Comments and a Response to “The Tyranny of the Diagnosis Code” by Slee and Colleagues

David C. Kibbe, MD, MBA

I’d like to heartily applaud the article, “Tyranny of the Diagnosis Code,” by Virgil Slee, Debora Slee, and HJ Schmidt, which offers us a wealth of insight and historical perspective on medical coding. I welcome this thoughtful essay especially because it comes to us just as the United States healthcare system is embarking on what I call the Era of Medical Measurement, and the authors remind us of just how much is at stake for physicians and medical practices in the choices that are made about information management and components of health information technology systems.

Let me digress slightly. The American Academy of Family Physicians (AAFP) has publicly gone on record as being supportive of plans to improve the quality of healthcare through the collection, storage, analysis, and reporting of clinical quality and efficiency measures. In a recent letter to leaders of Congress, for example, we stated: “We believe that the medical profession has a professional and ethical responsibility to engage in activities to continuously improve the quality of care provided to patients. The Institute of Medicine’s landmark report, Crossing the Quality Chasm, documented significant gaps in healthcare quality and issued a challenge to the profession to work collaboratively to improve quality, safety and access. Our organizations accept this challenge…” The AAFP was joined in this letter by the three other major primary care specialty organizations, the American Academy of Pediatrics (AAP), the American College of Physicians (ACP), and the American College of Obstetricians and Gynecologists (ACOG). Together, these physician membership organizations represent over 250,000 doctors in practice in the United States, and their members’ work accounts for over 60% of all ambulatory care and outpatient visits in the United States each year.

This unanimity of opinion from the primary care medical specialty societies on the issue of quality is but one of many indications and signals that lead me to believe we are fast approaching a national consensus that the time has come to implement a uniform, all-payer quality and performance measurement program for the multiple purposes of achieving quality improvements, supporting value-based purchasing (also known as pay-for-performance), and enhancing public accountability.

While not wishing to minimize the large-scale nature and complexity of such a program, I have observed growing optimism from the public, federal and state governments, physicians and their organizations, private sector health plans, the business community, and many others about the potential to accomplish this aim in the near term. There appears to be a new willingness of the parties to engage collaboratively in a new frontier for healthcare in which measurement of performance plays a central role and also a new confidence in the capability of the standards, methods, and technologies now available, or nearly within our grasp, which are so necessary to the accomplishment of such a complex effort.

One powerful example of this trend is the recent consensus reached by the Ambulatory Care Quality Alliance (AQA) on standardizing ambulatory care quality and performance measures. The AQA is a multi-stakeholder group representing physicians, health plans, employers, the Centers for Medicare and Medicaid Services (CMS), patient advocacy organizations, and many others. The AAFP was one of AQA’s initial conveners, along with the American College of Physicians (ACP), America’s Health Insurance Plans (AHIP), and the Agency for Healthcare Research and Quality (AHRQ). In June of 2005, within six months of being formed, the AQA reached consensus on a national “starter set” of 26 evidence-based clinical performance measures for the ambulatory care setting. The starter set of 26 measures is intended to provide clinicians, consumers, and purchasers with a single set of quality indicators that may be utilized for quality improvement, public reporting, and pay-for-performance programs. It is comprised of prevention measures for cancer screening and vaccinations; measures for chronic conditions, including coronary artery disease, heart failure, diabetes, asthma, depression, and prenatal care; and two efficiency measures that address overuse and misuse.

Uniformity with respect to the performance measures employed by Medicare, Medicaid, and the many private health insurance plans across the country is an essential foundation for making meaningful comparisons of physician performance across geographical regions and across different payer groups in any area and a prerequisite for standardized use of physician-
level data for value-based purchasing programs (also known as pay-for-performance) by Medicare. Having a single set of measures that can be reported by a practice to different health plans with which the practice is contracted is also critical to reducing the cost burden of reporting measures borne by individual medical practices. Therefore, the consensus agreement reached by AQA in June, 2005 was a genuine milestone in the progress toward a national quality and performance measurement program.

There are clear signs from CMS that a federal mandate for submission of quality and efficiency measures from medical practices, tied to differential reimbursements for care of Medicare beneficiaries, will be initiated as early as 2006. A spate of recent and mostly bi-partisan health information technology legislation, such as the “Wired for Health Care Quality Act,” co-sponsored by Senators Frist, Kennedy, Enzi, and Clinton, contain provisions that require medical practices to routinely submit quality and efficiency data as a consequence of receiving federal funding through grants or loans associated with local, state, or regional health information technology initiatives. Meanwhile, dozens of pay-for-performance programs have sprung up in the private sector, sponsored by individual and private health plans, insurance companies, and corporate payers of employee healthcare bills. The largest and most often cited of these are the Integrated Healthcare Association’s (IHA) pay-for-performance program in California—the largest of its kind to date involving seven major health plans, over 215 medical groups, and over six million HMO enrollees—and the Bridges to Excellence program, which has clearly demonstrated that EHRs in small and medium size medical practices can be effectively used to collect, analyze, and report quality and performance data. But there are many other similar programs in more than 25 states.

All of this measurement activity at the physician level rests on the foundation of existing medical vocabularies and coding systems, which the Slee article discusses. If I am correct that we are about to engage in a large-scale national effort to aggregate data from administrative, (e.g., billing and clinical information produced from health insurance claims systems and EHRs), then there is no question that the great majority of these data will be encoded in ICD-9-CM (and I would include mention of CPT-4 and SNOMED), a set of codes that Slee et al. characterize as classification output codes derived primarily for the purposes of financial transactions and hospital statistical analysis, rather than being input codes describing what doctors and patients actually experience.

Just the thought of building a national system of outcomes measurement, which not only may affect future health policy in this country, but may also determine how much doctors get paid, on such a shaky foundation as ICD-9-CM should give all physicians reason to pause and contemplate where we are headed. I expect that at the present time, there is neither the political nor economic will to reverse course and take up the solution that Slee et al. propose, namely the establishment of a new super coding set and standard vocabulary capable of diagnosis entity and episode identification. We may have yet to experience the untoward and unintended consequences of measuring things falsely or inaccurately in enough programs and projects to be able to see the wisdom in the course that Dr. Slee and his colleagues recommend.

However, I for one do think we’ll get there eventually. Britain’s pay-for-performance program for general practitioners using the Quality Management and Analysis System is beginning to generate anecdotes that suggest changes in physician and practice behaviors may not be precisely those sought after. For example, it may be that physicians and nurses are now shifting their focus of attention, perhaps only slightly, from caring for the problems patients bring to them, toward fulfilling the tests and other measures determined by the diagnoses already carried by patients, (eg diabetes, congestive heart failure, etc.). In other words, the measurement system may have the unintended consequence of encouraging more attention to those diagnoses for which measurement is attached, and less to those where there are no measurements. It is the latter activity that earns British physicians significant bonuses through the National Health Systems’ Service’s new pay-for-performance program. Since the ICD-9-CM and ICD-10-CM coding systems are firmly embedded in this exercise, we may ultimately learn from the British experience—and from our own—that which Slee et al. warn us about: measuring the wrong information will not make our healthcare system right.

REFERENCES


2 Information on the Bridges to Excellence Program is available at: http://www.bridgestoexcellence.org/bte/.

3 Personal correspondence with Kevin Peterson, MD, after his recent 2005 visit to England, during which he visited a number of primary care practices.