Research Paper ■

Evaluation of the Clinical LOINC (Logical Observation Identifiers, Names, and Codes) Semantic Structure as a Terminology Model for Standardized Assessment Measures

Suzanne Bakken, RN, DNSc, James J. Cimino, MD, Robert Haskell, BSEE, Rita Kukafka, DrPH, Cindi Matsumoto, RN, MS, Garrett K. Chan, RN, MS, Stanley M. Huff, MD

A b s t r a c t   **Objective:** The purpose of this study was to test the adequacy of the Clinical LOINC (Logical Observation Identifiers, Names, and Codes) semantic structure as a terminology model for standardized assessment measures.

Methods: After extension of the definitions, 1,096 items from 35 standardized assessment instruments were dissected into the elements of the Clinical LOINC semantic structure. An additional coder dissected at least one randomly selected item from each instrument. When multiple scale types occurred in a single instrument, a second coder dissected one randomly selected item representative of each scale type.

Results: The results support the adequacy of the Clinical LOINC semantic structure as a terminology model for standardized assessments. Using the revised definitions, the coders were able to dissect into the elements of Clinical LOINC all the standardized assessment items in the sample instruments. Percentage agreement for each element was as follows: component, 100 percent; property, 87.8 percent; timing, 82.9 percent; system/sample, 100 percent; scale, 92.6 percent; and method, 97.6 percent.

Discussion: This evaluation was an initial step toward the representation of standardized assessment items in a manner that facilitates data sharing and re-use. Further clarification of the definitions, especially those related to time and property, is required to improve inter-rater reliability and to harmonize the representations with similar items already in LOINC.

■ **J Am Med Inform Assoc.** 2000;7:529–538.

Tremendous progress has been made in the development of standardized sets of health care terms, yet experts agree that gaps and areas of overlap are present in existing standardized terminologies and that no single terminology can meet all needs.[1–5] Non-ambiguous concept representations have been proposed to achieve sharing and re-use of health care data across heterogeneous computer-based systems.[6,7] Recent reports have described representations for diseases,[4,8,9] surgical procedures,[10,11] clinical drugs,[12] nursing activities,[13,14] and laboratory tests.[15,16] Another significant type of information for the delivery and evaluation of health care is stan-

Table 1 ■

Use of Terminology Model and Modified KRSS Representation Language to Create Nursing Activity Molecules from Atomic Terms

Aggregate (Molecular) Expression	Representation of Nursing Activities (Molecules) from Atomic Terms in SNOMED*
Assist patient with feeding	(define-concept "assist patient with feeding" (and procedure
	(some "has object" patient)
	(some "has object" feeding)
	(some "uses technique" assist)))
Teach client about safe sex	(define-concept "teach client safe sex" (and procedure
	(some "has object" client)
	(some "has object" safe sex)
	(some "uses technique" teach)))

NOTE: KRSS indicates Knowledge Representation Specification Syntax (KRSS).

* "Uses technique" is a proposed role; it has not yet received final approval.

dardized assessment data, e.g., data related to functional status, behavioral risk, and quality of life. Toward the goal of data sharing and re-use of standardized assessment data, the purpose of this study was to test the adequacy of the Clinical LOINC (Logical Observation Identifiers, Names, and Codes)[15,16] semantic structure as a terminology model for standardized assessment measures.

Background

In this paper, "terminology model" is used to mean an explicit representation of a system of concepts that is optimized for terminology management and supports the intensional definition of concepts and the extensional mapping among terminologies.[17] A terminology model depicts the relationship between an aggregate (molecular) expression and more primitive (atomic) concepts.[18] In Table 1, for example, atomic terms are used to represent nursing activity molecules through the use of a terminology model[19] that specifies the relationships and a representation language, modified Knowledge Representation Specification Syntax (KRSS).[20] A type definition is a particular aspect of a terminology model that states the semantic links (associative relations) that must be specified for every concept of a particular type.[21] This is synonymous with the notion of a fully specified

name in LOINC. A terminology model is an essential component of a reference terminology[4] or third-generation language system[22] that also includes other types of information (e.g., hierarchic knowledge about represented concepts).[18]

Standardized Assessments

Standardized assessment measures have been developed for many aspects of health care. These include measures of dimensions such as knowledge (e.g., Omaha System[23]), attitudes (e.g., Marlowe-Crowne Social Desirability Scale[24]), beliefs (e.g., Self-efficacy Scale[25]), behavior (e.g., Behavioral Risk Factor Surveillance Survey[26]), function (e.g., Seattle Angina Questionnaire[27]), and status (e.g., Post-anesthesia Score[28]). Assessments are used to predict risk (e.g., Braden Pressure Ulcer Risk Assessment,[29] Center for Epidemiological Studies Depression Scale[30]), determine appropriate interventions (e.g., Abbreviated Injury Scale,[31] Sign and Symptom Checklist for Persons with HIV[32]), monitor health status (e.g., Medical Outcomes Study Short Form 36,[33] Seattle Angina Questionnaire[27]), and evaluate outcomes of health care (e.g., Nursing Outcomes Classification,[34] Karnofsky Performance Status[35]).

