Designation: E2522 – 07 (Reapproved 2013)

Standard Guide for Quality Indicators for Health Classifications[1]

This standard is issued under the fixed designation E2522; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ε) indicates an editorial change since the last revision or reapproval.

1. Scope

1.1 This international standard is intended to document principal ideas which are necessary and sufficient to assign value to a classification. The standard will serve as a guide for governments, funding agencies, terminology developers, terminology integration organizations, and the purchasers and users of classification systems toward improved terminological development and recognition of value in a classification. It is applicable to all areas of health about which information is kept or utilized. Appropriately, classifications should be evaluated within the context of their stated scope and purpose. It is intended to complement and utilize those notions already identified by other national and international standards bodies. This standard explicitly refers only to classifications. This international standard will also provide classification developers and authors with the quality guidelines needed to construct useful, maintainable classifications. These tenets do not attempt to specify all of the richness which can be incorporated into a classification. However, this standard does specify the minimal requirements, which if not adhered to will assure that the classification will have limited generalizability and will be very difficult if not impossible to maintain. We have used the word "Shall" to indicate mandatory requirements and the word "Should" to indicate those requirements which we feel are desirable but may not be widely achievable in current implementations. Classifications, which do not currently meet these criteria, can be in compliance with this standard by putting in place mechanisms to move toward these goals. This standard will provide classification developers with a sturdy starting point for the development of useful classifications. This foundation serves as the basis from which classification developers will build robust concept systems.

2. Referenced Documents

2.1 *Normative References*—The following normative documents contain provisions, which through reference in this text, constitute provisions of this Guide E2522. For dated references, subsequent amendments to, or revisions of, any of these publications do not apply. However, parties to agreements based on Guide E2522 are encouraged to investigate the possibility of applying the most recent editions of the normative documents indicated below. For undated references, the latest edition of the normative document referred to applies. Members of ISO and IEC maintain registers of currently valid International Standards.

2.2 *ASTM Standards:*[2]

E1238 Specification for Transferring Clinical Observations Between Independent Computer Systems (Withdrawn 2002)[3]

E1239 Practice for Description of Reservation/Registration-Admission, Discharge, Transfer (R-ADT) Systems for Electronic Health Record (EHR) Systems

E1284 Guide for Construction of a Clinical Nomenclature for Support of Electronic Health Records (Withdrawn 2007)[3]

E1384 Practice for Content and Structure of the Electronic Health Record (EHR)

E1633 Specification for Coded Values Used in the Electronic Health Record

2.3 *ISO Standards:*[4]

ISO 704 Principles and Methods of Terminology

ISO/DIS 860 International Harmonization of Concepts and Terms

ISO 1087-2 Terminology—Vocabulary—Part 2: Computer Applications

ISO 11179-3 Terminology—Data Registries

ISO 12200 Terminology—Computer Applications—Machine Readable Terminology Interchange Format

ISO 12620 Terminology—Computer Applications—Data Categories

ISO 15188 Project Management for Terminology Standardization

[1] This guide is under the jurisdiction of ASTM Committee E31 on Healthcare Informatics and is the direct responsibility of Subcommittee E31.35 on Healthcare Data Analysis.

Current edition approved March 1, 2013. Published March 2013. Originally approved in 2007. Last previous edition approved in 2007 as E2522–07. DOI: 10.1520/E2522-07R13.

[2] For referenced ASTM standards, visit the ASTM website, www.astm.org, or contact ASTM Customer Service at service@astm.org. For *Annual Book of ASTM Standards* volume information, refer to the standard's Document Summary page on the ASTM website.

[3] The last approved version of this historical standard is referenced on www.astm.org.

[4] Available from American National Standards Institute (ANSI), 25 W. 43rd St., 4th Floor, New York, NY 10036, http://www.ansi.org.

Copyright © ASTM International, 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428-2959. United States

ISO 2382-4 Information Technology—Vocabulary—Part 4: Organization of Data
TR 9789 Guidelines for the Organization and Representation of Data Elements for Data Interchange—Coding Methods and Principles
2.4 *CEN Standards:*[5]
ENV 12017 Medical Informatics—Vocabulary

3. Terminology

3.1 *For the purposes of this guide, the following terms and definitions apply:*

3.1.1 *canonical term*—a preferred atomic or pre-coordinated term for a particular medical concept.

