Problem-Specific Coding Systems

John H. Renner, MD
Ellis A. Bauman, MS
Madison, Wisconsin

The data found in medical records, particularly in problem-oriented records, could yield very useful information for many aspects of the health care delivery system. Unless the tools and methods used in the analysis of health care data are very carefully designed, the potential value of this information cannot be fully realized.

Coding is a critical step in the process of converting medical record data into useful information. This paper presents six criteria as design goals for a standard, comprehensive, coherent set of nomenclatures, classifications, taxonomies, and corresponding coding schemes for the analysis of medical and administrative record data. These goals are: compatibility with established routine operating procedure, health care problem specificity, hierarchical structure, economy through articulation of the components with one another and through the similarities in various aspects of health care delivery, and adaptability to change.

The Need

In recent years, the use of problem-oriented medical records has been recognized as a definite improvement in the documentation of medical care, particularly in family practice and other areas of primary care. While it may appear that Weed's "Word" is spreading very far very rapidly, this record-keeping revolution is, in turn, creating the need for a corresponding revolution in those corners of the health care world concerned with the collection and analysis of medical record data. To state the matter plainly: the proper processing of problem-oriented data requires problem-oriented tools and methods. While this dictum may seem simple – after all, problem-oriented records obviously contain problem-oriented data – the issue is a substantive one. It rests on the importance of our having access to problem-oriented information.*

No one denies that medical records represent a vast repository of data which, if analyzed, would yield an immense amount of useful information for the improvement of patient care, medical education, health planning, medical administration, health policies, patient education, preventive health maintenance techniques, and many other facets of our health care system. The accuracy, validity, and even the utility of this information, however, is directly dependent on the specificity of the medical record data regarding who does what to whom, where, when, why, and at the expense of what resources (time, money, personnel, supplies, space, etc) in the provision of health care. In fact, one of the strongest arguments behind the problem-oriented approach is that it provides the framework, both structural and philosophical, required for meaningful organization of these detailed data. Yet, unless the data collection and analysis process takes advantage of the potential specificity inherent in problem-oriented medical record data, much of the utility of the information resulting from that process will be lost.

Next to placing the raw data in the medical record, the coding of the data is perhaps the most critical step in the process of turning those data into useful information. It is here that the data are converted into a form amenable to automated manipulation. At the same time, the first aggregation of the data into the categories of a classification takes place, with the unavoidable result that some of the specificity originally present in the data evaporates. In order to minimize this loss of information content, we must put a great deal of time and effort into the design of the coding schemes we use.

In designing an information system, be it based on medical record data or otherwise, the place to begin is not with the input, but rather with the output. Before we ask what data need to be encoded, we must ask what information we seek, and why we want it. Only after these questions have been answered can we discuss the design goals for a problem classification, and only after agreement on goals has been reached can we hope to adopt standard classifications and coding schemes for the analysis of health care data.

There are many applications of the information one might obtain from medical record data. Many of these are in areas of direct concern to family medicine, and all ultimately result in improvements in patient care and in the health status of individuals and communities. For example, new knowledge of disease processes and etiology might be gained, leading to better preventive medicine and better patient education. It might be possible to document accurately, and hence evaluate, the quantity and quality of care delivered in different settings by differently organized health care teams. The information may prove

From the Department of Family Medicine and Practice, School of Medicine, University of Wisconsin, Madison, Wisconsin. Dr. Renner is Chairman of the Department of Family Medicine and Practice. Requests for reprints should be addressed to Mr. Ellis A. Bauman, Information Management Specialist, Department of Family Medicine and Practice, School of Medicine, University of Wisconsin, 777 South Mills Street, Madison, Wis 53715.