While most assessments relate to individual patients or clients, some are focused on other units of analysis, such as family caregivers (e.g., Caregiver Burden Scale,[36] Caregiver Quality of Life Index-Cancer Scale[37]) or communities (e.g., Omaha System[23]). Standardized assessments may reflect professional observation or judgment (e.g., Apgar Score,[38] Morse Fall Score[39]) or the perceptions of the recipient of health care (e.g., Life Satisfaction Index A,[40] Living with HIV[41]).

Although the purposes, dimensions, and units of analysis of standardized health care assessments vary, as do the respondents to them, consistencies are present in data structures at the item level. A standardized assessment is composed of one or more structured narrative items (attributes) with associated responses (values). The items may be expressed in different formats; for instance, as a question (e.g., "How frequently do you experience shortness of breath?") or request ("Please rate your ability to perform the following activities"). The response may be quantitative, as in visual analog scales, where the respondent is asked to mark an anchored line (typically 10 cm in length) according to a perception (e.g., amount of pain currently experienced), the length of the line between the anchor and the mark being

measured to determine the numeric value of the attribute.

Many standardized assessment items have response sets that are ordinal (e.g., absent, present; mild, moderate, or severe; rarely, occasionally, frequently, always). Responses to assessment items may also be nominal. For example, an item on the Behavioral Risk Factor Surveillance System states, "What type of arthritis did the doctor say you have?" Possible responses include "osteoarthritis/degenerative arthritis," "rheumatism," "rheumatoid arthritis," and "Lyme disease."[26]

The manner in which the value associated with an item is used in clinical practice or research varies with assessment instrument. In some instances, a single response is used (e.g., Visual Analog Scale—Quality of Life[25]). With assessments such as injury scores (e.g., Abbreviated Injury Scale,[31] Revised Trauma Score[42]), responses from multiple items are summed to provide a total score. Some assessments are associated with more complex scoring algorithms, and a score may be reported at the subscale or factor level rather than as a total score (e.g., Medical Outcomes Study Short Form 36[33]). Regardless, for purposes of data sharing and re-use, it is necessary to identify unambiguously the assessment item (attribute) and its associated value. In addition, attribute-value pairs for factor scores and total scores must be defined when appropriate.

LOINC/Clinical LOINC

In its initial state of development LOINC focused on a public set of codes and names for electronic reporting of laboratory test results.[15] The original aim of the LOINC Committee, when it began meeting in February 1994, was to produce a code system that would cover at least 98 percent of the tests performed in the average laboratory. One of the first tasks of the committee was to define a formal semantic structure for observation names that would distinguish tests that were clinically different and then to use this structure to create the database of clinically distinct names and related codes.

Influenced by the work of the International Union of Pure and Applied Chemistry (IUPAC),[43] the LOINC semantic structure is composed of six elements: analyte/component, kind of property measured or observed, time aspect of the measurement or observation, system/sample type that contains the analyte or component being observed, type of scale of the measurement or observation, and type of method

used to obtain the measurement. The semantic structure explicitly excludes items typically reported in separate attributes of a clinical laboratory test result message (e.g., priority of the testing, volume of sample, physical location of the testing).

With minor extensions of the original definitions, but not of the elements of the semantic structure, LOINC content continues to expand, especially in the area of direct patient measurements and clinical observations (e.g., blood pressure, symptoms).[16] LOINC has gained wide acceptance because of its content coverage, access (it is freely available on the World Wide Web and on CD-ROM), and most recently its relevance to the Health Insurance Portability and Accountability Act of 1996.[44]

Methods

Research Question

The research question addressed in the evaluation was "Can the sample of standardized assessment items be validly and reliably dissected into elements of the Clinical LOINC semantic structure?"

Sample of Standardized Assessment Items

The sample was composed of 1,088 items from 35 standardized assessments. As shown in Table 2, the sample was reflective of instruments with different primary purposes (e.g., risk screening, functional status measurement) and respondents (i.e., professional, patient, and caregiver).

