times, these agendas provide the rationale to help deans allocate resources necessary to best educate students.

QUALITY ASSURANCE. Invested with the philosophy of "total quality improvement," the curriculum committee should exercise a significant quality assurance function, monitoring student course evaluations and sponsoring course reviews with teams of faculty members that represent various disciplines.

This position springs from the recognition that the curriculum committee, if it is to be responsible for the curriculum, must have the authority necessary to exercise this responsibility. This authority must include the power to decide whether courses offered by a college of medicine are of high quality. Of course, the committee will work with the directors of marginal courses to improve them, but the authority to disapprove courses or clerkships is necessary to maintain quality when less draconian interventions are ineffective.

EMPOWERMENT. The curriculum committee should conduct its business without undue departmental influence. The committee should report directly to the college of medicine dean who implements all curricular decisions based on economic reality and with input from groups independent of the curriculum committee.

This policy speaks to the potential problem of "departmentalization" of the curriculum. It is necessary because the curriculum committee could be faced with decisions that are not in the best interest of a particular department. Because the responsibility of the curriculum committee is education, and not the welfare of individual departments, this provision diminishes the ability of powerful department chairs to dominate curricular matters to the detriment of the overall educational enterprise.

CONCLUSION

Curriculum committees from schools in the midwest CGEA are constituted in almost as many ways as there are schools. Our hypothesis is that certain structures and functions of curriculum committees are optimal for devising the best educational experiences in medical schools. This position paper represents what a group of educators in the Midwest believes is a reasonable set of guidelines for the organization of an effective curriculum committee.

Curriculum Committee SIG members who have approved the position paper are: C. M. Banerjee, MD, PhD, James E. Carter, MD, Linda H. Distlehorst, PhD, J. Kevin Dorsey, MD, PhD, Fred L. Ficklin, EdD, Alberto Galofré, MD, Linda K. Gunzburger, PhD, Albert E. Langley, PhD, Edward A. Lichter, MD, Michael M. Ravitch, PhD, Allen B. Rawitch, PhD, Murray Saffran, PhD, Ruth Andrea Seeler, MD, Mary R. Smith, MD, Robert J. Winter, MD, and Ernest L. Yoder, MD.

Dr. Whalen is associate professor of medicine and program director, Department of Medicine, at the University of Illinois College of Medicine at Chicago.

References

1. Cantor, J. C., Cohen, A. B., Barker, D. C., Shuster, A. L., and Reynolds, R. C. Medical Educators' Views on Medical Education Reform. *JAMA* **265**(1991):1002–1006.
2. Kettel, L. J. Assessing Change in Medical Education (ACME). *Acad. Med.* **64** (1989):378–379.
3. Hendricson, W. D., Katz, M. S., and Hoy, L. J. Survey on Curriculum Committees at U.S. and Canadian Medical Schools. *J. Med. Educ.* **63**(1988):762–774.

Medical Informatics

MARK FRISSE, MD
Associate Editor

Saying What You Mean and Meaning What You Say: Coupling Biomedical Terminology and Knowledge

JAMES J. CIMINO, MD

Effective information management promises to play a vital role in addressing many of the problems facing health care. Tasks such as quality assurance, documentation of procedures potentially liable to medical malpractice suits, cost containment, utilization review, and health care reimbursement can be managed more effectively if information readily available to individuals can be coordinated and processed in efficient ways. Health care providers empowered with this information should be able to deliver higher-quality, efficient care. In the ordering of medical procedures, for example, the propriety of a decision is dictated by factors such as contraindications, redundancy, timing, and efficacy. Inappropriate ordering of medical procedures diminishes the quality and efficiency of medical practice and opens the way for possible harm to the patient as well as denial of payment. Rules governing ordering of medical procedures are often simple: "administer platelets before performing a lumbar puncture when the patient's platelet count is too low"; "submitting an additional stool culture is unlikely to yield additional information when six cultures have already been submitted"; "do not perform a barium enema on the day after an upper GI series was performed unless the patient has been adequately prepped." When recorded accurately, information about medical decision making can, in turn, provide the documentation needed to support decisions and obtain appropriate compensation. When coupled with information about patient outcomes, information about—and leading to—decisions can be used to develop practice guidelines.

