AN EVALUATION METHODOLOGY FOR CLINICAL VOCABULARIES AND EVALUATION OF THE READ CODES

Lorraine Toews
Health Sciences Library
University of Saskatchewan
e-mail: toewsl@duke.usask.ca

Abstract
The purpose of this study was to develop a framework for evaluating clinical vocabularies and to use this framework to evaluate the suitability of the Read Codes for use in a clinical charting and decision-support system. This paper describes an evaluative framework based on a formal standard for thesauri and on criteria outlined in the literature. Version 3.0 of the Read Codes was used to code selected concepts from a clinical care map used for managing myocardial infarction. The evaluative framework was then used to assess the sample.

Introduction
Health care providers and administrators are under increasing pressure to evaluate the quality, effectiveness, and cost of clinical care. In order to accurately assess clinical performance, clinicians, government bodies, accrediting agencies, and regulatory bodies require consistent, adequate, timely and retrievable patient data. Although much of the data needed for quality assurance research, outcomes research, and utilization research is contained in patient records, accessing this information is often a problem. One of the major barriers to effective retrieval is not technological, but linguistic.

Most health care providers still document patient care in natural language narrative text. Although clinicians use a constrained vocabulary in patient records, there is considerable variation in how they express similar concepts. Because inconsistencies in the description of clinical data tend to arise when natural language is used, retrieval of information in patient records is often difficult. Although computers can rapidly manipulate, analyze, and retrieve large amounts of data, they currently have a limited capacity to compensate for the
problems inherent in retrieving information in natural language text. A range of clinical vocabularies have been developed in order to address the problem of language in the computerized patient record.

**Purpose of Study**
The purpose of this study was to develop a framework for evaluating clinical vocabularies and to use this framework to evaluate the suitability of the Read Codes for use in a clinical charting and decision-support system.

**Review of Related Literature**
Cimino et al.(1989) critiqued several controlled vocabularies designed for clinical use and outlined key qualities necessary for an effective clinical vocabulary: domain completeness, unambiguity, nonredundancy, explicit representation of synonyms, multiple classification of terms, and explicit relationships among terms. Huff and Warner (1990) noted some of the flaws in existing vocabularies: not comprehensive, no institutional support, not sufficiently detailed, not updated in a timely fashion, not proven in the clinical environment, and not flexible enough for decision support processing.

Lindberg and Humphreys (1992) outlined several guidelines for the use of controlled vocabulary in automated patient records. They recommended that controlled vocabulary be used in at least selected core sections of the record. They also recommended that the use of acronyms and abbreviations be kept to a minimum because their presence inhibited effective information retrieval. Additionally, they noted the importance of having machine-readable links to other clinical vocabularies. McDonald (1992) and McHugh (1992) delineated some of the key data elements that a clinical vocabulary would have to be capable of representing in order to meet the needs of physicians and nurses. The ANSI HISPP Working Group on Codes and Vocabularies (1994) developed a draft framework for evaluating clinical vocabularies that used four dimensions for evaluation of vocabularies: scope, structural characteristics, maintenance characteristics, and useability characteristics. The National Library of Medicine (1994) conducted a small preliminary study comparing version 3.1 of the Read
Codes with the SNOMED International clinical vocabulary. They found that Read had twice as many terms for disorders and clinical findings as SNOMED. However, they attributed this finding to Read's greater number of precoordinated terms. They also found cultural differences in the administrative terms in Read. They concluded that Read and SNOMED might provide useful complimentary coverage of clinical concepts.

**Read Codes**

The Read Codes are a comprehensive, hierarchically arranged, coded thesaurus of terms used in health care. The Read Codes were originally designed by Dr. James Read for the purpose of maintaining a computerized patient record. In 1990, the Read Codes were purchased by the British Secretary of State for Health and they became Crown Copyright. Since then, the Codes have been further developed and maintained by the British National Health Service Centre for Coding and Classification (NHS CCC). The NHS CCC is working with over forty-five medical and allied health colleges in Britain in order to expand the terms in the Codes. For example, terms projects for nurses, midwives, speech therapists, occupational therapists, physiotherapists, and dieticians are currently underway or nearing completion.

The developers of the Read Codes aim to produce a comprehensive terminology capable of reflecting all the concepts that may be written in the clinical record by any health care provider. The main chapters of the Read Codes cover administration, causes of injury and poisoning, history and observations, investigations, occupations, operations and procedures, staging and scales, tumour morphology, disorders, regimes and therapies, provision of appliances, and prevention. There are currently approximately 130,000 preferred terms, and 150,000 non-preferred terms.

The preferred terms for each concept are arranged in hierarchy that does not represent a true classification, but rather groups terms together in a way that is clinically meaningful. The hierarchy allows for expansion to incorporate additional terms at each level without disturbing the position of current terms. Each
preferred and non-preferred term in the Read Codes has a unique alpha-numeric code that acts as a unique identifier for that term and is fixed over time. The code does not represent a term's position in the hierarchy. The Read Codes are manually and electronically cross-referenced to many other national and international classifications such as ICD-9, ICD-9CD, OPCS-4, RCGP 1986, ICPC 1986, BNF, and the ATC drug classification. New releases of the Read Codes are issued quarterly. New terms are added to each release as a result of collaboration with relevant health care bodies, or as a result of feedback from users. Users can allocate new terms to holding areas pending their inclusion in the next release (NHS Centre for Clinical Classification 1993).

