Common Medical Terminology Comes of Age, Part Two: Current Code and Terminology Sets—Strengths and Weaknesses

Jeffrey S. Rose, MD; Bruce J. Fisch, MD;
William R. Hogan, MD; Brian Levy, MD;
Philip Marshall, MD, MPH; David R. Thomas, MD;
Debra Kirkley, PhD, RN

ABSTRACT

A number of clinical coding and vocabulary schemes are in use in health-care enterprises today. Most of them are weak in light of the qualities that characterize adequate controlled medical terminologies, as outlined in Part One of this review. Here we look at the major code and terminology sets with a critical eye, as well as suggest practical steps to enable health industry information system purchasers and users to move forward with their effort to use common terminology to improve the quality, service, and knowledge in their enterprise.

KEYWORDS
- Controlled medical terminology/vocabulary
- Clinical informatics
- Common medical terminology
- Standard clinical terminology
- Standard language
- Healthcare quality
- Sharing of data
- Outcomes

Part One of this review pointed out the need for controlled medical terminology (CMT) as infrastructure for healthcare information technology. It covered the characteristics and categories of terminology sets and detailed the necessary components of a comprehensive CMT and tools for managing it. In this part, we turn our attention to the coding and vocabulary sets in use

today. We take a critical look at these schema, some common and some not yet well known.

Diagnosis Coding Systems

The International Statistical Classification of Diseases, Injuries, and Causes of Death, formerly the International Classification of Deaths (ICD), originated in the London Bills of Mortality of the 1800s. The World Health Organization (WHO) assumed sponsorship of ICD in 1948 and since then has issued several versions, the most recent being ICD-10 in 1992 (now the formal name is International Statistical Classification of Diseases and Related Health Problems). The U.S. National Center for Health Statistics (NCHS) determined that the adopted version, ICD-9, with approximately seven thousand codes, was insufficient for use in the United States, and set about adding about six thousand additional codes to create ICD-9-CM (Clinical Modifications). The NCHS has similarly determined that ICD-10, with approximately thirteen thousand codes, is deficient and has commissioned development of clinical modifications to ICD-10 to replace ICD-9-CM. According to the NCHS, ICD-10-CM (about sixty thousand codes) will not be available until some time in 2001.[1] The United States is required by treaty to report its national health data to the WHO using ICD-10, although ICD-10 has not replaced ICD-9-CM as the administrative coding system for billing and reimbursement.

ICD-9 and ICD-10 are *classifications.* Thus one code typically represents a category onto which several diseases may be mapped. Several consequences result from this arrangement. One is that many categories are too broad to be clinically useful. Another is that a significant amount of clinical detail is lost when a paper-based medical record is coded. The only way to retrieve the detail in the medical record is to review the chart manually—an expensive and time-consuming task. A third consequence of using the ICD classification systems is that they contain many ambiguous and redundant catch-all categories. An example is ICD-9CM code 637.70, which essentially means "abortion, unspecified whether spontaneous, legally induced, or illegally induced, unspecified whether complete or incomplete, with other specified complications."

Accurately coding a patient's hospitalization or outpatient visit using ICD, with all the inherent flaws this discussion has revealed, is even more improbable than finding the proper code to represent the physician's written diagnoses. This results from the overwhelming rules for coding (in ICD-10, an entire volume is devoted to a rules "instruction manual"). The rules, which are intended to bring clarity, are complex enough that accuracy of use is dubious, and the likelihood of intercoder and intracoder variation is high.

Third-party payors in the United States place an additional burden on coders, in that certain codes are disallowed or "not reimbursable." Thus in the unlikely event that an ICD code properly describes the diagnosis or problem

in the medical record, it is doubtful that a coder will use it if it has been deemed nonreimbursable. This has led to providers' coding up to a reimbursable code to maximize reimbursement, and to payors' coding down to minimize payment. Such financial incentives have led to distortion of healthcare data in the United States.