Computers play a vital role in information management because they collect, process, and deliver information in an organized way. Unfortunately, although previous installments in this series have attested to the value of computers in clinical care, the status quo addresses neither the breadth nor the scale necessary to solve the broader set of problems posed by modern health care. The clinical data collected at a hospital in Boston, for example, are not readily reduced to practice guidelines in Washington. Nor can practice guidelines in Washington easily be included in decision-support systems used in Salt Lake City. The first installment of this column[1] addressed two technical difficulties that were reviewed in subsequent columns: terminology[2] and knowledge representation.[3] Although progress has been made in each of these two areas, the unification of terminology and knowledge representation remains a difficult challenge in medical informatics.[1-3]

TERMINOLOGY VERSUS KNOWLEDGE

To understand the need for unification of terminology and knowledge, it is important to be clear about their differences. Terminology is defined here as a set of recognized names (or terms) for concepts. Knowledge is information about the concepts. The concept names that constitute a terminology are one kind of knowledge. Other kinds of knowledge include the characteristics used for classification, the definitions of the concepts, and how the concepts relate to one another.

Physicians use terms with an implicit understanding of underlying medical knowledge. Whether the statement is "order a potassium," "the potassium is 2.3," "the potassium hemolyzed," or "administer some potassium," the human listener knows that the topic is a laboratory order, a test result value, a laboratory artifact, or a medication order, respectively. Computers do not do this very well. If it were within a computer's power to comply, it might respond to the last phrase with a dose of caustic metal instead of an appropriate electrolyte solution. A rule could be written for the computer to follow, such as "if the last potassium is less than 3, then suggest that the doctor order some potassium chloride." But it must be specified that by "last potassium" the meaning is "last result of a serum or plasma potassium ion test," with "last" meaning "the last one that isn't hemolyzed" and "within the last few days" (whatever "few" means).[4]

If we are explicit enough about what we say, we can get the computer to do what we mean. However, this task can get tedious, and the potential for being misunderstood is never eliminated. When dealing with computers, it is much better to have a finite list of terms with which to express oneself. Thus, one selects "serum electrolyte panel" when ordering a laboratory test, "serum potassium" when requesting a test result, and "potassium chloride elixir" when ordering a medication. These lists are the controlled terminology by which we speakers of natural language constrict our discourse to convey precise meaning. Controlled terminologies are easy to construct and use when dealing with well-defined collections of concepts such as tests in a laboratory, bones in the human skeleton, or elements in the periodic table. Difficulties arise when trying to enumerate all the terms in larger domains, such as "disease." Particularly troublesome are issues of *content*, *structure*, and *semantics*.

Providing adequate content seems simple enough: keep adding terms until all the necessary terms have been added. But what's necessary? Does "congestive heart failure" belong in the disease list or does it belong in a list of diagnostic findings? Should a "fracture of X" be listed for every bone in the body? If so, what about "greenstick fracture of X?" Should "greenstick fracture of the skull" be included? As additions are made to the terminology, care must be taken to avoid reintroducing a concept under a different name. It will do the quality assurance manager little good to retrieve all the ornithosis patients from a hospital database if they've all been coded as having psittacosis.

Order is usually imposed on controlled terminologies by providing an organizational structure. The most common approach is the use of classification hierarchies. Because we tend to think about technical issues in a hierarchical fashion,[5] arranging terminologies that way provides a natural method for browsing. So, when the manager looks under lung disease and can't find "ornithosis," "psittacosis" may be found in its place. A variety of classification schemes abound, and medical concepts often rightfully belong to more than one class. A disease, for example, can often be identified with one class based on its location in the body and with another class based on its etiology. Unfortunately, for technical reasons, entries in controlled terminologies are usually placed in only one class. If "psittacosis" is listed under infectious diseases, our manager may not find it by looking under lung diseases. Thus, the way a terminology is organized can be just as important as the terms it contains.

Simple enumeration of phrases in a terminology is not sufficient to convey the semantics (or meanings) of the terms. Organizational structure can provide some information; for example, if "Paget's disease" is listed under the bone diseases, one can infer that osteitis deformans is the intended meaning and not Paget's mammary disease. But when the term "Other Specified Conduction Disorders" is encountered, as it is in ICD9-CM[6] (code 426.89), its precise meaning is less clear. Knowing that the term appears under "Conduction Disorders" (426) is not sufficient to distinguish it from the term "Other Heart Block" (426.6), which also appears there. And when encountering "Other Psoriasis" under two different codes (696.1 and 696.8), how does one know which code is appropriate? Some of the intended semantics are provided in voluminous coding manuals; none are provided in a computer-usable fashion. Yet if the computer is to use controlled terms appropriately, well-defined semantics are crucial, since computers have no inherent medical knowledge.

