



Delivering
SNOMED CT

A large, stylized green leaf graphic with a white outline, positioned on the left side of the page.

Data Analytics with SNOMED CT – Case Studies

MAY 2015

Date: 2015-05-21

Version: 1.00

Status: FINAL

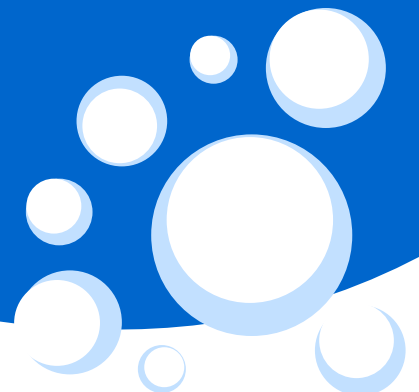


TABLE OF CONTENTS

1	INTRODUCTION.....	4
1.1	Background	4
1.2	Purpose	4
2	APPENDIX A – Project case studies	5
2.1	Overview	5
2.2	Data Migration Workbench (UK).....	5
2.3	Kaiser Permanente (USA)	8
2.4	National Medication Decision Support System (Denmark)	9
2.5	Semantic Search (Australia)	11
2.6	Radiology Activity (UK).....	13
3	APPENDIX B – Vendor Case Studies	15
3.1	Overview	15
3.2	3M Health Information Systems	15
3.3	Allscripts.....	17
3.4	Apelon.....	18
3.5	B2i Healthcare.....	19
3.6	Cambio	22
3.7	Caradigm	24
3.8	Cerner	25
3.9	Clinithink	29
3.10	EMIS	30
3.11	Epic.....	31
3.12	First Databank	32
3.13	Intelligent Medical Objects	33

Copyright and Acknowledgements

The document is a draft publication of the International Health Terminology Standards Development Organisation (IHTSDO), the association that owns and maintains SNOMED Clinical Terms. IHTSDO acknowledges the work undertaken by Malcolm Duncan in co-authoring this guide during his participation in the SNOMED CT Implementation Advisor (SIA) scheme.

© Copyright 2014 International Health Terminology Standards Development Organisation (IHTSDO), all rights reserved. SNOMED, SNOMED CT and IHTSDO are registered trademarks of the International Health Terminology Standards Development Organisation. SNOMED CT licensing information is available at www.ihtsdo.org/licensing. For more information about IHTSDO and IHTSDO Membership, please refer to www.ihtsdo.org or contact us at info@ihtsdo.org.

1 INTRODUCTION

1.1 BACKGROUND

SNOMED CT is widely recognized as the leading global clinical terminology for use in Electronic Health Records (EHRs). SNOMED CT is used within EHRs to support data capture, retrieval, and subsequent reuse for a wide range of purposes, from patient-based queries to operational reporting, public health reporting, strategic planning, predictive medicine and clinical research. As the SNOMED CT encoding of healthcare data increases, so too have the benefits being realized from analytics processes performed over this data.

1.2 PURPOSE

The purpose of this document is to present two sets of case studies that demonstrate data analytics with SNOMED CT:

- Firstly, a number of projects that implement or support analytics using SNOMED CT are described;
- Secondly, a variety of commercial tools which support analytics over SNOMED CT enabled data are presented.

These case studies form two appendices to accompany the report 'Data Analytics with SNOMED CT'. The purpose of this report is to describe current approaches, tools, and techniques for performing data analytics using SNOMED CT and to share developing practice in this area. This report aims to benefit members, vendors and users of SNOMED CT by promoting a greater awareness of both what has been achieved, and what can be achieved using SNOMED CT to enhance analytics services.

The 'Data Analytics with SNOMED CT' report can be found at <http://snomed.org/analytics.pdf>.

2 APPENDIX A – PROJECT CASE STUDIES

2.1 OVERVIEW

This appendix includes brief reviews of a variety of projects which implement or support analytics over SNOMED CT enabled data. The projects included in this review include:

- Data Migration Workbench (UK)
- Kaiser Permanente (USA);
- National Medication Decision Support System (Denmark);
- Semantic Search (Australia);
- Radiology Activity (UK).

We welcome additional input to this appendix and anticipate updates to this report as new information becomes available.

2.2 DATA MIGRATION WORKBENCH (UK)

The workbench is produced by the UK Terminology Centre (UKTC) as a DRAFT ‘proof of principle’ product. It demonstrates advanced functionalities, leveraging SNOMED CT as a sophisticated reference terminology together with the mappings across the current five NHS terminologies or classifications. It is not currently intended to develop this as a fully supported product.

[<https://isd.hscic.gov.uk/trud3/user/quest/group/0/pack/8>]

The UK Terminology Centre Data Migration Workbench (DMWB) is designed to support the NHS Primary Care Summary Care Record, Primary Care Systems of Choice and Data Migration programs. This tool demonstrates some of the properties and advanced uses of the data migration and mapping products published by the UKTC and the terminologies and classifications that they link.

The workbench uses SNOMED CT to perform novel and sophisticated analyses of patient data. It has immediate ‘off the shelf’ international utility despite the inclusion of the UK-only terminologies within the standard tool distribution.

2.2.1 OVERVIEW

The software contains SNOMED CT, Read Codes Version 2 and CTV3, maps between these and maps to ICD-10 International Edition (UK map not the same as the IHTSDO one) and OPCS Classification of Interventions and Procedures (OPCS-4). The Workbench modules support:

1. Searching and browsing the hosted code systems;
2. Viewing maps between the hosted code systems;
3. Authoring analytics subsets (i.e. terminology query predicates) and their testing, maintenance and ‘translation’ between code systems;
4. Electronic Patient Record (EPR) data quality analysis and data repair; and
5. EPR reporting and case mix analysis.

2.2.2 QUERIES TOOL

The Queries Tool offers advanced functionality for authoring, maintaining and testing query code sets (called ‘clusters’) or subset definitions within any of the supported terminologies or classifications. One major application is to produce query sets which will return comparable results from patient records encoded with different code systems. To assist with this, the tool translates subset definitions expressed using one terminology into subset definitions expressed using another, allowing refinement by manual editing (see Figure 2-1 and Figure 2-2).

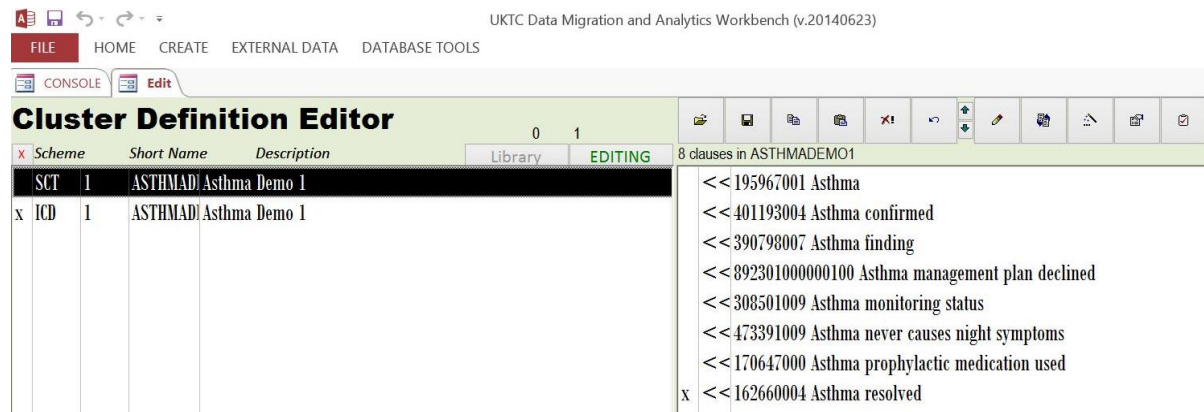


Figure 2-1: DMWB Queries Tool building a SNOMED CT asthma subset

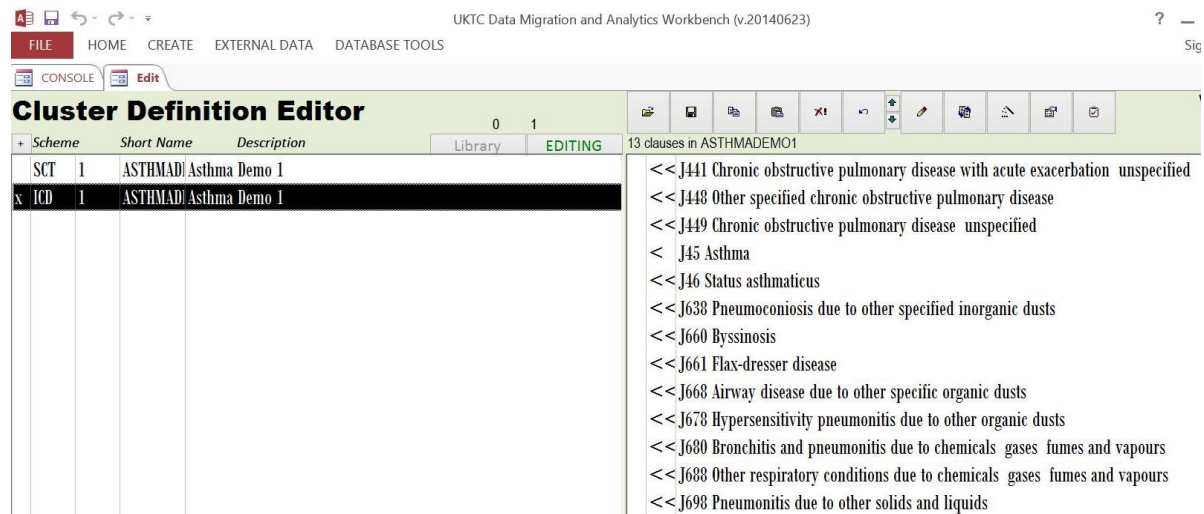


Figure 2-2: DMWB Queries Tool showing SNOMED CT asthma subset translated into ICD-10

2.2.3 ELECTRONIC PATIENT RECORD DATA

The EPR Data Tool provides an environment for loading and analyzing patient data, either to design query specifications or as part of data quality and case mix analytics. The analytics functions include detection and management of coding data quality issues such as:

- Records with inactive SNOMED CT codes;

- Records with codes from inappropriate SNOMED CT hierarchies e.g. a diagnosis recorded using a concept from the substance hierarchy (e.g. 419442005 |ethyl alcohol|) rather than a code from the disorder hierarchy (e.g. 25702006 |alcohol intoxication| or 7200002 |alcoholism|).