Despite all the shortcomings of ICD, no end to its use is in sight. The Healthcare Insurance Portability and Accountability Act of 1996 (HIPAA) mandates certain information exchange standards for electronic transactions in healthcare. Unfortunately, the code set proposed for this purpose by the National Committee on Vital and Health Statistics is ICD-10.

Procedure Coding Systems

Three procedure-coding systems are in use in the United States:

1. Volume 3 of ICD-9-CM
2. The American Medical Association's (AMA) Current Procedural Terminology (CPT)
3. The Health Care Financing Administration's Common Procedure Coding System (HCPCS)

ICD-9-CM Volume 3 is used by hospitals for reporting procedures in the Uniform Hospital Discharge Data Set (UHDDS) for reimbursement by Medicare and Medicaid. The procedure codes in ICD-9-CM were derived from a companion volume to ICD-9 (developed after ICD-9 was released) called International Classification of Procedures in Medicine. This volume never received much attention outside the United States, and no similar volume is associated with ICD-10.

The AMA first published CPT in 1966, with new versions appearing approximately every four years until 1984. Since then, the AMA has released revisions every year, with each new version titled CPT with the year when it was released (for example, CPT 2001). CPT is maintained by an editorial panel of approximately fifteen physicians who are advised by two committees, one consisting of seventy-five physicians and another of ten individuals representing other healthcare professions such as nursing, social work, and physical therapy.

CPT is a flat (that is, nonhierarchical) list of approximately eight thousand five-digit numeric codes. Because CPT is intended for administrative and reimbursement purposes only, it tends to group procedures by the relative amount of effort required for performing the procedure.

Thus, as with ICD, procedures are lumped into categories and the clinical detail is lost. For example, consider CPT code 11620: "excision, malignant lesion, except skin tag (unless listed elsewhere), scalp, neck, hands, feet, genitalia; lesion diameter 0.5 cm or less." Codes 11621, 11622, 11623, 11624,

and 11626 are the same except for the size of the lesion. As a result, a researcher trying to collate all cases where an excision of a malignant lesion of the skin of any one of those areas was performed would have to use all six codes in the search. Locating all such cases, regardless of body site, would require the use of eighteen codes.

Furthermore, if the researcher wanted to identify cases in which the excision was performed on sun-exposed regions of the skin, perhaps to study malignant melanoma, a return to the paper-based record would be required. This is because sun-exposed and skin sites not typically exposed to the sun are contained in a single category; the genitalia are usually not exposed (nor is the scalp if it has not experienced hair loss). In contrast, the rest of the body sites in the list for code 11620 are typically exposed to the sun.

In addition to the five-digit codes, CPT has a list of thirty two-digit modifiers that are usually applied to the five-digit code. However, the two-digit modifier is not applicable to every five-digit code, leading to a similar situation as with ICD, in which a complex set of rules must be followed for proper coding.

HCPCS is essentially CPT augmented with codes for medical and surgical devices, supplies, and drugs, and nonstandard codes that each Medicare carrier may use for its own purposes. HCPCS has additional modifiers that must be appended when reporting to Medicare or Medicaid for reimbursement.

Pharmacy Coding Systems

Pharmacy terminology is an integral component in a controlled medical terminology that encompasses all the vocabulary necessary to document, order and fill prescriptions, give medication decision support, and facilitate outcomes-based research.

Interest in pharmacy terminologies, especially as a component of electronic medical records, has increased recently, in large part thanks to the adverse drug event problem. An Institute of Medicine study reported that up to ninety-eight thousand hospitalized patients die annually from medical errors. Other studies have demonstrated that 30 percent of hospitalized patients may experience an adverse drug event, at a cost of more than $2,000 per patient.[2,3] Information technology that gives computer-based and accurate clinical drug information, drug interaction, allergy and correct dose checking has been shown effective in reducing these iatrogenic events.[4] Data integration, however, is one of the major barriers to deployment and use of electronic medication processes.[5] A controlled pharmacy terminology is paramount for this data integration; it is intuitively obvious that drug-allergy and drug-condition checking is as fundamentally dependent on accurate depiction of allergies and conditions as it is on precision in drug identification.

