposition. http://www.ihtsdo.org/snomed-ct/whysnomedct/snomedfeatures.

SNOMED CT. http://www.nlm.nih.gov/research/umls/Snomed/snomed_main.html.

SNOMED CT: A standard clinical terminology is essential for the interoperability of electronic health records across care settings. http://www.connectingforhealth.nhs.uk/systemsandservices/data/uktc/snomed.

SNOMED Endorsement. http://ww.ithealthboard.health.nz/sites/all/files/SNOMED CT Endorsement.doc. Last accessed: April 11, 2012.

Spackman, K. A. (2001). Normal forms for description logic expressions of clinical concepts in SNOMED RT. In Proceedings of the AMIA Symposium (p. 627-31). American Medical Informatics Association.

Spackman, K. A. (2005). Rates of Change in a Large Clinical Terminology: Three Years Experience with SNOMED Clinical Terms. In AMIA Annual Symposium Proceedings (Vol. 2005, p. 714-8). American Medical Informatics Association. http://www.ncbi.nlm.nih.gov/pubmed/16779133.

Takeda, H., Veerkamp, P., & Yoshikawa, H. (1990). Modelling design process. AI magazine, 11(4), 37.

Taylor, S. J., & Bogdan, R. (1984). Introduction to qualitative research methods: The search for meanings. New York: John Wiley & Sons.

Tripp, S. D., & Bichelmeyer, B. (1990). Rapid prototyping: An alternative instructional design strategy. Educational Technology Research and Development, 38(1), 31-44.

Tvede, I., Bredegaard, K., & Andersen, J. S. (2010). Quality improvements based on detailed and precise terminology. Studies in Health Technology and Informatics, 155, 71-7. http://www.ncbi.nlm.nih.gov/pubmed/20543312.

van Berkum, M. M. (2003). SNOMED CT® Encoded Cancer Protocols. In Amia Annual Symposium Proceedings (Vol. 2003, p. 1039). American Medical Informatics Association. http://www.ncbi.nlm.nih.gov/pubmed/14728542.

Wade, G., & Rosenbloom, S. T. (2008). Experiences mapping a legacy interface terminology to SNOMED CT. BMC medical informatics and decision making, 8(Suppl 1), S3. http://www.ncbi.nlm.nih.gov/pubmed/19007440.

Wade, G., & Rosenbloom, S. T. (2009). The impact of SNOMED CT revisions on a mapped interface terminology: terminology development and implementation issues. Journal of biomedical informatics, 42(3), 490-493. http://www.ncbi.nlm.nih.gov/pubmed/19285570.

Wang Y, Patrick J, Miller G, O'Hallaran J. A computational linguistics motivated mapping of ICPC-2 PLUS to SNOMED CT. BMC Med Inform Decis Mak. 2008 Oct 27;8 Suppl 1:S5. http://www.ncbi.nlm.nih.gov/pubmed/19007442.

Wang, W., & Brooks, R. J. (2007). Improving the understanding of conceptual modelling. Journal of Simulation, 1(3), 153-158.

Wang, Y., Halper, M., Min, H., Perl, Y., Chen, Y., & Spackman, K. A. (2007). Structural methodologies for auditing SNOMED. Journal of Biomedical Informatics, 40(5), 561-581. http://www.ncbi.nlm.nih.gov/pubmed/17276736.

Wang, Y., Patrick, J., Miller, G., & O'Hallaran, J. (2008). A computational linguistics motivated mapping of ICPC-2 PLUS to SNOMED CT. BMC medical informatics and decision making, 8(Suppl 1), S5. http://www.ncbi.nlm.nih.gov/pubmed/19007442.

What are EMR content standards for PHC? http://www.cihi.ca/CIHI-ext-portal/internet/EN/document/types+of+care/primary+health/faq_ph_emr_content_stdrds.

Windle, J., Van-Milligan, G., Duffy, S., McClay, J., & Campbell, J. (2003). Web-based physician order entry: an open source solution with broad physician involvement. In AMIA Annual Symposium Proceedings (Vol. 2003, p. 724-7). American Medical Informatics Association. http://www.ncbi.nlm.nih.gov/pubmed/14728268.

Wright, A., Pang, J., Feblowitz, J. C., Maloney, F. L., Wilcox, A. R., Ramelson, H. Z., ... & Bates, D. W. (2011). A method and knowledge base for automated inference of patient problems from structured data in an electronic medical record. Journal of the American Medical Informatics Association, 18(6), 859-867. http://www.ncbi.nlm.nih.gov/pubmed/21613643

Zaninelli, M., Campagnoli, A., Reyes, M., & Rojas, V. (2012). The O3-Vet project: Integration of a standard nomenclature of clinical terms in a veterinary electronic medical record for veterinary hospitals. Computer Methods and Programs in Biomedicine, 108(2), 760-772. http://www.ncbi.nlm.nih.gov/pubmed/22595264.

Zetterberg, C., Ahlzén, K., Ericsson, E., & Kron, B. (2012). An Example of a Multi-Professional Process-Oriented Structured Documentation Bound to SNOMED CT. Studies in Health Technology and Informatics, 180, 1215-7. http://www.ncbi.nlm.nih.gov/pubmed/22874405.

Zhu, X., Fan, J. W., Baorto, D. M., Weng, C., & Cimino, J. J. (2009). A review of auditing methods applied to the content of controlled biomedical terminologies. Journal of biomedical informatics, 42(3), 413-425. http://www.ncbi.nlm.nih.gov/pubmed/19285571.

12. APPENDICES

12.1 Appendix A: Ethics Approval

Human Research Ethics Board
Office of Research Services
Administrative Services Building
PO Box 1700 STN CSC
Victoria British Columbia V8W 2Y2 Canada
Tel 250-472-4545, Fax 250-721-8960
Email ethics@uvic.ca Web www.research.uvic.ca

University of Victoria

Certificate of Renewed Approval

PRINCIPAL INVESTIGATOR:	**Morgan Price**	**ETHICS PROTOCOL NUMBER**	**10-143**
UVic STATUS:	**Postdoctoral Fellow**	ORIGINAL APPROVAL DATE:	01-Apr-10
UVic DEPARTMENT:	**HEIS**	RENEWED ON:	27-Feb-13
SUPERVISOR:	**Francis Lau**	APPROVAL EXPIRY DATE:	31-Mar-14

PROJECT TITLE: **An Initial Data Analysis for an Electronic Medical Record (EMR) System**

RESEARCH TEAM MEMBERS: Francis Lau, Co-Investigator (UVic)
Dennis Lee, Research Assistant (UVic)

DECLARED PROJECT FUNDING: **CIHR (previous)**

CONDITIONS OF APPROVAL

This Certificate of Approval is valid for the above term provided there is no change in the protocol.

Modifications
To make any changes to the approved research procedures in your study, please submit a "Request for Modification" form. You must receive ethics approval before proceeding with your modified protocol.

Renewals
Your ethics approval must be current for the period during which you are recruiting participants or collecting data. To renew your protocol, please submit a "Request for Renewal" form before the expiry date on your certificate. You will be sent an emailed reminder prompting you to renew your protocol about six weeks before your expiry date.

Project Closures
When you have completed all data collection activities and will have no further contact with participants, please notify the Human Research Ethics Board by submitting a "Notice of Project Completion" form.

Certification

This certifies that the UVic Human Research Ethics Board has examined this research protocol and concluded that, in all respects, the proposed research meets the appropriate standards of ethics as outlined by the University of Victoria Research Regulations Involving Human Participants.

Dr. Rachael Scarth
Associate Vice-President, Research

10-143 Price, Morgan

Certificate Issued On: 27-Feb-13

Figure 12-1. Ethics Approval for Protocol 10-143 - "An Initial Data Analysis for an Electronic Medical Record (EMR) System."

Human Research Ethics Board
Office of Research Services
Administrative Services Building
PO Box 1700 STN CSC
Victoria British Columbia V8W 2Y2 Canada
Tel 250-472-4545, Fax 250-721-8960
Email ethics@uvic.ca Web www.research.uvic.ca

University of Victoria

Certificate of Approval

PRINCIPAL INVESTIGATOR	**Dennis Lee**	**ETHICS PROTOCOL NUMBER**	**11-535**
UVic STATUS:	**Ph.D. Student**	ORIGINAL APPROVAL DATE:	16-Dec-11
UVic DEPARTMENT:	**HEIS**	APPROVED ON:	16-Dec-11
SUPERVISOR:	**Dr. Francis Lau**	APPROVAL EXPIRY DATE:	15-Dec-12

PROJECT TITLE: **SNOMED CT Implementation Survey**

RESEARCH TEAM MEMBERS: Ronald Cornet, Co-Investigator (University of Amsterdam)
Nicolette de Keizer, Co-Investigator (University of Amsterdam)

DECLARED PROJECT FUNDING: **None**

CONDITIONS OF APPROVAL

This Certificate of Approval is valid for the above term provided there is no change in the protocol.

Modifications
To make any changes to the approved research procedures in your study, please submit a "Request for Modification" form. You must receive ethics approval before proceeding with your modified protocol.

Renewals
Your ethics approval must be current for the period during which you are recruiting participants or collecting data. To renew your protocol, please submit a "Request for Renewal" form before the expiry date on your certificate. You will be sent an emailed reminder prompting you to renew your protocol about six weeks before your expiry date.

Project Closures
When you have completed all data collection activities and will have no further contact with participants, please notify the Human Research Ethics Board by submitting a "Notice of Project Completion" form.

Certification

This certifies that the UVic Human Research Ethics Board has examined this research protocol and concluded that, in all respects, the proposed research meets the appropriate standards of ethics as outlined by the University of Victoria Research Regulations Involving Human Participants.

Dr. Rachael Scarth
Associate Vice-President, Research

11-535 Lee, Dennis

Certificate Issued On: 16-Dec-11

Figure 12-2. Ethics Approval for Protocol 11-535 - "SNOMED CT Implementation Survey."

Human Research Ethics Board
Office of Research Services
Administrative Services Building
PO Box 1700 STN CSC
Victoria British Columbia V8W 2Y2 Canada
Tel 250-472-4545, Fax 250-721-8960
ethics@uvic.ca www.research.uvic.ca

University of Victoria

Certificate of Approval

PRINCIPAL INVESTIGATOR	**Dennis Lee**	**ETHICS PROTOCOL NUMBER**	**12-529**
UVic STATUS:	**Ph.D. Student**	ORIGINAL APPROVAL DATE:	18-Dec-12
UVic DEPARTMENT:	**HEIS**	APPROVED ON:	18-Dec-12
SUPERVISOR:	**Dr. Francis Lau**	APPROVAL EXPIRY DATE:	17-Dec-13

PROJECT TITLE: **SNOMED CT Implementation Expert Feedback Sessions**

RESEARCH TEAM MEMBERS: None

DECLARED PROJECT FUNDING: **None**

CONDITIONS OF APPROVAL

This Certificate of Approval is valid for the above term provided there is no change in the protocol.

Modifications
To make any changes to the approved research procedures in your study, please submit a "Request for Modification" form. You must receive ethics approval before proceeding with your modified protocol.

Renewals
Your ethics approval must be current for the period during which you are recruiting participants or collecting data. To renew your protocol, please submit a "Request for Renewal" form before the expiry date on your certificate. You will be sent an emailed reminder prompting you to renew your protocol about six weeks before your expiry date.

Project Closures
When you have completed all data collection activities and will have no further contact with participants, please notify the Human Research Ethics Board by submitting a "Notice of Project Completion" form.

Certification

This certifies that the UVic Human Research Ethics Board has examined this research protocol and concluded that, in all respects, the proposed research meets the appropriate standards of ethics as outlined by the University of Victoria Research Regulations Involving Human Participants.

Dr. Rachael Scarth
Associate Vice-President, Research

Certificate Issued On: 19-Dec-12

12-529 Lee, Dennis

Figure 12-3. Ethics approval for Protocol 12-529 "SNOMED CT Implementation Expert Feedback Sessions."

12.2 Appendix B: For Chapter Three

12.2.1 Classification of Papers

Table 12-1. Papaers used in the literature review.

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
1.	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 Nov; 65(11): 1106-13	CT	Evaluation	Retrieve/analyse patient data	Anaesthesia	United Kingdom	2010
2.	Benoit SR, McDonald LC, English R, Tokars JI. Automated surveillance of Clostridium difficile infections using BioSense. Infection control and hospital epidemiology : the official journal of the Society of Hospital Epidemiologists of America 2011 Jan; 32(1): 26-33	CT	Evaluation	Retrieve/analyse patient data	Laboratory	United States	2011
3.	Elkin PL, Froehling D, Wahner-Roedler D, Trusko B, Welsh G, Rosenbloom ST, Speroff T, Brown SH. The Health Archetype Language (HAL-42): interface considerations. International journal of medical informatics 2010 Apr; 79(4): e71-5	CT	Evaluation	Retrieve/analyse patient data	Other/multiple/unknown/not applicable	United States	2010
4.	Hussain F., Muller F., Husain E. Under-reporting of invasive malignant melanomas in North East of Scotland. British Journal of Dermatology 2010; 163: 67-	CT	Evaluation	Retrieve/analyse patient data	Cancer	United Kingdom	2010
5.	Koopman B, Bruza P, Sitbon L, Lawley M Towards semantic search and inference in electronic medical records: An approach using concept--based information retrieval. The Australasian medical journal 2012;5(9):482-8	CT	Evaluation	Retrieve/analyse patient data	Hospital	Australia	2012
6.	Liaw ST, Chen HY, Maneze D, Taggart J, Dennis S, Vagholkar S, Bunker J. Health reform: is routinely collected electronic information fit for purpose? Emergency medicine Australasia : EMA 2012 Feb;24(1):57-63	CT	Evaluation	Retrieve/analyse patient data	Emergency department/room	Australia	2012
7.	Matheny ME, Fitzhenry F, Speroff T, Green JK, Griffith ML, Vasilevskis EE, Fielstein EM, Elkin PL, Brown SH. Detection of infectious symptoms from VA emergency department and primary care clinical documentation. International journal of medical informatics 2012 Mar;81(3):143-56	CT	Evaluation	Retrieve/analyse patient data	Other/multiple/unknown/not applicable	United States	2012

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
8.	Nguyen A, Moore J, Zuccon G, Lawley M, Colquist S. Classification of pathology reports for cancer registry notifications. Studies in health technology and informatics 2012;178:150-6	CT	Evaluation	Retrieve/analyse patient data	Cancer	Australia	2012
9.	Alecu I, Bousquet C, Degoulet P, Jaulent M.C PharmARTS: terminology web services for drug safety data coding and retrieval. Studies in health technology and informatics 2007; 129(Pt 1): 699-704	CT	Implementation	Description of implementation of SNOMED	Drugs	France	2007
10.	Aller RD. Making a case for computers to 'talk the talk'. CAP today / College of American Pathologists 2004 Feb; 18(2): 56-8, 60, 62 passim	Multiple	Implementation	Description of implementation of SNOMED	Pathology	United States	2004
11.	Bakhshi-Raiez F, de Keizer NF, Cornet R, Dorrepaal M, Dongelmans D, Jaspers MW. A usability evaluation of a SNOMED CT based compositional interface terminology for intensive care. International journal of medical informatics 2011 Oct 24;	CT	Implementation	Description of implementation of SNOMED	Intensive care	Netherlands	2011
12.	Cao F, Sun X, Wang X, Li B, Li J, Pan Y Ontology-based knowledge management for personalized adverse drug events detection. Studies in health technology and informatics 2011; 169: 699-703	CT	Implementation	Description of implementation of SNOMED	Drugs	South Korea	2011
13.	Cole CL, Kanter AS, Cummens M, Vostinar S, Naeymi-Rad F. Using a terminology server and consumer search phrases to help patients find physicians with particular expertise. Studies in health technology and informatics 2004; 107(Pt 1): 492-6	CT	Implementation	Description of implementation of SNOMED	Other/multiple/unknown/not applicable	United States	2004
14.	Farfán Sedano FJ, Terrón Cuadrado M, García Rebolledo EM, Castellanos Clemente Y, Serrano Balazote P, Gómez Delgado A. Implementation of SNOMED CT to the medicines database of a general hospital. Studies in health technology and informatics 2009; 148: 123-30	CT	Implementation	Description of implementation of SNOMED	Drugs	Spain	2009
15.	Gambarte ML, Osornio AL, Martinez M, Reynoso G, Luna D, de Quiros FG A practical approach to advanced terminology services in health information systems. Studies in health technology and informatics 2007; 129(Pt 1): 621-5	CT	Implementation	Description of implementation of SNOMED	Other/multiple/unknown/not applicable	Argentina	2007

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
16.	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. International journal of medical informatics 2011 Nov 10;	CT	Implementation	Description of implementation of SNOMED	Wounds	South Korea	2011
17.	Lee NJ, Bakken S. Development of a prototype personal digital assistant-decision support system for the management of adult obesity. International journal of medical informatics 2007 Oct; 76 Suppl 2: S281-92	CT	Implementation	Description of implementation of SNOMED	Other/multiple/unknown/not applicable	United States	2007
18.	Liu J, Lane K, Lo E, Lam M, Truong T, Veillette C. Addressing SNOMED CT implementation challenges through multi-disciplinary collaboration. Studies in health technology and informatics 2010; 160(Pt 2): 981-5	CT	Implementation	Description of implementation of SNOMED	Other/multiple/unknown/not applicable	Canada	2010
19.	Lusky K. Pilot points way to speedier cancer surveillance. CAP today / College of American Pathologists 2005 Feb; 19(2): 5-6, 8	CT	Implementation	Description of implementation of SNOMED	Cancer	United States	2005
20.	Navas H, Osornio AL, Baum A, Gomez A, Luna D, de Quiros FG. Creation and evaluation of a terminology server for the interactive coding of discharge summaries. Studies in health technology and informatics 2007; 129(Pt 1): 650-4	CT	Implementation	Description of implementation of SNOMED	Hospital	Argentina	2007
21.	Richesson R, Syed A, Guillette H, Tuttle MS, Krischer J. A web-based SNOMED CT browser: distributed and real-time use of SNOMED CT during the clinical research process. Studies in health technology and informatics 2007; 129(Pt 1): 631-5	CT	Implementation	Description of implementation of SNOMED	Other/multiple/unknown/not applicable	United States	2007
22.	Richesson R, Young K, Guillette H, Tuttle M, Abbondondolo M, Krischer J. Standard terminology on demand: facilitating distributed and real-time use of SNOMED CT during the clinical research process. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2006; : 1076	CT	Implementation	Description of implementation of SNOMED	Other/multiple/unknown/not applicable	United States	2006
23.	Robinson J, de Lusignan S, Kostkova P, Madge B, Marsh A, Biniaris C. The Primary Care Electronic Library: RSS feeds using SNOMED-CT indexing for dynamic content delivery. Informatics in primary care 2006; 14(4): 247-52	CT	Implementation	Description of implementation of SNOMED	Cardiology	United Kingdom	2006

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
24.	Robinson TJ, DuVall SL, Wiggins RH. Creation and storage of standards-based pre-scanning patient questionnaires in PACS as DICOM objects. Journal of digital imaging : the official journal of the Society for Computer Applications in Radiology 2011 Oct; 24(5): 823-7	CT	Implementation	Description of implementation of SNOMED	Radiology	United States	2011
25.	Ryan A, Patrick J, Herkes R. Introduction of enhancement technologies into the intensive care service, Royal Prince Alfred Hospital, Sydney. The HIM journal 2008; 37(1): 40-5	CT	Implementation	Description of implementation of SNOMED	Intensive care	Australia	2008
26.	Sherman S, Shats O, Fleissner E, Bascom G, Yiee K, Copur M, Crow K, Rooney J, Mateen Z, Ketcham MA, Feng J, Sherman A, Gleason M, Kinarsky L, Silva-Lopez E, Edney J, Reed E, Berger A, Cowan K. Multicenter breast cancer collaborative registry. Cancer informatics 2011; 10: 217-26	CT	Implementation	Description of implementation of SNOMED	Cancer	United States	2011
27.	Sherman S, Shats O, Ketcham MA, Anderson MA, Whitcomb DC, Lynch HT, Ghiorzo P, Rubinstein WS, Sasson AR, Grizzle WE, Haynatzki G, Feng J, Sherman A, Kinarsky L, Brand RE. PCCR: Pancreatic Cancer Collaborative Registry. Cancer informatics 2011; 10: 83-91	CT	Implementation	Description of implementation of SNOMED	Cancer	United States	2011
28.	Sundvall E, Nyström M, Petersson H, Ahlfeldt H. Interactive visualization and navigation of complex terminology systems, exemplified by SNOMED CT. Studies in health technology and informatics 2006; 124: 851-6	CT	Implementation	Description of implementation of SNOMED	Other/multiple/unknown/not applicable	Sweden	2006
29.	Van Berkum MM. SNOMED CT encoded Cancer Protocols. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2003; : 1039	CT	Implementation	Description of implementation of SNOMED	Cancer	United States	2003
30.	Yang K, Evens MW, Trace DA. Using a Java Dynamic Tree to manage the terminology in a suite of medical applications. Methods of information in medicine 2008; 47(6): 499-504	CT	Implementation	Description of implementation of SNOMED	Other/multiple/unknown/not applicable	United States	2008
31.	Zaninelli M, Campagnoli A, Reyes M, Rojas V. The O3-Vet project: Integration of a standard nomenclature of clinical terms in a veterinary electronic medical record for veterinary hospitals. Computer methods and programs in biomedicine 2012 Nov;108(2):760-72	CT	Implementation	Description of implementation of SNOMED	Veterinary	Italy	2012

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
32.	Zetterberg C, Ahlzén K, Ericsson E, Kron B. An Example of a Multi-Professional Process-Oriented Structured Documentation Bound to SNOMED CT. Studies in health technology and informatics 2012;180:1215-7	CT	Implementation	Description of implementation of SNOMED	Gynaecology	Sweden	2012
33.	Beach K., Le H.V., Powell G., Pattishall E., Ryan P., Mera R. Performance of a semi-automated process for estimation of risk in a claims database. Pharmacoepidemiology and Drug Safety (PDS) 2009; 18: S164-	CT	Implementation	Used to classify or code in a study	Drugs	United States	2009
34.	Brown SH, Elkin PL, Rosenbloom ST, Fielstein E, Speroff T. eQuality for all: Extending automated quality measurement of free text clinical narratives. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2008; : 71-5	CT	Implementation	Used to classify or code in a study	Hospital	United States	2008
35.	Elkin PL, Brown SH, Bauer BA, Husser CS, Carruth W, Bergstrom LR, Wahner-Roedler DL. A controlled trial of automated classification of negation from clinical notes. BMC medical informatics and decision making 2005; 5: 13	CT	Implementation	Used to classify or code in a study	Internal medicine	United States	2005
36.	Elkin PL, Froehling D, Wahner-Roedler D, Trusko B, Welsh G, Ma H, Asatryan AX, Tokars JI, Rosenbloom ST, Brown SH. NLP-based identification of pneumonia cases from free-text radiological reports. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2008; : 172-6	CT	Implementation	Used to classify or code in a study	Radiology	United States	2008
37.	Long W. Extracting diagnoses from discharge summaries. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2005; : 470-4	CT	Implementation	Used to classify or code in a study	Other/multiple/unknown/not applicable	United States	2005
38.	Lowe HJ, Huang Y, Regula DP. Using a statistical natural language Parser augmented with the UMLS specialist lexicon to assign SNOMED CT codes to anatomic sites and pathologic diagnoses in full text pathology reports. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2009; 2009: 386-90	CT	Implementation	Used to classify or code in a study	Pathology	United States	2009
39.	McBride SJ, Lawley MJ, Leroux H, Gibson S. Using Australian medicines terminology (AMT) and SNOMED CT-AU to better support clinical research Studies in health technology and informatics 2012;178:144-9	CT	Implementation	Used to classify or code in a study	Drugs	Australia	2012

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
40.	McCoy AB, Wright A, Laxmisan A, Singh H, Sittig DF. A prototype knowledge base and SMART app to facilitate organization of patient medications by clinical problems. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2011;2011:888-94	CT	Implementation	Used to classify or code in a study	Drugs	United States	2011
41.	Nguyen A, Moore J, Lawley M, Hansen D, Colquist S. Automatic extraction of cancer characteristics from free-text pathology reports for cancer notifications. Studies in health technology and informatics 2011; 168: 117-24	CT	Implementation	Used to classify or code in a study	Cancer	Australia	2011
42.	Nguyen A, Moore J, Zuccon G, Colquist S. Towards a smart cancer registry Asia-Pac. J. Clin. Oncol. 2012; 8: 327	CT	Implementation	Used to classify or code in a study	Cancer	Australia	2012
43.	Nguyen A.N., Lawley M.J., Hansen D.P., Bowman R.V., Clarke B.E., Duhig E.E., Colquist S. Surgical pathology cancer case checklist reporting from free text reports. Asia-Pacific Journal of Clinical Oncology 2009; 5: A218-A219	CT	Implementation	Used to classify or code in a study	Cancer	Australia	2009
44.	Nguyen AN, Lawley MJ, Hansen DP, Bowman RV, Clarke BE, Duhig EE, Colquist S. Symbolic rule-based classification of lung cancer stages from free-text pathology reports. Journal of the American Medical Informatics Association : JAMIA 2010 Jul-Aug; 17(4): 440-5	CT	Implementation	Used to classify or code in a study	Pathology	Australia	2010
45.	Patrick J, Nguyen D, Wang T, Paoloni R. Computational recognition of SNOMED CT codes from ED case notes. Studies in health technology and informatics 2012;178:175-9	CT	Implementation	Used to classify or code in a study	Emergency department/room	Australia	2012
46.	Ruch P, Gobeill J, Lovis C, Geissbühler A. Automatic medical encoding with SNOMED categories. BMC medical informatics and decision making 2008; 8 Suppl 1: S6	CT	Implementation	Used to classify or code in a study	Other/multiple/unknown/not applicable	Switzerland	2008
47.	Shah NH, Jonquet C, Chiang AP, Butte AJ, Chen R, Musen MA Ontology-driven indexing of public datasets for translational bioinformatics. BMC bioinformatics 2009; 10 Suppl 2: S1	CT	Implementation	Used to classify or code in a study	Genetics	United States	2009
48.	Tvede I, Bredegaard K, Andersen JS. Quality improvements based on detailed and precise terminology. Studies in health technology and informatics 2010; 155: 71-7	CT	Implementation	Used to classify or code in a study	Intensive care	Denmark	2010

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
49.	Wagholikar AS, Lawley MJ, Hansen DP, Chu K. Identifying symptom groups from Emergency Department presenting complaint free text using SNOMED CT. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2011;2011:1446-53	CT	Implementation	Used to classify or code in a study	Emergency department/room	Australia	2011
50.	Wahner-Roedler DL, Welsh GA, Trusko BE, Froehling DA, Froehling DA, Temesgen Z, Elkin PL Using natural language processing for identification of pneumonia cases from clinical records of patients with serologically proven influenza. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2008; : 1165	CT	Implementation	Used to classify or code in a study	Pathology	United States	2008
51.	Welsh G, Wahner-Roedler D, Froehling DA, Trusko B, Elkin P. Whole record surveillance is superior to chief complaint surveillance for predicting influenza. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2008; : 1173	CT	Implementation	Used to classify or code in a study	Pathology	United States	2008
52.	Windle J, Van-Milligan G, Duffy S, McClay J, Campbell J. Web-based physician order entry: an open source solution with broad physician involvement. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2003; : 724-7	CT	Implementation	Used to classify or code in a study	Other/multiple/unknown/not applicable	United States	2003
53.	DHHS to promote national paperless health system. Public health reports (Washington, D.C. : 1974) 2003 Nov-Dec; 118(6): 563	CT	Not applicable	As example	Other/multiple/unknown/not applicable	United States	2003
54.	Front matter. Studies in health technology and informatics 2012;179:i-xvii	CT	Not applicable	As example	Other/multiple/unknown/not applicable	Multiple countries	2012
55.	On not reinventing the wheel Nat. Genet. 2012; 44(3): 233	CT	Not applicable	As example	Genetics	United States	2012
56.	Bonney W. Is it appropriate, or ethical, to use health data collected for the purpose of direct patient care to develop computerized predictive decision support tools? Studies in health technology and informatics 2009; 143: 115-21	CT	Not applicable	As example	Other/multiple/unknown/not applicable	Canada	2009
57.	Brown PJ, Warmington V, Laurence M, Prevost AT. A methodology for the functional comparison of coding schemes in primary care. Informatics in primary care 2003; 11(3): 145-8	CT	Not applicable	As example	Primary care	United Kingdom	2003

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
58.	Case JT. Are we really communicating? Standard terminology versus terminology standards in veterinary clinical pathology. Veterinary clinical pathology / American Society for Veterinary Clinical Pathology 2005; 34(1): 5-6	CT	Not applicable	As example	Veterinary	United States	2005
59.	Ceusters W, Smith B, Goldberg L. A terminological and ontological analysis of the NCI Thesaurus. Methods of information in medicine 2005; 44(4): 498-507	CT	Not applicable	As example	Cancer	Multiple countries	2005
60.	Chen ES, Melton GB, Sarkar IN. Translating standards into practice: Experiences and lessons learned in biomedicine and health care J. Biomed. Informatics 2012; 45(4): 609-612	CT	Not applicable	As example	Other/multiple/unknown/not applicable	United States	2012
61.	Chute CG, Huff SM, Ferguson JA, Walker JM, Halamka JD. There are important reasons for delaying implementation of the new ICD-10 coding system. Health affairs (Project Hope) 2012 Apr;31(4):836-42	CT	Not applicable	As example	Other/multiple/unknown/not applicable	United States	2012
62.	Creek J. A standard terminology for occupational therapy Br. J. Occup. Ther. 2006; 69(5): 202-208	CT	Not applicable	As example	Other/multiple/unknown/not applicable	United Kingdom	2006
63.	Currie LM. Informatics and quality improvement/assurance. Nursing leadership (Toronto, Ont.) 2011 Apr; 24(1): 19-22	CT	Not applicable	As example	Nursing	Canada	2011
64.	de Lusignan S, Teasdale S. Achieving benefit for patients in primary care informatics: The report of a international consensus workshop at Medinfo 2007 Informatics Prim. Care 2007; 15(4): 255-261	CT	Not applicable	As example	Primary care	United Kingdom	2007
65.	Demir OM, Appleby R, Wang Y, Logan RPH. What factors might contribute to borderline faecal calprotectin levels? J. Crohn's Colitis 2012; 6: S56	CT	Not applicable	As example	Other/multiple/unknown/not applicable	United Kingdom	2012
66.	Dougherty M, Mitchell S. Getting better data from the MDS. Improving diagnostic data reporting in long-term care facilities. Journal of AHIMA / American Health Information Management Association 2004 Nov-Dec; 75(10): 28-33; quiz 35-6	CT	Not applicable	As example	Nursing	United States	2004
67.	Foulis P.R., Haley J.A. Standardization of bioinformatics. Annals of Clinical and Laboratory Science 2009; 39: 210-	CT	Not applicable	As example	Other/multiple/unknown/not applicable	United States	2009

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
68.	Franklin RC, Béland MJ, Krogmann ON. Mapping and coding of nomenclatures for paediatric and congenital heart disease. Cardiology in the young 2006 Apr; 16(2): 105-6	CT	Not applicable	As example	Cardiology	Unknown	2006
69.	Franklin RC, Jacobs JP, Krogmann ON, Béland MJ, Aiello VD, Colan SD, Elliott MJ, William Gaynor J, Kurosawa H, Maruszewski B, Stellin G, Tchervenkov CI, Walters Iii HL, Weinberg P, Anderson RH. Nomenclature for congenital and paediatric cardiac disease: historical perspectives and The International Pediatric and Congenital Cardiac Code. Cardiology in the young 2008 Dec; 18 Suppl 2: 70-80	CT	Not applicable	As example	Cardiology	United Kingdom	2008
70.	García Rojo M. State of the art and trends for digital pathology. Studies in health technology and informatics 2012;179:15-28	CT	Not applicable	As example	Pathology	Spain	2012
71.	Giannangelo K, Fenton S. EHR's effect on the revenue cycle management Coding function. Journal of healthcare information management : JHIM 2008; 22(1): 26-30	CT	Not applicable	As example	Other/multiple/unknown/not applicable	United States	2008
72.	Griffon N, Savoye-Collet C, Massari P, Daniel C, Darmoni SJ. An interface terminology for medical imaging ordering purposes. AMIA Annu Symp Proc 2012;2012:1237-43	CT	Not applicable	As example	Radiology	France	2012
73.	Halper M, Morrey CP, Chen Y, Elhanan G, Hripcsak G, Perl Y. Auditing hierarchical cycles to locate other inconsistencies in the UMLS. AMIA Annu Symp Proc 2011;2011:529-36	CT	Not applicable	As example	Other/multiple/unknown/not applicable	United States	2011
74.	Hammond WE. eHealth interoperability. Studies in health technology and informatics 2008; 134: 245-53	CT	Not applicable	As example	Other/multiple/unknown/not applicable	United States	2008
75.	Haroske G. Structured reports in tumor pathology-how and why. Virchows Archiv 2010; 457: 140-	CT	Not applicable	As example	Pathology	Germany	2010
76.	Haux R, Knaup P, Leiner F. On educating about medical data management: The other side of the electronic health record Methods Inf. Med. 2007; 46(1): 74-79	CT	Not applicable	As example	Other/multiple/unknown/not applicable	Germany	2007
77.	Ingenerf J [Terminologies or classifications. What does the future hold?]. Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz 2007 Aug; 50(8): 1070-83	CT	Not applicable	As example	Veterinary	Germany	2007

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
78.	Kamal J, Liu J, Ostrander M, Santangelo J, Dyta R, Rogers P, Mekhjian HS. Information warehouse - a comprehensive informatics platform for business, clinical, and research applications. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2010; 2010: 452-6	CT	Not applicable	As example	Other/multiple/unknown/not applicable	United States	2010
79.	Kapoor M. For the NHS Care Records Service the future's bright, it's SNOMED CT Pharm. J. 2008; 280(7500): 534	CT	Not applicable	As example	Other/multiple/unknown/not applicable	United Kingdom	2008
80.	Kloss L. The promise of ICD-10-CM. Health management technology 2005 Jul; 26(7): 48, 47	CT	Not applicable	As example	Other/multiple/unknown/not applicable	United States	2005
81.	Lependu P, Musen MA, Shah NH. Enabling enrichment analysis with the Human Disease Ontology. Journal of biomedical informatics 2011 Apr 29;	CT	Not applicable	As example	Genetics	United States	2011
82.	Lindblom A, Robinson PN. Bioinformatics for human genetics: promises and challenges. Human mutation 2011 May; 32(5): 495-500	CT	Not applicable	As example	Genetics	Sweden	2011
83.	Lloyd D, Tang P, Clark J, Hesketh J, Paget A, Baker B. Capturing nursing information in the United Kingdom through use of a conceptual model. Studies in health technology and informatics 2009; 146: 683-7	CT	Not applicable	As example	Nursing	United Kingdom	2009
84.	Long W. Lessons extracting diseases from discharge summaries. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2007; : 478-82	CT	Not applicable	As example	Other/multiple/unknown/not applicable	United States	2007
85.	Lowe HJ, Ferris TA, Hernandez PM, Weber SC. STRIDE--An integrated standards-based translational research informatics platform. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2009; 2009: 391-5	CT	Not applicable	As example	Other/multiple/unknown/not applicable	United States	2009
86.	Madden A. Health informatics and the importance of coding Anaesth. Intensive Care Med. 2010; 11(12): 500-501	CT	Not applicable	As example	Primary care	United Kingdom	2010
87.	Madden AP, Barham C. Clinical systems Anaesth. Intensive Care Med. 2007; 8(12): 518-519	CT	Not applicable	As example	Other/multiple/unknown/not applicable	United Kingdom	2007
88.	Marwede D, Fielding M. The epistemological-ontological divide in clinical radiology. Studies in health technology and informatics 2005; 116: 749-54	CT	Not applicable	As example	Radiology	Germany	2005

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
89.	Menárguez-Tortosa M, Fernández-Breis JT. Validation of the openEHR archetype library by using OWL reasoning. Studies in health technology and informatics 2011;169:789-93	CT	Not applicable	As example	Other/multiple/unknown/not applicable	Spain	2011
90.	Morrissey J .A defining moment. Database to help establish universally accepted clinical terminology, easing development of a standardized medical record. Modern healthcare 2003 Jul 28; 33(30): 30, 32	CT	Not applicable	As example	Other/multiple/unknown/not applicable	United States	2003
91.	Natale DA, Arighi CN, Barker WC, Blake J, Chang TC, Hu Z, Liu H, Smith B, Wu CH. Framework for a protein ontology. BMC bioinformatics 2007; 8 Suppl 9: S1	CT	Not applicable	As example	Other/multiple/unknown/not applicable	United States	2007
92.	Northey L, Bhardwaj G, Curran S, McGirr J. Epidemiology of eye trauma presenting to a regional hospital emergency department and ophthalmology service Clin. Exp. Ophthalmol. 2012; 40: 79	CT	Not applicable	As example	Ophthalmology	United States	2012
93.	Parker CG, Rocha RA, Campbell JR, Tu SW, Huff SM. Detailed clinical models for sharable, executable guidelines. Studies in health technology and informatics 2004; 107(Pt 1): 145-8	CT	Not applicable	As example	Other/multiple/unknown/not applicable	United States	2004
94.	Pathak J, Solbrig HR, Buntrock JD, Johnson TM, Chute CG. LexGrid: a framework for representing, storing, and querying biomedical terminologies from simple to sublime. Journal of the American Medical Informatics Association : JAMIA 2009 May-Jun; 16(3): 305-15	CT	Not applicable	As example	Other/multiple/unknown/not applicable	United States	2009
95.	Pugliano L, Fumagalli D, Irrthum A, Schenk B, Saini KS, Dolci SM, Arahmani A, Piccart MJ, Bucur A. Integrate: A new model in collaborative breast cancer research Breast 2011; 20: S26-S27	CT	Not applicable	As example	Cancer	United Kingdom	2011
96.	Rhodes H. Managing data content. Clinical data management programs improve reimbursement. Journal of AHIMA / American Health Information Management Association 2006 Feb; 77(2): 62-3, 72	CT	Not applicable	As example	Other/multiple/unknown/not applicable	United States	2006
97.	Rodrigues JM, Kumar A, Bousquet C, Trombert B. Using the CEN/ISO standard for categorial structure to harmonise the development of WHO international terminologies. Studies in health technology and informatics 2009; 150: 255-9	CT	Not applicable	As example	Other/multiple/unknown/not applicable	France	2009

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
98.	Rodrigues JM, Rosse C, Fogelberg M, Kumar A, Paviot BT. A road from health care classifications and coding systems to biomedical ontology: the CEN categorial structure for terminologies of human anatomy: Catanat. Studies in health technology and informatics 2007; 129(Pt 1): 735-40	CT	Not applicable	As example	Other/multiple/unknown/not applicable	Multiple countries	2007
99.	Rodriguez-Gonzalez A, Hernandez-Chan G, Colomo-Palacios R, Gomez-Berbis JM, Garcia-Crespo A, Alor-Hernandez G, Valencia-Garcia R. Towards an ontology to support Semantics enabled diagnostic decision support systems Curr. Bioinform. 2012; 7(3): 234-245	CT	Not applicable	As example	Other/multiple/unknown/not applicable	United States	2012
100.	Roemer LK, Richardson SJ, Sward K, Tilley C. Redundancy in a computer-generated order list: meeting the needs of nurses at various levels of practice expertise. Computers, informatics, nursing : CIN 2005 Mar-Apr; 23(2): 73-82; quiz 83-4	CT	Not applicable	As example	Nursing	United States	2005
101.	Röhrig R, Rüth R. [Intelligent telemedicine in intensive care units. Bed-side operation of medical technology devices and IT in intensive care medicine]. Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz 2009 Mar; 52(3): 279-86	CT	Not applicable	As example	Other/multiple/unknown/not applicable	Germany	2009
102.	Rojo MG Standardization efforts of digital pathology in Europe Anal. Cell. Pathol. 2011; 34(4): 172-173	CT	Not applicable	As example	Pathology	Multiple countries	2011
103.	Rojo MG, Castro AM, Gonçalves L. COST Action "EuroTelepath": digital pathology integration in electronic health record, including primary care centres. Diagnostic pathology 2011; 6 Suppl 1: S6	CT	Not applicable	As example	Pathology	Spain	2011
104.	Rosenbloom ST, Shultz E, Wright A. Managing the flood of codes: maintaining patient problem lists in the era of Meaningful Use and ICD10. AMIA Annu Symp Proc 2012;2012:8-10	CT	Not applicable	As example	Problem list / diagnoses	United States	2012
105.	Rosendal M, Falkø E. [Diagnostic classification in Denmark with emphasis on general practice]. Ugeskrift for laeger 2009 Mar 16; 171(12): 997-1000	CT	Not applicable	As example	Primary care	Denmark	2009
106.	Runy LA. IT standards. Speaking a common language. Hospitals & health networks / AHA 2004 Feb; 78(2): 63-8, 2	CT	Not applicable	As example	Other/multiple/unknown/not applicable	Unknown	2004

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
107.	Sim I, Detmer DE. Beyond trial registration: a global trial bank for clinical trial reporting. PLoS medicine 2005 Nov; 2(11): e365	CT	Not applicable	As example	Drugs	United States	2005
108.	Slee VN, Slee D, Schmidt HJ. The tyranny of the diagnosis code. North Carolina medical journal 2005 Sep-Oct; 66(5): 331-7	CT	Not applicable	As example	Other/multiple/unknown/not applicable	United States	2005
109.	Sullivan F, Wyatt JC. How computers help make efficient use of consultations. BMJ (Clinical research ed.) 2005 Oct 29; 331(7523): 1010-2	CT	Not applicable	As example	Other/multiple/unknown/not applicable	United States	2005
110.	Wallace S. Mean what you say. Before achieving IT interoperability, leaders must agree on a definition. Modern healthcare 2005 Mar 21; 35(12): 22	CT	Not applicable	As example	Other/multiple/unknown/not applicable	United States	2005
111.	Wang L, Wang J, Wang M, Li Y, Liang Y, Xu D. Using Internet search engines to obtain medical information: a comparative study. Journal of medical Internet research 2012 May 16;14(3):e74	CT	Not applicable	As example	Cancer	China	2012
112.	White JM, Kalenderian E, Stark PC, Ramoni RL, Vaderhobli R, Walji MF. Evaluating a dental diagnostic terminology in an electronic health record. Journal of dental education 2011 May; 75(5): 605-15	CT	Not applicable	As example	Dentistry	United States	2011
113.	Wilson ML, Windle PE, Krenzischek DA, Clifford TL. Perianesthesia data elements: from concept, to validation, to operations. Journal of perianesthesia nursing : official journal of the American Society of PeriAnesthesia Nurses / American Society of PeriAnesthesia Nurses 2008 Jun; 23(3): 198-203	CT	Not applicable	As example	Nursing	United States	2008
114.	de Lusignan S. SNOMED is coming, and more about using and interacting with technology in primary care. Informatics in primary care 2011;19(1):1-2	CT	Not applicable	Letter	Other/multiple/unknown/not applicable	United Kingdom	2011
115.	Jenders RA. Classification of psychiatric disorders. JAMA : the journal of the American Medical Association 2005 Oct 19; 294(15): 1899; author reply 1899-900	CT	Not applicable	Letter	Other/multiple/unknown/not applicable	Unknown	2005
116.	Moore H. Getting the SNOMED ball rolling in Australia. Journal of AHIMA / American Health Information Management Association 2004 Nov-Dec; 75(10): 6	CT	Not applicable	Letter	Other/multiple/unknown/not applicable	Australia	2004
117.	Nguyen A, Sun YK, Sitbon L, Geva S. Representation of assertions in clinical free-text using SNOMED CT Australas. Med. J. 2012; 5(9): 472-474	CT	Not applicable	Letter	Other/multiple/unknown/not applicable	Australia	2012

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
118.	Rajeev D, Staes CJ, Evans RS, Mortice S, Rolfs R, Samore MH, Whitney J, Kurzban R, Huff SM. In response to letter to the editor: 'Concerning SNOMED-CT content for public health case reports' J. Am. Med. Informatics Assoc. 2010; 17(5): 613-614	CT	Not applicable	Letter	Other/multiple/unknown/not applicable	Multiple countries	2010
119.	Rector A. AIM: A personal view of where I have been and where we might be going Artif. Intell. Med. 2001; 23(1): 111-127	CT	Not applicable	Letter	Other/multiple/unknown/not applicable	United Kingdom	2001
120.	Roumier J, Jamoulle M, Stichele RV, Romary L, Cardillo E. Towards a terminologies support system in primary care. Informatics in primary care 2011;19(4):257-8	CT	Not applicable	Letter	Primary care	United Kingdom	2012
121.	Wilcke JR, Green JM, Spackman KA, Martin MK, Case JT, Santamaria SL, Zimmerman K. Concerning SNOMED-CT content for public health case reports. Journal of the American Medical Informatics Association : JAMIA 2010 Sep-Oct; 17(5): 613; author reply 613-4	CT	Not applicable	Letter	Other/multiple/unknown/not applicable	United States	2010
122.	Ahmadian L, Cornet R, VAN Klei WA, DE Keizer NF. Data collection variation in preoperative assessment: a literature review. Computers, informatics, nursing : CIN 2011 Nov; 29(11): 662-70	CT	Not applicable	Reviews and surveys	Preoperative	Netherlands	2011
123.	Ahmadian L, Cornet R, van Klei WA, de Keizer NF. Diversity in preoperative-assessment data collection, a literature review. Studies in health technology and informatics 2008; 136: 127-32	CT	Not applicable	Reviews and surveys	Preoperative	Netherlands	2008
124.	Arts DG, Cornet R, de Jonge E, de Keizer NF. Methods for evaluation of medical terminological systems -- a literature review and a case study. Methods of information in medicine 2005; 44(5): 616-25	Multiple	Not applicable	Reviews and surveys	Other/multiple/unknown/not applicable	Netherlands	2005
125.	Cornet R, de Keizer N. Forty years of SNOMED: a literature review. BMC medical informatics and decision making 2008; 8 Suppl 1: S2	Multiple	Not applicable	Reviews and surveys	Other/multiple/unknown/not applicable	Multiple countries	2008
126.	Cornet R, Van Eldik A, De Keizer N. Inventory of tools for dutch clinical language processing. Studies in health technology and informatics 2012;180:245-9	CT	Not applicable	Reviews and surveys	Other/multiple/unknown/not applicable	Netherlands	2012

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
127.	Elhanan G, Perl Y, Geller J. A Survey of Direct Users and Uses of SNOMED CT: 2010 Status. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2010; 2010: 207-11	CT	Not applicable	Reviews and surveys	Other/multiple/unknown/not applicable	United States	2010
128.	Elhanan G, Perl Y, Geller J. A survey of SNOMED CT direct users, 2010: impressions and preferences regarding content and quality. Journal of the American Medical Informatics Association : JAMIA 2011 Aug 11;	CT	Not applicable	Reviews and surveys	Other/multiple/unknown/not applicable	Multiple countries	2011
129.	Giannangelo K, Fenton SH. SNOMED CT survey: an assessment of implementation in EMR/EHR applications. Perspectives in health information management / AHIMA, American Health Information Management Association 2008; 5: 7	CT	Not applicable	Reviews and surveys	Other/multiple/unknown/not applicable	United States	2008
130.	Lee D, Cornet R, Lau F, de Keizer N. A survey of SNOMED CT implementations. Journal of biomedical informatics 2012 Oct 3;	CT	Not applicable	Reviews and surveys	Other/multiple/unknown/not applicable	Multiple countries	2012
131.	Adamusiak T, Bodenreider O. Quality assurance in LOINC using Description Logic. AMIA Annu Symp Proc 2012;2012:1099-108	CT	Pre-development	Compare to/map with other TS	Laboratory	United States	2012
132.	Alecu I, Bousquet C, Jaulent MC. A case report: using SNOMED CT for grouping Adverse Drug Reactions Terms. BMC medical informatics and decision making 2008; 8 Suppl 1: S4	CT	Pre-development	Compare to/map with other TS	Drugs	France	2008
133.	Alecu I, Bousquet C, Mougin F, Jaulent MC. Mapping of the WHO-ART terminology on Snomed CT to improve grouping of related adverse drug reactions. Studies in health technology and informatics 2006; 124: 833-8	CT	Pre-development	Compare to/map with other TS	Drugs	France	2006
134.	Bakhshi-Raiez F, Ahmadian L, Cornet R, de Jonge E, de Keizer NF. Construction of an interface terminology on SNOMED CT. Generic approach and its application in intensive care. Methods of information in medicine 2010; 49(4): 349-59	CT	Pre-development	Compare to/map with other TS	Intensive care	Netherlands	2010
135.	Bakhshi-Raiez F, Cornet R, Bosman RJ, Joore H, de Keizer NF. Using SNOMED CT to identify a crossmap between two classification systems: a comparison with an expert-based and a data-driven strategy. Studies in health technology and informatics 2010; 160(Pt 2): 1035-9	CT	Pre-development	Compare to/map with other TS	Intensive care	Netherlands	2010

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
136.	Bakhshi-Raiez F, Cornet R, de Keizer NF. Cross-mapping APACHE IV "reasons for intensive care admission" classification to SNOMED CT. Studies in health technology and informatics 2008; 136: 779-84	CT	Pre-development	Compare to/map with other TS	Intensive care	Netherlands	2008
137.	Bakken S, Warren J, Lundberg C, Casey A, Correia C, Konicek D, Zingo C. An evaluation of the utility of the CEN categorical structure for nursing diagnoses as a terminology model for integrating nursing diagnosis concepts into SNOMED. Studies in health technology and informatics 2001; 84(Pt 1): 151-5	Multiple	Pre-development	Compare to/map with other TS	Nursing	United Kingdom	2001
138.	Bakken S, Warren JJ, Lundberg C, Casey A, Correia C, Konicek D, Zingo C. An evaluation of the usefulness of two terminology models for integrating nursing diagnosis concepts into SNOMED Clinical Terms. International journal of medical informatics 2002 Dec 18; 68(1-3): 71-7	CT	Pre-development	Compare to/map with other TS	Nursing	United States	2002
139.	Berman JJ Tumor taxonomy for the developmental lineage classification of neoplasms. BMC cancer 2004 Nov 30; 4: 88	CT	Pre-development	Compare to/map with other TS	Cancer	United States	2004
140.	Bodenreider O. Issues in mapping LOINC laboratory tests to SNOMED CT. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2008; : 51-5	CT	Pre-development	Compare to/map with other TS	Laboratory	United States	2008
141.	Bodenreider O. Using SNOMED CT in combination with MedDRA for reporting signal detection and adverse drug reactions reporting. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2009; 2009: 45-9	CT	Pre-development	Compare to/map with other TS	Drugs	United States	2009
142.	Bodenreider O, Zhang S. Comparing the representation of anatomy in the FMA and SNOMED CT. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2006; : 46-50	CT	Pre-development	Compare to/map with other TS	Anatomy	United States	2006
143.	Bousquet C, Gasperina P, Trombert B, Clavel L, Kumar A, Rodrigues JM. Ontological representation of adverse drug reactions using the Foundational Model of Anatomy. Studies in health technology and informatics 2009; 150: 507-11	CT	Pre-development	Compare to/map with other TS	Drugs	France	2009