The tool also enables the rapid repair of such data by substituting inappropriate codes with more appropriate ones. This service is performed in an offline reporting environment.

2.2.4 ANALYTICS

The workbench data analytics tool runs cluster queries, combined with demographic data, to perform clinically valuable case finding, case mix and caseload analysis.

The 'Overview' Report module includes:

- Basic demographics (population age, sex, ethnicity);
- Analyses of episodes with a SNOMED CT code;
- Counts by SNOMED CT supercategory;
- List of the 100 most frequently used individual SNOMED CT codes; and
- List of the 15 most common SNOMED CT codes for each age cohort.

The Trends module analyzes the frequency with which individual SNOMED CT codes are used in the EPR instance data, looking for those whose recording frequency has changed over the course of the data collection period.

The Induce module performs a more sophisticated analysis of case mix and caseload trends within a clinical department. Instead of returning the most frequently used individual codes, the Induce module attempts to identify the most frequently used types of codes. For example, an emergency department may use roughly 500 different SNOMED CT codes for a laceration in a particular anatomical location. While none of the site-specific codes may appear in a list of most common codes, the descendants of 312608009 |laceration| may collectively account for a significant part of the department's workload.

The Graphs tool performs fundamentally the same query and search operations, but generates graphs based on the patients or episodes identified, showing e.g. the age:sex distribution of patients in a defined casemix cohort, or the changing incidence of one or more specified clinical phenomena (e.g. disease presentation, or procedure performed) by year, quarter, month or day of the week. These graphs can be copied into documents.

2.3 KAISER PERMANENTE (USA)

Kaiser Permanente HealthConnect® is a comprehensive electronic health record and one of the largest private electronic health systems in the world. KP HealthConnect with its integrated model securely connects more than 611 medical offices and 37 hospitals, linking patients to their health care teams, their personal health information and the latest medical knowledge.

[<http://share.kaiserpermanente.org/total-health/connectivity/>]

Kaiser Permanente (KP) has been involved in the development of SNOMED CT since its inception. Preceding this, KP collaborated with the College of American Pathologists in the 1990's on the immediate predecessor of SNOMED CT (SNOMED-RT). Some of the very earliest deployments of SNOMED CT have been within KP electronic patient record systems.

The terminology services deployed within the KP HealthConnect electronic health record illustrate the practical use of SNOMED CT as a key reference terminology within a multi-coding system environment. KP is also at the forefront of realizing new possibilities offered by SNOMED CT using its description logic capabilities.

2.3.1 CONVERGENT MEDICAL TERMINOLOGY

Convergent Medical Terminology (CMT) is KP's Enterprise Terminology System. While the KP HealthConnect EHR system is built by Epic (see Section 3.11), the CMT is proprietary to Kaiser Permanente. CMT hosts several components:

- Standard reference terminologies
- End user terminology (e.g. the terms presented to clinicians or patients)
- Administrative codes and classifications (e.g. ICD-9-CM, ICD-10-CM, CPT4, HCPCS)
- Analytics services (querying and decision support)
- Request submission for new terms

CMT uses SNOMED CT as a reference terminology, taking advantage of its poly-hierarchy and definitional attributes to support advanced analytics – for example:

- Identifying patient cohorts with certain conditions for Population Care.
- Identifying subsets for use as “input criteria” for KPHC decision support modules, such as Best Practice Alerts, Reminders, etc.
- Performing queries such as “find all conditions where |causative agent| is |Aspergillus (organism)|”
- Performing large aggregate queries, such as “find all patients coded with concepts in the cardiovascular disorders subset”

In September 2010 Kaiser Permanente, IHTSDO and the US Department of Health and Human Services jointly announced KPs donation of their CMT content and related tooling to the IHTSDO. The donation consists of terminology content (including several CMT subsets), tools to help create, manage and quality control terminology.

2.3.2 COLLABORATION WITH OXFORD UNIVERSITY

KP in collaboration with the Information Systems Group at Oxford University are investigating how to perform complex queries efficiently across extremely large numbers of patient records. The team at Oxford University has developed an open source triple store (i.e. 'subject-predicate-object') database called RDFox. RDFox is highly scalable and performant 'Not Only SQL' database readily distributed across parallel processing units. RDFox is an implementation of the W3C Resource Description Framework (RDF) standard, which supports OWL-RL description logic.

In this collaborative project, clinical data is being represented in OWL-RL as 'entity-role-act' triples. This uses a logical model (with Entities in Roles participating in Acts) that is similar to HL7 V3's Reference Information Model. OWL-RL and Datalog rule language is being used to reason over hundreds of millions of patient data triples. While SNOMED CT expressions cannot be fully represented in OWL-RL, RDFox performs the preliminary large-scale clinical data retrieval to return a far smaller record set. This smaller set is then processed using a richer featured but less performant description logic reasoner supporting SNOMED CT.

Prototype work has been completed using real patient data, including observations for diabetes (coded using SNOMED CT) and observations of Hemoglobin A1C levels. Datalog instructions and SPARQL queries were used to calculate Healthcare Effectiveness Data and Information Set quality measures for diabetes management – for example, both numerators and denominators for the Diabetes HgB A1C report.

2.4 NATIONAL MEDICATION DECISION SUPPORT SYSTEM (DENMARK)

"Physicians often lack the time to familiarize themselves with the details of particular allergies or other drug restrictions. Clinical Decision Support (CDS), based on a structured terminology, such as SNOMED CT, can help physicians get an overview by automatically alerting allergy, interactions and other important information. The centralized CDS platform based on SNOMED CT controls Allergy, Interactions, Risk Situation Drugs and Max Dose restrictions with the help of databases developed for these specific purposes."

[<http://ebooks.iospress.nl/publication/34363>]

The National Release Centre of Denmark (National eHealth Authority) produces a SNOMED CT drug extension for medications. The Danish SNOMED CT drug extension was primed by data extraction, cleansing and conversion of content from the Danish Medicine Agency Database (DKMDB), which is primarily meant for pricing and stock handling. The DKMDB was then complemented with SNOMED CT substances and their unique IDs. The Danish SNOMED CT drug extension includes information such as trade names, substances, dose forms, strengths and units of measure.

Building upon the Danish drug extension, the National eHealth Authority is working to introduce centralized decision support (CDS) services for both primary care and hospital prescribing systems.

The CDS server will respond to web service requests from the various electronic medication systems and return alerts and other prescribing information (see Figure 2-3)

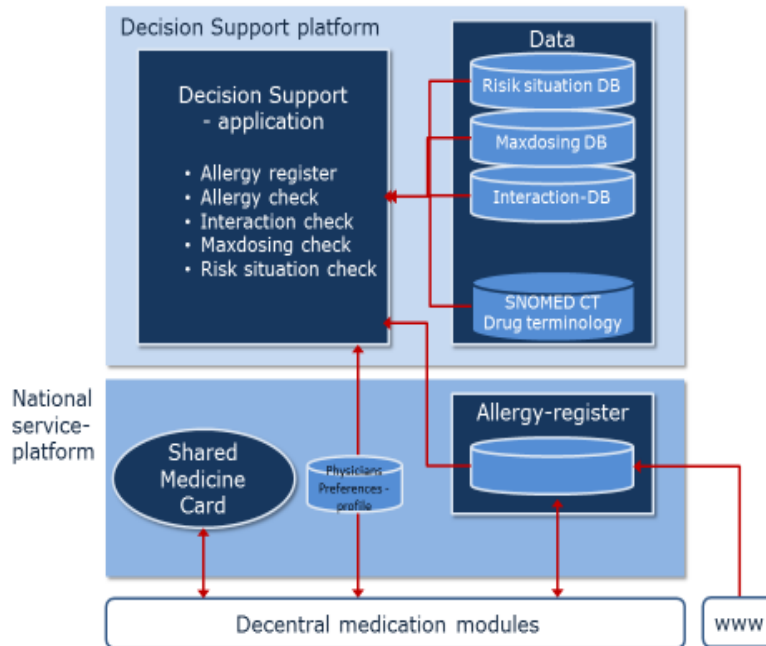


Figure 2-3: Denmark medication decision support services overview

2.4.1 ALLERGIES REGISTER

A group of allergy specialists, family practitioners and CDS experts are developing a standard set of information to be used in a patient drug allergy register. A SNOMED CT subset, from the Drug Allergy (disorder) sub hierarchy in the Findings hierarchy, is used to document allergies. .

2.4.2 ALLERGY ALERT SERVICE

Allergy alerts are enabled by the relationships in SNOMED CT between allergy disorders and substances (via the |causative agent| attribute), and relationships between drug products and substance concepts (via the |has active ingredient| attribute).

2.4.3 INTERACTIONS SERVICE

Based on an existing service, with data primarily drawn from peer-reviewed literature, the interaction database describes 2,500 interactions between different drugs based on their ingredients.