Unfortunately, a complete, standardized pharmacy terminology does not exist. National Drug Codes (NDC) were originally created for an out-of-hospital Medicare drug reimbursement program. These codes are ten- or eleven-digit

codes made up of three components: the labeler code (the first four or five digits), the product code (the next three or four digits), and the package code (the last two digits).[6] NDC codes are available from three government sources (FDA, HCFA, and VA). HIPAA chose NDC as a pharmacy standard, but these codes fall far short of a complete controlled pharmacy terminology. Some have been reused, and others have become obsolete. Discrepancies in codes exist among government agencies and vendors, and the codes are not descriptive enough for physician prescription order entry, clinical documentation, or decision support.

Currently, only specialized vendors, notably First Data Bank (merged with MediSpan), Multum (subsidiary of Cerner), and Micromedex, offer a pharmacy terminology. Traditionally, these companies' markets have been primarily pharmacy systems; medication orders written by hand (in the chart or on a prescription pad) are entered into the pharmacy system using the terminology specific to that system. After entry, the pharmacist may be able to view drug interaction, drug allergy, and dose range (acceptable low- and high-dosing parameters) for a prescription on the basis of the accuracy and completeness of information entered, as well as the rules in that particular system for such evaluations. The increasing use of electronic medical records and electronic prescription systems has made pharmacy terminology important throughout the entire computerized healthcare spectrum.

Ambulatory Patient Classification

As mentioned earlier, various administrative terminologies have been created to satisfy a specific need at a specific time. One terminology recently developed is the Ambulatory Patient Classification (APC). This was created by HCFA under the authority of the Balanced Budget Act of 1997 and the Balanced Budget Refinement Act of 1999 as part of the Hospital Outpatient Prospective Payment System for payment of outpatient services offered under Medicare.[7] The APC is made up of groups of procedures such that the services within each group have comparable resource usage and are clinically similar. The basis for the grouping was the Ambulatory Patient Groups developed by 3M Health Information Systems in the late 1980s and early 1990s.[8] The 451 APC groups can be categorized into five status types: significant procedures, surgical procedures, medical visits, ancillary tests and procedures, and partial hospitalization.[8]

HCFA defined the groups within the APC system according to five fundamental criteria: resource homogeneity, clinical homogeneity, provider concentration, frequency of service, and limited opportunities for upgrading of codes and code fragmentation.[7]

SNOMED and SNOMED-RT

SNOMED-RT (RT for reference terminology) and the foundations it was built upon have been under development by the College of American Pathology for more than thirty years. It was first known as the Systematized Nomenclature of Pathology (SNOP). SNOP became available in 1965; the majority of

pathologists adopted it over the ensuing ten years. Although SNOP addressed only nomenclature and coding (not classification), its four-axis arrangement of medical terms became a valuable tool for organizing and retrieving free-text pathology reports. SNOP's success led to SNOMED (Systematized Nomenclature of Medicine), which was first released in 1977 and featured a multiaxial and hierarchical classification of medically useful terms. Two years later, SNOMED II was released; at the time, seven core axes were included to cover topography, morphology, etiology, function, disease, procedure, and occupation. With the advent of SNOMED version 3.x in 1993, the etiology axis was split into four sections to represent living organisms, chemicals, physical agents and forces, and general linkage modifiers. Veterinary terms were added and the overall content was iteratively enhanced through version 3.5 in 1998.

Several investigations confirmed SNOMED as a source with one of the best overall coverages of clinical content.[9,10] A major evolution for SNOMED, from SNOMED 3.x to SNOMED-RT, moves the terminology into the realm of true ontology through use of explicit hierarchies, description logic concept definitions, and inference-based relationships. The underlying granularity of SNOMED-RT, which enables expressiveness, and the organization of these granular terms into hierarchies with logical relationships, permits computational categorization and management of the vocabulary based on strict concept definitions. In addition, the codes assigned by SNOMED editors no longer carry explicit meaning or implied relationships. All definitions are now handled regardless of the code and hierarchies made explicit by using *is-a* or *part-of* statements.