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
144.	Bowman S. Coordinating SNOMED-CT and ICD-10. Journal of AHIMA / American Health Information Management Association 2005 Jul-Aug; 76(7): 60-1	CT	Pre-development	Compare to/map with other TS	Other/multiple/unknown/not applicable	United States	2005
145.	Brouch K. AHIMA project offers insights into SNOMED, ICD-9-CM mapping process. Journal of AHIMA / American Health Information Management Association 2003 Jul-Aug; 74(7): 52-5	CT	Pre-development	Compare to/map with other TS	Other/multiple/unknown/not applicable	United States	2003
146.	Brown SH, Husser CS, Wahner-Roedler D, Bailey S, Nugent L, Porter K, Bauer BA, Elkin PL. Using SNOMED CT as a reference terminology to cross map two highly pre-coordinated classification systems. Studies in health technology and informatics 2007; 129(Pt 1): 636-9	CT	Pre-development	Compare to/map with other TS	Problem list / diagnoses	United States	2007
147.	Burton MM, Simonaitis L, Schadow G. Medication and indication linkage: A practical therapy for the problem list? AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2008; : 86-90	CT	Pre-development	Compare to/map with other TS	Problem list / diagnoses	United States	2008
148.	Cantor MN, Lussier YA. Putting data integration into practice: using biomedical terminologies to add structure to existing data sources. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2003; : 125-9	CT	Pre-development	Compare to/map with other TS	Genetics	United States	2003
149.	Dykes PC, DaDamio RR, Goldsmith D, Kim HE, Ohashi K, Saba VK. Leveraging standards to support patient-centric interdisciplinary plans of care. AMIA Annu Symp Proc 2011;2011:356-63	CT	Pre-development	Compare to/map with other TS	Nursing	United States	2011
150.	Fung KW, Bodenreider O. Utilizing the UMLS for semantic mapping between terminologies. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2005; : 266-70	CT	Pre-development	Compare to/map with other TS	Other/multiple/unknown/not applicable	United States	2005
151.	Fung KW, Bodenreider O, Aronson AR, Hole WT, Srinivasan S. Combining lexical and semantic methods of inter-terminology mapping using the UMLS. Studies in health technology and informatics 2007; 129(Pt 1): 605-9	CT	Pre-development	Compare to/map with other TS	Problem list / diagnoses	United States	2007

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
152.	Fung KW, Hole WT, Nelson SJ, Srinivasan S, Powell T, Roth L. Integrating SNOMED CT into the UMLS: an exploration of different views of synonymy and quality of editing. Journal of the American Medical Informatics Association : JAMIA 2005 Jul-Aug; 12(4): 486-94	CT	Pre-development	Compare to/map with other TS	Other/multiple/unknown/not applicable	United States	2005
153.	Fung KW, Xu J. Synergism between the mapping projects from SNOMED CT to ICD-10 and ICD-10-CM. AMIA Annu Symp Proc 2012;2012:218-27	CT	Pre-development	Compare to/map with other TS	Other/multiple/unknown/not applicable	United States	2012
154.	Giannangelo K, Millar J. Mapping SNOMED CT to ICD-10. Studies in health technology and informatics 2012;180:83-7	CT	Pre-development	Compare to/map with other TS	Problem list / diagnoses	United States	2012
155.	Goldberg LJ, Ceusters W, Eisner J, Smith B. The Significance of SNODENT. Studies in health technology and informatics 2005; 116: 737-42	CT	Pre-development	Compare to/map with other TS	Dentistry	United States	2005
156.	Grabar N, Hamon T. Exploitation of linguistic indicators for automatic weighting of synonyms induced within three biomedical terminologies. Studies in health technology and informatics 2010; 160(Pt 2): 1015-9	Multiple	Pre-development	Compare to/map with other TS	Other/multiple/unknown/not applicable	France	2010
157.	Hardiker NR, Casey A, Coenen A, Konicek D. Mutual enhancement of diverse terminologies. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2006; : 319-23	CT	Pre-development	Compare to/map with other TS	Nursing	United Kingdom	2006
158.	Harman TL, Seeley RA, Oliveira IM, Sheide A, Kartchner T, Woolstenhulme RD, Wilson PS, Lau LM, Matney SA. Standardized mapping of nursing assessments across 59 U.S. military treatment facilities. AMIA Annu Symp Proc 2012;2012:331-9	CT	Pre-development	Compare to/map with other TS	Nursing	United States	2012
159.	Hogan WR, Slee VN. Measuring the Information Gain of Diagnosis vs. Diagnosis Category Coding. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2010; 2010: 306-10	CT	Pre-development	Compare to/map with other TS	Other/multiple/unknown/not applicable	United States	2010
160.	Imel M. A closer look: the SNOMED clinical terms to ICD-9-CM mapping. Journal of AHIMA / American Health Information Management Association 2002 Jun; 73(6): 66-9; quiz 71-2	CT	Pre-development	Compare to/map with other TS	Other/multiple/unknown/not applicable	United States	2002

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
161.	Jacobs AK, Quinn TA, Nelson SJ. Mapping SNOMED-CT concepts to MeSH concepts. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2006; : 965	CT	Pre-development	Compare to/map with other TS	Other/multiple/unknown/not applicable	United States	2006
162.	Jaulent MC, Alecu I. Evaluation of an ontological resource for pharmacovigilance. Studies in health technology and informatics 2009; 150: 522-6	CT	Pre-development	Compare to/map with other TS	Drugs	France	2009
163.	Kalenderian E, Ramoni RL, White JM, Schoonheim-Klein ME, Stark PC, Kimmes NS, Zeller GG, Willis GP, Walji MF. The development of a dental diagnostic terminology. Journal of dental education 2011 Jan; 75(1): 68-76	CT	Pre-development	Compare to/map with other TS	Dentistry	Multiple countries	2011
164.	Kim SY, Kim HH, Lee IK, Kim HS, Cho H. Proposed Algorithm with Standard Terminologies (SNOMED and CPT) for Automated Generation of Medical Bills for Laboratory Tests. Healthcare informatics research 2010 Sep; 16(3): 185-90	CT	Pre-development	Compare to/map with other TS	Laboratory	South Korea	2010
165.	Kim TY, Coenen A, Hardiker N. Semantic mappings and locality of nursing diagnostic concepts in UMLS. Journal of biomedical informatics 2011 Sep 18;	CT	Pre-development	Compare to/map with other TS	Nursing	United States	2011
166.	Kumar A, Smith B, Novotny DD. Biomedical informatics and granularity. Comparative and functional genomics 2004; 5(6-7): 501-8	CT	Pre-development	Compare to/map with other TS	Other/multiple/unknown/not applicable	United States	2004
167.	Le H.V., Beach K.J. Advantages of mapping ICD-9-CM into SNOMED CT concepts and hierarchy. Pharmacoepidemiology and Drug Safety (PDS) 2009; 18: S33-S34	CT	Pre-development	Compare to/map with other TS	Problem list / diagnoses	United States	2009
168.	Le H.V., Beach K.J. Automated robust mapping GPRD medical diagnosis read codes to SNOMED CT for OMOP project. Pharmacoepidemiology and Drug Safety 2010; 19: S106-	CT	Pre-development	Compare to/map with other TS	Drugs	United Kingdom	2010
169.	Li W, Tolson J, Horan TC. Creating public health standard vocabularies: mapping a set of CDC's pathogen codes to SNOMED concepts. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2003; : 907	CT	Pre-development	Compare to/map with other TS	Internal medicine	United States	2003

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
170.	Lu DF, Eichmann D, Konicek D, Park HT, Ucharattana P, Delaney C. Standardized nursing language in the systematized nomenclature of medicine clinical terms: A cross-mapping validation method. Computers, informatics, nursing : CIN 2006 Sep-Oct; 24(5): 288-96	CT	Pre-development	Compare to/map with other TS	Nursing	United States	2006
171.	Lu DF, Park HT, Ucharattana P, Konicek D, Delaney C. Nursing outcomes classification in the systematized nomenclature of medicine clinical terms: a cross-mapping validation. Computers, informatics, nursing : CIN 2007 May-Jun; 25(3): 159-70	CT	Pre-development	Compare to/map with other TS	Nursing	United States	2007
172.	Lussier YA, Li J. Terminological mapping for high throughput comparative biology of phenotypes. Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing 2004; : 202-13	CT	Pre-development	Compare to/map with other TS	Biology	United States	2004
173.	Marquet G, Mosser J, Burgun A. A method exploiting syntactic patterns and the UMLS semantics for aligning biomedical ontologies: the case of OBO disease ontologies. International journal of medical informatics 2007 Dec; 76 Suppl 3: S353-61	CT	Pre-development	Compare to/map with other TS	Pathology	France	2007
174.	Merabti T, Letord C, Abdoune H, Lecroq T, Joubert M, Darmoni SJ. Projection and inheritance of SNOMED CT relations between MeSH terms. Studies in health technology and informatics 2009; 150: 233-7	CT	Pre-development	Compare to/map with other TS	Other/multiple/unknown/not applicable	France	2009
175.	Montella D, Brown SH, Elkin PL, Jackson JC, Rosenbloom ST, Wahner-Roedler D, Welsh G, Cotton B, Guillamondegui OD, Lew H, Taber KH, Tupler LA, Vanderploeg R, Speroff T. Comparison of SNOMED CT versus Medcin terminology concept coverage for mild Traumatic Brain Injury. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2011;2011:969-78	CT	Pre-development	Compare to/map with other TS	Internal medicine	United States	2011
176.	Nadkarni PM, Darer JA. Migrating existing clinical content from ICD-9 to SNOMED. Journal of the American Medical Informatics Association : JAMIA 2010 Sep-Oct; 17(5): 602-7	CT	Pre-development	Compare to/map with other TS	Other/multiple/unknown/not applicable	United States	2010
177.	Nadkarni PM, Darer JD. Determining correspondences between high-frequency MedDRA concepts and SNOMED: a case study. BMC medical informatics and decision making 2010; 10: 66	CT	Pre-development	Compare to/map with other TS	Drugs	United States	2010

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
178.	Nash S Integrating Osteopathic terminology into SNOMED CT. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2005; : 1061	CT	Pre-development	Compare to/map with other TS	Osteopathy	United States	2005
179.	Nyström M, Vikström A, Nilsson G, Orman H, Ahlfeldt H. Visualization of disease distribution with SNOMED CT and ICD-10. Studies in health technology and informatics 2010; 160(Pt 2): 1100-3	CT	Pre-development	Compare to/map with other TS	Primary care	United Kingdom	2010
180.	Nyström M, Vikström A, Nilsson GH, Ahlfeldt H, Orman H. Enriching a primary health care version of ICD-10 using SNOMED CT mapping. Journal of biomedical semantics 2010; 1(1): 7	CT	Pre-development	Compare to/map with other TS	Primary care	Sweden	2010
181.	Park HA, Lundberg C, Coenen A, Konicek D. Evaluation of the content coverage of SNOMED CT representing ICNP seven-axis version 1 concepts. Methods of information in medicine 2011; 50(5): 472-8	CT	Pre-development	Compare to/map with other TS	Nursing	United States	2011
182.	Park HA, Lundberg C, Coenen A, Konicek D. Mapping ICNP Version 1 concepts to SNOMED CT. Studies in health technology and informatics 2010; 160(Pt 2): 1109-13	CT	Pre-development	Compare to/map with other TS	Nursing	South Korea	2010
183.	Park HA, Lundberg CB, Coenen A, Konicek DJ. Evaluation of the content coverage of SNOMED-CT to represent ICNP Version 1 catalogues. Studies in health technology and informatics 2009; 146: 303-7	CT	Pre-development	Compare to/map with other TS	Nursing	South Korea	2009
184.	Park HT, Lu DF, Konicek D, Delaney C. Nursing interventions classification in systematized nomenclature of medicine clinical terms: a cross-mapping validation. Computers, informatics, nursing : CIN 2007 Jul-Aug; 25(4): 198-208; quiz 209-10	CT	Pre-development	Compare to/map with other TS	Nursing	United States	2007
185.	Patel CO, Cimino JJ. A scale-free network view of the UMLS to learn terminology translations. Studies in health technology and informatics 2007; 129(Pt 1): 689-93	CT	Pre-development	Compare to/map with other TS	Problem list / diagnoses	United States	2007
186.	Rath A, Olry A, Dhombres F, Brandt MM, Urbero B, Ayme S. Representation of rare diseases in health information systems: the Orphanet approach to serve a wide range of end users. Human mutation 2012 May;33(5):803-8	CT	Pre-development	Compare to/map with other TS	Other/multiple/unknown/not applicable	Multiple countries	2012

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
187.	Reich C, Ryan PB, Stang PE, Rocca M. Evaluation of alternative standardized terminologies for medical conditions within a network of observational healthcare databases. Journal of biomedical informatics 2012 Aug;45(4):689-96	CT	Pre-development	Compare to/map with other TS	Other/multiple/unknown/not applicable	United States	2012
188.	Richesson RL, Fung KW, Krischer JP. Heterogeneous but "standard" coding systems for adverse events: Issues in achieving interoperability between apples and oranges. Contemporary clinical trials 2008 Sep; 29(5): 635-45	CT	Pre-development	Compare to/map with other TS	Drugs	United States	2008
189.	Robinson J, de Lusignan S, Kostkova P, Madge B. Using UMLS to map from a library to a clinical classification: Improving the functionality of a digital library. Studies in health technology and informatics 2006; 121: 86-95	CT	Pre-development	Compare to/map with other TS	Other/multiple/unknown/not applicable	United Kingdom	2006
190.	Rosenbloom ST, Brown SH, Froehling D, Bauer BA, Wahner-Roedler DL, Gregg WM, Elkin PL. Using SNOMED CT to represent two interface terminologies. Journal of the American Medical Informatics Association : JAMIA 2009 Jan-Feb; 16(1): 81-8	CT	Pre-development	Compare to/map with other TS	Other/multiple/unknown/not applicable	United States	2009
191.	Saitwal H, Qing D, Jones S, Bernstam EV, Chute CG, Johnson TR. Cross-terminology mapping challenges: a demonstration using medication terminological systems. Journal of biomedical informatics 2012 Aug;45(4):613-25	CT	Pre-development	Compare to/map with other TS	Drugs	United States	2012
192.	Sarntivijai S, Xiang Z, Shedden KA, Markel H, Omenn GS, Athey BD, He Y. Ontology-based combinatorial comparative analysis of adverse events associated with killed and live influenza vaccines. PLoS One 2012;7(11):e49941	CT	Pre-development	Compare to/map with other TS	Drugs	United States	2012
193.	So EY, Park HA. Exploring the Possibility of Information Sharing between the Medical and Nursing Domains by Mapping Medical Records to SNOMED CT and ICNP. Healthcare informatics research 2011 Sep; 17(3): 156-61	CT	Pre-development	Compare to/map with other TS	Gastroenterology	South Korea	2011
194.	Steindel SJ. A comparison between a SNOMED CT problem list and the ICD-10-CM/PCS HIPAA code sets. Perspectives in health information management / AHIMA, American Health Information Management Association 2012;9:1b	CT	Pre-development	Compare to/map with other TS	Problem list / diagnoses	United States	2012

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
195.	Torres-Urquidy MH, Schleyer T. Evaluation of the Systematized Nomenclature of Dentistry using case reports: preliminary results. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2006; : 1124	CT	Pre-development	Compare to/map with other TS	Dentistry	United States	2006
196.	Uribe GA, López DM, Blobel B. Architectural analysis of clinical ontologies for pHealth interoperability. Studies in health technology and informatics 2012;177:176-82	CT	Pre-development	Compare to/map with other TS	Other/multiple/unknown/not applicable	Multiple countries	2012
197.	Vikström A, Nyström M, Ahlfeldt H, Strender LE, Nilsson GH. Views of diagnosis distribution in primary care in 2.5 million encounters in Stockholm: a comparison between ICD-10 and SNOMED CT. Informatics in primary care 2010; 18(1): 17-29	CT	Pre-development	Compare to/map with other TS	Primary care	Sweden	2010
198.	Vikström A, Skånér Y, Strender LE, Nilsson GH. Mapping the categories of the Swedish primary health care version of ICD-10 to SNOMED CT concepts: rule development and intercoder reliability in a mapping trial. BMC medical informatics and decision making 2007; 7: 9	CT	Pre-development	Compare to/map with other TS	Primary care	Sweden	2007
199.	Wang X, Quek HN, Cantor M, Kra P, Schultz A, Lussier YA. Automating terminological networks to link heterogeneous biomedical databases. Studies in health technology and informatics 2004; 107(Pt 1): 555-9	CT	Pre-development	Compare to/map with other TS	Other/multiple/unknown/not applicable	United States	2004
200.	Wang Y, Patrick J, Miller G, O'Hallaran J. A computational linguistics motivated mapping of ICPC-2 PLUS to SNOMED CT. BMC medical informatics and decision making 2008; 8 Suppl 1: S5	CT	Pre-development	Compare to/map with other TS	Primary care	Australia	2008
201.	Warden GI, Lacson R, Khorasani R. Leveraging terminologies for retrieval of radiology reports with critical imaging findings. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2011;2011:1481-8	CT	Pre-development	Compare to/map with other TS	Radiology	United States	2011
202.	Westra BL, Bauman R, Delaney CW, Lundberg CB, Petersen C. Validation of concept mapping between PNDS and SNOMED CT. AORN journal 2008 Jun; 87(6): 1217-29	CT	Pre-development	Compare to/map with other TS	Nursing	United States	2008

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
203.	Wu ST, Liu H, Li D, Tao C, Musen MA, Chute CG, Shah NH. Unified Medical Language System term occurrences in clinical notes: a large-scale corpus analysis. Journal of the American Medical Informatics Association : JAMIA 2012 Jun 1;19(e1):e149-e156	CT	Pre-development	Compare to/map with other TS	Problem list / diagnoses	United States	2012
204.	Xu J, Fung KW. Handling age specification in the SNOMED CT to ICD-10-CM cross-map. AMIA Annu Symp Proc 2012;2012:1014-22	CT	Pre-development	Compare to/map with other TS	Other/multiple/unknown/not applicable	United States	2012
205.	Berman JJ. Nomenclature-based data retrieval without prior annotation: facilitating biomedical data integration with fast doublet matching. In silico biology 2005; 5(3): 313-22	CT	Pre-development	Design considerations	Other/multiple/unknown/not applicable	United States	2005
206.	Ceusters W. SNOMED CT revisions and coded data repositories: when to upgrade? AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2011;2011:197-206	CT	Pre-development	Design considerations	Cancer	United States	2011
207.	Cruanes J, Romá-Ferri MT, Lloret E. Measuring lexical similarity methods for textual mapping in nursing diagnoses in Spanish and SNOMED-CT. Studies in health technology and informatics 2012;180:255-9	CT	Pre-development	Design considerations	Nursing	Spain	2012
208.	Cuggia M, Bayat S, Garcelon N, Sanders L, Rouget F, Coursin A, Pladys P. A full-text information retrieval system for an epidemiological registry. Studies in health technology and informatics 2010; 160(Pt 1): 491-5	Multiple	Pre-development	Design considerations	Paediatrics	France	2010
209.	Daniel C, Buemi A, Mazuel L, Ouagne D, Charlet J. Functional requirements of terminology services for coupling interface terminologies to reference terminologies. Studies in health technology and informatics 2009; 150: 205-9	Multiple	Pre-development	Design considerations	Hospital	France	2009
210.	de Lusignan S, Chan T, Jones S. Large complex terminologies: more coding choice, but harder to find data--reflections on introduction of SNOMED CT (Systematized Nomenclature of Medicine--Clinical Terms) as an NHS standard. Informatics in primary care 2011;19(1):3-5	CT	Pre-development	Design considerations	Primary care	United Kingdom	2011

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
211.	Dinh D, Tamine L, Boubekeur F. Factors affecting the effectiveness of biomedical document indexing and retrieval based on terminologies. Artificial intelligence in medicine 2012 Oct 20;	CT	Pre-development	Design considerations	Other/multiple/unknown/not applicable	Multiple countries	2012
212.	Dolin RH. SNOMED at the point of service: the challenges and opportunities of SNOMED-encoded data. Journal of AHIMA / American Health Information Management Association 2010 Mar; 81(3): 42-3	CT	Pre-development	Design considerations	Other/multiple/unknown/not applicable	United States	2010
213.	Dolin RH, Spackman KA, Markwell D. Selective retrieval of pre- and post-coordinated SNOMED concepts. Proceedings / AMIA ... Annual Symposium. AMIA Symposium 2002; : 210-4	CT	Pre-development	Design considerations	Other/multiple/unknown/not applicable	United States	2002
214.	Fung KW, Xu J, Rosenbloom ST, Mohr D, Maram N, Suther T. Testing Three Problem List Terminologies in a simulated data entry environment. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2011;2011:445-54	CT	Pre-development	Design considerations	Problem list / diagnoses	United States	2011
215.	Garcia Rojo M., Laurinavicius A., Punys V., Gehring S. Encoding pathology diagnosis with SNOMED CT: Feasibility of migration from SNOMED II to SNOMED CT. Virchows Archiv 2011; 459: S309-	CT	Pre-development	Design considerations	Pathology	Multiple countries	2011
216.	Green JM, Wilcke JR, Abbott J, Rees LP. Development and evaluation of methods for structured recording of heart murmur findings using SNOMED-CT post-coordination. Journal of the American Medical Informatics Association : JAMIA 2006 May-Jun; 13(3): 321-33	CT	Pre-development	Design considerations	Cardiology	Unknown	2006
217.	Handler J, Feied C, Smith M, Gillam M. Improving a UMLS based allergy list for use in live electronic medical record systems. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2006; : 942	CT	Pre-development	Design considerations	Allergies	United States	2006
218.	Hansen DP, Kemp ML, Mills SR, Mercer MA, Frosdick PA, Lawley MJ. Developing a national emergency department data reference set based on SNOMED CT. The Medical journal of Australia 2011 Feb 21; 194(4): S8-10	CT	Pre-development	Design considerations	Emergency department/room	Australia	2011

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
219.	Herbert I. CLICSIG report: issues around compositional terminologies, SNOMED-CT in particular. Informatics in primary care 2007; 15(3): 193-7	CT	Pre-development	Design considerations	Other/multiple/unknown/not applicable	Unknown	2007
220.	Herbert I, Hawking M. Bringing SNOMED-CT into use within primary care. Informatics in primary care 2005; 13(1): 61-4	CT	Pre-development	Design considerations	Primary care	United Kingdom	2005
221.	Ingenerf J, Beisiegel T. A version management system for SNOMED CT. Studies in health technology and informatics 2008; 136: 827-32	CT	Pre-development	Design considerations	Other/multiple/unknown/not applicable	Germany	2008
222.	Jarman J, Berndt DJ. Throw the bath water out, keep the baby: keeping medically-relevant terms for text mining. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2010; 2010: 336-40	CT	Pre-development	Design considerations	Hospital	United States	2010
223.	Kanter AS, Wang AY, Masarie FE, Naeymi-Rad F, Safran C. Interface terminologies: bridging the gap between theory and reality for Africa. Studies in health technology and informatics 2008; 136: 27-32	CT	Pre-development	Design considerations	Other/multiple/unknown/not applicable	Africa	2008
224.	Kemp M, Walker S, Scott P. Coding of thoughts, words and things. The HIM journal 2005; 34(2): 54-6	CT	Pre-development	Design considerations	Other/multiple/unknown/not applicable	Australia	2005
225.	Kokkinakis D. Semantic Relations of Binary Compounds Annotated with SNOMED CT. Studies in health technology and informatics 2012;180:169-73	CT	Pre-development	Design considerations	Other/multiple/unknown/not applicable	Sweden	2012
226.	Lee D, Cornet R, Lau F. Implications of SNOMED CT versioning. International journal of medical informatics 2011 Jun; 80(6): 442-53	CT	Pre-development	Design considerations	Problem list / diagnoses	Canada	2011
227.	Lee DH, Lau FY, Quan H. A method for encoding clinical datasets with SNOMED CT. BMC medical informatics and decision making 2010; 10: 53	CT	Pre-development	Design considerations	Palliative care	Canada	2010
228.	Leroux H, McBride S, Lefort L, Kemp M, Gibson S. A method for the semantic enrichment of clinical trial data. Studies in health technology and informatics 2012;178:111-6	CT	Pre-development	Design considerations	Other/multiple/unknown/not applicable	Australia	2012
229.	Lieberman MI, Ricciardi TN, Masarie FE, Spackman KA. The use of SNOMED CT simplifies querying of a clinical data warehouse. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2003; : 910	CT	Pre-development	Design considerations	Cardiology	United States	2003

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
230.	López-García P, Boeker M, Illarramendi A, Schulz S. Usability-driven pruning of large ontologies: the case of SNOMED CT. Journal of the American Medical Informatics Association : JAMIA 2012 Jun 1;19(e1):e102-e109	CT	Pre-development	Design considerations	Cardiology	Spain	2012
231.	Mougin F, Bodenreider O, Burgun A. Looking for Anemia (and Other Disorders) in SNOMED CT: Comparison of Three Approaches and Practical Implications. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2010; 2010: 527-31	CT	Pre-development	Design considerations	Internal medicine	France	2010
232.	Nachimuthu SK, Lau LM. Practical issues in using SNOMED CT as a reference terminology. Studies in health technology and informatics 2007; 129(Pt 1): 640-4	CT	Pre-development	Design considerations	Other/multiple/unknown/not applicable	United States	2007
233.	Nachimuthu SK, Woolstenhulme RD. Generalizability of hybrid search algorithms to map multiple biomedical vocabulary domains. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2006; : 1042	CT	Pre-development	Design considerations	Other/multiple/unknown/not applicable	United States	2006
234.	Nadkarni PM, Marenco LA. Implementing description-logic rules for SNOMED-CT attributes through a table-driven approach. Journal of the American Medical Informatics Association : JAMIA 2010 Mar-Apr; 17(2): 182-4	CT	Pre-development	Design considerations	Other/multiple/unknown/not applicable	United States	2010
235.	Navas H, Lopez Osornio A, Gambarte L, Elías Leguizamón G, Wasserman S, Orrego N, Luna D, de Quirós FG. Implementing rules to improve the quality of concept post-coordination with SNOMED CT. Studies in health technology and informatics 2010; 160(Pt 2): 1045-9	CT	Pre-development	Design considerations	Other/multiple/unknown/not applicable	Argentina	2010
236.	Patel CO, Cimino JJ. Mining cross-terminology links in the UMLS. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2006; : 624-8	CT	Pre-development	Design considerations	Other/multiple/unknown/not applicable	United States	2006
237.	Pathak J, Jiang G, Dwarkanath SO, Buntrock JD, Chute CG. Adopting Graph Traversal Techniques for Context-Driven Value Sets Extraction from Biomedical Knowledge Sources. Proceedings / IEEE International Conference on Semantic Computing ... ICSC. IEEE International Conference on Semantic Computing 2008 Aug 12; 2008: 460-467	CT	Pre-development	Design considerations	Other/multiple/unknown/not applicable	United States	2008

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
238.	Pathak J, Jiang G, Dwarkanath SO, Buntrock JD, Chute CG, Chute C. LexValueSets: an approach for context-driven value sets extraction. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2008; : 556-60	CT	Pre-development	Design considerations	Other/multiple/unknown/not applicable	United Kingdom	2008
239.	Patrick J, Li M. High accuracy information extraction of medication information from clinical notes: 2009 i2b2 medication extraction challenge. Journal of the American Medical Informatics Association : JAMIA 2010 Sep-Oct; 17(5): 524-7	CT	Pre-development	Design considerations	Other/multiple/unknown/not applicable	Australia	2010
240.	Randorff Højen A, Rosenbeck Gøeg K. SNOMED CT Implementation. Mapping Guidelines Facilitating Reuse of Data. Methods of information in medicine 2012 Oct 1;51(6)	CT	Pre-development	Design considerations	Other/multiple/unknown/not applicable	Denmark	2012
241.	Randorff Rasmussen A, Markussen D. Methods for testing SNOMED CT in a defined clinical domain. Studies in health technology and informatics 2009; 150: 248	CT	Pre-development	Design considerations	Pathology	Denmark	2009
242.	Rasmussen AR, Rosenbeck K. SNOMED CT implementation: implications of choosing clinical findings or observable entities. Studies in health technology and informatics 2011; 169: 809-13	CT	Pre-development	Design considerations	SNOMED CT Hierarchy	Denmark	2011
243.	Rojo M.G. SNOMED CT in Europe. Virchows Archiv 2010; 457: 140-	CT	Pre-development	Design considerations	Pathology	Spain	2010
244.	Rosenbloom ST, Miller RA, Johnson KB, Elkin PL, Brown SH. Interface terminologies: facilitating direct entry of clinical data into electronic health record systems. Journal of the American Medical Informatics Association : JAMIA 2006 May-Jun; 13(3): 277-88	CT	Pre-development	Design considerations	Other/multiple/unknown/not applicable	United States	2006
245.	Sevenster M, Aleksovski Z. SNOMED CT Saves Keystrokes: Quantifying Semantic Autocompletion. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2010; 2010: 742-6	CT	Pre-development	Design considerations	Other/multiple/unknown/not applicable	Netherlands	2010
246.	Sevenster M, van Ommering R, Qian Y. Algorithmic and user study of an autocompletion algorithm on a large medical vocabulary. Journal of biomedical informatics 2011 Oct 11;	CT	Pre-development	Design considerations	Other/multiple/unknown/not applicable	Netherlands	2011

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
247.	Stenzhorn H, Pacheco EJ, Nohama P, Schulz S. Automatic mapping of clinical documentation to SNOMED CT. Studies in health technology and informatics 2009; 150: 228-32	CT	Pre-development	Design considerations	Cardiology	Germany	2009
248.	Wade G, Rosenbloom ST. The impact of SNOMED CT revisions on a mapped interface terminology: terminology development and implementation issues. Journal of biomedical informatics 2009 Jun; 42(3): 490-3	CT	Pre-development	Design considerations	Other/multiple/unknown/not applicable	United States	2009
249.	Wei WQ, Tao C, Jiang G, Chute CG. A high throughput semantic concept frequency based approach for patient identification: a case study using type 2 diabetes mellitus clinical notes. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2010; 2010: 857-61	CT	Pre-development	Design considerations	Other/multiple/unknown/not applicable	United States	2010
250.	Wright A, Feblowitz J, McCoy AB, Sittig DF. Comparative analysis of the VA/Kaiser and NLM CORE problem subsets: an empirical study based on problem frequency. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2011;2011:1532-40	CT	Pre-development	Design considerations	Problem list / diagnoses	United States	2011
251.	Ahmadian L, Cornet R, de Keizer NF. Facilitating pre-operative assessment guidelines representation using SNOMED CT. Journal of biomedical informatics 2010 Dec; 43(6): 883-90	CT	Pre-development	Prospective content coverage	Preoperative	Netherlands	2010
252.	Ahmadian L, De Keizer NF, Cornet R. The use of SNOMED CT for representing concepts used in preoperative guidelines. Studies in health technology and informatics 2009; 150: 658-62	CT	Pre-development	Prospective content coverage	Preoperative	Netherlands	2009
253.	Bodenreider O, Fushman DD. Investigating drug classes in biomedical terminologies from the perspective of clinical decision support. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2010; 2010: 56-60	CT	Pre-development	Prospective content coverage	Drugs	United States	2010
254.	Brown SH, Bauer BA, Wahner-Roedler DL, Elkin PL. Coverage of oncology drug indication concepts and compositional semantics by SNOMED-CT. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2003; : 115-9	CT	Pre-development	Prospective content coverage	Drugs	United States	2003

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
255.	Brown SH, Elkin PL, Bauer BA, Wahner-Roedler D, Husser CS, Temesgen Z, Hardenbrook SP, Fielstein EM, Rosenbloom ST. SNOMED CT: utility for a general medical evaluation template. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2006; : 101-5	CT	Pre-development	Prospective content coverage	Hospital	United States	2006
256.	Brown SH, Rosenbloom ST, Bauer BA, Wahner-Roedler D, Froehling DA, Bailey KR, Lincoln MJ, Montella D, Fielstein EM, Elkin PL. Direct comparison of MEDCIN and SNOMED CT for representation of a general medical evaluation template. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2007; : 75-9	CT	Pre-development	Prospective content coverage	Hospital	United States	2007
257.	Campbell JR, Xu J, Fung KW. Can SNOMED CT fulfill the vision of a compositional terminology? Analyzing the use case for problem list. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2011;2011:181-8	CT	Pre-development	Prospective content coverage	Problem list / diagnoses	United States	2011
258.	Chiang MF, Casper DS, Cimino JJ, Starren J. Representation of ophthalmology concepts by electronic systems: adequacy of controlled medical terminologies. Ophthalmology 2005 Feb; 112(2): 175-83	CT	Pre-development	Prospective content coverage	Ophthalmology	United States	2005
259.	Daniel-Le Bozec C, Steichen O, Dart T, Jaulent MC. The role of local terminologies in electronic health records. The HEGP experience. Studies in health technology and informatics 2007; 129(Pt 1): 780-4	CT	Pre-development	Prospective content coverage	Other/multiple/unknown/not applicable	France	2007
260.	De Silva TS, MacDonald D, Paterson G, Sikdar KC, Cochrane B. Systematized nomenclature of medicine clinical terms (SNOMED CT) to represent computed tomography procedures. Computer methods and programs in biomedicine 2011 Mar; 101(3): 324-9	CT	Pre-development	Prospective content coverage	Radiology	Canada	2011
261.	Deitzer JR, Payne PR, Starren JB. Coverage of clinical trials tasks in existing ontologies. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2006; : 903	CT	Pre-development	Prospective content coverage	Cancer	United States	2006

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
262.	Dykes PC, Currie LM, Cimino JJ. Adequacy of evolving national standardized terminologies for interdisciplinary coded concepts in an automated clinical pathway. Journal of biomedical informatics 2003 Aug-Oct; 36(4-5): 313-25	CT	Pre-development	Prospective content coverage	Cardiology	United States	2003
263.	Elkin PL, Brown SH, Balas A, Temesgen Z, Wahner-Roedler D, Froehling D, Liebow M, Trusko B, Rosenbloom ST, Poland G .Biosurveillance evaluation of SNOMED CT's terminology (BEST Trial): coverage of chief complaints. Studies in health technology and informatics 2008; 136: 797-802	CT	Pre-development	Prospective content coverage	Problem list / diagnoses	United States	2008
264.	Elkin PL, Brown SH, Husser CS, Bauer BA, Wahner-Roedler D, Rosenbloom ST, Speroff T. Evaluation of the content coverage of SNOMED CT: ability of SNOMED clinical terms to represent clinical problem lists. Mayo Clinic proceedings. Mayo Clinic 2006 Jun; 81(6): 741-8	CT	Pre-development	Prospective content coverage	Problem list / diagnoses	United States	2006
265.	Elkin PL, Trusko BE, Koppel R, Speroff T, Mohrer D, Sakji S, Gurewitz I, Tuttle M, Brown SH. Secondary use of clinical data. Studies in health technology and informatics 2010; 155: 14-29	CT	Pre-development	Prospective content coverage	Problem list / diagnoses	United States	2010
266.	Fung KW, McDonald C, Srinivasan S. The UMLS-CORE project: a study of the problem list terminologies used in large healthcare institutions. Journal of the American Medical Informatics Association : JAMIA 2010 Nov-Dec; 17(6): 675-80	CT	Pre-development	Prospective content coverage	Problem list / diagnoses	United States	2010
267.	Hrabak KM, Campbell JR, Tu SW, McClure R, Weida RT. Creating interoperable guidelines: requirements of vocabulary standards in immunization decision support. Studies in health technology and informatics 2007; 129(Pt 2): 930-4	CT	Pre-development	Prospective content coverage	Primary care	United States	2007
268.	James AG, Ng E, Shah PS. A reference set of SNOMED terms for the structured representation of respiratory disorders of the newborn infant. Studies in health technology and informatics 2009; 150: 243-7	CT	Pre-development	Prospective content coverage	Newborn	Canada	2009
269.	James AG, Spackman KA. Representation of disorders of the newborn infant by SNOMED CT. Studies in health technology and informatics 2008; 136: 833-8	CT	Pre-development	Prospective content coverage	Newborn	United States	2008

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
270.	James AG, Spackman KA. SNOMED CT as the clinical terminology for the representation of the clinical care of the newborn infant. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2007; : 989	CT	Pre-development	Prospective content coverage	Newborn	United States	2007
271.	Kim H, Harris MR, Savova G, Chute CG. Content coverage of SNOMED-CT toward the ICU nursing flowsheets and the acuity indicators. Studies in health technology and informatics 2006; 122: 722-6	CT	Pre-development	Prospective content coverage	Intensive care	United States	2006
272.	Kim HY, Cho IS, Lee JH, Kim JH, Sim DH, Kim Y. Matching between the concepts of knowledge representation for a hypertension guideline and SNOMED CT. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2008; : 1005	CT	Pre-development	Prospective content coverage	Other/multiple/unknown/not applicable	South Korea	2008
273.	Kokkinakis D. What is the coverage of SNOMED CT®on scientific medical corpora? Studies in health technology and informatics 2011; 169: 814-8	CT	Pre-development	Prospective content coverage	Other/multiple/unknown/not applicable	Sweden	2011
274.	Lathrop SL, Davis WL, Nolte KB. Medical terminology coding systems and medicolegal death investigation data: searching for a standardized method of electronic coding at a statewide medical examiner's office. Journal of forensic sciences 2009 Jan; 54(1): 207-11	CT	Pre-development	Prospective content coverage	Other/multiple/unknown/not applicable	United States	2009
275.	Lee S, Tsirbas A, Goldberg RA, McCann JD. Standardized terminology for aesthetic ophthalmic plastic surgery. Ophthalmic plastic and reconstructive surgery 2006 Sep-Oct; 22(5): 371-4	CT	Pre-development	Prospective content coverage	Anaesthesia	United States	2006
276.	Liu H, Burkhart Q, Bell DS. Evaluation of the NCPDP Structured and Codified Sig Format for e-prescriptions. Journal of the American Medical Informatics Association : JAMIA 2011 Sep-Oct;18(5):645-51	CT	Pre-development	Prospective content coverage	Drugs	United States	2011
277.	Liu H, Wagholikar K, Wu ST. Using SNOMED-CT to encode summary level data - a corpus analysis. AMIA Summits on Translational Science proceedings AMIA Summit on Translational Science 2012;2012:30-7	CT	Pre-development	Prospective content coverage	Problem list / diagnoses	United States	2012

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
278.	Mantena S, Schadow G. Evaluation of the VA/KP problem list subset of SNOMED as a clinical terminology for electronic prescription clinical decision support. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2007; : 498-502	CT	Pre-development	Prospective content coverage	Problem list / diagnoses	United States	2007
279.	Matney SA, Warren JJ, Evans JL, Kim TY, Coenen A, Auld VA. Development of the nursing problem list subset of SNOMED CT® Journal of biomedical informatics 2011 Dec 20;	CT	Pre-development	Prospective content coverage	Nursing	United States	2011
280.	McClay JC, Campbell J. Improved coding of the primary reason for visit to the emergency department using SNOMED. Proceedings / AMIA ... Annual Symposium. AMIA Symposium 2002; : 499-503	Multiple	Pre-development	Prospective content coverage	Emergency department/room	United States	2002
281.	McCormick M., Windle T.A., Windle J.R. A new schema for problem lists: Towards real-time decision support. Journal of the American College of Cardiology 2010; 55: A132-	CT	Pre-development	Prospective content coverage	Hospital	United States	2010
282.	Min Z, Baofen D, Weeber M, van Ginneken AM. Mapping OpenSDE domain models to SNOMED CT. Applied to the domain of cardiovascular disease. Methods of information in medicine 2006; 45(1): 4-9	CT	Pre-development	Prospective content coverage	Cardiology	Netherlands	2006
283.	Osornio AL, Luna D, Gambarte ML, Gomez A, Reynoso G, de Quirós FG. Creation of a local interface terminology to SNOMED CT. Studies in health technology and informatics 2007; 129(Pt 1): 765-9	CT	Pre-development	Prospective content coverage	Other/multiple/unknown/not applicable	Argentina	2007
284.	Pathak J, Wang J, Kashyap S, Basford M, Li R, Masys DR, Chute CG. Mapping clinical phenotype data elements to standardized metadata repositories and controlled terminologies: the eMERGE Network experience. Journal of the American Medical Informatics Association : JAMIA 2011 Jul-Aug; 18(4): 376-86	CT	Pre-development	Prospective content coverage	Genetics	United States	2011
285.	Penz JF, Brown SH, Carter JS, Elkin PL, Nguyen VN, Sims SA, Lincoln MJ. Evaluation of SNOMED coverage of Veterans Health Administration terms. Studies in health technology and informatics 2004; 107(Pt 1): 540-4	CT	Pre-development	Prospective content coverage	Other/multiple/unknown/not applicable	United States	2004

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
286.	Přečková P, Zvárová J, Zvára K. Measuring diversity in medical reports based on categorized attributes and international classification systems. BMC medical informatics and decision making 2012 Apr 12;12:31	CT	Pre-development	Prospective content coverage	Cardiology	Czech Republic	2012
287.	Ravvaz K, Senk P, Patrick TB, Coenen A, Kim TY, Zhao H, Gaudioso C, Jansen KR, Lang NM. Mapping nursing concepts to ontologies for evidence-based nursing. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2008; : 1105	CT	Pre-development	Prospective content coverage	Nursing	United States	2008
288.	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. Journal of the American Medical Informatics Association : JAMIA 2006 Sep-Oct; 13(5): 536-46	CT	Pre-development	Prospective content coverage	Rheumatology	United States	2006
289.	Rollason W, Khunti K, de Lusignan S. Variation in the recording of diabetes diagnostic data in primary care computer systems: implications for the quality of care. Informatics in primary care 2009; 17(2): 113-9	CT	Pre-development	Prospective content coverage	Chronic care	United Kingdom	2009
290.	Sampalli T, Shepherd M, Duffy J. Clinical vocabulary as a boundary object in multidisciplinary care management of multiple chemical sensitivity, a complex and chronic condition. Journal of multidisciplinary healthcare 2011; 4: 91-102	CT	Pre-development	Prospective content coverage	Chronic care	Canada	2011
291.	Sampalli T, Shepherd M, Duffy J, Fox R. An evaluation of SNOMED CT in the domain of complex chronic conditions. International journal of integrated care 2010; 10: e038	CT	Pre-development	Prospective content coverage	Chronic care	Canada	2010
292.	Sarkar IN. Leveraging biomedical ontologies and annotation services to organize microbiome data from Mammalian hosts. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2010; 2010: 717-21	CT	Pre-development	Prospective content coverage	Genetics	United States	2010
293.	Shah N., Pittelkow M. Assessment of dermatology-domain specific term coverage in ICD-9 and SNOMED-CT using a dermatology textbook index. Journal of Investigative Dermatology 2009; 129: S60-	CT	Pre-development	Prospective content coverage	Dermatology	Unknown	2009

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
294.	Shah NH, Rubin DL, Supekar KS, Musen MA. Ontology-based annotation and query of tissue microarray data. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2006; : 709-13	CT	Pre-development	Prospective content coverage	Pathology	United States	2006
295.	Smith CA, Wicks PJ. PatientsLikeMe: Consumer health vocabulary as a folksonomy. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2008; : 682-6	CT	Pre-development	Prospective content coverage	Problem list / diagnoses	United States	2008
296.	So EY, Park HA. Mapping medical records of gastrectomy patients to SNOMED CT. Studies in health technology and informatics 2011; 169: 764-8	CT	Pre-development	Prospective content coverage	Gastroenterology	South Korea	2011
297.	Thomas N, Minard JP, Lougheed M. Primary care asthma program's asthma care map: Use of SNOMED-CT and LOINC to standardize terminology for electronic health records Am. J. Respir. Crit. Care Med. 2010; 181(1 MeetingAbstracts): -	CT	Pre-development	Prospective content coverage	Primary care	Canada	2010
298.	Travers DA, Haas SW. Unified medical language system coverage of emergency-medicine chief complaints. Academic emergency medicine : official journal of the Society for Academic Emergency Medicine 2006 Dec; 13(12): 1319-23	CT	Pre-development	Prospective content coverage	Emergency department/roo m	United States	2006
299.	Trusko B, Rosenbloom ST, Montella D, Jackson JC, Fitzhenry F, Brown SH, Elkin PL, Fielstein E, Kotter K, Tuttle M, Iannelli RJ, Speroff T. Are posttraumatic stress disorder mental health terms found in SNOMED-CT medical terminology. Journal of traumatic stress 2010 Dec; 23(6): 794-801	CT	Pre-development	Prospective content coverage	Mental	United States	2010
300.	van der Kooij J, Goossen WT, Goossen-Baremans AT, de Jong-Fintelman M, van Beek L. Using SNOMED CT codes for coding information in electronic health records for stroke patients. Studies in health technology and informatics 2006; 124: 815-23	CT	Pre-development	Prospective content coverage	Other/multiple/u nknown/not applicable	Netherlands	2006
301.	Vázquez-Naya JM, Martínez-Romero M, Porto-Pazos AB, Novoa F, Valladares-Ayerbes M, Pereira J, Munteanu CR, Dorado J. Ontologies of drug discovery and design for neurology, cardiology and oncology. Current pharmaceutical design 2010; 16(24): 2724-36	CT	Pre-development	Prospective content coverage	Drugs	Spain	2010

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
302.	Vikström A, Hägglund M, Nyström M, Strender LE, Koch S, Hjerpe P, Lindblad U, Nilsson GH. Coding of procedures documented by general practitioners in Swedish primary care-an explorative study using two procedure coding systems. BMC family practice 2012 Jan 9;13:2	CT	Pre-development	Prospective content coverage	Primary care	Sweden	2012
303.	Vikström A, Nilsson G. SNOMED CT in multidisciplinary clinical practice--evaluation of usefulness for classification and coding of care-planning procedures. Studies in health technology and informatics 2006; 124: 824-9	CT	Pre-development	Prospective content coverage	Other/multiple/unknown/not applicable	Sweden	2006
304.	Wade G, Gotlieb EM, Weigle C, Warren R. Assessing voids in SNOMED CT for pediatric concepts. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2008; : 1164	CT	Pre-development	Prospective content coverage	Paediatrics	United States	2008
305.	Wade G, Rosenbloom ST. Experiences mapping a legacy interface terminology to SNOMED CT. BMC medical informatics and decision making 2008; 8 Suppl 1: S3	CT	Pre-development	Prospective content coverage	Other/multiple/unknown/not applicable	United States	2008
306.	Warren JJ, Wilson RP. Representing cardiovascular concepts in an electronic health record using SNOMED CT. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2006; : 1135	CT	Pre-development	Prospective content coverage	Cardiology	United States	2006
307.	Wasserman H, Wang J. An applied evaluation of SNOMED CT as a clinical vocabulary for the computerized diagnosis and problem list. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2003; : 699-703	CT	Pre-development	Prospective content coverage	Problem list / diagnoses	United States	2003
308.	Wyke S, Orford R, Duarte-Davidson R, Pelclova D, Edwards N, Kennedy K, Sutton N, Good AM, Desel H, Schaper A, Bronstein A, Dragelyte G, Mathieu-Nolf M, Kupferschmidt H. The evaluation of standard medical terminology systems to describe symptoms of poisoning, An output of the ASHTII project Clin. Toxicol. 2010; 48(3): 271	CT	Pre-development	Prospective content coverage	Other/multiple/unknown/not applicable	Multiple countries	2010
309.	Yu A, Ballow M, Shvarts A, Lehman H. Evaluation of coverage of allergy concepts in electronic health records J. Allergy Clin. Immunol. 2012; 129(2 SUPPL. 1): AB108	CT	Pre-development	Prospective content coverage	Allergies	United States	2012

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
310.	Andrews JE, Patrick TB, Richesson RL, Brown H, Krischer JP. Comparing heterogeneous SNOMED CT coding of clinical research concepts by examining normalized expressions. Journal of biomedical informatics 2008 Dec; 41(6): 1062-9	CT	Pre-development	Prospective inter-rater agreement	Other/multiple/unknown/not applicable	United States	2008
311.	Andrews JE, Richesson RL, Krischer J. Variation of SNOMED CT coding of clinical research concepts among coding experts. Journal of the American Medical Informatics Association : JAMIA 2007 Jul-Aug; 14(4): 497-506	CT	Pre-development	Prospective inter-rater agreement	Pathology	United States	2007
312.	Chiang MF, Hwang JC, Yu AC, Casper DS, Cimino JJ, Starren JB. Reliability of SNOMED-CT coding by three physicians using two terminology browsers. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2006; : 131-5	CT	Pre-development	Prospective inter-rater agreement	Ophthalmology	United States	2006
313.	Hwang JC, Yu AC, Casper DS, Starren J, Cimino JJ, Chiang MF. Representation of Ophthalmology Concepts by Electronic Systems. Intercoder Agreement among Physicians Using Controlled Terminologies Ophthalmology 2006; 113(4): 511-519	CT	Pre-development	Prospective inter-rater agreement	Ophthalmology	United States	2006
314.	Hwang JC, Yu AC, Casper DS, Starren J, Cimino JJ, Chiang MF. Representation of ophthalmology concepts by electronic systems: intercoder agreement among physicians using controlled terminologies. Ophthalmology 2006 Apr; 113(4): 511-9	CT	Pre-development	Prospective inter-rater agreement	Ophthalmology	United States	2006
315.	Kim SY, Kim HH, Shin KH, Kim HS, Lee JI, Choi BK. Comparison of Knowledge Levels Required for SNOMED CT Coding of Diagnosis and Operation Names in Clinical Records. Healthcare informatics research 2012 Sep;18(3):186-90	CT	Pre-development	Prospective inter-rater agreement	Other/multiple/unknown/not applicable	South Korea	2012
316.	Patrick TB, Richesson R, Andrews JE, Folk LC. SNOMED CT coding variation and grouping for "other findings" in a longitudinal study on urea cycle disorders. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2008; : 11-5	CT	Pre-development	Prospective inter-rater agreement	Internal medicine	United States	2008

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
317.	Zimmerman KL, Wilcke JR, Robertson JL, Feldman BF, Kaur T, Rees LR, Spackman KA SNOMED representation of explanatory knowledge in veterinary clinical pathology. Veterinary clinical pathology / American Society for Veterinary Clinical Pathology 2005; 34(1): 7-16	CT	Pre-development	Prospective inter-rater agreement	Pathology	United States	2005
318.	Abhyankar S, Lloyd-Puryear MA, Goodwin R, Copeland S, Eichwald J, Therrell BL, Zuckerman A, Downing G, McDonald CJ. Standardizing newborn screening results for health information exchange. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2010; 2010: 1-5	CT	Pre-development	Standard for EHR	Newborn	United States	2010
319.	Bernstein K, Andersen U. Managing care pathways combining SNOMED CT, archetypes and an electronic guideline system. Studies in health technology and informatics 2008; 136: 353-8	CT	Pre-development	Standard for EHR	Hospital	Denmark	2008
320.	Bernstein K, Bruun-Rasmussen M, Vingtoft S. A method for specification of structured clinical content in electronic health records. Studies in health technology and informatics 2006; 124: 515-21	CT	Pre-development	Standard for EHR	Cardiology	Denmark	2006
321.	Bird L, Brooks C, Cheong YC, Tun NN. A logical approach to semantic interoperability in healthcare. Studies in health technology and informatics 2011; 168: 1-9	CT	Pre-development	Standard for EHR	Other/multiple/unknown/not applicable	Singapore	2011
322.	Bouhaddou O, Warnekar P, Parrish F, Do N, Mandel J, Kilbourne J, Lincoln MJ. Exchange of computable patient data between the Department of Veterans Affairs (VA) and the Department of Defense (DoD): terminology mediation strategy. Journal of the American Medical Informatics Association : JAMIA 2008 Mar-Apr; 15(2): 174-83	CT	Pre-development	Standard for EHR	Other/multiple/unknown/not applicable	United States	2008
323.	Bozkurt S, Kahn CE Jr. An open-standards grammar for outline-style radiology report templates. Journal of digital imaging : the official journal of the Society for Computer Applications in Radiology 2012 Jun;25(3):359-64	CT	Pre-development	Standard for EHR	Radiology	Multiple countries	2012

