Nursing Informatics
U. Gerdin et al. (Eds.)
IOS Press 1997

Evaluating Standardized Coding and Classification Systems for Clinical Practice: A Critical Review of the Nursing Literature in the United States

S.B. Henry[a] and C.N. Mead[a,b]
[a]*Department of Community Health Systems, School of Nursing, University of California, San Francisco, CA,* [b]*CarecCentric Solutions, Inc., Atlanta, GA*

Developers of healthcare information systems are challenged by the difficulty of meeting the simultaneous goals of 1) capturing and electronically representing the broad array of data related to healthcare with sufficient expressibility to provide adequate documentation of the patient encounter, and 2) utilizing standardized coding and classification systems to facilitate linkages among computer-based systems. The evaluation studies to date have primarily focused on matching actual clinical data with terms in the recognized classification systems. These studies have provided evidence that the classification systems are relatively domain complete for categorizing patient problems and nursing interventions. Although several of the published criteria for evaluation relate to structure, it is noteworthy that this has not yet been a major focus of study in nursing. There are several areas of critical need that must be addressed. First, additional work is needed to develop and refine a standardized set of atomic-level terms relevant to nursing, including those for assessments, problems, and activities. Second, knowledge representations must be developed to support the building of complex concepts from atomic-level data.

Background

Developers of healthcare information systems are challenged by the difficulty of meeting the simultaneous goals of 1) capturing and electronically representing the broad array of data related to healthcare with sufficient expressibility to provide adequate documentation of the patient encounter, and 2) utilizing standardized coding and classification systems to facilitate linkages to knowledge-based resources such as bibliographic databases, clinical practice guidelines, therapeutic protocols, and decision support systems, as well as for abstraction to clinical data repositories. Significant pioneering research has focused on the development of coding and classification systems for nursing (e.g. Nursing Interventions Classification,[1, 2] International Classification of Nursing Practice);[3] the creation of architectures, e.g., the Unified Medical Language System (UMLS),[4] to link standardized coding and classification systems; and the testing of standardized coding and classification systems with clinical data.[5-9]

The purposes of this paper are to review criteria for the evaluation of standardized coding and classification systems, to critically examine the published evaluation studies related to standardized coding and classifications applicable to nursing, and to suggest future directions for research and development.

Evaluation criteria

While a "gold standard" has not been identified, a number of authors have proposed evaluation criteria for a standardized coding and classification system designed to support clinical practice. The authors and the resulting criteria represent a variety of perspectives.

Cimino identified nine criteria for a multi-purpose controlled vocabulary for clinical information systems.[10] These criteria were aimed at increasing the sensitivity and specificity of information retrieval queries. Clark and Lang described criteria from the perspective of the development of the International Classification of Nursing Practice (ICNP).[3] McCloskey and Bulechek generated criteria specifically for the evaluation of the taxonomic structure of the Nursing Intervention Classification (NIC).[2] Several authors have focused on criteria related to the clinical expressiveness of classification.[6, 8] For recognition by the American Nurses Association Committee on Databases to Support Nursing Practice, a system must meet criteria for clinical usefulness, reliability and validity, and processes for revision and extension of the classification system.[11]

The evaluation criteria reported in the literature fall into five broad areas: 1) domain completeness, 2) conceptual clarity and coherence, and 3) data structures and relationships among terms, 4) clinical concept capture, and 5) utility. The criteria are defined in the following section.