With its official release in 2000, SNOMED-RT includes approximately 180,000 terms linked to 110,000 clinically relevant concepts organized into eleven axes.[11] The new standard creates 260,000 hierarchical and semantic relationships, and the entirety of the content (terms, codes, concepts, and relationships) will be made available in computer-optimized formats.

The heart of SNOMED-RT is its breadth of terminological coverage, logical classification system, and compositional nature arising from medically relevant and granular concept elements. Examined critically, it meets most of the requirements of an "ideal" CMT as outlined in Part One of this report. In comparison to other medical terminology coding and classification systems, SNOMED-RT is far superior in satisfying the desiderata and covering the breadth and depth of healthcare terminology. It therefore represents the most advanced medical vocabulary for developers and enterprises looking to reconcile clinical medical language with other workflows that require coded data.

UMLS

The Unified Medical Language System (UMLS) was developed by the National Library of Medicine in an attempt to unify disparate medical vocabularies and facilitate sharing medical knowledge across information systems.[12]

Initial efforts were directed at providing a linguistic framework for integrating clinical knowledge, but additions covering more arcane coding systems have complemented its content, such that now the eleventh edition of the UMLS Metathesaurus contains more than sixty source vocabularies. Because of its massive scope and linguistic challenge, some developers have attempted to use its core unit of meaning, the concept unique identifier (CUI), as a coding standard itself, despite the warnings of informatics experts who point out that UMLS CUIs have external meaning strictly in the context of their source vocabularies and coding schemes.[13,14] Moreover, official UMLS documentation warns the user at the outset that "the UMLS Knowledge Sources are created for developers, and are not end-user applications. Significant intellectual effort and computing resources may be required to understand and use them correctly!"[15] For these reasons, the UMLS Metathesaurus is discussed here as a coding and language support tool, rather than a coding system per se.

The UMLS is composed of several knowledge sources. (The Metathesaurus is described later.) The semantic network gives general categories or semantic types to which all concepts in the Metathesaurus have been assigned. Finally, the SPECIALIST lexicon creates a container for syntactic information about terms in the Metathesaurus. Additionally, the UMLS is distributed with lexical programs designed for searching, indexing, and processing.

Because the Metathesaurus captures the actual codes, relationships, terminology, and meaning of its source vocabularies and adds additional structure and basic information to these schema, it is the key ingredient when considering the role of the UMLS in codifying medical data. A complete listing of its component sources is available.[16] Organization of items within the Metathesaurus is based on *concepts* that link alternate names and offer a mechanism for interrelation within the content. A CUI, which has no intrinsic meaning, is assigned to each concept. String identifiers are created for each unique linguistic representation of a CUI. In the case of English, an additional object known as a common term identifier is created for every possible lexical variant. Thus, strings are linked to terms, and both strings and terms are linked to concepts. Tools help developers build applications to retrieve all strings or concept names that include specific words or word groups; among the tools are a word index, a normalized word index, and a normalized string index. The latest version of the UMLS contains 730,000 concepts and 1.5 million concept names. Specific code sets of interest are ICD (9 and 10), CPT, and NDC.

Much has been published in the informatics literature regarding evaluation and use of the UMLS, but little is known about its commercial value. Natural language processing (NLP) research has shown the UMLS to be of some use in parsing, extracting, and storing data from free-text reports involving asthma.[12] The authors of that research note that elaborate grammatical analysis and semantic decompositions are required for each diagnostic domain being subjected to NLP processes and that "off the shelf" NLP may be many years in

coming. Debate regarding the notion of UMLS CUIs has led some to conclude that the principal meaning of a CUI should be derived from the source terminology rather than from the structure of the UMLS itself. This so-called extensional meaning within the UMLS is created from the terms that are linked to a CUI and may be different from the intentional meaning of a particular term as it exists in its source vocabularies.[14] Detailed analysis of relationships in the Metathesaurus—many created by a lexical matching process—has highlighted the continuing need for comprehensive manual auditing.[14,17–19] The point is that the Metathesaurus is a valuable resource, but it should not be considered a candidate for becoming a controlled terminology.