3.1.2 *classification*—collection of terms grouped by a common characteristic. Usually not intended to represent the full content of a knowledge domain. Classifications are an aggregation of a nomenclature. A classification is a terminology which aggregates data at a prescribed level of abstraction for a particular domain. This fixing of the level of abstraction that can be expressed using the classification system is often fixed to enhance consistency when the classification is to be applied across a diverse user group, such as is the case with some of the current billing classification schemes. Examples are ICD9-CM and CPT.

3.1.3 *controlled health vocabulary*—a terminology intended for clinical use. This implies enough content and structure to provide a representation capable of encoding comparable data, at a granularity consistent with that generated by the practice within the domain being represented, within the purpose and scope of the terminology.

3.1.4 *index term*—a pointer to a concept in a classification. This can be a synonym, abbreviation, acronym or some mnemonic which can be used to indicate the correct code to use from the classification. Such use is important for mapping from the way clinicians tend to speak to the classification. These have been referred to as Entry terms by some authors.

3.1.5 *modifier*—a string which, when added to a term, changes the meaning of the term in the Clinical sense (for example, clinical stage or severity of illness).

3.1.6 *nomenclature*—the canonical set of terms comprising a given controlled vocabulary; their structure, relationships and, if existing, systematic and formal definitions; and the code, meaning formal rules and general principles, guiding how the controlled vocabulary may be changed.

3.1.7 *ontology*—an organization of concepts by relationships for which one can make a rational argument. Colloquially, this term is used to describe a hierarchy constructed for a specific purpose. For example, a hierarchy of qualifiers would be a Qualifier Ontology.

3.1.8 *qualifier*—a string which, when added to a term, changes the meaning of the term in a Temporal or Administrative sense (for example, "History of" or "Recurrent").

3.1.9 *term*—a word or words corresponding to one or more concepts.

3.1.10 *terminology*—set of terms representing a system of concepts within a specified domain.

3.1.10.1 *Discussion*—This implies a published purpose and scope from which one can determine the degree to which this representation adequately covers the domain specified.

4. General

4.1 *Basics*—Basic characteristics of a terminology influence its utility and appropriateness in clinical applications.

4.2 *Concept Orientation*—The basic unit of a terminology shall be a concept, which is the embodiment of some specific meaning and not a code or character string. Identifiers of a Concept shall correspond to one and only one meaning and in a well-ordered vocabulary only one concept may have that same meaning (DIS 860). However, multiple terms (linguistic representations) may have the same meaning if they are explicit representations of the same concept. This implies non-redundancy, non-ambiguity, non-vagueness and internal consistency.

4.2.1 *Non-Redundancy*—Terminologies shall be internally normalized. There shall not be more than one concept identifier in the terminology with the same meaning (ISO 704, Guide E1284). This does not exclude synonymy; rather, it requires that this be explicitly represented.

4.2.2 *Non-Ambiguity*—No concept identifier should have more than one meaning. However, an entry term (some authors have referred to this as an "interface terminology") can point to more than one concept (for example, MI as Myocardial Infarction and Mitral Insufficiency).

4.2.3 *Non-Vagueness*—Concept names shall be context free (some authors have referred to this as "context laden"). For example, "diabetes mellitus" should not have the child concept "well controlled"; instead, the child concept's name should be "diabetes mellitus, well controlled."

4.2.4 *Internal Consistency*—Relationships between concepts should be uniform across parallel domains within the terminology. For example, if heart valve structures are specified anatomically, the diagnosis related to each structure should also be specified using the same relationships.

4.3 *Purpose and Scope*—Any classification shall have its purpose and scope clearly stated in operational terms so that its fitness for particular purposes can be assessed and evaluated (ISO 15188). Where appropriate, it may be useful to illustrate the scope by examples or 'use cases' as in database models and other specification tools. Criteria such as coverage and comprehensiveness can only be judged relative to the intended use and scope. For example, a classification might be comprehensive and detailed enough for aggregation of billing codes from a hospital admission (for example, DRGs), but inadequate for specifying the indication for a surgical procedure.