Procedures

Extension of the Model Definitions

The elements of the Clinical LOINC semantic structure were examined by a domain expert (S.B.) for their potential utility in representing standardized assessments, and extensions of components definitions were proposed. The extensions were reviewed by four LOINC experts (J.J.C., R.H., S.M.H., and C. J. McDonald) and then presented (by S.B.) to the Clinical LOINC Committee for additional suggestions for extension.

The definitions of Clinical LOINC elements were extended in several areas. The definitions used for dissecting the standardized assessment items are shown in Table 3. The primary extension of the definitions related to System/Sample; the definition was extended to include aggregate units of analysis (e.g.,

Table 2 ■

Standardized Assessments Used in Evaluation

Assessment	Primary Focus	Rater	Items
Antepartum Questionnaire[51]	Depression screen	Patient	24
Apgar Score[38]	Physical status	Professional	5
Beck Depression Inventory[52]	Symptom status	Patient	21
Behavioral Risk Factor Surveillance System[26]	Risk screening	Patient	76
Braden Pressure Ulcer Risk Assessment[29]	Risk screening	Professional	6
Caregiver Burden Scale[36]	Caregiver burden	Caregiver	29
Caregiver Quality of Life Index—Cancer Scale[37]	Quality of life	Caregiver	35
Center for Epidemiologic Studies Depression Scale[30]	Depression screen	Patient	20
Denver II[53]	Developmental status	Professional	125
Douglas Ward Risk Calculator[54]	Pressure ulcer risk screen	Professional	6
Gosnell Pressure Ulcer Risk Assessment[55]	Pressure ulcer risk screen	Professional	5
Home Health Care Classification[56]	Nursing-sensitive	Professional	145
Illness Intrusiveness Ratings Scale[57]	Illness impact	Patient	13
Index of Activities of Daily Living[46]	Functional status	Professional	6
Karnofsky Performance Status Scale[35]	Functional status	Professional	1
Life Satisfaction Index A[40]	Life satisfaction	Patient	20
Living with HIV[41]	Quality of life	Patient	32
Marlowe-Crowne Social Desirability Scale[24]	Attitudes and traits	Patient	33
Medical Outcomes Study Short Form 36[33]	Health status	Patient	36
Morse Fall Score[39]	Falls risk screen	Professional	6
Nursing Outcomes Classification[34]	Nursing-sensitive	Professional	190
Norton Scale for Decubiti[58]	Decubitus risk screen	Professional	5
Omaha System[23]	Nursing-sensitive	Professional	123
Post-anesthesia Score[28]	Physical status	Professional	5
Post-anesthesia Score (modified for patients having anesthesia on an ambulatory basis)[59]	Physical status	Professional	10
Problematic Back Pain Screening Questionnaire[60]	Back pain risk screen	Patient	24
Quality Audit Marker[61]	Functional status	Professional	10
Revised Trauma Score[42]	Physical status	Professional	3
Seattle Angina Questionnaire[27]	Functional status	Patient	19
Self-Efficacy Scale[25]	Beliefs	Patient	10
Self-Reported Medication Taking Scale[62]	Adherence behavior	Patient	4
Sign and Symptom Checklist for Persons with HIV[32]	Symptom status	Patient	26
Trauma Score[42]	Physical status	Professional	7
Visual Analog Scale—General Health, Pain, Quality of Life[25]	Health status	Patient	1
Waterlow Pressure Sore Risk Assessment[63]	Risk screening	Professional	7

family). The refinements of other definitions were primarily the addition of examples for purposes of clarification.

Dissection of Assessment Items

Following training on a small sample of items to establish initial inter-coder reliability, each assessment item was entered into a Microsoft Access database and dissected into the elements of the Clinical LOINC semantic structure by one of four coders (S.B., C.M., G.C., R.K.). One coder (S.B.) was involved in the extension of the LOINC definitions; the other three were domain experts in the types of scales they were coding but were unfamiliar with LOINC. To assess inter-coder reliability of the dissections across all instruments in the evaluation, two coders dissected a sample of items.

The sample for inter-coder reliability was composed of at least one randomly selected item per assessment instrument. For the assessment instruments in the evaluation, a single assessment instrument reflected only a single type of Property, Timing, System/Sample, and Method for all items, but some had multiple scale types. Thus, when the instrument had a single scale type, one item was randomly selected. More than one item per instrument was placed into the inter-coder reliability sample when multiple types of scales (e.g., nominal and ordinal) were present in a single instrument. In that instance, one item of each scale type was randomly selected for inclusion, resulting in a sample of 40 items. Inter-coder reliability was calculated using percentage agreement.