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
324.	Bucur A., Fumagalli D., Desmedt C., Vdovjak R., Loi S., Saini K., Doci S.M., Schenk B., Sotiriou C., Piccart M.J. Integrate: Driving excellence in integrative cancer research through innovative biomedical infrastructures. European Journal of Cancer 2011; 47: S215-S216	CT	Pre-development	Standard for EHR	Other/multiple/unknown/not applicable	Multiple countries	2011
325.	Cheong YC, Bird L, Tun NN, Brooks C. Using a logical information model-driven design process in healthcare. Studies in health technology and informatics 2011; 169: 804-8	CT	Pre-development	Standard for EHR	Other/multiple/unknown/not applicable	Singapore	2011
326.	Daniel C, Booker D, Beckwith B, Della Mea V, García-Rojo M, Havener L, Kennedy M, Klossa J, Laurinavicius A, Macary F, Punys V, Scharber W, Schrader T. Standards and specifications in pathology: image management, report management and terminology. Studies in health technology and informatics 2012;179:105-22	CT	Pre-development	Standard for EHR	Pathology	Multiple countries	2012
327.	Daniel C, Macary F, Rojo MG, Klossa J, Laurinavičius A, Beckwith BA, Della Mea V. Recent advances in standards for Collaborative Digital Anatomic Pathology. Diagnostic pathology 2011; 6 Suppl 1: S17	CT	Pre-development	Standard for EHR	Pathology	France	2011
328.	Dolin RH, Mattison JE, Cohn S, Campbell KE, Wiesenthal AM, Hochhalter B, LaBerge D, Barsoum R, Shalaby J, Abilla A, Clements RJ, Correia CM, Esteva D, Fedack JM, Goldberg BJ, Gopalarao S, Hafeza E, Hendler P, Hernandez E, Kamangar R, Kahn RA, Kurtovich G, Lazzareschi G, Lee MH, Lee T, Levy D, Lukoff JY, Lundberg C, Madden MP, Ngo TL, Nguyen BT, Patel NP, Resneck J, Ross DE, Schwarz KM, Selhorst CC, Snyder A, Umarji MI, Vilner M, Zer-Chen R, Zingo C. Kaiser Permanente's Convergent Medical Terminology. Studies in health technology and informatics 2004; 107(Pt 1): 346-50	CT	Pre-development	Standard for EHR	Other/multiple/unknown/not applicable	United States	2004
329.	Fidahussein M, Friedlin J, Grannis S. Practical challenges in the secondary use of real-world data: the notifiable condition detector. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2011;2011:402-8	CT	Pre-development	Standard for EHR	Laboratory	United States	2011

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
330.	Garde S, Hovenga E, Buck J, Knaup P. Expressing clinical data sets with openEHR archetypes: a solid basis for ubiquitous computing. International journal of medical informatics 2007 Dec; 76 Suppl 3: S334-41	CT	Pre-development	Standard for EHR	Paediatrics	Germany	2007
331.	Hamm RA, Knoop SE, Schwarz P, Block AD, Davis WL. Harmonizing clinical terminologies: driving interoperability in healthcare. Studies in health technology and informatics 2007; 129(Pt 1): 660-3	CT	Pre-development	Standard for EHR	Other/multiple/unknown/not applicable	United States	2007
332.	Heymans S, McKennirey M, Phillips J. Semantic validation of the use of SNOMED CT in HL7 clinical documents. Journal of biomedical semantics 2011; 2(1): 2	CT	Pre-development	Standard for EHR	Other/multiple/unknown/not applicable	United States	2011
333.	Hovenga E, Garde S, Heard S. Nursing constraint models for electronic health records: a vision for domain knowledge governance. International journal of medical informatics 2005 Dec; 74(11-12): 886-98	CT	Pre-development	Standard for EHR	Nursing	Australia	2005
334.	Hunscher D, Boyd A, Green LA, Clauw DJ. Representing natural-language case report form terminology using Health Level 7 Common Document Architecture, LOINC, and SNOMED-CT: lessons learned. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2006; : 961	CT	Pre-development	Standard for EHR	Other/multiple/unknown/not applicable	United States	2006
335.	Hurrell M.J., Monk T.G., Nicol A., Norton A.N., Reich D.L., Walsh J.L. Implementation of a standards-based, CDA-compliant anesthesia record. Journal of Clinical Monitoring and Computing 2011; 25: 15-16	CT	Pre-development	Standard for EHR	Anaesthesia	Multiple countries	2011
336.	Hwang KH, Chung KI, Chung MA, Choi D. Review of semantically interoperable electronic health records for ubiquitous healthcare. Healthcare informatics research 2010 Mar; 16(1): 1-5	CT	Pre-development	Standard for EHR	Other/multiple/unknown/not applicable	South Korea	2010
337.	Kostick K. SNOMED CT integral part of quality EHR documentation. Journal of AHIMA / American Health Information Management Association 2012 Oct;83(10):72-5	CT	Pre-development	Standard for EHR	Other/multiple/unknown/not applicable	United States	2012
338.	Lezcano L, Sicilia MA, Rodríguez-Solano C. Integrating reasoning and clinical archetypes using OWL ontologies and SWRL rules. Journal of biomedical informatics 2011 Apr; 44(2): 343-53	CT	Pre-development	Standard for EHR	Anaesthesia	Spain	2011

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
339.	Liu S, Ni Y, Mei J, Li H, Xie G, Hu G, Liu H, Hou X, Pan Y. iSMART: Ontology-based Semantic query of CDA documents. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2009; 2009: 375-9	CT	Pre-development	Standard for EHR	Other/multiple/unknown/not applicable	China	2009
340.	Lougheed MD, Minard J, Dworkin S, Juurlink MA, Temple WJ, To T, Koehn M, Van Dam A, Boulet LP. Pan-Canadian REspiratory STandards INitiative for Electronic Health Records (PRESTINE): 2011 national forum proceedings. Canadian respiratory journal : journal of the Canadian Thoracic Society 2012 Mar-Apr;19(2):117-26	CT	Pre-development	Standard for EHR	Other/multiple/unknown/not applicable	Canada	2012
341.	Meizoso García M, Iglesias Allones JL, Martínez Hernández D, Taboada Iglesias MJ. Semantic similarity-based alignment between clinical archetypes and SNOMED CT: an application to observations. International journal of medical informatics 2012 Aug;81(8):566-78	CT	Pre-development	Standard for EHR	Other/multiple/unknown/not applicable	Spain	2012
342.	Mollura DJ, Carrino JA, Matuszak DL, Mnatsakanyan ZR, Eng J, Cutchis P, Babin SM, Sniegoski C, Lombardo JS. Bridging radiology and public health: the emerging field of radiologic public health informatics. Journal of the American College of Radiology : JACR 2008 Mar; 5(3): 174-81	CT	Pre-development	Standard for EHR	Radiology	United States	2008
343.	Moner D, Maldonado JA, Angulo C, Bosca D, Perez D, Abad I, Reig E, Robles M. Standardization of discharge reports with the ISO 13606 norm. Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference 2008; 2008: 1470-3	CT	Pre-development	Standard for EHR	Hospital	Spain	2008
344.	Nagy M, Hanzlícek P, Precková P, Ríha A, Dioszegi M, Seidl L, Zvárová J. Semantic interoperability in Czech healthcare environment supported by HL7 version 3. Methods of information in medicine 2010; 49(2): 186-95	CT	Pre-development	Standard for EHR	Other/multiple/unknown/not applicable	Czech Republic	2010
345.	Ouagne D, Hussain S, Sadou E, Jaulent MC, Daniel C. The Electronic Healthcare Record for Clinical Research (EHR4CR) information model and terminology. Studies in health technology and informatics 2012;180:534-8	CT	Pre-development	Standard for EHR	Other/multiple/unknown/not applicable	France	2012

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
346.	Pantazatos SP, Li J, Pavlidis P, Lussier YA. Integration of Neuroimaging and Microarray Datasets through Mapping and Model-Theoretic Semantic Decomposition of Unstructured Phenotypes. Cancer informatics 2009 Jun 8; 8: 75-94	CT	Pre-development	Standard for EHR	Pathology	United States	2009
347.	Patil NG, Doll C, Ghatge P, Konschuh S, Temple W, Craighead P. Web-based synoptic reporting for gynecologic radiation oncologists: An alberta initiative Radiother. Oncol. 2010; 96: S52	CT	Pre-development	Standard for EHR	Other/multiple/unknown/not applicable	Canada	2010
348.	Qamar R, Kola J, Rector AL. Unambiguous data modeling to ensure higher accuracy term binding to clinical terminologies. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2007; : 608-13	CT	Pre-development	Standard for EHR	Other/multiple/unknown/not applicable	United Kingdom	2007
349.	Qamar R, Rector A. Semantic issues in integrating data from different models to achieve data interoperability. Studies in health technology and informatics 2007; 129(Pt 1): 674-8	CT	Pre-development	Standard for EHR	Other/multiple/unknown/not applicable	United States	2007
350.	Rajeev D, Staes CJ, Evans RS, Mottice S, Rolfs R, Samore MH, Whitney J, Kurzban R, Huff SM. Development of an electronic public health case report using HL7 v2.5 to meet public health needs. Journal of the American Medical Informatics Association : JAMIA 2010 Jan-Feb; 17(1): 34-41	CT	Pre-development	Standard for EHR	Other/multiple/unknown/not applicable	United States	2010
351.	Ryan A, Eklund P A. framework for semantic interoperability in healthcare: a service oriented architecture based on health informatics standards. Studies in health technology and informatics 2008; 136: 759-64	CT	Pre-development	Standard for EHR	Hospital	Australia	2008
352.	Ryan A, Eklund P. The health service bus: An architecture and case study in achieving interoperability in healthcare Stud. Health Technol. Informatics 2010; 160(PART 1): 922-926	CT	Pre-development	Standard for EHR	Other/multiple/unknown/not applicable	Australia	2010
353.	Ryan A, Eklund P, Esler B. Toward the interoperability of HL7 v3 and SNOMED CT: a case study modeling mobile clinical treatment. Studies in health technology and informatics 2007; 129(Pt 1): 626-30	CT	Pre-development	Standard for EHR	Other/multiple/unknown/not applicable	Australia	2007

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
354.	Sundvall E, Qamar R, Nyström M, Forss M, Petersson H, Karlsson D, Ahlfeldt H, Rector A. Integration of tools for binding archetypes to SNOMED CT. BMC medical informatics and decision making 2008; 8 Suppl 1: S7	CT	Pre-development	Standard for EHR	Other/multiple/unknown/not applicable	Sweden	2008
355.	Wurtz R, Cameron BJ. Electronic laboratory reporting for the infectious diseases physician and clinical microbiologist. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America 2005 Jun 1; 40(11): 1638-43	Multiple	Pre-development	Standard for EHR	Laboratory	United States	2005
356.	Xiao L, Cousins G, Courtney B, Hederman L, Fahey T, Dimitrov BD. Developing an electronic health record (EHR) for methadone treatment recording and decision support. BMC medical informatics and decision making 2011; 11: 5	CT	Pre-development	Standard for EHR	Drugs	Ireland	2011
357.	Yu S, Berry D, Bisbal J. Clinical coverage of an archetype repository over SNOMED-CT. Journal of biomedical informatics 2012 Jun;45(3):408-18	CT	Pre-development	Standard for EHR	Other/multiple/unknown/not applicable	Ireland	2011
358.	Abdoune H, Merabti T, Darmoni SJ, Joubert M. Assisting the translation of the CORE subset of SNOMED CT into French. Studies in health technology and informatics 2011; 169: 819-23	CT	Pre-development	Translation	Problem list / diagnoses	France	2011
359.	Deléger L, Merkel M, Zweigenbaum P. Contribution to terminology internationalization by word alignment in parallel corpora. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2006; : 185-9	CT	Pre-development	Translation	Other/multiple/unknown/not applicable	France	2006
360.	Deleger L, Merkel M, Zweigenbaum P. Enriching medical terminologies: an approach based on aligned corpora. Studies in health technology and informatics 2006; 124: 747-52	CT	Pre-development	Translation	Other/multiple/unknown/not applicable	France	2006
361.	Deléger L, Merkel M, Zweigenbaum P. Translating medical terminologies through word alignment in parallel text corpora. Journal of biomedical informatics 2009 Aug; 42(4): 692-701	CT	Pre-development	Translation	Other/multiple/unknown/not applicable	France	2009

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
362.	Joubert M, Abdoune H, Merabti T, Darmoni S, Fieschi M. Assisting the translation of SNOMED CT into French using UMLS and four representative French-language terminologies. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2009; 2009: 291-5	CT	Pre-development	Translation	Other/multiple/unknown/not applicable	France	2009
363.	Klein GO, Chen R. Translation of SNOMED CT - strategies and description of a pilot project. Studies in health technology and informatics 2009; 146: 673-7	CT	Pre-development	Translation	Other/multiple/unknown/not applicable	Sweden	2009
364.	Zhu Y, Pan H, Zhou L, Zhao W, Chen A, Andersen U, Pan S, Tian L, Lei J. Translation and localization of SNOMED CT in China: a pilot study. Artificial intelligence in medicine 2012 Feb;54(2):147-9	CT	Pre-development	Translation	SNOMED CT Hierarchy	China	2012
365.	Implementation of SNOMED-CT needed to facilitate interoperable exchange of health information. Journal of the American Health Information Management Association 2005; 76: 30-	CT	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	United States	2005
366.	Benson T. The history of the Read Codes: the inaugural James Read Memorial Lecture 2011. Informatics in primary care 2011;19(3):173-82	Multiple	Theoretic	Description of SNOMED	Primary care	United Kingdom	2012
367.	Casey A, Spisla C, Konicek D, Warren JJ. Practical definition of SNOMED CT concepts: The case of education, advice and counselling. Studies in health technology and informatics 2006; 122: 742-5	CT	Theoretic	Description of SNOMED	Nursing	United States	2006
368.	Ceusters W. SNOMED CT's RF2: Is the future bright? Studies in health technology and informatics 2011; 169: 829-33	CT	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	United States	2011
369.	Ceusters W, Elkin P, Smith B. Negative findings in electronic health records and biomedical ontologies: a realist approach. International journal of medical informatics 2007 Dec; 76 Suppl 3: S326-33	CT	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	United States	2007
370.	Ceusters WM, Spackman KA, Smith B. Would SNOMED CT benefit from realism-based ontology evolution? AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2007; : 105-9	CT	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	United States	2007
371.	Chalmers RJ. Health care terminology for the electronic era. Mayo Clinic proceedings. Mayo Clinic 2006 Jun; 81(6): 729-31	Multiple	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	Multiple countries	2006

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
372.	Cimino JJ, Zhu X. The practical impact of ontologies on biomedical informatics. Yearbook of medical informatics 2006; : 124-35	CT	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	United States	2006
373.	Cook J, Foley MM, Giannangelo K, Paterno MD, Scichilone RA, Schwarz KM. Universal adapters. Terminology standards enable meaningful data exchange. Journal of AHIMA / American Health Information Management Association 2009 Jan; 80(1): 36-40; quiz 41-2	CT	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	United States	2009
374.	Coonan KM. Medical informatics standards applicable to emergency department information systems: making sense of the jumble. Academic emergency medicine : official journal of the Society for Academic Emergency Medicine 2004 Nov; 11(11): 1198-205	CT	Theoretic	Description of SNOMED	Emergency department/room	United States	2004
375.	Cornet R. Definitions and qualifiers in SNOMED CT. Methods of information in medicine 2009; 48(2): 178-83	CT	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	Netherlands	2009
376.	Cornet R. Do SNOMED CT relationships qualify? Studies in health technology and informatics 2008; 136: 785-90	CT	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	Netherlands	2008
377.	Cornet R. Information-content-based measures for the structure of terminological systems and for data recorded using these systems. Studies in health technology and informatics 2010; 160(Pt 2): 1075-9	CT	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	Netherlands	2010
378.	Cornet R, de Keizer NF. Recording associated disorders using SNOMED CT. Studies in health technology and informatics 2011; 169: 824-8	CT	Theoretic	Description of SNOMED	Problem list / diagnoses	Netherlands	2011
379.	Cornet R, Schulz S. Relationship groups in SNOMED CT. Studies in health technology and informatics 2009; 150: 223-7	CT	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	Netherlands	2009
380.	de Lusignan S. Codes, classifications, terminologies and nomenclatures: definition, development and application in practice. Informatics in primary care 2005; 13(1): 65-70	Multiple	Theoretic	Description of SNOMED	Primary care	United Kingdom	2005
381.	de Lusignan S, van Weel C. The use of routinely collected computer data for research in primary care: opportunities and challenges. Family practice 2006 Apr; 23(2): 253-63	Multiple	Theoretic	Description of SNOMED	Primary care	United Kingdom	2006
382.	Donnelly K. Multilingual documentation and classification. Studies in health technology and informatics 2008; 134: 235-43	CT	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	United States	2008

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
383.	Donnelly K. SNOMED-CT: The advanced terminology and coding system for eHealth. Studies in health technology and informatics 2006; 121: 279-90	CT	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	United States	2006
384.	Elevitch FR. SNOMED CT: electronic health record enhances anesthesia patient safety. AANA journal 2005 Oct; 73(5): 361-6	CT	Theoretic	Description of SNOMED	Anaesthesia	United States	2005
385.	Foley MM, Glenn RM, Meli PL, Scichilone RA. Policy agenda for the next decade: creating a path for graceful evolution and harmonized classifications and terminologies used for encoding health information in electronic environments. Perspectives in health information management / AHIMA, American Health Information Management Association 2009; 6: 1c	CT	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	United States	2009
386.	García-Rojo M, Daniel C, Laurinavicius A. SNOMED CT in Pathology. Studies in health technology and informatics 2012;179:123-40	CT	Theoretic	Description of SNOMED	Pathology	Spain	2012
387.	Giannangelo K, Berkowitz L. SNOMED CT helps drive EHR success. Journal of AHIMA / American Health Information Management Association 2005 Apr; 76(4): 66-7	CT	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	United States	2005
388.	Grain H. Clinical terminology. Studies in health technology and informatics 2010; 151: 70-83	CT	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	Australia	2010
389.	Hastings J, Schulz S. Ontologies for human behavior analysis and their application to clinical data. Int Rev Neurobiol 2012;103:89-107	CT	Theoretic	Description of SNOMED	Mental	Switzerland	2012
390.	Hoskins HD, Hildebrand PL, Lum F. The American Academy of Ophthalmology adopts SNOMED CT as its official clinical terminology. Ophthalmology 2008 Feb; 115(2): 225-6	CT	Theoretic	Description of SNOMED	Ophthalmology	United States	2008
391.	Ingenerf J, Pöppl SJ. Biomedical vocabularies--the demand for differentiation. Studies in health technology and informatics 2007; 129(Pt 1): 610-5	CT	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	Germany	2007
392.	Kilbourne J, Williams T. Unicode, UTF-8, ASCII, and SNOMED CT. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2003; : 892	CT	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	United States	2003

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
393.	Kim JM, Frosdick P. Description of a drug hierarchy in a concept-based reference terminology. Proceedings / AMIA ... Annual Symposium. AMIA Symposium 2001; : 314-8	CT	Theoretic	Description of SNOMED	Drugs	United States	2001
394.	Kudla KM, Blakemore M. SNOMED takes the next step. Journal of AHIMA / American Health Information Management Association 2001 Jul-Aug; 72(7): 62, 64-8; quiz 69-70	CT	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	United States	2001
395.	Levy B. Evolving to clinical terminology. Journal of healthcare information management : JHIM 2004; 18(3): 37-43	CT	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	United States	2004
396.	Lunney M, Delaney C, Duffy M, Moorhead S, Welton J. Advocating for standardized nursing languages in electronic health records. The Journal of nursing administration 2005 Jan; 35(1): 1-3	CT	Theoretic	Description of SNOMED	Nursing	United States	2005
397.	Massey KA, Ansermino JM, von Dadelszen P, Morris TJ, Liston RM, Magee LA. What is SNOMED CT and why should the ISSHP care? Hypertension in pregnancy : official journal of the International Society for the Study of Hypertension in Pregnancy 2009 Feb; 28(1): 119-21	CT	Theoretic	Description of SNOMED	Gynaecology	Canada	2009
398.	Nadkarni PM. Drug safety surveillance using de-identified EMR and claims data: issues and challenges. Journal of the American Medical Informatics Association : JAMIA 2010 Nov-Dec; 17(6): 671-4	CT	Theoretic	Description of SNOMED	Drugs	United States	2010
399.	Nash SK. Nonsynonymous synonyms: correcting and improving SNOMED CT. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2003; : 949	CT	Theoretic	Description of SNOMED	SNOMED CT Hierarchy	United States	2003
400.	Rector AL, Brandt S. Why do it the hard way? The case for an expressive description logic for SNOMED. Journal of the American Medical Informatics Association : JAMIA 2008 Nov-Dec; 15(6): 744-51	CT	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	United Kingdom	2008
401.	Rollins G. Information technology. Help is on the way. Hospitals & health networks / AHA 2003 Sep; 77(9): 24-5	CT	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	Unknown	2003
402.	Rothschild AS, Lehmann HP, Hripcsak G. Inter-rater agreement in physician-coded problem lists. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2005; : 644-8	CT	Theoretic	Description of SNOMED	Internal medicine	United States	2005

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
403.	Schulz S, Hahn U, Rogers J. Semantic Clarification of the Representation of Procedures and Diseases in SNOMED((R))CT. Studies in health technology and informatics 2005; 116: 773-8	CT	Theoretic	Description of SNOMED	SNOMED CT Hierarchy	Germany	2005
404.	Schulz S, Hanser S, Hahn U, Rogers J. The semantics of procedures and diseases in SNOMED CT. Methods of information in medicine 2006; 45(4): 354-8	CT	Theoretic	Description of SNOMED	SNOMED CT Hierarchy	Germany	2006
405.	Schulz S, Klein GO. SNOMED CT - advances in concept mapping, retrieval, and ontological foundations. Selected contributions to the Semantic Mining Conference on SNOMED CT (SMCS 2006). BMC medical informatics and decision making 2008; 8 Suppl 1: S1	CT	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	Germany	2008
406.	Schulz S, Markó K, Suntisrivaraporn B. Formal representation of complex SNOMED CT expressions. BMC medical informatics and decision making 2008; 8 Suppl 1: S9	CT	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	Germany	2008
407.	Schulz S, Schober D, Daniel C, Jaulent MC. Bridging the semantics gap between terminologies, ontologies, and information models. Studies in health technology and informatics 2010; 160(Pt 2): 1000-4	CT	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	Germany	2010
408.	Schulz S, Suntisrivaraporn B, Baader F. SNOMED CT's problem list: ontologists' and logicians' therapy suggestions. Studies in health technology and informatics 2007; 129(Pt 1): 802-6	CT	Theoretic	Description of SNOMED	Problem list / diagnoses	Germany	2007
409.	Schulz S, Suntisrivaraporn B, Baader F, Boeker M. SNOMED reaching its adolescence: ontologists' and logicians' health check. International journal of medical informatics 2009 Apr; 78 Suppl 1: S86-94	CT	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	Germany	2009
410.	Scichilone RA. The benefits of using SNOMED CT and LOINC in assessment instruments. Journal of AHIMA / American Health Information Management Association 2008 Jul; 79(7): 56-7	CT	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	United States	2008
411.	Shahpori R, Doig C. Systematized Nomenclature of Medicine-Clinical Terms direction and its implications on critical care. Journal of critical care 2010 Jun; 25(2): 364.e1-9	CT	Theoretic	Description of SNOMED	Critical care	Canada	2010

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
412.	Shamoun D, Livesay L. Organizing the animal hierarchy into a Linnean Taxonomy in SNOMED CT. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2003; : 1005	CT	Theoretic	Description of SNOMED	SNOMED CT Hierarchy	United States	2003
413.	Showell C, Cummings E, Turner P. Language games and patient-centred eHealth. Studies in health technology and informatics 2010; 155: 55-61	CT	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	Australia	2010
414.	Simpson CR, Anandan C, Fischbacher C, Lefevre K, Sheikh A. Will Systematized Nomenclature of Medicine-Clinical Terms improve our understanding of the disease burden posed by allergic disorders? Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology 2007 Nov; 37(11): 1586-93	CT	Theoretic	Description of SNOMED	Allergies	United Kingdom	2007
415.	Southwick K. Information please--SNOMED answers call. CAP today / College of American Pathologists 2001 Jan; 15(1): 54-6	Multiple	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	United States	2001
416.	Spackman K. SNOMED CT unlocks the power of clinical data for pathologists Lab. Med. 2002; 33(4): 263-266	CT	Theoretic	Description of SNOMED	Pathology	United States	2002
417.	Spackman KA. Rates of change in a large clinical terminology: three years experience with SNOMED Clinical Terms. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2005; : 714-8	CT	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	United States	2005
418.	Spackman KA. SNOMED CT milestones: endorsements are added to already-impressive standards credentials. Healthcare informatics : the business magazine for information and communication systems 2004 Sep; 21(9): 54, 56	CT	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	United States	2004
419.	Spackman KA, Dionne R, Mays E, Weis J. Role grouping as an extension to the description logic of Ontylog, motivated by concept modeling in SNOMED. Proceedings / AMIA ... Annual Symposium. AMIA Symposium 2002; : 712-6	CT	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	United States	2002
420.	Stearns MQ, Price C, Spackman KA, Wang AY. SNOMED clinical terms: overview of the development process and project status. Proceedings / AMIA ... Annual Symposium. AMIA Symposium 2001; : 662-6	CT	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	United States	2001

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
421.	Vanopstal K, Vander Stichele R, Laureys G, Buysschaert J. Vocabularies and retrieval tools in biomedicine: disentangling the terminological knot. Journal of medical systems 2011 Aug; 35(4): 527-43	CT	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	Belgium	2011
422.	Venkat-Raman G, Tomson CR, Gao Y, Cornet R, Stengel B, Gronhagen-Riska C, Reid C, Jacquelinet C, Schaeffner E, Boeschoten E, Casino F, Collart F, De Meester J, Zurriaga O, Kramar R, Jager KJ, Simpson K. New primary renal diagnosis codes for the ERA-EDTA. Nephrol Dial Transplant 2012 Dec;27(12):4414-4419	CT	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	Multiple countries	2012
423.	Wagner L. Data sharing on the horizon. Provider (Washington, D.C.) 2004 Jan; 30(1): 20-4, 26-7, 30-1 passim	Multiple	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	United States	2004
424.	Wang AY, Barrett JW, Bentley T, Markwell D, Price C, Spackman KA, Stearns MQ. Mapping between SNOMED RT and Clinical terms version 3: a key component of the SNOMED CT development process. Proceedings / AMIA ... Annual Symposium. AMIA Symposium 2001; : 741-5	CT	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	United States	2001
425.	Wang AY, Sable JH, Spackman KA. The SNOMED clinical terms development process: refinement and analysis of content. Proceedings / AMIA ... Annual Symposium. AMIA Symposium 2002; : 845-9	CT	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	United States	2002
426.	Warren JJ, Casey A, Konicek D, Lundberg C, Correia C, Zingo C. Where is the nursing in SNOMED CT? CTGFN has the answer! AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2003; : 1047	CT	Theoretic	Description of SNOMED	Nursing	United States	2003
427.	Wei D, Wang Y, Perl Y, Xu J, Halper M, Spackman KA. Complexity measures to track the evolution of a SNOMED hierarchy. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2008; : 778-82	CT	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	United States	2008
428.	Yu AC. Methods in biomedical ontology. Journal of biomedical informatics 2006 Jun; 39(3): 252-66	CT	Theoretic	Description of SNOMED	Other/multiple/unknown/not applicable	United States	2006

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
429.	Al-Mubaid H, Nguyen HA. A cluster-based approach for semantic similarity in the biomedical domain. Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference 2006; 1: 2713-7	CT	Theoretic	Illustrate TS theory	Other/multiple/unknown/not applicable	United States	2006
430.	Bakhshi-Raiez F, Cornet R, de Keizer NF. Development and application of a framework for maintenance of medical terminological systems. Journal of the American Medical Informatics Association : JAMIA 2008 Sep-Oct; 15(5): 687-700	CT	Theoretic	Illustrate TS theory	Other/multiple/unknown/not applicable	Netherlands	2008
431.	Batet M, Sánchez D, Valls A. An ontology-based measure to compute semantic similarity in biomedicine. Journal of biomedical informatics 2011 Feb; 44(1): 118-25	CT	Theoretic	Illustrate TS theory	Other/multiple/unknown/not applicable	Spain	2011
432.	Bodenreider O. Biomedical ontologies in action: role in knowledge management, data integration and decision support. Yearbook of medical informatics 2008; : 67-79	CT	Theoretic	Illustrate TS theory	Other/multiple/unknown/not applicable	United States	2008
433.	Burgun A. Desiderata for domain reference ontologies in biomedicine. Journal of biomedical informatics 2006 Jun; 39(3): 307-13	CT	Theoretic	Illustrate TS theory	Anatomy	France	2006
434.	Burkhart L, Androwich I. Measuring spiritual care with informatics. ANS. Advances in nursing science 2009 Jul-Sep; 32(3): 200-10	CT	Theoretic	Illustrate TS theory	Nursing	United States	2009
435.	Burkhart L, Konicek R, Moorhead S, Androwich I. Mapping parish nurse documentation into the nursing interventions classification: a research method. Computers, informatics, nursing : CIN 2005 Jul-Aug; 23(4): 220-9	CT	Theoretic	Illustrate TS theory	Nursing	United States	2005
436.	Ceusters W, Smith B. Strategies for referent tracking in electronic health records. Journal of biomedical informatics 2006 Jun; 39(3): 362-78	CT	Theoretic	Illustrate TS theory	Other/multiple/unknown/not applicable	Germany	2006
437.	Chabalier J, Mosser J, Burgun A. Integrating biological pathways in disease ontologies. Studies in health technology and informatics 2007; 129(Pt 1): 791-5	CT	Theoretic	Illustrate TS theory	Pathology	France	2007

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
438.	Ciolko E, Lu F, Joshi A. Intelligent clinical decision support systems based on SNOMED CT. Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference 2010; 2010: 6781-4	CT	Theoretic	Illustrate TS theory	Other/multiple/unknown/not applicable	Canada	2010
439.	Cornet R, de Keizer NF, Abu-Hanna A. A framework for characterizing terminological systems. Methods of information in medicine 2006; 45(3): 253-66	CT	Theoretic	Illustrate TS theory	Other/multiple/unknown/not applicable	Netherlands	2006
440.	Elkin PL, Carter JS, Nabar M, Tuttle M, Lincoln M, Brown SH. Drug knowledge expressed as computable semantic triples. Studies in health technology and informatics 2011; 166: 38-47	CT	Theoretic	Illustrate TS theory	Drugs	United States	2011
441.	Elkin PL, Tuttle MS, Trusko BE, Brown SH. BioProspecting: novel marker discovery obtained by mining the bibleome. BMC bioinformatics 2009; 10 Suppl 2: S9	CT	Theoretic	Illustrate TS theory	Genetics	United States	2009
442.	Garla VN, Brandt C. Semantic similarity in the biomedical domain: an evaluation across knowledge sources. BMC bioinformatics 2012 Oct 10;13(1):261	CT	Theoretic	Illustrate TS theory	Other/multiple/unknown/not applicable	United States	2012
443.	Hibbert M, Lohrey J, Melnikoff S. Integration of data for research. Studies in health technology and informatics 2010; 151: 461-75	CT	Theoretic	Illustrate TS theory	Other/multiple/unknown/not applicable	Australia	2010
444.	Ingenerf J, Linder R. Assessing applicability of ontological principles to different types of biomedical vocabularies. Methods of information in medicine 2009; 48(5): 459-67	CT	Theoretic	Illustrate TS theory	Other/multiple/unknown/not applicable	Germany	2009
445.	Lee WN, Shah N, Sundlass K, Musen M. Comparison of ontology-based semantic-similarity measures. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2008; : 384-8	CT	Theoretic	Illustrate TS theory	Pathology	United States	2008
446.	Marenco L, Wang R, Nadkarni P. Automated database mediation using ontological metadata mappings. Journal of the American Medical Informatics Association : JAMIA 2009 Sep-Oct; 16(5): 723-37	CT	Theoretic	Illustrate TS theory	Other/multiple/unknown/not applicable	United States	2009
447.	Martínez S, Sánchez D, Valls A. A semantic framework to protect the privacy of electronic health records with non-numerical attributes. Biomed Inform 2012 Dec 8	CT	Theoretic	Illustrate TS theory	Other/multiple/unknown/not applicable	Spain	2012

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
448.	Melton GB, Parsons S, Morrison FP, Rothschild AS, Markatou M, Hripcsak G. Inter-patient distance metrics using SNOMED CT defining relationships. Journal of biomedical informatics 2006 Dec; 39(6): 697-705	CT	Theoretic	Illustrate TS theory	Other/multiple/unknown/not applicable	United States	2006
449.	Pantazatos SP, Li J, Pavlidis P, Lussier YA. Integration of Neuroimaging and Microarray Datasets through Mapping and Model-Theoretic Semantic Decomposition of Unstructured Phenotypes. Summit on translational bioinformatics 2009; 2009: 85-9	CT	Theoretic	Illustrate TS theory	Pathology	United States	2009
450.	Pedersen T, Pakhomov SV, Patwardhan S, Chute CG. Measures of semantic similarity and relatedness in the biomedical domain. Journal of biomedical informatics 2007 Jun; 40(3): 288-99	CT	Theoretic	Illustrate TS theory	Other/multiple/unknown/not applicable	United States	2007
451.	Pivovarov R, Elhadad N. A hybrid knowledge-based and data-driven approach to identifying semantically similar concepts. Journal of biomedical informatics 2012 Jun;45(3):471-81	CT	Theoretic	Illustrate TS theory	Nephrology	United States	2012
452.	Rodrigues JM, Trombert Paviot B, Martin C, Vercherin P, Samuel O. Co-ordination between clinical coding systems and pragmatic clinical terminologies based on a core open system: the role of ISO/TC215/WG3 and CEN/TC2511/WG2 standardisation? Studies in health technology and informatics 2002; 90: 401-5	CT	Theoretic	Illustrate TS theory	Other/multiple/unknown/not applicable	France	2002
453.	Rosenbloom ST, Miller RA, Johnson KB, Elkin PL, Brown SH. A model for evaluating interface terminologies. Journal of the American Medical Informatics Association : JAMIA 2008 Jan-Feb; 15(1): 65-76	CT	Theoretic	Illustrate TS theory	Other/multiple/unknown/not applicable	United States	2008
454.	Schriml LM, Arze C, Nadendla S, Chang YW, Mazaitis M, Felix V, Feng G, Kibbe WA. Disease Ontology: a backbone for disease semantic integration. Nucleic acids research 2011 Nov 12;	CT	Theoretic	Illustrate TS theory	Genetics	United States	2011
455.	Silverstein JC, Dech F, Kouchoukos PL. Enhancing radiological volumes with symbolic anatomy using image fusion and collaborative virtual reality. Studies in health technology and informatics 2004; 98: 347-52	CT	Theoretic	Illustrate TS theory	Radiology	United States	2004

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
456.	Smith B. From concepts to clinical reality: an essay on the benchmarking of biomedical terminologies. Journal of biomedical informatics 2006 Jun; 39(3): 288-98	CT	Theoretic	Illustrate TS theory	Other/multiple/unknown/not applicable	United States	2006
457.	Smith B, Brochhausen M. Putting biomedical ontologies to work. Methods of information in medicine 2010; 49(2): 135-40	CT	Theoretic	Illustrate TS theory	Other/multiple/unknown/not applicable	Germany	2010
458.	Smith B, Ceusters W. An ontology-based methodology for the migration of biomedical terminologies to electronic health records. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2005; : 704-8	CT	Theoretic	Illustrate TS theory	Other/multiple/unknown/not applicable	United States	2005
459.	Souvignet J, Declerck G, Trombert B, Rodrigues JM, Jaulent MC, Bousquet C. Evaluation of automated term groupings for detecting anaphylactic shock signals for drugs. AMIA Annu Symp Proc 2012;2012:882-90	CT	Theoretic	Illustrate TS theory	Drugs	France	2012
460.	Wilson PS. What mapping and modeling means to the HIM professional. Perspectives in health information management / AHIMA, American Health Information Management Association 2007; 4: 2	CT	Theoretic	Illustrate TS theory	Other/multiple/unknown/not applicable	United States	2007
461.	Wu S, Liu H. Semantic characteristics of NLP-extracted concepts in clinical notes vs. biomedical literature. AMIA Annu Symp Proc 2011;2011:1550-8	CT	Theoretic	Illustrate TS theory	Other/multiple/unknown/not applicable	United States	2011
462.	Agrawal A, Elhanan G, Halper M. Dissimilarities in the Logical Modeling of Apparently Similar Concepts in SNOMED CT. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2010; 2010: 212-6	CT	Theoretic	Terminology audit	Wounds	United States	2010
463.	Bodenreider O, Smith B, Kumar A, Burgun A Investigating subsumption in SNOMED CT: an exploration into large description logic-based biomedical terminologies. Artificial intelligence in medicine 2007 Mar; 39(3): 183-95	CT	Theoretic	Terminology audit	Other/multiple/unknown/not applicable	United States	2007
464.	Burgun A, Bodenreider O, Mougin F Classifying diseases with respect to anatomy: a study in SNOMED CT. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2005; : 91-5	CT	Theoretic	Terminology audit	Other/multiple/unknown/not applicable	France	2005

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
465.	Ceusters W. Applying Evolutionary Terminology Auditing to SNOMED CT. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2010; 2010: 96-100	CT	Theoretic	Terminology audit	Other/multiple/unknown/not applicable	United States	2010
466.	Ceusters W, Smith B, Kumar A, Dhaen C. Mistakes in medical ontologies: where do they come from and how can they be detected? Studies in health technology and informatics 2004; 102: 145-63	CT	Theoretic	Terminology audit	Other/multiple/unknown/not applicable	United States	2004
467.	Ceusters W, Smith B, Kumar A, Dhaen C. Ontology-based error detection in SNOMED-CT. Studies in health technology and informatics 2004; 107(Pt 1): 482-6	CT	Theoretic	Terminology audit	Other/multiple/unknown/not applicable	United States	2004
468.	Geller J, Ochs C, Perl Y, Xu J. New abstraction networks and a new visualization tool in support of auditing the SNOMED CT content. AMIA Annu Symp Proc 2012;2012:237-46	CT	Theoretic	Terminology audit	SNOMED CT Hierarchy	United States	2012
469.	Halper M, Wang Y, Min H, Chen Y, Hripcsak G, Perl Y, Spackman KA. Analysis of error concentrations in SNOMED. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2007; : 314-8	CT	Theoretic	Terminology audit	Other/multiple/unknown/not applicable	United States	2007
470.	Héja G, Surján G, Varga P. Ontological analysis of SNOMED CT. BMC medical informatics and decision making 2008; 8 Suppl 1: S8	CT	Theoretic	Terminology audit	Other/multiple/unknown/not applicable	Hungary	2008
471.	Huang KC, Geller J, Elhanan G, Perl Y, Halper M. Auditing SNOMED Integration into the UMLS for Duplicate Concepts. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2010; 2010: 321-5	CT	Theoretic	Terminology audit	Other/multiple/unknown/not applicable	United States	2010
472.	Jiang G, Chute CG. Auditing the semantic completeness of SNOMED CT using formal concept analysis. Journal of the American Medical Informatics Association : JAMIA 2009 Jan-Feb; 16(1): 89-102	CT	Theoretic	Terminology audit	SNOMED CT Hierarchy	United States	2009
473.	Kreuzthaler M, Schulz S. Metonymies in medical terminologies. A SNOMED CT case study. AMIA Annu Symp Proc 2012;2012:463-7	CT	Theoretic	Terminology audit	Other/multiple/unknown/not applicable	Australia	2012
474.	Mikroyannidi E, Stevens R, Iannone L, Rector A. Analysing Syntactic Regularities and Irregularities in SNOMED-CT. J Biomed Semantics 2012 Dec 17;3(1):8	CT	Theoretic	Terminology audit	Other/multiple/unknown/not applicable	United Kingdom	2012

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
475.	Pacheco E, Stenzhorn H, Nohama P, Paetzold J, Schulz S. Detecting Underspecification in SNOMED CT concept definitions through natural language processing. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2009; 2009: 492-6	CT	Theoretic	Terminology audit	Other/multiple/unknown/not applicable	Brazil	2009
476.	Rector A, Iannone L. Lexically suggest, logically define: Quality assurance of the use of qualifiers and expected results of post-coordination in SNOMED CT. Journal of biomedical informatics 2011 Oct 14;	CT	Theoretic	Terminology audit	Problem list / diagnoses	United Kingdom	2011
477.	Rector AL, Brandt S, Schneider T. Getting the foot out of the pelvis: modeling problems affecting use of SNOMED CT hierarchies in practical applications. Journal of the American Medical Informatics Association : JAMIA 2011 Jul-Aug; 18(4): 432-40	CT	Theoretic	Terminology audit	Problem list / diagnoses	United Kingdom	2011
478.	Sable JH, Nash SK, Wang AY. Culling a clinical terminology: a systematic approach to identifying problematic content. Proceedings / AMIA ... Annual Symposium. AMIA Symposium 2001; : 578-82	CT	Theoretic	Terminology audit	Other/multiple/unknown/not applicable	United States	2001
479.	Schulz S, Rector A, Rodrigues JM, Spackman K. Competing interpretations of disorder codes in SNOMED CT and ICD. AMIA Annu Symp Proc 2012;2012:819-27	CT	Theoretic	Terminology audit	Problem list / diagnoses	Austria	2012
480.	Schulz S, Spackman K, James A, Cocos C, Boeker M. Scalable representations of diseases in biomedical ontologies. Journal of biomedical semantics 2011; 2 Suppl 2: S6	CT	Theoretic	Terminology audit	Pathology	Austria	2011
481.	Wang J, Day R, Visweswaran S, Hogan W. The use of semantic distance metrics to support ontology audit. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2010; 2010: 842-6	CT	Theoretic	Terminology audit	Anatomy	United States	2010
482.	Wang Y, Halper M, Min H, Perl Y, Chen Y, Spackman KA Structural methodologies for auditing SNOMED. Journal of biomedical informatics 2007 Oct; 40(5): 561-81	CT	Theoretic	Terminology audit	Other/multiple/unknown/not applicable	United States	2007
483.	Wang Y, Halper M, Wei D, Gu H, Perl Y, Xu J, Elhanan G, Chen Y, Spackman KA, Case JT, Hripcsak G. Auditing complex concepts of SNOMED using a refined hierarchical abstraction network. Journal of biomedical informatics 2011 Sep 1;	CT	Theoretic	Terminology audit	Other/multiple/unknown/not applicable	United States	2011

No	Reference	Version	Focus Category	Usage Category	Domain	Country	Year
484.	Wang Y, Halper M, Wei D, Perl Y, Geller J. Abstraction of complex concepts with a refined partial-area taxonomy of SNOMED. Journal of biomedical informatics 2011 Aug 25;	CT	Theoretic	Terminology audit	Other/multiple/unknown/not applicable	United States	2011
485.	Wang Y, Wei D, Xu J, Elhanan G, Perl Y, Halper M, Chen Y, Spackman KA, Hripcsak G. Auditing complex concepts in overlapping subsets of SNOMED. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2008; : 273-7	CT	Theoretic	Terminology audit	Other/multiple/unknown/not applicable	United States	2008
486.	Wei D, Bodenreider O. Using the abstraction network in complement to description logics for quality assurance in biomedical terminologies - a case study in SNOMED CT. Studies in health technology and informatics 2010; 160(Pt 2): 1070-4	CT	Theoretic	Terminology audit	Other/multiple/unknown/not applicable	United States	2010
487.	Wei D, Halper M, Elhanan G, Chen Y, Perl Y, Geller J, Spackman KA. Auditing SNOMED relationships using a converse abstraction network. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2009; 2009: 685-9	CT	Theoretic	Terminology audit	SNOMED CT Hierarchy	United States	2009
488.	Zhang GQ, Bodenreider O. Large-scale, Exhaustive Lattice-based Structural Auditing of SNOMED CT. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium 2010; 2010: 922-6	CT	Theoretic	Terminology audit	Other/multiple/unknown/not applicable	United States	2010

12.3 Summary of Key Findings from Abstracts

* Refers to an id that was assigned to each paper that was retrieved.

12.3.1 Usage Category: Used to classify or code in a study

Table 12-2. Summary of key findings: Used to classify or code in a study.

Id*	Summary
218	Require multidisciplinary content experts to develop subset
280	Use of SNOMED CT to extract diagnosis and procedures from discharge summary
285	Automated assignment of negation to concepts identified in health records based on review of the text is feasible and practical. SNOMED-CT had overall coverage of 88.7% of the concepts being negated.
456	Using SNOMED CT to annotate free text and feed into electronic quality monitoring has the potential to improve healthcare quality and safety.
462	We conclude that SNOMED CT based computable rules are accurate enough for the automated biosurveillance of pneumonias from radiological reports.
465	Annotating data with SNOMED CT can help to identify influenza
466	No implications

Id*	Summary
470	SNOMED CT could represent an improvement compared to existing medical terminologies such as MeSH
527	Used to annotate anatomic and diagnostic noun phrases in pathology reports with SNOMED CT.
531	No implicates, early stage of research. Other papers by same authors show results.
533	By annotating free text with NIC thesaurus and SNOMED CT, ontology-based queries can be executed. Enable users to locate biomedical data related to ontology concepts.
562	SNOMED CT used as a standard to facilitate decision support
635	A system to classify lung TNM stages from free-text pathology reports was developed. SNOMED CT can be used for the extraction of key lung cancer characteristics from free-text reports. Future work will investigate the applicability of using the proposed methodology for extracting other cancer characteristics and types.
662	Use of a more granulated, precise and well defined terminology (SNOMED CT vs ICD-10) can help improve quality monitoring and quality development, reduce time and cost in quality management. SNOMED CT can facilitate international benchmarking and research.
717	Use SNOMED CT to automatically classify cancer from free text, can lessen reliance on expert clinical staff, improve efficiency and available of cancer info
744	SNOMED CT used for capturing symptoms at ED, can be used to help classify symptoms into groups
745	No implications
760	No implications
764	Mapped medications to SNOMED CT, potential to improve analysis
858	SNOMED CT can be used to classify cancers in reports using medical free-text processing. Developing decision support systems that are integrated with the free-text processing can reduce costs, enable decision support, enhance efficiency and timeliness

12.3.2 Usage Category: Description of SNOMED CT Implementation

Table 12-3. Summary of key findings: Description of SNOMED CT Implementation.

Id*	Summary
215	No implications
244	No implications in abstract
272	Concept-based searching is helpful for searching for physician referrals
288	No abstract
347	Tools are needed to locate appropriate SNOMED CT concepts quickly
355	SNOMED CT concepts can be used to retrieve literature indexed with MeSH via the UMLS
389	Visualisation techniques can aid in exploring terminologies like SNOMED CT
420	SNOMED CT could represent 83% of concepts used in clinical practice guidelines for the management of obesity
433	SNOMED CT provides standardization and facilitates decision support systems
436	Tools to aid coding of clinical research data using SNOMED CT can aid the implementation of data standards to facilitate high quality research data
440	No implications in abstract
444	No implications in abstract
498	SNOMED CT provides standard for enterprise applications
519	SNOMED CT enables consistency of recording patient conditions, allowing retrieval and analysis from narratives, facilitate answering questions on data, potentially. improve quality and efficiency of care

Id*	Summary
569	No findings, but SNOMED CT has potential benefit to promote patient safety, provide a standard to interface to knowledge base, potential economic advantage, promote interoperability
626	No findings
653	Four implementation challenges: user interface, validating templates that use SNOMED CT, handling SNOMED CT subsets and extensions, creating fast, meaningful, non-redundant search results. Usage of SNOMED CT in tertiary care is promising, requires collaboration.
673	SNOMED CT facilitates standardized approach to data collection and reporting, which can be used in developing diagnostic, prevention, treatment and survivorship strategies against breast cancer.
674	SNOMED CT captured data can facilitate research studies (new approaches to disease prognosis, risk factor assessment, and therapeutic interventions)
685	Usability is an important aspect of implementing a clinical system. Complexities such as the hierarchy should be hidden from the user. Subsets, terms familiar to clinicians and well-designed system are key success factors.
686	Coded structured data forms mapped to SNOMED CT is useful for supporting clinical decision making on pressure ulcer wound management.
720	Rich semantics in SNOMED CT can help support detection of adverse drug events via semantic query and reasoning.
762	Structured forms coded with SNOMED CT enables data entry to be done just once and facilitates advanced data retrieval. Reduced total time for documentation, reporting and follow up. Use of structured documentation with SNOMED CT improves documentation, supports advanced retrieval of data and reduces resource utilization.
788	SNOMED CT used to help standardize data entry and improve results of retrieval.

12.3.3 Usage Category: Retrieve or analyse patient data

Table 12-4. Summary of key findings: Retrieve or analyse patient data.

Id*	Summary
600	Searched database (free text) using SNOMED CT search terms, no implications
604	Search database (free text) using SNOMED CT terms. Search can be used to help reconcile underreported stats
619	Search SNOMED CT encoded database. SMEs are competent to perform both easy and complex searches.
684	SNOMED CT not include in abstract. Synonyms used to identify patients. Although not semantic searches, can still be used to retrieve relevant patients.
765	SNOMED CT synonyms used as keywords to train NLP. Can help index free text, which in turns facilitates decision support
767	Can be used to help audit completeness of discharge summaries by extracting concepts
781	Concept based searches perform better than keyword searches
785	SNOMED CT is used to annotate free text, which can then classify cancers

12.3.4 Search Strategy

PubMed Search Strategy:

("systematized nomenclature of medicine"[MeSH Terms] OR ("systematized"[All Fields] AND "nomenclature"[All Fields] AND "medicine"[All Fields]) OR "systematized nomenclature of medicine"[All Fields] OR "snomed"[All Fields]) AND ("2001-01-01"[PDAT] : "2012-12-31"[PDAT])

Last run: March 13, 2013
Number of results: 537

Embase Search Strategy:

1 SNOMED.mp. or exp "Systematized Nomenclature of Medicine"/ (823)
2 limit 1 to (english language and yr="2001 - 2012") (593)
3 SNOMED.af. (686)
4 limit 3 to (english language and yr="2001 - 2012") (472)
5 2 or 4 (594)

Last run: March 13, 2013
Number of results: 594

12.3.5 Comparison between PubMed and JAMIA

PubMed Search Strategy:

(("systematized nomenclature of medicine"[MeSH Terms] OR ("systematized"[All Fields] AND "nomenclature"[All Fields] AND "medicine"[All Fields]) OR "systematized nomenclature of medicine"[All Fields] OR "snomed"[All Fields]) AND ("J Am Med Inform Assoc"[Journal] OR "jamia"[All Fields])) AND ("2001/01/01"[PDAT] : "2012/12/31"[PDAT])

JAMIA Search Strategy (title or abstract):

JAMIA Search Strategy (full text):

Paper by Paper Comparison

Table 12-5. Paper by paper comparison between JAMIA and PubMed.