4.3.1 *Coverage*—Each segment of the healthcare process shall have explicit in-depth coverage and not rely on broad leaf node categories that lump specific clinical concepts together. For example, it is often important to distinguish specific diagnosis from categories presently labeled "Not Elsewhere Classified" (NEC) or to differentiate disease severity such as indolent prostate cancer from widely metastatic disease. The

[5] Available from European Committee for Standardization (CEN), 36 rue de Stassart, B-1050, Brussels, Belgium, http://www.cenorm.be.

extent to which the depth of coverage is incomplete shall be explicitly specified for each domain (scope) and purpose as indicated in 4.3. (1)[6]

4.3.2 *Comprehensiveness*—The extent to which the degree of comprehensiveness is incomplete shall be explicitly specified for each domain (scope) and purpose as indicated in 4.3. Within the scope and purpose, all aspects of the healthcare process shall be addressed for all related disciplines, such as physical findings, risk factors, or functional status—across the breadth of medicine, surgery, nursing, and dentistry. This criterion applies because decision support, risk adjustment, outcomes research, and useful guidelines require more than diagnoses and procedures. Examples include existing Agency for Healthcare Research and Quality guidelines, and the CMS mortality model. (2)

4.4 *Mapping:*

4.4.1 Government and payers mandate the form and classification schema for much clinical data exchange. Thus, comprehensive and detailed representations of patient data within computer-based patient records should be able to be mapped to those classifications, such as ICD-9. This need for multiple granularities is needed for clinical healthcare as well (ISO TR 9789). For example, an endocrinologist may specify more detail about a patient's Diabetes Mellitus than a generalist working in an urgent care setting or a nurse assessing the extent to which the individual is coping with their disorder, even though all may be caring for the same patient. The degree to which the terminology is mappable to other classifications shall be explicitly stated. (3)

4.4.2 The rules for mapping from a classification to a nomenclature are the same rules that are defined for the formation of compositional expressions from the reference terminology and therefore are by definition terminology dependent. As a well-formed classification is defined as a pre-coordination from the reference terminology, different constructs (codes in the classification) developed are built for a specific purpose. Therefore, if a classification is useful for a purpose and it is fully specified, then by definition it is coordinate with the reference terminology. Two classifications defined by the same reference terminology are coordinate and interoperable. Mapping here is accomplished at the level of the reference terminology and therefore is correct at the lowest level of computable meaning.

4.5 *Systematic Definitions*—In order for users of the terminology to be certain that the meaning that they assign to concepts is identical to the meaning which the authors of the vocabulary have assigned, these definitions will need to be explicit and available to the users. Further, as relationships are built into vocabularies, multiple authors will need these definitions to ensure consistency in authorship. For example, the concept "Hypertension" might be defined as a consistently elevated Blood Pressure and not "BP > 140/85."

4.6 *Explicitness of Relations*—The logical definition of subsumption should be defined. The formal behavior of all links/relations/attributes should be explicitly defined. If a looser meaning such as "broader than/narrower than" is used, it should be explicitly stated. For example, the primary hierarchical relation should be subsumption as exemplified by logical implication: "B is a kind of A" means "All B's are A's."

4.7 *Multiple Hierarchies*—Concepts should be accessible through all reasonable hierarchical paths (that is, they shall allow multiple semantic parents). For example, stomach cancer can be viewed as a neoplasm or as a gastrointestinal disease. A balance between number of parents (as siblings) and number of children in a hierarchy should be maintained. This feature assumes obvious advantages for natural navigation of terms (for retrieval and analysis) as a concept of interest can be found by following intuitive paths (that is, users should not have to guess where a particular concept was instantiated). (4)

4.8 *Consistency of View*—A concept in multiple hierarchies shall be the same concept in each case. Our example of stomach cancer shall not have changes in nuance or structure when arrived at via the cancer hierarchy as opposed to GI diseases. Inconsistent views could have catastrophic consequences for retrieval and decision support by inadvertently introducing variations in meaning which may be unrecognized and therefore be misleading to users of the system. (5)

4.9 *Explicit Uncertainty*—Notions of "probable," "suspected," "history of," or differential possibilities (that is, a Differential Diagnosis list) shall be supported. The impact of certain versus very uncertain information has obvious impact on decision support and other secondary data uses. Similarly, in the case of incomplete syndromes, clinicians should be able to record the partial criteria consistent with the patient's presentation. This criterion is listed separately as many current terminological systems fail to address this adequately.