Evaluative Framework

A two-part framework was used to evaluate a sample of the Read Codes. First, the Codes were evaluated using the Proposed American Standard Guidelines for the Construction, Format, and Management of Monolingual Thesauri (ANSI/NISO Z39.19-199X). This standard, developed by the National Information Standards Organization in the United States, was chosen because it is the most current standard available. This standard is very similar in content to the corresponding British standard (BS 5723:1987) for thesauri. Second, the Read Codes were evaluated by posing the following series of questions derived from criteria outlined in the literature:

Coverage

Scope

Is the vocabulary capable of representing all of the concepts found in the complete patient record? Does the vocabulary have the terms necessary to represent the full range of health problems in various health care settings ie. acute care, long term care, community care? Does the vocabulary encompass the terminology used to describe the various diagnostic and therapeutic procedures performed by different care providers and specialty groups? Does the vocabulary use terms that are commonly used by care providers? Does the vocabulary include related terms, as well as synonyms and variant forms of
terms? Does the vocabulary include modifiers or qualifiers that express the certainty, degree, or severity of a process? Is the vocabulary able to represent time intervals? Are users able to add terms to the vocabulary in order to meet local needs?

*Specificity*

Is the vocabulary specific enough to accurately represent the many aspects of health care reality? Is there minimal loss of clinical detail when data are encoded in the vocabulary? Does the vocabulary capture information in sufficient detail to support efficient statistical reporting for research and policy development purposes? What is the proportion of atomic to precoordinated terms in the vocabulary?

*Structure*

Are the vocabulary hierarchies logical and complete? Are the meanings of terms clearly defined, either by their position in a hierarchy or by a scope note? Does the vocabulary divorce the hierarchical arrangement of a concept from its unique identifier? Does the vocabulary contain redundant terms? Are there explicit rules for combining terms, or for combining terms and qualifiers? Does the vocabulary allow for multiple classification of terms, that is, can terms appear in more than one hierarchy?

*Maintenance*

Does the vocabulary have ongoing institutional support? Does the institution or body that developed the vocabulary have stable funding? Does this institution or body regularly evaluate and update the vocabulary? Does this agency regularly consult with users of the vocabulary on a formal or informal basis in order to obtain feedback?

*Useability*

Is the vocabulary electronically mapped to other major clinical vocabularies? Does the vocabulary meet the needs of a range of end users? Does the user interface facilitate optimal use of the vocabulary with minimal training?

*Sample*
Sixty concepts selected from a manual care map used for managing myocardial infarction were coded using version 3.0 of the Read Codes. The care map is currently in use in the Intensive Care Unit of the Sturgeon General Hospital in St. Albert, Alberta.

**Results**

**ANSI/NISO Z39.19-199X Evaluation**

Section 3.1 recommends that each descriptor in a thesaurus should represent a single concept or unit of thought. The concept may be expressed by a single-word term, or a multi-word term. Some of the descriptors in the Read sample contained more than one concept: "pulmonary artery catheter insertion via jugular vein." This descriptor should be separated into two or three separate descriptors.

Section 3.2 describes a number of methods used to clarify the meaning of potentially ambiguous terms. For example, scope notes are used to clarify the meaning of descriptors when their meaning may not be obvious from their place in a hierarchy. None of the Read descriptors had scope notes. Consequently, users are left to guess at the intended meaning of some descriptors. Inconsistent coding may result.

Section 3.4.1 recommends that where possible, descriptors should be nouns or noun phrases, and that noun phrases should exclude prepositions. A number of Read terms were not nouns: "seen in hospital casualty" and "seen by dietician." Prepositions were included in descriptors: "advice on diet." Section 3.6.2.2 deals with the preferred spelling for descriptors. A few descriptors did not reflect common North American spelling practices: "ischaemic" rather than ischemic, and "gynaecology" rather than gynecology.

Section 4 provides guidelines for developing multiword descriptors. The standard recommends that multiword terms should be used as descriptors if users in the domain commonly use the compound term, or if splitting the term into its component parts would lead to a loss of meaning or ambiguity. The Read Codes contain numerous lengthy compound terms: "extrinsic coagulation pathway"
observation", "pulmonary artery catheter insertion via subclavian vein", and "adverse reaction to agents mainly affecting blood constituents." All of these descriptors express several concepts, and would function more effectively if they were divided into their component parts. Section 4.3 of the standard provides guidelines for splitting terms. Section 4.3.2.4 states that multiword terms that serve only to group a set of narrower terms should not be used as descriptors, but should rather be clearly designated as node labels. The Read descriptor "adverse reaction to drugs/medicines/biological substances" would be a likely candidate for conversion to a node label.

Section 5 of the standard deals with the display of equivalence, hierarchical, and associative relationships in thesauri. The Read Codes displayed no associative relationships. The usefulness of the Read Codes as a coding and retrieval tool would be enhanced by including associative terms, since users would have access to a wider range of terms in the vocabulary.