Results

The coders were able to dissect all items from the 35 standardized assessment instruments into the elements of the Clinical LOINC semantic structure. Sample dissections are shown in Table 4. There was agreement between both coders on all elements in 25 of the 40 items composing the sample for inter-rater reliability. Percentage agreement for each element was as follows: component, 100 percent; property, 87.8 percent; timing, 82.9 percent; system/sample, 100 percent; scale, 92.6 percent; and method, 97.6 percent.

Discussion

The results of this evaluation support the adequacy of the Clinical LOINC semantic structure as a terminology model for standardized assessments. Using the revised definitions, the coders were able to dissect all standardized assessment items in the instruments

Table 3 ◾

Definitions for Dissecting Assessment Items

1. LOINC CODE (TO BE ASSIGNED; LEAVE BLANK)

2. COMPONENT—Attribute of a patient or an organ system within a patient; **name of the scale and item**

3. PROPERTY—kind of quantity related to a substance

   3.1. Finding—atomic clinical observation, not a summary statement as an impression; **can be professional or non-professional; can be of any scale type**

   3.2. Impression—a diagnostic statement, always an interpretation or abstraction of some other observations and almost always generated by a professional

4. TIMING—interval of time to which the measurement applies

   4.1. Point—single point in time

   4.2. Interval—more than a single point; specified in minutes, hours, days, weeks, months, etc.

5. SYSTEM (SAMPLE)—**individual or group who is the object of the measurement**

   5.1. Patient/client

   5.2. Family

   5.3. Caregiver

   5.4. Child

   5.5. Community

   5.6. Parent-child dyad

   5.7. Patient-caregiver dyad

6. SCALE—type of scaling used in the measurement of the item

   6.1. Quantitative—numeric value that relates to a continuous numeric scale **(e.g., visual analog scale)**

   6.2. Ordinal—reported either an integer, ratio, a real number or range; ordered categorical responses **(e.g., semantic differentials, likert-type scales; yes/no; positive/negative)**

   6.3. Nominal—nominal or categorical responses that do not have a natural ordering; typically have a coded value **(e.g., diagnosis)**

   6.4. Narrative—free text narrative

7. METHOD—**method of completing the measurement**

   7.1. Observed **(professional's rating)**

   7.2. Reported **(patient/client self-report)**

NOTE: Bold font indicates changes or extensions to LOINC.

composing the sample into the elements of Clinical LOINC. The initial results for inter-coder reliability are encouraging. The primary limitation of the evaluation was the number of standardized assessments in the sample, which may not be representative of the full scope of assessments of interest. In addition, the evaluation explicated some significant issues. Following discussion of these issues, the implications of the findings for the LOINC database and for the incorporation of standardized assessments into reference terminologies are summarized.

Table 4 ■

Dissection of Items from Standardized Assessment Scales Using the Clinical LOINC Specification

Standardized Assessment Scale/Item	Component	Property	Timing	System/Sample	Scale	Method
Omaha System/Bowel function:status	OMAHA.BOWEL FUNCTION:STATUS	IMPRESSION	POINT	PATIENT	ORDINAL	OBSERVED
Medical Outcomes Study - Short Form 36/How much bodily pain have you had during the past week?	SF-36.BODILY PAIN	FINDING	1 WK	PATIENT	ORDINAL	REPORTED
Nursing Outcomes Classification/Caregiver emotional health	NOC.CAREGIVER EMOTIONAL HEALTH	IMPRESSION	POINT	CAREGIVER	ORDINAL	OBSERVED
Apgar/Heart rate	APGAR.HEART RATE	FINDING	POINT	PATIENT	ORDINAL	OBSERVED
Home Health Care Classification/Knowledge deficit of medication regimen	HHCC.KNOWLEDGE DEFICIT MEDICATION REGIMEN	IMPRESSION	POINT	PATIENT	ORDINAL	OBSERVED
Braden Pressure Ulcer Risk Assessment/Friction and shear	BRADEN.FRICTION AND SHEAR	IMPRESSION	POINT	PATIENT	ORDINAL	OBSERVED
Karnofsky Performance Status Scale	KARNOFSKY.TOTAL	IMPRESSION	POINT	PATIENT	QUANTITATIVE	OBSERVED
Beck Depression Inventory/Sense of failure	BECK.SENSE OF FAILURE	FINDING	POINT	PATIENT	ORDINAL	REPORTED
Behavioral Risk Surveillance System/How long has it been since your last mammogram?	BFRSS.LAST MAMMOGRAM	FINDING	POINT	PATIENT	ORDINAL	REPORTED
Activities of Daily Living/Continence	ADL.CONTINENCE	IMPRESSION	POINT	PATIENT	ORDINAL	OBSERVED
Seattle Angina Questionnaire/Over the past 4 weeks, on average, how many times have you had chest pain, chest tightness, or angina?	SEATTLE ANGINA.FREQUENCY CHEST PAIN/ CHEST TIGHTNESS/ANGINA	FINDING	4 WK	PATIENT	ORDINAL	REPORTED

Issues

Some inconsistencies were the result of simple errors in which the coding of the items was not consistent with the definitions (e.g., scale type); however, the evaluation revealed several issues that require clarification and development of a consistent approach. These include temporal aspects, differentiation between finding and impression, representation of nominal responses as attributes or values, and algorithmic information.