4.10 *Representational Form*—The representational form of the identifiers within the terminology should be meaningless. Computer coding of concept identifiers shall not place arbitrary restrictions on the terminology, such as numbers of digits, attributes, or composite elements. To do so subverts meaning and content of a terminology to the limitations of format, which in turn often results in the assignment of concepts to the wrong location because it might no longer "fit" where it belongs in a hierarchy. These reorganizations confuse people and machines alike, as intelligent navigation agents are led astray for arbitrary reasons. The long, sequential, alphanumeric tags used as concept identifiers in the UMLS project of the National Library of Medicine exemplify well this principle.

5. Use Cases for Bounding Classifications

5.1 *Principles:*

5.1.1 Classifications should serve a stated purpose.

5.1.2 Classifications should have an explicit scope.

5.1.3 The quality of the classification should only be evaluated within the context of its purpose and scope.

5.1.4 Classifications should be created and maintained in response to a real world need.

5.1.5 Classification should be able to be derived from aggregations of data specified in detailed health nomenclatures.

5.1.6 Classification should evolve gracefully.

[6] The boldface numbers in parentheses refer to the list of references at the end of this standard.

5.2 *Use Cases for Classifications*—These use cases set the boundary definitions for the types of uses which are more appropriate for a controlled health terminology or a classification. The scale is from 0 to 10 with ten being most appropriate for a nomenclature and zero being most appropriate for a classification. These categories are slices across a general problem that needs to address the boundaries and overlap between these two types of knowledge representation.

Specifying categories of health problems or procedures for the purpose of determining reimbursement	3
Public Health Surveillance	6
Utilization Review	3
Managing the financial side of Medical Practice. Budget Planning, Purchasing, etc.	3
Patient Safety	8
Hospital Discharge Planning	6
Effectiveness of Care (Planned interventions)	7
Education (practitioner, patient)	6
Quality Assurance	8
Credentialing (competencies)	5
EBM	8
Decision Support	9
Information Retrieval (Querying)	7
Data Representation	9
Mandatory Reporting	3
NLP	9
Outcomes Analysis	7
Morbidity Coding	7
Mortality Coding	5
Health Indicators	5
Healthcare provider roles	1
Partition of Health Information on the web	6
Data Warehousing	7

5.3 *Questions to be Answered*—We have attempted to answer these questions within the body of the standard.

5.3.1 What domains are valid to represent in a classification?

5.3.2 What is the boundary between classifications and detailed nomenclatures?

5.3.3 What are the rules for mapping between nomenclatures and classifications?

5.3.4 How can multiple consistent views allow classifications to serve multiple purposes? For example, a classification which needs to provide administrative aggregation for public health purposes as well as reimbursement information.

5.3.5 Are there indicators of quality for mapping between and among classifications?

5.3.6 What are the quality indicators for building, organizing, distributing, and maintaining classifications?

5.3.7 What are the international concerns regarding the methodology identified in 5.3.6?

6. Maintenance

NOTE 1—Technical choices can impact the capacity of a terminology to evolve, change, and remain usable over time.

6.1 *Context Free Identifiers*—Unique codes attached to concepts shall not be tied to hierarchical position or other contexts; their format shall not carry meaning. Because health knowledge is being constantly updated, how we categorize health concepts is likely to change (for example, Peptic Ulcer Disease is now understood as an infectious disease, but this was not always so). For this reason, the "code" assigned to a concept shall not be inextricably bound to a hierarchy position in the terminology so that we need not change the code as we update our understanding of, in this case, the disease. Changing the code may make historical patient data confusing or erroneous. This notion is the same as Non-Semantic Identifiers. **(6)**

6.2 *Persistence of Identifiers*—Codes shall not be reused when a concept is obsolete or superseded. Consistency of patient description over time is not possible when concepts change codes; the problem is worse when codes can change meaning. This practice not only disrupts historical analyses of aggregate data, but can be dangerous to the management of individual patients whose data might be subsequently misinterpreted. This encompasses the notion of Concept Permanence.

6.3 *Version Control*—Updates and modifications shall be referable to consistent version identifiers. Usage in patient records should carry this version information. This is true because the interpretation of coded patient data is a function of terminologies that exist at a point in time (for example, AIDS patients were coded inconsistently before the introduction of the term AIDS). Terminology representations should specify the state of the terminology system at the time a term is used; version information most easily accomplishes this and may be hidden from ordinary review (ISO 15188, ISO 12620, ISO 1087-2, ISO 11179-3, ISO 2382-4). **(7, 8)**

6.3.1 *Editorial Information*—New and revised terms, concepts, and synonyms shall have their date of entry or effect in the system along with pointers to their source or authority, or both. Previous ways of representing a new entry should be recorded for historical retrieval purposes.