HL7 Vocabulary

Health Level Seven (HL7) is an ANSI-accredited, nonprofit, standard-developing organization that creates methods for interoperability of healthcare data interchange. HL7 focuses on clinical and administrative data; its members include providers, vendors, consultants, government groups, and others. HL7 is organized into fourteen technical committees and fourteen special interest groups. The HL7 flagship work is the *Application Protocol for Electronic Data Exchange in Healthcare Environments,* which offers standardized message formats for interchange between healthcare-related applications.

The HL7 standard includes message layouts for such patient administration services as admissions, discharges and transfers; order entry; querying; patient accounting; observation reporting; scheduling; referrals; medical records management; and many more.[20]

Both the 2.3 and 3.0 versions of HL7 use messages to transmit information. A message, as specified by the protocol, includes multiple segments that are sent in a particular order among systems. Each segment represents a collection of data elements that are in turn composed of domains (the set of appropriate coded values for a data element). This HL7 standard specifies the order of the data elements and segments, thus enabling applications to share data.

HL7 acknowledges that the version 2.3 standard falls short of achieving true interoperability among application systems in part for lack of a standardized vocabulary.[21] To effectively communicate, applications need to share the same sets of coded values. Otherwise, disparate applications could send messages to each other but not understand the meaning of the messages and be unable to take appropriate action. The standards group will eventually locate already defined standard vocabularies instead of using only new values coded by HL7.

The Vocabulary Technical Committee of HL-7 also developed a set of principles to help organizations using HL7 determine which standard vocabulary sets to use:

- The terminology must be compliant with the HL7 message structure.
- The terminology provider must be willing to participate with HL7 efforts.
- An organization needs to be responsible for maintenance and update of the terminology.
- Terminology license fees should be reasonable and proportional to their use.
- The terminology should be comprehensive.
- The use of certain terminologies may be mandated by regulatory agencies.

The HL7 model includes several tables to help an organization choose and use a specific vocabulary domain within its HL7 framework. These tables include vocabulary domain definition, vocabulary domain relationship, source vocabulary domain representation, and observation identifier to vocabulary domain linking. They define vocabulary domains, concepts, and relationships.

Logical Observation Identifier Names and Codes

Logical Observation Identifier Names and Codes (LOINC, pronounced "LOW-ink"[22]) is an example of how a modern, concept-based terminology can solve system integration problems and thus find widespread acceptance in the healthcare IT community.[23]

LOINC creates a set of standard codes for observations made by providers during the process of care. By far, its largest and most important domain is that of laboratory tests, but it also includes such clinical observations such as vital signs, intake and output measurements, temperature, EKG tracing measurements, and echocardiogram measurements. The LOINC code is used to identify the particular observation made.

The terminology is constructed along six axes: component measured, specimen type (or system), precision of measurement (qualitative, quantitative, or ordinal), the method by which the measurement was made (important and thus specified for many LOINC concepts), timing (point in time, twenty-four-hour collection, and so on) and the kind of property being measured (mass concentration, mass rate, and others). (A sample code for a serum sodium test is 2951–2.) The developers of LOINC have made an agreement with the College of American Pathologists to coordinate their efforts.[24] The fruits of this collaboration are evident in the latest beta release of SNOMED-RT, in which a supplementary file of LOINC concepts are defined and related to other concepts in the SNOMED-RT hierarchy.

The Medical Entities Dictionary

The Medical Entities Dictionary (MED) is largely the work of a pioneer at the Columbia Medical Center in New York City. The MED is a luminary example of an early CMT.[25] The original purpose of the MED, not surprisingly, was to integrate disparate, departmental legacy systems with a clinical data repository

at Columbia-Presbyterian Medical Center. The design criteria are, in fact, as much involved with medical information processing requirements as they are with those of terminology maintenance.[26] One of the main functions of the MED is translation of nonstandard, proprietary, legacy-system–specific codes into uniform, standard, nonproprietary codes.