Temporal Aspects

Temporal aspects were the most difficult to code consistently. For example, some time spans (e.g., during pregnancy, have you ever had "X"?) did not fall into clear temporal categories such as number of hours, weeks, or years. When items included an explicit time (e.g., within the last four weeks) or were observations by a professional at a specific point in time, the dissections were consistent across coders.

Findings versus Impressions

Differentiating between findings and impressions was problematic only for assessments in which the professional was the respondent, since by definition all assessments completed by the client were coded as findings. Lack of agreement between the two coders occurred primarily in the instances of physiologic variables measured by ordinal scales, such as respiratory rate and functional aspects like the ability to open the eyes or move the limbs.

Attributes versus Values

Consistent with the approach for Clinical LOINC, responses were treated as values associated with attributes (i.e., the standardized assessment items). In items for which only one response is possible, representation of a nominal response as a value is not problematic. However, when the responses made up a list (e.g., medical diagnoses as responses to questions on the Behavior Risk Factor Surveillance System[26]) and it was possible for an individual to select more than one response (i.e., multiple values like diagnoses), each response could be considered an attribute associated with the ordinal value of either *present* or *absent*.

In either instance, coded data elements from a terminology such as SNOMED have a potential role.[4] In the former the terminology codes would serve as the values, and in the latter as the attributes. The first approach is more consistent with the current rela-

tionship that exists between LOINC codes and SNOMED codes for laboratory names and results,[45] but the second reflects many electronic health record user interfaces. A consistent strategy is necessary to facilitate data sharing and re-use among computer-based systems.

Algorithmic Knowledge

Terminology models are not intended to handle algorithmic knowledge such as scoring instructions associated with many standardized assessments. For example, selected items on the Medical Outcomes Study Short Form 36[33] and the Sign and Symptom Checklist for Persons with HIV[32] require computation and are reported as subscale scores (i.e., values) rather than as individual item values. The terminology model should be capable of representing the attribute of subscale score and its value (e.g., physical functioning from the Short Form 36 score) and, similarly, the total assessment score, if appropriate.

Implications

The findings of this evaluation have implications for the expansion of the LOINC database and for the incorporation of standardized assessments into reference terminologies.

LOINC Database

Representation of standardized assessment items as fully specified Clinical LOINC names is necessary for incorporation into the LOINC database, and the results of this evaluation support the fit of the data to the semantic structure. In addition, beyond the information associated with the semantic structure, the LOINC database provides a structure for the inclusion of other useful information, such as the textual descriptions of the assessment items and the response sets.

Discussion with members of the Clinical LOINC Committee about the evaluation findings resulted in a recommendation for the inclusion of these types of standardized assessments into LOINC. However, for actual integration, several issues must be addressed. First, the dissections must be compared for consistency with other similar measures already included or slated for inclusion in the LOINC database. For example, in contrast to other items in LOINC, organ or body system was not used as a potential System/Sample in this evaluation, even though it might be appropriate for some assessments.

Also, after review of the evaluation results, the Committee suggested that the name of the standard-

ized assessment instrument be moved from Component to Method, (e.g., REPORTED.SF36, OBSERVED.NOC) for consistency with other items slated for incorporation into the next release of the LOINC database.

Second, permission to incorporate those assessments not in the public domain into LOINC database must be sought from the copyright holders. Assessments in the public domain and those copyrighted instruments that are frequently used in health care will be the highest priorities for inclusion.

Third, an approach to linking the items in an assessment is needed. Preliminary discussion supports the notion of a battery of assessment items analogous to a battery of laboratory tests; for example, the LOINC code 18729-4 represents a complete urinalysis battery that includes a set of observations, each with its own LOINC code (e.g., urine color, urine appearance). Likewise, an Index of Activities of Daily Living battery could be composed of the individual items related to a specific activity of daily living (e.g., dressing, bathing).[46]

Reference Terminologies

This evaluation supported the adequacy of the Clinical LOINC semantic structure as a terminology model for standardized assessments, but a terminology model (e.g., Clinical LOINC specification) and set of terms (in this evaluation, standardized assessment items) are only two components of a reference terminology. Also needed are other types of information (e.g., hierarchic knowledge), a language for expressing the instantiated terminology model in a computable form,[47] and software tools for processing the representations.[48] Representation language and software tools (which have been recently reviewed in detail elsewhere[12,49]) are not specific to the type of data being represented.