The database contains a short description of all interactions and a recommendation of how the physician can handle the interaction. The ingredients have been linked to SNOMED CT substances to directly inform the decision support service.

2.4.4 RISK SITUATION DATABASE

The risk situation database contains drugs evaluated by experts as being potentially dangerous in specific situations. Drug products, ingredients and dose forms are converted to SNOMED CT concepts, which thus contribute to the decision support service.

2.4.5 MAXIMUM DOSE DATABASE

An existing database contains maximum doses for all drugs and recommended doses for patients with impaired renal function. In the decision support service ingredients are once again expressed as SNOMED CT substances.

2.4.6 ALERT FILTERING

The decision support platform will incorporate an alert filtering service in which physicians can set up their personal preferences for the displaying of alerts. For example, the dose form hierarchy of SNOMED CT will be used to enable filtering of unwanted alerts for specific dose forms (such as cutaneous dose forms).

2.5 SEMANTIC SEARCH (AUSTRALIA)

In his thesis, Bevan Koopman presents models for semantic search: Information Retrieval models that elicit the meaning behind the words found in documents and queries rather than simply matching keywords. This is achieved by the integration of structured domain knowledge [from SNOMED CT] and data-driven information retrieval methods...

[http://eprints.qut.edu.au/71385/1/Bevan_Koopman_Thesis.pdf]

2.5.1 OVERVIEW

A major application for Natural Language Processing technologies is indexing collections of free text transcripts or documents such that topic specific searches may be run on them. The challenge is to return ranked matches which permit selection of texts with high sensitivity and high specificity (i.e. that relevant documents are rarely overlooked and that irrelevant documents are rarely returned).

Clinical searches may be performed over transcripts or documents that reside in an electronic library, within medical records, or the Internet. Examples of searches include:

- “Show me articles on this website concerned with inflammatory bowel disease”
- “Does this patient have transcripts in their record suggesting a heart rhythm disturbance?”

Bevan Koopman’s PhD thesis explores semantic and statistical approaches to search. The intention is to move beyond the limitations of plain keyword searching strategies for medical document retrieval. Characterizing these limitations as the ‘semantic gap’ Bevan identifies and addresses several issues including:

- Vocabulary mismatch: hypertension vs. high blood pressure
- Granularity mismatch: antipsychotic vs. haloperidol
- Conceptual implication: e.g. from hemodialysis infer kidney failure
- Inferences of similarity e.g. comorbidities (anxiety and depression)

His specific aim was to determine whether graph-based features and the propagation of information over a graph can provide an inference mechanism to bridge this semantic gap. As part of this work, he assessed the contribution of using SNOMED CT data within the graphs used to drive inferences.

2.5.2 MATERIALS AND METHODS

The specific application in the thesis to find patients who match certain inclusion criteria for recruitment into clinical trials based on the analysis of free text transcripts from clinical records.

Queries included

- Patients with depression on antidepressant medication
- Patients treated for lower extremity chronic wound
- Patients with AIDS who develop pancytopenia

Indexing methods were applied to the TREC MedTrack corpus - a standard collection of electronic texts containing de-identified reports from multiple hospitals in the United States. It includes nine types of transcripts: history and physical examinations, consultations, reports, progress notes, discharge summaries and emergency department operation reports, radiology, surgical pathology and cardiology reports. The collection as used contained around 100,000 reports within around 17,000 unique 'visits'.

Graphs have a number of characteristics that align with the requirements of semantic search as inference. The edges in a graph capture interdependence between concepts – which is identified as one of the semantic gap problems. Graphs are a common feature of both ontologies and retrieval models. The propagation of information over a graph — such as the popular PageRank algorithm used in Internet Search engines— provides a powerful means of identifying relevant information items (be they terms, concepts or documents). Ontologies such as SNOMED CT may also be represented as graphs.

The Graph Inference model developed by Bevan Koopman specifically addresses a number of semantic gap problems. Regarding vocabulary mismatch, the Graph Inference model utilizes a concept-based representation as this helps to overcome vocabulary mismatches (i.e. missed synonymy). The Graph Inference model specifically addresses granularity mismatch by traversing parent-child (i.e. 'is a') relationships.

The semantic gap problem of 'conceptual implication' is where the presence of certain terms in the document infer the query terms. For example, an organism may imply the presence of a certain disease. Such associations are captured in SNOMED CT and thus the Graph Inference model can specifically address conceptual implication by traversing those relationships.

Finally, the semantic gap problem of 'inference of similarity', where the strength of association between two entities is critical, is specifically addressed by the diffusion factor, which assigns a measure of similarity to each domain knowledge-based relationship. In the case of SNOMED CT the diffusion factor was derived from SNOMED CT relationships. It was noted that some relationships contributed to search sensitivity or conversely could lead to noise (loss of specificity) for the purpose of document retrieval. A weighting was applied (empirically) to each SNOMED CT relationship type and used as part of the relationship type component of the diffusion factor. For example, relationship type weightings included:

- |is a| = 1.0
- |active ingredient| = 1.0
- |definitional manifestation| = 0.8
- |associated finding| = 0.6
- |severity| = 0.2
- |laterality| = 0.2

Documents were parsed and analyzed using Lemur – a highly versatile and customizable open source information retrieval package developed at the University of Massachusetts. The construction of the graph was done using the open source LEMON graph library. The graph was serialized using LEMON and stored inside the Lemur index directory. For the MedTrack corpus, which was found to have a vocabulary size of 36,467 SNOMED CT concepts, the resulting graph was 4.4MB.

2.5.3 DISCUSSION

The findings of the thesis demonstrated that the graph based retrieval approaches using SNOMED CT derived data performed better than other approaches on ‘hard queries’. A number of additional insights were also revealed. First, hard queries require inference and easy queries do not. Hard queries tended to be verbose and often contained multiple dependent aspects to the query (for example, a procedure and a diagnosis concept). Re-ranking using the Graph Inference model was effective here. Easy queries tended to have a small number of relevant documents and an unambiguous query concept. For these queries, inference was not required and the Bag-of concepts model was most effective. Overall, when valuable domain knowledge was provided by SNOMED CT, the Graph Inference model was effective – either by returning new relevant documents or by effectively re-ranking those selected. This again highlights the dependence on the underlying domain knowledge.

Regarding residual lack of sensitivity of all the IR strategies, Koopman suggests that an ideal ontology for information retrieval would not only contain definitional but also assertional data – for example “captopril can be used as a treatment of hypertension”, “myocardial infarction [may] cause heart block” and “diabetes mellitus may lead to renal failure”.

2.6 RADIOLOGY ACTIVITY (UK)

“The National Interim Clinical Imaging Procedure Code set is a list of codes and descriptions for the coded and textual representation of Clinical Imaging Procedures in electronic systems in the NHS. It supports the consistent and unambiguous representation of imaging procedures in electronic information systems so that treatment options can be based on a common understanding of what procedures have been performed or are planned and activity can be directly comparable between all service providers.” [<http://systems.hscic.gov.uk/data/uktc/imaging/nicipfaqs>]

Migration to native SNOMED CT electronic patient records is in progress in the United Kingdom National Health Service (NHS). In order to promote interoperability, usability and activity reporting, the NHS introduced a national standard set of imaging codes in 2005 – the National Clinical Imaging Procedure code set (NCIP).

While SNOMED CT was the prime candidate for populating the NCIP, many Radiology Information Systems (RIS) and Picture Archiving and Communication (PAC) systems at the time could not accommodate SNOMED CT 18-digit concept identifiers or (up to) 255 character descriptions without disruptive and costly software changes. There was also no consistent way to represent laterality of procedures, and some legacy systems required the creation of separate orderable items for each laterality – for example ‘Plain X-ray left wrist’, ‘Plain X-ray right wrist’, and ‘Plan X-ray both wrists’. For these reasons, the NCIP code set was developed based on SNOMED CT, but with the addition of unique identifiers compatible with legacy system’s character limitations (6 alphabetic characters), up to 40

character human readable descriptions, and additional laterality metadata. For example, 60027007 | Radiography of wrist | is represented within NCIP as:

SCT ID	Laterality_ID	Laterality	Short_Code	Preferred
60027007	51440002	Right and left	XWRIB	XR Wrist Both
60027007	7771000	Left	XWRIL	XR Wrist Lt
60027007	24028007	Right	XWRIR	XR Wrist Rt

NCIP short codes are ‘meaningful’, in that the modality of the procedure is defined by the first character of the code, and the finding site and laterality are both explicitly represented in the code.

Each hospital submits mandatory data extracts using NCIP from both legacy and SNOMED CT capable RIS. In addition to details of the imaging procedures, information about the referral source, patient type, demographics and times of each imaging related event are also collected centrally. The data from all sites is then combined and multiple reports are extracted. Hospitals can view their activity data via the iView web based reporting tool (see Figure 2-4) and compare their activity with other centers.