TERMINOLOGIES IN THE 1990s

Besides highly customized local terminologies, a few standard, relatively simple (albeit large) terminologies are used today to encode medical information for a variety of purposes. Table 1, based on a review by Masys,[2] lists the characteristics of the most widely used of these terminologies. While each has been successful for its specific purpose, none is adequate for representing medical information in ways that are needed for automating sophisticated tasks such as quality assurance, cost containment, reimbursement, and outcome analysis.

Limitations of the content, structure, and semantics are to blame.[7] Current work to build new terminologies, or modify existing ones, is giving more consideration to these issues. For example, the latest incarnation of the *International Classification of Diseases*, ICD-10, is due to be adopted in Europe this year and in

Table 1
Characteristics of Common Medical Terminologies

Terminology	Acronym	Publication Year	Organization	No. of Terms	Purpose
International Classification of Diseases, Ninth Revision, Clinical Modifications*	ICD-9-CM	1980†	U.S. Department of Health and Human Services	18,630	Indexing hospital charts by morbidity and mortality information for use in billing and epidemiology
Systematized Nomenclature of Medicine, Second Edition	SNOMED II	1982	College of American Pathologists	30,800	Encoding signs, symptoms, diagnoses, and procedures in hospital records
Current Procedural Terminology, Fourth Edition	CPT-4	1984†	American Medical Association	7,300	Encoding procedures for billing
Medical Subject Headings	MeSH	1966†	National Library of Medicine	15,890	Indexing medical literature citations

*Based on the World Health Organization's *International Classification of Diseases, 9th Revision* (ICD-9), published in 1977.
†Updated annually.

the United States by 1996. The developers of ICD-10, aware of the deficiencies in ICD9, have expanded the classification scheme and extended its content. However, the basic structure is unchanged. For example, each term appears in one class, even if more than one class might be appropriate.[8,9] The new edition may address some needs for computer-based terminologies (such as expanded content) but not others (such as knowledge representation). The *Systematized Nomenclature of Medicine* is nearing release of its third edition, SNOMED III.[10] Despite its flexibility, SNOMED II found little acceptance in medical computing. A great deal of attention has been paid to develop the classification scheme and the content, but these were not usually cited as faults of SNOMED II. Rather, it was the semantic representation that seemed to cause problems.[7,11] The basic model of SNOMED III is that of SNOMED II. Whether the expanded content will make SNOMED III suitable for future work in medical knowledge representation remains to be seen.

TERMINOLOGIES FOR THE TWENTY-FIRST CENTURY

Much effort has gone into the development of terminologies for the representation of medical knowledge, whether it be for patient records, health statistics, or expert systems. For some time, medical informatics researchers have been proposing that the process be inverted: that medical knowledge should be used to represent the terminologies.[12-15] The rationale is that by providing explicit information about the terms used to code patient data or to write clinical guidelines, computers can help manipulate them.[16]

Several projects are under way to bring terms and their underlying knowledge together, for the purposes of unifying existing terminologies and expanding them to serve the needs of medical computing. The Unified Medical Language System (UMLS), being developed by the National Library of Medicine, contains a metathesaurus of terms from numerous medical terminologies, a semantic network with information about how the terms relate to each other, and a map showing how the terms are used in information sources such as bibliographic databases, patient databases, and expert systems.[17] In Europe, the Advanced Informatics in Medicine (AIM) program is sponsoring the GALEN project. GALEN includes an effort to represent medical concepts for a multilingual coding system that can be used to integrate clinical information systems.[18] Also in Europe, the Comitė Européen de Normalisation (CEN) is sponsoring the Technical Committee on Medical Informatics (TC251) to provide content that conforms to GALEN's structure.[19]

The workers involved in all of these efforts are coming to some similar conclusions. First, existing terminologies such as those in Table 1, as flawed as they are for general usage, are here to stay. Rather than trying to usurp them or to simply ignore their presence, efforts are being made to subsume them, taking advantage of whatever value their content might provide, while developing new representational schemes that overcome the disadvantages of the original terminologies.