No	Reference	JAMIA	PubMed
1.	Strauss JA, Chao CR, Kwan ML, Ahmed SA, Schottinger JE, Quinn VP. Identifying primary and recurrent cancers using a SAS-based natural language processing algorithm. J Am Med Inform Assoc. 2013 Mar 1;20(2):349-55.		Y
2.	Wu ST, Liu H, Li D, Tao C, Musen MA, Chute CG, Shah NH. Unified Medical Language System term occurrences in clinical notes: a large-scale corpus analysis. J Am Med Inform Assoc. 2012 Jun;19(1e):e149-56.		Y
3.	López-García P, Boeker M, Illarramendi A, Schulz S. Usability-driven pruning of large ontologies: the case of SNOMED CT. J Am Med Inform Assoc. 2012 Jun;19(1e):e102-9.	Y	Y
4.	Elhanan G, Perl Y, Geller J. A survey of SNOMED CT direct users, 2010: impressions and preferences regarding content and quality. J Am Med Inform Assoc. 2011 Dec;18 Suppl 1:i36-44.	Y	Y
5.	Liu H, Burkhart Q, Bell DS. Evaluation of the NCPDP Structured and Codified Sig Format for e-prescriptions. J Am Med Inform Assoc. 2011 Sep-Oct;18(5):645-51.		Y
6.	Pathak J, Wang J, Kashyap S, Basford M, Li R, Masys DR, Chute CG. Mapping clinical phenotype data elements to standardized metadata repositories and controlled terminologies: the eMERGE Network experience. J Am Med Inform Assoc. 2011 Jul-Aug;18(4):376-86.		Y
7.	Rector AL, Brandt S, Schneider T. Getting the foot out of the pelvis: modeling problems affecting use of SNOMED CT hierarchies in practical applications. J Am Med Inform Assoc. 2011 Jul-Aug;18(4):432-40.	Y	Y
8.	Fung KW, McDonald C, Srinivasan S. The UMLS-CORE project: a study of the problem list terminologies used in large healthcare institutions. J Am Med Inform Assoc. 2010 Nov-Dec;17(6):675-80.		Y
9.	Nadkarni PM. Drug safety surveillance using de-identified EMR and claims data: issues and challenges. J Am Med Inform Assoc. 2010 Nov-Dec;17(6):671-4		Y
10.	Wilcke JR, Green JM, Spackman KA, Martin MK, Case JT, Santamaria SL, Zimmerman K. Concerning SNOMED-CT content for public health case reports. J Am Med Inform Assoc. 2010 Sep-Oct;17(5):613; author reply 613-4.	Y	Y
11.	Nadkarni PM, Darer JA. Migrating existing clinical content from ICD-9 to SNOMED. J Am Med Inform Assoc. 2010 Sep-Oct;17(5):602-7	Y	Y
12.	Rajeev D, Staes CJ, Evans SR, Mottice S, Rolfs R, Samore MH, Whitney J, Kurzban R, Huff SM. Concerning SNOMED-CT content for public health case reports. J Am Med Inform Assoc 2010;17:5 613-614.	Y	
13.	Nguyen AN, Lawley MJ, Hansen DP, Bowman RV, Clarke BE, Duhig EE, Colquist S. Symbolic rule-based classification of lung cancer stages from free-text pathology reports. J Am Med Inform Assoc. 2010 Jul-Aug;17(4):440-5.		Y
14.	Nadkarni PM, Marenco LA. Implementing description-logic rules for SNOMED-CT attributes through a table-driven approach. J Am Med Inform Assoc. 2010 Mar-Apr;17(2):182-4.	Y	Y
15.	Rajeev D, Staes CJ, Evans RS, Mottice S, Rolfs R, Samore MH, Whitney J, Kurzban R, Huff SM. Development of an electronic public health case report using HL7 v2.5 to meet public health needs. J Am Med Inform Assoc. 2010 Jan-Feb;17(1):34-41		Y
16.	Marenco L, Wang R, Nadkarni P. Automated database mediation using ontological metadata mappings. J Am Med Inform Assoc. 2009 Sep-Oct;16(5):723-37		Y
17.	Pathak J, Solbrig HR, Buntrock JD, Johnson TM, Chute CG. LexGrid: a framework for representing, storing, and querying biomedical terminologies from simple to sublime. J Am Med Inform Assoc. 2009 May-Jun;16(3):305-15.		Y
18.	Jiang G, Chute CG. Auditing the semantic completeness of SNOMED CT using formal concept analysis. J Am Med Inform Assoc. 2009 Jan-Feb;16(1):89-102.	Y	Y
19.	Rosenbloom ST, Brown SH, Froehling D, Bauer BA, Wahner-Roedler DL, Gregg WM, Elkin PL. Using SNOMED CT to represent two interface terminologies. J Am Med Inform Assoc. 2009 Jan-Feb;16(1):81-8.	Y	Y

No	Reference	JAMIA	PubMed
20.	Rector AL, Brandt S. Why do it the hard way? The case for an expressive description logic for SNOMED. J Am Med Inform Assoc. 2008 Nov-Dec;15(6):744-51.	Y	Y
21.	Bouhaddou O, Warnekar P, Parrish F, Do N, Mandel J, Kilbourne J, Lincoln MJ. Exchange of computable patient data between the Department of Veterans Affairs (VA) and the Department of Defense (DoD): terminology mediation strategy. J Am Med Inform Assoc. 2008 Mar-Apr;15(2):174-83.		Y
22.	Rosenbloom ST, Miller RA, Johnson KB, Elkin PL, Brown SH. A model for evaluating interface terminologies. J Am Med Inform Assoc. 2008 Jan-Feb;15(1):65-76.		Y
23.	Andrews JE, Richesson RL, Krischer J. Variation of SNOMED CT coding of clinical research concepts among coding experts. J Am Med Inform Assoc. 2007 Jul-Aug;14(4):497-506.	Y	Y
24.	Chen Y, Perl Y, Geller J, Cimino JJ. Analysis of a study of the users, uses, and future agenda of the UMLS. J Am Med Inform Assoc. 2007 Mar-Apr;14(2):221-31.		Y
25.	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 Sep-Oct;13(5):536-46.	Y	Y
26.	Rosenbloom ST, Miller RA, Johnson KB, Elkin PL, Brown SH. Interface terminologies: facilitating direct entry of clinical data into electronic health record systems. J Am Med Inform Assoc. 2006 May-Jun;13(3):277-88.		Y
27.	Green JM, Wilcke JR, Abbott J, Rees LP. Development and evaluation of methods for structured recording of heart murmur findings using SNOMED-CT post-coordination. J Am Med Inform Assoc. 2006 May-Jun;13(3):321-33. Epub 2006 Feb 24.	Y	Y
28.	Fung KW, Hole WT, Nelson SJ, Srinivasan S, Powell T, Roth L. Integrating SNOMED CT into the UMLS: an exploration of different views of synonymy and quality of editing. J Am Med Inform Assoc. 2005 Jul-Aug;12(4):486-94.	Y	Y

12.4 Appendix C: For Chapter Four

12.4.1 Towards a Successful SNOMED CT Implementation

The following table provides selected references from the summary of statements made in the interviews that were used to help develop the 10 steps.

Step	Survey
#1 - Understand Process of SNOMED CT Implementation	▪ Let clinicians drive the process, collaborate with subject matter experts, clinicians validate, terminology team provides (#6) ▪ Terminology analyst, terminology expert, clinicians (F#1 ▪ IT specialist, terminology specialist (medical background), able to understand doctors (#8)
#2 - Encode Local Terms to SNOMED CT	▪ Subsets developed by reviewing over 10 million free text records (#10) ▪ Develop own palliative care subset, about 2,000 concepts, derived from historical patient records, mainly disorders and findings (#1) ▪ Mapped Apache diagnoses to SNOMED CT and included all subordinate concepts, subset was 83k (#9)
#3 - Create Extensions	▪ Need for proper channels to submit extensions (#2) ▪ No formal extensions yet, just through interface terminology (#1, #6, #4, #9, #3)
#4 - Compile Subsets	▪ Over 1,500 concepts used in forms (#8) ▪ Drop down lists and auto-complete, use subsets (#6) ▪ 5,000 different subsets/value sets (e.g., blood pressure site, contain five to 20) (#4)
#5 - Design Intuitive Data Entry Interfaces	▪ Simple system, try to reduce mouse clicks (#2) ▪ Auto-complete, using algorithms and indexed tables (#2, #1, #9, #10, #11) ▪ Hotlist, 20 of most frequency occurring diagnosis/procedures selected by clinicians (#9) ▪ Simplicity of system, learnable in a day, add complexities later (#2) ▪ Ease of use, speed, performance, reliability, functionality, smart software, aid clinicians, enable decision support, secondary use (#4) ▪ Ability to find exactly what you want (#11)
#6 - Select a Data Storage Method	▪ Store concept ids directly (#2, #8) ▪ Store using interface terminology (#6, #1, #4, #3) ▪ Store description ids of extension, no concept ids or version (#10)
#7 - Incorporate Cross Maps	▪ Currently using National Heath Services cross map, more mature (#2) ▪ Uses maps for ICD-9, UMLS, RxNorm (#3) ▪ May use cross maps in the future once SNOMED CT diagnoses are used (#8) ▪ Want to use cross map but not available (#11) ▪ Uses ICD-9-CM cross map developed by EMIS (#12)
#8 - Design Retrieval Functions	▪ While subsets are important, focus on developing algorithms that can retrieve the appropriate concepts (#10) ▪ For reporting purposes (#2, #6) ▪ No ability to do subsumption queries even though requested from vendors (#11) ▪ Use enumerated list rather than subsumption hierarchy (#7)
#9 - Conduct Training Sessions	▪ Provide a lot of training (#2) ▪ Do not assume individuals computer literacy levels (#2) ▪ Went slowly, trained in very small areas (#2) ▪ Provide training for terms that cannot be found (#2) ▪ 30 to 45 minute training session (#9)
#10 - Develop Maintenance Policies	▪ Frequent updates to SNOMED CT, inactive concepts (#1) ▪ SNOMED CT is large, difficulty to stay synchronised and updated (#4) ▪ Challenge setting up the terminology server and maintaining it every six months (#7) ▪ Have not solved migration issue especially if SNOMED CT is used directly in patient records (#7) ▪ Not sure how to update inactive concepts (#7) ▪ Difficulty in getting vendors to upgrade their version of SNOMED CT as vendor state that very few clinicians are using it (#11)

12.5 Appendix D: For Chapter Six

12.5.1 Overview of Extensions

12.5.1.1 Concepts

The frequency count of concept attributes included concept status and is primitive.

12.5.1.1.1 Concept Status

The frequency count of extension concepts by concept status are shown in **Table 12-6**. The majority of the CA extension concepts were pending move (92.4%), the majority of the US extension were current (95.4%), while the UK extension consisted primarily of duplicate (33.5%), limited (27.4%) and current (25.5%).

Table 12-6. Frequency counts of CA, US and UK extension concepts by concept status.

No	Concept Status	CA		US		UK	
		Total	%	Total	%	Total	%
1.	Current (0)	64	4.6%	1,266	95.4%	18,335	25.5%
2.	Retired without a stated reason (1)					142	0.2%
3.	Duplicate (2)	4	0.3%	8	0.6%	24,090	33.5%
4.	Outdated (3)					11	0.0%
5.	Ambiguous (4)					1,957	2.7%
6.	Erroneous (5)	1	0.1%			1,176	1.6%
7.	Limited (6)					19,714	27.4%
8.	Moved elsewhere (10)	36	2.6%	53	4.0%	795	1.1%
9.	Pending move (11)	1,273	92.4%			5,651	7.9%
	Total	**1,378**	**100.0%**	**1,327**	**100.0%**	**71,871**	**100.0%**

12.5.1.1.2 Is Primitive

The frequency count of extension descriptions by description are shown in **Table 12-7**. The majority of active concepts were primitive for all three extensions. In addition, there were 677 fully defined inactive concepts in the UK extension.

Table 12-7. Frequency count of CA, US and UK extensions by is primitive status.

No	Is Primitive	CA		US		UK	
		Total	%	Total	%	Total	%
1.	Active concepts (including pending move)						
	Primitive	1,333	96.7%	1,030	77.6%	23,940	33.3%
	Fully defined	4	0.3%	236	17.8%	46	0.1%
2.	Inactive concepts (excluding pending move)						
	Primitive	41	3.0%	61	4.6%	47,208	65.7%
	Fully defined					677	

12.5.1.2 Descriptions

The frequency count of description attributes included description type and description status.

12.5.1.2.1 Description Type

The frequency count of extension descriptions by description are shown in **Table 12-8**. It is worth noting that only the UK extension included unspecified (**DescriptionType=0**) descriptions.

Table 12-8. Frequency count of CA, US and UK extensions by is primitive status.

No	Description Type	CA		US		UK	
		Total	%	Total	%	Total	%
1.	Preferred (1)	1,380	47.9%	1,634	38.6%	74,049	42.8%
2.	Synonym (2)	120	4.2%	1,179	27.9%	16,244	9.4%
3.	Fully specified name (3)	1,378	47.9%	1,415	33.5%	82,615	47.8%
4.	Unspecified (0)					29	0.0%
	Total	**2,878**	**100.0%**	**4,228**	**100.0%**	**172,937**	**100.0%**

12.5.1.2.2 Description Status

The frequency count of extension descriptions by description status are shown in **Table 12-9**. The majority of the CA extension descriptions were pending move (92.1%), the majority of the US extension were current (86.8%), while the UK extension consisted primarily of concept inactive (36.0%), current (24.9%) and limited (23.2%).

Table 12-9. Frequency counts of CA, US and UK extension descriptions by description status.

No	Description Status	CA		US		UK	
		Total	%	Total	%	Total	%
1.	Current (0)	139	4.8%	3,668	86.8%	43,105	24.9%
2.	Retired without a stated reason (1)	8	0.3%			13,277	7.7%
3.	Duplicate (2)					17	0.0%
4.	Outdated (3)					1,583	0.9%
5.	Erroneous (5)	2	0.1%	282	6.7%	33	0.0%
6.	Limited (6)					40,160	23.2%
7.	Inappropriate (7)					3	0.0%
8.	Concept Inactive (8)			278	6.6%	62,329	36.0%
9.	Moved elsewhere (10)	77	2.7%				
10.	Pending move (11)	2,652	92.1%			12,430	7.2%
	Total	**2,878**	**100.0%**	**4,228**	**100.0%**	**172,937**	**100.0%**

12.5.1.3 Relationships

The frequency count of relationship attributes included relationship type, characteristic type and how the core and extension concepts were linked to each other.

12.5.1.3.1 Relationship Type

The frequency count of extension relationships by relationship type are shown in **Table 12-10**. The majority of all three extensions were **116680003|Is a (attribute)|** relationships. The numbers of unique relationship types were 31 for the CA extension, 38 for the US extension and 61 for the UK extension. It should be noted that the UK extension included an extension relationship type **84971000000100|PBCL flag true (attribute)|**.

Table 12-10. Frequency counts of CA, US and UK extension relationships by relationship type.

No	Relationship Type	CA		US		UK	
		Total	%	Total	%	Total	%
1.	260507000|Access (attribute)|					1,618	0.7%
2.	255234002|After (attribute)|					85	0.0%
3.	246090004|Associated finding (attribute)|			41	1.8%	4,392	2.0%
4.	116676008|Associated morphology (attribute)|	1	0.0%	84	3.7%	1,410	0.6%
5.	363589002|Associated procedure (attribute)|	1	0.0%	51	2.2%	843	0.4%
6.	47429007|Associated with (attribute)|			6	0.3%	109	0.0%
7.	246075003|Causative agent (attribute)|	4	0.1%	23	1.0%	759	0.3%
8.	263502005|Clinical course (attribute)|			8	0.4%	4,811	2.2%
9.	246093002|Component (attribute)|					1,346	0.6%
10.	363699004|Direct device (attribute)|	5	0.1%	2	0.1%	206	0.1%
11.	363700003|Direct morphology (attribute)|	7	0.2%	1	0.0%	154	0.1%
12.	363701004|Direct substance (attribute)|	33	0.9%	4	0.2%	285	0.1%
13.	42752001|Due to (attribute)|			12	0.5%	34	0.0%
14.	246456000|Episodicity (attribute)|					4,749	2.1%
15.	408729009|Finding context (attribute)|			40	1.8%	1,349	0.6%
16.	419066007|Finding informer (attribute)|					60	0.0%
17.	418775008|Finding method (attribute)|	5	0.1%	32	1.4%	464	0.2%
18.	363698007|Finding site (attribute)|	6	0.2%	87	3.8%	2,441	1.1%
19.	127489000|Has active ingredient (attribute)|	10	0.3%	1	0.0%		
20.	363705008|Has definitional manifestation (attribute)|	3	0.1%	7	0.3%	133	0.1%
21.	411116001|Has dose form (attribute)|	1	0.0%				
22.	363702006|Has focus (attribute)|	3	0.1%	44	1.9%	370	0.2%
23.	363703001|Has intent (attribute)|	105	3.0%	1	0.0%	450	0.2%
24.	363713009|Has interpretation (attribute)|	1	0.0%	4	0.2%	147	0.1%
25.	116686009|Has specimen (attribute)|					51	0.0%
26.	363709002|Indirect morphology (attribute)|			2	0.1%	4	0.0%
27.	363714003|Interprets (attribute)|	14	0.4%	4	0.2%	1,261	0.6%
28.	116680003|Is a (attribute)|	1,625	46.2%	1,465	64.6%	72,027	32.4%
29.	272741003|Laterality (attribute)|	4	0.1%	4	0.2%	60	0.0%

No	Relationship Type	CA		US		UK	
		Total	%	Total	%	Total	%
30.	149016008|MAY BE A (attribute)|					2,355	1.1%
31.	260686004|Method (attribute)|	806	22.9%	12	0.5%	4,806	2.2%
32.	384598002|MOVED FROM (attribute)|	32	0.9%			12,107	5.4%
33.	370125004|MOVED TO (attribute)|	32	0.9%	53	2.3%	795	0.4%
34.	246454002|Occurrence (attribute)|	4	0.1%	8	0.4%	338	0.2%
35.	370135005|Pathological process (attribute)|	3	0.1%	7	0.3%	429	0.2%
36.	84971000000100|PBCL flag true (attribute)|					3,088	1.4%
37.	260870009|Priority (attribute)|					5,903	2.7%
38.	408730004|Procedure context (attribute)|	4	0.1%	50	2.2%	548	0.2%
39.	405815000|Procedure device (attribute)|					25	0.0%
40.	405816004|Procedure morphology (attribute)|					14	0.0%
41.	363704007|Procedure site (attribute)|	5	0.1%			191	0.1%
42.	405813007|Procedure site - Direct (attribute)|	627	17.8%	3	0.1%	1,316	0.6%
43.	405814001|Procedure site - Indirect (attribute)|	18	0.5%	2	0.1%	439	0.2%
44.	370130000|Property (attribute)|					86	0.0%
45.	370131001|Recipient category (attribute)|					8	0.0%
46.	370124000|REPLACED BY (attribute)|					1,188	0.5%
47.	246513007|Revision status (attribute)|					61	0.0%
48.	410675002|Route of administration (attribute)|					13	0.0%
49.	168666000|SAME AS (attribute)|			8	0.4%	60,330	27.1%
50.	246112005|Severity (attribute)|			7	0.3%	4,811	2.2%
51.	118171006|Specimen procedure (attribute)|					25	0.0%
52.	118170007|Specimen source identity (attribute)|			1	0.0%	4	0.0%
53.	118168003|Specimen source morphology (attribute)|			1	0.0%	19	0.0%
54.	118169006|Specimen source topography (attribute)|			2	0.1%	44	0.0%
55.	370133003|Specimen substance (attribute)|			6	0.3%	27	0.0%
56.	408732007|Subject relationship context (attribute)|	3	0.1%	91	4.0%	1,394	0.6%
57.	424876005|Surgical approach (attribute)|					52	0.0%
58.	408731000|Temporal context (attribute)|	4	0.1%	92	4.1%	1,901	0.9%
59.	425391005|Using access device (attribute)|	3	0.1%			188	0.1%
60.	424226004|Using device (attribute)|	11	0.3%	1	0.0%	164	0.1%
61.	424244007|Using energy (attribute)|					16	0.0%
62.	424361007|Using substance (attribute)|	139	3.9%			197	0.1%
63.	159083000|WAS A (attribute)|					19,915	9.0%
	Total	**3,519**	**100.0%**	**2,267**	**100.0%**	**222,415**	**100.0%**

12.5.1.3.2 Characteristic Type

The frequency counts of relationships by characteristic type are shown in **Table 12-11**. All three extensions included defining relationships (**CharacteristicType=0**) while only the US and UK extensions included historical (**CharacteristicType=3**) relationships. The UK extension was the only extension that included qualifier (**CharacteristicType=1**) and additional relationships (**CharacteristicType=3**).

Table 12-11. Frequency counts of CA, US and UK extension relationships by characteristic type.

No	Characteristic Type	CA		US		UK	
		Total	%	Total	%	Total	%
1.	Defining (0)	3,519		2,206		96,658	
2.	Qualifier (1)					25,979	
3.	Historical (2)			61		96,690	
4.	Additional (3)					3,088	

12.5.2 Verification Rules

12.5.2.1 SNOMED CT Identifiers

SNOMED CT extension identifiers consist of an item identifier, namespace, partition identifier and check digit (refer to **Figure 12-4**). The verification rules for the namespace, partition identifier and check digit are described in this section. As the identifications do not contain any semantic information, having an incorrect namespace, partition-identifier or check digit is not a critical error.

All SNOMED CT identifiers (SCTIDs) that use a seven-digit namespace identifier should be 11-18 digits in length. Leading zeros are not permitted. Leading zeros impact the calculation of the check digit and may result in ambiguity in validating an SCTID; hence they are not allowed.

SCTID for Extensions: 243815557777107

Item identifier: Uniquely identifies a particular component. Each organization can assign numbers from 1-8 digits and should devise a scheme to ensure the numbers are never repeated.

Namespace: The seven-digit namespace assigned by the IHTSDO. Use the same namespace for all identifiers created by your organization.

Partition-identifier: A two-digit identifier that ensures that extension components are distinguished from the main body of SNOMED CT. Extensions must use one of the following partition identifiers:

- 10 A concept in an extension
- 11 A description in an extension
- 12 A relationship in an extension
- 13 A subset in an extension
- 14 A cross map in an extension
- 15 A cross map target in an extension

Check digit: Calculate this value using Verhoeff's dihedral check algorithm.

Figure 12-4. Anatomy of SNOMED CT extension identifiers.

12.5.2.1.1 Namespace

The namespace used in SNOMED CT identifiers refers to the seven digits from the third right. The namespace verification checks that the component identifiers created use the correct namespace. The verification rules are shown in **Table 12-12**.

Table 12-12. Namespace verification rules (where X refers to the namespace required).

No	Component	SQL
1.	Concepts	SELECT * FROM TBL_MyExtensionConcepts WHERE LEFT(RIGHT(ConceptId, 10), 7)!='X'
2.	Descriptions	SELECT * FROM TBL_MyExtensionDescriptions WHERE LEFT(RIGHT(DescriptionId, 10), 7)!='X'
3.	Relationships	SELECT * FROM TBL_MyExtensionRelationships WHERE LEFT(RIGHT(RelationshipId, 10), 7)!='X'

12.5.2.1.2 Partition Identifier

The partition identifiers for extension concepts, descriptions and relationships should be **10**, **11** and **12** respectively. The verification rules are shown in **Table 12-13**.

Table 12-13. Partition identifier verification rules.

No	Component	Core	Extension	SQL
1.	Concepts	00	10	SELECT * FROM TBL_MyExtensionConcepts WHERE LEFT(RIGHT(ConceptId, 3), 2)!='10'
2.	Descriptions	01	11	SELECT * FROM TBL_MyExtensionDescriptions WHERE LEFT(RIGHT(DescriptionId, 3), 2)!='11'
3.	Relationships	02	12	SELECT * FROM TBL_MyExtensionRelationships WHERE LEFT(RIGHT(RelationshipId, 3), 2)!='12'

12.5.2.1.3 Check Digit

The check digit is used to verify the validity of a SNOMED CT identifier and is calculated using Verhoeff's dihedral check algorithm. The function in PHP (a web scripting language) that is used to validate the check digit is shown in **Figure 12-5**. A result of **true** indicates the component identifier is valid while a result of **false** indicates the component identifier is invalid.

```
function CheckVerhoeff($SNOMEDCT) {

     $FnF = array();
     $FnF[0] = array(0, 1, 2, 3, 4, 5, 6, 7, 8, 9);
     $FnF[1] = array(1, 5, 7, 6, 2, 8, 3, 0, 9, 4 );

     for ($i=2; $i<8; $i++) {
          for ($j=0; $j<10; $j++) {
               $FnF[$i][$j] = $FnF [$i-1][$FnF[1][$j]];
          }
     }

     $Dihedral = array();
     $Dihedral[0] = array(0, 1, 2, 3, 4, 5, 6, 7, 8, 9);
     $Dihedral[1] = array(1, 2, 3, 4, 0, 6, 7, 8, 9, 5);
     $Dihedral[2] = array(2, 3, 4, 0, 1, 7, 8, 9, 5, 6);
     $Dihedral[3] = array(3, 4, 0, 1, 2, 8, 9, 5, 6, 7);
     $Dihedral[4] = array(4, 0, 1, 2, 3, 9, 5, 6, 7, 8);
     $Dihedral[5] = array(5, 9, 8, 7, 6, 0, 4, 3, 2, 1);
     $Dihedral[6] = array(6, 5, 9, 8, 7, 1, 0, 4, 3, 2);
     $Dihedral[7] = array(7, 6, 5, 9, 8, 2, 1, 0, 4, 3);
     $Dihedral[8] = array(8, 7, 6, 5, 9, 3, 2, 1, 0, 4);
     $Dihedral[9] = array(9, 8, 7, 6, 5, 4, 3, 2, 1, 0);
```

```
    $InverseD5 = array(0, 4, 3, 2, 1, 5, 6, 7, 8, 9);

    $Check = 0;

    for ($i=strlen($SNOMEDCT)-1; $i>=0; $i--) {
        $Check = $Dihedral[$Check][$FnF[(strlen($SNOMEDCT)-$i-1) % 8][substr($SNOMEDCT, $i, 1)]];
    }

    if ($Check != 0) {
        return "false";
    } else {
        return "true";
    }

}
```
Figure 12-5. Verhoeff's check digit verification algorithm in PHP.

12.5.2.1.4 Identifier Duplication

As extensions that are submitted for inclusion into the core use the same extension identifiers, the verification rules in **Table 12-14** are to ensure that there are no duplicates between the identifiers contained in the core and the extension.

Table 12-14. Identifier duplication verification rules.

No	Component	SQL
1.	Concepts	SELECT * FROM TBL_MyExtensionConcepts MEC, SCT_Concepts C WHERE C.ConceptId=MEC.ConceptId
2.	Descriptions	SELECT * FROM TBL_MyExtensionDescriptions MED, SCT_Descriptions D WHERE D.DescriptionId=MED.DescriptionId
3.	Relationships	SELECT * FROM TBL_MyExtensionRelationships MER, SCT_Relationships R WHERE R.RelationshipId=MER.RelationshipId

12.5.2.2 Value Sets

Eight data elements in SNOMED CT use value sets. Two are used in concepts, four in descriptions and two in relationships. While the relationship group in the relationships table (**SCT_Relationships.RelationshipGroup**) does not use values from a value set, it must be a numeral. The verification rules are shown in **Table 12-15**. Having the correct values are important to ensure that textual descriptions can be displayed correctly and that semantic information are properly assigned to concepts. For example, a description with a fully specified name that is defined as **en-US** rather than **en** may not be retrieved if the query explicitly looks for fully specified names with an **en** language code. In another example, if historical relationships use the wrong characteristic type, it may not be possible to locate an active concept for an inactive concept.

Table 12-15. Value set verification rules.

No	Table	Data Element	Permissible Values (RF1)	SQL
1.	Concepts	ConceptStatus	0, 1, 2, 3, 4, 5, 6, 8, 10, 11	SELECT * FROM TBL_MyExtensionConcepts WHERE ConceptStatus NOT IN (0, 1, 2, 3, 4, 5, 6, 8, 10, 11)
2.	Concepts	IsPrimitive	0, 1	SELECT * FROM TBL_MyExtensionConcepts WHERE IsPrimitive NOT IN (0, 1)
3.	Descriptions	DescriptionStatus	0, 1, 2, 3, 5, 6, 7, 8, 10, 11	SELECT * FROM TBL_MyExtensionDescriptions WHERE DescriptionStatus NOT IN (0, 1, 2, 3, 5, 6, 7, 8, 10, 11)

No	Table	Data Element	Permissible Values (RF1)	SQL
4.	Descriptions	DescriptionType	0, 1, 2, 3	SELECT * FROM TBL_MyExtensionDescriptions WHERE DescriptionType NOT IN (0, 1, 2, 3)
5.	Descriptions	InitialCapitalStatus	0, 1	SELECT * FROM TBL_MyExtensionDescriptions WHERE InitialCapitalStatus NOT IN (0, 1)
6.	Descriptions	LanguageCode	en, en-CA, en-GB, en-US	SELECT * FROM TBL_MyExtensionDescriptions WHERE LanguageCode NOT IN ('en', 'en-CA', 'en-GB', 'en-US')
7.	Relationships	CharacteristicType	0, 1, 2, 3	SELECT * FROM TBL_MyExtensionRelationships WHERE CharacteristicType NOT IN (0, 1, 2, 3)
8.	Relationships	Refinability	0, 1	SELECT * FROM TBL_MyExtensionRelationships WHERE Refinability NOT IN (0, 1, 2, 3)
9.	Relationships	RelationshipGroup	IsNumeric	SELECT * FROM TBL_ MyExtensionRelationships WHERE CEIL(RelationshipGroup)!= RelationshipGroup

12.5.2.3 Dependencies

In this context, dependencies refer to the permissible value of a data element based on the value contained in another data element. Two types of dependencies were identified: intra component and inter component.

12.5.2.3.1 Intra Component Dependencies

Intra component dependencies are situations in which the value of a data element is dependent on the value of another data element within the same table. For example, only concepts that are current or pending move can be considered fully-defined. Therefore the **IsPrimitive** field is dependent on the **ConceptStatus** field within the concepts table. Intra component dependencies were identified in all three core components.

12.5.2.3.1.1 Concept Dependencies

The only intra dependency for concepts is that the **IsPrimitive** flag must be set to true (**IsPrimitive=1**) when the concept is not current (**ConceptStatus=0**) or pending move (**ConceptStatus=11**) as inactive concepts cannot be considered fully defined. The concept intra dependency verification rule is shown in Table 12-16. This verification rule is not critical because designating an inactive concept as a fully defined concept has not bearing on any data capture or retrieval.

Table 12-16. Concept intra dependency verification rule.

No	Description	SQL Auditing
1.	Fully defined concepts must be active concepts	SELECT * FROM TBL_MyExtensionConcepts WHERE ConceptStatus NOT IN (0, 11) AND IsPrimitive IN (0)

12.5.2.3.1.2 Description Dependencies

The only intra dependency for descriptions is that current (**DescriptionStatus=0**), pending (**DescriptionStatus=11**) and concept inactive (**DescriptionStatus=8**) descriptions that are designated as fully specified name (**DescriptionType=3**) should have a language code of English (**LanguageCode=en**). The description intra dependency verification rule is shown in **Table 12-17**. This verification rule is critical to ensuring that fully specified name descriptions can be retrieved.

Table 12-17. Description intra dependency verification rule.

No	Description	SQL Auditing
1.	Active and pending move fully specified name descriptions must have EN as the language code	SELECT * FROM TBL_MyExtensionDescriptions WHERE DescriptionStatus IN (0, 8, 11) AND DescriptionType IN (3) AND LanguageCode NOT IN ('en')

12.5.2.3.1.3 Relationship Dependencies

There are four data elements (**CharacteristicType**, **Refinability**, **RelationshipType** and **RelationshipGroup**) for relationships that have mutual dependencies. Eight sets of verification rules were identified and are shown in **Table 12-18**. The relationships between **ConceptId1**, **RelationshipType** and **ConceptId2** have additional dependencies that are discussed in **Section 12.5.2.4**. These verification rules are critical as they ensure that defining attributes, qualifying relationships, historical relationships and part of relationships are assigned correctly. Errors may result in an incorrect construction of the subsumption hierarchy and definition of concepts and the inability to retrieve active concepts through the historical relationships.

Table 12-18. Relationships intra dependency verification rules.

No	Description	SQL
1.	"Is a" attribute must be defining (CharacteristicType=0), not refinable (Refinability=0) and must be ungrouped (RelationshipGroup=0)	SELECT * FROM TBL_MyExtensionRelationships WHERE RelationshipType IN (116680003) AND (CharacteristicType NOT IN (0) OR Refinability NOT IN (0) OR RelationshipGroup NOT IN (0))
2.	Concept model attributes (with the exception of "associated finding," "associated procedure," "is a" and "part of") not linked to qualifier value ranges must have be defining (CharacteristicType=0) and have an optional refinability (Refinability=1)	SELECT * FROM TBL_MyExtensionRelationships WHERE RelationshipType IN (255234002, 116676008, 47429007, 246075003, 42752001, 419066007, 418775008, 363698007, 363705008, 363714003, 246093002, 116686009, 370129005, 127489000, 363699004, 363700003, 363701004, 363702006, 363710007, 363709002, 405815000, 405816004, 363704007, 405813007, 405814001, 370131001, 425391005, 424244007, 424361007, 408732007, 118171006, 118170007, 118168003, 118169006, 370133003) AND (CharacteristicType NOT IN (0) OR Refinability NOT IN (1))
3.	Concept model attributes linked the qualifier value ranges (plus "associated finding" and "associated procedure", minus "surgical approach" and "using device") must be defining with optional refinability (CharacteristicType=0, Refinability=1) or qualifying with mandatory refinability (CharacteristicType=1, Refinability=2)	SELECT * FROM TBL_MyExtensionRelationships WHERE RelationshipType IN (246090004, 363589002, 263502005, 246456000, 363713009, 246454002, 370135005, 246112005, 370130000, 370132008, 370134009, 411116001, 260507000, 363703001, 260686004, 260870009, 246513007, 410675002, 424876005, 408729009, 408730004, 408731000) AND ((CharacteristicType NOT IN (0) AND Refinability NOT IN (1)) AND (CharacteristicType NOT IN (1) AND Refinability NOT IN (2)))
4.	"Surgical approach" attribute must be defining or qualifying (CharacteristicType=0 or CharacteristicType=1) and optional refinability (Refinability=1)	SELECT * FROM TBL_MyExtensionRelationships WHERE RelationshipType IN (424876005) AND (CharacteristicType NOT IN (0, 1) OR Refinability NOT IN (1))

No	Description	SQL
5.	"Using device" attribute must have defining (with optional refinability) or qualifying (with either non refinable or optional refinability)	SELECT * FROM TBL_MyExtensionRelationships WHERE RelationshipType IN (424226004) AND ((CharacteristicType NOT IN (0) AND Refinability NOT IN (1)) AND (CharacteristicType NOT IN (1) AND Refinability NOT IN (0, 1)))
6.	"Laterality" attribute must have a defining (with refinability optional) or qualifying (with refinability mandatory) characteristic type and relationship group must be 0	SELECT * FROM TBL_MyExtensionRelationships WHERE RelationshipType IN (272741003) AND (((CharacteristicType NOT IN (0) AND Refinability NOT IN (1)) AND (CharacteristicType NOT IN (1) AND Refinability NOT IN (2))) OR RelationshipGroup NOT IN (0))
7.	Concept history attributes must have a historical characteristic type, refinability is not refinable and relationship group must be 0	SELECT * FROM TBL_MyExtensionRelationships WHERE RelationshipType IN (178066000, 149016008, 384598002, 370125004, 370124000, 168666000, 159083000) AND (CharacteristicType NOT IN (2) OR Refinability NOT IN (0) OR RelationshipGroup NOT IN (0))
8.	"Part of" attribute must have an additional characteristic type, refinability is not refinable and relationship group must be 0	SELECT * FROM TBL_MyExtensionRelationships WHERE RelationshipType IN (123005000) AND (CharacteristicType NOT IN (3) OR Refinability NOT IN (0) OR RelationshipGroup NOT IN (0))

12.5.2.3.2 Inter Component Dependencies

Inter component dependencies are situations in which the value of a data element is dependent on the value of another data element in a separate table. Three inter component dependencies were defined. First, SNOMED CT identifiers used in the descriptions and relationships tables. Second, the corresponding values between concept status and description status. Third, the corresponding values between concept status and relationship types.

12.5.2.3.2.1 SNOMED CT Identifiers

There are four data elements that refer to SNOMED CT identifiers from other components. The **ConceptId** in the extension descriptions table must exist either as a core or extension concept. Similarly, the **ConceptId1**, **RelationshipType** and **ConceptId2** from the extension relationships table must exist either as a core or extension concept. The verification rules are shown in Table 12-19 and ensure that there are no orphaned descriptions or relationships (i.e., all descriptions and relationships belong to existing concepts).

Table 12-19. SNOMED CT identifiers inter component dependencies verification rules.

No	Description	SQL
1.	ConceptId in the Descriptions table must exist in the Concepts table	SELECT MED.* FROM TBL_MyExtensionDescriptions MED LEFT JOIN TBL_MyExtensionConcepts MEC ON MED.ConceptId=MEC.ConceptId LEFT JOIN SCT_Concepts C ON MED.ConceptId=C.ConceptId WHERE MEC.ConceptId IS NULL AND C.ConceptId IS NULL
2.	ConceptId1 in the Relationships table must exist in the Concepts table	SELECT MER.* FROM TBL_MyExtensionRelationships MER LEFT JOIN TBL_MyExtensionConcepts MEC ON MER.ConceptId1=MEC.ConceptId LEFT JOIN SCT_Concepts C ON MER.ConceptId1=C.ConceptId WHERE MEC.ConceptId IS NULL AND C.ConceptId IS NULL
3.	RelationshipType in the Relationships table must exist in the Concepts table	SELECT MER.* FROM TBL_MyExtensionRelationships MER LEFT JOIN SCT_TransitiveClosure TC ON MER.RelationshipType=TC.SubTypeId AND TC.SuperTypeId IN (410663007, 410662002, 116680003) WHERE TC.SubTypeId IS NULL
4.	ConceptId2 in the Relationships table must exist in the Concepts table	SELECT MER.* FROM TBL_MyExtensionRelationships MER LEFT JOIN TBL_MyExtensionConcepts MEC ON MER.ConceptId2=MEC.ConceptId LEFT JOIN SCT_Concepts C ON MER.ConceptId2=C.ConceptId WHERE MEC.ConceptId IS NULL AND C.ConceptId IS NULL

12.5.2.3.2.2 Concept Status and Description Status

There are dependencies between the concept statuses and description statuses. For example, a description with a description status of current (**DescriptionStatus=0**) can only belong to a concept with a concept status of current (**ConceptStatus=0**). Although there are 10 different description statuses, only six are used in the core. Duplicate (**DescriptionStatus=2**), outdated (**DescriptionStatus=3**), erroneous (**DescriptionStatus=5**) are not currently used as retired (**DescriptionStatus=1**) is generally used instead while pending move (**DescriptionStatus=11**) is only used in extensions. Based on the "Permitted DescriptionStatus values for possible ConceptStatus values" table in the TIG, the combination of description and concept status are shown in **Table 12-20**.

Table 12-20. Permitted description status values for possible concept status values.

Concept (Rows) / Description Status (Columns)	0 – Current	1 - Retired	2 - Duplicate	3 - Outdated	5 - Erroneous	6 - Limited	7 - Inappropriate	8 – Concept Inactive	10 – Moved Elsewhere	11 – Pending Move
0 – Current	✓	✓	✓	✓	✓		✓			
1 – Retired		✓	✓	✓	✓		✓	✓		
2 – Duplicate		✓	✓	✓	✓		✓	✓		
3 – Outdated		✓	✓	✓	✓		✓	✓		
4 – Ambiguous		✓	✓	✓	✓		✓	✓		
5 – Erroneous		✓	✓	✓	✓		✓	✓		
6 – Limited		✓	✓	✓	✓	✓	✓			
10 – Moved Elsewhere		✓	✓	✓	✓		✓	✓	✓	
11 – Pending Move		✓	✓	✓	✓		✓			✓

A review of **Table 12-20** shows that there are six main patterns: (1) Current; (2) Retired, Duplicate, Outdated, Erroneous and Inappropriate; (3) Limited; (4) Concept inactive; (5) Moved elsewhere and (6) Pending move. The verification rules are shown in **Table 12-21** and ensure that the appropriate descriptions can be displayed for concepts of different concept statuses.

Table 12-21. Description status and concept status inter component dependencies verification rules.

No	Description Status	SQL
1.	Current descriptions can only belong to current concepts	SELECT MED.* FROM TBL_MyExtensionConcepts MEC, TBL_MyExtensionDescriptions MED WHERE MEC.ConceptId=MED.ConceptId AND MED.DescriptionStatus IN (0) AND MEC.ConceptStatus NOT IN (0)
2.	Retired, duplicate, outdated, erroneous, and inappropriate descriptions can belong to any concept	SELECT MED.* FROM TBL_MyExtensionConcepts MEC, TBL_MyExtensionDescriptions MED WHERE MEC.ConceptId=MED.ConceptId AND MED.DescriptionStatus IN (1, 2, 3, 5, 7) AND MEC.ConceptStatus NOT IN (0, 1, 2, 3, 4, 5, 6, 10, 11)
3.	Limited descriptions can only	SELECT MED.* FROM TBL_MyExtensionConcepts MEC, TBL_MyExtensionDescriptions MED

No	Description Status	SQL
	belong to limited concepts	WHERE MEC.ConceptId=MED.ConceptId AND MED.DescriptionStatus IN (6) AND MEC.ConceptStatus NOT IN (6)
4.	Concept inactive descriptions can belong to any inactive concept except limited	SELECT MED.* FROM TBL_MyExtensionConcepts MEC, TBL_MyExtensionDescriptions MED WHERE MEC.ConceptId=MED.ConceptId AND MED.DescriptionStatus IN (8) AND MEC.ConceptStatus NOT IN (1, 2, 3, 4, 5, 10)
5.	Moved elsewhere descriptions can only belong to moved elsewhere concepts	SELECT MED.* FROM TBL_MyExtensionConcepts MEC, TBL_MyExtensionDescriptions MED WHERE MEC.ConceptId=MED.ConceptId AND MED.DescriptionStatus IN (10) AND MEC.ConceptStatus NOT IN (10)
6.	Pending move descriptions can only belong to pending move concepts	SELECT MED.* FROM TBL_MyExtensionConcepts MEC, TBL_MyExtensionDescriptions MED WHERE MEC.ConceptId=MED.ConceptId AND MED.DescriptionStatus IN (11) AND MEC.ConceptStatus NOT IN (11)

12.5.2.3.3 Concept Status and Relationship Type

There are two main types of dependencies between the concept status and relationship type. First, non-historical relationships (**CharacteristicType=0, 1, 3**) should be linked to current (**ConceptStatus=0**) or pending move (**ConceptStatus=11**). The verification rule is shown in **Table 12-17** and ensures that all defining, qualifying and part of attributes are comprised of current concepts. Having inactive defining attributes can cause unexpected results when testing for equivalency and subsumption.

Table 12-22. Concept status and relation type for non-historical relationships verification rule.

No	Description	SQL Auditing
1.	All destination concepts in the relationships table are current	SELECT MER.*, MEC.ConceptStatus, C.ConceptStatus FROM TBL_MyExtensionRelationships MER LEFT JOIN TBL_MyExtensionConcepts MEC ON MEC.ConceptId=MER.ConceptId2 LEFT JOIN SCT_Concepts C ON C.ConceptId=MER.ConceptId2 WHERE MER.CharacteristicType IN (0, 1, 3) AND ((MEC.ConceptStatus IS NOT NULL AND MEC.ConceptStatus NOT IN (0, 11)) OR (C.ConceptStatus IS NOT NULL AND C.ConceptStatus NOT IN (0, 11)))

Second, there are certain historical relationships that are applicable to certain inactive concepts. Although the TIG indicates that historical relationships can be used with multiple concept statuses, a review of the historical relationships in the core indicates that it is not being used in that manner. In the January 2012 release version, historical relationships were used exclusively with a single concept status with the exception of limited (**ConceptStatus=6**). It was only in July 2012 and January 2013 that **159083000|WAS A (attribute)|** was used with concept statuses such as retired without a stated reason, outdated and erroneous (refer to **Table 12-23**). The low frequencies indicate these are probably erroneous and that each inactive concept status should be used exclusively with one historical relationship with the exception of limited (**ConceptStatus=6**).

Table 12-23. Historical relationships and concept statuses for the past three release versions of SNOMED CT.

No	ConceptId1 ConceptStatus	RelationshipType	ConceptId2 ConceptStatus	January 2012	July 2012	January 2013
1.	1 (Retired)	370124000|REPLACED BY (attribute)|	0 (Current)	2,178	2,180	2,186
2.		159083000|WAS A (attribute)|				2
3.	2 (Duplicate)	168666000|SAME AS (attribute)|	0 (Current)	37,815	37,892	37,933
4.	3 (Outdated)	370124000|REPLACED BY (attribute)|	0 (Current)	1,431	1,444	1,458

No	ConceptId1 ConceptStatus	RelationshipType	ConceptId2 ConceptStatus	January 2012	July 2012	January 2013
5.		159083000|WAS A (attribute)|				1
6.	4 (Ambiguous)	149016008|MAY BE A (attribute)|	0 (Current)	16,110	30,370	30,429
7.	5 (Erroneous)	370124000|REPLACED BY (attribute)|	0 (Current)	1,152	1,165	1,165
8.		159083000|WAS A (attribute)|			2	3
9.	6 (Limited)	168666000|SAME AS (attribute)|	6 (Limited)	20,930	21,155	21,154
10.	6 (Limited)	159083000|WAS A (attribute)|	0 (Current)	5,987	6,135	6,135
11.	10 (Moved elsewhere)	370125004|MOVED TO (attribute)|	0 (Current)	14,457	14,457	14,457
	Total			**100,060**	**114,800**	**114,923**

Excluding the low frequency count of corresponding concept status used in conjunction with historical relationships, there are eight main rules for using historical relationships for inactive concepts (refer to **Table 12-24**). Note that although pending move is not considered an active concept, it is technically not an inactive concept.

Table 12-24. Verification rules for concept history attributes and inactive concepts.

No	ConceptId1 ConceptStatus	RelationshipType	ConceptId2 ConceptStatus
1.	1 (Retired)	370124000|REPLACED BY (attribute)|	0 (Current)
2.	2 (Duplicate)	168666000|SAME AS (attribute)|	0 (Current)
3.	3 (Outdated)	370124000|REPLACED BY (attribute)|	0 (Current)
4.	4 (Ambiguous)	149016008|MAY BE A (attribute)|	0 (Current)
5.	5 (Erroneous)	370124000|REPLACED BY (attribute)|	0 (Current)
6.	6 (Limited)	168666000|SAME AS (attribute)|	6 (Limited)
7.	6 (Limited)	159083000|WAS A (attribute)|	0 (Current)
8.	10 (Moved elsewhere)	370125004|MOVED TO (attribute)|	0 (Current)

The verification rules for historical relationships and inactive concepts are shown in **Table 12-25** and ensures that active concepts are retrieved when using historical relationships

Table 12-25. Concept status and relationship type for historical relationships verification rules.

No	Definition	SQL
1.	370124000|REPLACED BY (attribute)| should be used to link retired (1), outdated (3) or erroneous (5) to current (0) concepts	SELECT MER.*, MEC1.ConceptStatus, MEC2.ConceptStatus, C1.ConceptStatus, C2.ConceptStatus FROM TBL_MyExtensionRelationships MER LEFT JOIN TBL_MyExtensionConcepts MEC1 ON MEC1.ConceptId=MER.ConceptId1 LEFT JOIN TBL_MyExtensionConcepts MEC2 ON MEC2.ConceptId=MER.ConceptId2 LEFT JOIN SCT_Concepts C1 ON C1.ConceptId=MER.ConceptId1 LEFT JOIN SCT_Concepts C2 ON C2.ConceptId=MER.ConceptId2 WHERE MER.RelationshipType=370124000 AND (((MEC1.ConceptStatus IS NOT NULL AND MEC1.ConceptStatus NOT IN (1, 3, 5)) OR (C1.ConceptStatus IS NOT NULL AND C1.ConceptStatus NOT IN (1, 3, 5))) AND ((MEC2.ConceptStatus IS NOT NULL AND MEC2.ConceptStatus NOT IN (0)) OR (C2.ConceptStatus IS NOT NULL AND C2.ConceptStatus NOT IN (0))))

No	Definition	SQL
2.	168666000|SAME AS (attribute)| should be used to link duplicate (2) concepts to current (0) concepts, or limited (6) concepts to other limited (6) concepts	SELECT MER.*, MEC1.ConceptStatus, MEC2.ConceptStatus, C1.ConceptStatus, C2.ConceptStatus FROM TBL_MyExtensionRelationships MER LEFT JOIN TBL_MyExtensionConcepts MEC1 ON MEC1.ConceptId=MER.ConceptId1 LEFT JOIN TBL_MyExtensionConcepts MEC2 ON MEC2.ConceptId=MER.ConceptId2 LEFT JOIN SCT_Concepts_20120731 C1 ON C1.ConceptId=MER.ConceptId1 LEFT JOIN SCT_Concepts_20120731 C2 ON C2.ConceptId=MER.ConceptId2 WHERE MER.RelationshipType=168666000 AND ((((MEC1.ConceptStatus IS NOT NULL AND MEC1.ConceptStatus NOT IN (2, 6)) OR (C1.ConceptStatus IS NOT NULL AND C1.ConceptStatus NOT IN (2, 6))) OR ((MEC2.ConceptStatus IS NOT NULL AND MEC2.ConceptStatus NOT IN (0, 6)) OR (C2.ConceptStatus IS NOT NULL AND C2.ConceptStatus NOT IN (0, 6)))) OR (((MEC1.ConceptStatus IN (2) OR C1.ConceptStatus IN (2)) AND (MEC2.ConceptStatus NOT IN (0) AND C1.ConceptStatus NOT IN (0))) OR ((MEC1.ConceptStatus NOT IN (2) AND C1.ConceptStatus NOT IN (2)) AND (MEC2.ConceptStatus IN (0) OR C1.ConceptStatus IN (0))) OR ((MEC1.ConceptStatus IN (6) OR C1.ConceptStatus IN (6)) AND (MEC2.ConceptStatus NOT IN (6) AND C1.ConceptStatus NOT IN (6))) OR ((MEC1.ConceptStatus NOT IN (6) AND C1.ConceptStatus NOT IN (6)) AND (MEC2.ConceptStatus IN (6) OR C1.ConceptStatus IN (6)))))
3.	149016008|MAY BE A (attribute)| should be used to link ambiguous (4) concepts to current (0) concepts	SELECT MER.*, MEC1.ConceptStatus, MEC2.ConceptStatus, C1.ConceptStatus, C2.ConceptStatus FROM TBL_MyExtensionRelationships MER LEFT JOIN TBL_MyExtensionConcepts MEC1 ON MEC1.ConceptId=MER.ConceptId1 LEFT JOIN TBL_MyExtensionConcepts MEC2 ON MEC2.ConceptId=MER.ConceptId2 LEFT JOIN SCT_Concepts C1 ON C1.ConceptId=MER.ConceptId1 LEFT JOIN SCT_Concepts C2 ON C2.ConceptId=MER.ConceptId2 WHERE MER.RelationshipType=149016008 AND (((MEC1.ConceptStatus IS NOT NULL AND MEC1.ConceptStatus NOT IN (4)) OR (C1.ConceptStatus IS NOT NULL AND C1.ConceptStatus NOT IN (4))) AND ((MEC2.ConceptStatus IS NOT NULL AND MEC2.ConceptStatus NOT IN (0)) OR (C2.ConceptStatus IS NOT NULL AND C2.ConceptStatus NOT IN (0))))
4.	159083000|WAS A (attribute)| should be used to link limited (6) concepts to current concepts (0)	SELECT MER.*, MEC1.ConceptStatus, MEC2.ConceptStatus, C1.ConceptStatus, C2.ConceptStatus FROM TBL_MyExtensionRelationships MER LEFT JOIN TBL_MyExtensionConcepts MEC1 ON MEC1.ConceptId=MER.ConceptId1 LEFT JOIN TBL_MyExtensionConcepts MEC2 ON MEC2.ConceptId=MER.ConceptId2 LEFT JOIN SCT_Concepts C1 ON C1.ConceptId=MER.ConceptId1 LEFT JOIN SCT_Concepts C2 ON C2.ConceptId=MER.ConceptId2 WHERE MER.RelationshipType=159083000 AND (((MEC1.ConceptStatus IS NOT NULL AND MEC1.ConceptStatus NOT IN (6)) OR (C1.ConceptStatus IS NOT NULL AND C1.ConceptStatus NOT IN (6))) AND ((MEC2.ConceptStatus IS NOT NULL AND MEC2.ConceptStatus NOT IN (0)) OR (C2.ConceptStatus IS NOT NULL AND C2.ConceptStatus NOT IN (0))))
5.	370125004|MOVED TO (attribute)| should be used to link moved elsewhere (10) concepts to current (0) concepts.	SELECT MER.*, MEC1.ConceptStatus, MEC2.ConceptStatus, C1.ConceptStatus, C2.ConceptStatus FROM TBL_MyExtensionRelationships MER LEFT JOIN TBL_MyExtensionConcepts MEC1 ON MEC1.ConceptId=MER.ConceptId1 LEFT JOIN TBL_MyExtensionConcepts MEC2 ON MEC2.ConceptId=MER.ConceptId2 LEFT JOIN SCT_Concepts C1 ON C1.ConceptId=MER.ConceptId1 LEFT JOIN SCT_Concepts C2 ON C2.ConceptId=MER.ConceptId2 WHERE MER.RelationshipType=370125004 AND (((MEC1.ConceptStatus IS NOT NULL AND MEC1.ConceptStatus NOT IN (10)) OR (C1.ConceptStatus IS NOT NULL AND C1.ConceptStatus NOT IN (10))) AND ((MEC2.ConceptStatus IS NOT NULL AND MEC2.ConceptStatus NOT IN (0)) OR (C2.ConceptStatus IS NOT NULL AND C2.ConceptStatus NOT IN (0))))

12.5.2.4 Machine Readable Concept Model

The Machine Readable Concept Model (MRCM) defines how concepts can be modelled and how post-coordinated expressions can be created. Three types of verification rules were created. First, to ensure that the attributes used can be applied to the domains. Second, to ensure that the ranges used can be applied to the attributes. Third, to ensure that incorrect combinations of attributes are not used. The only known dependencies of attributes are with concepts from the **243796009|Situation with explicit context (situation)|** hierarchy. Concepts that are subtypes of **243796009|Situation with explicit context (situation)|** can be post-coordinated with six attributes (i.e., **246090004|Associated finding (attribute)|**, **363589002|Associated procedure (attribute)|**, **408729009|Finding context (attribute)|**, **408730004|Procedure context (attribute)|**, **408732007|Subject relationship context (attribute)|** and **408731000|Temporal context (attribute)|**). The context attributes should occur only once while **408729009|Finding context (attribute)|** should only be used in conjunction with **246090004|Associated finding (attribute)|**, while **408730004|Procedure context (attribute)|** should only be used with **363589002|Associated procedure (attribute)|**. The verification rules are shown in **Table 12-26** and ensure that concepts are modelled consistently by conforming to the MRCM.