Analytics on this central platform are wholly SNOMED CT based. SNOMED CT hierarchies support sophisticated reports – for example, the monthly waiting times for Magnetic Resonance Imaging excluding Cardiac MRI and MRI guided procedures is specified as:

- **Includes** hierarchy << 113091000 | Magnetic resonance imaging |
- **Excludes** hierarchy << 258177008 | Magnetic resonance imaging guidance |
- **Excludes** hierarchy << 241620005 | Magnetic resonance imaging of heart |

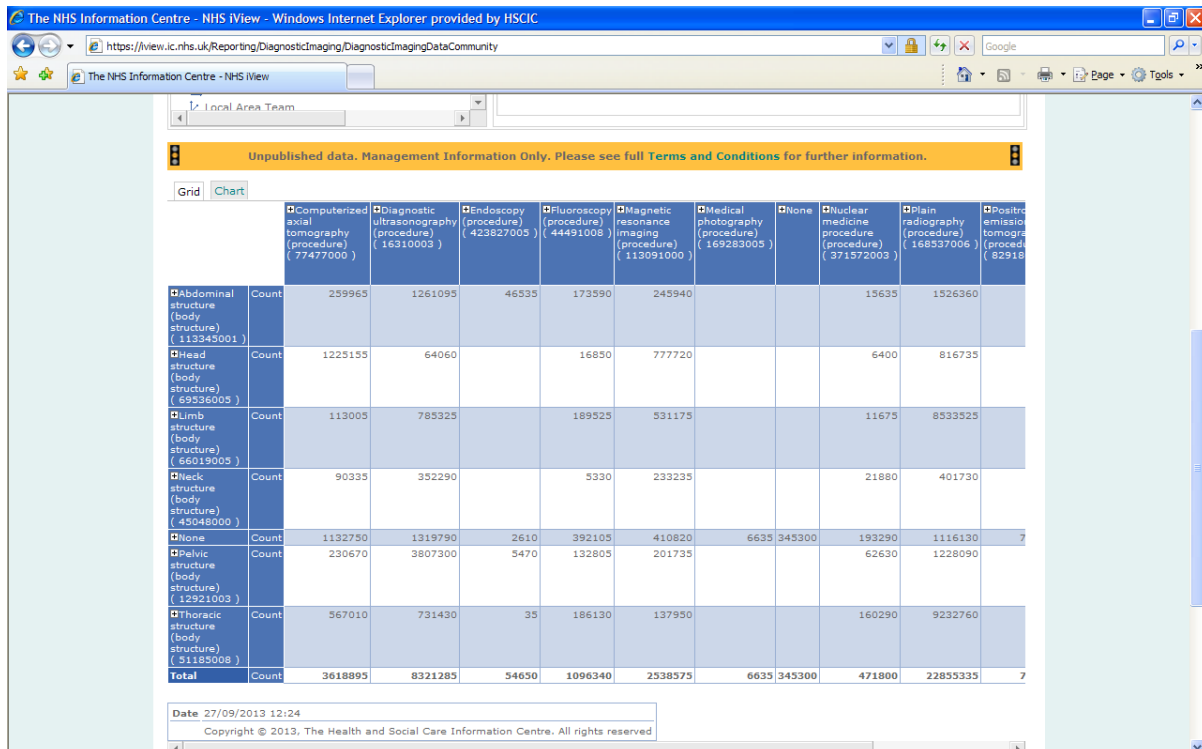


Figure 2-4: Detail of SNOMED CT based report on the NHS iView platform

3 APPENDIX B – VENDOR CASE STUDIES

3.1 OVERVIEW

This appendix includes brief reviews of a variety of commercial tools which support analytics over SNOMED CT enabled data. These reviews focus predominantly on those tooling features that are relevant to supporting analytics services. The vendors who have contributed to this review include:

- 3M Health Information Systems;
- Allscripts Healthcare Solutions;
- Apelon;
- B2i Healthcare;
- Cambio Healthcare Systems;
- Caradigm;
- Cerner Corporation;
- Clinithink;
- EMIS;
- Epic;
- First Databank; and
- Intelligent Medical Objects.

We welcome additional input to this appendix and anticipate updates to this report as new information becomes available.

3.2 3M HEALTH INFORMATION SYSTEMS

“3M Health Information Systems provides intelligent tools to help compile and use health information for better clinical and financial performance. Best known for market-leading coding system and ICD-10 expertise, 3M Health Information Systems also delivers innovative software and consulting services for clinical documentation improvement, computer-assisted coding, case mix and quality outcomes reporting, mobile physician solutions, and a robust healthcare data dictionary and terminology services to support the Electronic Healthcare Record.”

[\[http://solutions.3m.com/wps/portal/3M/en_US/Health-Information-Systems/HIS/\]](http://solutions.3m.com/wps/portal/3M/en_US/Health-Information-Systems/HIS/)

The 3M Healthcare Data Dictionary (HDD) is a controlled medical vocabulary server. The HDD has been continuously expanded and maintained for over 15 years, both as a standalone product and embedded within several of 3M’s core products and services. The 3M HDD enables mapping and management of medical terminologies, integration of content and standardization of healthcare data. The 3M Healthcare Data Dictionary incorporates a selection of standard healthcare terminologies, including (but not limited to) SNOMED CT, LOINC, RxNorm, ICD-9-CM and ICD-10-CM.

Concepts in the HDD are grouped and organized using both hierarchical and non-hierarchical relationships. One of the hierarchical relationships in the HDD is SNOMED CT’s ‘is a’ relationship which allows users to programmatically use and analyze SNOMED CT concepts captured at various levels of

granularity. The analytics capabilities of the HDD are also extended through the use of other relationship types.

3.2.1 DATA WAREHOUSING

The content within the HDD makes a key contribution to analytics in several settings. For example one large academic research institution uses the HDD to integrate over 100,000 medication concepts from disparate systems for comprehensive data assimilation. Many of the medication concepts are mapped to RxNorm codes and linked through a 'has ingredient' relationship to SNOMED CT codes.

The 3M HDD has a knowledge base and poly-hierarchical structure that defines the relationships between each clinical drug. Figure 3-1 shows the relationships that exist for Ramipril including the links to SNOMED CT content, which can be used to query the data warehouse. The knowledge base allows the hospital's researchers to customize their searches by various levels of granularity and organize their clinical content into meaningful relationships.

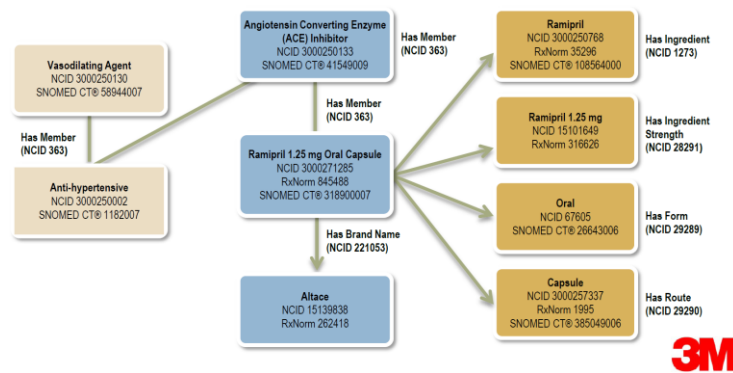


Figure 3-1: 3M HDD - Application of a knowledge base and hierarchies

The HDD supports researchers in performing data mining by:

- Extracting and mapping clinical metadata using a streamlined, systematic approach;
- Translating diverse clinical terminologies using a coded medical vocabulary;

3.3 ALLSCRIPTS

“Allscripts Healthcare Solutions, Inc. (Allscripts) is a provider of clinical, financial, connectivity and information solutions and related professional services to hospitals, physicians and post-acute organizations. The Company provides a variety of integrated clinical software applications for hospitals, physician practices and post-acute organizations. For hospitals and health systems these applications include its Sunrise Enterprise suite of clinical solutions, consisting of a range of acute care Electronic Health Record (EHR), integrated with financial/administrative solutions, including performance management and revenue cycle/access management. The Company’s acute care solutions include Emergency Department Information System (EDIS), care management and discharge management.

[<http://www.reuters.com/finance/stocks/companyProfile?symbol=MDRX.O>]

Allscripts released their first version of Vocabulary Management utilizing SNOMED CT in 2005. Since then Allscripts systems have been able to utilize SNOMED CT for clinical decision support and reporting. Allscripts uses a common terminology platform for all three electronic health record systems: Sunrise Clinical Manager™, Sunrise Acute Care™ and Sunrise Ambulatory Care™.

When a query is performed over a health record that requires clinical terminology, the terminology service always returns a SNOMED CT code. If the primary code stored in the health record is not SNOMED CT (e.g. ICD-9 or ICD-10), then the terminology service performs the mapping to SNOMED CT, saves the SNOMED CT code in the health record next to the original code (to make future queries more efficient), and returns the SNOMED CT code.

Allscripts Sunrise applications are able to link SNOMED CT to all orders, order form pick lists, observations pick lists and results.

3.3.1 POINT OF CARE DECISION SUPPORT

Sunrise Clinical Manager integrates reference content from medical publishers into clinician workflow. The Sunrise InfoButton™ feature provides clinicians access to relevant medical reference content wherever patient care decisions are made, without requiring them to log into or visit another site. Sunrise InfoButton uses encoded patient problem lists and medication data elements to query third-party medical content selected by the clinician. Several healthcare reference content providers index their content using SNOMED CT. This enables the delivery of on-topic information without manual searching. Resources with SNOMED CT indexed content include Wolters Kluwer Clin-eguide which provides context-specific evidence on particular medical issues and diseases and Lexi-Comp, providing drug and drug-interaction information, diagnosis and disease management, formulary services, patient-education resources and clinical support tools.

3.3.2 REPORTING

The Allscripts Clinical Quality Management (CQM) is an automated chart abstraction and analytics system. CQM is able to create population sets with SNOMED CT encoded patient records and use these patient sets for reporting. CQM is a flexible, powerful reporting and analytics system presenting information in a variety of formats ranging from simple list style reports, to Online Analytical Processing (OLAP) Data Cubes with Pivot reports.

Allscripts Clinical Performance Management (CPM) is a business intelligence solution for monitoring clinical performance, improving patient outcomes and reducing costs. With prebuilt or customized reporting and dashboards, healthcare leaders have powerful access to performance information enabling quality improvement across the health enterprise.

