Validity and Reliability of Standard Terms and Codes for Patient Care Data

Judy G. Ozbolt, PhD, RN, University of Virginia Health Sciences Center
Marguerite Russo, RN, University of Virginia Medical Center
Melissa P. Stultz, RN, Martha Jefferson Hospital
Charlottesville, Virginia

A set of standard terms and codes for patient care data was derived from care planning and documentation materials submitted by 9 hospitals. The set contained 329 terms for Patient Problems, 308 terms for Patient Outcomes, and 1261 terms for Patient Care Actions. Six of the hospitals participated in a test of validity and reliability of the standard terms and codes. Manual audits were conducted on 465 patient records from two services in each hospital. All auditors achieved acceptable accuracy in coding. The auditors identified 18,995 items in the patient records as representing statements of Patient Problems, Patient Outcomes, or Patient Care Actions. The standard terms and codes matched 99.1% of these items. Thus, for the services audited, the standard terms and codes provided a valid representation of the Patient Problems, Patient Outcomes, and Patient Care Actions in the patient records.

INTRODUCTION

In the competitive health care market, hospitals are squeezed between cutting costs and maintaining or improving quality. Consequently, they seek to identify and provide "best value" services: those that yield good patient outcomes at reasonable cost. Discovering the "relationships of investments in health care resources and procedures to health outcomes" requires studying data from large populations of patients across numerous institutions.[1] Aggregating such data for analysis presupposes standard terminology and coding systems. The state of the art of standard terms and codes for patient care data, however, leaves much to be done.

Standard Vocabularies and Clinical Data

Henry and associates described classification schemes for nursing data and for general health care data.[2] Summarizing nine evaluation studies, they concluded that "while each vocabulary served the purpose for which it was designed, no one existing vocabulary or unifying scheme was adequate to represent the broad array of clinical data in the patient record." Their own evaluation of the Systematized Nomenclature of Medicine (SNOMED III)[3] showed that 69% of the nursing terms used to represent patient problems or conditions were represented with one or more SNOMED III terms.

Similarly, we have found efforts to develop standard vocabularies for nursing commendable but not well suited to capturing the full range of data needed for outcome studies in acute care settings. Some, like the Home Health Care Classification System,[4] the Nursing Intervention Lexicon and Taxonomy,[5,6,7] and the Omaha Community Health System,[8,9] arise from non-acute settings and lack the specific vocabulary of acute care. The North American Nursing Diagnosis Association (NANDA) Taxonomy I,[10] although widely used, does not cover all fields of specialty practice. And the Nursing Interventions Classification (NIC)[11] is not always sufficiently fine-grained to capture differences in practice that may affect outcomes. No up-to-date vocabularies exist for patient outcomes.

Prior Research

The need to capture data from nursing and other non-physician aspects of patient care for studies of outcomes and cost-effectiveness led to the compilation of terms from care planning and documentation materials used in hospitals.[12] Rather than a "top-down" effort to impose a standard vocabulary on hospitals, this approach sought to represent the clinical vernacular already in use. The hypothesis was that nurses and other health professionals in different hospitals and geographic regions shared a common vocabulary. Initially the categories in which terms were collected were identified as *Nursing Diagnoses, Expected Patient Outcomes,* and *Nursing Interventions,* reflecting three of the Nursing Care Elements of the Nursing Minimum Data Set.[13] With the recognition that hospitals were increasingly using integrated, interdisciplinary care planning and documentation, more inclusive labels were adopted: *Patient Problems, Patient Outcomes, and Patient Care Actions.*

To provide a basis for comparison across hospitals, terms were collected from member hospitals of the University Hospital Consortium (UHC), a voluntary association of some 68 academic health science centers throughout the United States. Likewise, codes were developed consistently with the standards of UHC's Clinical Information Network (CIN), a patient-level database that contains the data of the Uniform Hospital Discharge Data Set. This includes patient demographics, medical diagnoses and procedures, cost of care, and length of stay, among other data. Merging patient care data with the data already in the CIN would make it possible to compare costs and effectiveness of care practices across hospitals to identify the "best value" services.

Aims of This Paper

Before recommending the adoption of these terms and codes into a national database such as the CIN, however, it was necessary to assess their validity and reliability. This paper will describe the process of developing the standard terms and codes and the methods of testing their reliability and validity. It will present results and discuss the relationship of this work to other national efforts to develop standard vocabularies for patient care information.