Table 12-26. Machine Readable Concept Model (MRCM) verification rules.

No	Definition	SQL
1.	Appropriateness of attribute with domain	SELECT * FROM TBL_MyExtensionRelationships WHERE CharacteristicType IN (0, 1, 3) AND RelationshipType!=116680003 AND RelationshipId NOT IN (SELECT MER.RelationshipId FROM TBL_MyExtensionRelationships MER, TBL_MyExtensionTransitiveClosure METC, SCT_ConceptModel CM WHERE CM.DomainId=METC.SuperTypeId AND METC.SubTypeId=MER.ConceptId1 AND MER.RelationshipType=CM.AttributeId AND MER.CharacteristicType IN (0, 1, 3) AND MER.RelationshipType!=116680003)
2.	Appropriateness of range with attribute	SELECT * FROM TBL_MyExtensionRelationships WHERE CharacteristicType IN (0, 1, 3) AND RelationshipType!=116680003 AND RelationshipId NOT IN (SELECT MER.RelationshipId FROM TBL_MyExtensionRelationships MER, TBL_MyExtensionTransitiveClosure METC, SCT_ConceptModel CM WHERE CM.RangeId=METC.SuperTypeId AND METC.SubTypeId=MER.ConceptId2 AND MER.RelationshipType=CM.AttributeId AND MER.CharacteristicType IN (0, 1, 3) AND MER.RelationshipType!=116680003) ORDER BY RelationshipType
3.	Concepts should either be defined with (associated finding and finding context) or (associated procedure and procedure context) but not a combination of the two	SELECT MEC.* FROM TBL_MyExtensionConcepts MEC, TBL_MyExtensionRelationships MER1, TBL_MyExtensionRelationships MER2 WHERE MEC.ConceptId=MER1.ConceptId1 AND MEC.ConceptId=MER2.ConceptId1 AND MER1.RelationshipId!=MER2.RelationshipId AND MER1.RelationshipType IN (246090004, 408729009) AND MER2.RelationshipType IN (363589002, 408730004)
4.	Concepts should either be defined with only one each of the following: associated finding, finding context, associated procedure, procedure context, temporal context and subject relationship context	SELECT * FROM TBL_MyExtensionRelationships WHERE RelationshipType IN (246090004, 408729009, 363589002, 408730004, 408731000, 408732007) GROUP BY ConceptId1, RelationshipType HAVING COUNT(*) > 1

12.5.2.5 Inferred Relationships

As the records in the relationships table contain inferred relationships, this verification refers to checking to ensure that all children concepts contain at least the same defining attributes of their parent concepts or a refinement of those defining attributes. The method involves retrieving the defining attributes of both the parent and child concept, normalising the concepts and testing them via structural subsumption as described in the TIG. Missing inferred relationships will cause problems when testing for equivalency and subsumption as the long normal forms generated will be missing these relationships.

12.5.2.6 Occurrence

While the previous verification rules centred on using the appropriate values, the verification rules in this section focus on the number of occurrences of description or relationship extensions. Having the appropriate number of descriptions and relationships is critical to ensure that fully specified names and preferred terms are available to be displayed, mandatory "is a" relationships exist, concepts belong to only one top-level hierarchy and historical relationships to locate active concepts exist.

12.5.2.6.1 Descriptions

For descriptions, there were two types of occurrence verification rules, one for the fully specified name and one for the preferred term.

12.5.2.6.1.1 Fully Specified Name

Each concept should have one and only fully specified name in English and not any other English dialect. The corresponding description status will also depend on the concept status. The verification rules are shown in **Table 12-27**.

Table 12-27. Occurrence of fully specified name verification rules.

No	Description	SQL
1.	Each current extension concept (ConceptStatus=0) should have one current (DescriptionStatus=0) fully specified name (DescriptionType=3) in English (LanguageCode=en)	SELECT MEC.* FROM TBL_MyExtensionConcepts MEC LEFT JOIN TBL_MyExtensionDescriptions MED ON MEC.ConceptId=MED.ConceptId AND MED.DescriptionStatus=0 AND MED.DescriptionType=3 AND MED.LanguageCode='en' WHERE MEC.ConceptStatus=0 AND MED.ConceptId IS NULL
2.	Each current extension concept (ConceptStatus=0) should have only one current (DescriptionStatus=0) fully specified name (DescriptionType=3) in English (LanguageCode=en)	SELECT MEC.* FROM TBL_MyExtensionConcepts MEC, TBL_MyExtensionDescriptions MED WHERE MEC.ConceptId=MED.ConceptId AND MED.DescriptionStatus=0 AND MED.DescriptionType=3 AND MED.LanguageCode='en' AND MEC.ConceptStatus=0 GROUP BY MEC.ConceptId HAVING COUNT(*)>1
3.	Each pending move extension concept (ConceptStatus=11) should have one pending move (DescriptionStatus=11) fully specified name (DescriptionType=3) in English (LanguageCode=en)	SELECT MEC.* FROM TBL_MyExtensionConcepts MEC LEFT JOIN TBL_MyExtensionDescriptions MED ON MEC.ConceptId=MED.ConceptId AND MED.DescriptionStatus=11 AND MED.DescriptionType=3 AND MED.LanguageCode='en' WHERE MEC.ConceptStatus=11 AND MED.ConceptId IS NULL
4.	Each pending move extension concept (ConceptStatus=11) should have only one pending move (DescriptionStatus=11) fully specified name (DescriptionType=3) in English (LanguageCode=en)	SELECT MEC.* FROM TBL_MyExtensionConcepts MEC, TBL_MyExtensionDescriptions MED WHERE MEC.ConceptId=MED.ConceptId AND MED.DescriptionStatus=0 AND MED.DescriptionType=3 AND MED.LanguageCode='en' AND MEC.ConceptStatus=11 GROUP BY MEC.ConceptId HAVING COUNT(*)>1
5.	Each inactive extension concept (ConceptStatus!=0 AND ConceptStatus!=6 AND ConceptStatus!=11) should have one concept inactive (DescriptionStatus=8) fully specified name (DescriptionType=3) in English (LanguageCode=en)	SELECT MEC.* FROM TBL_MyExtensionConcepts MEC LEFT JOIN TBL_MyExtensionDescriptions MED ON MEC.ConceptId=MED.ConceptId AND MED.DescriptionStatus=8 AND MED.DescriptionType=3 AND MED.LanguageCode='en' WHERE MEC.ConceptStatus NOT IN (0, 6, 11) AND MED.ConceptId IS NULL
6.	Each inactive extension concept (ConceptStatus!=0 AND ConceptStatus!=6 AND ConceptStatus!=11) should have only one concept inactive (DescriptionStatus=8) fully specified name (DescriptionType=3) in English (LanguageCode=en)	SELECT MEC.* FROM TBL_MyExtensionConcepts MEC, TBL_MyExtensionDescriptions MED WHERE MEC.ConceptId=MED.ConceptId AND MED.DescriptionStatus=8 AND MED.DescriptionType=3 AND MED.LanguageCode='en' AND MEC.ConceptStatus NOT IN (0, 6, 11) GROUP BY MEC.ConceptId HAVING COUNT(*)>1

No	Description	SQL
7.	Each limited extension concept (ConceptStatus=6) should have one pending move (DescriptionStatus=6) fully specified name (DescriptionType=3) in English (LanguageCode=en)	SELECT MEC.* FROM TBL_MyExtensionConcepts MEC LEFT JOIN TBL_MyExtensionDescriptions MED ON MEC.ConceptId=MED.ConceptId AND MED.DescriptionStatus=6 AND MED.DescriptionType=3 AND MED.LanguageCode='en' WHERE MEC.ConceptStatus=6 AND MED.ConceptId IS NULL
8.	Each limited extension concept (ConceptStatus=6) should have only one pending move (DescriptionStatus=6) fully specified name (DescriptionType=3) in English (LanguageCode=en)	SELECT MEC.* FROM TBL_MyExtensionConcepts MEC, TBL_MyExtensionDescriptions MED WHERE MEC.ConceptId=MED.ConceptId AND MED.DescriptionStatus=6 AND MED.DescriptionType=3 AND MED.LanguageCode='en' AND MEC.ConceptStatus=6 GROUP BY MEC.ConceptId HAVING COUNT(*)>1

12.5.2.6.1.2 Preferred Term

Each concept should have at least one preferred term but only one in each language code. Similar to the fully specified name, different description statuses are required for different concept status. The verification rules are shown in **Table 12-28**.

Table 12-28. Occurrence of preferred term auditing rules.

No	Description	SQL Validate
1.	Each current extension concept (ConceptStatus=0) should have at least one current (DescriptionStatus=0) preferred term (DescriptionType=1)	SELECT MEC.* FROM TBL_MyExtensionConcepts MEC LEFT JOIN TBL_MyExtensionDescriptions MED ON MEC.ConceptId=MED.ConceptId AND MED.DescriptionStatus=0 AND MED.DescriptionType=1 WHERE MEC.ConceptStatus=0 AND MED.ConceptId IS NULL
2.	Each current extension concept (ConceptStatus=0) should have at only one current (DescriptionStatus=0) preferred term (DescriptionType=1) in each language code	SELECT MEC.* FROM TBL_MyExtensionConcepts MEC, TBL_MyExtensionDescriptions MED WHERE MEC.ConceptId=MED.ConceptId AND MED.DescriptionStatus=0 AND MED.DescriptionType=1 AND MEC.ConceptStatus=0 GROUP BY MED.ConceptId, MED.LanguageCode HAVING COUNT(*)>1
3.	Each pending move extension concept (ConceptStatus=11) should have at least one pending move (DescriptionStatus=11) preferred term (DescriptionType=1)	SELECT MEC.* FROM TBL_MyExtensionConcepts MEC LEFT JOIN TBL_MyExtensionDescriptions MED ON MEC.ConceptId=MED.ConceptId AND MED.DescriptionStatus=11 AND MED.DescriptionType=1 WHERE MEC.ConceptStatus=11 AND MED.ConceptId IS NULL
4.	Each pending move extension concept (ConceptStatus=11) should have at only one pending move (DescriptionStatus=11) preferred term (DescriptionType=1) in each language code	SELECT MEC.* FROM TBL_MyExtensionConcepts MEC, TBL_MyExtensionDescriptions MED WHERE MEC.ConceptId=MED.ConceptId AND MED.DescriptionStatus=11 AND MED.DescriptionType=1 AND MEC.ConceptStatus=11 GROUP BY MED.ConceptId, MED.LanguageCode HAVING COUNT(*)>1
5.	Each inactive extension concept (ConceptStatus!=0 AND ConceptStatus!=6 AND ConceptStatus!=11) should have at least one concept inactive (DescriptionStatus=8) preferred term (DescriptionType=1)	SELECT MEC.* FROM TBL_MyExtensionConcepts MEC LEFT JOIN TBL_MyExtensionDescriptions MED ON MEC.ConceptId=MED.ConceptId AND MED.DescriptionStatus=8 AND MED.DescriptionType=1 WHERE MEC.ConceptStatus NOT IN (0, 6, 11) AND MED.ConceptId IS NULL
6.	Each inactive extension concept (ConceptStatus!=0 AND ConceptStatus!=6 AND ConceptStatus!=11) should have only one concept inactive (DescriptionStatus=8) preferred term (DescriptionType=1) in each language code	SELECT MEC.* FROM TBL_MyExtensionConcepts MEC, TBL_MyExtensionDescriptions MED WHERE MEC.ConceptId=MED.ConceptId AND MED.DescriptionStatus=8 AND MED.DescriptionType=1 AND MEC.ConceptStatus NOT IN (0, 6, 11) GROUP BY MED.ConceptId, MED.LanguageCode HAVING COUNT(*)>1
7.	Each limited extension concept (ConceptStatus=6) should have at least one current (DescriptionStatus=6) preferred term (DescriptionType=1)	SELECT MEC.* FROM TBL_MyExtensionConcepts MEC LEFT JOIN TBL_MyExtensionDescriptions MED ON MEC.ConceptId=MED.ConceptId AND MED.DescriptionStatus=6 AND MED.DescriptionType=1 WHERE MEC.ConceptStatus=6 AND MED.ConceptId IS NULL

No	Description	SQL Validate
8.	Each limited extension concept (ConceptStatus=6) should have only one current (DescriptionStatus=6) preferred term (DescriptionType=1) in each language code	SELECT MEC.* FROM TBL_MyExtensionConcepts MEC, TBL_MyExtensionDescriptions MED WHERE MEC.ConceptId=MED.ConceptId AND MED.DescriptionStatus=6 AND MED.DescriptionType=1 AND MEC.ConceptStatus=6 GROUP BY MED.ConceptId, MED.LanguageCode HAVING COUNT(*)>1

12.5.2.6.2 Relationships

The verification rules only apply to the occurrences of defining relationships (**CharacteristicType=0**) and historical relationships (**CharacteristicType=2**). Qualifying (**CharacteristicType=1**) and additional relationships (**CharacteristicType=3**) are optional and therefore excluded from the verification.

12.5.2.6.2.1 Defining Relationships

Six verification rules were identified for defining relationships (refer to **Table 12-29**).

Table 12-29. Occurrence of defining relationships verification rules.

No	Description	SQL
1.	Each extension relationship added to a core concept should not duplicate what already exists in the core	SELECT MER.* FROM TBL_MyExtensionRelationships MER, SCT_Relationships R WHERE MER.CharacteristicType=R.CharacteristicType AND MER.RelationshipType=R.RelationshipType AND MER.ConceptId1=R.ConceptId1 AND MER.ConceptId2=R.ConceptId2 AND MER.RelationshipGroup=R.RelationshipGroup
2.	Each extension "is a" relationship should not be redundant to what already exists in the core	SELECT DISTINCT(MER.ConceptId1) FROM TBL_MyExtensionRelationships MER, TBL_MyExtensionTransitiveClosure METC1, TBL_MyExtensionTransitiveClosure METC2 WHERE METC1.SuperTypeId!=METC1.SubTypeId AND METC1.SubTypeId=METC2.SuperTypeId AND METC2.SuperTypeId!=METC2.SubTypeId AND MER.RelationshipType=116680003 AND METC1.SuperTypeId=MER.ConceptId2 AND METC2.SubTypeId=MER.ConceptId1
3.	Each extension concept should belong to only one top-level hierarchy	SELECT MEC.* FROM TBL_MyExtensionConcepts MEC, TBL_MyExtensionTransitiveClosure METC, SCT_Relationships R WHERE MEC.ConceptId=METC.SubTypeId AND METC.SuperTypeId=R.ConceptId1 AND R.RelationshipType=116680003 AND R.ConceptId2=138875005 GROUP BY MEC.ConceptId HAVING COUNT(*) > 1
4.	Each current extension concept should have at least one "is a" relationship	SELECT MEC.* FROM TBL_MyExtensionConcepts MEC LEFT JOIN TBL_MyExtensionRelationships MER ON MEC.ConceptId=MER.ConceptId1 AND MER.RelationshipType=116680003 WHERE MER.RelationshipId IS NULL
5.	Concepts should only have one of the following: associated finding, associated procedure, finding context, procedure context, temporal context and subject relationship context	SELECT * FROM TBL_MyExtensionRelationships WHERE RelationshipType IN (246090004,363589002,363589002,408729009,408730004,408732007,408731000) GROUP BY ConceptId1, RelationshipType HAVING COUNT(*) > 1
6.	Each active extension concept should not have a historical relationship except for 384598002|MOVED FROM (attribute)|	SELECT MEC.* FROM TBL_MyExtensionConcepts MEC, TBL_MyExtensionRelationships MER WHERE MEC.ConceptStatus IN (0, 11) AND MEC.ConceptId=MER.ConceptId1 AND MER.CharacteristicType=2 AND MER.RelationshipType IN (SELECT ConceptId1 FROM SCT_Relationships WHERE RelationshipType=116680003 AND ConceptId2=410663007 AND ConceptId1!=384598002)

12.5.2.6.2.2 Historical Relationships

There are two sets of historical relationships verification rules. The first set ensures that inactive concepts are subtypes of **362955004|Inactive concept (inactive concept)|** (Rule 1 to 8 in **Table 12-30**). The second set ensures each inactive concept has a historical relationship to an active concept (Rule 9 to 15 in **Table 12-30**).

Table 12-30. Occurrence of historical relationships verification rules.

No	Description	SQL
1.	Each inactive extension concept should only have an "is a" relationship and concept history attributes	SELECT MEC.*, MER.RelationshipType FROM TBL_MyExtensionConcepts MEC, TBL_MyExtensionRelationships MER WHERE MEC.ConceptId=MER.ConceptId1 AND MEC.ConceptStatus NOT IN (0, 11) AND MER.RelationshipType NOT IN (SELECT ConceptId1 FROM SCT_Relationships WHERE RelationshipType=116680003 AND ConceptId2=410663007) AND MER.RelationshipType NOT IN (116680003)
2.	Retired without a stated reason concepts (ConceptStatus=1) should be a subtype of 363661006|Reason not stated concept (inactive concept)|	SELECT * FROM TBL_MyExtensionConcepts MEC LEFT JOIN TBL_MyExtensionRelationships MER ON MER.RelationshipType=116680003 AND MER.ConceptId1=MEC.ConceptId AND MER.ConceptId2=363661006 WHERE MEC.ConceptStatus=1 AND MER.RelationshipId IS NULL SELECT * FROM TBL_MyExtensionConcepts MEC, TBL_MyExtensionRelationships MER WHERE MER.RelationshipType=116680003 AND MER.ConceptId1=MEC.ConceptId AND MER.ConceptId2=363661006 AND MEC.ConceptStatus!=1
3.	Duplicate concepts (ConceptStatus=2) should be a subtype of 363662004|Duplicate concept (inactive concept)|	SELECT * FROM TBL_MyExtensionConcepts MEC LEFT JOIN TBL_MyExtensionRelationships MER ON MER.RelationshipType=116680003 AND MER.ConceptId1=MEC.ConceptId AND MER.ConceptId2=363662004 WHERE MEC.ConceptStatus=2 AND MER.RelationshipId IS NULL SELECT * FROM TBL_MyExtensionConcepts MEC, TBL_MyExtensionRelationships MER WHERE MER.RelationshipType=116680003 AND MER.ConceptId1=MEC.ConceptId AND MER.ConceptId2=363662004 AND MEC.ConceptStatus!=2
4.	Outdated concepts (ConceptStatus=3) should be a subtype of 363663009|Outdated concept (inactive concept)|	SELECT * FROM TBL_MyExtensionConcepts MEC LEFT JOIN TBL_MyExtensionRelationships MER ON MER.RelationshipType=116680003 AND MER.ConceptId1=MEC.ConceptId AND MER.ConceptId2=363663009 WHERE MEC.ConceptStatus=3 AND MER.RelationshipId IS NULL SELECT * FROM TBL_MyExtensionConcepts MEC, TBL_MyExtensionRelationships MER WHERE MER.RelationshipType=116680003 AND MER.ConceptId1=MEC.ConceptId AND MER.ConceptId2=363663009 AND MEC.ConceptStatus!=3
5.	Ambiguous concepts (ConceptStatus=4) should be a subtype of 363660007|Ambiguous concept (inactive concept)|.	SELECT * FROM TBL_MyExtensionConcepts MEC LEFT JOIN TBL_MyExtensionRelationships MER ON MER.RelationshipType=116680003 AND MER.ConceptId1=MEC.ConceptId AND MER.ConceptId2=363660007 WHERE MEC.ConceptStatus=4 AND MER.RelationshipId IS NULL SELECT * FROM TBL_MyExtensionConcepts MEC, TBL_MyExtensionRelationships MER WHERE MER.RelationshipType=116680003 AND MER.ConceptId1=MEC.ConceptId AND MER.ConceptId2=363660007 AND MEC.ConceptStatus!=4

No	Description	SQL
6.	Erroneous concepts (ConceptStatus=5) should be a subtype of 363664003|Erroneous concept (inactive concept)|	SELECT * FROM TBL_MyExtensionConcepts MEC LEFT JOIN TBL_MyExtensionRelationships MER ON MER.RelationshipType=116680003 AND MER.ConceptId1=MEC.ConceptId AND MER.ConceptId2=363664003 WHERE MEC.ConceptStatus=5 AND MER.RelationshipId IS NULL SELECT * FROM TBL_MyExtensionConcepts MEC, TBL_MyExtensionRelationships MER WHERE MER.RelationshipType=116680003 AND MER.ConceptId1=MEC.ConceptId AND MER.ConceptId2=363664003 AND MEC.ConceptStatus!=5
7.	Limited concepts (ConceptStatus=6) should be a subtype of 443559000|Limited status concept (inactive concept)|	SELECT * FROM TBL_MyExtensionConcepts MEC LEFT JOIN TBL_MyExtensionRelationships MER ON MER.RelationshipType=116680003 AND MER.ConceptId1=MEC.ConceptId AND MER.ConceptId2=443559000 WHERE MEC.ConceptStatus=6 AND MER.RelationshipId IS NULL SELECT * FROM TBL_MyExtensionConcepts MEC, TBL_MyExtensionRelationships MER WHERE MER.RelationshipType=116680003 AND MER.ConceptId1=MEC.ConceptId AND MER.ConceptId2=443559000 AND MEC.ConceptStatus!=6
8.	Moved elsewhere concepts (ConceptStatus=10) should be a subtype of 370126003|Moved elsewhere (inactive concept)|	SELECT * FROM TBL_MyExtensionConcepts MEC LEFT JOIN TBL_MyExtensionRelationships MER ON MER.RelationshipType=116680003 AND MER.ConceptId1=MEC.ConceptId AND MER.ConceptId2=370126003 WHERE MEC.ConceptStatus=10 AND MER.RelationshipId IS NULL SELECT * FROM TBL_MyExtensionConcepts MEC, TBL_MyExtensionRelationships MER WHERE MER.RelationshipType=116680003 AND MER.ConceptId1=MEC.ConceptId AND MER.ConceptId2=370126003 AND MEC.ConceptStatus!=10
9.	Each "1 (Retired)" concept must be "370124000|REPLACED BY (attribute)|" by one, and only one "0 (Current)" concept	SELECT C1.* FROM TBL_MyExtensionConcepts C1 LEFT JOIN TBL_MyExtensionRelationships R ON C1.ConceptId=R.ConceptId1 AND R.RelationshipType=370124000 LEFT JOIN TBL_MyExtensionConcepts C2 ON C2.ConceptId=R.ConceptId2 AND C2.ConceptStatus IN (0) WHERE C1.ConceptStatus IN (4) GROUP BY C1.ConceptId HAVING COUNT(*)!=1
10.	Each "2 (Duplicate)" concept must be the "168666000|SAME AS (attribute)|" one, and only one "0 (Current)" concept	SELECT C1.* TBL_MyExtensionConcepts C1 LEFT JOIN TBL_MyExtensionRelationships R ON C1.ConceptId=R.ConceptId1 AND R.RelationshipType=168666000 LEFT JOIN TBL_MyExtensionConcepts C2 ON C2.ConceptId=R.ConceptId2 AND C2.ConceptStatus IN (0) WHERE C1.ConceptStatus IN (2) GROUP BY C1.ConceptId HAVING COUNT(*)!=1
11.	Each "3 (Outdated)" concept must be "370124000|REPLACED BY (attribute)|" by one, and only one "0 (Current)" concept	SELECT C1.* FROM TBL_MyExtensionConcepts C1 LEFT JOIN TBL_MyExtensionRelationships R ON C1.ConceptId=R.ConceptId1 AND R.RelationshipType=370124000 LEFT JOIN TBL_MyExtensionConcepts C2 ON C2.ConceptId=R.ConceptId2 AND C2.ConceptStatus IN (0) WHERE C1.ConceptStatus IN (3) GROUP BY C1.ConceptId HAVING COUNT(*)!=1
12.	Each "4 (Ambiguous)" concept 149016008|MAY BE A (attribute)|" one or more "0 (Current)" concepts	SELECT C1.* FROM TBL_MyExtensionConcepts C1 LEFT JOIN TBL_MyExtensionRelationships R ON C1.ConceptId=R.ConceptId1 AND R.RelationshipType=149016008 LEFT JOIN TBL_MyExtensionConcepts C2 ON C2.ConceptId=R.ConceptId2 AND C2.ConceptStatus IN (0) WHERE C1.ConceptStatus IN (4) GROUP BY C1.ConceptId HAVING COUNT(*)<1
13.	Each "5 (Erroneous)" concept must be "370124000|REPLACED BY (attribute)|" by one, and only one "0 (Current)" concept	SELECT C1.* FROM TBL_MyExtensionConcepts C1 LEFT JOIN TBL_MyExtensionRelationships R ON C1.ConceptId=R.ConceptId1 AND R.RelationshipType=370124000 LEFT JOIN TBL_MyExtensionConcepts C2 ON C2.ConceptId=R.ConceptId2 AND C2.ConceptStatus IN (0) WHERE C1.ConceptStatus IN (5) GROUP BY C1.ConceptId HAVING COUNT(*)!=1
14.	Each "6 (Limited)" concept "159083000|WAS A (attribute)|" "0 (Current)" concept	SELECT C1.* FROM TBL_MyExtensionConcepts C1 LEFT JOIN TBL_MyExtensionRelationships R ON C1.ConceptId=R.ConceptId1 AND R.RelationshipType=159083000 LEFT JOIN TBL_MyExtensionConcepts C2 ON C2.ConceptId=R.ConceptId2 AND C2.ConceptStatus IN (0) WHERE C1.ConceptStatus IN (6) GROUP BY C1.ConceptId HAVING COUNT(*)<1

No	Description	SQL
15.	Each "10 (Moved elsewhere)" concept must be "370125004|MOVED TO (attribute)|" to one and only "0 (Current)" concept	SELECT C1.* FROM TBL_MyExtensionConcepts C1 LEFT JOIN TBL_MyExtensionRelationships R ON C1.ConceptId=R.ConceptId1 AND R.RelationshipType=370125004 LEFT JOIN TBL_MyExtensionConcepts C2 ON C2.ConceptId=R.ConceptId2 AND C2.ConceptStatus IN (0) WHERE C1.ConceptStatus IN (10) GROUP BY C1.ConceptId HAVING COUNT(*)!=1

12.5.3 Extension Concepts in the Core

Table 12-31. Extension concepts that have been incorporated into the core.

No	Concept	Concept Status	Is Primitive
1.	1051000119103|Acquired penile adhesion (disorder)|	0	1
2.	551000119107|Acquired pronation deformity of ankle (disorder)|	0	1
3.	521000119104|Acute cystitis in pregnancy, antepartum (disorder)|	0	1
4.	1801000119106|Anemia, pre-end stage renal disease on erythropoietin protocol (disorder)|	0	1
5.	1771000119109|Arthritis of wrist (disorder)|	0	0
6.	211000119109|Atypical ductal hyperplasia of breast (disorder)|	0	1
7.	1731000119106|Atypical mycobacterial infection of lung (disorder)|	0	0
8.	351000119100|Childhood failure to gain weight (finding)|	0	1
9.	1621000119101|Chlamydia trachomatis infection in pregnancy (disorder)|	0	1
10.	451000119106|Closed injury of head (disorder)|	0	0
11.	1141000119101|Congenital labial adhesion (disorder)|	0	1
12.	1641000119107|Coronary arteriosclerosis in native artery (disorder)|	0	1
13.	141000119100|Daytime somnolence (finding)|	0	1
14.	931000119107|Dependence on supplemental oxygen (finding)|	0	1
15.	1061000119101|Excess panniculus of abdomen (finding)|	0	1
16.	181000119105|Fetal growth retardation, antenatal (disorder)|	0	1
17.	541000119105|Hypertension complicating pregnancy, childbirth and the puerperium, antepartum (disorder)|	0	1
18.	531000119101|Hypothyroidism in pregnancy, antepartum (disorder)|	0	1
19.	1181000119106|Inappropriate diet and eating habits (finding)|	0	1
20.	1411000119106|Influenza vaccine needed (situation)|	0	0
21.	1031000119109|Insufficient prenatal care (finding)|	0	1
22.	61000119108|Irritant rhinitis (disorder)|	0	1
23.	581000119100|Long term methotrexate user (finding)|	0	1
24.	171000119107|Maternal obesity complicating pregnancy, childbirth and the puerperium, antepartum (disorder)|	0	1
25.	1611000119108|Mild dehydration (disorder)|	0	1
26.	1601000119105|Moderate dehydration (disorder)|	0	1
27.	301000119104|Muscle spasm of cervical muscle of neck (disorder)|	0	1
28.	221000119102|Never smoked any substance (finding)|	0	1

No	Concept	Concept Status	Is Primitive
29.	311000119101|Non-allergic rhinitis (disorder)|	0	1
30.	111000119104|Nonischemic congestive cardiomyopathy (disorder)|	0	1
31.	101000119102|Numbness and tingling sensation of skin (finding)|	0	1
32.	1101000119103|Obstructive sleep apnea of adult (disorder)|	0	0
33.	1091000119108|Obstructive sleep apnea of child (disorder)|	0	0
34.	1081000119105|Opioid dependence, on agonist therapy (disorder)|	0	1
35.	1041000119100|Partner relationship problem (finding)|	0	1
36.	1721000119108|Patient awaiting renal transplant (finding)|	0	1
37.	1021000119106|Peripherally inserted central venous catheter in situ (finding)|	0	1
38.	161000119101|Placental abnormality, antepartum (finding)|	0	1
39.	1701000119104|Primary adenocarcinoma of colon (disorder)|	0	0
40.	1651000119109|Primary adenocarcinoma of pancreas (disorder)|	0	0
41.	991000119106|Reactive airway disease (disorder)|	0	1
42.	951000119101|Repetitive strain injury of arm or forearm (disorder)|	0	1
43.	471000119102|Sacroiliac joint somatic dysfunction (finding)|	0	1
44.	1111000119100|Sebaceous nevus (disorder)|	2	1
45.	261000119107|Severe depressed bipolar I disorder (disorder)|	0	1
46.	251000119105|Severe major depression, single episode (disorder)|	0	1
47.	271000119101|Severe mixed bipolar I disorder (disorder)|	0	1
48.	681000119101|Sibling relationship problem (finding)|	0	1
49.	981000119108|Single epileptic seizure (finding)|	0	1
50.	961000119104|Sprain of shoulder rotator cuff (disorder)|	0	1
51.	371000119109|Stenosis of intervertebral foramina (disorder)|	0	0
52.	41000119109|Strain of foot (disorder)|	0	1
53.	481000119104|Strain of hamstring muscle (disorder)|	0	1
54.	941000119103|Strain of knee (disorder)|	0	1
55.	341000119102|Tattoo of skin (finding)|	0	1
56.	971000119105|Telangiectasia of limb (disorder)|	0	1
57.	331000119106|Tendinitis of elbow or forearm (disorder)|	0	1
58.	1011000119104|Vaginal pessary in situ (finding)|	0	1

12.5.4 Core Concepts to Core Concepts

Table 12-32. Full core concepts to core concepts United States extension.

No	Parent Concept	Concept Status	Child Concept	Concept Status
1.	443189004|Enteropathogenic Escherichia coli serogroup O18 (organism)|	0	112283007|Escherichia coli (organism)|	0

No	Parent Concept	Concept Status	Child Concept	Concept Status
2.	406627005|Enterotoxigenic Escherichia coli serogroup O6 (organism)|	0	112283007|Escherichia coli (organism)|	0
3.	407156003|Enterotoxigenic Escherichia coli, serotype O8 (organism)|	0	112283007|Escherichia coli (organism)|	0
4.	407282001|Escherichia coli serogroup O104 (organism)|	0	112283007|Escherichia coli (organism)|	0
5.	407205001|Escherichia coli serogroup O114 (organism)|	0	112283007|Escherichia coli (organism)|	0
6.	407208004|Escherichia coli serogroup O119 (organism)|	0	112283007|Escherichia coli (organism)|	0
7.	407232006|Escherichia coli serogroup O124 (organism)|	0	112283007|Escherichia coli (organism)|	0
8.	407210002|Escherichia coli serogroup O125 (organism)|	0	112283007|Escherichia coli (organism)|	0
9.	407212005|Escherichia coli serogroup O126 (organism)|	0	112283007|Escherichia coli (organism)|	0
10.	407215007|Escherichia coli serogroup O127 (organism)|	0	112283007|Escherichia coli (organism)|	0
11.	407181000|Escherichia coli serogroup O128 (organism)|	0	112283007|Escherichia coli (organism)|	0
12.	407296007|Escherichia coli serogroup O137 (organism)|	0	112283007|Escherichia coli (organism)|	0
13.	407259003|Escherichia coli serogroup O14 (organism)|	0	112283007|Escherichia coli (organism)|	0
14.	407223009|Escherichia coli serogroup O142 (organism)|	0	112283007|Escherichia coli (organism)|	0
15.	407238005|Escherichia coli serogroup O143 (organism)|	0	112283007|Escherichia coli (organism)|	0
16.	407162008|Escherichia coli serogroup O15 (organism)|	0	112283007|Escherichia coli (organism)|	0
17.	407242008|Escherichia coli serogroup O152 (organism)|	0	112283007|Escherichia coli (organism)|	0
18.	407166006|Escherichia coli serogroup O157 (organism)|	0	112283007|Escherichia coli (organism)|	0
19.	427294000|Escherichia coli serogroup O158 (organism)|	0	112283007|Escherichia coli (organism)|	0
20.	407187001|Escherichia coli serogroup O159 (organism)|	0	112283007|Escherichia coli (organism)|	0
21.	407244009|Escherichia coli serogroup O164 (organism)|	0	112283007|Escherichia coli (organism)|	0
22.	407304004|Escherichia coli serogroup O165 (organism)|	0	112283007|Escherichia coli (organism)|	0
23.	407193009|Escherichia coli serogroup O169 (organism)|	0	112283007|Escherichia coli (organism)|	0
24.	407261007|Escherichia coli serogroup O22 (organism)|	0	112283007|Escherichia coli (organism)|	0

No	Parent Concept	Concept Status	Child Concept	Concept Status
25.	407167002|Escherichia coli serogroup O25 (organism)|	0	112283007|Escherichia coli (organism)|	0
26.	131260002|Escherichia coli serogroup O26 (organism)|	0	112283007|Escherichia coli (organism)|	0
27.	407266002|Escherichia coli serogroup O45 (organism)|	0	112283007|Escherichia coli (organism)|	0
28.	407195002|Escherichia coli serogroup O55 (organism)|	0	112283007|Escherichia coli (organism)|	0
29.	407178005|Escherichia coli serogroup O78 (organism)|	0	112283007|Escherichia coli (organism)|	0
30.	407274001|Escherichia coli serogroup O83 (organism)|	0	112283007|Escherichia coli (organism)|	0
31.	279953009|Familial neonatal seizures (disorder)|	0	32895009|Hereditary disease (disorder)|	0
32.	442111003|Intermittent pneumatic compression stockings (physical object)|	0	303476002|Cardiovascular equipment (physical object)|	0
33.	301913002|Lesion of eyelid (finding)|	0	404684003|Clinical finding (finding)|	0
34.	59008008|Leukoencephalomyelopathy of Rottweilers (disorder)|	0	127326005|Non-human disorder (disorder)|	0
35.	246829005|Lid adhesions (finding)|	0	301913002|Lesion of eyelid (finding)|	0
36.	258560004|Oral saliva sample (specimen)|	0	441620008|Oral fluid specimen (specimen)|	0
37.	119342007|Saliva specimen (specimen)|	0	309051001|Body fluid sample (specimen)|	0
38.	1111000119100|Sebaceous nevus (disorder)|	2	239107007|Epidermal nevus (disorder)|	0

12.6 Appendix E: For Chapter Seven

12.6.1 Inconsistent or Incomplete Acronyms, Abbreviations and Synonyms

Table 12-33. Descriptions of the concept “13645005|Chronic obstructive lung disease (disorder)|”.

No	Acronym & Expanded Term	Acronym Only	Expanded Term Only
1.	475422013|CAFL - Chronic airflow limitation|		475425010|Chronic airflow limitation|
2.	475426011|CAL - Chronic airflow limitation|		
3.	475424014|CAO - Chronic airflow obstruction|		475421018|Chronic airflow obstruction|
4.	475423015|COAD - Chronic obstructive airways disease|		475428012|Chronic obstructive airway disease|
5.	475430014|COLD - Chronic obstructive lung disease|	23289016|COLD|	23287019|Chronic obstructive lung disease|
6.	475427019|COPD - Chronic obstructive pulmonary disease|	23290013|COPD|	475431013|Chronic obstructive pulmonary disease|
7.			475420017|Chronic irreversible airway obstruction|

Table 12-34. Different ways of describing cancer.

No	Type	363406005|Malignant tumor of colon (disorder)|	254837009|Malignant tumor of breast (disorder)|	363358000|Malignant tumor of lung (disorder)|	363484005|Malignant tumor of pelvis (disorder)|	363516004|Malignant tumor of penis (disorder)|
1.	Malignant tumour of x	482612014|Malignant tumor of colon|	379661016|Malignant tumor of breast|	482515017|Malignant tumor of lung|	482755017|Malignant tumor of pelvis|	482820014|Malignant tumor of penis|
2.	Cancer of x	1228536014|Cancer of colon|				1228631014|Cancer of penis|
3.	CA - Cancer of x	1228535013|CA - Cancer of colon|	379663018|CA - Breast cancer|	1228498010|CA - Lung cancer|		1228634018|CA - Cancer of penis|
4.	Ca x					1228633012|Ca penis|
5.	X cancer		379662011|Breast cancer|			
6.	X ca					1228632019|Penile Ca|

12.6.2 Frequency of Same Descriptions

Table 12-35. Frequency of descriptions with more than one occurrence by top-level hierarchy.

Hierarchy	Body structure	Clinical finding	Environment	Event	Linkage concept	Observable entity	Organism	Product	Physical force	Physical object	Procedure	Qualifier	Situation	Social context	Special concept	Specimen	Staging scales	Substance
Body structure	13,203	999	1	3	5	4	8			7	13	39				2		32
Clinical finding	999	1,506	1	8	6	111	30	2	1		177	53	7	3	1			11
Environmen t	1	1	4		1		4			1		2						1
Event	3	8		16			1				1	1						
Linkage	5	6	1		2	8	2			3	7	30		1				3
Observable	4	111			8	58	1			2	37	14		3	4		11	
Organism	8	30	4	1	2	1	502	8		2		4						97
Product		2					8	98			8	4						3,932
Physical force		1							2	3	2	2						
Physical object	7		1		3	2	2		3	38	10	33		1		1		6
Procedure	13	177		1	7	37		8	2	10	471	54	2		1		3	225
Qualifier	39	53	2	1	30	14	4	4	2	33	54	166	2	6		1	7	32
Situation		7									2	2	4					
Social context		3			1	3				1		6		22				
Special concept		1				4					1							
Specimen	2									1		1				4		4
Staging scale						11					3	7						
Substance	32	11	1		3		97	3,932		6	225	32				4		181

12.6.3 Examples Errors in Post-coordinated Expressions in the Literature

12.6.3.1 Syntax Error

In **Figure 12-6**,[5] there are multiple issues with the examples of post-coordination but the error that is highlighted here is that when nesting a refinement or qualification, parentheses should be used, and that there should only be one target concept per attribute. For the nesting error, **(** should follow **47429007|associated with|=** and **)** should follow **276339004|environment|** to indicate that that particular portion of the expression is being refined. For the target concept error, **410658008|triggered by|** should only have a single target and not two (i.e., **44190009|chemical|** and **276339004|environment|**. Without the parentheses and the multiple targets for a single attribute, the expression may be parsed incorrectly and the nested refinement or qualification may be omitted or misinterpreted.

Source terms on MCS	Normalized term	Post co-ordination
Multiple chemical sensitivity (disorder)	Chemical; multiple; senstivity	281867008|Multisystem Disorder|: {47429007|associated with|= 35209006|sensitivity|: 410658008|triggered by|= 441900009|chemical| 276339004|environment|}
Heightened sensitivity to environment (clinical finding)	Environment; heightened; sensitivity	35209006|sensitivity|: {37135001|tolerance related finding|=260365005|heightened|: 44190009|triggered by|= 276339004|environment|}

Figure 12-6. Examples of post-coordination in "An evaluation of SNOMED CT in the domain of complex chronic conditions."

12.6.3.2 Incorrect Type of Post-coordination

In **Figure 12-7**,[8] instead of using post-coordination combination, **128045006|Cellulitis (disorder)|** should be linked to **61685007|Lower limb structure (body structure)|** via **363698007|Finding site (attribute)|**. While there is little documentation on how post-coordinated combination expressions should be constructed, creating a post-coordinated expression between concepts from the **404684003|Clinical finding (finding)|** and **123037004|Body structure (body structure)|** should be through the use of the Concept Model attribute **363698007|Finding site (attribute)|**. Without the attribute, the relationship between **128045006|Cellulitis (disorder)|** and **61685007|Lower limb structure (body structure)|** may not be understood correctly.

> Where a term doesn't exist in the terminology it can be created by joining concepts together to represent the idea. For example:
>
> 128045008 = Cellulitis (disorder)
>
> 61685007 = Lower extremity = lower limb structure (body structure)
>
> *Therefore cellulitis of leg could be represented as*
>
> 128045008+61685007
>
> This is called Post-Coordination. The capacity to build concepts in this way is one of the strengths of a terminological system such as SNOMED-CT. However when representing clinical information it is important that the information be represented consistently and that consistency is highly supported by pre-coordination of the terms.

Figure 12-7. Example of post-coordination from "Clinical terminology."

In **Figure 12-8**,[11] the qualifier **255274008|Pre-bronchodilation (qualifier value)|** is not accessible via the Concept Model. The guidelines on post-coordination combination are limited but the use of a concept from the **362981000|Qualifier value (qualifier value)|** hierarchy as part of a combination expression is unlikely as they are usually used in post-coordination qualification. Similar to the previous example, without a Concept Model attribute linking the **71388002|Procedure (procedure)|** to the **362981000|Qualifier value (qualifier value)|**, the relationship between **127783003|Spirometry (procedure)|** and **255274008|Pre-bronchodilation (qualifier value)|** may not be understood correctly.

Clinical expression	Candidate SNOMED CT mappings	Assessment
Spirometry	Spirometry (procedure)	Precise expression mapping
	Measurement of respiratory function (procedure)	Lacking semantics
Without bronchodilation	Pre bronchodilation (qualifier value)	Exact expression mapping, however inadequate
	Spirometry (procedure) + Pre bronchodilation (qualifier value)	Precise and sufficient mapping due to the organizing element and the clinical expression

Figure 12-8. Example of post-coordination from "Construction of an interface terminology on SNOMED CT."

12.6.3.3 Incorrect Domain

In Figure 12-9,[12] **405175009|Parenting behavior (observable entity)|** cannot be post-coordinated with **254648000|Effective (qualifier value)|** because according to the Concept Model, concepts from the **363787002|Observable entity (observable entity)|** cannot be qualified.

> There were two types of ICNP catalogue concepts not mapped to SNOMED-CT concepts. The first group of concepts had a use case and the second group of concepts had no use case. For the first group of concepts, we proposed a parent concept of the SNOMED-CT under which the ICNP catalogue concept could be added. . However, these catalogue concepts can be post-coordinated by combining a concept from the Qualifier value hierarchy with a concept from either the Observable entity or Finding hierarchy. An example of post-coordination of a qualifier value concept and an observable entity concept is 'Effective Parenting' which can be post-coordinated with 'Effective (qualifier value)' and 'Parenting behavior (observable entity)'. An example of post-coordination of a qualifier value concept and a Finding concept is 'Enhanced Ability to Manage Regime' which can be post-coordinated with 'Increased (qualifier value)' and 'Finding related to therapeutic regimen management (finding)'.

Figure 12-9. Example of post-coordination in "Evaluation of the content coverage of SNOMED-CT to represent ICNP Version 1 catalogues."

12.6.3.4 Incorrect Attribute

In Figure 12-10,[13] **370131001|Recipient category (attribute)|** is used to link **3457005|Patient referral (procedure)|** to **66862007|Radiologist (occupation)|**. However, the ranges for **370131001|Recipient category (attribute)|** are **133928008|Community (social concept)|**, **105455006|Donor for medical or surgical procedure (person)|**, **35359004|Family (social concept)|**, **389109008|Group (social concept)|** or **125676002|Person (person)|**. In this case **66862007|Radiologist (occupation)|** is not a subtype of any of these concepts. Moreover, there are no Concept Model attributes that can be used to link to the range of **14679004|Occupation (occupation)|**.

Source Term	Post-coordinated Concept
Pain;mouth	22253000 pain (clinical finding) + 21082005 entire mouth region (body structure): relationship type = 363698007 finding site (attribute)
Referral;radiologist	3457005 patient referral (procedure) + 66862007 radiologist (occupation): relationship type = 370131001 recipient category (attribute)
Abuse;verbal ;relative	125677006 Relative (person) + 225825002 Victim of verbal abuse (clinical finding): relationship type = indeterminate
Dislocation;knee;simple	13673007 Simple (qualifier value) + 129156001 Traumatic dislocation of knee joint (clinical finding): relationship type = 246100006 onset (attribute)

Figure 12-10. Examples of post-coordination from "A computational linguistics motivated mapping of ICPC-2 PLUS to SNOMED CT."

12.6.3.5 Incorrect Range

In **Figure 12-11**,[14] in addition to **363819003|Drug therapy observable (observable entity)|** not being allowed to be qualified due to the incorrect domain, **246454002|Occurrence (attribute)|** has the range of **282032007|Periods of life (qualifier value)|** and **255216001|First (qualifier value)|** is not a subtype concept.

SNOMED CT
Figure 7 shows the number of DEs for each eMERGE site that were mapped to the SNOMED concepts via post-coordination that were not initially mapped via pre-coordination (figure 5). Since eleMAP does not support post-coordination at present, our process involved manually determining the best possible post-coordinated concept using the CliniClue SNOMED browser (http://www.cliniclue.com). Continuing with the above example, we can represent age of first statin use as follows:
363819003|drug therapy observable|:
{24645002|occurrence|=255216001|first|},
{127489000|has active ingredient|=6302009|HMG-CoA reductase inhibitor|},

Figure 12-11. Example of post-coordination in "Mapping clinical phenotype data elements to standardized metadata repositories and controlled terminologies: the eMERGE Network experience."

12.6.4 Generating Synonyms from Implicit Clinical Findings

Table 12-36. Examples of clinical terms that implicitly refer to a body structure.