Applications include:

- Alert usage analysis: User-customizable reports drill down into clinical decision support usage data revealing the reasons and circumstances for bypassed or overridden alerts. Seeing the impact of decision support enables a sharper focus on patient outcomes and supports the refinement of rule logic.
- Order-set usage analysis: Organizations can evaluate their order set usage patterns of computerized provider order entry. By observing order set configuration, deployment and use, organizations can revise them to enhance their effectiveness and usability.
- Clinician utilization analysis: Clinicians can examine the vast array of health issues and clinical observations for discharged patients to support patient treatment decisions and protocol implementation. In addition, patient cohorts can be tracked over time to determine if the proper treatments are being delivered.

3.4 APELON

“Apelon is an international informatics company focusing on data standardization and interoperability. Leading healthcare organizations use Apelon’s products and services to better manage terminology assets. Apelon solutions help healthcare application vendors, biomedical researchers, providers, biotech companies and government agencies improve the quality, comparability, and accessibility of clinical information.”

[<http://www.apelon.com/>]

SNOMED CT plays a central role in many Apelon products and projects. Apelon tools feature navigation and visualization tools to support SNOMED CT in a variety of ways. Apelon also undertakes bespoke content development and consultancy work in healthcare and biomedicine using SNOMED CT.

The Apelon Terminology Development Environment (TDE) software was used by the College of American Pathologists to build and maintain the SNOMED CT International Edition prior to the formation of IHTSDO. Apelon software continues to be used by major healthcare organizations and some National Release Centers to maintain SNOMED CT extensions, maps and subsets.

3.4.1 APELON DISTRIBUTED TERMINOLOGY SYSTEM

The Apelon Distributed Terminology System (DTS) offers a variety of human and computer interfaces to navigate, visualize and query SNOMED CT. DTS allows users to create custom extensions to SNOMED CT and perform incremental description logic classification to ensure that the extensions are consistent with the base version of SNOMED CT. DTS permits navigation, and side-by-side comparison of concepts across multiple SNOMED CT versions. Features of Apelon DTS supporting analytics and data retrieval include:

- **Subsetting:** DTS allows users to create customized SNOMED CT subsets using advanced logic techniques. The user is able to create extensional and intentional value sets of concepts for queries based on both hierarchical and non-hierarchical relationships.
- **Data normalization:** DTS supports the matching of text input to standardized terms and concepts via word order analysis, word stemming, spelling correction and term completion
- **Code translation:** DTS supports the mapping of clinical data to standard coding systems such as SNOMED CT, ICD-9, 10 and CPT

3.4.2 PROJECTS USING SNOMED CT

Apelon frequently choose to rely upon SNOMED CT in their consulting work as the overarching reference terminology. Recent projects using SNOMED CT include:

- Work with a major performance measure developer to create a large number of SNOMED CT value sets representing the inclusion and exclusion criteria for quality measures. SNOMED CT supplies the expressivity for the detailed distinctions amongst disorders and patient characteristics that is essential for this work.
- Use of SNOMED CT as the “backbone” terminology in a number of mapping projects for Health Information Exchanges. A small value set of SNOMED CT concepts serves as the “source of truth” for determining appropriate maps, and then codes from other terminologies are assigned based on whether they are a good fit with the SNOMED CT concept. This strategy provides a way to capture the precise intent behind the often-fuzzy language found in clinical documents.
- Use of SNOMED CT to index patient education materials for a major content provider. Documents are retrieved via an ‘Infobutton’ request in the EMR based on codes found in the patient record.

3.5 B2I HEALTHCARE

B2i Healthcare provides tools and services to maximize SNOMED CT’s utility.

B2i Healthcare Pte Ltd (B2i) is a boutique software engineering firm specialized in SNOMED CT and healthcare information standards and exchange. B2i provide products to simplify SNOMED CT adoption and offer software development services to support healthcare IT needs.

[\[http://www.b2international.com/\]](http://www.b2international.com/)

Snow Owl is a clinical terminology platform developed by B2i Healthcare. The Snow Owl technology family is deployed in over 2,500 locations in 83+ countries worldwide. The Snow Owl® terminology server has been licensed by the IHTSDO to form the basis of the IHTSDO Terminology Server.

3.5.1 SNOW OWL TERMINOLOGY SERVER

The Snow Owl® terminology server scales from a small kernel embedded in single-user products to n-tier clusters supporting hundreds of concurrent users. Clients can easily access and query SNOMED CT, LOINC, ATC, ICD-10, and dozens of additional terminologies via REST or Java APIs. Collaborative distributed

authoring is also supported, including creating and maintaining local code systems, mapping between terminologies, and creating terminology subsets.

Terminology server features include:

- Extensive support for expression constraints and semantic queries, including Extended SNOMED CT Compositional Grammar and Groovy scripts.
- Distributed revision control system supports large teams of authors and reviewers working on independent branches.
- Full support for SNOMED CT logical definitions (OWL 2 EL) with extended support for extensions using advanced description logic features (OWL 2 DL) including datatype properties, universal restriction, disjunction, etc.
- Standard distribution formats (e.g. SNOMED RF2, ICD-10 ClaML, LOINC csv)
- Traditional, white-label (embedded within client product), and source code licenses available.

The Singapore Drug Dictionary (SDD) is the biggest SNOMED CT extension - larger than SNOMED CT International release itself. To support medication safety initiatives like medication management and adverse drug event surveillance, the drug ontology makes use of Snow Owl's extended description logic support.

3.5.2 SNOW OWL IDE

The Snow Owl IDE (Integrated Development Environment) simplifies developer tasks related to terminology tooling. The architecture allows customized extensions to integrate tooling needs within a single platform.

The IDE embeds a terminology server and simplifies common terminology maintenance, ETL, and other tasks. Customized authoring environments support developing a library of queries (SNOMED CT expression constraints) using the Simple or Extended SNOMED CT Compositional Grammars and Groovy scripting. Files can be exported in a variety of formats like OWL 2, SNOMED CT RF1 and RF2, ClaML, spreadsheets and text files. Custom formats can also be created that support direct import and export to proprietary EHR and terminology applications.

Typical vendor deployments: EHR vendors use Snow Owl to create and maintain their local terminologies and mappings to reference terminologies like SNOMED CT. Snow Owl IDE allows exporting this in a format consumable by the proprietary EHR system format. The Snow Owl IDE has been built into proprietary tooling combining information modelling with ontology development.

3.5.3 SNOW OWL COLLABORATIVE AUTHORING PLATFORM

Snow Owl's collaborative terminology authoring platform maintains terminology artefacts developed by a team and supported by business workflows. The platform consists of the terminology server with remote clients collaborating with independent authoring workflows. The platform integrates with external task management systems like Bugzilla and JIRA.

Features:

- Full support for creating SNOMED CT extensions, including RF1 and RF2, all subset and mapping RF2 reference set types, modules, and full change history to 2002.
- Support for dozens of terminologies and any local code systems.
- Creation of value sets including mixing and matching codes from different code systems.

- Import existing value sets from the USA National Library of Medicine’s Value Set Authority Center.
- Creation of mapping sets between any two terminologies or mapping local code systems to reference terminologies like SNOMED CT and LOINC.
- Configurable workflow support for authoring use cases like single, dual, and dual independent authoring.
- Support for terminology-specific workflows and editing restrictions.

The Singapore Ministry of Health Holdings uses Snow Owl to maintain their national SNOMED CT extension and local code systems as well as multi-terminology value sets and mappings used in their National Healthcare Data Dictionary.

3.5.4 SNOW OWL MEANINGFUL QUERY

The international adoption of SNOMED CT and related healthcare ontologies has provided the logical definitions that enable a new breed of queries. Unfortunately, it’s challenging to run ad hoc queries that make use of the full semantics of the underlying EHRs. Operational stores have the data, but in a variety of structures that can’t act on the semantic relationships between healthcare terms. Data warehouses can query only aggregated data that has been placed into predefined buckets which don’t provide the scale of complexity inherent in the original data. And multiple information models represent the same semantic meaning in different ways.

Snow Owl Meaningful Query (MQ) allows semantic EHR queries on operational data stores without requiring predefined structures like data warehouses or the presence of a single unified healthcare information model. The system is optimized specifically for *ad hoc* queries of hundreds of millions of electronic health records. We combine ontological reasoning over the EHRs with more traditional query methods to incorporate demographic and ancillary data.

This query interface is being rolled out to all Singapore public hospitals and the national procurement office to allow search and retrieval of pharmaceuticals contained within the Singapore Drug Dictionary ontology. All lexical and semantic properties can be searched, including datatype properties and mappings to local code systems, external terminologies like ATC, and internal procurement codes.

3.6 CAMBIO

“Cambio Healthcare Systems is a market leading Electronic Patient Record (EPR) company headquartered in Stockholm, Sweden with offices in the UK, Sweden, Denmark and Sri-Lanka. Cambio COSMIC® is a patient-centered integrated EPR system for comprehensive and clinical healthcare solutions with a focus on patient safety. Cambio COSMIC® offers solutions within all healthcare sectors and is used by over 100,000 clinicians and healthcare professionals.”

[\[http://www.cambiohealthcare.co.uk/\]](http://www.cambiohealthcare.co.uk/)

The Cambio COSMIC® Electronic Patient Record system has been under continuous development since 1993. Cambio has applied innovations within healthcare informatics in areas such as information models, clinical terminology and formal languages for expressing clinical decision support rules. The COSMIC® EPR combines openEHR archetypes, SNOMED CT terminology and Guideline Definition Language rules in implementations which benefit patients, clinical staff and healthcare enterprise management. Using these technologies, their system is able to incorporate advanced analytics capabilities.