Despite significant contributions to the theory of how modern CMTs should be developed, the MED has not yet become a standard itself because it is a solution unique to a single medical center. An experiment to determine whether the MED can be used at other medical centers essentially suggests that a standard CMT rather than the MED is needed.[27]

Consumer Terminology

Despite the growing trend in consumer access to health information through use of the Internet, and despite a growing volume of medical reference material available for consumers, little systematic effort has been made to develop a complete consumer-oriented medical terminology, even though the need for a consumer-oriented terminology has been debated. What happens when patients are able to access their clinical records online? Will the problem list, treatment plan, and test results be represented in a way that is understandable to them? What happens when patients begin to use online applications that enable them to report their own health history? Will they report this information in a way that is clinically acceptable? Will it be machine-readable? Will the information be linked to accepted clinical terminologies?

The differences of education, socioeconomic status, culture, and language among patients and healthcare professionals can be barriers to the process of health information gathering, management, and care.[28–30] Patients often present quite a different perspective from that of healthcare professionals on what is important in the healthcare encounter.[31] In addition, the vocabulary used by laypeople can differ significantly from that of healthcare professionals—even when they are both referring to the same health concept.[28,32]

With the new Internet-enabled e-health environment, patients are realizing the promise of access to their own health records, access to insurance information, credible health content, and greater clinical and billing efficiency. To enable patient participation, however, the words of the patient must be treated with as much respect as the words of the healthcare professional.

Conclusion

One can see from this review that a CMT is essential for moving healthcare into an automated, IT-based system that can actually improve quality, reduce errors, and enhance knowledge about well-being and disease. Current coding systems (ICD, CPT, APC), with the exception of SNOMED and LOINC, fall far short of the ideal characteristics of CMT but will be around for some time to facilitate

billing and satisfaction of HIPAA requirements for transaction processing. Careful analysis of any clinical information system to be built or purchased should include in-depth investigation of the nature of its underlying vocabulary structure and code representation. This review is intended to give a basis upon which to begin that very important investigation.

References

1. http://www.cdc.gov/nchs/.
2. Kohn, L., and others (eds.). *To Err Is Human: Building a Safer Health System.* (Committee on Quality of Health Care in America, Institute of Medicine.) Washington, DC: National Academy Press, 1999.
3. Classen, D. C., and others. "Adverse Drug Events in Hospitalized Patients: Excess Length of Stay, Extra Costs, and Attributable Mortality." *Journal of the American Medical Association,* 1997, *277,* 301–307.
4. Bates, D. M., and others, for Adverse Drug Event Prevention Study Group. "Incidence of Adverse Drug Events and Potential Adverse Drug Events: Implications for Prevention." *Journal of the American Medical Association,* 1995, *274,* 35–43.
5. Wendel, S. "Medical Errors." *Healthcare Informatics, 17*(2), 37.
6. http://www.fda.gov/cder/drug/default.htm.
7. *Federal Register,* Apr. 7, 2000, 65(68), 18433–18482.
8. CHPS Consulting. http://www.chpsconsulting.com.
9. Chute, C., Cohn, S., Campbell, K., Olive, D., and Campbell, J. R. "The Content Coverage of Clinical Classifications." *Journal of the American Informatics Association,* 1996, *3,* 224–233.
10. Campbell, J. R., Carpenter, P., Sneiderman, C., Cohn, S., Chute, C., and Warren, J. "Phase II Evaluations of Clinical Coding Schemes: Completeness, Taxonomy, Mapping, Definitions, and Clarity." *Journal of the American Informatics Association,* 1997, *3,* 238–251.
11. Kudla, K., and Rallins, M. "SNOMED: A Controlled Vocabulary for Computer-Based Patient Records." *Journal of AHIMA,* 1998, *69*(5).
12. Sager, N., Lyman, M., Bucknall, C., Nhan, N., and Tick, L. "Natural Language Processing and the Representation of Clinical Data." *Journal of the American Informatics Association,* 1994, *1,* 142–160.
13. Campbell, K., and others. "Representing Thoughts, Words and Things in the UMLS." *Journal of the American Informatics Association,* 1998, *5*(5), 421–431.
14. Campbell, K., Oliver, D., and Shortliffe, E. "The Unified Medical Language System: Toward a Collaborative Approach for Solving Terminologic Problems." *Journal of the American Informatics Association,* 1998, *5,* 12–16.
15. *UMLS Knowledge Sources.* (11th ed.) Washington, D.C.: National Library of Medicine, National Institutes of Health, U.S. Department of Health and Human Services, Jan. 2000, p. 1.
16. http://www.nlm.nih.gov/research/umls/metab.html#b2.
17. Cimino, J. "Auditing the Unified Medical Language System by Semantic Methods." *Journal of the American Informatics Association,* 1998, *5,* 41–51.
18. Hales, J., Schoeffler, K., and Kessler, D. "Extracting Medical Knowledge for a Coded Problem List Vocabulary from the UMLS Knowledge Sources." *Proceedings of the American Medical Informatics Association Annual Symposium.* Bethesda, Md.: American Informatics Association, 1998.
19. Humphreys, B., McCray, A., and Cheh, M. "Evaluating the Coverage of Controlled Health Data Terminologies: Report on the Results of the NLM/AHCPR Large Scale Vocabulary Test." *Journal of the American Informatics Association,* 1997, *4*(6), 484–500.
20. http://www.hl7.org/about.

