SNOMED Takes the Next Step

Save to myBoK

by Karen M. Kudla, MBA, MT (ASCP), and Margo Blakemore, BS, RHIT

Healthcare terminologies SNOMED and the Read Codes have been around for a long time. Now, a new terminology combining the two promises to take the industry one step closer to realizing the vision of an electronic patient record. Here's a sneak peek at SNOMED® CT.

As technology evolves, healthcare takes slow but sure steps toward the development of a computer-based patient record. An important part of this process is the development of a reference terminology that makes healthcare knowledge more accessible wherever and whenever it is needed. With the upcoming release of SNOMED CT, the industry will take another step toward a computerized record.

The evolution of a concept-based clinical terminology like SNOMED CT provides new opportunities for HIM professionals and promises to benefit clinicians, patients, administrators, and payers as well. This article will examine the "new era" of the SNOMED terminology, look at its structure, and provide perspective on its evolution.

A Different Kind of Merger

Through the combined efforts of standards development organizations, the government, and private industry (including organizations that deliver healthcare and those that develop the systems and technology for it), the prospect of a computer-based patient record is beginning to be realized. The next step will be taken with the combination of two established terminologies.

SNOMED, the Systemized Nomenclature of Medicine, is a copyrighted work of the College of American Pathologists. Clinical Terms Version 3 (formerly known as the Read Codes) is a crown copyright work of the National Health Service in the United Kingdom. These two systems, introduced in the 1970s and 1980s, respectively, were the products of similar visions—to facilitate the exchange, retrieval, and analysis of key data in the medical record.

The current SNOMED iteration, SNOMED® RT, includes an elementary mapping to ICD-9-CM, which offers the flexibility to see how the terms are cross-referenced to ICD-9-CM. However, as SNOMED RT and Clinical Terms Version 3 become SNOMED Clinical Terms (CT), refinement of mapping criteria and guidelines of the "initial" mapping will begin to evolve.

SNOMED CT, currently in development and scheduled for release in 2002, is the merger of SNOMED RT and Clinical Terms Version 3. For background about these two terminologies, see "Four Decades Tell Story of SNOMED, Read Codes."

The development of the SNOMED CT system is relevant to HIM professionals because of the many opportunities and challenges it presents to those who will be managing a relationship between a clinical terminology and the classification systems used today for reporting activities. While this may seems like an elementary issue, the challenge is to create a relationship between a very granular terminology and a broad-based classification system such as ICD-9-CM.

The evolution of a concept-based clinical terminology like SNOMED CT provides growing opportunities for HIM professionals in data management, not only to use recognized classification systems such as ICD-9-CM, but to use the clinical terminology as its own entity in research analysis, statistical reporting, disease state management, and outcome measurements.

The New Era

SNOMED RT

Development of the Systematized Nomenclature of Medicine Reference Terminology (SNOMED RT) began in earnest in 1997 with the combined efforts of the College of American Pathologists and a team of physicians and nurses from the Kaiser Permanente Convergent Medical Terminology project. SNOMED RT, released in 2000, is a concept-based reference terminology. Although commonly viewed as coded terminology for electronic health records, SNOMED RT is being implemented in a variety of healthcare systems, including radiology, clinical laboratory, and surgical pathology information systems. Uses include retrieval of information for cancer registries, structuring data repositories, infectious disease reporting, and literature search and retrieval.

Clinical Terms Version 3

In the early 1990s, the British government, through its National Health Service, obtained the rights to the Read Codes and in 1992 embarked on a three-year project to expand the work from a primary care focus to a comprehensive terminology representing the entire scope of clinical practice.

The development of Clinical Terms Version 3 was a product of the Clinical Terms Project. Under the leadership of the NHS Information Authority, the project involved 55 clinical specialty working groups in the United Kingdom, representing more than 2,000 individuals from the Royal Medical Colleges, the nursing professions, and allied health professions. These working groups proposed terms and developed and reviewed the content of what was to become the core terminology for electronic health records throughout Britain.