6.3.2 *Obsolete Marking*—Superseded entries should be so marked, together with their preferred successor. Because data may still exist in historical patient records using obsolete terms, their future interpretation and aggregation are dependent upon that term being carried and cross-referenced to subsequent terms (for example, HTLV III to HIV).

6.4 *Recognize Redundancy*—Authors of these large-scale vocabularies will need mechanisms to identify redundancy when it occurs. This is essential for the safe evolution of any such vocabulary.

6.5 *Language Independence*—It would be desirable for classifications to support multilingual presentations. As healthcare confronts the global economy and multiethnic practice environments, routine classification maintenance shall incorporate multilingual support. While substantially lacking the power and utility of machine translation linguistics, this simplistic addition will enhance understanding and use globally. Have there been translations? What is the expected cost of translation?

6.6 *Responsiveness*—The frequency of updates, or subversions, should be sufficiently short to accommodate new codes and repairs quickly, ideally on the order of weeks.

6.7 *Distribution Format:*

```
<Classification name = " ">
  <Code>
    <Text-Long, Language = " ">
    <Text-Short, Language = " ">
    <Entry Term, Language = " ">
    <Relation name = " " code = " ">
```

<Status {Hx, Active, Inactive}>
<Scope>
<Purpose>
<RT-Comp Expr>
<Boolean connector>
<Code, Coding Scheme>
<Relation_Code, Coding Scheme>
<Code, Coding Scheme>
<Date>
<Author>
<Systematic Definition>

7. Evaluation

NOTE 2—As we seek to understand quality in the classifications that we create or use, we need standard criteria for the evaluation of these systems. All evaluations shall reflect and specifically identify the purpose and scope of the classification being evaluated. (**9**)

7.1 *Purpose and Scope*—Important dimensions along which scope should be defined include:

7.1.1 *Clinical Area*—What is the clinical area of use of the classification, the disease area of patients addressed or the expected profession of users, or both? Within what parts of healthcare is it intended to be used and by whom?

7.1.2 *Primary Use*—What is the primary intended usage of the classification? Examples include: reporting for remuneration, management planning, epidemiological research, indexing for bibliographic, web-based retrieval, recording of clinical details for direct patient care, use for decision support, linking of record to decision support, etc.

7.1.3 *Persistence and Extent of Use*—While some classifications are intended, at least initially, primarily for a specific study or a specific site, others are not. If intended to be persistent, what is the mechanism for effecting change management? For example, marking codes as obsolete and then pointing those codes where appropriate to a new concept.

7.1.4 *Transformations (Mappings) to Other Vocabularies*—What transformations/mappings are supported for what intended purpose? For example, transformation for purposes of bibliographic retrieval may require less precision than transformation for clinical usage? What is the sensitivity and specificity of the mappings?

7.1.5 *User/Developer Extensibility*—Is it intended that the classification be extended by users or application developers? If so, within what limits? If not, what mechanisms are available for meeting new needs as they arise?

7.1.6 *Natural Language*—Is natural language input or output supported (for analysis or input)? To what level of accuracy?

7.1.7 *Other Functions*—What other functions are intended? For example, linkage to specific decision support systems, linkage to post-marketing surveillance, etc.

7.1.8 *Current Status*—To what extent is the system intended to be "finished" or work in progress? If different components of the classification are at different stages of completion, how is this indicated?

7.2 *Measures of Quality—Terminological Tools:*

7.2.1 *Interconnectivity (Mapping):*

7.2.1.1 *Classification and Other Coding Systems*—To what extent is the classification mappable to other coding systems or reference terminologies?

7.2.1.2 *Classification and Terminological Enhancements*—To what extent can the classification accommodate local terminological enhancements?

7.2.1.3 *Classification and Networking*—Can the vocabulary server respond to queries sent over a network (LAN, WAN)?