METHODS AND RESULTS

Development of Standard Terms and Codes

The standard terms and codes for Patient Problems, Patient Outcomes, and Patient Care Actions have been developed as appendices to a data dictionary. The data dictionary provides a definition of each data element and specifies how it is to be coded. The specific terms and codes that represent instances of the data element follow in the appendix.

The first set of standard terms and codes, the development of which has been described elsewhere,[12] contained terms from care planning and documentation materials developed on 5 patient care units at the University of Virginia Hospital and 2 patient care units at Thomas Jefferson University Hospital. In 1994, the set of terms was expanded to include terms from all patient care units at the first two hospitals. In addition, care planning and documentation materials were sampled from 7 other UHC member hospitals. Because of limited resources, it was not possible to review the materials from every unit of every hospital. Instead, each hospital provided materials from 8 to 12 patient care units, selected in such a way that there were at least two examples of each type of unit, and the full

range of units was included. The hospitals that contributed care planning and documentation materials were: Indiana University Medical Center, the Medical University of South Carolina Medical Center, Penn State's Milton S. Hershey Medical Center, Thomas Jefferson University Hospital, the University of Arizona Medical Center, the University of Missouri Hospital and Clinics, the University of North Carolina Hospitals, the University of Virginia Medical Center, and the University of Wisconsin Hospital and Clinics. The expanded version of terms and codes included 329 Patient Problems, 308 Patient Outcomes, and 1261 Patient Care Actions.

Testing Validity and Reliability

Sample. Six member hospitals of UHC volunteered to conduct audits of patient records at their own expense. Five hospitals audited records in the Orthopedics and Thoracic / Cardiovascular Surgery Services. These hospitals were: Indiana University Medical Center, Penn State's Milton S. Hershey Medical Center, Thomas Jefferson University Hospital, the University of Arizona Medical Center, and the University of Virginia Medical Center. One hospital (the University of Missouri Hospital and Clinics) audited records of coronary artery bypass patients and of patients in the Adolescent unit.

Each hospital sampled patient records in the selected services for one month. As of March 3, 1995, a total of 465 patient records had been audited.

Table 1 shows the numbers of instances of the key data elements identified in the patient records and recorded on the audit forms:

Table 1. Instances of Data Elements

Problems	4,880
Outcomes	5,237
Actions	8,878
TOTAL	18,995

Reliability of auditors. Reliability was defined in terms of the accuracy of the auditor in matching a statement in the patient record to a term in the code book and recording the corresponding numerical code on the audit sheet. To evaluate this ability, a 60-item test was constructed containing statements from patient care planning and documentation materials from all the participating hospitals. Some of the test items were exact matches to terms in the

code books; others represented slight variations in wording (such as *Discontinue I.V.* vs. *Remove I.V.*), as might be found in the participating hospitals. The scores of the 11 auditors ranged from 76.7% to 95.8%, with a mean of 86.3%.

Validity of the set of terms: comprehensiveness. An important dimension of validity is comprehensiveness. Does the set of standard terms match the universe of terms actually used to record the data elements in patient records? Auditors in the participating hospitals were able to assign codes to 17,278 items (91%). Auditors had been instructed that if they were unable to match an item in the patient record to a standard term, they should record the item verbatim on the audit sheet. Of the 1717 items submitted verbatim (not coded) by the auditors, the University of Virginia investigators matched 1538 (8.1% of total items) to existing terms. For 174 items submitted verbatim (0.9% of the total items), the investigators determined that it would be necessary to modify an existing term to make it more inclusive. In only 5 instances out of 18,995 items recorded on the audit forms was it necessary to create a new term to accommodate the item.

In summary, standard terms existed in the code books for 99.1% of the Patient Problems, Patient Care Actions, and Patient Outcomes identified in the patient records. Nearly all the remaining items from the patient records could be accommodated by minor modifications to an existing term.

Validity: diversity of terms selected. The set of terms is intended to cover all clinical services. In this study, most records were sampled from Orthopedics or Thoracic / Cardiovascular Surgery. If the total set of terms is adequate and relevant, one would expect a diverse selection of terms associated with the types of patients sampled, with many other terms not being selected because of their association with other patient types.

As of this writing, data entry for the total sample is not complete. Data were analyzed for a subsample of 80 records distributed approximately equally among 4 hospitals and the two services, Orthopedics and Thoracic / Cardiovascular Surgery, that comprised most of the total sample. Table 2 compares the number of instances of each data element in the subsample, the number of different terms those instances represented, and the number of different terms for the data element in the code book.