Separated at Birth? Notable Similarities

Both terminologies, although developed independently, followed similar guiding principles that were beginning to take shape in the medical informatics community in the 1990s. These principles were formally described in two papers, both published in 1998, and represent the foundation upon which not only SNOMED RT and Clinical Terms Version 3 were built, but SNOMED CT as well.[1, 2] Between the two papers, nearly 20 principles in total are outlined. Among these, three core features are worth special mention:

- **Both terminologies are concept-based.** A concept is a word or phrase with specific meaning. Concepts can be nouns or verbs that describe a disease or diagnosis, a part of the anatomy, a procedure, a body function, or an etiology.

The word "cast," for example, can carry a number of different meanings. Among other things, it can refer to a rigid material molded to the body, a precipitated product from the kidney tubules found in urine, or, in the US, the name of a proprietary drug product. Each of these concepts represent unique expressions of thought and are assigned their own unique concept codes. An example showing concepts from the SNOMED RT concepts table is shown in "Sample SNOMED RT Concepts."

- **In both terminologies, concepts are represented for everyday usage by terms.** A term is the commonly used name for a concept. For example, the concept of "cast" is fully specified in the concepts table as Cast (device), Cast (morphologic abnormality), and Cast (substance) to clearly communicate and differentiate their distinct meanings. However, in everyday usage they are all simply known as the term "cast."

In SNOMED RT (as well as Clinical Terms Version 3) the meaning of each of these terms is differentiated by its corresponding concept identifier. A single concept can often be represented by a variety of terms or synonyms. For example, synonyms of the disease pruritus hiemalis include "chapped skin," "dry skin," or "winter skin."

"Sample SNOMED RT Descriptions," above, provides an excerpt from the SNOMED RT descriptions table showing how terms and synonyms are represented and associated with their corresponding concept.

- Finally, both terminologies employ explicit computer-readable semantic relationships. In early versions of SNOMED and the Read Codes, the meaning of a concept was embedded in the code or identifier. (See "The Search for Meaning.") Identifiers carry no meaning in SNOMED RT or Clinical Terms Version 3. Instead, the computer-readable meaning of a concept is represented in a relationships table (SNOMED RT) or hierarchy and templates tables (Clinical Terms Version 3). "Sample SNOMED RT Relationships," shows how concepts are semantically defined in SNOMED RT.

A Team Effort

In April 1999, the National Health Service and the College of American Pathologists announced a historic collaboration to unite SNOMED and Clinical Terms Version 3 and create "a unified international terminology that supports the integrated electronic medical record." Shortly following that announcement, representatives from the NHS and the CAP began work designing and building SNOMED CT.

When it is released in 2002, SNOMED CT will incorporate all content from both SNOMED RT and Clinical Terms Version 3 in a merged data structure that is expected to total an estimated 250,000 to 300,000 concepts, 400,000 terms, and more than 1,000,000 semantic relationships. (Documents on the file structure and other details of SNOMED CT development can be viewed online at http://www.snomed.org/snomedct.)

Benefits of SNOMED CT

A concept-based terminology like SNOMED RT and Clinical Terms Version 3 and the future SNOMED CT will standardize medical terms, allowing the transmission of patient-related data across diverse information systems. It allows healthcare practitioners, allied health professionals, and institutions to collect and analyze data more effectively, compare the quality of healthcare being administered, develop effective treatment guidelines, and conduct important outcomes research.

These benefits will be realized with the adoption of SNOMED CT in electronic health records. In the United Kingdom, the National Health Service produced a mandate that such systems begin to implement SNOMED CT in 2003. Without a similar mandate in the United States, adoption may be more gradual, but it promises to increase with the growth of electronic records throughout the healthcare industry.

With this growth, the industry may see a number of related benefits. A concept-based terminology helps to provide structure essential to computerizing the patient medical record, benefiting both healthcare providers and patients. It allows providers easily accessible and more complete information pertaining to the entire care process (medical history, illnesses, treatments, laboratory results, etc.) and can contribute to improved patient outcomes. And it allows for the identification of patients based on certain coded information in their records and thereby facilitates follow-up.