7.2.2 *Precision and Recall:*

7.2.2.1 *Classification*—What are the classification's precision and recall for mapping Diagnoses, Procedures, Manifestations, Anatomy, Organisms, etc. against an established and nationally recognized standard query test set, using a standard well-principled method? This should be evaluated only within the intended scope and purpose of the vocabulary system.

7.2.2.2 *Search Engine*—Is a standard search engine used in the mapping process?

7.2.2.3 *Inter-rater Reliability*—To what degree can users reproduce the same definition for the same concept (for a valid random sample of concepts from the classification)?

7.2.3 *Usability:*

7.2.3.1 *Validation of Usability*—Has the usability of the classification been verified?

7.2.3.2 *Interface Considerations*—How have interface considerations been separated from classification evaluation?

7.2.3.3 *Prototypes*—Has an effective user interface been built? Has the classification been shown to have an effective user interface for its intended use? If not, what are the questions or issues outstanding? Evidence for speed of entry, accuracy, comprehensiveness in practice, etc. with different approaches? If not, is there a proof of concept?

7.2.3.4 *Application Programmer Interfaces*—Is there support for computer interfaces and system implementers? Is there a demonstrated proof of concept implementation in software? Can it be shown to be usable for the primary purpose indicated? Have there been failed implementations?

7.2.4 *Feasibility*—If it is intended for use in an Electronic Patient Record (EPR), what are the options for information storage? Has feasibility been demonstrated?

7.3 *Measures of Quality*—The generalizability (applicability) of any Study Design reported (Evaluating Reported Evaluations) should be able to be evaluated.

7.3.1 *Healthcare/Clinical Relevance*—What is the classification's Healthcare/Clinical Relevance?

7.3.2 *Gold Standard*—What was the Gold Standard used in the evaluation?

7.3.3 *Specific Aims*—Were the Specific Aims clear?

7.3.4 *Blinding*—Was the study appropriately blinded?

7.3.5 *Randomization*—Was the Test Set Selection randomized or shown in some sense to be a representative sample of the end user population?

7.3.6 *Test Location:*

7.3.6.1 *Independence*—Was it different from the developer's location?

7.3.6.2 *Appropriate for Study Design*—How was the test site suited to the study design (tools, resources, etc.)?

7.3.6.3 *Principle Investigator:*

(1) Was the Principle Investigator independent of the vocabulary being evaluated?

(2) Does the principle investigator have a track record of publication in this field of study?

(3) Have there been any conflicts of interest in performing this research?

7.3.7 *Project Completion*—Was the project completed in a reasonable period of time?

7.3.8 *Sample Size:*

7.3.8.1 *Power*—Was the sample size of sufficient size to show the anticipated effect, should one exist?

7.3.8.2 *Statistics*—Who reviewed the Statistical Methods?

7.3.9 *Personnel:*

7.3.9.1 *Training Level*—What is the average level of training of the study personnel?

7.3.9.2 *Reviewers:*

(1) Variability—What is the inter-reviewer variability?

(2) Type—What was the type of reviewer (physician, nurse, other clinician, coder, knowledge engineer) used in the study?

(3) Independence—Were the reviewers blinded to the other reviewers' judgments (that is, reviewer independence)?

APPENDIX

(Nonmandatory Information)

X1. RATIONALE

X1.1 Introduction

X1.1.1 In 1839 William Farr stated in his First Annual Report of the Registrar-General of Births, Deaths, and Marriages in England, "The nomenclature is of as much importance in this department of inquiry, as weights and measures in the physical sciences, and should be settled without delay." Since that time, this theme has been heard resounding from an increasingly large group of scientists. Today, the need for controlled vocabularies to support health record systems has been widely recognized (Specification E1238, Guide E1239, Guide E1384, Specification E1633, ENV 12017). Classifications provide systems with the means to aggregate data. This aggregation of data can be done at multiple levels of granularity and therefore can enhance the clinical retrieval of a problem-oriented record, data pertaining to a classification for billing purposes, or outcomes data for a given population. Maintenance of large-scale classifications has become a burdensome problem as the size of term sets has escalated (ISO 15188). Without a well-structured backbone, classifications cannot scale to provide the level of accuracy required by today's electronic health record and epidemiologic applications.