21. Huff, S. M., and others. "A Proposal for Incorporating Health Level Seven (HL7) Vocabulary in the UMLS Metathesaurus."

22. Huff, S. M., and others. "Development of the Logical Observation Identifier Names and Codes (LOINC) Vocabulary." *Journal of the American Informatics Association,* 1998, *5*(3), 276–292.

23. http://www.cdc.gov/od/hissb/docs/elr-1999.pdf.

24. http://www.cap.org/html/publications/archive/side1_699.html.

25. Cimino, J. J., Hripsack, G., Johnson, S. B., and Clayton, P. B. "Designing an Introspective, Multipurpose, Controlled Medical Vocabulary." *Proceedings of the 1989 Symposium on Computer Applications in Medical Care,* 1989, 513–518.

26. Cimino, J. J. "Knowledge-Based Approached to the Maintenance of a Controlled-Medical Terminology." *Journal of the American Informatics Association,* 1994, *1.*

27. Kannry, J. L., and others. "Portability Issues for a Structured Clinical Vocabulary: Mapping from Yale to the Columbia Medical Entities Dictionary." *Journal of the American Informatics Association,* 1996, *3*(1), 66–78.

28. Wood, M. L. "Naming the Illness: The Power of Words." *Family Medicine,* 1991, *23*(7), 534–538.

29. Hadlow, J., and Pitts, M. "The Understanding of Common Health Terms by Doctors, Nurses and Patients." *Social Science and Medicine,* 1991, *32*(2), 193–196.

30. Shapiro, J., and Lenahan, P. "Family Medicine in a Culturally Diverse World." *Family Medicine,* 1996, *28,* 249–255.

31. Henry, S. B., and others. "A Comparison of Problem Lists Generated by Physicians, Nurses, and Patients: Implications for CPR Systems." *Proceedings of the Annual Symposium on Computer Applications in Medical Care,* 1995, 382–386.

32. Spiro, D., and Heidrich, F. "Lay Understanding of Medical Terminology." *Journal of Family Practice,* 1983, *17*(2), 277–279.

About the Authors

Jeffrey S. Rose, MD, is a physician executive with the Care Transformation Consulting Practice, Cerner Corp.

Bruce J. Fisch, MD, is chief medical officer of Health Language Inc., Aurora, Colorado.

William R. Hogan, MD, is medical director of Health Language, Inc.

Brian Levy, MD, is vice president of Health Language, Inc.

Philip Marshall, MD, MPH, is director of products, WellMed, Inc.

David R. Thomas, MD, is physician program manager, Clinical Information Solutions, at IBM.

Debra Kirkley, PhD, RN, is vice president of product management at Health Language, Inc.