3.6.1 DECISION SUPPORT

Cambio uses the Guideline Definition Language (GDL) to combine archetypes, terminologies and clinical decision support rules. GDL provides:

- Bindings between archetype elements and variables in the rules;
- Rule expressions that are easily converted to industry rule engine languages;
- Bindings between local concepts used in the rules and concepts from SNOMED CT.

GDL rules can be used to trigger a variety of system actions, including pre-filling a form, proposing a test or prescription, or sending a notification to the system user. The criteria for triggering actions from GDL rules may be based on demographics data, the context of care (e.g. clinic or inpatient), current medications and diagnoses, or observation values (e.g. lab results).

Decision support rules created in COSMIC® are authored using an editor. Figure 3-2 shows the high level view of a rule for calculating a complex clinical risk-score (CHA₂DS₂-VASc Score for stroke risks stratification in atrial fibrillation).

affects a particular condition, and the effects of different medications, researchers need trial subjects to meet specific criteria.

3.6.2 OFF-LINE REPORTING AND DATA WAREHOUSING

COSMIC Intelligence is a data warehouse and reporting application. Analyses and reports that do not require real time information are produced within this separate analysis system. COSMIC Intelligence is a data store optimized for queries, retrieval and output of data. Data is periodically retrieved from the 'live' clinical system, transformed and loaded into the data store.

3.7 CARADIGM

“Caradigm is a population health company dedicated to helping organizations improve care, reduce costs and manage risk. Caradigm analytics solutions provide insight into patients, populations and performance, enabling healthcare organizations to understand their clinical and financial risk and identify the actions needed to address it. Caradigm population health solutions enable teams to deliver the appropriate care to patients through effective coordination and patient engagement, helping to improve outcomes and financial results.” [<http://www.caradigm.com/en-us/about/>]

Caradigm is a joint venture between Microsoft and GE Healthcare, which is dedicated to population health management. Caradigm’s cornerstone product is the Intelligence Platform. This platform can connect over 295 types of source systems, including Allscripts, Athenahealth, Cerner, Epic, GE, McKesson and Meditech. Data from disparate systems within one or more healthcare organizations is collected, normalized and standardized to enable applications to leverage this data in a unified and consistent way.

Caradigm’s solutions provide explorative, comparative, predictive and guided elements aimed at analyzing the disparate data and driving the insight that is gained into action. Caradigm’s three main solution areas are:

- Healthcare analytics (including clinical, operational and financial analytics);
- Coordination management; and
- Wellness promotion and patient engagement.

Some of Caradigm’s customers use SNOMED CT natively in their clinical systems, while others use natural language and other code systems. In order to aggregate data from disparate sources, it must first be standardized by mapping into a common code system. Code systems used to standardize the disparate data include SNOMED CT, ICD-9 and ICD-10. By standardizing the data, users are able to leverage the analytics tools - for example, to understand trends within different diagnoses, to look at a comprehensive list of everything that has happened to a patient in a longitudinal patient record, and to support care management by displaying the different diagnoses or problems of a patient in a consistent manner.

Caradigm currently implements an approach to SNOMED CT based analytics using clinical value sets. These value sets are developed manually by a team of clinical analysts for topics such as diabetes and heart disease. Clinical users are then able to create queries in a user friendly interface, which allows

them to (for example) define cohorts built on criteria such as age, gender, specific diseases, conditions, medication or other treatments. These queries are then converted behind the scenes into SQL statements which are executed against a SQL Server database and return records containing data in the selected clinical value sets.

Caradigm also has Natural Language Processing (NLP) tools, which are able to extract and encode data, such as problems and medications, from natural language notes within documents such as discharge summaries and radiology reports.

As part of their strategic roadmap, Caradigm are exploring ways to enhance the capabilities of their tooling platforms by leveraging the architecture of the terminology sets that they are using for analytics. In particular, they are planning to start utilizing the hierarchical and non-hierarchical relationships of SNOMED CT to enable more powerful query capabilities and to extend NLP processing options.

3.8 CERNER

“Cerner Corporation is a supplier of healthcare information technology solutions, services, devices and hardware. Cerner solutions optimize processes for healthcare organizations. These solutions are licensed by approximately 9,300 facilities globally, including more than 2,650 hospitals; 3,750 physician practices 40,000 physicians; 500 ambulatory facilities, such as laboratories, ambulatory centers, cardiac facilities, radiology clinics and surgery centers; 800 home health facilities; 40 employer sites and 1,600 retail pharmacies. The Company operates in two segments: domestic and global. The domestic segment includes revenue contributions and expenditures associated with business activity in the United States. The global segment includes revenue contributions and expenditures linked to business activity in Argentina, Aruba, Australia, Austria, Canada, Cayman Islands, Chile, China (Hong Kong), Egypt, England, France, Germany, Guam, India, Ireland, Italy, Japan, Malaysia, Morocco, Puerto Rico, Qatar, Saudi Arabia, Singapore, Spain, Sweden, Switzerland and the United Arab Emirates.”

[<http://www.reuters.com/finance/stocks/companyProfile?symbol=CERN.O>]

3.8.1 OVERVIEW

Cerner Corporation’s Millennium healthcare system manages terminologies, classifications and other code systems within a terminology service - the Cerner Millennium Terminology (CMT) package. CMT accommodates and integrates SNOMED CT International Release data and National extension content - such as concepts, relationships, descriptions, subsets, maps etc.

At the Cerner Millennium user interface, content can be captured at point of care as SNOMED CT codes. Modules used with SNOMED CT include: Problems and Diagnoses, Allergies, Procedures, Pharmacy Orders, Radiology Orders and Cellular Pathology Reports.

Either Cerner or third party clinical encoding software can process SNOMED CT Diagnoses and Procedures captured in Millennium and suggest ICD-10 and other classification codes to Clinical Coders for activity reporting and billing.

SNOMED CT is also used extensively behind the scenes to support more sophisticated analytic facilities within their Natural Language Processing (NLP) tools and reporting tools. These Cerner products and services exploit unique features and content of SNOMED CT to extend the power of these applications.

3.8.2 DATA WAREHOUSING

The Cerner product suite includes two data warehousing applications. These applications share much of their terminology-related technology, including supporting subsumption queries with the CMT Concept Explode/Transitive Closure facility, which utilizes the SNOMED ‘is a’ relationships.

The PowerInsight® Data Warehouse (PIDW) is an enterprise level data warehouse which updates on a nightly basis from the live electronic patient record. PIDW services standard operational reporting, mandatory reports (e.g. for National or State governments and regulatory bodies), and *ad hoc* queries e.g. individual lists of patients treated as requested by clinicians for audit.

The Health Facts® Reporting supports the pooling of anonymized data from different healthcare organizations. Health Facts data warehouse represents information of electronic records from millions of inpatient, emergency department, and outpatient visits from participating U.S. health care organizations. (Data are encrypted and secured to ensure patient confidentiality in compliance with HIPAA privacy regulations.) The reporting facilities enable the analysis of patient care and process trends within a facility and provide comparisons to other Health Facts contributors.

The Cerner data warehouse query tools include simple graphical user interfaces which directly create powerful reports using the SNOMED CT hierarchy content. The screenshot below in Figure 3-4 shows a report of attendances with diagnoses which are a descendent of the SNOMED CT concept 417746004 |traumatic injury|.

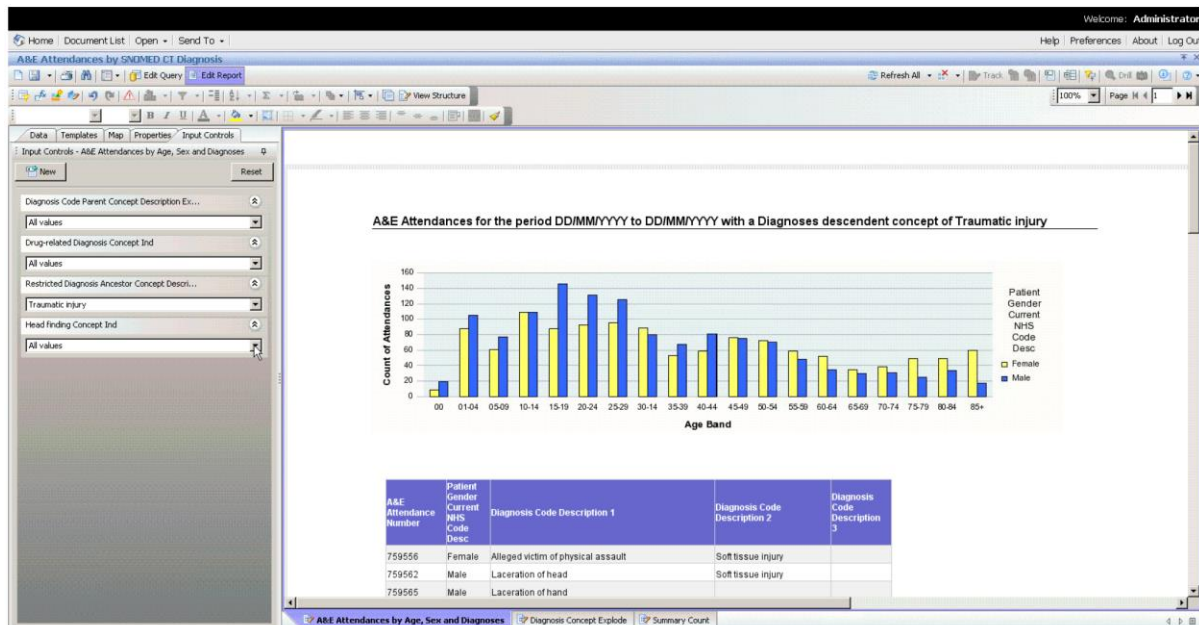


Figure 3-4: Report produced by Cerner’s data warehouse query tool

3.8.3 OTHER APPLICATIONS

Cerner's Natural Language Processing (NLP) technology interprets the content of clinical notes through a complex understanding of grammar, syntax, synonymy and phraseology. SNOMED CT's semantic content and concept model enriches the analysis of the text in several ways including

- Concept recognition using synonyms – for example: | heart attack | is a synonym of | myocardial infarction |;
- The hierarchical relationships between concepts – for example: | pneumonia | is a | respiratory disease |;
- The identification of context, such as negation, certainty, subject and timing;

Computer Assisted Coding allows the extraction of appropriate SNOMED CT codes for automating coding and billing processes.