In contrast, the hierarchic knowledge used for classification varies with the information being classified. In SNOMED,[4] for example, hierarchic knowledge related to classification of diseases includes topography (e.g., organ system) and morphology. For standardized assessment items, other types of hierarchic knowledge may be appropriate. At least two aspects (component and dimension of the component measured) would be useful for hierarchic classification purposes. Taking medication and activities of daily living are examples of components that occurred in multiple standardized assessments. Examples of potential dimensions are beliefs, attitudes, knowledge, and behavior.

The notion of dimension was originally proposed (by S.B.) as an extension to the LOINC semantic structure; however, the Clinical LOINC Committee recommended alternatively that dimension be included as descriptive information in a field in the LOINC database. Classification by component and dimension would make it possible to retrieve a variety of assessments (e.g., both knowledge and behavior) related to taking medication or to multiple components (e.g., medications and bowel continence). Examples are shown in Figure 1.

Figure 1 Examples of using Clinical LOINC semantic structure and hierarchic information to represent standardized assessment items in a reference terminology.

Organ systems, care components, and functional health patterns are examples of other types of information that might be useful in the classification of standardized assessments. Minimally, a hierarchic classification for standardized assessments should support the linkage of a standardized assessment with the individual items it comprises. In some instances, the same item is part of a number of standardized assessments, which would support data re-use (e.g., the various versions of the Medical Outcomes Study questionnaire[50]).

This evaluation was an initial step toward the representation of standardized assessment items in a manner that facilitates data sharing and re-use. The ability to incorporate standardized assessments into an existing semantic structure such as LOINC is an important step toward simplifying their assimilation into computer-based systems for multiple purposes, such as documenting assessments in usable, shareable, and analyzable form; linking process and outcome data; and transmitting claims attachments.

The analysis supported the validity of the Clinical LOINC semantic structure for representing standardized assessment measures. Further clarification of the definitions, especially those related to time and property, is required to improve inter-coder reliability and to harmonize the representations with similar items already in the LOINC database. Potential areas for further research include evaluating the terminology model with additional standardized assessments and the critical examination of potential hierarchic structures (e.g., dimensions, care components) that would support the integration of standardized assessment measures into reference terminologies.