*The International Organization for Standardization/ Technical Committee 97/ Subcommittee 1 has agreed on the following definitions: Data: a representation of facts, concepts, or instructions in a formalized manner suitable for communication, interpretation, or processing by humans or automatic means. Information: the meaning that a human assigns to data by means of the known conventions used in their representation.[1]

useful in evaluating and improving educational programs for undergraduate, graduate, and postgraduate medical students, as well as other health care personnel. It may also help us to evaluate these students. Moreover, it is very likely that the information we gain through the analysis of medical record data will be invaluable in areas of health services research: comprehensive health planning, medical facilities planning and management, public health, community medicine, and health program administration. Peer review and care audit will also be enhanced, as will epidemiologic studies based on follow-up data for large, well-stratified samples of the patient population. The list of applications goes on and on.

In the past, workers in each of these fields have had to develop and redevelop coding systems geared to their particular interests within health care delivery. Thus, we now have H-ICDA for coding diagnoses, CRVS for procedures, SNOP or SNOMed for pathology. Unfortunately, because these were all developed with narrow purposes in mind, their usefulness is very limited. Each addresses only small portions of the application areas listed above. Yet, because of the amount of work required to develop such codes, there has been a tendency to apply those available beyond their intended use rather than to create more suitable ones. The result is a great deal of information of doubtful value. Recently, some workers in the field, for example Schneider and Kilpatrick,[2] have attempted to link these codes together, a task of monumental proportions. These experiences all point out the need for a standard, comprehensive, coherent set of nomenclatures, classifications, taxonomies, and associated coding schemes that would allow comparisons of samples, methodologies, and results to be meaningfully made between studies.

The Design

Let us consider the design goals for such a set of data analysis tools. First of all, these tools should be capable of handling data that are routinely collected as an integral part of health care delivery, that is, medical and administrative record data. The ability to code and classify such data would obviate the need for special data collection instruments (eg, encounter forms) and the attendant wasteful duplication in time and effort on the part of providers and clerical personnel. In addition, the disturbance to office routine would be reduced, leading to greater compliance in the data collection process. Moreover, codes designed to be compatible with administrative and medical record forms and procedures could be used to support the health care delivery system, as well as to study it. They could be used by managers and planners, providers and provider educators, clinical researchers and social science researchers alike. Thus, the first goal is *compatibility with established routine.*

A second design goal for a set of codes is *specificity*. Coding at the most elemental level not only reduces coder training requirements, thereby increasing accuracy and reproducibility in the coding process, but, more important, it is demanded by the need to make decisions based on the information content of the data. Peer review entails a determination of the appropriateness of the intervention modality used (observation, treatment, referral, etc) as well as of the intervention itself. Unless the coding schemes are sufficiently detailed to separate the different types of health care problems (diagnoses, signs, symptoms, social problems, laboratory findings, etc), differences in significance arising from the seriousness of the problem or the skill of the provider cannot be taken into account. Similarly, judgments of resource allocation depend on knowledge of the services required. If the elements of a coding scheme are too general, they cannot provide the detailed information necessary for the setting of health care policies or for the formulation of problems or their solutions. Ideally, then, coding should take place at the nomenclature level in order to lower cost, raise validity, and avoid loss of information through aggregation of data too early in the analysis process.

On the other hand, in order to facilitate the aggregation of data at appropriate points in the analysis process, the classifications and taxonomies should be organized in parallel hierarchies wherever feasible. For example, the nomenclatures of physical findings, symptoms, diagnoses, laboratory test results, and x-ray study results might all be organized according to major body systems and organs. Reflecting such a hierarchical structure in the corresponding codes can greatly reduce the time and effort involved in writing the computer software to manipulate the data. Moreover, both gross and fine studies can be based on the same sets of data, with a high degree of comparability. A third design goal, then, is to *base the codes on parallel structures.*