Domain completeness. The classification system must include all the terms necessary to describe the domain[2, 10] In addition, from the perspective of the ICNP, domain completeness means that the classification system is broad enough to serve the multiple purposes required by different countries. McCloskey and Bulechek apply the notion of completeness at the class level in the NIC taxonomy, that is, the intervention class will include all the interventions belonging to that class.[2]

Conceptual clarity and coherence. The classification system should be consistent with a clearly defined conceptual framework, but not dependent upon a particular theory or model.[2, 3] Clark and Lang propose that the conceptual framework should be reflective of the common value system of nursing across the world as expressed in the International Council (ICN) Code for Nurses.[3] Other criteria related to conceptual clarity and coherence include: 1) clear, understandable definitions;[11] 2) only one way to express each concept (non-redundancy);[10] 3) terms should refer to only one concept (unambiguous);[10] and 4) all terms within a category are members of the same class (homogeneity).[2]

Data structures and relationships among terms. The relationships among terms should be explicit.[10] For instance, in the Nursing Interventions Classification,(1) Bowel Incontinence Care IS-A Elimination Management intervention. IS-A is a statement of explicit relationship; other types of explicit relationships among terms include EQUIVALENT-TO, PART-OF, and ASSOCIATED-WITH. Another useful perspective on structure is that of Ingenerf who has explicated four types of taxonomic vocabularies or standardized coding and classification systems for health care based on the underlying structure and related knowledge representation formalism.[12] *Thesauri* are defined as lexical vocabularies containing definitions and cross references (e.g. UMLS Metathesaurus). *Classification systems* are vocabularies that can be represented as hierarchies or decision trees, and that have as a main emphasis the disjunctive and exhaustive classification of terms. *Nomenclatures* are combinatorial taxonomic vocabularies containing more complex polyhierarchies or axes. Terms within a nomenclature may be combined into complex concepts using semantic grammars; however, explicit rules for canonical (disambiguated) representation of terms is lacking. *Formal terminologies*, such as the GRAIL representation language developed in conjunction with the GALEN project, are systems that are based on concepts, rather than on terms and that include explicit rules for sensible composition of primitive concepts into complex concepts.[13] The concepts are represented using knowledge formalisms such as description logic or conceptual graphs.

Clinical concept capture. Classification systems for clinical practice should be clinically expressive, that is, include the types of natural language terms used to describe patient problems and health care interventions in the medical record.[6, 8, 14] To do this, the classification system should include modifiers such as those related to time and severity.

Utility. Three criteria related to utility of a classification system have been described in the context of the ICNP.[3] First, the system is "...simple enough to be seen by the ordinary practitioner of nursing as a meaningful description of practice and a useful means of structuring practice" (p. 111). Second, the classification system is complementary with the family of disease and health-related classification systems. Third, the classification system is based on a central core which can be updated through a continual process of development and refinement. Others have also noted the significance of the last criteria.[11] Ozbolt recently emphasized the importance of having a standardized set of terms that can capture the varied and evolving clinical practice in addition to formal vocabularies as exemplified in nursing by NIC.[14]

Evaluation strategies in nursing studies in the US

A variety of strategies have been utilized to evaluate the standardized coding and classification systems for use by nursing in the US. Excluded from the studies in this review are those done by system developers for the purposes of creating, validating, and refining the systems and studies aimed a validating a single entity within a system, e.g., validating the defining characteristics for a particular nursing diagnosis. The studies are listed individually in Table 1. n the following section the studies are discussed from a chronological perspective which relates to the type of evaluation strategies utilized. In the early 1990's, Griffith and Robinson conducted two provider surveys focused on the degree to which Physician's Current Procedural Terminology (CPT) coded services were provided by nurses in a variety of nursing specialties.[15, 16] These studies provided evidence that nurses do perform a limited number of interventions that can be represented using the CPT codes, however, the determination of whether or not the CPT codes can represent the scope of nursing was not an intent of the study. While Griffith and Robinson identified the potential overlapping functions of physicians and nurses in some areas (as identified by CPT-coded procedures), Zielstorff et al.'s study highlighted the differences among systems in the UMLS and the nursing classification systems that were not at the time included in the UMLS.[22] Subsequently, the nursing classification systems that have been recognized by the ANA Steering Committee on Databases to Support Clinical Practice have been added to the UMLS.