SNOMED CT will help software developers throughout the world eliminate redundancy in coding and simplify mapping techniques. These developers will benefit from being able to rely on one primary source for reference terminology.

This means:

- less effort and cost to develop or customize vocabularies for different medical specialties or geographic areas
- less worry and time spent in keeping the vocabulary complete, up to date, and reflective of contemporary clinical practice
- greater compatibility across software applications. And because of the structure of SNOMED CT, building systems to support consistent reporting, outcomes analysis and clinical decision support will be made simpler

Third-party payers will also benefit from the use of a controlled terminology. By standardizing data that is currently heterogeneous, payers will be able to review and process claims in a more timely manner. Standardization will also be a means to simplify administrative tasks and save costs in the process.

Finally, if integrated into various quality programs, such as the National Committee on Quality Assurance's HEDIS measures, it can streamline the collection and analysis of the clinical information that forms the basis for these quality assessments.

Furthermore, a concept-based terminology will be critical for research. SNOMED encompasses key elements necessary for research such as granularity, comprehensive coverage, clinical detail, and the ability to retrieve and analyze clinical data at a more abstract level. These and other features position SNOMED RT and Clinical Terms Version 3 today—and SNOMED CT tomorrow—as invaluable tools.

The Search for Meaning

In early versions of SNOMED and the Read Codes, the meaning of a concept was embedded in the code or identifier.

SNOMED 3.5 example: Each letter or number in a SNOMED code represents a specific placement in a hierarchy. Codes beginning with DC-1 through DC-3 are red blood cell disorders.

Example:

DC-10106   Acquired hemolytic anemia
DC-10107   Congenital hemolytic anemia
DC-10109   Chronic idiopathic hemolytic anemia

Problem: How to code anemias that occur with other disorders

Unsatisfactory Solution: Code it under one disorder only

D6-90324 Hemolytic anemia with emphysema and cutis laxa *(coded as a metabolic disease)*
D8-74300
          Hemolytic disease of fetus or newborn *(coded as a neonatal disease)*

Read Code example: Each digit in a Read Code represents a specific level in the hierarchy (three-digit code = third level, four-digit code = fourth level, etc., to five digits or levels). The coding system breaks down when a sixth level is needed.

Example:

4th level code: 7130 Mastectomy
5th level code: 71304 Subcutaneous mastectomy

Problem: Where to put "Subcutaneous mastectomy, gynaecomastia"—a more specific form (i.e., a child) of subcutaneous mastectomy?

Unsatisfactory solution: Code it as a fifth level code 71307, which incorrectly represents this term as a sibling and not a child of subcutaneous mastectomy

Sample SNOMED RT Descriptions

DescriptionType 1 represents the preferred term for the concept while DescriptionType 2 represents synonyms. For each concept there must be one (and only one) preferred term and zero to many synonyms. Note that while each term has its own unique description identifier, all terms associated with a particular concept are assigned the same concept ID. SNOMED CT will also include a descriptions table similar to this:

DescriptionID	ConceptID	Term	Description Type
14003019	7867008	Pruritus hiemalis	1
14004013	7867008	Winter itch	2
14005014	7867008	Xerotic eczema	2
14006010	7867008	Asteatotic eczema	2
14007018	7867008	Eczema craquelé	2
14008011	7867008	Dry skin dermatitis	2
14009015	7867008	Chapped skin	2
14010013	7867008	Eczema craquele	2
173702013	109167006	CAST	1

36444016	21718009	Cast	1
177056014	118643004	Cast	1
101041018	60805002	Hemolytic anemia with emphysema AND cutis laxa	1
178708014	33130007	Hemolytic disease of fetus OR newborn	1
189556016	33130007	Alloimmune hemolytic disease of newborn	2
189547012	33130007	Erythroblastosis fetalis	2
189548019	33130007	Erythroblastosis neonatorum	2
189549010	33130007	Neonatal anemia	2

Sample SNOMED RT Concepts

SNOMED CT will represent concepts in a table similar to this:

Concept ID	Fully Specified Name
7867008	Pruritus hiemalis (disorder)
109167006	CAST (substance)
21718009	Cast (morphologic abnormality)
118643004	Cast device
60805002	Hemolytic anemia with emphysema AND cutis laxa (disorder)
33130007	Hemolytic disease of fetus OR newborn

Sample SNOMED RT Relationships

The relationships table carries explicit semantic definitions for a concept. Note that the relationships table enables the representation of multiple parents (see Pruritus hiemalis and Hemolytic disease of the newborn). SNOMED CT relationships will be presented in a similar format.