Closely related to the goals of specificity and hierarchical organization is that of *articulation of the component codes with each other.* If all aspects of health care delivery are to be studied, if we are to discover who (the provider) does or could do what (the services, treatments, therapies) to whom (the patient population), where (the facilities), when (the stages in the health care process), why (the decision process), and at the expense of what resources (the money, personnel, supplies, space, time, etc), then the codes used must dovetail cleanly with each other. Again, the recent work of Schneider and Kilpatrick[2] illustrates the extra effort required to make up for the lack of articulation. On the other hand, good articulation not only provides economy in the number of component codes, but also facilitates linkage between the several aspects of health care mentioned above. For example, the model clinics of the University of Wisconsin Department of Family Medicine and Practice use a patient identifier system which also automatically provides a family identifier, both patient and family account numbers, a chart-filing location, and the sex and position of the patient within the household social structure, including the birth order of natural children.[3] Similarly, our provider identifiers distinguish between skill levels as well as health occupations. We also append one character to the zip code to indicate census tract, thereby linking the data of our medical and administrative records to the socioeconomic data available from the Census Bureau. To be sure, the linkage allowed by such articulation is often provided directly by the forms and procedures we use, but in order to realize the benefits, the forms and procedures and the codes must be designed with linkage in mind.

Another goal of code design which results in economy in the number of codes is to *take advantage of similar-*

ities in various aspects of health care. A scheme for coding chief complaints ought to play a role in the routine operation of an appointment system and in the recording of telephone contacts. Prefixes and suffixes can provide common modifiers indicating body part, duration, severity, etc. A medication code can be used for prescriptions as well as inventory control in a pharmacy. While the economy thus provided arises in much the same way as that gained from articulation, capitalizing on similarities extends the flexibility of a code by permitting it to be used for several different applications, and it enhances linkage.

Finally, each of the component codes, as well as the full comprehensive set of classifications, must be designed to permit graceful expansion and modification. Although we cannot accurately predict the nature of the changes the future has in store for us, we do know with certainty that there will be change. Unless we *provide for ongoing adaptation on the basis of experience,* trends in the patient population and in the organization of health care delivery will make our tools and methods obsolete.

In summary, the components of a comprehensive information system for health care must coalesce into a coherent whole. The nomenclature, classification schemes, taxonomies, and system of corresponding codes should support the routine operations of the health care delivery system as well as provide tools for the study of that system without interfering with its efficiency of operation. The components of the data analysis process must be problem-oriented, for only then will they be sufficiently specific to provide the information necessary for decision-making and evaluation of the decisions made. Moreover, it is only through specificity that wide applicability can be obtained. If the codes can be used by HMO's as well as fee-for-service delivery systems, by neurosurgeons as well as family physicians, by public health officials as well as intensive care unit personnel, then meaningful comparisons of practice patterns across specialties and across delivery organizations can be made. For us in family medicine, this information could be very useful in documenting the differences between family physicians and others engaged in ambulatory and primary care. Even within a single specialty, the comparison of practice profiles requires a specificity that only problem-oriented information can provide. If both the data and the information are to be problem-oriented, then so must the tools and methods of analysis. Finally, if the information gained is to be used for the improvement of health care administration and management, as well as to guide the education and training of health care providers and be the basis for sound policy and planning, then it must be sufficiently specific to make a thorough and detailed systems analysis possible. It follows that the collected data must form an accurate image of the real world of health care. While the process of analysis necessarily entails an unavoidable loss of detail, we must make every effort to keep the picture as sharp as possible. In short, the tools of research must be carefully fitted to the task. In the context of family medicine, this principle means we need problem-specific coding systems to deal with medical and administrative record data, if we are to benefit from the study of problem-oriented medical practice and the operation of its care delivery system.

References

1. American National Standard Vocabulary for Information Processing. New York, American National Standards Institute, 1970
2. Schneider DP, Kilpatrick KE: A medical classification system for ambulatory-care manpower planning. Health Serv Res 9:221-233, 1974
3. Renner JH, Piernot RW, Bauman EA, Treleven DE: A patient and family number assignment and chart filing system for family physicians. Med Care 13:346-359, 1975