X1.1.2 The solution rests with standards (**10**). Over the past ten or more years, Medical Informatics researchers have been studying concept representation issues directly. They have examined the structure and content of existing classifications to determine why they seem unsuitable for particular needs and they have proposed solutions. In some cases, proposed solutions have been carried forward into practice and new experience has been gained (**11**). As we prepare to enter the twenty-first century, it seems appropriate to pause to reflect on this experience, and publish a standard set of goals for the development of comparable, reusable, multipurpose, and maintainable controlled health vocabularies (ISO 12200, ISO 12620).

X1.2 History of Classification

X1.2.1 The present coding practices rely on data methods and principles for terminology maintenance that have changed little since the adoption of the statistical bills of mortality in the mid-17th century (**12**). The most widely accepted standard for representing patient conditions, ICD-9-CM (**13**), is an intellectual descendent of this tradition. ICD-9-CM relies overwhelmingly on a tabular data structure with limited concept hierarchies and no explicit mechanism for synonymy, value restrictions, inheritance or semantic and non-semantic linkages. The maintenance environment for this healthcare classification is a word processor and its distribution is nearly exclusively paper-based.

X1.2.2 The first edition of Physicians' Current Procedural Terminology (CPT) terminology appeared in 1966. In the United States, CPT is the coding system used by Medicare and virtually all third-party payers, including workers compensation and Medicaid. As part of the Medicare Part B physician payment schedule, CPT codes are associated with the Resource Based Relative Value Scale (RBRVS) and used to determine payment for services. The CPT code set is Level I of the Health Care Financing Administration Common Procedure Coding System (HCPCS). The CPT code set, currently in its fourth edition, contains numeric modifiers, notes, guidelines and an index designed to provide explanatory information and facilitate the correct usage of the coding system. The American Medical Association (AMA) is currently working to develop the next generation of CPT (that is, CPT-5).

X1.2.3 Significant cognitive advances in disease and procedure representation took place in 1928 at the New York Academy of Medicine, resulting in industry-wide support for what became the Standard Nomenclature of Diseases and Operations. The profound technical innovation was the adoption of a multiaxial classification scheme. (**2, 13**) Now a pathologic process (for example, Inflammation) could be combined with an anatomic site (for example, Oropharynx Component: Tonsil) to form a diagnosis (for example, Tonsillitis). The expressive power afforded by the compositional nature of a multiaxial terminological coding system tremendously increased the scope of tractable terminology and additionally the level of granularity that diagnosis could be encoded about our patients. (**2**)

X1.2.4 The College of American Pathology (CAP) carried the torch further by creating the Systematized Nomenclature of Pathology (SNOP), and subsequently the Systemized Nomenclature of Medicine (SNOMED). In these systems, the number, scope, and size of the compositional structures has increased to the point where an astronomical number of terms can be synthesized from SNOMED atoms. One well-recognized limitation of this expressive power is the lack of syntactic grammar, compositional rules, and normalization of both the concepts and the semantics. Normalization is the process by which the system knows that two compositional constructs with the same meaning are indeed the same (for example, that the term "Colon Cancer" is equivalent to the composition of "Malignant Neoplasm" and the site "Large Bowel"). These are issues addressed by CAP in their efforts to make SNOMED a robust reference terminology for healthcare. **(4, 14)**

X1.2.5 Other initiatives of importance are the Clinical Terms v3 (Read Codes), which are maintained and disseminated by the National Health Service in the United Kingdom and the Galen effort, which expresses a very detailed formalism for term description. The Read Codes are a large corpus of terms, which is now in its third revision that is hierarchically designed and is slated for use throughout Great Britain. A development of interesting note is the joint effort of CAP and the NHS to merge the content of SNOMED-RT and Clinical Terms Version 3 into a derivative work (Announced 4/99), which has been released and named SNOMED Clinical Terms (SNOMED-CT). SNOMED-CT has been adopted for the representation of clinical problems by NCVHS.

X1.2.6 For medications, the National Drug Formulary–Reference Terminology (NDF-RT) is a reference terminology for medications and has been adopted by NCVHS for physiological effects. The clinical drug names are distributed in the UMLS by the national library of medicine as RxNorm. This work was a joint effort by the Veterans Health Administration, the FDA and the NLM.

X1.2.7 Classifications need to exist where the patterns of aggregation needed for a purpose are not adequately supported by the hierarchies of a reference terminology (the operative reference terminology). Classifications shall be exclusively derived as a set of pre-coordinations from a reference terminology.