No	Phrase	49755003|Morphologically abnormal structure (morphologic abnormality)|	404684003|Clinical finding (finding)|
1.	Basal cell carcinoma	1338007|Basal cell carcinoma (morphologic abnormality)| ▪ 3345018|Basal cell carcinoma| ▪ 738536018|Basal cell carcinoma (morphologic abnormality)| ▪ 3347014|Basal cell epithelioma| ▪ 3350012|Basiloma| ▪ 1220948017|BCC - Basal cell carcinoma| ▪ 3348016|Pigmented basal cell carcinoma| ▪ 3349012|Rodent ulcer|	254701007|Basal cell carcinoma of skin (disorder)| ▪ 379314010|Basal cell carcinoma of skin| ▪ 645632010|Basal cell carcinoma of skin (disorder)| ▪ 379316012|Basalioma| ▪ 379317015|BCC - Basal cell carcinoma of skin| ▪ 379315011|Rodent ulcer| ▪ 379318013|RU - Rodent ulcer|
2.	Tubulovillous adenoma	61722000|Tubulovillous adenoma (morphologic abnormality)| ▪ 102555015|Papillotubular adenoma| ▪ 102552017|Tubulovillous adenoma| ▪ 800816012|Tubulovillous adenoma (morphologic abnormality)| ▪ 102554016|Villoglandular adenoma|	449410007|Tubulovillous adenoma of anorectum (disorder)| ▪ 2902059017|Anorectal tubulovillous adenoma| ▪ 2902060010|Tubulovillous adenoma of anorectum| ▪ 2899234018|Tubulovillous adenoma of anorectum (disorder)|

No	Phrase	49755003|Morphologically abnormal structure (morphologic abnormality)|	404684003|Clinical finding (finding)|
3.	Intraductal carcinoma	86616005|Intraductal carcinoma, noninfiltrating, no ICD-O subtype (morphologic abnormality)| ▪ 201292017|DCIS| ▪ 201291012|DIN 3| ▪ 201290013|Ductal carcinoma in situ| ▪ 201289016|Ductal intraepithelial neoplasia 3| ▪ 143634017|Intraductal adenocarcinoma, noninfiltrating| ▪ 143635016|Intraductal carcinoma| ▪ 143630014|Intraductal carcinoma, noninfiltrating| ▪ 2774154015|Intraductal carcinoma, noninfiltrating, no ICD-O subtype| ▪ 1772492014|Intraductal carcinoma, noninfiltrating, no ICD-O subtype (morphologic abnormality)| ▪ 1235157017|Non-infiltrating intraductal adenocarcinoma| ▪ 1235156014|Non-infiltrating intraductal carcinoma|	278053004|Intraductal carcinoma of breast (disorder)| ▪ 2871985014|Intraductal carcinoma of breast| ▪ 2869248010|Intraductal carcinoma of breast (disorder)|
4.	Intussusception	35327006|Intussusception (morphologic abnormality)| ▪ 58951019|Introsusception| ▪ 58950018|Intussusception| ▪ 766954011|Intussusception (morphologic abnormality)|	49723003|Intussusception of intestine (disorder)| ▪ 1230768010|Intestinal intussusception| ▪ 82816014|Intussusception of intestine| ▪ 787433010|Intussusception of intestine (disorder)| ▪ 1230767017|Intussusception of the intestine| ▪ 82817017|Invagination of intestine| ▪ 1230766014|ISN - Intussusception|

12.6.5 Example of Concepts with Multiple Body Structures

427061007|Computerized tomography of both ankles (procedure)|

427061007|Computerized tomography of both ankles (procedure)|:
{260686004|Method (attribute)|=
312251004|Computed tomography imaging - action (qualifier value)|,
405813007|Procedure site - Direct (attribute)|=
51636004|Structure of left ankle (body structure)|}
{260686004|Method (attribute)|=
312251004|Computed tomography imaging - action (qualifier value)|,
405813007|Procedure site - Direct (attribute)|=
6685009|Structure of right ankle (body structure)|}

271473006|Benign neoplasm of bronchus and lung (disorder)|

271473006|Benign neoplasm of bronchus and lung (disorder)|:
{116676008|Associated morphology (attribute)|=
3898006|Neoplasm, benign (morphologic abnormality)|,
363698007|Finding site (attribute)|=
39607008|Lung structure (body structure)|}
{116676008|Associated morphology (attribute)|=
3898006|Neoplasm, benign (morphologic abnormality)|,
363698007|Finding site (attribute)|=
955009|Bronchial structure (body structure)|}

112763005|Osteotomy of tibia and fibula (procedure)|

112763005|Osteotomy of tibia and fibula (procedure)|:
{260686004|Method (attribute)|=

```
        299713009|Osteotomy - action (qualifier value)|,
    405813007|Procedure site - Direct (attribute)|=
        12611008|Bone structure of tibia (body structure)|}
    {260686004|Method (attribute)|=
        299713009|Osteotomy - action (qualifier value)|,
    405813007|Procedure site - Direct (attribute)|=
        87342007|Bone structure of fibula (body structure)|}
```

Figure 12-12. Examples of concepts that refer to multiple body structures.

12.6.6 Frequency Count of Concepts with the Same Description within the Same Hierarchy

Table 12-37. Frequency count of concepts with the same descriptions within the same hierarchy.

No	Hierarchy	Have "Is A" Relationship	Do Not Have "Is A" Relationship	Total
1.	123037004|Body structure (body structure)|	6,349	204	6,553
2.	404684003|Clinical finding (finding)|	210	533	743
3.	308916002|Environment or geographical location (environment / location)|	0	2	2
4.	272379006|Event (event)|	3	5	8
5.	106237007|Linkage concept (linkage concept)|	0	1	1
6.	363787002|Observable entity (observable entity)|	4	24	28
7.	410607006|Organism (organism)|	122	119	241
8.	373873005|Pharmaceutical / biologic product (product)|	6	43	49
9.	78621006|Physical force (physical force)|	1	0	1
10.	260787004|Physical object (physical object)|	9	10	19
11.	71388002|Procedure (procedure)|	106	126	232
12.	362981000|Qualifier value (qualifier value)|	4	77	81
13.	419891008|Record artifact (record artifact)|	0	0	0
14.	243796009|Situation with explicit context (situation)|	2	0	2
15.	48176007|Social context (social concept)|	1	10	11
16.	370115009|Special concept (special concept)|	0	0	0
17.	123038009|Specimen (specimen)|	2	0	2
18.	254291000|Staging and scales (staging scale)|	0	0	0
19.	105590001|Substance (substance)|	15	74	89
	Totals	**6,834**	**1,228**	**8,062**

12.6.7 Supplementary Modeling Scenario Details

12.6.7.1 Candidate Is Laterality and Predicate Includes Body Structure

According to the Concept Model, laterality can only be applied to any concepts from the **123037004|Body structure (body structure)|** hierarchy. While that may be true, it does not necessarily mean that laterality must be applied to all defining attributes that have a range of **123037004|Body structure (body structure)|**. Take for example the term "monoparesis of right arm." As there is no pre-coordinated concept, it is necessary to encode "monoparesis of arm" and then apply the laterality of "right." The concept **249944006|Monoparesis - arm (disorder)|** has two sets of defining attributes (refer to **Figure 12-13**). In this case, laterality should be applied to **53120007|Upper limb structure (body structure)|** but probably not to **25087005|Structure of nervous system (body structure)|**. There does not appear to be any rules in place about which **123037004|Body structure (body structure)|** concepts laterality should be applied to. One possibility, although not verified, may be to apply laterality to the **123037004|Body structure (body structure)|** concepts that have a qualifying relationship of **272741003|Laterality (attribute)|=182353008|Side (qualifier value)|**.

```
249944006|Monoparesis - arm (disorder)|:
    {363698007|Finding site (attribute)|=
        25087005|Structure of nervous system (body structure)|,
    363698007|Finding site (attribute)|=
        53120007|Upper limb structure (body structure)|}
```

Figure 12-13. Defining attributes for "249944006|Monoparesis - arm (disorder)|".

12.6.7.2 Candidate is a Predicate's Defining Attribute

For example, the term "muscle spasm" does not exist in SNOMED CT and could be individually encoded as **71616004|Skeletal and/or smooth muscle structure (body structure)|**, which has a synonym of **118966016|Muscle|**, and **45352006|Spasm (finding)|**. In this case, it is implicit that **45352006|Spasm (finding)|** refers to "muscle" as evident of the defining attribute **71616004|Skeletal and/or smooth muscle structure (body structure)|**. Therefore in this case there is no need to merge **71616004|Skeletal and/or smooth muscle structure (body structure)|** with **45352006|Spasm (finding)|** because it is already part of the defining attributes.

12.6.7.3 Candidate is a Refinement of Predicate's Defining Attribute

For example, the concept "fracture of sixth rib" does not exist in SNOMED CT and could be individually encoded as **125605004|Fracture of bone (disorder)|** (using description **473534016|Fracture|**) and **59558009|Bone structure of sixth rib (body structure)|** (using description **98922014|Sixth rib|**). In this case, it can be determined that **59558009|Bone structure of sixth rib (body structure)|** is from the **123037004|Body structure (body structure)|** hierarchy and it can only be linked to a **404684003|Clinical finding (finding)|** via the attribute **363698007|Finding site (attribute)|**. **125605004|Fracture of bone (disorder)|** should be queried to determine if there is a defining attribute using **363698007|Finding site (attribute)|**, which in this case is **272673000|Bone structure (body structure)|**. **59558009|Bone structure of sixth rib (body structure)|** is then checked to see if it is a subtype of **272673000|Bone structure (body structure)|**.

12.6.7.4 Candidate Uses Same Attribute of Predicate's Defining Attribute

For example, the concept "tingling lip" does not exist in SNOMED CT and could be individually encoded as **62507009|Pins and needles (finding)|** (using description **103892015|Tingling|**) and **48477009|Lip structure (body**

structure)| (using description **80770013|Lip|**). The defining attributes for **62507009|Pins and needles (finding)|** are shown in **Table 12-19**.

```
62507009|Pins and needles (finding)|:
    363698007|Finding site (attribute)|=
        39937001|Skin structure (body structure)|
```

Figure 12-14. Defining attributes for "62507009|Pins and needles (finding)|".

In this case, **48477009|Lip structure (body structure)|** is not a subtype concept of **39937001|Skin structure (body structure)|**. Therefore a new search should be conducted for "lip" that is also subtype concept of **39937001|Skin structure (body structure)|**. The search in Figure 12-15 resulted in 10 concepts, which when aggregated, resulted in **88089004|Skin structure of lip (body structure)|** and **85193007|Mucocutaneous junction of lip (body structure)|**. As there were two matches, one possibility would be to compare the difference between the fully specified name of **39937001|Skin structure (body structure)|** with **88089004|Skin structure of lip (body structure)|** and **85193007|Mucocutaneous junction of lip (body structure)|**. In this case, **88089004|Skin structure of lip (body structure)|** is a closer match. Refer to **Section 7.3.1.3** on how to sort search results.

```
SELECT C.ConceptId, C.FullySpecifiedName, D.DescriptionId, D.Term FROM SCT_Concepts C,
SCT_Descriptions D, SCT_TransitiveClosure TC WHERE TC.SuperTypeId IN (39937001) AND
TC.SubTypeId=C.ConceptId AND C.ConceptStatus=0 AND C.ConceptId=D.ConceptId AND D.DescriptionStatus=0
AND MATCH(D.Term) AGAINST ('lip')
```

Figure 12-15. Search for "lip" in subtype of "39937001|Skin structure (body structure)|".

12.6.7.5 Candidate is a Body Structure and Predicate is a Finding or Procedure

To illustrate the use of the Extended Concept Model, take the concept **399963005|Abrasion (disorder)|** for example, which has only one defining attribute of **116676008|Associated morphology (attribute)|=400061001|Abrasion (morphologic abnormality)|**. Since **399963005|Abrasion (disorder)|** does not have a defining attribute of **363698007|Finding site (attribute)|**, if we want to specify a site, we will use the Concept Model (SNOMED CT User Guide, page 26) attributes for the **404684003|Clinical finding (finding)|** hierarchy, which is **363698007|Finding site (attribute)|<< 442083009|Anatomical or acquired body structure (body structure)|.** The range is very broad and contains nearly 26,000 possible concepts, most of which are not applicable to represent an abrasion. What we can do is analyse how other "abrasion" concepts have been modelled to determine what type of body structures are used.

By comparing **116676008|Associated morphology (attribute)|=400061001|Abrasion (morphologic abnormality)|** to the concepts used in conjunction with the range of **363698007|Finding site (attribute)|**, we can determine that there are 121 concepts that are made up of 85 unique body structures. By aggregating the body structures, we can determine that there are basically six concepts from which a concept that is defined using **400061001|Abrasion (morphologic abnormality)|** uses a **363698007|Finding site (attribute)|**.

Table 12-38. Finding sites used in accordance with "400061001|Abrasion (morphologic abnormality)|" as part of the defining attributes.

No	Concept	Rolled Up	Total Subtypes
1.	30291003|Chin structure (body structure)|	1	11
2.	371398005|Eye region structure (body structure)|	5	898
3.	74262004|Oral cavity structure (body structure)|	15	927
4.	31389004|Oropharyngeal structure (body structure)|	2	75
5.	39937001|Skin structure (body structure)|	38	962
6.	22943007|Trunk structure (body structure)|	24	9054

The use of these constraints becomes apparent when trying to represent a clinical statement that does not have a pre-coordinated concept. For example, if we were to encode the phrase "abrasion of temple," we could try an exact match or match all and not return any result as there is no concept is SNOMED CT that contains both the words "abrasion" and "temple" in a single description. By breaking down the phrase into single words or by using the word "of" as a delimiter, we could encode both words individually and then locate an appropriate Concept Model attribute to link both concepts together. A search for the word "abrasion" will yield three exact matches from three different hierarchies: **8420001|Abrasion (procedure)|**, **400061001|Abrasion (morphologic abnormality)|** and **399963005|Abrasion (disorder)|**. Assuming that the context is a problem (as opposed to a procedure), we would select the **404684003|Clinical finding (finding)|** concept (as opposed to the **71388002|Procedure (procedure)|** or **123037004|Body structure (body structure)|**). If we take the same approach to "temple," the exact match would yield the description **52325016|Temple|** from the concept **31287004|Temporal region structure (body structure)|**. Using the Concept Model, we can determine that there is only one attribute that can link a **404684003|Clinical finding (finding)|** to a **123037004|Body structure (body structure)|**, and that is through a **363698007|Finding site (attribute)|** to form the expression **399963005|Abrasion (disorder)|:363698007|Finding site (attribute)|=31287004|Temporal region structure (body structure)|**.

While the encoding conforms to the Concept Model and is a valid post-coordinated expression, it is semantically incorrect because the wrong "type" of temple is used. As we have previously constrained the finding sites that are used in relation to **400061001|Abrasion (morphologic abnormality)|** to one of the five supertype concepts, we can perform an exact or partial search (refer to **Figure 12-16**).

```
SELECT C.ConceptId, C.FullySpecifiedName, D.DescriptionId, D.Term FROM SCT_Concepts C,
SCT_Descriptions D, SCT_TransitiveClosure TC WHERE TC.SuperTypeId IN (371398005, 74262004, 31389004,
39937001, 22943007) AND TC.SubTypeId=C.ConceptId AND C.ConceptStatus=0 AND C.ConceptId=D.ConceptId
AND D.DescriptionStatus=0 AND MATCH(D.Term) AGAINST ('temple')
```

Figure 12-16. Search for "temple"

In this case, there is only match, **244081009|Skin of temple (body structure)|**. Therefore we can encode "abrasion of temple" using the appropriate body structure, **399963005|Abrasion (disorder)|:363698007|Finding site**

(attribute)|=244081009|Skin of temple (body structure)|. When there are multiple matches, a similar set of rules as laid out in **Section 7.3.1.3** may be applied.

12.6.7.6 Encoding Template Exists for Predicate and Candidate

12.6.7.6.1 Abnormal, Normal, Increased and Decreased

The words "abnormal," "normal," "increased/elevated" and "decreased" are often used to describe the results of tests but encoding them is not straightforward as there are not always pre-coordinated concepts and the qualifier values such as **263654008|Abnormal (qualifier value)|** can only be applied to concepts from the **404684003|Clinical finding (finding)|** hierarchy while the test is usually matched to a concept from the **71388002|Procedure (procedure)|** hierarchy. For example, encoding "physical exam abnormal" could be broken down to **410182004|Physical examination assessment (procedure)|** and **263654008|Abnormal (qualifier value)|** but the latter concept cannot be affixed to the former concept. In this case **442618008|Abnormal finding on evaluation procedure (finding)|** should be used as the focus concept and **363714003|Interprets (attribute)|** should be refined from **386053000|Evaluation procedure (procedure)|** to **410182004|Physical examination assessment (procedure)|**. Another example is "Dilantin level elevated." "Dilantin level" can be encoded with **33748008|Phenytoin measurement (procedure)|**. "Elevated" or "high" **75540009|High (qualifier value)|** can be used in association with a **404684003|Clinical finding (finding)|** via **363713009|Has interpretation (attribute)|**. However, this is not appropriate even though it is an exact match. When dealing with measurements, if the result is "elevated" or "high," it should be qualified with **281302008|Above reference range (qualifier value)|** and if it is "decreased" or "low," it should be qualified with **281300000|Below reference range (qualifier value)|** as opposed to **1250004|Decreased (qualifier value)|** or **62482003|Low (qualifier value)|** as that is how other concepts with "high" or "low" measurements are defined in SNOMED CT. To simplify this process, templates are needed. There are three supertype concepts that can be used in assisting in the post-coordination (refer to **Table 12-39**). All three concepts use **260245000|Findings values (qualifier value)|** as the range and is linked to **363713009|Has interpretation (attribute)|**.

Table 12-39. Representing evaluation results.

No	441742003|Evaluation finding (finding)|	363713009|Has interpretation (attribute)|	363714003|Interprets (attribute)|
1.	442618008|Abnormal finding on evaluation procedure (finding)|	263654008|Abnormal (qualifier value)|	386053000|Evaluation procedure (procedure)|
2.	442756004|Measurement finding above reference range (finding)|	281302008|Above reference range (qualifier value)|	122869004|Measurement procedure (procedure)|
3.	442686002|Measurement finding below reference range (finding)|	281300000|Below reference range (qualifier value)|	122869004|Measurement procedure (procedure)|

Unlike "abnormal" findings, there is no supertype concept that represents "normal" results. There are only 15 different concepts that have the word "normal" in one of their descriptions, with 12 of them occurring in the **404684003|Clinical finding (finding)|** hierarchy (refer to **Figure 12-17**).

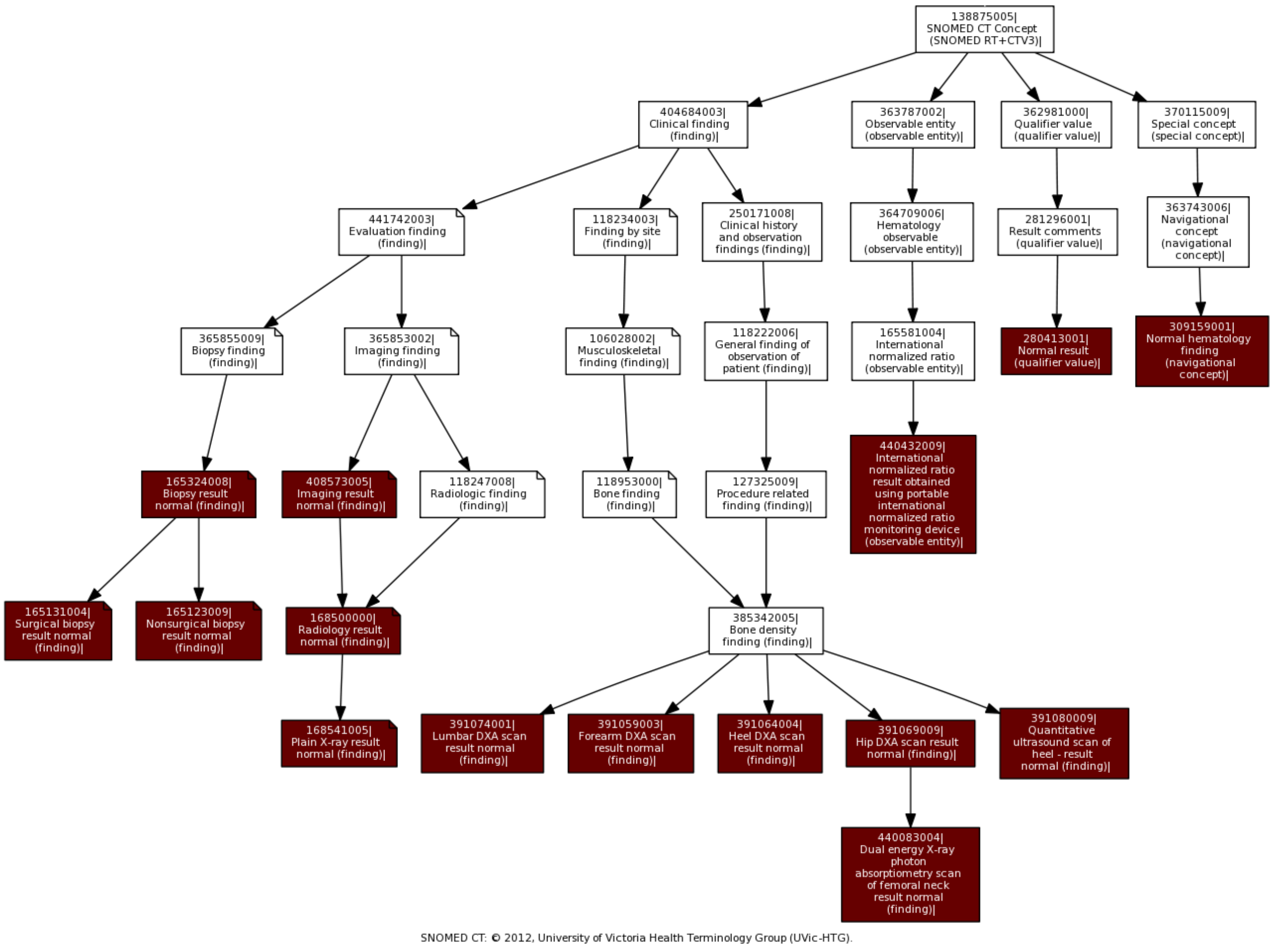


Figure 12-17. Location of concepts in the hierarchy with the word "normal" in the description.

The concept **442618008|Abnormal finding on evaluation procedure (finding)|** is a fully-defined concept and sets a precedent on how to model "normal" evaluation procedures (refer to **Figure 12-18**). As **17621005|Normal (qualifier value)|** is the opposite of **263654008|Abnormal (qualifier value)|**, it is reasonable to assume that that is the appropriate concept for representing "normal" findings.

```
404684003|Clinical finding (finding)|:
    {363713009|Has interpretation (attribute)|=
        263654008|Abnormal (qualifier value)|,
    363714003|Interprets (attribute)|=(
        71388002|Procedure (procedure)|:
            260686004|Method (attribute)|=
                129265001|Evaluation - action (qualifier value)|)}
```

Figure 12-18. Long normal form of "442618008|Abnormal finding on evaluation procedure (finding)|".

Therefore a template to encode abnormal, normal, elevated and decreased findings would be as shown in **Figure 12-19.**

Abnormal	442618008|Abnormal finding on evaluation procedure (finding)|: 363714003|Interprets (attribute)|<<

	386053000|Evaluation procedure (procedure)|
Normal	404684003|Clinical finding (finding)|: {363713009|Has interpretation (attribute)|= 17621005|Normal (qualifier value)|, 363714003|Interprets (attribute)|<< 386053000|Evaluation procedure (procedure)|}
Increased (Elevated)	442756004|Measurement finding above reference range (finding)|: 363714003|Interprets (attribute)|<< 122869004|Measurement procedure (procedure)|
Decreased	442686002|Measurement finding below reference range (finding)|: 363714003|Interprets (attribute)|<< 122869004|Measurement procedure (procedure)|

Figure 12-19. Template for encoding abnormal, normal, increased and decreased findings.

12.6.7.6.2 Allergies

Allergies are generally recorded using the concept **420134006|Propensity to adverse reactions (disorder)|** or one of its subtypes. In order to determine how to represent an allergy using SNOMED CT, it is important to understand how allergies are current defined by SNOMED CT. There are 1,644 subtype concepts and are defined with a combination of eight different attributes (refer to **Table 12-40**).

Table 12-40. Defining attributes used in concept definition of subtype concepts of "420134006|Propensity to adverse reactions (disorder)|".

No	Attribute	Frequency	Range
1.	255234002|After (attribute)|	1	▪ 40001004|Smallpox vaccination (procedure)|
2.	116676008|Associated morphology (attribute)|	3	▪ 75889009|Acute and chronic inflammation (morphologic abnormality)| (2) ▪ 23583003|Inflammation (morphologic abnormality)|
3.	47429007|Associated with (attribute)|	7	▪ 300910009|Allergy to pollen (disorder)| ▪ 108241001|Dialysis procedure (procedure)| (4) ▪ 68332000|Digoxin immune fab (substance)| ▪ 41778007|Patch test (procedure)|
4.	246075003|Causative agent (attribute)|	1,708	▪ 1533
5.	263502005|Clinical course (attribute)|	4	▪ 21864008|Seasonal (qualifier value)|
6.	42752001|Due to (attribute)|	10	▪ 418367004|Allergic reaction to grass pollen (disorder)| ▪ 418364006|Allergic reaction to pollen (disorder)| ▪ 418943003|Allergic reaction to tree pollen (disorder)| ▪ 55985003|Atopic reaction (disorder)| (3) ▪ 418925002|Immune hypersensitivity reaction (disorder)| ▪ 12263007|Type 1 hypersensitivity response (disorder)| (3)
7.	363698007|Finding site (attribute)|	10	▪ 122865005|Gastrointestinal tract structure (body structure)| ▪ 39937001|Skin structure (body structure)| ▪ 30315005|Small intestinal structure (body structure)| (5) ▪ 34508005|Structure of mucous membrane of nose (body structure)| (3)
8.	363705008|Has definitional manifestation (attribute)|	1,473	▪ 416093006|Allergic reaction to drug (disorder)| (1,473)

A closer inspection of the 35 concepts with lower frequency counts of the defining attributes shows that it can be accounted for by 20 unique concepts (refer to **Table 12-48**). All other concepts are defined using

246075003|Causative agent (attribute)| and **363705008|Has definitional manifestation (attribute)|**. Therefore if a term refers to an "allergy of (substance)," where substance is one of the subtype concepts of **105590001|Substance (substance)|**, the allergy template in **Figure 12-20** may be used.

```
420134006|Propensity to adverse reactions (disorder)|:
        246075003|Causative agent (attribute)|<<
                105590001|Substance (substance)|
```

Figure 12-20. Proposed template to record general allergies.

Drug allergies require a slightly different template because they are currently also defined using **363705008|Has definitional manifestation (attribute)|=416093006|Allergic reaction to drug (disorder)|**. Therefore a suitable template shown in **Figure 12-21** may be used. The way to differentiate between a "regular" allergy and a drug allergy is to determine whether the **105590001|Substance (substance)|** is a subtype of **410942007|Drug or medicament (substance)|**. If it is, the drug allergy template should be used instead.

```
416098002|Drug allergy (disorder)|:
        246075003|Causative agent (attribute)|<<
                410942007|Drug or medicament (substance)|
```

Figure 12-21. Proposed template to record drug allergies.

12.6.7.6.3 Screening

The screening for clinical conditions is also common in clinical records (e.g., "TB screen," "STD screen, "MRSA screen," "STI screen"). If encoded incorrectly, these terms may indicate that a clinical condition is present whereas only a screening procedure is merited. An example of such a template can be seen in **Figure 12-22.** In this case, the trigger words are "screen" and "screening," and are followed by a concept that is a **404684003|Clinical finding (finding)|**.

```
20135006|Screening procedure (procedure)|:
        363702006|Has focus (attribute)|<<
                404684003|Clinical finding (finding)|
```

Figure 12-22. Proposed template to record that screening procedure.

12.6.7.6.4 Declined

There are 72 SNOMED CT concepts with "declined" in the textual descriptions and they are spread out over three hierarchies: **404684003|Clinical finding (finding)|**, **71388002|Procedure (procedure)|** and **243796009|Situation with explicit context (situation)|**. There is only one concept from the **71388002|Procedure (procedure)|** hierarchy, **413162002|Learning disabilities health action plan declined (procedure)|**, which is a primitive concept and offers no defining attributes that can help in the modelling process. The 11 concepts from the **404684003|Clinical finding (finding)|** are shown in **Table 12-41.** Nine of the 11 concepts are subtypes of **307824009|Administrative statuses (finding)|** while two are subtypes of **309298003|Drug therapy finding (finding)|**. All 11 concepts are primitive and also do not include any defining attributes.

Table 12-41. Concepts from the 404684003|Clinical finding (finding)| hierarchy that have a description of "declined."

No	Concept
1.	390845001|Breast screening declined (finding)|
2.	310881009|Cholesterol reduction program - declined (finding)|
3.	408775001|Consent for postmortem declined (finding)|
4.	407653007|Environmental risk assessment declined (finding)|
5.	391145007|HPV - Human papillomavirus test declined (finding)|
6.	414657005|Medication dosage increase declined (finding)|
7.	373812008|Patient declined anti-cancer treatment (finding)|
8.	399250008|Patient declined clinical trial (finding)|
9.	416135000|Patient held diabetic record declined (finding)|
10.	407662009|Personal risk assessment declined (finding)|
11.	415523003|Shared care prescribing declined (finding)|

The focus is now on the remaining 61 concepts that are from the **243796009|Situation with explicit context (situation)|** hierarchy. A closer inspection of the hierarchy shows that they are all subtype concepts of **183944003|Procedure refused (situation)|**. An inspection of the 99 subtype concepts shows that the negation keywords are "refused" and declined" and that the defining attribute used to indicate this are **443390004|Refused (qualifier value)|** or **385648002|Rejected by recipient (qualifier value)|**. The proposed template is shown in **Figure 12-23**.

```
183944003|Procedure refused (situation)|:
        363589002|Associated procedure (attribute)| <<
                71388002|Procedure (procedure)|
```

Figure 12-23. Proposed template to record that a procedure has been "declined."

12.6.7.6.5 Resolved

The word "resolved" is only used in eleven SNOMED CT concepts, eight times as a **404684003|Clinical finding (finding)|** and once as a **64572001|Disease (disorder)|**, **71388002|Procedure (procedure)|** and **243796009|Situation with explicit context (situation)|**. As the disorder, procedure and situation with explicit context only appear once, it is difficult to determine how a "resolved" concept can be modelled. Therefore the focus will centre on the finding concepts (refer to **Figure 12-24**).

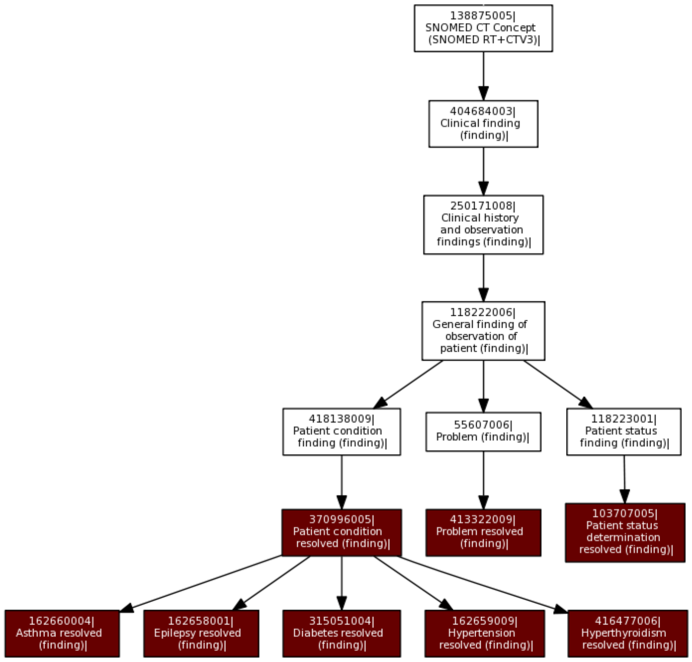


Figure 12-24. Finding SNOMED CT concepts that have the word "resolved" in the textual description.

The concepts of interest are the five subtype concepts of **370996005|Patient condition resolved (finding)|**. Unfortunately none of the five concepts have any defining attributes that can help shed light on how to model the defining attributes. In this case, the condition that is resolved is likely a **404684003|Clinical finding (finding)|**. Since **370996005|Patient condition resolved (finding)|** is also a **404684003|Clinical finding (finding)|**, there are only four Concept Model attributes that can be used to link a **404684003|Clinical finding (finding)|** with another **404684003|Clinical finding (finding)|** (i.e., **255234002|After (attribute)|**, **47429007|Associated with (attribute)|**, **42752001|Due to (attribute)|** and **363705008|Has definitional manifestation (attribute)|**). In this instance, it would appear that only **47429007|Associated with (attribute)|** is suitable. Therefore to record a clinical condition that has been resolved, the proposed template is shown in **Figure 12-25**, whereby **404684003|Clinical finding (finding)|** is replaced with the relevant concept. It should be noted that this is a temporary solution until the IHTSDO proposes a different solution.

```
370996005|Patient condition resolved (finding)|:
    47429007|Associated with (attribute)| <<
        404684003|Clinical finding (finding)|
```

Figure 12-25. Proposed template to record a clinical condition that has been "resolved."

12.6.7.6.6 Obstetrics History – Gravida Para Abortus

When encoding primary care data, obstetrics history is often encountered and is usually stored using the format of GNPNAN where G refers to gravida, P to Para, A to Abortus and N to the number of occurrences. For example, G2P1A1 means that the patient was pregnant twice with one live birth and one abortion or miscarriage. Typical encoding algorithms will not be able to encode the short form, therefore the use of regular expressions is necessary (e.g., **\b(g([\d]|∅){1,2} ?,? ?p([\d]|∅){1,2}(?,? ?(a([\d]|∅){1,2}))?)|(p([\d]|∅){1,2} ?,? ?g([\d]|∅){1,2} ?,? ?(a([\d]|∅){1,2}))\b**). Each part of the short form can then be encoded individually (e.g., G2 can be encoded with **127365008|Gravida 2 (finding)|**). Refer to **Table 12-42** for the list of gravida, para and abortus concepts.

Table 12-42. Gravida, para and abortus concepts.

<table>
<tr><th>No</th><th>Number</th><th>Gravida</th><th>Para</th><th>Abortus</th></tr>
<tr><td>1.</td><td>0 / Null</td><td>53881005|Gravida 0 (finding)|</td><td>102877006|Nulliparous (finding)|</td><td>44638009|Abortion 0 (finding)|</td></tr>
<tr><td>2.</td><td>One</td><td>127364007|Primigravida (finding)|</td><td>14209008|Para 1 (finding)|</td><td>38835003|Abortion number 1 (finding)|</td></tr>
<tr><td>3.</td><td>Two</td><td>127365008|Gravida 2 (finding)|</td><td>48395006|Para 2 (finding)|</td><td>28493005|Abortion number 2 (finding)|</td></tr>
<tr><td>4.</td><td>Three</td><td>127366009|Gravida 3 (finding)|</td><td>49955008|Para 3 (finding)|</td><td>56029000|Abortion number 3 (finding)|</td></tr>
<tr><td>5.</td><td>Four</td><td>127367000|Gravida 4 (finding)|</td><td>42279001|Para 4 (finding)|</td><td>76491002|Abortion number 4 (finding)|</td></tr>
<tr><td>6.</td><td>Five</td><td>127368005|Gravida 5 (finding)|</td><td>13868004|Para 5 (finding)|</td><td>9402004|Abortion number 5 (finding)|</td></tr>
<tr><td>7.</td><td>Six</td><td>127369002|Gravida 6 (finding)|</td><td>26886008|Para 6 (finding)|</td><td>62331000|Abortion number 6 (finding)|</td></tr>
<tr><td>8.</td><td>Seven</td><td>127370001|Gravida 7 (finding)|</td><td>26611006|Para 7 (finding)|</td><td>44108004|Abortion number 7 (finding)|</td></tr>
<tr><td>9.</td><td>Eight</td><td>127371002|Gravida 8 (finding)|</td><td>36614009|Parity eight (finding)|</td><td>83712005|Abortion number 8 (finding)|</td></tr>
<tr><td>10.</td><td>Nine</td><td>127372009|Gravida 9 (finding)|</td><td>81514003|Parity nine (finding)|</td><td>57197002|Abortion number 9 (finding)|</td></tr>
<tr><td>11.</td><td>Ten</td><td>127373004|Gravida 10 (finding)|</td><td>23037000|Parity ten or more (finding)|</td><td>4426000|Abortion number 10 (finding)|</td></tr>
<tr><td>12.</td><td>Eleven</td><td>127374005|Gravida more than 10 (finding)|</td><td></td><td>81131001|Abortion number 11 (finding)|</td></tr>
<tr><td>13.</td><td>Twelve</td><td></td><td></td><td>31227001|Abortion number 12 (finding)|</td></tr>
</table>

12.6.7.7 Concept Model Contains Appropriate Attribute

Table 12-43. Domains and ranges that have multiple attributes that link the two together.

<table>
<tr><th>No</th><th>Domain</th><th>Attributes</th><th>Range</th><th>Frequency</th><th>Percent</th><th>Default</th></tr>
<tr><td rowspan="4">1.</td><td rowspan="4">404684003|Clinical finding (finding)|</td><td>363705008|Has definitional manifestation (attribute)|</td><td rowspan="4">404684003|Clinical finding (finding)|</td><td>5,287</td><td>59.9%</td><td rowspan="4">N/A</td></tr>
<tr><td>42752001|Due to (attribute)|</td><td>1,661</td><td>18.8%</td></tr>
<tr><td>47429007|Associated with (attribute)|</td><td>1,141</td><td>12.9%</td></tr>
<tr><td>255234002|After (attribute)|</td><td>731</td><td>8.3%</td></tr>
<tr><td rowspan="2">2.</td><td rowspan="2">404684003|Clinical finding (finding)|</td><td>42752001|Due to (attribute)|</td><td rowspan="2">272379006|Event (event)|</td><td>120</td><td>92.3%</td><td>Y</td></tr>
<tr><td>47429007|Associated with (attribute)|</td><td>10</td><td>7.7%</td><td></td></tr>
<tr><td rowspan="2">3.</td><td rowspan="2">404684003|Clinical finding (finding)|</td><td>246075003|Causative agent (attribute)|</td><td rowspan="2">410607006|Organism (organism)|</td><td>5,942</td><td>99.98%</td><td>Y</td></tr>
<tr><td>47429007|Associated with (attribute)|</td><td>1</td><td>0.02%</td><td></td></tr>
</table>

No	Domain	Attributes	Range	Frequency	Percent	Default
4.	404684003|Clinical finding (finding)|	47429007|Associated with (attribute)|	373873005|Pharmaceutical / biologic product (product)|	0	0.0%	N/A
		246075003|Causative agent (attribute)|		0	0.0%	
5.	404684003|Clinical finding (finding)|	246075003|Causative agent (attribute)|	78621006|Physical force (physical force)|	629	100.0%	Y
		47429007|Associated with (attribute)|		0	0.0%	
6.	404684003|Clinical finding (finding)|	246075003|Causative agent (attribute)|	260787004|Physical object (physical object)|	202	93.5%	Y
		47429007|Associated with (attribute)|		14	6.5%	
7.	404684003|Clinical finding (finding)|	363714003|Interprets (attribute)|	71388002|Procedure (procedure)|	5,578	50.4%	N/A
		418775008|Finding method (attribute)|		3,843	34.7%	
		255234002|After (attribute)|		942	8.5%	
		47429007|Associated with (attribute)|		703	6.4%	
8.	404684003|Clinical finding (finding)|	47429007|Associated with (attribute)|	138875005|SNOMED CT Concept (SNOMED RT+CTV3)|	0	0.0%	N/A
		246075003|Causative agent (attribute)|		0	0.0%	
9.	404684003|Clinical finding (finding)|	246075003|Causative agent (attribute)|	105590001|Substance (substance)|	10,118	99.3%	Y
		47429007|Associated with (attribute)|		74	0.7%	
10.	272379006|Event (event)|	42752001|Due to (attribute)|	404684003|Clinical finding (finding)|	20	95.2%	Y
		47429007|Associated with (attribute)|		1	4.8%	
		255234002|After (attribute)|		0	0.0%	
11.	272379006|Event (event)|	47429007|Associated with (attribute)|	272379006|Event (event)|	0	0.0%	N/A
		42752001|Due to (attribute)|		5	100.0%	
12.	272379006|Event (event)|	47429007|Associated with (attribute)|	410607006|Organism (organism)|	36	100.0%	N/A
		246075003|Causative agent (attribute)|		0	0.0%	
13.	272379006|Event (event)|	47429007|Associated with (attribute)|	373873005|Pharmaceutical / biologic product (product)|	0	0.0%	N/A
		246075003|Causative agent (attribute)|		0	0.0%	
14.	272379006|Event (event)|	246075003|Causative agent (attribute)|	78621006|Physical force (physical force)|	3	100.0%	N/A
		47429007|Associated with (attribute)|		0	0.0%	
15.	272379006|Event (event)|	47429007|Associated with (attribute)|	260787004|Physical object (physical object)|	0	0.0%	N/A
		246075003|Causative agent (attribute)|		0	0.0%	
16.	272379006|Event (event)|	255234002|After (attribute)|	71388002|Procedure (procedure)|	0	0.0%	N/A
		47429007|Associated with (attribute)|		0	0.0%	
17.	272379006|Event (event)|	47429007|Associated with (attribute)|	138875005|SNOMED CT Concept (SNOMED RT+CTV3)|	0	0.0%	N/A
		246075003|Causative agent (attribute)|		0	0.0%	
18.	272379006|Event (event)|	47429007|Associated with (attribute)|	105590001|Substance (substance)|	8	100.0%	N/A
		246075003|Causative agent (attribute)|		0	0.0%	
19.	71388002|Procedure (procedure)|	405813007|Procedure site - Direct (attribute)|	442083009|Anatomical or acquired body structure (body	31,092	72.0%	N/A
		405814001|Procedure site - Indirect		7,945	18.4%	

No	Domain	Attributes	Range	Frequency	Percent	Default
		(attribute)|	structure)|			
		363704007|Procedure site (attribute)|		3,449	8.0%	
		363700003|Direct morphology (attribute)|		378	0.9%	
		246093002|Component (attribute)|		217	0.5%	
		363709002|Indirect morphology (attribute)|		47	0.1%	
		405816004|Procedure morphology (attribute)|		36	0.1%	
20.	71388002|Procedure (procedure)|	424226004|Using device (attribute)|	49062001|Device (physical object)|	4,289	43.6%	N/A
		363699004|Direct device (attribute)|		3,866	39.3%	
		425391005|Using access device (attribute)|		1,349	13.7%	
		405815000|Procedure device (attribute)|		269	2.7%	
		363710007|Indirect device (attribute)|		71	0.7%	
21.	71388002|Procedure (procedure)|	363700003|Direct morphology (attribute)|	49755003|Morphologically abnormal structure (morphologic abnormality)|	7,641	91.3%	Y?
		363709002|Indirect morphology (attribute)|		538	6.4%	
		405816004|Procedure morphology (attribute)|		179	2.1%	
		405813007|Procedure site - Direct (attribute)|*		11	0.1%	
		246093002|Component (attribute)|*		1	0.0%	
22.	71388002|Procedure (procedure)|	246093002|Component (attribute)|	105590001|Substance (substance)|	7886		N/A
		363701004|Direct substance (attribute)|		4217		
		424361007|Using substance (attribute)|		2641		
23.	243796009|Situation with explicit context (situation)|	246090004|Associated finding (attribute)|	363787002|Observable entity (observable entity)|	0	0.0%	N/A
		363589002|Associated procedure (attribute)|		0	0.0%	

12.6.7.8 Candidate is Contextual Qualifier and Predicate is a Finding or Procedure

For example, if "backache in the past" has been encoded separately as **161891005|Backache (finding)|** and **410513005|In the past (qualifier value)|**, the situation wrapper, **243796009|Situation with explicit context (situation)|**, will be created and the predicate and candidate will be linked using **246090004|Associated finding (attribute)|=161891005|Backache (finding)|** and **408731000|Temporal context (attribute)|= 410513005|In the past (qualifier value)|** respectively.

12.6.7.9 Candidate is Explicit Concept Model Attribute

The descriptions of the 59 Concept Model attributes were used to determine how often they occur in textual descriptions and concept definitions. Of the 59, textual descriptions occur in only nine different Concept Model attributes (refer to **Table 12-44**).

Table 12-44. Concept Model attributes that are used in SNOMED CT descriptions (TD=Textual Description; CD=Concept Definition).

No	Fully Specified Name	Term	Domain	Yes TD / Yes CD	Yes TD/ No CD	No TD/ Yes CD
1.	255234002|After (attribute)|	380400010|After|	404684003|Clinical finding (finding)|	58	95	1,538
2.	116676008|Associated morphology (attribute)|	1220306019|Morphology|	404684003|Clinical finding (finding)|	8	31	41,585
3.	47429007|Associated with (attribute)|	79074014|Associated with|	404684003|Clinical finding (finding)|	344	150	1,533
4.	246093002|Component (attribute)|	367805018|Component|	386053000|Evaluation procedure (procedure)|	14	0	8,063
5.	42752001|Due to (attribute)|	71336013|Caused by|	404684003|Clinical finding (finding)|	38	40	1,618
6.	42752001|Due to (attribute)|	71335012|Due to|	404684003|Clinical finding (finding)|	507	2,324	1,149
7.	42752001|Due to (attribute)|	71335012|Due to|	272379006|Event (event)|	2	398	23
8.	363703001|Has intent (attribute)|	1490773016|Intent|	71388002|Procedure (procedure)|	1	1	3,982
9.	260686004|Method (attribute)|	388315014|Method|	71388002|Procedure (procedure)|	185	336	46,547
10.	260870009|Priority (attribute)|	388521012|Priority|	71388002|Procedure (procedure)|	1	1	181
11.	424226004|Using device (attribute)|	2643091013|Using device|	71388002|Procedure (procedure)|	1	0	3,967

The domains and ranges associated with the use of the attributes **255234002|After (attribute)|**, **47429007|Associated with (attribute)|**, **42752001|Due to (attribute)|** and **246075003|Causative agent (attribute)|** are shown in **Table 12-45.**

Table 12-45. Domains and ranges that use "255234002|After (attribute)|", "47429007|Associated with (attribute)|", "42752001|Due to (attribute)|" and "246075003|Causative agent (attribute)|".

	Domains (Columns) / Range (Rows)	404684003|Clinical finding (finding)|	272379006|Event (event)|
1.	404684003|Clinical finding (finding)|	▪ 255234002|After (attribute)| ▪ 47429007|Associated with (attribute)| ▪ 42752001|Due to (attribute)|	▪ 255234002|After (attribute)| ▪ 47429007|Associated with (attribute)| ▪ 42752001|Due to (attribute)|
2.	272379006|Event (event)|	▪ 47429007|Associated with (attribute)| ▪ 246075003|Causative agent (attribute)|	▪ 47429007|Associated with (attribute)| ▪ 42752001|Due to (attribute)|
3.	410607006|Organism (organism)|	▪ 47429007|Associated with (attribute)| ▪ 246075003|Causative agent (attribute)|	▪ 47429007|Associated with (attribute)| ▪ 246075003|Causative agent (attribute)|
4.	373873005|Pharmaceutical / biologic product (product)|	▪ 47429007|Associated with (attribute)| ▪ 246075003|Causative agent (attribute)|	▪ 47429007|Associated with (attribute)| ▪ 246075003|Causative agent (attribute)|
5.	78621006|Physical force (physical force)|	▪ 47429007|Associated with (attribute)| ▪ 246075003|Causative agent (attribute)|	▪ 47429007|Associated with (attribute)| ▪ 246075003|Causative agent (attribute)|
6.	260787004|Physical object (physical object)|	▪ 47429007|Associated with (attribute)| ▪ 246075003|Causative agent (attribute)|	▪ 47429007|Associated with (attribute)| ▪ 246075003|Causative agent (attribute)|

	Domains (Columns) / Range (Rows)	404684003|Clinical finding (finding)|	272379006|Event (event)|
7.	71388002|Procedure (procedure)|	▪ 255234002|After (attribute)| ▪ 47429007|Associated with (attribute)| ▪ 418775008|Finding method (attribute)|	▪ 255234002|After (attribute)| ▪ 47429007|Associated with (attribute)|
8.	138875005|SNOMED CT Concept (SNOMED RT+CTV3)|	▪ 255234002|After (attribute)| ▪ 47429007|Associated with (attribute)|	▪ 47429007|Associated with (attribute)| ▪ 246075003|Causative agent (attribute)|
9.	105590001|Substance (substance)|	▪ 47429007|Associated with (attribute)| ▪ 246075003|Causative agent (attribute)|	▪ 47429007|Associated with (attribute)| ▪ 246075003|Causative agent (attribute)|

12.6.8 Additional Post-coordination

12.6.8.1Acute Pneumonia

While it may seem reasonable to encode a condition such as "acute pneumonia" using **233604007|Pneumonia (disorder)|:263502005|Clinical course (attribute)|=373933003|Acute onset (qualifier value)|**, such qualification may be incomplete. A closer investigation revealed that **373933003|Acute onset (qualifier value)|** is never used as part of a concept's defining attributes. A search for concepts that have the word "acute" in the fully specified name and that are subtypes of **362981000|Qualifier value (qualifier value)|** showed there are seven concepts, only three of which are used as defining attributes (refer to Table 12-46). It should be noted that of the four that were excluded, only **373933003|Acute onset (qualifier value)|** can be accessed via the Concept Model.

Table 12-46. "362981000|Qualifier value (qualifier value)|" concepts that have "acute" in the fully specified name and their frequency of use in defining attributes.

No	Concept	Total
1.	255212004|Acute-on-chronic (qualifier value)|	20
2.	255228009|Recurrent acute (qualifier value)|	8
3.	18131002|Acute fulminating (qualifier value)|	6
4.	255251009|Acute phase (qualifier value)|	0
5.	310072004|Acute pain service (qualifier value)|	0
6.	445449000|Acute care hospice service (qualifier value)|	0
7.	373933003|Acute onset (qualifier value)|	0

Concepts such as **274663001|Acute pain (finding)|** and **29050005|Acute iritis (disorder)|** that have the word "acute" in the fully specified name use **424124008|Sudden onset AND/OR short duration (qualifier value)|** instead of **373933003|Acute onset (qualifier value)|**. In addition to using **424124008|Sudden onset AND/OR short duration (qualifier value)|** in the defining attributes, the **116676008|Associated morphology (attribute)|** is often refined to include a **49755003|Morphologically abnormal structure (morphologic abnormality)|** concept with the word "acute." For example, **65074000|Iritis (disorder)|** had a defining attribute of **116676008|Associated morphology (attribute)|=23583003|Inflammation (morphologic abnormality)|** while **29050005|Acute iritis (disorder)|** had a defining attribute of **4532008|Acute inflammation (morphologic abnormality)|**.

A query was conducted to determine the number of concepts that have the word "acute" in the description, **116676008|Associated morphology (attribute)|** and **362981000|Qualifier value (qualifier value)|**. The results are shown in **Table 12-47**.

Table 12-47. "362981000|Qualifier value (qualifier value)|" concepts that have "acute" in a description and their frequency of use in defining attributes.

No	Description	Primitive	Fully Defined	Total
1.	Clinical finding concepts with "acute" as part of a description	597	668	1265
2.	Clinical finding concepts with "acute" as part of its qualifier defining attributes or "sudden onset and/or short duration"	427	552	979
3.	Clinical finding concepts with "acute" as part of its associated morphology defining attributes	345	544	889

A total of 774 unique fully defined concepts either had "acute" in the fully specified name, **116676008|Associated morphology (attribute)|** or **362981000|Qualifier value (qualifier value)|**. Of the 668 concepts that had a description that include "acute," 331 (49.6%) had both defining attributes of **116676008|Associated morphology (attribute)|** and **362981000|Qualifier value (qualifier value)|** that referred "acute," 159 (23.8%) only had a **362981000|Qualifier value (qualifier value)|** that referred to "acute," 158 (23.7%) only had a **116676008|Associated morphology (attribute)|**, that referred to "acute," while 20 (3.0%) did not have either. A review of the concepts did not reveal a pattern of whether "acute" should be specified as a **116676008|Associated morphology (attribute)|** and/or **362981000|Qualifier value (qualifier value)|**. A general rule may be to always encode a term that refers to "acute" with **424124008|Sudden onset AND/OR short duration (qualifier value)|**. In addition, if the concept has a defining attribute that refers to **116676008|Associated morphology (attribute)|** that has a subtype concept with "acute" in the description, it should also be refined.

12.6.8.2 Allergic Concepts

Table 12-48. Allergic concepts that have additional defining attributes in addition to "246075003|Causative agent (attribute)|" and "363705008|Has definitional manifestation (attribute)|".