Chart Search/Semantic Search is a tool that enables clinicians at the point of care to search in real time through a patient's multiple charts, pathology reports and other documents, for topics such as 'heart disease' and 'diabetes'. The interface, as shown below in Figure 3-5 has the look and feel of a World Wide Web search engine.

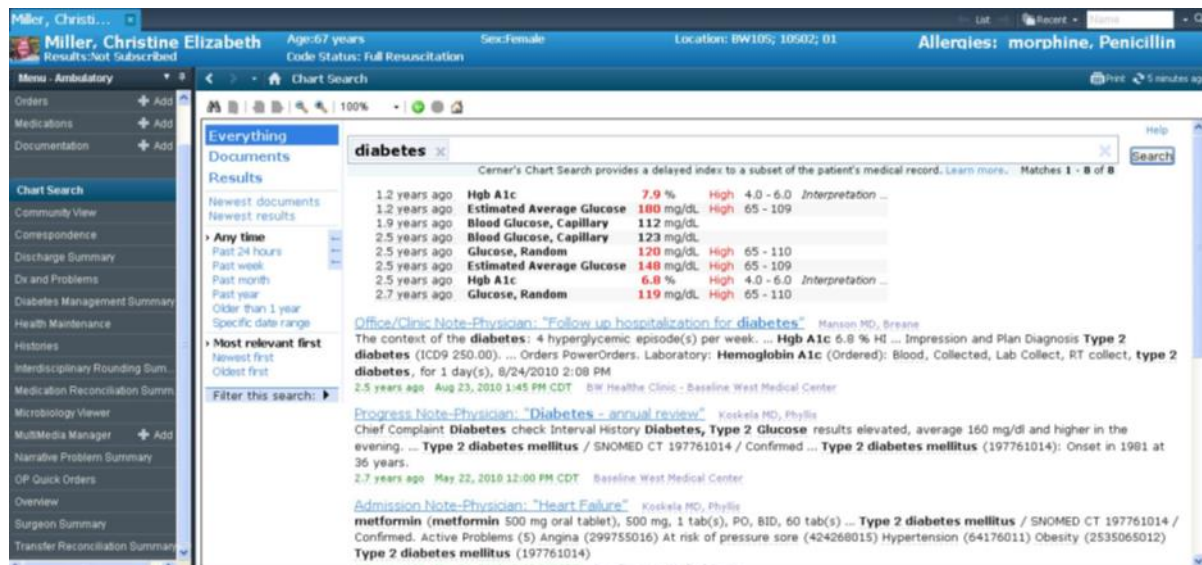


Figure 3-5: User interface of Cerner's Chart Search/Semantic tool

The searches can be filtered by date, document type etc. However the power of this approach is extended beyond conventional search engine indexing by using SNOMED CT. Cerner's tools index SNOMED CT findings (including diseases and symptoms) and procedures to make searches and queries over these domains faster. Cerner also hand curates exceptions and associations between related concepts. A Clinical Significance Score is assigned to each concept to allow documents to be sorted based on the probable relevance of the concepts in the document given their context of use.

As shown in Figure 3-6 for example, when given the search term 'heart disease', the 'is a' hierarchies of SNOMED CT enable recognition and return of documents which reference 'sinus bradycardia' and 'dilated cardiomyopathy'.

heart disease x

Searched items through Oct 8, 2010 8:22 AM CDT Matches 1 - 10 of 24

[Search](#) [Reset](#)

[Office/Clinic Note-Physician: "Heart Disease Management"](#)
REASON FOR VISIT: Management of **atherosclerotic heart disease**. ... Multiple attempts were made to open the **occluded right coronary artery**. ...
PROBLEMS/DIAGNOSES: 1. **Atherosclerotic heart disease**, severe.
9.3 years ago Jul 11, 2001 2:45 PM CDT Pleasant Valley Medical Group

[History and Physical Office/Clinic: "Coronary Artery Disease, New Patient Evaluation"](#)
Medical: 1. Type 2 diabetes mellitus. 2. Hyperlipidemia. 3. **Coronary artery disease** (as noted above). ... His father died at age 80 of **heart disease** and a Staph infection. ... Labs: EKG shows **sinus bradycardia** at a rate of 58, normal axis, normal intervals, and no evidence of remote **myocardial infarction**.
9.0 years ago Oct 26, 2001 11:39 AM CDT Pleasant Valley Medical Group

[Office/Clinic Note-Physician: "Routine Follow-up, Multiple Problems"](#)
She also had a little bit of **tachycardia**. I had ordered a Doppler study, and the patient says it was done. ... It is also possible that she could be having **angina, since ischemic heart disease** is increased in patients with lupus and also increased in people with rheumatoid arthritis.
7.5 years ago Apr 30, 2003 1:40 PM CDT BW Endocrine Clinic - Baseline West Medical Center

[History and Physical Office/Clinic](#)
The remaining review of systems is negative. Social History: As noted, her sister died of a **myocardial infarction** or "**heart attack**". ... Her mother now has **coronary artery disease**, and she is going to be her caregiver.
9.8 years ago Dec 19, 2000 2:00 PM CST East Clinic

[Office/Clinic Note-Physician: "aortic disease management"](#)
Nonischemic **dilated cardiomyopathy** followed by Dr. Y. 3. ... Status post aortic root repair in 1996. 6. Status post pacemaker for **third-degree heart block** in 02/98. ... Family History: Negative for **myocardial infarction**, cerebrovascular accident, cancers.
8.0 years ago Oct 15, 2002 11:50 AM CDT Baseline West Medical Pavilion

[History and Physical Office/Clinic: "Cardiology Consult"](#)
His 2 children are both male, ages 25 and 21. Family History: His father had an **MI at age 56**. ... Mr. X has a family history of premature coronary artery disease (His father died of **myocardial infarction** at the age of 56.
14.8 years ago Dec 14, 1995 9:50 AM CST BWCH Cardiology - Baseline West Medical Center

Figure 3-6: Search results from Cerner's Chart Search tool

3.9 CLINITHINK

“Clinithink’s tools enable analytics and querying over SNOMED CT encoded patient data ... The CLiX CNLP platform transforms clinical narrative into rich structured data for healthcare providers and solution vendors. CLiX ENRICH, powered by CLiX CNLP to support analytics, converts unstructured clinical data into actionable data required to help solve today’s toughest healthcare business problems.

[\[www.clinithink.com\]](http://www.clinithink.com)

Building on the capabilities of its CLiX Clinical Natural Language Processing (CNLP) platform, Clinithink has created a solution to enable the analysis of healthcare data sourced directly from clinical narrative called CLiX ENRICH. This technology can be integrated into existing healthcare solutions to process any relevant narrative and encode the key clinical elements – medications, diagnoses, procedures, symptoms and findings – using SNOMED CT. These features can then be queried using powerful, user-definable SNOMED CT queries to present structured data in a form that can be easily consumed by existing BI platforms.

For example, when the narrative text below is typed into diagnosis and observation fields (respectively), Clinithink is able to provide a list of possible coding options for manual confirmation (as shown below).

HX: She presented with occipital headache for last six hours. She has been intermittently nauseated but has not vomited. Mild fever. She has no past history of headaches but has a past history of asthma. Mother has a severe allergy to penicillin but she has no known penicillin allergy. Currently taking paracetamol prn.

OE: BP is 102/60, pulse 70, respiratory rate 20, temperature 37.2 C. Nasopharynx normal, both tympanic membranes normal.

History

- Occipital headache
 - six hours
- Nauseated
- not** Vomiting
- Fever
 - Mild

Past History

- no** History of headache
- History of - asthma

Family History

- Allergy to penicillin
 - Severe
 - Mother

Allergies, Risks, Warnings

- no known** Penicillin allergy

Medications

- Paracetamol
 - As required

Observations and Findings

- Blood pressure
 - 102/60 units: none
- Pulse finding
 - 70 units: none
- Respiratory rate
 - 20 units: none
- Body temperature
 - 37.2 units: none
- Nasopharynx normal
- Tympanic membrane normal
 - Right and left

3.10 EMIS

“EMIS clinical systems are used by over 5,000 healthcare organizations across the United Kingdom, from Primary Care and out-of-hours services, to community care and sexual health services. By using the same system, everyone can access the same information about their patients - no matter where they are treated - making the prospect of integrated care a reality. With over 25 years experience of working with the NHS, EMIS are entrusted with over 40 million patient records.”

[<http://www.emis-online.com/company-profile>]

Egton Medical Information Systems (EMIS) began in the 1980s in a rural practice in Egton in North Yorkshire, United Kingdom. The founders, Dr Peter Sowerby and Dr David Stables, wrote the software and adopted the NHS Read Code system during the 1980s. A series of systems have since been deployed in over half of England’s primary care practices. The latest product (EMIS Web) moved to a data center based architecture, thin client front end and built in secure web based patient access facilities.