REFERENCES

(1) Cimino, J. J., "Desiderata for Controlled Medical Vocabularies in the Twenty-First Century," *Methods of Information in Medicine*, 37(4/5), 1998, pp. 394–403.

(2) Cote, R. A., and Rothwell, D. J., "The Classification-Nomenclature Issues in Medicine: A Return to Natural Language," *Medical Informatics*, 14(1), 1989, pp. 25–41.

(3) Rocha, R. A., Rocha, B. H., and Huff, S. M., "Automated Translation Between Medical Vocabularies Using a Frame-based Interlingua," *Proceedings of the Annual Symposium of Computer Applications in Medical Care*, 1993, pp. 690–694.

(4) Campbell, K. E., and Musen, M. A., "Representation of Clinical Data Using SNOMED III and Conceptual Graphs," *Proceedings of the Annual Symposium on Computer Applications in Medical Care*, 1992, pp. 354–358.

(5) Rossi Mori, A., Galeazzi, E., Gangemi, A., Pisanelli, D. M., Thornton, A. M., "Semantic Standards for the Representation of Medical Records," *Medical Decision Making*, 4(Suppl), 1991, pp. S76–S80.

(6) Tuttle, M. S., Olson, N. E., Campbell, K. E., Sherertz, D. D., Nelson, S. J., and Cole, W. G., "Formal Properties of the Metathesaurus," *Proceedings of the Annual Symposium on Computer Applications in Medical Care*, 1994, pp. 145–149.

(7) Campbell, K. E., Cohn, S. P., Chute, C. G., Rennels, G., and Shortliffe, E. H., "Galapagos: Computer-based Support for Evolution of a Convergent Medical Terminology," *JAMIA*, SympSuppl, 1996, pp. 269–273.

(8) Cimino, J. J., "Formal Descriptions and Adaptive Mechanisms for Changes in Controlled Medical Vocabularies," *Methods of Information in Medicine*, 35(3), 1996, pp. 211–217.

(9) Elkin, P. L., ANSI-HISB Code Set Evaluation Criterion Survey, 1998; Minutes ANSI-HISB meeting 4/98.

(10) Masys, D. R., "Of Codes and Keywords: Standards for Biomedical Nomenclature," *Acad Med*., 65, 1990, pp. 627–629.

(11) Solbrig, H. R., Final submission to the CorbaMED Request for Proposals on Lexical Query Services (CorbaLex), OMG, 1998. http://www.omg.org/cgi-bin/doc?formal/99-3-6.pdf or http://www.omg.org/cgi-bin/doc?formal/99-3-1.pdf

(12) Farr, W., "Regarding the Cullenian system of 1785," *First Annual Report of the Registrar-General of Births, Deaths, and Marriages in England*, London, 1839, p. 99.

(13) Evans, D. A., Cimino, J. J., Hersh, W. R., Huff, S. M., and Bell, D. S., for the Canon Group, "Toward a Medical-Concept Representation Language," *JAMIA*, 1, 1994, pp. 207–217.

(14) Musen, M. A., Wieckert, K. E., Miller, E. T., Campbell, K. E., and Fagan, L. M., "Development of a Controlled Medical Terminology: Knowledge Acquisition and Knowledge Representation," *Methods of Information in Medicine*, 34(1-2), March 1995, pp. 85–95.

ASTM International takes no position respecting the validity of any patent rights asserted in connection with any item mentioned in this standard. Users of this standard are expressly advised that determination of the validity of any such patent rights, and the risk of infringement of such rights, are entirely their own responsibility.

This standard is subject to revision at any time by the responsible technical committee and must be reviewed every five years and if not revised, either reapproved or withdrawn. Your comments are invited either for revision of this standard or for additional standards and should be addressed to ASTM International Headquarters. Your comments will receive careful consideration at a meeting of the responsible technical committee, which you may attend. If you feel that your comments have not received a fair hearing you should make your views known to the ASTM Committee on Standards, at the address shown below.

This standard is copyrighted by ASTM International, 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428-2959, United States. Individual reprints (single or multiple copies) of this standard may be obtained by contacting ASTM at the above address or at 610-832-9585 (phone), 610-832-9555 (fax), or service@astm.org (e-mail); or through the ASTM website (www.astm.org). Permission rights to photocopy the standard may also be secured from the ASTM website (www.astm.org/COPYRIGHT/).