The majority of recent evaluation studies have tested existing classification systems with clinical data to examine the extent to which the systems capture clinical concepts and are domain complete.[5, 17, 18, 20, 21] The systems examined were Systematized Nomenclature of Medicine (SNOMED), North American Nursing Diagnosis Association (NANDA) Taxonomy 1, NIC, CPT, and Home Health Care Classification (HHCC). The earliest study used a semi-automated lexical matching approach to determine if exact matches could be found between terms used by nurses to describe patient problems in the patient record and standardized terms in SNOMED which includes NANDA terms in its functional axis. The study found that 69% of the terms could be matched. Of particular note was the fact that terms in SNOMED other than NANDA accounted for 35% of the matches; that is, nurses frequently used symptoms, signs, and medical diagnoses to describe patient problems in their documentation.

Recognizing that the standardized systems developed for nursing to date were aimed at classification or aggregation of atomic-level data into categories, four studies focused on the domain completeness of existing nursing classification systems. Henry et al. compared the

ability of NIC and CPT codes for categorizing nursing activity terms from three acute care hospitals and reported the superiority of NIC to CPT in representing the domain of nursing activities.[17] Holzemer and associates used a related data set to examine the utility of the HHCC for categorizing patient problems and nursing interventions in the hosptial environment.[18] The findings were that the HHCC was useful beyond the home health care setting for which it was designed. Ozbolt noted that the care components of the HHCC were a useful organizing framework, but that a standardized set of more atomic-level terms was needed.[20] Parlocha tested the utility of the HHCC to abstract chart data related to psychiatric home care with the intent of developing a critical path for Major Depressive Disorder.[21] While the HHCC problem scheme worked well for this data set, subcategories for psychiatric nursing interventions were added to adequately capture that area of nursing activities.

Congruent with the shift of research focus in the larger health care vocabulary arena to an examination of data structures and representations in addition to content coverage, Henry and Mead recently critically analyzed three nursing intervention schema related to two sets of criteria described earlier in this paper; the typology of taxonomic vocabularies published by Ingenerf[12] and the Cimino[10] criteria for a multi-purpose controlled vocabulary.[23] Their analysis demonstrated that the recognized systems in the US have classification as their primary purpose and that there are no nomenclatures or formal terminologies for nursing that meet the definitions proposed by Ingenerf.

The evaluation studies to date have primarily focused on matching actual clinical data with terms in the recognized classification systems. These studies have validated that the classification systems are relatively domain complete for categorizing patient problems and nursing interventions. Only one study had a nomenclature, rather than a classification system, as its focus and had exact lexical matches rather than categorization as an aim.[5] The study finding that SNOMED terms other than NANDA diagnoses were exact matches for terms used by nurses to describe patient problems in the patient record, suggests that nomenclatures, as well as classification systems, have a role in representing terms for computer-based systems. Studies are needed that focus upon the utility of the existing standardized coding and classification systems for purposes other than classification, e.g., how useful are the systems for representing nursing terms in a multi-purpose clinical information system and what are the implementation barriers related to the existing standardized coding and classification systems? Additionally, although several of the criteria for evaluation relate to structure, it is noteworthy that this has not yet been a major focus of study in nursing as compared to medicine.[8, 24]

Directions for future development and study

In addition to the ongoing significant work related to the recognized nursing classification systems in the US and elsewhere, several research teams are working on the gaps identified in this review of evaluation studies. For example, Ozbolt and associates have developed and continue to refine a more atomic-level set of standardized problem and activity terms for the acute care environment.[14, 19, 20] Grobe and associates are utilizing complex natural language processing techniques to examine both content and structure of nursing documentation as an extension of the work on the Nursing Intervention Lexicon and Taxonomy (NILT).[25, 26] In an effort to compare the relative domain completeness of five existing classification systems (NANDA, NIC, HHCC, Omaha, and SNOMED) across three care settings (acute care, skilled nursing facility, and home care), Henry and associates are testing each system with the same data sets. Another aspect of the same research project is aimed at comparing the frequencies of matches along a 9-point scale[6] to essentially provide a fingerprint of each system related to its proportion of abstract as compared to atomic-level terms.