Second, construction of the terminologies is not simply a matter of selecting an organizational structure and then convening a committee of medical experts to provide the content. Instead, each group is compiling well-defined rules that, when applied by different cohorts of experts, will provide reproducible results regarding what terms are added, how they are arranged, and what terms are discarded. It is hoped that the efforts invested in developing these rules and in their rigorous application will pay off in terms of the utility, assessment, and, most important, adoption of the results.

Third, it is recognized that shoehorning the rich expression of medical language into a paltry few thousand terms is inappropriate. For example, when a patient's diagnosis of "intermittent isorhythmic atrioventricular dissociation" is represented as merely "Other Specified Conduction Disorders," so much detail is lost that the remaining information is almost useless. Instead of discarding the additional qualifying information, researchers are seeking ways to represent it to maintain as much clinical information as possible in coded form.

Fourth, the days of large collections of medical terms, cleverly arranged into some form that might be read by tenacious humans, are past. A clear need has been recognized for the codification of an additional layer of information in terminologies, one that includes knowledge about the terms themselves. Whether this layer is called a Metathesaurus[17] or Structured Meta Knowledge,[18] the notion is that this "meta-information" will provide the capabilities that have been absent from medical terminologies. Some of this meta-information is linguistic, some is historical, but the interesting parts are medical. Instead of simply asking medical experts what to put in the terminology, the expert knowledge is being identified and placed in the terminologies along with the terms. This is more than a matter of including free-text definitions of each term. That step alone would be of enormous value to human users, who must puzzle over the terms and wonder exactly what their inventors had in mind. But the most novel approach is the inclusion of medical knowledge in ways that can be used by the computers themselves. There are considerable potential benefits to both the builders and the users of these next-generation terminologies.

For the builders, there will be opportunities to develop computer-based tools for assisting in maintenance. Construction rules can be implemented as computer programs that can request specific medical knowledge about terms and then help with the inclusion of new terms.

Users of these terminologies will find countless ways to take advantage of meta-information. A pharmacy system with information about chemical components and therapeutic classes of medications can detect contraindications based on a patient's history and laboratory findings, and it can suggest substitutes that are more appropriate (quality assurance) and cheaper (cost containment). A patient database that has encoded detailed demographic information and procedure complications can help predict subgroups of patients at increased risk from particular operations (outcome analysis). A medical record system that has coded physician notes can help determine which patient encounters are comprehensive visits (provider reimbursement). The same system can warn when documentation of informed consent of an administered procedure appears inadequate (reducing liability).

Medical informaticians, hospital executives, health care policymakers, and record system vendors around the world are watching these developments carefully, with the hope that the current problems with computer-based medical terminologies will be solved. Projects touched on here and many other similar efforts are being examined closely by the Computer-Based Patient Record Institute (CPRI),[20] with the expectation that one or more will serve as or lead the way to a terminology on which a computerized record system can be based.

CONCLUSION

For computers to provide assistance with the problems facing modern medicine, attention must be paid to how patient information is coded. A current challenge for biomedicine can be exemplified by the Chinese saying: the beginning of wisdom is getting things by their right name. Medical informatics research is moving to meet this challenge with the development of knowledge-based approaches to medical terminology that will give us the flexibility to say what we mean and a scientific basis on which to verify that we mean what we say.

Dr. Cimino is assistant professor of medicine, Center for Medical Informatics, Columbia University College of Physicians and Surgeons, New York, New York.

Correspondence should be addressed to Dr. Cimino, Center for Medical Informatics, Atchley Pavilion, Room 1310, Columbia-Presbyterian Medical Center, New York, NY 10032.

This study was supported in part by a Unified Medical Language System (UMLS) contract from the National Library of Medicine.

The author thanks Mark Frisse, George Hripcsak, and Leslie Juceam for their editorial assistance.