Section 6.3.2 of the standard gives recommendations for basic information about the thesaurus that should be included in the introduction to a thesaurus. While most of this information was present in the print manual that accompanied the Read Code browser, it would be useful to include this information online as well. Section 7 of the standard gives guidelines for the electronic screen display of a thesaurus, outlining the different needs of three categories of users: thesaurus maintainers, expert searchers and indexers, and end-users. In the Version 3.0 Read Code browser, the difference between preferred terms and non-preferred terms might not be evident to a novice end-user, since the preferred terms were not clearly marked as such. Bold typeface, or some similar indicator would help to clarify the distinction. Alternate displays, such as an alphabetic or permuted display would also be useful to all classes of users.

Questions

Coverage

The Read Codes are not presently comprehensive in their coverage. Some of the existing hierarchies are incomplete, and terms for various allied health
professions have yet to be incorporated into the Codes. For example, version 3.0 does not contain drugs, chemicals, or nursing and allied health terminology. Studies using large, representative samples from a variety of domains need to be conducted in order to fully assess Read coverage. Provisions for consultation with professional bodies and actual users will help to ensure that the terms included in Read will reflect common usage. While most of the terms in the Read Codes are commonly used in North America, some of the administrative terms and the terms for care providers are distinctly British. For example, the emergency department is referred to as the hospital casualty. Home care nurses are referred to as health visitors or district nurses. In some instances, terms commonly used in the clinical setting are not represented in Read. For example, there is no term in Read for "vital signs." There are terms for each specific vital sign: "blood pressure", "respiratory observation", "pulse characteristics", and "body temperature observation."

The Read Codes contain a high proportion of multiword, precoordinated descriptors. As already stated earlier, precoordinated terms can facilitate precision in representing and retrieving clinical data. At a certain point however, lengthy compound terms become counterproductive. They are difficult for users to read, particularly when they appear in lengthy picking lists (NHS Centre for Coding and Classification 1993b, 3). They also increase the size and complexity of the thesaurus, making it more difficult and time consuming to maintain the vocabulary. The inclusion of qualifiers/modifiers in version 3.1 of Read will enable users to express concepts at a greater level of detail (NHS Centre for Coding and Classification 1993b, 3). The provision for users to add local terms to the vocabulary increases flexibility, however it also dilutes the benefits of standardized vocabulary.

**Structure**

As already stated, the addition of scope notes to clarify the meaning of potentially ambiguous descriptors would be useful. The Read Codes divorce the arrangement of a term in a hierarchy from its unique identifier. That is, Read term
identifiers do not indicate the place of a term in a hierarchy. This enables vocabulary developers to move terms within or between hierarchies in response to changes in scientific knowledge (Humphreys 1994, 474). This also obviates the problem of running out of room for new codes. Additionally, because the Read Code unique identifiers are fixed over time, users can continue to access the vocabulary via a consistent code (Humphreys 1994, 474).

The Read Codes have explicit rules for combining terms, and for combining terms and qualifiers (NHS Centre for Coding and Classification 1993b, 15). This provision should reduce instances of unintended redundancy (National Library of Medicine 1994). Read also allows for multiple classification of terms. For example, pulmonary tuberculosis is classified under infectious diseases and under respiratory diseases.

Maintenance & Useability
As mentioned previously, ongoing responsibility for maintaining and developing the Read Codes rests with the British National Health Service Centre for Coding and Classification. Relative to other clinical vocabularies, the Read Codes have good institutional support. New versions of the Read Codes are released on a quarterly basis, and changes are made to the Codes as a result of collaboration with national bodies or end users (NHS Centre for Coding and Classification 1993, 53). The Read Codes are electronically mapped to several other major clinical vocabularies as indicated earlier. Useability is a difficult quality to assess in an automated vocabulary, since it is not always possible to draw distinct boundaries between the interface and the vocabulary itself. A variety of displays, rather than just the hierarchical display, and more distinct identification of preferred terms would help to make the existing browser more user-friendly.

Conclusion
In many respects the Read Codes are a promising clinical vocabulary. Some of the key strengths of the Read Codes are: specificity, multiple classification, ongoing institutional support, consultation with users, separation of term hierarchy arrangement from unique identifiers, and explicit rules for combination
of terms. Weaknesses of Read include: precoordination levels that are unwieldy, term ambiguity, and absence of associative relationship displays. Further studies will be needed in order to obtain more conclusive evidence on the relative merits and deficiencies of the Read Codes. The National Library of Medicine is considering including the Read Codes in a large scale study of several clinical vocabularies. The purpose of this study is to determine the suitability of these vocabularies as a base for an eventual American standard health care vocabulary (National Library of Medicine 1994). An English study to compare the clinical coverage of Read 3 with ICD-9, ICD-10, and the UK procedure code is currently in the planning stages (National Library of Medicine 1994).

**Recommendations for Further Study**

Larger studies are needed that will examine the capacity of Read to represent the content of the entire patient record, including parts of the record generated by nurses and allied health professionals. Retrieval studies and user studies, preferably in a clinical environment, are also needed.

**References**