ConceptId1	Concept Id1 Name	Relationship	ConceptId2	ConceptId2 Name
118643004	Cast device	Is a	7704007	Stabilizing appliance (physical device)
109167006	CAST (substance)	Is a	96391005	United States of America specific proprietary drug (substance)
21718009	Cast (morphologic abnormality)	Is a	36850000	Retention (morphologic abnormality)
7867008	Pruritus hiemalis (disorder)	Is a	105963001	Noninfectious, erythematous, papular AND/OR squamous disease
7867008	Pruritus hiemalis (disorder)	Is a	43116000	Eczema (disorder)
7867008	Pruritus hiemalis (disorder)	Is a	78934009	Pruritus of skin (finding)
7867008	Pruritus hiemalis (disorder)	Associated topography	39937001	Skin (organ) (body structure)
7867008	Pruritus hiemalis (disorder)	Associated morphology	23583003	Inflammation (morphologic abnormality)
7867008	Pruritus hiemalis (disorder)	Associated morphology	4538007	Papulovesicular rash (morphologic abnormality)

33130007	Hemolytic disease of fetus OR newborn	Is a	111472004	Transitory neonatal endocrine AND/OR metabolic disorder
33130007	Hemolytic disease of fetus OR newborn	Is a	37097005	Erythroblastosis (disorder)
33130007	Hemolytic disease of fetus OR newborn	Is a	61261009	Hemolytic anemia (disorder)
33130007	Hemolytic disease of fetus OR newborn	Associated topography	83418008	Fetus (body structure)
33130007	Hemolytic disease of fetus OR newborn	Associated topography	41898006	Erythrocyte (cell)

Mapping From SNOMED RT to ICD-9-CM

ICD code 366.12, Incipient cataract, maps to the following more granular SNOMED RT concepts:

52421005	Incipient cataract
12195004	Coronary cataract
40714009	Punctate cataract

Four Decades Tell Story of SNOMED, Read Codes

Think back to 1976. IBM mainframes and VAX minicomputers were state of the art. Personal computers were still ideas in the minds of innovators like Steve Jobs and Bill Gates. Hospitals (not yet known as "healthcare organizations"), if they were computerized at all, used mainframes primarily for financial transactions. Departmental computing was in its infancy and was found only in technologically advanced clinical laboratories where laboratory information systems were used for reporting lab results via paper printouts.

SNOMED was developed in this era as a "logical comprehensive approach to medical information management." It was built upon an even earlier coded terminology, the Systematized Nomenclature of Pathology (SNOP), which was introduced in 1965 and was widely used by pathologists for coding surgical pathology cases. SNOP used an alphanumeric coding scheme in which the first letter of the code represented the primary characteristic of the element being coded. Codes beginning with "T" identified the topography (anatomic source) of the surgical pathology case, while codes beginning with "M" identified the morphologic characteristics of the case. "T" and "M" codes were further partitioned by a numbering system that represented increasingly detailed topography or morphology terminology. For example,

M-7200	Hypertrophy
M-7201	Hypertrophy, focal
M-7202	Hypertrophy, compensatory
M-7203	Hypertrophy, diffuse

SNOMED adopted SNOP's alphanumeric coding scheme and expanded it to facilitate the encoding of potentially all terms needed for healthcare data handling, with a primary focus on diagnoses and procedures. SNOMED, and its successor SNOMED II introduced in 1980, contained six primary axes or sections: topography, morphology, etiology, functions, diseases, and procedures. In addition, SNOMED II also included a list of occupations.