No	Concept	255234002|After (attribute)|	116676008|Associated morphology (attribute)|	47429007|Associated with (attribute)|	263502005|Clinical course (attribute)|	42752001|Due to (attribute)|	363698007|Finding site (attribute)|
1.	294657002|Smallpox vaccine allergy (disorder)|	✓					
2.	444316004|Seasonal allergy (disorder)|				✓		
3.	418968001|Allergy to gauze (disorder)|					✓	
4.	91926002|Allergic rhinitis due to grass pollens (disorder)|		✓		✓	✓	✓
5.	91927006|Allergic rhinitis due to tree pollens (disorder)|		✓		✓	✓	✓
6.	91928001|Allergic rhinitis due to weed pollens (disorder)|		✓		✓	✓	✓
7.	294268009|Digoxin-specific-antibody allergy (disorder)|			✓			

No	Concept	255234002|After (attribute)|	116676008|Associated morphology (attribute)|	47429007|Associated with (attribute)|	263502005|Clinical course (attribute)|	42752001|Due to (attribute)|	363698007|Finding site (attribute)|
8.	294299005|Dialysis fluid allergy (disorder)|			✓			
9.	294300002|Peritoneal dialysis solution allergy (disorder)|			✓			
10.	294301003|Hemodialysis fluid allergy (disorder)|			✓			
11.	294302005|Hemofiltration solution allergy (disorder)|			✓			
12.	402246000|Allergic sensitization by patch test (disorder)|			✓			
13.	432807008|Pollen-food allergy (disorder)|			✓			
14.	21626009|Cutaneous hypersensitivity (disorder)|						✓
15.	197493001|Malabsorption due to intolerance to soya protein (disorder)|						✓
16.	197494007|Intestinal malabsorption of fat (disorder)|						✓
17.	235720008|Malabsorption due to intolerance to carbohydrate (disorder)|						✓
18.	302921006|Malabsorption due to intolerance to protein (disorder)|						✓
19.	303062005|Malabsorption due to intolerance to fat (disorder)|						✓
20.	414314005|Gastrointestinal food allergy (disorder)|						✓

12.6.9 Contextual Values

12.6.9.1408731000|Temporal context (attribute)|

Table 12-49. 410510008|Temporal context value (qualifier value)|.

No	408731000|Temporal context (attribute)|	Pre-coordinated Concept
1.	410589000|All times past (qualifier value)|	
2.	15240007|Current (qualifier value)|	
3.	410584005|Current - time specified (qualifier value)|	
4.	410585006|Current - time unspecified (qualifier value)|	
5.	410511007|Current or past (actual) (qualifier value)|	
6.	410512000|Current or specified time (qualifier value)|	
7.	410513005|In the past (qualifier value)|	417662000|History of clinical finding in subject (situation)|
8.	410587003|Past - time specified (qualifier value)|	
9.	410588008|Past - time unspecified (qualifier value)|	
10.	6493001|Recent (qualifier value)|	
11.	410586007|Specified time (qualifier value)|	

12.6.9.2408729009|Finding context (attribute)|

Table 12-50. 18410514004|Finding context value (qualifier value)|

No	Finding Context Value	Pre-coordinated Concept with Finding Context
1.	410519009|At risk context (qualifier value)|	
2.	410605003|Confirmed present (qualifier value)|	395098000|Disorder confirmed (situation)|
3.	410594000|Definitely NOT present (qualifier value)|	
4.	410591008|Definitely present (qualifier value)|	
5.	410517006|Expectation context (qualifier value)|	
6.	410518001|Goal context (qualifier value)|	
7.	71033007|Impending (qualifier value)|	
8.	36692007|Known (qualifier value)|	
9.	410516002|Known absent (qualifier value)|	373572006|Clinical finding absent (situation)|
10.	410590009|Known possible (qualifier value)|	443859009|Possible clinical finding (situation)|
11.	410515003|Known present (qualifier value)|	373573001|Clinical finding present (situation)|
12.	410596003|Likely outcome (qualifier value)|	
13.	428263003|NOT suspected (qualifier value)|	444436002|Clinical finding not suspected (situation)|
14.	410593006|Probably NOT present (qualifier value)|	
15.	410592001|Probably present (qualifier value)|	
16.	410595004|Prognosis context (qualifier value)|	
17.	415684004|Suspected (qualifier value)|	41769001|Disease suspected (situation)|
18.	261665006|Unknown (qualifier value)|	X

12.6.9.3408730004|Procedure context (attribute)|

Table 12-51. 288532009|Context values for actions (qualifier value)|

No	Context Values for Actions	Pre-coordinated Concept
1.	385657008|Abandoned (qualifier value)|	180863000|Operation abandoned (situation)|
2.	385645004|Accepted (qualifier value)|	
3.	410537005|Action status unknown (qualifier value)|	399714002|Procedure status unknown (situation)|
4.	44996008|Approved and scheduled (qualifier value)|	
5.	410542002|Attended (qualifier value)|	
6.	385649005|Being organized (qualifier value)|	
7.	89925002|Canceled (qualifier value)|	
8.	441898007|Consented (qualifier value)|	
9.	385661002|Considered and not done (qualifier value)|	
10.	410536001|Contraindicated (qualifier value)|	183932001|Procedure contraindicated (situation)|
11.	441889009|Denied (qualifier value)|	
12.	410543007|Did not attend (qualifier value)|	
13.	410546004|Discontinued (qualifier value)|	416406003|Procedure discontinued (situation)|

No	Context Values for Actions	Pre-coordinated Concept
14.	385658003|Done (qualifier value)|	443938003|Procedure carried out on subject (situation)|
15.	385656004|Ended (qualifier value)|	
16.	385651009|In progress (qualifier value)|	
17.	410535002|Indicated (qualifier value)|	
18.	442633000|Legal agent consented (qualifier value)|	
19.	410525008|Needed (qualifier value)|	417451006|Procedure needed (situation)|
20.	385660001|Not done (qualifier value)|	416237000|Procedure not done (situation)|
21.	410534003|Not indicated (qualifier value)|	428119001|Procedure not indicated (situation)|
22.	410529002|Not needed (qualifier value)|	
23.	410530007|Not offered (qualifier value)|	416064006|Procedure not offered (situation)|
24.	410521004|Not to be done (qualifier value)|	
25.	385653007|Not to be stopped (qualifier value)|	
26.	410528005|Not wanted (qualifier value)|	416432009|Procedure not wanted (situation)|
27.	410527000|Offered (qualifier value)|	
28.	385650005|Organized (qualifier value)|	416662009|Procedure organized (situation)|
29.	398166005|Performed (qualifier value)|	
30.	397943006|Planned (qualifier value)|	183976008|Operative procedure planned (situation)|
31.	10523001|Post-starting action status (qualifier value)|	
32.	410522006|Pre-starting action status (qualifier value)|	
33.	442681007|Recipient consented (qualifier value)|	
34.	443390004|Refused (qualifier value)|	183944003|Procedure refused (situation)|
35.	385647007|Rejected by performer (qualifier value)|	
36.	385648002|Rejected by recipient (qualifier value)|	
37.	385644000|Requested (qualifier value)|	400999005|Procedure requested (situation)|
38.	443942000|Requested by recipient (qualifier value)|	183995001|Patient requested procedure (situation)|
39.	385646003|Schedule rejected (qualifier value)|	
40.	416151008|Scheduled - procedure status (qualifier value)|	
41.	385652002|Started (qualifier value)|	394906002|Procedure started (situation)|
42.	410545000|Stopped before completion (qualifier value)|	394908001|Procedure stopped (situation)|
43.	385655000|Suspended (qualifier value)|	
44.	385643006|To be done (qualifier value)|	
45.	385654001|To be stopped (qualifier value)|	
46.	385642001|Under consideration (qualifier value)|	
47.	410531006|Under consideration, not wanted yet (qualifier value)|	
48.	410532004|Under consideration, not yet offered (qualifier value)|	
49.	410526009|Wanted (qualifier value)|	
50.	410524007|Was not started (qualifier value)|	

12.7 Appendix F: For Chapter Eight

12.7.1 Fundamentals of SNOMED CT Queries

When testing two expressions (either pre-coordinated concepts or post-coordinated expressions) for equivalency and subsumption, the expression that is being tested to see if it subsumes another expression is called the *predicate* while the expression that is being tested to see if it is subsumed by another expression is called the *candidate*. To demonstrate the types of relationships concepts have with each other (or lack of relationships), seven concepts in red have been plotted out in Figure 12-26. The defining attributes of the seven concepts are shown in Error! Reference source not found.. **73211009|Diabetes mellitus (disorder)|** is used as the predicate in the following examples. When testing for equivalency and subsumption, there are generally four possible results. First, *equivalent*, whereby both expressions contain exactly the same semantic information. For example, the candidate **73211009|Diabetes mellitus (disorder)|** is the same concept as the predicate and is equivalent. Second, *subsumption*, or subtype, whereby the predicate subsumes the candidate. In this case the candidate may contain the same or refined definitions as well as additional definitions. For example, **73211009|Diabetes mellitus (disorder)|** is a parent of **46635009|Diabetes mellitus type 1 (disorder)|**. In this example, both concepts have the same definitions **363698007|Finding site (attribute)|=113331007|Structure of endocrine system (body structure)|**. Third, *reverse subsumption*, or supertype, whereby the candidate subsumes the predicate. For example, **362969004|Disorder of endocrine system (disorder)|** is a parent of **73211009|Diabetes mellitus (disorder)|**. Fourth, *unrelated*, whereby the predicate and candidate do not have a supertype/subtype relationship. For example, **3218000|Mycosis (disorder)|**, does not have a supertype/subtype relationship with **73211009|Diabetes mellitus (disorder)|**.

While the hierarchy is useful for aggregating concepts, not all related concepts are arranged strictly using the subsumption hierarchy. Concepts may be related by being siblings, via defining attributes or in a different context. For example, when retrieving concepts related to **73211009|Diabetes mellitus (disorder)|**, it may also be useful to look up concepts that are related to **73211009|Diabetes mellitus (disorder)|** but are not subtypes of **73211009|Diabetes mellitus (disorder)|**. For example, a diagnosis of **371087003|Diabetic foot ulcer (disorder)|** implies that **73211009|Diabetes mellitus (disorder)|** is present in a patient even though it is not a subtype concept of **73211009|Diabetes mellitus (disorder)|** but instead has a defining attribute of **47429007|Associated with (attribute)|=73211009|Diabetes mellitus (disorder)|**. In this example, **74627003|Diabetic complication (disorder)|** is the "container" concept for all concepts that are associated with diabetes mellitus. This type of query, via defining attributes, may be useful when retrieving related conditions. Similarly, **161445009|History of - diabetes mellitus (situation)|** is related to **73211009|Diabetes mellitus (disorder)|** because it has a defining attribute of **246090004|Associated finding (attribute)|=73211009|Diabetes mellitus (disorder)|**. However, it is used in a different context (**410513005|In the past (qualifier value)|** vs **410512000|Current or specified time (qualifier value)|**). This type of query, context free, may be useful when trying to determine if a condition exists but in a different context. Finally,

80394007|Hyperglycemia (disorder)| does not have a supertype or subtype relationship with **73211009|Diabetes mellitus (disorder)|**, nor is it defined with the latter concept, but they are sibling concepts as they are both children concepts of **362969004|Disorder of endocrine system (disorder)|**. There is probably less value in determining that two concepts are sibling concepts, although it may be more informative to say that two concepts are siblings than to say two concepts are unrelated.

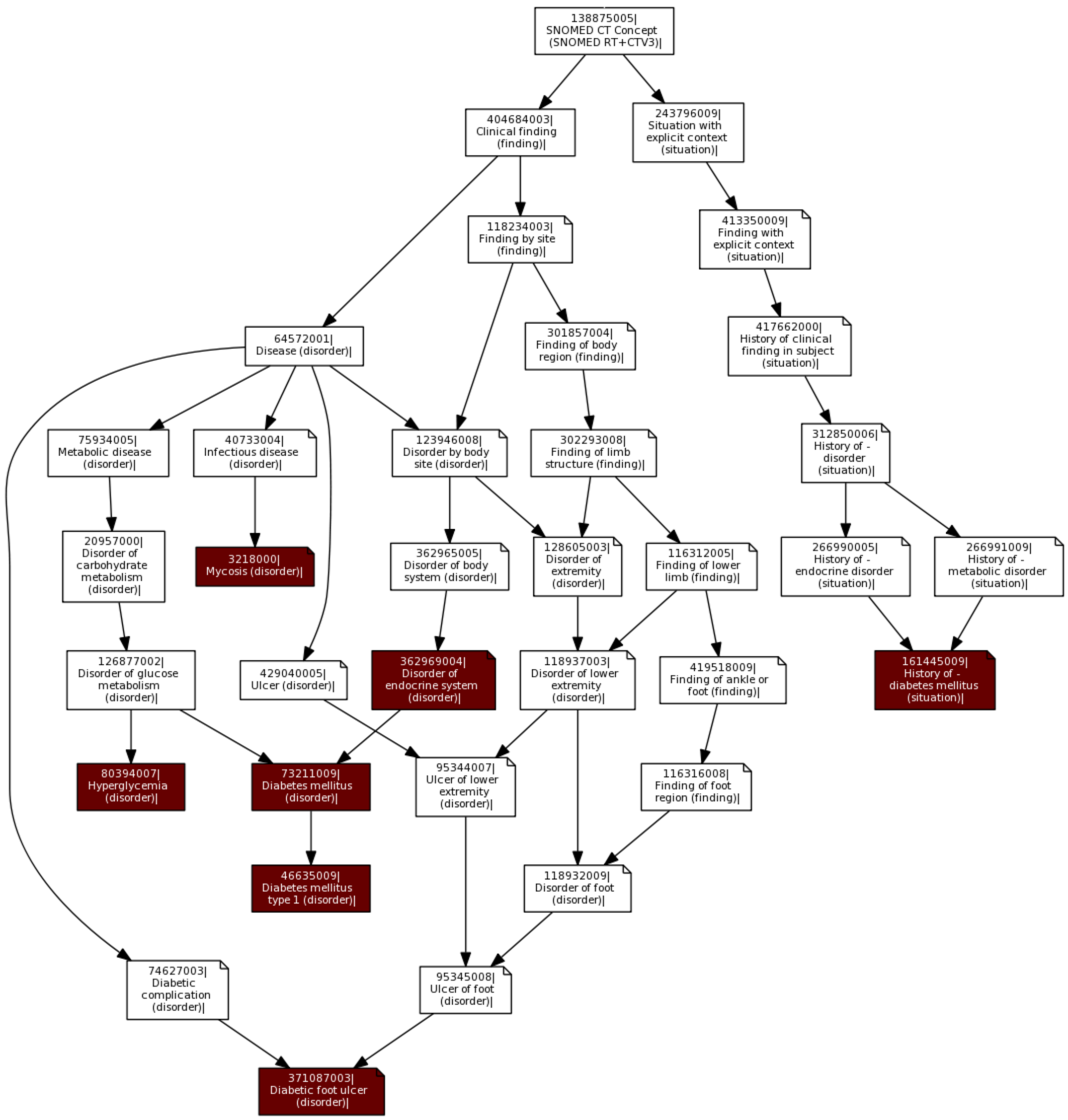


Figure 12-26. Concepts used to demonstrate the types of SNOMED CT queries that can be conducted.

Table 12-52. Defining attributes of concepts used to demonstrate the types of SNOMED CT queries that can be conducted.

No	Concept	Defining Attributes
1.	73211009|Diabetes mellitus (disorder)|	73211009|Diabetes mellitus (disorder)|: 363698007|Finding site (attribute)|= 113331007|Structure of endocrine system (body structure)|
2.	46635009|Diabetes mellitus type 1 (disorder)|	46635009|Diabetes mellitus type 1 (disorder)|: 363698007|Finding site (attribute)|=

		113331007|Structure of endocrine system (body structure)|
3.	362969004|Disorder of endocrine system (disorder)|	362969004|Disorder of endocrine system (disorder)|: 363698007|Finding site (attribute)|= 113331007|Structure of endocrine system (body structure)|
4.	371087003|Diabetic foot ulcer (disorder)|	371087003|Diabetic foot ulcer (disorder)|: 47429007|Associated with (attribute)|= 73211009|Diabetes mellitus (disorder)| {116676008|Associated morphology (attribute)|= 56208002|Ulcer (morphologic abnormality)|, 363698007|Finding site (attribute)|= 56459004|Foot structure (body structure)|}
5.	80394007|Hyperglycemia (disorder)|	80394007|Hyperglycemia (disorder)|
6.	161445009|History of - diabetes mellitus (situation)|	161445009|History of - diabetes mellitus (situation)|: {246090004|Associated finding (attribute)|= 73211009|Diabetes mellitus (disorder)|, 408729009|Finding context (attribute)|= 410515003|Known present (qualifier value)|, 408731000|Temporal context (attribute)|= 410513005|In the past (qualifier value)|, 408732007|Subject relationship context (attribute)|= 410604004|Subject of record (person)|}
7.	3218000|Mycosis (disorder)|	3218000|Mycosis (disorder)|: 246075003|Causative agent (attribute)|= 414561005|Kingdom Fungi (organism)|, 370135005|Pathological process (attribute)|= 441862004|Infectious process (qualifier value)|

It should be noted that equivalency can occur at two levels. First, two pre-coordinated concepts can be tested for equivalency. In general, the long normal forms of pre-coordinated concepts with context should all be unique. However, there are redundancies between concepts that are subtypes of **404684003|Clinical finding (finding)|** and **373573001|Clinical finding present (situation)|**, and **71388002|Procedure (procedure)|** and **443938003|Procedure carried out on subject (situation)|**. When a concept is normalised and the concept is from the **404684003|Clinical finding (finding)|** or **71388002|Procedure (procedure)|** hierarchies, a "situation" wrapper is included and the "default contexts" are added (refer to Table 12-53). For example, when a concept from the **404684003|Clinical finding (finding)|** hierarchy is used, it is assumed that it means that the **408729009|Finding context (attribute)|** refers to **410515003|Known present (qualifier value)|**, the **408731000|Temporal context (attribute)|** refers to **410512000|Current or specified time (qualifier value)|** and the **408732007|Subject relationship context (attribute)|** refers to **410604004|Subject of record (person)|**. So if we use the concept **422587007|Nausea (finding)|** in a clinical record, we are implying that nausea is the condition that we know is present, refers to a time period of now or a specified period of time, and refers to the individual whose record we are recording.

Table 12-53. Default contexts for "373573001|Clinical finding present (situation)|" and "443938003|Procedure carried out on subject (situation)|".

Context / Situations	373573001|Clinical finding present (situation)|	443938003|Procedure carried out on subject (situation)|
408729009|Finding context (attribute)|	410515003|Known present (qualifier value)|	
408730004|Procedure context (attribute)|		385658003|Done (qualifier value)|
408731000|Temporal context (attribute)|	410512000|Current or specified time (qualifier value)|	410512000|Current or specified time (qualifier value)|
408732007|Subject relationship context (attribute)|	410604004|Subject of record (person)|	410604004|Subject of record (person)|

The number of subtype concepts from **373573001|Clinical finding present (situation)|** and **443938003|Procedure carried out on subject (situation)|** are shown in Table 12-54. This means there are 82 and 468 concepts from the **404684003|Clinical finding (finding)|** and **71388002|Procedure (procedure)|** hierarchies respectively that are redundantly represented in the **243796009|Situation with explicit context (situation)|** hierarchy. The implication is that we cannot just compare if two pre-coordinated concepts have a supertype/subtype relationship. If one concept is from the **404684003|Clinical finding (finding)|** hierarchy while the other is a subtype of **373573001|Clinical finding present (situation)|** or if one concept is from the **71388002|Procedure (procedure)|** hierarchy while the other is a subtype of **443938003|Procedure carried out on subject (situation)|**, we have to conduct additional checks using the long normal form.

Table 12-54. Number of subtype concepts from the concepts "373573001|Clinical finding present (situation)|" and "443938003|Procedure carried out on subject (situation)|."

Number of Subtype Concepts	Primitive	Fully-defined	Total
373573001|Clinical finding present (situation)|	7	75	82
443938003|Procedure carried out on subject (situation)|	276	192	468

12.7.1.1 Equivalency of Pre-coordinated Concepts

Examples of pre-coordinated concepts that belong to different hierarchies that are semantically equivalent are **162057007|Nausea present (situation)|** and **422587007|Nausea (finding)|**, and **182833002|Medication given (situation)|** and **18629005|Administration of drug or medicament (procedure)|** (refer to Error! Reference source not found.). The former are equivalent because **162057007|Nausea present (situation)|** is defined with **246090004|Associated finding (attribute)|=422587007|Nausea (finding)|** while the latter are equivalent because **182833002|Medication given (situation)|** is defined with **363589002|Associated procedure (attribute)|=18629005|Administration of drug or medicament (procedure)|.**

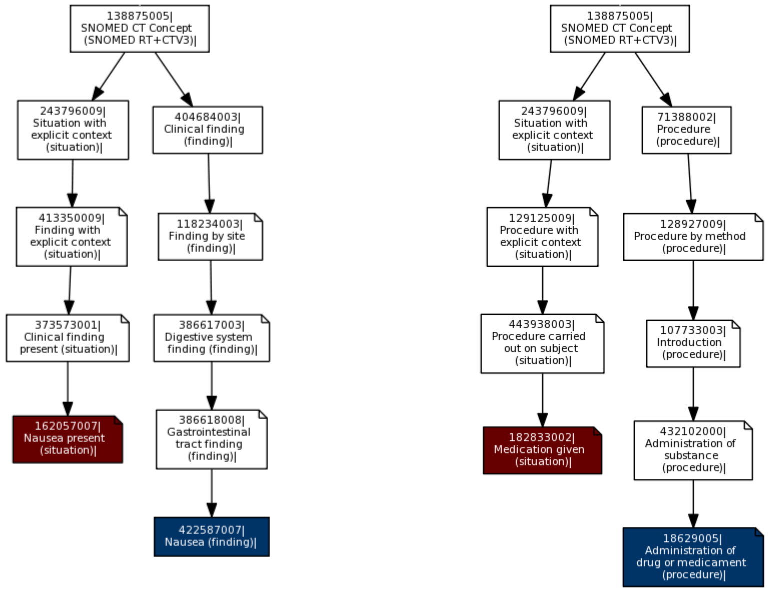


Figure 12-27. Location of concepts "162057007|Nausea present (situation)|", "422587007|Nausea (finding)|", "182833002|Medication given (situation)|" and "18629005|Administration of drug or medicament (procedure)|" in the hierarchy.

12.7.1.2 Equivalency of a Pre-coordinated Concept and Post-coordinated Expression

The term "lung cancer" can be represented using the pre-coordinated concept **363358000|Malignant tumor of lung (disorder)|** or a variety of post-coordinated expressions (refer to Error! Reference source not found.). The three post-coordinated expressions along with the pre-coordinated concept are all considered equivalent because they all have the same long normal form (refer to Example 3 in Error! Reference source not found.).

Example #1
19829001|Disorder of lung (disorder)|:
    116676008|Associated morphology (attribute)|=
        367651003|malignant neoplasm of primary, secondary, or uncertain origin (morphologic abnormality)|

Example #2
363346000|Malignant neoplastic disease (disorder)|:
    363698007|Finding site (attribute)|=
        39607008|Lung structure (body structure)|

Example #3
64572001|Disease (disorder)|:
    {116676008|Associated morphology (attribute)|=
        367651003|Malignant neoplasm of primary, secondary, or uncertain origin (morphologic abnormality)|,
    363698007|Finding site (attribute)|=
        39607008|Lung structure (body structure)|}

Figure 12-28. Different ways of represent "lung cancer" using post-coordinated expressions.

12.8 Appendix G: For Chapter Nine

Note: The SNOMED CT Clinical Value Framework was initially called SNOMED CT Meaningful Use Framework but was relabelled after feedback from external reviewers.

12.8.1 Clinical Feedback Session Slides

Towards Demonstrating the Meaningful Use of SNOMED CT in Primary Care: Clinical Feedback Session

Dennis Lee Hon Kit
dlhk@uvic.ca

1

Outline

1. Review US Meaningful Use Stages and Criteria
2. Proposed SNOMED CT Meaningful Use Framework
3. Materials, Methods, Results
   - Completeness of Problem List
   - Completeness of Encounters and Diagnostic Codes
   - Completeness of Clinical Summaries
   - Adherence of Care Guidelines
4. Summary of Meaningful Use of SNOMED CT
5. Application (EMR Prototype Demo)
6. Discussion
   - Does it make sense?
   - Does it add value?
   - What are the potential implementation issues?

2

US Meaningful Use Stages and Criteria

Stage 1: Data Capture and Sharing	Stage 2: Advance Clinical Processes	Stage 3: Improved Outcomes
Electronically capturing health information in a standardised format Using that information to track key clinical conditions Communicating that information for care coordination processes Initiating the reporting of clinical quality measures and public health information Using information to engage patients and their families in their care	More rigorous health information exchange (HIE) Increased requirements for e-prescribing and incorporating lab results Electronic transmission of patient care summaries across multiple settings More patient-controlled data	Improving quality, safety, and efficiency, leading to improved outcomes Decision support for national high-priority conditions Patient access to self-management tools Access to comprehensive patient data through patient-centred HIE

http://www.healthit.gov/policy-researchers-implementers/meaningful-use

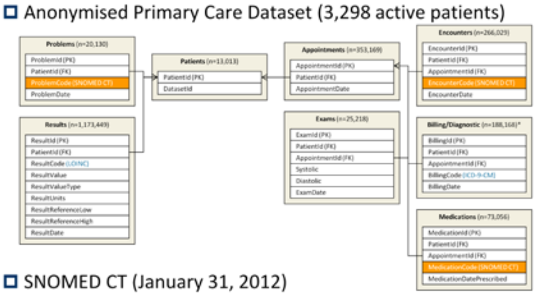
3

Proposed SNOMED CT Meaningful Use Framework

Meaningful Use Stage / Data Use Target Audience	Patient	Practice	Population
Stage 1: Data Capture and Sharing Electronically capturing health information in a standardised format Using that information to track key clinical conditions Initiating the reporting of clinical quality measures and public health information	Suggest diagnostic code based on encounter diagnosis Suggest encounter diagnosis based on diagnostic code Suggest adding encounter diagnosis to problem list	Audit (critique) completeness of problem list and encounter diagnosis Audit (critique) accuracy of diagnostic codes Patient recall / patient case queries	Generate statistics on key clinical quality measures (CQM)
Stage 2: Advance Clinical Processes Electronic transmission of patient care summaries across multiple settings	Generate patient clinical summary in the form of the Continuity Care Document (CCD)	Generate practice quality reports in the form of the Quality Reporting Document Architecture (QRDA) Category II and III	Generate practice quality reports in the form of the Quality Reporting Document Architecture (QRDA) Category III
Stage 3: Improved Outcomes Improving quality, safety, and efficiency, leading to improved outcomes Decision support for national high-priority conditions	Alert clinicians to drug-allergy interactions Remind clinicians to order routine tests Suggest medications based on out of range test results	Audit (critique) evidence-based guideline adherence Assess patient conditions	Compare clinical outcomes between practice and population

Materials

- Anonymised Primary Care Dataset (3,298 active patients)



- SNOMED CT (January 31, 2012)
- SNOMED CT to ICD-9-CM Reimbursement Cross Map
- Continuity of Care Document (CCD)

5

Methods

No	Method	Meaningful Use Criteria	Problem	Encounter	Lab Result	Diag. (ICD-9-CM)	Medication	Examination
1.	Completeness of problem list[1]	Stage 1 – Electronically capturing health information in a standardised format	✓		✓	✓	✓	
2.	Completeness of encounter diagnoses and diagnostic codes	Stage 1 – Electronically capturing health information in a standardised format		✓		✓		
3.	Completeness of clinical summary	Stage 2 – Electronic transmission of patient care summaries across multiple settings	✓	✓				
4.	Adherence of care guidelines[2]	Stage 3 – Improving quality, safety, and efficiency, leading to improved outcomes Stage 3 – Decision support for national high-priority conditions	✓	✓	✓		✓	✓

1. Wright A, et al. **A method and knowledge base for automated inference of patient problems from structured data in an electronic medical record.** J Am Med Inform Assoc. 2011 Nov-Dec;18(6):859-67.
2. Hahn KA, et al. **Diabetes flow sheet use associated with guideline adherence.** Ann Fam Med. 2008 May-Jun;6(3):235-8.

6

Encoding Results

Encoding	Problems		Encounters		Medications	
	Total	Unique	Total	Unique	Total	Unique
Number of Records	**20,140**	**2,880**	**266,029**	**15,122**	**73,056**	**22,230**
Fully encoded record	18,298	2,246	231,963	8,541	69,513	19,564
Partially encoded record	1,666	582	27,198	5,794	0	0
Fully unencoded record	177	43	6,149	779	3,543	2,666

- Problem
  - COPD
  - 13645005|Chronic obstructive lung disease (disorder)|
- Encounter
  - Acute pneumonia
  - 233604007|Pneumonia (disorder)|:
    263502005|Clinical course (attribute)|=
    373933003|Acute onset (qualifier value)|
- Medication
  - Metformin (Glucophage) 500 mg PO BID
  - 109081006|Metformin (product)|

7

Encoding Results – Diabetes

- Examples of SNOMED CT concepts and local terms
  - 73211009|Diabetes mellitus (disorder)|
    - DIABETIS
    - Diabetes
    - Diabetes Mellitus
  - 46635009|Diabetes mellitus type 1 (disorder)|
    - Diabetes Type 1
    - Diabetes Insulin Dependent
    - Diabetes Mellitus -Type 1
    - Diabetes Mellitus Insulin Dependent
    - Diabetes Mellitus -Type 1- Insulin Dependent
  - 44054006|Diabetes mellitus type 2 (disorder)|
    - Diabetes Type 2
    - Diabetes Mellitus Non Insulin Dependent
    - Diabetes Mellitus Non Insulin Dependent (Type 2)
    - Diabetes Mellitus -Type 2- Insulin Dependent
    - Diabetes Mellitus -Type 2- Non Insulin Dependent

8

Completeness of Problem List

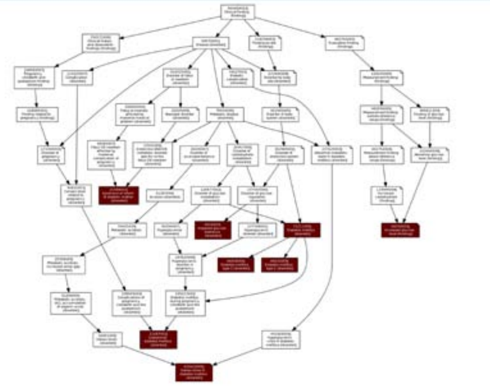
	Pt	Prac	Pop
#1			
#2			
#3			

- Definitions
  - Definition 1 – Laboratory Results (≥1)
    - Haemoglobin A1c/ Haemoglobin.total in Blood - (LOINC: 4548-4) ≥7
    - Haemoglobin A1c/ Haemoglobin.total in Blood by HPLC - (LOINC: 17856-6) ≥7
    - Haemoglobin A1c/ Haemoglobin.total in Blood by Electrophoresis - (LOINC: 4549-2) ≥7
  - Definition 2 – Billing (Diagnostic) Codes (≥2)
    - Diabetes mellitus - (ICD-9-CM: 250) or one of its subtypes
  - Definition 3 – Medications (≥1)
    - ETC: 5886 – Injectable Antidiabetic Agents & 154 – Oral Antidiabetic Agents
    - SNOMED CT: 384953001|Antidiabetic preparation (product)|& 373245004|Antidiabetic agent (substance)|
- And do not have any of these problems
  - 21584002|Syndrome of infant of diabetic mother (disorder)|
  - 9414007|Impaired glucose tolerance (disorder)|
  - 73211009|Diabetes mellitus (disorder)|
- Then suggest adding these problems:
  - Default problem
    - 73211009|Diabetes mellitus (disorder)|
  - Additional options as related terms
    - 46635009|Diabetes mellitus type 1 (disorder)|
    - 44054006|Diabetes mellitus type 2 (disorder)|
    - 9414007|Impaired glucose tolerance (disorder)|
    - 420422005|Ketoacidosis in diabetes mellitus (disorder)|
    - 68256003|Increased glucose level (finding)|
    - 11687002|Gestational diabetes mellitus (disorder)|

Wright A, et al. **A method and knowledge base for automated inference of patient problems from structured data in an electronic medical record.** J Am Med Inform Assoc. 2011 Nov-Dec;18(6):859-67.

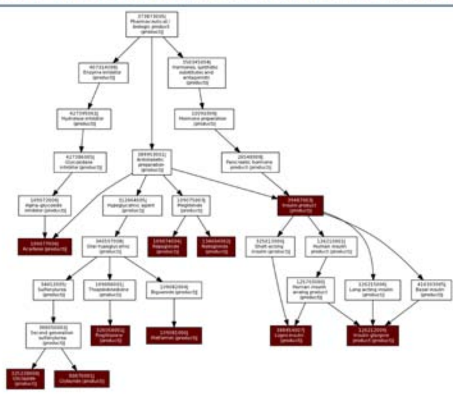
9

Closer Look at Diabetes Concepts



10

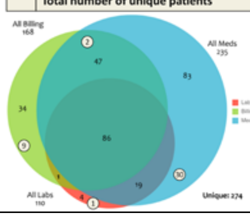
Closer Look at Antidiabetic Medications



11

Completeness of Problem List Results

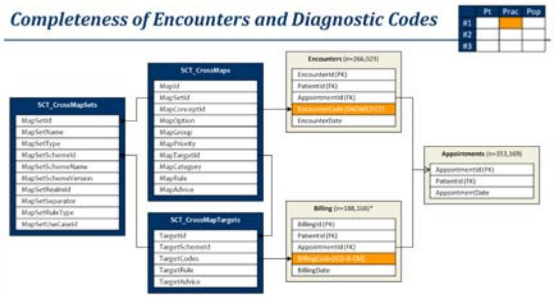
No	Definition	Patients identified in definition	Patients without problem	Percent without problem
1.	Definition 1 (Laboratory results ≥ 1) Haemoglobin A1c ≥ 7 (4548-4, 17856-6 or 4549-2)	110	1	0.9%
2.	Definition 2 (Billing codes ≥ 2) Diabetes mellitus (ICD-9-CM: 250)	168	11	6.5%
3.	Definition 3 (Medications ≥ 1) Antidiabetic medication (384953001 / 373245004)	235	32	13.6%
	Total number of unique patients	**274**	**42**	**15.3%**



- 42 (15%) patients have at least one Haemoglobin A1c ≥7, at least two billing codes of 250, or at least one antidiabetic medication, do not have diabetes on the problem list.

12

Completeness of Encounters and Diagnostic Codes



13

Completeness of Encounters and Diagnostic Codes Results

- SNOMED CT to ICD-9-CM Reimbursement Cross Map

No	SNOMED CT	ICD-9-CM	Number of Appointments	Exact Code	Main Code	No Matching Code	Percent No Matching Code
1.	73211009 | Diabetes mellitus (disorder) |	**250.00**	125	0	14	111	88.8%
2.	46635009 | Diabetes mellitus type 1 (disorder) |	**250.01**	84	33	15	36	42.9%
3.	44054006 | Diabetes mellitus type 2 (disorder) |	**250.00**	267	0	78	189	70.8%
	Totals		**476**	**33**	**107**	**336**	**70.6%**

- ICD-9-CM to SNOMED CT

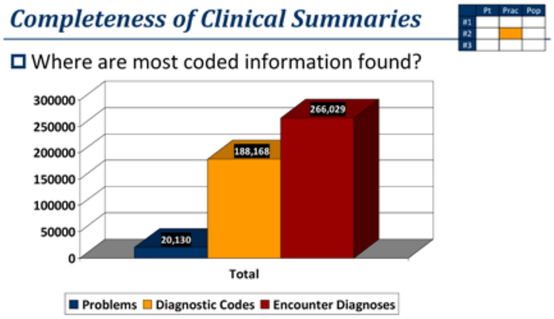
No	ICD-9-CM	Number of Appointments	Matching Diagnoses	No Matching Diagnoses	Percent No Matching Diagnoses
1.	**250** - Diabetes mellitus	174	108	66	37.9%
2.	**250.01** - Diabetes mellitus without mention of complication, type I [juvenile type], not stated as uncontrolled	48	35	13	27.1%
	Totals	**222**	**143**	**79**	**35.6%**

- 79 (35.6%) patients have 250/250.01 but not diabetes in encounter diagnoses

14

Completeness of Clinical Summaries

- Where are most coded information found?



15

Completeness of Clinical Summaries, 2

Pt Prac Pop
#1
#2
#3

2.16.840.1.113883.10.20.1.13 **Summary Purpose**	2.16.840.1.113883.10.20.1.9 **Payers**	2.16.840.1.113883.10.20.1.1 **Advance Directive** • **Include:** 365870005 | Finding of resuscitation status (finding) |	2.16.840.1.113883.10.20.1.5 **Functional Status** • **Include:** 105501005 | Dependence on enabling machine or device (finding) |
2.16.840.1.113883.10.20.1.11 **Problems** • **Include:** 64572001 | Disease (disorder) | • **Exclude:** 281647001 | Adverse reaction (disorder) |	2.16.840.1.113883.10.20.1.4 **Family History** • **Include:** 57177007 | Family history with explicit context (situation) |	2.16.840.1.113883.10.20.1.15 **Social History** • **Include:** 365949003 | Health-related behavior finding (finding) |	2.16.840.1.113883.10.20.1.2 **Allergies, Adverse Reactions & Alerts** • **Include:** 281647001 | Adverse reaction (disorder) | • **Include:** 420134006 | Propensity to adverse reactions (disorder) |
2.16.840.1.113883.10.20.1.8 **Medications** • **Include:** 373873005 | Pharmaceutical / biologic product (product) |	2.16.840.1.113883.10.20.1.7 **Medical Equipment** • **Include:** 303607000 | Biomedical equipment (physical object) |	2.16.840.1.113883.10.20.1.6 **Immunisations** • **Include:** 127785005 | Administration of substance to produce immunity, either active or passive (procedure) |	2.16.840.1.113883.10.20.1.16 **Vital Signs** • **Include:** 118227000 | Vital signs finding (finding) |
2.16.840.1.113883.10.20.1.14 **Results** • **Include:** 442618008 | Abnormal finding on evaluation procedure (finding) |	2.16.840.1.113883.10.20.1.12 **Procedures** • **Include:** 387713003 | Surgical procedure (procedure) |	2.16.840.1.113883.10.20.1.3 **Encounters**	2.16.840.1.113883.10.20.1.10 **Plan of Care** • **Include:** 243120004 | Regimes and therapies (regime/therapy) |

HL7. **HL7 Implementation Guide: CDA Release 2 – Continuity of Care Document (CCD).** April 1, 2007.

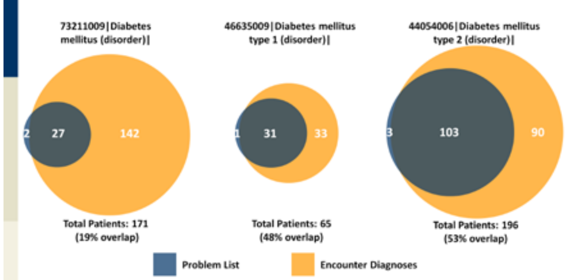
16

Completeness of Clinical Summaries Examples

Summary Purpose	**Payers**	**Advance Directive (30)** 304253006 | Not for resuscitation (finding) |	**Functional Status** N/A
Problems (*) 73211009 | Diabetes mellitus (disorder) |	**Family History (80)** 312824007 | Family history of cancer of colon (situation) |	**Social History (4.2k)** 65568007 | Cigarette smoker (finding) |	**Allergies, Adverse Reactions & Alerts (130)** 292055008 | Codeine adverse reaction (disorder) |
Medications N/A	**Medical Equipment (100)** 87405001 | Cane, device (physical object) |	**Immunisations (2.4k)** 86198006 | Influenza vaccination (procedure) |	**Vital Signs (500)** 61086009 | Pulse irregular (finding) |
Results (~70) 313172000 | Colonoscopy abnormal (finding) |	**Procedures (800)** 52734007 | Total replacement of hip (procedure) |	**Encounters**	**Plan of Care (2.2k)** 67516001 | Detoxification therapy (regime/therapy) |

17

Closer Look at Problems and Encounters



18

Sample Encounter Diagnoses

- 66590003|Alcohol dependence (disorder)|
- 128869009|Infestation by Sarcoptes scabiei var hominis (disorder)|
- 223475005|Consulting with (procedure)|
- 54150009|Upper respiratory infection (disorder)|
- 34014006|Viral disease (disorder)|
- 225358003|Wound care (regime/therapy)|
- 117614009|Laboratory procedure, review by pathologist (procedure)|
- 48532005|Muscle strain (disorder)|
- 18070006|Impacted cerumen (disorder)|
- 16932000|Nausea and vomiting (disorder)|
- 62315008|Diarrhea (finding)|
- 103742009|Renewal of prescription (procedure)|
- 427898007|Infection of tooth (disorder)|
- 271737000|Anemia (disorder)|
- 8943002|Weight gain finding (finding)|
- 43116000|Eczema (disorder)|
- 19943007|Cirrhosis of liver (disorder)|
- 235595009|Gastroesophageal reflux disease (disorder)|
- 39898005|Sleep disorder (disorder)|
- 191997003|Persistent insomnia (disorder)|
- 49727002|Cough (finding)|
- 10743008|Irritable bowel syndrome (disorder)|
- 299709002|Dental abscess (disorder)|
- 4556007|Gastritis (disorder)|
- 422587007|Nausea (finding)|
- 61582004|Allergic rhinitis (disorder)|
- 418363000|Itching of skin (finding)|
- 191813001|Chronic alcoholism in remission (disorder)|
- 89362005|Weight loss finding (finding)|
- 182782007|Dermatitis (disorder)|
- 248342006|Underweight (finding)|
- 74506000|Bereavement due to life event (finding)|
- 60442001|Perforation of tympanic membrane (disorder)|
- 79890006|Loss of appetite (finding)|
- 45007003|Low blood pressure (disorder)|
- 236071009|Chronic diarrhea (disorder)|
- 25374005|Gastroenteritis (disorder)|
- 276444007|Generalized pruritus (finding)|
- 69322001|Psychotic disorder (disorder)|
- 30037006|Anal fissure (disorder)|
- 371088008|Reactive airways dysfunction syndrome (disorder)|
- 249519007|Diarrhea and vomiting (finding)|
- 426232007|Environmental allergy (disorder)|
- 95891005|Influenza-like illness (finding)|
- 75694006|Pancreatitis (disorder)|
- 228281002|Problem drinker (finding)|
- 196746003|Persistent vomiting (disorder)|
- 235675006|Gastroparesis (disorder)|
- 200936003|Lupus erythematosus (disorder)|
- 396550006|Blood test (procedure)|
- 429196001|Partial obstruction of small bowel (disorder)|
- 68154008|Chronic cough (finding)|
- 14760008|Constipation (disorder)|
- 68566005|Urinary tract infectious disease (disorder)|
- 417746004|Traumatic injury (disorder)|
- 266261006|Raynaud's phenomenon (finding)|
- 89627008|Hyponatremia (disorder)|
- 271823003|Tachypnea (finding)|
- 4800001|Pyuria (finding)|
- 55464009|Systemic lupus erythematosus (disorder)|
- 292055008|Codeine adverse reaction (disorder)|
- 28560003|Hypovolemia (disorder)|
- 33688009|Cholestasis (finding)|
- 43275000|Otitis (disorder)|
- 37610005|Inflammation of cervix (disorder)|
- 281255004|Small bowel obstruction (disorder)|
- 271739002|Wax in ear canal (finding)|
- 235494005|Chronic pancreatitis (disorder)|
- 83414005|Macrocytic anemia (disorder)|
- 48813009|Lymphocytopenia (disorder)|
- 243114000|Support (regime/therapy)|
- 281032009|Well woman health check (procedure)|
- 385955003|Foot care (regime/therapy)|
- 363490009|Malignant tumor of anus (disorder)|
- 3135009|Otitis externa (disorder)|
- 414521009|Internal hordeolum (disorder)|
- 20928004|Ulcer of anus (disorder)|
- 68653001|Anal pain (finding)|
- 77880009|Rectal pain (finding)|
- 276610007|Low birth weight infant (disorder)|
- 414941008|Onychomycosis (disorder)|
- 86198006|Influenza vaccination (procedure)|
- 12866006|Pneumococcal vaccination (procedure)|
- 235760009|Radiation proctitis (disorder)|
- 3951002|Proctitis (disorder)|

19

Sample Encounter Diagnoses

20

Improved Organisation

- Allergies, Adverse Reactions and Alerts
  - 292055008|Codeine adverse reaction (disorder)|
  - 426232007|Environmental allergy (disorder)|
- Social History
  - 228281002|Problem drinker (finding)|
- Immunisation
  - 86198006|Influenza vaccination (procedure)|
  - 12866006|Pneumococcal vaccination (procedure)|
- Vital Signs
  - 271823003|Tachypnea (finding)|
- Plan of Care
  - 225358003|Wound care (regime/therapy)|
  - 385955003|Foot care (regime/therapy)|
- Problem (Sample)
  - 54150009|Upper respiratory infection (disorder)|
  - 19943007|Cirrhosis of liver (disorder)|
  - 10743008|Irritable bowel syndrome (disorder)|
  - 371088008|Reactive airways dysfunction syndrome (disorder)|
  - 68566005|Urinary tract infectious disease (disorder)|
  - 281255004|Small bowel obstruction (disorder)|

21

Adherence of Care Guidelines

	Pt	Prac	Pop
#1			
#2			
#3			

- Diabetes assessment scores were based on documentation in the medical record of five assessments (20 points each):
  - Haemoglobin A1c level tested in the past six months
  - Urine microalbumin level tested in the past year
  - Low-density lipoprotein (LDL-cholesterol) level tested in the past year
  - Smoking status
  - Blood pressure
- Diabetes treatment scores consisted of documentation in the medical record of 4 measures (25 points each):
  - Haemoglobin A1c level at or below 8% or use of a hypoglycemic agent
  - Urine microalbumin level below 30 mg/g (3.39 mg/mmol) creatine or use of an angiotensin-converting enzyme (ACE) inhibitor or angiotensin II receptor blocker (ARB)
  - LDL-cholesterol below 100 mg/dL (2.59 mmol/L) or use of lipid-lowering agent
  - Blood pressure at or below 130/85 mm Hg or use of an antihypertensive agent
- Diabetes target attainment scores consisted of documentation in the medical record of attainment of 3 targets (33.3 points each):
  - Haemoglobin A1c level less than 8%
  - LDL-cholesterol level less than or equal to 100 mg/dL
  - Blood pressure less than or equal to 130/85 mm Hg

Hahn KA, et al. **Diabetes flow sheet use associated with guideline adherence.** Ann Fam Med. 2008 May-Jun;6(3):235-8.

22

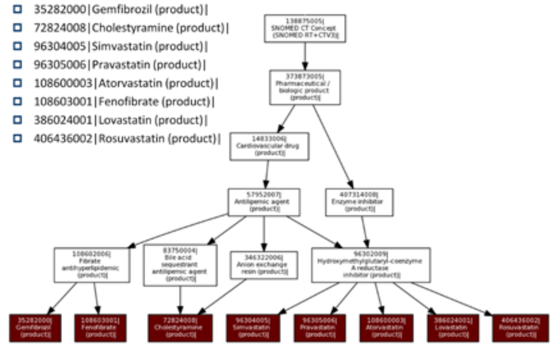
Smoking Status

- 160616005|Trying to give up smoking (finding)|
- 160617001|Stopped smoking (finding)|
- 365981007|Finding of tobacco smoking behavior (finding)|
- 65568007|Cigarette smoker (finding)|
- 77176002|Smoker (finding)|
- 8517006|Ex-smoker (finding)|

23

Lipid Lowering Agent

- 35282000|Gemfibrozil (product)|
- 72824008|Cholestyramine (product)|
- 96304005|Simvastatin (product)|
- 96305006|Pravastatin (product)|
- 108600003|Atorvastatin (product)|
- 108603001|Fenofibrate (product)|
- 386024001|Lovastatin (product)|
- 406436002|Rosuvastatin (product)|



24

Adherence of Care Guidelines Results

	Criteria	Total	Percent
Assessment Scores (20 points per criteria; Overall average: 43.2)			
1.	Haemoglobin A1c level tested in the past six months	86	39.4%
2.	Urine microalbumin level tested in the past year	73	33.5%
3.	Low-density lipoprotein (LDL-cholesterol) level tested in the past year	129	59.2%
4.	Smoking Status	23	10.6%
5.	Blood pressure	160	73.4%
	Criteria met: **5** (n=4, 1.8%), **4** (n=45, 20.6%), **3** (n=47, 21.6%), **2** (n=47, 21.6%), **1** (n=36, 16.5%), **0** (n=39, 17.9%)		
Treatment Scores (25 points per criteria; Overall average: 39.8)			
1.	Haemoglobin A1c level at or below 8% or use of a hypoglycemic agent	78	35.8%
2.	Urine microalbumin level below 30 mg/g creatine or use of an angiotensin-converting enzyme (ACE) inhibitor or angiotensin II receptor blocker (ARB)	58	26.6%
3.	LDL-cholesterol below 100 mg/dL or use of lipid-lowering agent	97	44.5%
4.	Blood pressure at or below 130/85 mm Hg or use of an antihypertensive agent	114	52.3%
	Criteria met: **4** (n=21, 9.6%), **3** (n=34, 15.6%), **2** (n=57, 26.1%), **1** (n=47, 21.6%), **0** (n=59, 27.1%)		
Target Attainment Scores (33.3 points per criteria; Overall average: 46.1)			
1.	Haemoglobin A1c level less than 8%	67	30.7%
2.	LDL-cholesterol level less than or equal to 100 mg/dL	126	57.8%
3.	Blood pressure less than or equal to 130/85 mm Hg	109	50.0%
	Criteria met: **3** (n=34, 15.5%), **2** (n=75, 34.4%), **1** (n=50, 22.9%) and **0** (n=59, 27.1%)		

25

Summary of Meaningful Use of SNOMED CT

- Improved Standardisation
- Improved Completeness
  - Problem List
  - Encounter Diagnoses
  - Diagnostic Codes
- Simplified and Accurate Retrieval
- Facilitated Decision Support
  - E.g.
    - Care Guidelines
    - Allergies

26

Benefits of Using SNOMED CT Versus ***

***	Challenges of ***	SNOMED CT Benefits	Facilitator
Free Text	Spelling mistakes; Local jargon, abbreviations and acronyms	Standardised data; Codified data	Encoding Algorithms; Extensions
Local or Enterprise Data Dictionary	May not be concept-based; May contain duplicates	Semantic and hierarchical relationships; International standard; Semantic interoperability	Encoding Algorithms; Extensions
Classification System (e.g., ICD-9-CM)	Lack of granularity; Groups of concepts	Granularity	Cross maps
Other Standardised Terminologies		Fulfils US Meaningful Use criteria; Scope; Multilingual; Compositional; Extensible	Harmonisation; Unified Medical Language System (UMLS)

- LOINC, MeDRA, RxNorm

27

Application (EMR Prototype Demo)

Target / Stage	Patient	Practice	Population
Stage 1 (Capture)			
Stage 2 (Processes)			
Stage 3 (Outcomes)			

- Focus
  - All 3 stages of Meaningful Use
  - Patient/Practice
- Demonstrate
  - Automated generation of diagnostic codes
  - Completeness of problem list
  - Improved organisation of clinical record
  - Decision support
    - Reminders for ordering laboratory tests
    - Suggestions for prescribing medications
    - Alerts for allergies and/or adverse reactions
  - Patient recall / patient case queries

28

Patient – Stage 1: Data Capture

Target / Stage	Patient	Practice	Population
Stage 1 (Capture)			
Stage 2 (Processes)			
Stage 3 (Outcomes)			

- Importance
  - Completeness of problem list (US Meaningful Use criteria)
  - Completeness of billing/diagnostic code (financial/statistical implications)
  - Completeness of encounter diagnoses
- Materials/Methods
  - SNOMED CT to ICD-9-CM reimbursement cross map
- Demo (Patient Id: 6280)
  - Suggest billing code (backache)
  - Suggest adding to problem list (hypertension)
  - Suggest encounter diagnosis (250.01)

29

Patient – Stage 2: Clinical Summaries

Target / Stage	Patient	Practice	Population
Stage 1 (Capture)			
Stage 2 (Processes)			
Stage 3 (Outcomes)			

- Importance
  - Improve data quality
  - Facilitate decision support
  - Semantically interoperable patient clinical summary
- Materials/Methods
  - SNOMED CT Queries
  - Continuity of Care Document (CCD)
- Demo (PatientId: 3966)
  - Suggest re-organisation (allergies)
  - Suggest re-organisation (immunisations)

30

Patient – Stage 3: Decision Support

- Importance
  - Adherence to evidence-based care guidelines
  - Avoidance of adverse event
- Materials/Methods
  - SNOMED CT defining attributes
  - Diabetes care guidelines
- Scenarios
  - Muscle ache, prescribe ibuprofen
- Demo
  - Alerts (Patient Id: 2753, 5252)

Target / Stage	Patient	Practice	Population
Stage 1 (Capture)			
Stage 2 (Processes)			
Stage 3 (Outcomes)			

31

Practice – Stage 1: Patient Case Queries

- Importance
  - Patient recall
  - Patient case queries
  - Reporting of quality measures
- Materials/Methods
  - SNOMED CT defining attributes
  - Clinical Quality Measures (CQM)

Target / Stage	Patient	Practice	Population
Stage 1 (Capture)			
Stage 2 (Processes)			
Stage 3 (Outcomes)			

32

Practice – Stage 2: Standardised Reporting

- Importance
  - Standardised, comparable reports
- Materials/Methods
  - Quality Reporting Document Architecture (QRDA)

Target / Stage	Patient	Practice	Population
Stage 1 (Capture)			
Stage 2 (Processes)			
Stage 3 (Outcomes)			

33

Practice – Stage 3: Auditing (Critiquing)

- Importance
  - Audit (critique) evidence-based guideline adherence
  - Assess patient conditions
- Materials/Methods
  - Diabetes care guidelines

Target / Stage	Patient	Practice	Population
Stage 1 (Capture)			
Stage 2 (Processes)			
Stage 3 (Outcomes)			

34

Discussion

- Does it make sense?
- Does it add value?
- What are the potential implementation issues?
- Feedback on SNOMED CT Meaningful Use Framework

Target / Stage	Patient	Practice	Population
Stage 1 (Capture)			
Stage 2 (Processes)			
Stage 3 (Outcomes)			

35

12.8.2 Technical Feedback Session Slides

Towards Demonstrating the Meaningful Use of SNOMED CT in Primary Care: Technical Feedback Session

Dennis Lee Hon Kit
dlhk@uvic.ca

1

Objectives

- Proposed SNOMED CT Meaningful Use Framework
- Retrieval Basics
- SNOMED CT retrieval objectives
  - How to identify potential drug-allergy interactions
  - How to improve the completeness of a problem list and clinical summary
  - How to structure expressions for patient case queries and patient recall
  - How to demonstrate the value of SNOMED CT over free text queries
- SNOMED CT API
- Challenges
- Discussion/Feedback
  - Does it make sense?
  - Does it add value?
  - What are the potential implementation issues?

