The Case For National Health Data Standards

State data systems illustrate how far we’ve come and the challenges that lie ahead in the development of national health data standards.

by Rosanna M. Coffey, Judy K. Ball, Meg Johantgen, Anne Elixhauser, Patrick Purcell, and Roxanne Andrews

PROLOGUE: “Knowledge,” observed Samuel Johnson, “is of two kinds. We know a subject ourselves, or we know where we can find information upon it.” By this definition, many key players in health care might have to admit to some ignorance on the subject, particularly those who depend on the disparate data sources that constitute the “system” of health care information. A number of promising efforts are now under way that would standardize data collection and reporting, thereby improving the usefulness of health information. However, Rosanna Coffey and colleagues urge caution: They note that the development of standards is but one step in a much longer journey.

The authors should know as well as anyone what challenges lie ahead. All were deeply involved in the design and implementation of the Healthcare Cost and Utilization Project, a national database developed by the Agency for Health Care Policy and Research (AHCPR) that assembles information on hospital inpatient services from state governments and private organizations in almost twenty states. Coffey, who is currently a vice-president at The MEDSTAT Group, recently left AHCPR after twenty-three years of government service. Judy Ball is a senior researcher at AHCPR; she is also serving with the Office of the Assistant Secretary for Planning and Evaluation on the implementation of health data standards mandated by the Health Insurance Portability and Accountability Act. Meg Johantgen is currently a clinical associate with HSS, Inc., a Bethesda, Maryland–based software engineering firm. Anne Elixhauser recently joined MEDTAP International as a senior researcher after five years with AHCPR. Patrick Purcell is an analyst with the Congressional Research Service. Roxanne Andrews is a research scientist at the California Office of Statewide Health Planning and Development.
ABSTRACT: The Health Insurance Portability and Accountability Act of 1996 (HIPAA) contains groundbreaking provisions to encourage the development of a national health information system through the establishment of standards. This paper compares statewide inpatient data systems to one standard—the Uniform Bill (UB)—to understand how standards have been used and how they can be improved. We recommend changes to the UB, note the need for better compliance, and suggest new standards for common, derived elements.

The great Baltimore fire of 1904 burned for two days, destroying more than 1,500 buildings over seventy city blocks. Although water was plentiful and fire fighters with equipment from surrounding cities responded quickly, their efforts were thwarted when few hoses would fit the Baltimore hydrants. Later the same year, a fire erupted near the National Bureau of Standards. Although a night watchman retrieved hoses from two Bureau buildings, the hoses would not fit together because of differences in the threads.

Emarrassed that their hoses lacked standard couplings, the Bureau investigated and found over 600 variations in fire hose couplings across the country. Fire associations, long advocates of standards, welcomed the Federal support, and agreement on national standards was reached quickly. However, adoption of the standards took decades. The expense of conversion or replacement and the limited need to combine equipment contributed to the slow adoption. Thirteen years after standards were established, only a handful of cities in the U.S. had complied with them.¹

Like the fire hydrants and hoses at the turn of the century, health data today adhere to no single standard. For three decades rising national health expenditures have consumed an ever-greater share of the U.S. gross domestic product, and every new approach to control costs and improve quality has accentuated the value of health care data.

As the varied uses of health care data have grown, so has the need for standards. Groups that have actively contributed to the development of more consistent health care data include the National Committee on Vital and Health Statistics (NCVHS), which has advised the secretary of the U.S. Department of Health and Human Services (HHS) on data issues for decades; the National Uniform Billing Committee (NUBC), which has been responsible for the standards of the Uniform Bill (UB), first in 1982 and then in 1992; and the American National Standards Institute (ANSI) X12N Committee, which has coordinated industry agreements on standards for electronic data exchange for third-party reimbursement.² However, these standards have accommodated different rules for different payers and were modified often to meet local information needs. There has been no single standard.

The purpose of this paper is to advance the development of health data standards. We examine statewide hospital inpatient data systems to better understand how health care data deviate from accepted standards.³ This study was stimulated by statewide data organizations requesting guidance on how to achieve greater uni-
formity in data across states, an interest that arose after these organizations began to realize benefits from combining data from multiple states. Lessons from these data systems highlight the challenges facing those who will implement the newly enacted federal mandate for the development of health data standards.

**The Federal Mandate For Health Data Standards**

The Health Insurance Portability and Accountability Act (HIPAA), which was enacted 21 August 1996, is best known for health insurance reform: It guarantees portability of health insurance between jobs and restricts denials of coverage based on preexisting conditions. However, it also contains lesser-known groundbreaking provisions for “administrative simplification” (Title II, Subtitle F): “to improve the . . . efficiency and effectiveness of the health care system, by encouraging the development of a health information system through the establishment of standards and requirements for the electronic transmission of certain health information.”

To accomplish this, the secretary of HHS is directed to adopt standards for financial and administrative transactions and associated data elements, including code sets for clinical nomenclature. The transactions include health care claims and equivalent encounter information, enrollment, disenrollment, eligibility, payments, remittance advice, premium payments, first report of injury, health claim status, referral certification and authorization, and other transactions to be determined by the secretary (sec. 1173[a][1]).