The large scale efforts of the National Library of Medicine related to identifying the meta-set of terms needed for multiple uses in health care and supporting the linkage among terms through the structure of the UMLS has the potential to significantly accelerate the developments in coding and classification systems.[27] However, there are several areas of critical need that must also be addressed. First, additional work is needed to identify and refine a standardized set of atomic-level terms relevant to nursing, including those for assessments, problems, and activities. The ongoing work related to the ICNP and TELENURSE has provided an excellent foundation to be built upon. Second, knowledge formalisms must be developed to build more complex concepts from primitive concepts (atomic-level data). Several investigators have reported the applicability of conceptual graphs as a type of knowledge representation for medicine,[8, 28] but little is known about similarities or differences in structure between nursing and medical knowledge or whether the same strategies for representing knowledge are appropriate or feasible across disciplines. The significant work in Europe within the GALEN project on this topic[29] must be critically examined to determine the scope of its applicability. During the last decade excellent progress has been made in the development and validation of classification systems for nursing. The rapidly evolving nature of computer-based system implementation in health care has highlighted the need for nomenclatures and formal terminologies in addition to classification systems to support nursing practice with a variety of systems including those for decision support. The profession of nursing must address these areas of critical need while continuing the refinement of the existing classifications.

References

1. Iowa Intervention Project. The NIC taxonomy structure. *Image* 1993;25:187-192.
2. McCloskey JC, Bulechek GM. Nursing Interventions Classification. (2nd ed.) St. Louis: C. V. Mosby, 1996.
3. Clark J, Lang NM. Nursing's next advance: An international classification for nursing practice. *Int Nurs Rev* 1992;39:109-112.
4. Humphreys B, Lindberg DAB. Building the unified medical language system. In: Kingsland L, ed. *Symposium on Computer Applications in Medical Care:* IEEE Computer Society Press, 1989:475-480.
5. Henry SB, Holzemer WL, Reilly CA, Campbell KE. Terms used by nurses to describe patient problems: Can SNOMED III represent nursing concepts in the patient record? *J Am Med Informatics Assoc* 1994;1:61-74.
6. Chute CG, Atken GE, Ihrke DM. An empirical evaluation of concept capture by clinical classifications.. In: Lun KC, Degoulet P, Piemme TE, Riehoff O, eds. *MedInfo92*. Geneva, Switzerland: North-Holland, 1992
7. Chute CG, Cohn SP, E. CK, Oliver DE, Campbell JR. The content coverage of clinical classifications. *J Am Med Informatics Assoc* 1996;3(3):224-233.
8. Campbell KE, Musen M. Representation of clinical data using SNOMED III and conceptual graphs. In: Frisse ME, ed. *Symposium on Computer Applications in Medical Care*. Baltimore, MD: McGraw-Hill, Inc., 1992.
9. Campbell JR, Kallhenberg GA, Sherrick RC. The clinical utility of META: An analysis for hypertension.. In: Frisse M, ed. *Proceedings of the 16th Annual Symposium of Computer Applications in Medical Care*. New York: McGraw-Hill, 1993
10. Cimino JJ, Hripsak G, Johnson SB, Clayton PD. Designing an introspective, multi-purpose, controlled medical vocabulary. In: Kingsland III LC, ed. *Symposium on Computer Applications in Medical Care*. Washington, D.C.: IEEE Computer Society Press, 1989.
11. McCormick K, Lang N, Zielstorff R, Milholland DK, Saba V, Jacox A. Toward standard classification schemes for nursing language: Recommendations of the American Nurses Association Steering Committee on Databases to Support Nursing Practice. *J Am Med Informatics Assoc* 1994;1:421-427.
12. Ingenerf J. Taxonomic vocabularies in medicine: the intention of usage determines different established structures. In: Greenes RA, Peterson HE, Protti DJ, eds. *MedInfo95*. Vancouver, British Columbia: HealthCare Computing & Communications, Canada, Inc., 1995.
13. Rector AL, Nowlon WA, Kay S, Horan B, Wilson A. Foundations of an electronic medical record. Methods *Information Med* 1991;30:179-186.