References ■

1. Rector AL. Thesauri and formal classifications: terminologies for people and machines. Methods Inf Med. 1998;37(4–5):501–9.
2. Cimino JJ. The concepts of language and the language of concepts. Methods Inf Med. 1998;37(4–5):311.
3. Chute CG, Cohn SP, Campbell JR for the ANSI Healthcare Informatics Standards Board Vocabulary Working Group and the Computer-based Patient Records Institute Working Group on Codes and Structures. A framework for comprehensive terminology systems in the United States: development guidelines, criteria for selection, and public policy implications. J Am Med Inform Assoc. 1998;5(6):503–10.
4. Spackman KA, Campbell KE, Cote RA. SNOMED RT: a reference terminology for health care. Proc AMIA Annu Fall Symp. 1997;640–4.
5. Humphreys BL, Lindberg DAB, Schoolman HM, Barnett GO. The Unified Medical Language System: an informatics research collaboration. J Am Med Inform Assoc. 1998;5(1):1–11.
6. Beeler GW. Taking HL7 to the next level. MD Comput. 1999;21–4.
7. Bakken S, Campbell KE, Cimino JJ, Huff SM, Hammond WE. Toward vocabulary domain specifications for Health Level 7–coded data elements. J Am Med Inform Assoc. 2000;7:333–42.
8. Rector AL, Bechhofer S, Goble CA, Horrocks I, Nowlan WA, Solomon WD. The GRAIL concept modeling language for medical terminology. Artif Intell Med. 1997;9:139–71.
9. Masarie FE, Miller RA, Bouhaddou O, Guise NB, Warner HR. An interlingua for electronic interchange of medical information: using frames to map between clinical vocabularies. Comput Biomed Res. 1991;24:379–400.
10. Price C, O'Neil M, Bentley TE, Brown PJB. Exploring the ontology of surgical procedures in the Read Thesaurus. Methods Inf Med. 1998;37(4–5):420–5.
11. Cimino J, Barnett GO. Automated translation between medical terminologies using semantic definitions. MD Comput. 1990;7(2):104–9.
12. Cimino JJ, McNamara TJ, Meredith T. Evaluation of a proposed method for representing drug terminology. Proc AMIA Annu Symp. 1999:47–51.
13. Bakken S, Cashen MS, Mendonca E, O'Brien A, Zieniewicz J. Representing nursing activities within a concept-based terminological system: evaluation of a type definition. J Am Med Inform Assoc. 2000;7(1):81–90.
14. Hardiker NR, Rector AL. Modeling nursing terminology using the GRAIL representation language. J Am Med Inform Assoc. 1998;5(1):120–8.
15. Forrey AW, McDonald CJ, DeMoor G, et al. Logical Observation Identifiers, Names, and Codes (LOINC) database: a public use set of codes and names for electronic reporting of clinical laboratory results. Clin Chem. 1996;42:81–90.
16. Huff SM, Rocha RA, McDonald CJ, et al. Development of the LOINC (Logical Observation Identifiers, Names, and Codes) vocabulary. J Am Med Inform Assoc. 1998;5(3):276–92.
17. Bakken S, Button P, Hardiker NR, Mead CN, Ozbolt JG, Warren JJ. On the path toward a reference terminology for nursing concepts. In: Chute C (ed). Proceedings of the International Medical Informatics Association Working Group 6 Conference on Natural Language and Medical Concept Representation; Dec 16–19, 1999; Phoenix, Arizona; pp 98–123.
18. Huff SM, Carter JS. A characterization of terminology models, clinical templates, message models, and other kinds of clinical information models. In: Chute C (ed). Proceedings of the International Medical Informatics Association Working Group 6 Conference on Natural Language and Medical Concept Representation; Dec 16–19, 1999; Phoenix, Arizona; pp 74–82.
19. National Convergent Medical Terminology. Concept Modeling Style Guide. Oakland, Calif: Kaiser Permanente; 1999.
20. Mays E, Weida R, Dionne R, et al. Scalable and expressive medical terminologies. Proc AMIA Annu Fall Symp. J Am Med Inform Assoc. 1996:259–63.

21. Sowa J. Conceptual Structures. Reading, Mass.: Addison Wesley, 1984.

22. Rossi Mori A, Consorti F, Galeazzi E. Standards to support development of terminological systems for healthcare telematics. Methods Inf Med. 1998;37(4–5):551–63.

23. Martin KS, Scheet NJ. The Omaha System: applications for community health nursing. Philadelphia, Pa.: Saunders, 1992.

24. Crowne DP, Marlowe D. A new scale of social desirability independent of psychopathology. J Consult Psychol. 1960;24(4):349–54.

25. Lorig K, Stewart A, Ritter P, Gonzalez V, Laurent D, Lynch J. Outcome measures for health education and other health care interventions. Thousand Oaks, Calif.: Sage, 1996.

26. Centers for Disease Control. Behavioral Risk Factor Surveillance System. Atlanta, Ga.: CDC, 1999.

27. Spertus JA, Winder JA, Dewhurst TA, et al. Development and evaluation of the Seattle Angina Questionnaire: a new functional status measure for coronary artery disease. J Am Coll Cardiol. 1995;25(2):333–41.

28. Aldrete JA, Kroulik D. A postanesthetic recovery score. Anesth Anal. 1970;49(6):924–34.

29. Braden BI, Bergstrom N. Clinical utility of the Braden scale for predicting pressure sores. Decubitus. 1989;2:44–51.

30. Radoff LS. The CES-D scale: A self-report depression scale for research in the general population. Appl Psychol Measure.1977;1:385–401.

31. Civil ID, Schwab CW. The Abbreviated Injury Scale, 1985 Revision: a condensed chart for clinical use. J Trauma. 1988;28(1):87–90.

32. Holzemer WL, Henry SB, Nokes KM, et al. Validation of the Sign and Symptom Checklist for Persons with HIV disease (SSC-HIV). J Adv Nurs. 1999;30(5):1041–9.

33. McHorney CA, Ware JE, Raczek AE. The MOS 36-item short-form health survey (SF-36), part II: psychometric and clinical tests of validity in measuring physical and mental health constructs. Med Care. 1993;31(3):247–63.

34. Johnson M, Maas M, Moorhead S (eds). Nursing Outcomes Classification (NOC). 2nd Ed. St. Louis: Mosby, 2000.

35. Karnofsky DA, Abelmann WH, Craver LF, Burchenal JH. The use of nitrogen mustards in the palliative treatment of carcinoma with particular reference to bronchogenic carcinoma. Cancer. 1948;1:634–56.