The use of SNOMED II increased considerably in the 1980s, predominantly in surgical pathology laboratories where computerization was becoming more widespread. Systems that stored surgical pathology reports in computer files began incorporating encoders that would encode key aspects of the reports into SNOMED codes. These codes would be stored in a database and enable retrieval of cases based on topography, morphology, diagnosis, procedure—whatever was encoded.

Into the 1990s, plans were under way to publish a new edition of SNOMED that would more than double the 44,000 terms in SNOMED II. To accomplish this, the SNOMED II coding scheme needed to be expanded to accommodate nearly 60,000 new terms as well as future expansion. SNOMED International, also known as SNOMED 3, was introduced in 1993 with 12 axes, 100,000 terms, and a new hexadecimal coding scheme.

The alphanumeric code remained a core component of SNOMED. It carried intrinsic meaning, identifying not only the term but the axis to which it belonged, and, to the extent that the numbering system allowed, the term's placement in a hierarchy of related terms.

The Read Codes: an Early Evolution for Primary Care

In the 1980s, James Read, a family practice physician in England, developed a coded terminology for the recording and retrieval of primary care patient data in a computer. Like SNOMED, the Read Codes were also based on a strict coding scheme where each term was carefully assigned an identifier that represented its position in a hierarchy.

Whereas SNOMED, with roots in pathology, grew in popularity in pathology departments across the United States, the Read Codes, with roots in primary care, grew in popularity in ambulatory care clinics across the United Kingdom.

The Identifier Problem

Both SNOMED and the Read Codes were developed at a time when computers were beginning to make inroads in healthcare but were not well established. The coding scheme employed in the two terminologies enabled humans to easily understand and use the code's hierarchical information in simple computer applications. However, as the number of terms in the Read Codes and SNOMED continued to grow, along with the ability of computers to handle more complex tasks, the limitations of a code to carry meaning and the potential for misinterpretation became more and more evident. (See "The Search for Meaning.")

The NHS eliminated the use of "meaningful" codes when it introduced Clinical Terms Version 3. The College of American Pathologists did the same with the introduction of SNOMED RT. This switch from "meaningful" to "meaningless" identifiers, in conjunction with the introduction of an improved structure to represent meaning, marked the beginning of a new era in the development of clinical terminology for the future.

The People Behind the Codes

The creation of SNOMED CT involves the combined effort of teams of clinical informaticists, database experts, physicians, nurses, pharmacists, and other clinical healthcare professionals from both the US and UK. Although its initial release will be in the English language, work is under way to provide a Spanish translation.

The SNOMED International Editorial Board, which includes representatives from the College of American Pathologists, National Health Service, and numerous US and UK-based clinical specialty organizations and US government agencies (such as the CDC, Department of Veterans Affairs, and National Center for Health Statistics), provides scientific direction for the development of SNOMED CT. The SNOMED International Authority, composed of representatives from the CAP and NHS, provides leadership in business and strategic development issues.

Before it is released, SNOMED CT will undergo alpha and beta testing in the US and UK with software developers who will evaluate the technical aspects of incorporating and using SNOMED CT in clinical information systems and clinical domain experts who will review the clinical content.

Notes

1. Cimino, J.J. "Desiderata for Controlled Medical Vocabularies in the Twenty-first Century." *Methods of Information in Medicine* 37 (1998): 394-403.

2. Chute, C.G., S.P. Cohn, and J.R. Campbell. "A Framework for Comprehensive Health Terminology Systems in the United States." *Journal of the American Medical Informatics Association* 4, no. 6 (1998): 503-9.

Karen Kudla (kkudla@cap.org) is director of SNOMED operations at the College of American Pathologists in Northfield, IL. Margo Blakemore (mblakem@cap.org) is a research analyst at SNOMED International.

Article citation:
Kudla, Karen M. and Margo Blakemore. "SNOMED Takes the Next Step." *Journal of AHIMA* 72, no.7 (2001): 62-68.

Driving the Power of Knowledge