14. Ozbolt JG. From minimum data to maximum impact: using clinical data to strengthen patient care. *Adv Prac Nurs Q* 1996;1(4):62-69.
15. Griffith HM, Robinson KR. Survey of the degree to which critical care nurses are performing Current Procedural Terminology-coded services. *Am J Crit Care* 1992;1:91-98.
16. Griffith HM, Robinson KR. Current Procedural Terminology (CPT) coded services provided by nurse specialists. *Image* 1993;25:178-186.
17. Henry SB, Holzemer WL, Randell C, Hsieh S-F, Miller TJ. Comparison of Nursing Interventions Classification and Current Procedural Terminology codes for categorizing nursing activities. *Image*, In press.
18. Holzemer WL, Henry SB, Dawson C, Sousa K, Bain C, Hsieh S-F. An evaluation of the utility of the Home Health Care Classification for categorizing patient problems and nursing interventions from the hospital setting. NI97. Stockholm, Sweden, 1997.
19. Ozbolt J, Graves J. Clinical nursing informatics. Developing tools for knowledge workers. *Nurs Clin N Am* 1993;28(2):407-425.
20. Ozbolt J, Fruchtnicht JN, Hayden JR. Toward data standards for clinical nursing information. *J Am Med Informatics Assoc* 1994;1:175-185.
21. Parlocha PK. Examination of a Critical Path for Psychiatric Home Care Patients with a Diagnosis of Major Depressive Disorder [PhD]. San Francisco, CA: University of California, San Francisco, 1995.
22. Zielstorff RD, Cimino C, Barnett GO, Hassan L, Blewett DR. Representation of nursing terminology in the UMLS Metathesaurus: a pilot study. In: Frisse M, ed. Proceedings of the 16th Annual Symposium on Computer Applications in Medical Care. New York: McGraw-Hill, 1993:392-396.
23. Henry SB, Mead CN. Standardized nursing classification systems: necessary, but not sufficient for representing what nurses do. In: Cimino J, ed. Fall Symposium of the American Medical Informatics Association. Washington, DC: Hanley & Belfus, 1996.
24. Evans D, Chute C, Cimino JJ, et al. CANON: Toward a medical concept representation language. *Proceedings of the American Medical Informatics Association Spring Congress*. Bethesda, MD: American Medical Informatics Association, 1993.
25. Grobe SJ. Nursing intervention lexicon and taxonomy: Methodological aspects. In: Hovenga EJS, Hannah KJ, McCormick KA, Ronald JS, eds. Nursing Informatics 91: *Proceedings of the Fourth International Conference on Nursing Use of Computers and Information Science.* Berlin: Springer-Verlag, 1991
26. Grobe SJ. Personal communication, June 1996.
27. Humphreys B. Vocabularies for computer-based patient records: identifying candidates for large scale testing. Invitational Meeting at the National Library of Medicine. December, 1994:
28. Moorman PW, van Ginneken AM, van der Lei J, van Bemmel JH. A model for structured data entry based on explicit descriptional knowledge. *Methods Information Med* 1994;33:454-463.
29. Hardiker NR, Kirby J. Overcoming terminological barriers in nursing. In Brendies J et al, eds. *Proceedings of the 13th International Congress of the European Federation for Medical Informatics*, Copenhagen, 1996:227-231.

Acknowledgments

The preparation of this paper was supported by NIH-NR03874, Representing Nursing Concepts for Computer-Based Systems, S B Henry, Principal Investigator. The first author thanks the participants in the 1994 European Summer School of Nursing Informatics for their thoughtful discussions regarding evaluation criteria for nursing coding and classification systems during the Summer School.