36. Zarit SH, Reever KE, Bach-Peterson J. Relatives of the impaired elderly: correlates of feelings of burden. Gerontologist. 1980;20(6):649–53.

37. Weitzner MA, Jacobsen PB, Wagner H Jr, Friedland J, Cox C. The Caregiver Quality of Life Index—Cancer (CQOLC) scale: development and validation of an instrument to measure quality of life of the family caregiver of patients with cancer. Qual Life Res.1999;8:55–63.

38. Apgar V. Proposal for new method of evaluation of newborn infants. Anesth Analg. 1953;32:260–7.

39. Morse J, Morse R, Tylko S. Development of a scale to identify the fall-prone patient. Can J Aging. 1989;8:366–77.

40. Neugarten BL, Havighurst RJ, Tobin SS. The measurement of life satisfaction. J Gerontol. 1961;16:141.

41. Holzemer WL, Spicer JG, Wilson HS, Kemppainen J, Coleman C. Validation of the quality of life scale: living with HIV. J Adv Nurs. 1998;28:622–30.

42. Champion HR, Sacco WJ, Copes WS, Gann DS, Gennarelli TA, Flannagan ME. A revision of the Trauma Score. J Trauma. 1989;29(5):623–9.

43. International Union of Pure and Applied Chemistry/ International Federation of Clinical Chemistry. List of Particular Properties in Clinical Laboratory Sciences. Copenhagen: IUPAC/IFCC, 1994.

44. Braithwaite W. HIPAA and the Administration Simplification Law. MD Comput. 1999;16(5):13–6.

45. Spackman K. Relationship between SNOMED and LOINC. SNOMED Editorial Board Minutes, Feb 19, 2000.

46. Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW. Studies of illness in the aged: the Index of ADL, a standardized measure of biological and psychosocial function. JAMA. 1963;185:914–9.

47. Rector AL, Nowlan WA. The GALEN Representation and Integration Language (GRAIL) Kernel, version 1. In: The GALEN Consortium for the EC. Manchester, UK: University of Manchester, 1993.

48. Campbell KE, Cohn SP, Chute CG, Shortliffe EH, Rennels G. Scalable methodologies for distributed development of logic-based convergent medical terminology. Med Inf Med. 1998;37(4–5):426–39.

49. Chute CG. Terminology services as software components: An architecture and preliminary efforts. In: Chute C (ed). Proceedings of the International Medical Informatics Association Working Group 6 Conference on Natural Language and Medical Concept Representation; Dec 16–19, 1999; Phoenix, Arizona; pp 62–9.

50. Stewart AL, Ware JE Jr. Measuring functioning and well-being: The Medical Outcomes Study approach. Durham, NC: Duke University Press, 1992.

51. Posner NA, Unterman RR, Williams KN, Williams GH. Screening for postpartum depression: an antepartum questionnaire. J Reprod Med. 1997;42:207–15.

52. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. Arch Gen Psychol. 1961;4:561–71.

53. Frankenburg WK, Dodds J, Archer P, Shapiro H, Bresnick B. The Denver II: a major revision and restandardization of the Denver Developmental Screening Test. Pediatrics. 1992;89(1):91–7.

54. Pritchard V. Calculating the risk. Nurs Times. 1986;82:59–61.

55. Gosnell DJ. An assessment tool to identify pressure sores. Nurs Res. 1973;22:55–9.

56. Saba VK, Zuckerman AE. A new home health classification method. Caring Mag. 1992;11(9):27–34.

57. Devins GM, Mandin H, Hons RB, et al. Illness intrusiveness and quality of life in end-stage renal disease: comparison and stability across treatment modalities. Health Psychol. 1990;9(2):117–42.

58. Norton D, McLaren R, Exton-Smith. An investigation of geriatric nursing problems in the hospital. London, UK: National Corporation for the Care of Old People (now Centre for Policy on Aging), 1962.

59. Aldrete JA. The Post-anesthesia Recovery Score revisited. J Clin Anesth.1995;7:89–91.

60. Linton SJ, Hallden K. Can we screen for problematic back pain? A screening questionnaire for predicting outcome in acute and subacute back pain. Clin J Pain. 1998;14(3):209–15.

61. Holzemer WL, Henry SB, Stewart A, Janson-Bjerklie S. The HIV Quality Audit Marker (HIV-QAM): an outcome measure for hospitalized AIDS patients. Qual Life Res. 1993;2:99–107.

62. Morisky DE, Green LW, Levine DM. Concurrent and predictive validity of a self-reported measure of medication adherence. Med Care. 1986;24(1):67–74.

63. Waterlow J. A risk assessment card. Nurs Res. 1986;81:49–55.