2

Proposed SNOMED CT Meaningful Use Framework

Meaningful Use Stage / Data Use Target Audience	Patient	Practice	Population
Stage 1: Data Capture and Sharing - Electronically capturing health information in a standardised format - Using that information to track key clinical conditions - Initiating the reporting of clinical quality measures and public health information	- Suggest diagnostic code based on encounter diagnosis - Suggest encounter diagnosis based on diagnostic code - Suggest adding encounter diagnosis to problem list	- Audit (critique) completeness of problem list and encounter diagnosis - Audit (critique) accuracy of diagnostic codes - Patient recall / patient case queries	- Generate statistics on key clinical quality measures (CQM)
Stage 2: Advance Clinical Processes - Electronic transmission of patient care summaries across multiple settings	- Generate patient clinical summary in the form of the Continuity Care Document (CCD)	- Generate practice quality reports in the form of the Quality Reporting Document Architecture (QRDA) Category II and III	- Generate practice quality reports in the form of the Quality Reporting Document Architecture (QRDA) Category III
Stage 3: Improved Outcomes - Improving quality, safety, and efficiency, leading to improved outcomes - Decision support for national high-priority conditions	- Alert clinicians to drug-allergy interactions - Remind clinicians to order routine tests - Suggest medications based on out of range test results	- Audit (critique) evidence-based guideline adherence - Assess patient conditions	- Compare clinical outcomes between practice and population

3

Retrieval Basics

Introduction to SNOMED CT Retrieval

4

Type of Retrieval

- 73211009|Diabetes mellitus (disorder)|

Concept	Related
73211009|Diabetes mellitus (disorder)|	Equivalent
75934005|Metabolic disease (disorder)|	SuperType
313435000|Type I diabetes mellitus without complication (disorder)|	SubType
80394007|Hyperglycemia (disorder)|	Sibling
74627003|Diabetic complication (disorder)|	Defining Attributes
161445009|History of - diabetes mellitus (situation)|	Context

5

Retrieval Components

- Predicate
  - Concept or Expression
  - Core or Extension
- Types of Retrieval
  - Equivalent
  - Subsumption
  - Reverse Subsumption
  - Sibling
  - Defining Attributes
  - Context
- Include/Exclude
  - Include
  - Exclude
- Candidate
  - Single Expression
  - List of Expressions (Subset)
  - All of SNOMED CT
- Release Version*
  - SNOMED CT Release Version
- Extension*
  - SNOMED CT Namespace

- Predicate
  - 73211009|Diabetes mellitus (disorder)|
- Type of Retrieval
  - EquivalentAndSubsumption
- Include/Exclude
  - Include
- Candidate
  - Problem List Subset
- Release Version
  - January 31, 2013 International Release Version
- Extension
  - Canadian Extension 1000087

6

Prerequisite Tables

No	Table	Rationale
1.	SCT_Concepts	IsPrimitive (RF2: definitionStatusId)
2.	SCT_Relationships	Defining attribute(s)
3.	SCT_Canonical	Proximal primitive(s)
4.	SCT_TransitiveClosure	Efficient check for supertype/subtype relationship
5.	SCT_ConceptModel	Machine Readable Concept Model (MRCM)

7

Transitive Closure Table

Relationships Table

ConceptId1	RelationshipType	ConceptId2
279001004|Pain finding at anatomical site (finding)|	116680003|Is a (attribute)|	22253000|Pain (finding)|
161891005|Backache (finding)|	116680003|Is a (attribute)|	279001004|Pain finding at anatomical site (finding)|
279039007|Low back pain (finding)|	116680003|Is a (attribute)|	161891005|Backache (finding)|
278860009|Chronic low back pain (finding)|	116680003|Is a (attribute)|	279039007|Low back pain (finding)|

Transitive Closure Table

SuperTypeId	SubTypeId
22253000|Pain (finding)|	279001004|Pain finding at anatomical site (finding)|
22253000|Pain (finding)|	161891005|Backache (finding)|
22253000|Pain (finding)|	279039007|Low back pain (finding)|
22253000|Pain (finding)|	278860009|Chronic low back pain (finding)|

395,346 concepts
539,245 "is a" relationships
5,891,916 transitive closure

Note: The Technical Implementation Guide includes a MySQL procedure to generate the transitive closure table, which takes about 5 minutes to generate.

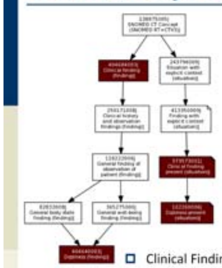
8

Pre-coordination Predicate and Candidate

- SQL Query
  - SELECT *
    FROM
    SCT_TransitiveClosure
    WHERE
    SuperTypeId=Predicate AND
    SubTypeId=Candidate
- Possible Results
  - 1 record returned (equivalent or subsumption)
  - 0 records returned (not related)
- Example
  - SELECT *
    FROM
    SCT_TransitiveClosure
    WHERE
    SuperTypeId=22253000 AND
    SubTypeId=278860009

9

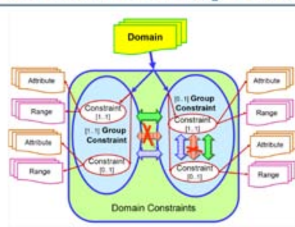
Be Aware of Redundancies



- Clinical Findings
  - 404684003|Clinical finding (finding)|
  - 373573001|Clinical finding present (situation)|
  - 410515003|Known present (qualifier value)|
- Procedures
  - 71388002|Procedure (procedure)|
  - 443938003|Procedure carried out on subject (situation)|
  - 385658003|Done (qualifier value)|

10

Machine Readable Concept Model

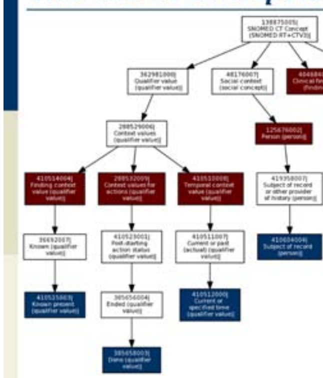


Domain	Attribute	Range	Permissible
123037004|Body structure (body structure)|	272741003|Laterality (attribute)|	182353008|Side (qualifier value)|	<=
404684003|Clinical finding (finding)|	246075003|Causative agent (attribute)|	105590001|Substance (substance)|	<<
404684003|Clinical finding (finding)|	42752001|Due to (attribute)|	272379006|Event (event)|	<<
404684003|Clinical finding (finding)|	363698007|Finding site (attribute)|	442083009|Anatomical or acquired body structure (body structure)|	<<

Note: The Machine Readable Concept Model is not part of the standard SNOMED CT distribution but can be obtained from the IHTSDO.

11

SNOMED CT Concept Model and Context



- 246090004|Associated finding (attribute)|
  - 404684003|Clinical finding (finding)|
- 363589002|Associated procedure (attribute)|
  - 71388002|Procedure (procedure)|
- 408729009|Finding context (attribute)|
  - 410514004|Finding context value (qualifier value)|
- 408730004|Procedure context (attribute)|
  - 288532009|Context values for actions (qualifier value)|
- 408731000|Temporal context (attribute)|
  - 410510008|Temporal context value (qualifier value)|
- 408732007|Subject relationship context (attribute)|
  - 125676002|Person (person)|

12

SNOMED CT Concept Model and Context, 2

Context Attribute	408729009|Finding context (attribute)|	408730004|Procedure context (attribute)|	408731000|Temporal context (attribute)|	408732007|Subject relationship context
Context Range	410514004|Finding context value (qualifier value)|	288532009|Context values for actions (qualifier value)|	410510008|Temporal context value (qualifier value)|	125676002|Person (person)|
Context Default	410515003|Known present (qualifier value)|	385658003|Done (qualifier value)|	410512000|Current or specified time (qualifier value)|	410604004|Subject of record (person)|
Context SubTypes	19	51	7	426
Context Examples	410516002|Known absent (qualifier value)|	443390004|Refused (qualifier value)|	410513005|In the past (qualifier value)|	444301002|Mother of subject (person)|

13

Normalisation (for Post-coordination)

- To facilitate complete and accurate retrieval of pre- and post-coordinated SNOMED CT expressions
- Normalisation Steps
  - Retrieve proximal primitive concept
  - Retrieve defining attributes
  - Merge defining attributes with any post-coordinated attributes
  - Recursively normalisation
- Closer-to-user form
  - 90834002|Pain in limb (finding)|
- Long normal form
  - 22253000|Pain (finding)|:
    363698007|Finding site (attribute)|=
    66019005|Limb structure (body structure)|

14

Testing for Subsumption

- 243796009|Situation with explicit context (situation)|:
  {246090004|Associated finding (attribute)|=
  38341003|Hypertensive disorder, systemic arterial (disorder),
  408729009|Finding context (attribute)|=
  410515003|Known present (qualifier value)|,
  408731000|Temporal context (attribute)|=
  410513005|In the past (qualifier value)|,
  408732007|Subject relationship context (attribute)|=
  410604004|Subject of record (person)|}
- 243796009|Situation with explicit context (situation)|:
  {246090004|Associated finding (attribute)|=
  38341003|Hypertensive disorder, systemic arterial (disorder)|,
  408729009|Finding context (attribute)|=
  410516002|Known absent (qualifier value)|,
  408731000|Temporal context (attribute)|=
  410589000|All times past (qualifier value)|,
  408732007|Subject relationship context (attribute)|=
  444148008|Person in family of subject (person)|}

15

Testing for Equivalency

- Cancer, left lung
  - 363346000|Malignant neoplastic disease (disorder)|:
    363698007|Finding site (attribute)|=
    44029006|Left lung structure (body structure)|
- Lung cancer, left
  - 363358000|Malignant tumor of lung (disorder)|:
    363698007|Finding site (attribute)|=
    44029006|Left lung structure (body structure)|
- Disease, cancer, left lung
  - 64572001|Disease (disorder)|:
    {116676008|Associated morphology (attribute)|=
    367651003|Malignant neoplasm of primary, secondary, or uncertain origin (morphologic abnormality)|,
    363698007|Finding site (attribute)|=(
    39607008|Lung structure (body structure)|:
    272741003|Laterality (attribute)|=
    7771000|Left (qualifier value)|
)}
- Why are they equivalent?
  - 243796009:{246090004=(64572001:{116676008=367651003,363698007=(39607008:272741003=7771000)}),408729009=410515003,408731000=410512000,408732007=410604004}

16

Subsumption Queries

- Type of Predicate Expression

	Primitive	Fully defined
Pre-coordination	#1	#3
Post-coordination	#2	

1. Pre-coordinated concept that is primitive
   - Retrieve only candidates with focus concept that is subtype of predicate concept
2. Post-coordinated expression with primitive focus concept
   - Retrieve supertype concepts of focus concept
   - Additional refinements may be subsumed by supertype concept
3. Pre-coordinated concept that is fully-defined or post-coordinated expression with a fully-defined concept
   - Do not exclude candidates with a focus concept that is a supertype of the focus concept of the predicate

17

Retrieval Objective #1

How to identify potential
drug-allergy interactions

18

Coding Drugs: Substance or Product?

- Ibuprofen
  - 105590001|Substance (substance)|
    - 387207008|Ibuprofen (substance)|
      - 425516000|Ibuprofen lysine (substance)|
  - 373873005|Pharmaceutical / biologic product (product)|
    - 38268001|Ibuprofen (product)|
      - 350322005|Topical form ibuprofen (product)|
        - 330202001|Ibuprofen 5% gel (product)|
        - 330232009|Ibuprofen 5% cream (product)|
      - 350321003|Oral form ibuprofen (product)|
        - 329652003|Ibuprofen 200mg tablet (product)|
        - 370195008|Ibuprofen 200mg capsule (product)|
  - 38268001|Ibuprofen (product)|:
    127489000|Has active ingredient (attribute)|=
    387207008|Ibuprofen (substance)|

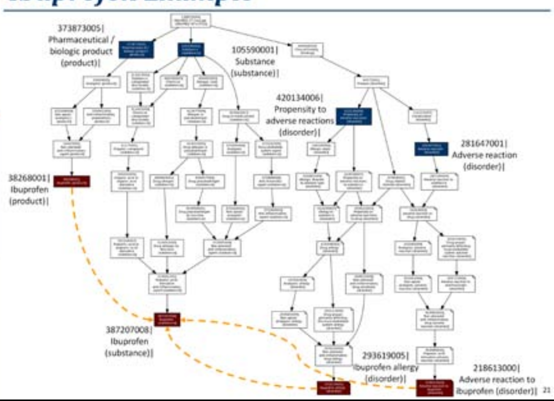
19

Adverse Reactions in SNOMED CT

Note: The July 2013 international release version of SNOMED CT will include improvements to the allergies hierarchy (June 2013 IHTSDO Newsletter)

20

Ibuprofen Example



21

Defining Attributes of Ibuprofen Concepts

- 105590001|Substance (substance)|
  - 387207008|Ibuprofen (substance)|
- 373873005|Pharmaceutical / biologic product (product)|
  - 38268001|Ibuprofen (product)|:
    127489000|Has active ingredient (attribute)|=
    387207008|Ibuprofen (substance)|
- 420134006|Propensity to adverse reactions (disorder)|
  - 293619005|Ibuprofen allergy (disorder)|:
    246075003|Causative agent (attribute)|=
    387207008|Ibuprofen (substance)|,
    363705008|Has definitional manifestation (attribute)|=
    416093006|Allergic reaction to drug (disorder)|
- 281647001|Adverse reaction (disorder)|
  - 218613000|Adverse reaction to ibuprofen (disorder)|:
    246075003|Causative agent (attribute)|=
    387207008|Ibuprofen (substance)|

22

Deriving Propensity From Adverse Reaction

- Identify adverse reaction
  - 218613000|Adverse reaction to ibuprofen (disorder)|:
- Identify 246075003|Causative agent (attribute)|
  - SELECT ConceptId2 FROM SCT_Relationships WHERE RelationshipType= 246075003 AND ConceptId1=218613000
  - 218613000|Adverse reaction to ibuprofen (disorder)|:
    246075003|Causative agent (attribute)|=
    387207008|Ibuprofen (substance)|
- Search for concept that
  - Is subtype of:
    - 420134006|Propensity to adverse reactions (disorder)|
  - Has defining attribute:
    - 246075003|Causative agent (attribute)|=
      387207008|Ibuprofen (substance)|
  - Sample SQL:
    - SELECT R.ConceptId1 FROM SCT_TransitiveClosure TC, SCT_Relationships R WHERE TC.SuperTypeId=420134006 AND TC.SubTypeId=R.ConceptId1 AND R.RelationshipType=246075003 AND R.ConceptId2=387207008

23

Deriving Propensity From Adverse Reaction, 2

- If found
  - Multiple concepts, use supertype
    - 294707002|Prednisolone allergy (disorder)|
  - Single concept
    - 293619005|Ibuprofen allergy (disorder)|
- If not found, post-coordinate
  - 419511003|Propensity to adverse reactions to drug (disorder)|:
    246075003|Causative agent (attribute)|=
    255632006|Anticonvulsant (substance)|
- Always check if pre-coordinated concept exists
  - 293857007|Antiepileptic allergy (disorder)|
- Caution
  - Primitive concepts
  - Under-defined concepts

24

Applying Drug-Allergy Alerts

- Identify drug prescribed
  - 329652003|Ibuprofen 200mg tablet (product)|
- Identify 127489000|Has active ingredient (attribute)|
  - SELECT ConceptId2 FROM SCT_Relationships WHERE RelationshipType=127489000 AND ConceptId1=329652003
  - 329652003|Ibuprofen 200mg tablet (product)|:
    127489000|Has active ingredient (attribute)|=
    387207008|Ibuprofen (substance)|
- Create predicate expression with substance
  - 419511003|Propensity to adverse reactions to drug (disorder)|:
    246075003|Causative agent (attribute)|=
    387207008|Ibuprofen (substance)|

25

Applying Drug-Allergy Alerts, 2

- Compare expressions
  - Predicate (Patient's allergy)
    - 293619005|Ibuprofen allergy (disorder)|
  - Candidate (Potential medication allergy)
    - 419511003|Propensity to adverse reactions to drug (disorder)|:
      246075003|Causative agent (attribute)|=
      387207008|Ibuprofen (substance)|

26

Why Use Subsumption?

- Allergy
  - 91936005|Allergy to penicillin (disorder)|:
    246075003|Causative agent (attribute)|
    373270004|Penicillin -class of antibiotic- (substance)|
- Medication
  - 39359008|Penicillin V (product)|:
    127489000|Has active ingredient (attribute)|=
    372725003|Penicillin V (substance)|
- Reasons
  - Allergy may be recorded at broader category
  - Exact concept id search in expression is inadequate

27

Suggest Medications

- Hahn Treatment Guidelines
  - Haemoglobin A1c level ≤8% or
  - 312064005|Hypoglycemic agent (product)|
- Use Guidelines as Decision Support
  - If Haemoglobin A1c level >8%
  - Suggest Medication Sans Allergy
    - 80870001|Glyburide (product)|
    - ~~109081006|Metformin (product)|~~
    - 325238000|Gliclazide (product)|
    - 326058001|Pioglitazone (product)|

Hahn KA, et al. **Diabetes flow sheet use associated with guideline adherence.** Ann Fam Med. 2008 May-Jun;6(3):235-8.

28

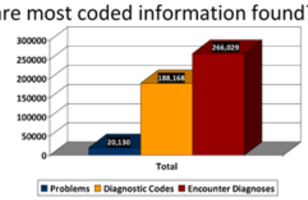
Retrieval Objective #2

How to improve the completeness of a problem list and clinical summary

29

Problem List and Clinical Summaries

- What should go into a problem list?
  - Non-transitive illnesses
- What should go into a clinical summary?
  - Advance directive, allergies, family history, functional status, immunisations, medical equipment, medications, plan of care, problems, procedures, results, social history, vital signs
- Where are most coded information found?



30

Assigning Predicate Expressions

2.16.840.1.113883.10.20.1.13 **Summary Purpose**	2.16.840.1.113883.10.20.1.9 **Payers**	2.16.840.1.113883.10.20.1.1 **Advance Directive** • **Include:** 365870005|Finding of resuscitation status (finding)|	2.16.840.1.113883.10.20.1.5 **Functional Status** • **Include:** 105501005| Dependence on enabling machine or device (finding)|
2.16.840.1.113883.10.20.1.11 **Problems** • **Include:** 64572001|Disease (disorder)| • **Exclude:** 281647001|Adverse reaction (disorder)|	2.16.840.1.113883.10.20.1.4 **Family History** • **Include:** 57177007|Family history with explicit context (situation)|	2.16.840.1.113883.10.20.1.15 **Social History** • **Include:** 365949003|Health-related behavior finding (finding)|	2.16.840.1.113883.10.20.1.2 **Allergies, Adverse Reactions & Alerts** • **Include:** 281647001|Adverse reaction (disorder)| • **Include:** 420134006|Propensity to adverse reactions (disorder)|
2.16.840.1.113883.10.20.1.8 **Medications** • **Include:** 373873005| Pharmaceutical / biologic product (product)|	2.16.840.1.113883.10.20.1.7 **Medical Equipment** • **Include:** 303607000| Biomedical equipment (physical object)|	2.16.840.1.113883.10.20.1.6 **Immunisations** • **Include:** 127785005| Administration of substance to produce immunity, either active or passive (procedure)|	2.16.840.1.113883.10.20.1.16 **Vital Signs** • **Include:** 118227000|Vital signs finding (finding)|
2.16.840.1.113883.10.20.1.14 **Results** • **Include:** 442618008|Abnormal finding on evaluation procedure (finding)|	2.16.840.1.113883.10.20.1.12 **Procedures** • **Include:** 387713003|Surgical procedure (procedure)|	2.16.840.1.113883.10.20.1.3 **Encounters**	2.16.840.1.113883.10.20.1.10 **Plan of Care** • **Include:** 243120004|Regimes and therapies (regime/therapy)|

HL7. **HL7 Implementation Guide: CDA Release 2 – Continuity of Care Document (CCD).** April 1, 2007.

31

Example of Results of Clinical Summary

Summary Purpose	**Payers**	**Advance Directive (30)** 304253006|Not for resuscitation (finding)|	**Functional Status** N/A
Problems (*) 73211009|Diabetes mellitus (disorder)|	**Family History (80)** 312824007|Family history of cancer of colon (situation)|	**Social History (4.2k)** 65568007|Cigarette smoker (finding)|	**Allergies, Adverse Reactions & Alerts (130)** 292055008|Codeine adverse reaction (disorder)|
Medications N/A	**Medical Equipment (100)** 87405001|Cane, device (physical object)|	**Immunisations (2.4k)** 86198006|Influenza vaccination (procedure)|	**Vital Signs (500)** 61086009|Pulse irregular (finding)|
Results (70) 313172000|Colonoscopy abnormal (finding)|	**Procedures (800)** 52734007|Total replacement of hip (procedure)|	**Encounters**	**Plan of Care (2.2k)** 67516001|Detoxification therapy (regime/therapy)|

32

Improved Completeness

- Allergies, Adverse Reactions and Alerts
  - 292055008|Codeine adverse reaction (disorder)|
  - 426232007|Environmental allergy (disorder)|
- Social History
  - 228281002|Problem drinker (finding)|
- Immunisation
  - 86198006|Influenza vaccination (procedure)|
  - 12866006|Pneumococcal vaccination (procedure)|
- Vital Signs
  - 271823003|Tachypnea (finding)|
- Plan of Care
  - 225358003|Wound care (regime/therapy)|
  - 385955003|Foot care (regime/therapy)|
- Problem (Sample)
  - 54150009|Upper respiratory infection (disorder)|
  - 19943007|Cirrhosis of liver (disorder)|
  - 10743008|Irritable bowel syndrome (disorder)|
  - 371088008|Reactive airways dysfunction syndrome (disorder)|
  - 68566005|Urinary tract infectious disease (disorder)|
  - 281255004|Small bowel obstruction (disorder)|

33

Adding Encounter Diagnosis to Problem List

- Criteria #1 – Include all diseases
  - 64572001|Disease (disorder)|
  - EquivalentAndSubsumption
  - Include
- Criteria #2 – Exclude adverse reactions
  - 281647001|Adverse reaction (disorder)|
  - EquivalentAndSubsumption
  - Exclude
- Criteria #3 – Exclude propensity to adverse reactions
  - 420134006|Propensity to adverse reactions (disorder)|
  - EquivalentAndSubsumption
  - Exclude
- Criteria #4 – Exclude if more specific disease is in problem list
  - Problem List: 73211009|Diabetes mellitus (disorder)|
  - Encounter Diagnosis: 46635009|Diabetes mellitus type 1 (disorder)|

34

Retrieval Objective #3

How to structure expressions for patient case queries and patient recall

35

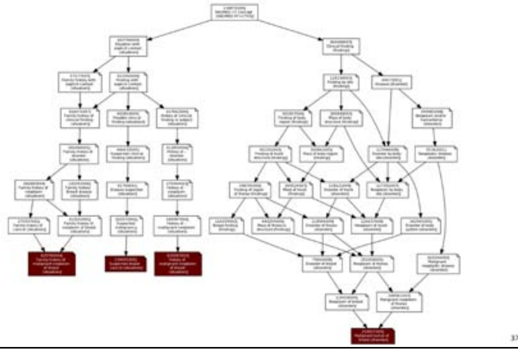
Question #1

- Patients at risk of breast cancer
  - Have a family history of breast cancer
  - Do not have a history of breast cancer
- Structuring the Query
  - Query #1 – include patients
    - 429740004|Family history of malignant neoplasm of breast (situation)|
  - Query #2 – include patients
    - 373572006|Clinical finding absent (situation)|
      246090004|Associated finding (attribute)|:
      254837009|Malignant tumor of breast (disorder)|
  - Query #3 – exclude patients
    - 254837009|Malignant tumor of breast (disorder)|

36

Context Free Queries

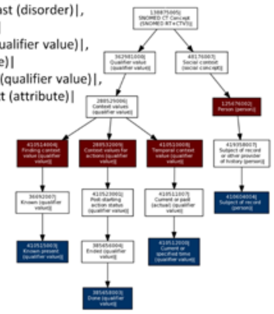
- What breast cancer expressions are in the dataset?



37

Structuring a Context Free Expression

- 243796009|Situation with explicit context (situation)|:
  246090004|Associated finding (attribute)|
  254837009|Malignant tumor of breast (disorder)|,
  408729009|Finding context (attribute)|
  410514004|Finding context value (qualifier value)|,
  408731000|Temporal context (attribute)|
  410510008|Temporal context value (qualifier value)|,
  408732007|Subject relationship context (attribute)|
  125676002|Person (person)|



38

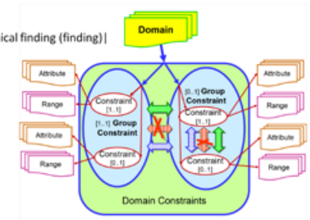
Question #2

- Number of diabetic patients in clinic
- Structuring the Query
  - Query #1 – include patients
    - 73211009|Diabetes mellitus (disorder)|
  - Query #2 – include patients
    - 404684003|Clinical finding (finding)|:
      47429007|Associated with (attribute)|=
      73211009|Diabetes mellitus (disorder)|
  - Query #3 – include patients
    - 404684003|Clinical finding (finding)|:
      363705008|Has definitional manifestation (attribute)|=
      73211009|Diabetes mellitus (disorder)|
  - Query #4 – include patients
    - 71388002|Procedure (procedure)|:
      363702006|Has focus (attribute)|=
      73211009|Diabetes mellitus (disorder)|
  - Query #5 – exclude patients?
    - 11687002|Gestational diabetes mellitus (disorder)|

39

Defining Attributes Queries

- Step 1 – Determine where focus concept occurs in Concept Model range
  - SQL Statement
    - SELECT CM.RangeId FROM SCT_ConceptModel CM, SCT_TransitiveClosure TC WHERE CM.RangeId=TC.SuperTypeId AND TC.SubTypeId=73211009|Diabetes mellitus (disorder)|
  - Results
    - 404684003|Clinical finding (finding)|



40

Defining Attributes Queries, 2

- Step 2 – Determine applicable domains and attributes
  - SQL Statement
    - SELECT CM.DomainId, CM.AttributeId FROM SCT_ConceptModel CM WHERE CM.RangeId=404684003|Clinical finding (finding)|
  - Results
    - 404684003|Clinical finding (finding)|
      47429007|Associated with (attribute)|
    - 404684003|Clinical finding (finding)|
      42752001|Due to (attribute)|
    - 404684003|Clinical finding (finding)|
      255234002|After (attribute)|
    - 404684003|Clinical finding (finding)|
      363705008|Has definitional manifestation (attribute)|
    - 71388002|Procedure (procedure)|
      363702006|Has focus (attribute)|
    - 272379006|Event (event)|
      47429007|Associated with (attribute)|
    - 272379006|Event (event)|
      42752001|Due to (attribute)|
    - 272379006|Event (event)|
      255234002|After (attribute)|

41

Defining Attributes Queries, 3

- Step 3 – Consolidate subtype domains and attributes
  - Attributes
    - 47429007|Associated with (attribute)|
      - 42752001|Due to (attribute)|
      - 255234002|After (attribute)|
  - Results
    - 404684003|Clinical finding (finding)|:
      47429007|Associated with (attribute)|=
      404684003|Clinical finding (finding)|
    - 404684003|Clinical finding (finding)|:
      363705008|Has definitional manifestation (attribute)|=
      404684003|Clinical finding (finding)|
    - 71388002|Procedure (procedure)|:
      363702006|Has focus (attribute)|=
      404684003|Clinical finding (finding)|
    - 272379006|Event (event)|:
      47429007|Associated with (attribute)|=
      404684003|Clinical finding (finding)|

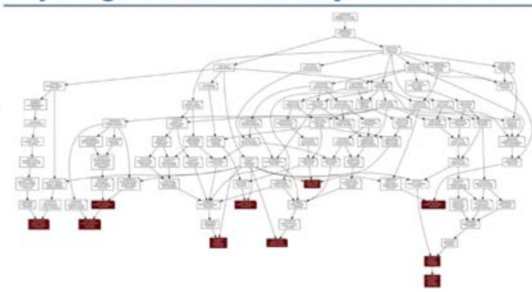
42

Defining Attributes Queries, 4

- Locate range in Concept Model
  - SELECT CM.DomainId, CM.AttributeId FROM SCT_TransitiveClosure TC, SCT_ConceptModel CM WHERE CM.RangeId=TC.SuperTypeId AND TC.SubTypeId=73211009|Diabetes mellitus (disorder)|
- Remove subtype attributes
  - 47429007|Associated with (attribute)|
    - 246075003|Causative agent (attribute)|
    - 255234002|After (attribute)|
    - 42752001|Due to (attribute)|
- Structuring predicate expression
  - 404684003|Clinical finding (finding)|:
    47429007|Associated with (attribute)|=
    73211009|Diabetes mellitus (disorder)|
  - 404684003|Clinical finding (finding)|:
    363705008|Has definitional manifestation (attribute)|=
    73211009|Diabetes mellitus (disorder)|
  - 71388002|Procedure (procedure)|:
    363702006|Has focus (attribute)|=
    73211009|Diabetes mellitus (disorder)|

43

Defining Attributes Example: Diabetes



- 127013003|Diabetic renal disease (disorder)|
- 230572002|Diabetic neuropathy (disorder)|
- 238982009|Diabetic dermopathy (disorder)|
- 280137006|Diabetic foot (disorder)|
- 34140002|Diabetic gastroparesis (disorder)|
- 371087003|Diabetic foot ulcer (disorder)|
- 420422005|Ketoacidosis in diabetes mellitus (disorder)|
- 424736006|Diabetic peripheral neuropathy (disorder)|
- 4855003|Diabetic retinopathy (disorder)|
- 59276001|Proliferative diabetic retinopathy (disorder)|

44

Question #3

- Meaningful Use Clinical Quality Measures (CQM)
  - 0062: Diabetes: Urine Protein Screening
  - Patients with a screening for nephropathy or evidence of nephropathy during the measurement period
- Structuring the Query
  - Query #1
    - 171234009|Nephropathy screening (procedure)|
  - Query #2
    - 90708001|Kidney disease (disorder)|
  - Query #3
    - 373572006|Clinical finding absent (situation)|:
      246090004|Associated finding (attribute)|=
      90708001|Kidney disease (disorder)|

45

Retrieval Objective #4

How to demonstrate the value of
SNOMED CT over free text queries

46

Comparing Data Retrieval Methods

- Materials
  - Encounters from primary care EMR
    - Unique descriptions: 15,123
    - Total descriptions: 266,032
- Methods
  - Free text search string
    - Wildcard (like) – WHERE Diagnosis LIKE '%xxx%'
    - Full-text index – WHERE MATCH (Diagnosis) AGAINST ('xxx')
  - SNOMED CT
    - Equivalent
    - Subsumption
    - Defining Attributes

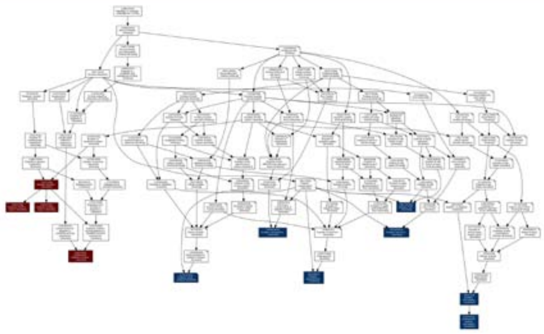
47

Data Retrieval Example 1: Diabetes

- Free Text Search
  - Diabetes – 3,894 (66 descriptions)
    - Diabetes Mellitus Non Insulin Dependent (Type 2) – 1,046
    - Diabetes Mellitus Insulin Dependent – 670
    - ...
    - Diabetes Mellitus Risk – 68 (false positive)
    - Diabetes Insipidus – 10 (false positive)
    - Pre-diabetes – 1 (false positive)
- SNOMED CT Search
  - 73211009|Diabetes mellitus (disorder)| – 3,837
    - 44054006|Diabetes mellitus type 2 (disorder)| – 2,002
    - 73211009|Diabetes mellitus (disorder)| – 1,052
    - 46635009|Diabetes mellitus type 1 (disorder)| – 775
    - 11687002|Gestational diabetes mellitus (disorder)| – 8
  - 404684003|Clinical finding (finding)|:
    47429007|Associated with (attribute)|=
    73211009|Diabetes mellitus (disorder)| – 60
    - 127013003|Diabetic renal disease (disorder)| – 25
    - 238982009|Diabetic dermopathy (disorder)| – 1
    - 280137006|Diabetic foot (disorder)| – 12
    - 34140002|Diabetic gastroparesis (disorder)| – 1
    - 371087003|Diabetic foot ulcer (disorder)| – 9
    - 4855003|Diabetic retinopathy (disorder)| – 10
    - 59276001|Proliferative diabetic retinopathy (disorder)| – 2
- Benefits of SNOMED CT
  - Ability to retrieve disease specialisation
  - Ability to retrieve implicit conditions
  - Ability to retrieve terms with appropriate context

48

Implicit Diabetes Mellitus Concepts



49

False Positive Diabetes

- Pre-diabetes - (1)
  - 9414007|Impaired glucose tolerance (disorder)|
- Diabetes Insipidus - (10)
  - 15771004|Diabetes insipidus (disorder)|
- Diabetes Mellitus Risk - (68)
  - 243796009|Situation with explicit context (situation)|:
    {246090004|Associated finding (attribute)|= 73211009|Diabetes mellitus (disorder)|,
    408729009|Finding context (attribute)|= 410519009|At risk context (qualifier value)|}

50

Data Retrieval Example 2: Colon Cancer

- Free Text Search – 34 total
  - Colon Cancer Family History – 59 (false positive)
  - Colon Cancer – 29
  - Colon Cancer Family History Strong – 8 (false positive)
  - Colon Cancer Liver Metastasis – 5
  - Colon Cancer - positive family history – 1 (false positive)
- SNOMED CT Search – 40 total
  - 363406005|Malignant tumor of colon (disorder)| – 35
    - Colon Cancer – 29
    - Colon Cancer Liver Metastasis – 5
  - 363410008|Malignant tumor of sigmoid colon (disorder)| – 6
    - Colon sigmoid cancer adenocarcinoma – 6
- Benefits of SNOMED CT
  - Ability to retrieve disease specialisation
  - Ability to retrieve terms with appropriate context

51

Data Retrieval Example 3+

- SNOMED CT Searches
  - **69878008|Polycystic ovaries (disorder)|**
    - PCOS - (4)
    - Leventhal syndrome - (2)
    - Polycystic ovaries - (1)
    - Ovarian polycystic disease - (2)
    - Polycystic ovarian syndrome - (1)
  - **75478009|Poisoning (disorder)|**
    - Intoxication - (7)
    - Toxicity - (5)
    - Poisoning - (2)
  - **48500005|Delusional disorder (disorder)|**
    - Paranoia - (10)
    - Delusional disorder - (9)
    - Paranoid psychosis - (2)
  - **235595009|Gastroesophageal reflux disease (disorder)|**
    - GERD - (16)
    - Gastroesophageal reflux - (9)
    - Esophageal reflux - (2)
    - Acid reflux - (2)
- Benefits of SNOMED CT
  - Ability to retrieve terms with different synonyms

52

Retrieval via SNOMED CT vs Free Text

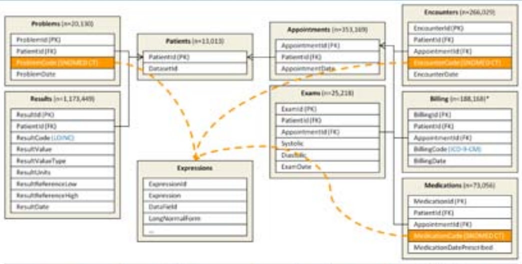
- Advantages
  - Ability to use single (or more) expression for query
  - Ability to retrieve disease specialisation
  - Ability to retrieve implicit conditions
  - Ability to retrieve terms with appropriate context
  - Ability to retrieve terms with different synonyms
- Challenges
  - More complicated algorithms than string searches
  - Ensure appropriate predicate expression used
  - Potentially slower

53

SNOMED CT API

54

Anonymised Primary Care Dataset



Encoding	Problems			Encounters			Medications		
	Total	%	Unique	Total	%	Unique	Total	%	Unique
Number of Records	**20,140**		**2,880**	**266,029**		**15,122**	**73,056**		**22,230**
Fully encoded	18,298	91	2,246	231,963	87	8,541	69,513	95	19,564
Partially encoded	1,666	8	582	27,198	10	5,794	0	0	0
Fully unencoded	177	1	43	6,149	2	779	3,543	5	2,666

55

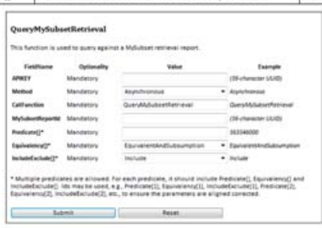
Application Programming Interface (API)

- Rational
  - Simplify implementation, separate terminology from patient records
- Method
  - Synchronous
  - Asynchronous
- General Functions
  - SCTGetAncestors, SCTGetDescendants
  - SCTGetConcept, SCTGetDescription, SCTGetRelationship
  - SCTGetCrossMap
  - SCTExpressionToExpression
- Normalisation and Compare Functions
  - SCTCompareExpression
  - SCTNormaliseExpression
- Subset Retrieval Functions
  - QueryMySubsetRetrieval

56

QueryMySubsetRetrieval API

No	Field Name	Description	Example
1.	APIKEY	Unique 36-character id to identify user	
2.	Method	Asynchronous or synchronous	Synchronous
3.	CallFunction	API function	QueryMySubsetRetrieval
4.	MySubsetReportId	Unique 36-character id to identify subset	
5.	Predicate[]	SNOMED CT predicate expression	73211009
6.	Equivalency[]	Type of result expected (e.g., equivalent, subsumption)	EquivalentAndSubsumption
7.	IncludeExclude[]	To include or exclude the results of the predicate	Include



57

API Data Flow

EMR

Clinician creates a new search in the problem list for patients with "diabetes." EMR identifies that "diabetes" is mapped to SNOMED CT concept 73211009 and sends a query to the SNOMED CT API for 73211009 in the problem list for subsumption

```
{"CallFunction":"QueryMySubsetRetrieval",
"MySubsetReportId":"XXXXX",
"Expressions":[{
   "Predicate":"73211009",
   "Equivalency":"EquivalentAndSubsumption",
   "IncludeExclude":"Include"}]
}
```

SNOMED CT Server

SNOMED CT retrieval service computes 73211009 against the problem list subset and returns four expressions: 73211009, 46635009, 44054006 and 11687002

```
{"CallFunction":"QueryMySubsetRetrieval",
"MySubsetReportId":"XXXXX",
"Expressions":[{
 73211009:"Equivalent",
 46635009:"Subsumption",
 44054006:"Subsumption",
11687002:"Subsumption"}
}
```

EMR receives the four SNOMED CT expressions and checks which patients records have been mapped

Assumption: Master problem list has been mapped to SNOMED CT and the SNOMED CT service has a copy of the problem list

58

Data Storage Options

No	Data Storage Method	Description	Support for Post-coordination	Impleme ntation Difficulty	Support for Data Retrieval	Recommendation
1.	Parsable Text Representation	Each expression is stored in a single row as a string	Unlimited	Easy	Inefficient	Recommended if only pre-coordination
2.	Unrestricted Relational Representation	Each expression is stored as multiple rows as attribute/value pairs	Unlimited	Complex	Somewhat Efficient	
3.	Restricted Relational Representation	Each expression is stored in a single row with multiple attribute/value pairs	Limited	Easy	Somewhat Efficient	Not recommended as limited support for post-coordination
4.	XML Representations (Plus JSON, etc)	Each expression is stored using XML	Unlimited	Moderate	Inefficient	Recommended only if implementers are familiar with XML and/or use an XML database

59

Proposed Enhanced Parasable Representation

No	Data Element	Description	Rationale	Example*
1.	HumanReadable	Human readable description		Family history of lung cancer
2.	CloserToUserForm	SNOMED CT expression		429011007|Family history of malignant neoplasm of lung (situation)|
3.	LongNormalForm	The long normal form of the expression.	To enable testing for equivalency quickly.	243796009:{246090004=(64572001:{116676008=367651003,363698007=39607008}),408729009=410515003,408731000=410511007,408732007=444148008}
4.	FocusConcept	The first concept before any post-coordination	Facilitates query efficiency	363358000|Malignant tumor of lung (disorder)|
5.	Proximal Primitive		Facilitates query efficiency	
6.	AssociatedFindingProcedure	Records the associated finding or procedure	Facilitate query efficiency	363358000|Malignant tumor of lung (disorder)|
7.	FindingProcedureContext	Records the finding or procedure context	Facilitates query efficiency	410515003|Known present (qualifier value)|
8.	TemporalContext	Records the temporal context	Facilitates query efficiency	410511007|Current or past (actual) (qualifier value)|
9.	SubjectRelationshipContext	Records the subject relationship context	Facilitates query efficiency	444148008|Person in family of subject (person)|
10.	ReleaseVersion	Release version used in the expression	Facilitate version control	20120131

60

Challenges

Challenges

1. Relationship Groups
2. Primitive Concepts
3. Modelling of Concepts
4. Errors in Modelling
5. Errors in Hierarchical Structure
6. Encoding Challenges

Challenge 1: Relationship Groups

Expression #1	Expression #2
211334007|Abrasion, ankle (disorder)|	399963005|Abrasion (disorder)|: 363698007|Finding site (attribute)|= 67269001|Skin structure of ankle (body structure)|
64572001|Disease (disorder)|: {116676008|Associated morphology (attribute)|= 400061001|Abrasion (morphologic abnormality)|, 363698007|Finding site (attribute)|= 67269001|Skin structure of ankle (body structure)|}	64572001|Disease (disorder)|: 116676008|Associated morphology (attribute)|= 400061001|Abrasion (morphologic abnormality)|, 363698007|Finding site (attribute)|= 67269001|Skin structure of ankle (body structure)|

- Expected Result
  - Expression #1 is equivalent to Expression #2
- Actual Result
  - Expression #1 subsumes Expression #2

Challenge 2: Primitive Concepts

Expression #1	Expression #2
53057004|Hand pain (finding)|	22253000|Pain (finding)|: 363698007|Finding site (attribute)|= 85562004|Hand structure (body structure)|
53057004|Hand pain (finding)|: 363698007|Finding site (attribute)|= 85562004|Hand structure (body structure)|	22253000|Pain (finding)|: 363698007|Finding site (attribute)|= 85562004|Hand structure (body structure)|

- Expected Result
  - Expression #1 is equivalent to Expression #2
- Actual Result
  - Expression #2 subsumes Expression #1

Challenge 3: Modelling of Concepts

Expression #1	Expression #2
274663001|Acute pain (finding)|	22253000|Pain (finding)|: 263502005|Clinical course (attribute)|= 373933003|Acute onset (qualifier value)|
301369003|Finding of pattern of pain (finding)|: 263502005|Clinical course (attribute)|= 424124008|Sudden onset AND/OR short duration (qualifier value)|	22253000|Pain (finding)|: 263502005|Clinical course (attribute)|= 373933003|Acute onset (qualifier value)|

- Expected Result
  - Expression #1 is equivalent to Expression #2
- Actual Result
  - Expression #1 is unrelated to Expression #2

Challenge 4: Errors in Modelling of Concepts

Expression #1	Expression #2
110168002|Abrasion of chin (disorder)|	399963005|Abrasion (disorder)|: 363698007|Finding site (attribute)|= 23747009|Skin structure of chin (body structure)|
64572001|Disease (disorder)|: {116676008|Associated morphology (attribute)|= 400061001|Abrasion (morphologic abnormality)|, 363698007|Finding site (attribute)|= 30291003|Chin structure (body structure)|} {116676008|Associated morphology (attribute)|= 400061001|Abrasion (morphologic abnormality)|, 363698007|Finding site (attribute)|= 73897004|Skin structure of face (body structure)|}	64572001|Disease (disorder)|: {116676008|Associated morphology (attribute)|= 400061001|Abrasion (morphologic abnormality)|, 363698007|Finding site (attribute)|= 23747009|Skin structure of chin (body structure)|}

- Expected Result
  - Expression #1 is equivalent to Expression #2
- Actual Result
  - Expression #1 subsumes Expression #2

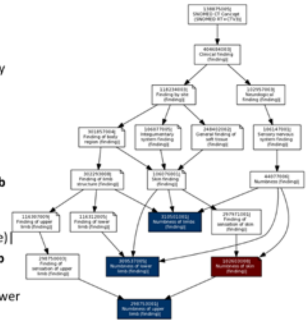
Challenge 5: Errors in Hierarchical Structure

- **102603008|Numbness of skin (finding)|**
  - 400199006|Structure of skin and/or surface epithelium (body structure)|
- **310501001|Numbness of limbs (finding)|**
  - 116370005|Skin structure of extremity (body structure)|
- **298753001|Numbness of upper limb (finding)|**
  - 371311000|Skin structure of upper extremity (body structure)|
- **309537005|Numbness of lower limb (finding)|**
  - 371304004|Skin structure of lower extremity (body structure)|

67

Challenge 5: Errors in Hierarchical Structure, 2

- **102603008|Numbness of skin (finding)|**
  - 400199006|Structure of skin and/or surface epithelium (body structure)|
- **310501001|Numbness of limbs (finding)|**
  - 116370005|Skin structure of extremity (body structure)|
- **298753001|Numbness of upper limb (finding)|**
  - 371311000|Skin structure of upper extremity (body structure)|
- **309537005|Numbness of lower limb (finding)|**
  - 371304004|Skin structure of lower extremity (body structure)|



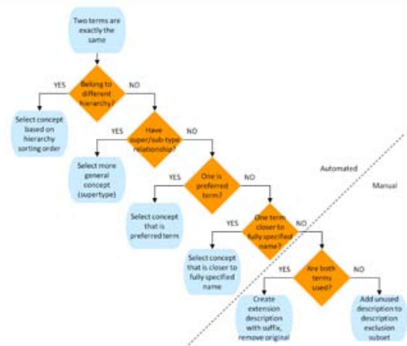
68

Challenge 6: Encoding Challenges

- Different Encoding Methods (e.g., "childhood asthma"
  - 161527007|History of - asthma (situation)|
  - 233678006|Childhood asthma (disorder)|
- Multiple Exact Matches (e.g. "abrasion")
  - 1779330015|Abrasion| - 400061001|Abrasion (morphologic abnormality)|
  - 1779244012|Abrasion| - 399963005|Abrasion (disorder)|
  - 195131013|Abrasion| - 8420001|Abrasion (procedure)|
- Inappropriate Body Structure (e.g. "abrasion of shin")
  - 399963005|Abrasion (disorder)|:
    363698007|Finding site (attribute)|=
    78234002|Shin structure (body structure)|
    or 18096008|Skin structure of shin (body structure)|
- Multiple Conditions (e.g., "bone, brain, liver metastasis)
  - 272673000|Bone structure (body structure)|
  - 12738006|Brain structure (body structure)|
  - 94381002|Secondary malignant neoplasm of liver (disorder)|
- Incomplete/Incorrect Post-coordination
  - Haemoglobin A1c normal
    - 43396009|Hemoglobin A1c measurement (procedure)|
    - 17621005|Normal (qualifier value)|
  - STD Screening
    - 8098009|Sexually transmitted infectious disease (disorder)|
    - 360156006|Screening - procedure intent (qualifier value)|

69

Sorting Algorithms



70

Encoding Templates

- Example #1: Procedure → Finding
  - Haemoglobin A1c normal
    - 43396009|Hemoglobin A1c measurement (procedure)|
    - 17621005|Normal (qualifier value)|
  - Normal Templates
    - 442082004|Measurement finding within reference range (finding)|:
      363714003|Interprets (attribute)|=
      122869004|Measurement procedure (procedure)|
    - 441742003|Evaluation finding (finding)|:
      363714003|Interprets (attribute)|=
      71388002|Procedure (procedure)|,
      363713009|Has interpretation (attribute)|=
      17621005|Normal (qualifier value)|
- Example #2: Disorder → Procedure
  - STD Screening
    - 8098009|Sexually transmitted infectious disease (disorder)|
    - 360156006|Screening - procedure intent (qualifier value)|
  - Screening Template
    - 312851005|Screening for disorder (procedure)|
      363702006|Has focus (attribute)|<<
      64572001|Disease (disorder)|

71

Discussion

- Does it make sense?
- Does it add value?
- What are the potential implementation issues?

72