EMIS software looks after the patient records of nearly 40 million people in the UK (30 million using EMIS Web). More than 2 in 3 of those patients can book appointments and order repeat medications online, and more than 1 in 3 can view their own medical record.

EMIS Web features significant advances in terminology use with EMIS adopting a phased approach to SNOMED CT. EMIS Web displays a familiar coding structure based on the construction of a Read Version 2 navigational hierarchy within SNOMED CT. The principle design objective has been to enable SNOMED CT within the clinical system to meet specific requirements, including:

- Supporting advanced decision support capabilities;
- Supporting interoperability within healthcare through the sharing of coded data;
- Supporting standards required in NHS General Practice Systems of Choice (e.g. the NHS mandates SNOMED CT coding within the National Summary Care Record service);
- Broadening the scope of terminology use to support the recording of encounters in disciplines such as dentistry and community healthcare;
- Supporting the mandatory requirement for the Electronic Prescription Service to natively use the UK SNOMED CT drug extension (i.e. NHS dictionary of medicines and devices, dm+d).

By using coded structured records and providing access to the specialist domain terminology available in SNOMED CT, EMIS has been able to extend the user base of EMIS Web by more than 20,000 new NHS users over the last year. These include practice nurses, community matrons, child health and mental health nurses, palliative care clinicians, diabetes specialists, physiotherapists and psychologists.

3.11 EPIC

“Epic makes software for mid-size and large medical groups, hospitals and integrated healthcare organizations – working with customers that include community hospitals, academic facilities, children's organizations, safety net providers and multi-hospital systems. Epic's integrated software spans clinical, access and revenue functions and extends into the home...”

*Epic's integrated analytics and reporting – collectively named **Cogito** – delivers current clinical intelligence and business intelligence based on role and workflow... Epic provides a combination of flexible tools, content, data sources, distribution, training, and process to support decisions throughout the health system with the best information available.”*

[<http://www.epic.com/about-index.php>]

3.11.1 OVERVIEW

Epic's electronic patient record systems use SNOMED CT as a reference terminology through the following mechanisms:

- Mappings between a subset of Epic's standard data elements and SNOMED CT concepts;
- Mappings between diagnoses imported from other code systems (e.g. those used in Intelligent Medical Object's and Health Language's products) and SNOMED CT concepts;
- Mappings from additional data elements to SNOMED CT concepts that can be created by clients using an External Concept Mapping activity.

These mechanisms provide linking behind the scenes, so when clinicians add a diagnosis to the problem list by selecting a familiar term, for example, they automatically select the corresponding SNOMED CT concept. This SNOMED CT encoding creates powerful possibilities within Epic's decision support and reporting facilities.

3.11.2 DECISION SUPPORT

The Epic system calls its decision support alerts 'Best Practice Advisories'. These customized alerts are programmed to fire according to predetermined triggers, such as specific chief complaints, vital signs, diagnoses or medications, either individually or in combination using inclusionary or exclusionary logic. Best Practice Advisories can thus be used to notify clinicians to tend to important tasks, such as reviewing a patient's allergies, writing orders, and completing charting. They can also present order sets and links to third party information sources refined using the clinical context of the patient record being reviewed.

Epic customers can use SNOMED CT's hierarchical structure to group related records, making the setup for clinical decision support much simpler than would be possible if users had to select records or clinical concepts individually. For example, an administrator creating Best Practice Advisories for diabetic patients could use 73211009 |diabetes mellitus| within the SNOMED CT hierarchy as one of the criteria instead of identifying every subtype of diabetes individually.

3.11.3 REPORTING

Within Epic's integrated analytics and reporting suite (i.e. Cogito) customers have achieved benefits by using SNOMED CT's clinical finding hierarchy to aggregate local diagnosis concepts.

The capability has for example been used by oncologists working with cancer-related ICD codes which are unsuited to grouping diagnoses by stage. Using the mapped SNOMED CT codes they are able to facilitate the reporting of staging data by utilizing the SNOMED CT hierarchy.

3.12 FIRST DATABANK

"The Multilex drug knowledge base is widely used throughout the UK and is integrated into clinical systems across the whole healthcare community. The Multilex drug terminology holds clinical and commercial information on more than 75,000 pharmaceutical products and packs and provides active clinical decision support and referential medicines information for all healthcare professionals."
[<http://www.fdbhealth.co.uk/solutions/multilex/>]

3.12.1 OVERVIEW

First DataBank (FDB) were in the first wave of suppliers to recognize the potential of SNOMED CT and begin to integrate support for SNOMED CT into their existing clinical decision support solutions. Their primary use of SNOMED CT in the patient's electronic health record (EHR) is to detect safety issues arising from certain combinations of medications, diagnoses and drug adverse reaction histories. In 2006 FDB introduced support for products and packs encoded using the NHS SNOMED CT UK Drug Extension. In the following year FDB launched new modules within the Multilex drug knowledge base supporting Drug-Condition Checking and Drug Sensitivity (Allergy) checking for the SNOMED CT EHR.

System vendors implementing Multilex decision support within SNOMED CT-enabled medical record applications include CSC (Lorenzo system), EPIC and JAC in secondary care, and CSE Servelec (*RiO system*) in community/mental health. Currently only pre-coordinated expressions are supported by the live Multilex SNOMED CT based decision support solutions.

3.12.2 DRUG-CONDITION CONTRAINDICATION CHECKS

The contraindications module alerts the clinician when a medication proposed to treat a disorder is incompatible with another of the patient's disorders or clinical states. For example a beta blocker like propranolol might be prescribed to treat someone with high blood pressure. However if that patient also has asthma, their asthma might significantly worsen or a dangerous acute attack might be produced by the drug.

Thousands of such drug-condition contraindications exist and nearly all medications have at least one. Without point of care decision support, the clinician must rely on memory or search reference sources for each drug prescribed. Also there is a risk that a contraindicating condition may be in the record but unknown to the prescribing clinician.

In a SNOMED CT enabled EHR, both the drugs (e.g. 318353009 | propranolol hydrochloride 40mg tablet |) and the conditions (e.g. 370219009 | moderate asthma |) are encoded.

Internally FDB maintain their own local ontology representing only those conditions relevant to prescribing decision support (e.g. asthma, gastric ulcer, heart disease, pregnancy). The items in this ontology are linked to SNOMED CT codes as required to support this (contraindication checking) use case. These SNOMED CT links range from the obvious, such as linking 195967001 |asthma| to FDB's 'asthma', to the more subtle, such as linking 447413000 |drainage of amniotic fluid using ultrasound guidance| to FDB's 'pregnancy'.

FDB reviews the relevant SNOMED CT domains (i.e. |clinical findings|, |procedures| and |situation with explicit context|) for concepts applicable to drug-condition checking. The FDB linking tool uses the SNOMED CT |is a| hierarchy and a SNOMED CT derived transitive closure table to locate and suggest links from the FDB ontology to SNOMED CT concepts. Other SNOMED CT relationships also help find related concepts via the browser but discovery is mainly by clinical knowledge combined with description based searches assisted by the rich synonym content of SNOMED CT.

3.12.3 DRUG SENSITIVITY (ALLERGY) CHECKS

The sensitivities module alerts the clinician when a proposed drug for a patient is *either* stated in that patient's record to have caused a previous adverse reaction *or* when an adverse reaction has occurred to a similar drug and thus likely to elicit a similar adverse response. For example, a patient allergic to penicillin is likely to react to most other drugs containing a β -lactam ring in their molecular structures.

In a similar way to how FDB links SNOMED CT conditions to its own internal ontology, SNOMED CT concepts which suggest allergy or previous adverse reactions to a medication are also linked to an internal FDB ontology for representing medication ingredients. This ontology is designed specifically to support allergic and adverse reaction cross-reactivity.

3.13 INTELLIGENT MEDICAL OBJECTS

"Intelligent Medical Objects (IMO) develops, manages, and licenses medical terminology and healthcare IT software applications that allow clinicians to capture their clinical intent at the point-of-care. IMO's comprehensive medical terminology of physician-friendly terms is mapped to the preferred billing and reference codes enabling clinicians to use the terms they are familiar with while ensuring improved coding accuracy."

[<https://www.e-imo.com/>]

IMO produces a medical terminology service for healthcare solutions, allowing over 2,500 hospitals and 350,000 clinicians to focus on patient care. IMO bridges the information gap between clinicians, coders, and patients in the US and internationally. IMO enable and support the accurate capture and preservation of clinical intent for clinical documentation, decision support, reimbursement, reporting, data analysis, research, and health education.'

IMO's clinical interface terminology is designed pragmatically to capture clinical intent at point of care. However it is also intended to enable and simplify the adoption of standard ontologies by vendor partners.

By choice, the editorial process requires all IMO interface terms to have one or many qualified maps to SNOMED CT. Clients can then use SNOMED CT to drive reporting, analytics, clinical decision support, and research.

The following examples demonstrate how IMO uses SNOMED CT for analytical purposes:

1. Helping patients find health professionals who have expertise or interest in specific areas of medicine. These areas include disorders, procedures, devices, medications, patient demographics, and medical specialties. These areas of expertise or interest include those that are self-reported by clinicians and those documented in clinical encounters. The search algorithms use hierarchies in SNOMED CT to retrieve and rank search results.
2. Helping clinicians use patient diagnoses and procedures documented at varying levels of granularity to find appropriate patient education materials using SNOMED CT is-a hierarchies.
3. Grouping together related clinical concepts in patient records for creating focused patient reports and driving clinical workflows.
4. Forming subsumption queries for cohort selection within patient data repositories and document libraries.
5. IMO uses natural language processing (NLP) to extract information coded in SNOMED CT from clinical narratives.