Thus, in this subsample, the auditors selected about a third of the possible terms for Patient Problems and Patient Outcomes, and about a fourth of the possible terms for Patient Care Actions. This suggests that the auditors found many relevant terms for patients in the two services, and that there were many other terms not relevant to patients in these two services that would presumably be relevant to patients in other services.

DISCUSSION

The results of data analysis so far are encouraging. The reliability of auditors is acceptable but could probably be improved by making the code books easier to use, as by adding indexes. This would also help to remedy the problem of auditors' failing to find an existing term to match an item in the patient record, as happened for 8.1% of the total items. Validity was excellent, in terms of both comprehensiveness and diversity, for the types of patients sampled. Further work is needed to test validity across the full range of clinical services.

In the long run, the issues of auditor reliability and ease of use may become moot. The set of standard terms and codes would be ideally employed by installing them in computer-based systems for

Table 2: Diversity of Terms Selected

Domain	No. of Items in Subsample	No. of Different Items Found	No. of Different Items in Code book
Problems	660	95	329
Outcomes	660	105	308
Actions	1,519	312	1,261

patient care planning and documentation and attaching a program to copy data items that are recorded in the course of care to a relational database for subsequent analysis. Collecting each patient's data in real time and copying it electronically to the database would avoid the expense and potential transcription inaccuracies of manual audits.

The analyses reported here have not addressed the question of the usefulness of the data for studies of cost and effectiveness. In collaboration with the University Hospital Consortium, a prototype database is being constructed that merges data collected in this audit with the same patients' data in the Clinical Information Network. The database will be used to generate descriptive reports and comparisons of care practices and costs, and the utility of these reports for management decisions and discovery of clinical knowledge will be assessed.

How does this research relate to other national efforts to develop standard vocabularies? The usefulness of any vocabulary is determined by the degree to which it serves the function for which it was designed. This set of terms aims to standardize the clinical vernacular to facilitate capturing data from acute care settings for studies of cost and effectiveness. Results to date suggest that it may serve that purpose well. Other vocabularies have been developed for other purposes and settings. Such diversity is probably necessary to meet diverse needs for vocabularies. It is to be hoped that those vocabularies that meet criteria of rigor and utility, such as those established by the American Nurses Association[14] may become part of a unified (as opposed to uniform) language system.

References

1. Board of Directors of the American Medical Informatics Association. Standards for medical identifiers, codes, and messages needed to create an efficient computer-stored medical record. J Am Med Informatics Assoc. 1994; 1:1-7.

2. 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.

3. Coté RA, Rothwell DJ, Palotay JL, Beckett RS. SNOMED International. Northfield, IL: College of American Pathologists, 1993.

4. Saba VK, Zuckerman AE. A home health care classification system. In: Lun KC, Degoulet P, Piemme TE, Rienhoff O, eds. MedInfo92. Geneva: North-Holland, 1992; 344-348.

5. Grobe, SJ. Nursing intervention lexicon and taxonomy study: language and classification methods. Adv Nurs Sci. 1990;13(2):22-34.

6. 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, 1991:126-31.

7. Grobe SJ. Nursing intervention lexicon and taxonomy: preliminary categorization. In: Lun KC, Degoulet P, Piemme TE, Rienhoff O, eds. MedInfo92. Geneva: North-Holland, 1992:981-6.

8. Martin KS, Sheet NJ. The Omaha System: Applications for Community Health Nursing. Philadelphia: W. B. Saunders, 1992.

9. Martin KS, Sheet NJ. A community health nursing data management model. In: Lun KC, Degoulet P, Piemme TE, Rienhoff O, eds. MedInfo92. Geneva: North-Holland, 1992:1036-40.

10. Kim MJ, McFarland G, McLane A. A Pocket Guide to Nursing Diagnosis. 3rd ed. St. Louis: C. V. Mosby, 1989.

11. McCloskey, JC, Bulechek GM. Nursing Intervention Classification. St. Louis: C. V. Mosby, 1992.

12. Ozbolt JG, Fruchtnicht JN, Hayden JR. Toward data standards for clinical nursing information. J Am Med Informatics Assoc. 1994;1:175-185.

13. Werley HH, Lang NM (eds). Identification of the Nursing Minimum Data Set. New York: Springer, 1988.

14. McCormick KA, 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 Clinical Nursing Practice. J Am Med Informatics Assoc. 1994;1:421-427.