References

1. Masys, D. R. Medical Informatics: Glimpses of the Promised Land. *Acad. Med.* **64**(1989):13–14.
2. Masys, D. R. Of Codes and Keywords: Standards for Biomedical Nomenclature. *Acad. Med.* **65**(1990):627–629.
3. Miller, R. A., and Giuse, N. B. Medical Knowledge Bases. *Acad. Med.* **66**(1991): 15–17.
4. Hripcsak, G., et al. The Arden Syntax for Medical Logic Modules. In *Proceedings of the Fourteenth Annual Symposium on Computer Applications in Medical Care*, R. A. Miller, ed., pp. 200–204. Los Alametos, California: IEEE Computer Society Press, 1990.
5. Blois, M. S. Medicine and the Nature of Vertical Reasoning. *N. Engl. J. Med.* **318**(1988):847–851.
6. United States National Center for Health Statistics. *International Classification of Diseases, Ninth Revision, with Clinical Modifications*, DHHS Publication No. (PHS) 80–1260. Washington, D.C.: Government Printing Office, 1980.
7. Cimino, J. J., Hripcsak, G., Johnson, S. B., and Clayton, P. D. Designing an Introspective, Controlled Medical Vocabulary. In *Proceedings of the Thirteenth Annual Symposium on Computer Applications in Medical Care*, L. W. Kingsland, ed., pp. 513–518. Washington, D.C.: IEEE Computer Society Press, 1989.
8. Bramer, G. R. International Statistical Classification of Diseases and Related Health Problems—Tenth Revision. *World Health Stat. Q.* **41**(1988):32–36.
9. Hirs, W. M. The Use of Terminological Principles and Methods in Medicine. In *MEDINFO 92*; K. C. Lun, P. Degoulet, T. E. Piemme, and O. Rienhoff, eds., pp. 1452–1457. Amsterdam, The Netherlands: Elsevier North-Holland Press, 1992.
10. Coté, R. A., and Rothwell, D. J. *Systematized Nomenclature of Medicine—3rd Edition* (SNOMED III). Skokie, Illinois: College of American Pathologists (in press).
11. Dunham, G., Henson, D., and Pacak, M. Three Solutions to Problems of Categorizing Medical Terminology. *Meth. Inform. Med.* **23***(1984):87–95.*
12. Huff, S. M., Craig, R. B., Gould, B. L., Castagno, D. L., and Smilan, R. E. Medical Data Dictionary for Decision Support Applications. In *Proceedings of the Eleventh Annual Symposium on Computer Applications in Medical Care*, W. W. Stead, ed., pp. 310–317, Los Angeles, California: IEEE Computer Society Press, 1987.
13. Cimino, J. J., and Barnett, G. O. Automated Translations between Medical Terminologies Using Semantic Definitions. *MD Comput.* **7**(1990):104–109.
14. Rector, A. L., Nowlan, W. A., and Kay, S. Unifying Medical Information Using an Architecture Based on Descriptions. In *Proceedings of the Fourteenth Annual Symposium on Computer Applications in Medical Care*, R. A. Miller, ed., pp. 190–194. Los Alametos, California, 1990.
15. Masarie, F. E., Miller, R. A., Bouhaddou, O., Giuse, N. B., and Warner, H. R. An Interlingua for Electronic Interchange of Medical Information: Using Frames to Map between Clinical Vocabularies. *Comput. Biomed. Res.* **24**(1991):379–400.
16. Blois, M. S. Information Holds Medicine Together. *MD Comput.* **4**(1987):42–46.
17. Humphreys, B. L., and Lindberg, D. A. B. Building of the Unified Medical Language System. In *Proceedings of the Thirteenth Annual Symposium on Computer Applications in Medical Care*, L. W. Kingsland, ed., pp. 475–480. Washington, D.C.: IEEE Computer Society Press, 1989.
18. Rector, A. L., Nowlan, W. A., and Kay, S. Conceptual Knowledge: the Core of Medical Information Systems. In *MEDINFO 92*, K. C. Lun, P. Degoulet, T. E. Piemme, and O. Rienhoff, eds., pp. 1420–1426. Amsterdam, The Netherlands, Elsevier North-Holland Press, 1992.
19. Rossi Mori, A., Galeazzi, E., Gangemi, A., Pisanelli, D. M., and Thornton, A. M. Semantic Standards for the Representation of Medical Records. *Med. Decis. Making* **11**, Supplement (October–December 1991):S76–S80.
20. Frisse, M. The Health of the Computer-based Patient Record. *Acad. Med.* **67**(1992):441–443.