To facilitate information sharing, the secretary is also directed to adopt standards for a “unique health identifier for each individual, employer, health plan, and health care provider” (sec. 1173[b][1]) and for security procedures to protect the integrity and confidentiality of information and to protect against unauthorized uses and disclosures (sec. 1173[d]).

The secretary has eighteen months (until February 1998) to adopt standards that will apply to health plans, health care clearinghouses, and health care providers that transmit health care transactions in electronic form (sec. 1172[a]). Compliance with the standards by health plans is required within two years of adoption of the standards (three years for small health plans) (sec. 1175[b][1]). HIPAA preempts state laws to the contrary (sec. 1178). The NCVHS is designated specifically as adviser to the secretary on these matters (sec. 1172[f]).

These provisions are the culmination of a long period of bipartisan agreement about the need for standards as a way to reduce the administrative costs of health care reimbursement, improve competition through disclosure of information, enhance performance measurement, and improve quality of care. This initiative will expe-
dite the development of a national health information system, which has been inhibited, in part, by the lack of standards for defining data elements, coding data, defining data file structures, and exchanging data electronically.

**Study Methods**

State governments and private organizations in more than forty states collect relatively uniform health data on inpatient services provided by hospitals. Despite varying purposes for collecting statewide data, the types of data collected are similar enough that these statewide data systems have been used to assemble a national database, the Healthcare Cost and Utilization Project (HCUP-3), sponsored by the Agency for Health Care Policy and Research. The challenge for HCUP-3 was similar to that of the HHS secretary: how to get data from diverse systems to follow one standard model.

For this study, we used data layouts and manuals for 1993 inpatient data from ten state data organizations and two statewide hospital associations participating in HCUP-3. These twelve statewide data systems are not typical. Rather, they represent some of the most advanced and largest inpatient data systems in the country. Nine of the twelve statewide systems have been in operation for ten or more years; the twelve states collectively represent 51 percent of all U.S. discharges and 37 percent of all U.S. hospitals. Information abstracted into detailed summary tables was reviewed by representatives of the twelve data sources to ensure accuracy of the information for their databases in 1993.

We compare the content of statewide systems with that of the UB-92 because that is the standard on which the state systems were based. The ANSI X12N standards for inpatient insurance claims have adopted the content of the UB-92. It is important to understand that we are analyzing the content of inpatient claims and not the architecture or format for electronic data interchange (EDI).

**Results**

Our comparisons of specific inpatient data elements from statewide systems to UB-92 standards revealed three areas for improvement. (1) Modifications to some UB data elements would better serve the information needs of health care policymakers and improve the uniformity of health data nationally. (2) Compliance with UB standards by statewide data systems would make data more comparable and easier to share. (3) Consensus on common, derived elements that are not part of the UB would enhance the ability to combine data across state systems. For each of these areas, we describe why the specific data elements are important to health care policymak-
ers, how states deviate from UB standards, and recommendations for improvements. This is followed by a discussion of patient and provider identifiers, which were addressed separately by HIPAA.

**Modifications.** Patient gender, ZIP code, race, and ethnicity are important components of health databases for analytic and administrative purposes. They are recorded in various ways by statewide data systems. In most cases, minor modifications to the UB categories plus improvements in coding instructions could reduce this variation so that data from states could be combined without the large investments required today to make the data uniform.

*Patient gender.* Patient gender is important because male/female differences in incidence of certain medical conditions and use of health services make risk adjustment essential. Recording patient gender should be straightforward, but we found five different coding schemes across the twelve states. The UB standard—M (male), F (female), and U (unknown)—which is probably a relic of paper forms, is followed by only three states. Most states use numeric values, but absent a numeric standard have adopted different approaches. Consideration should be given to requiring numeric values, which are clearly preferred.

*ZIP codes.* Health care policymakers use patients’ ZIP codes to define health care markets; to examine geographic variations in health care costs, quality, and access; and to track public health problems. In addition, ZIP codes can be linked to other data (for example, Census data) to obtain demographic characteristics of patients’ communities, which are generally available only from household surveys.

ZIP code is one of the most consistently defined data elements because it is an established part of the postal address system with no widely recognized alternative. All twelve statewide data systems report patients’ five-digit ZIP codes; five states report nine-digit ZIP codes, which are permitted but not required by the UB. Inconsistencies arise when statewide systems require foreign nationals or homeless persons to be identified in the ZIP-code field. As a result, no single method is used in more than two states.

We recommend adoption of a method for identifying foreign nationals and homeless persons in the ZIP-code field. Administrative data could then be used to assess the foreign demand for U.S. health care and the health care needs of the homeless, including rising public health problems such as tuberculosis and human immunodeficiency virus (HIV) infection.

*Race and ethnicity.* Definition of race and ethnicity generates much discussion and controversy. Many believe that racial categories are archaic, unscientific, and even racist. Others believe that advances
toward equality in health care could not have happened without measurements of inequality by race or ethnic group. Yet others believe that the concept of ethnicity is too complex to be captured by the few choices commonly used. Despite the opposing views, many government agencies that collect health data have legislative mandates to collect data on race and ethnicity, primarily for use in analyses of social, economic, and health trends.

The UB-92 does not include race or ethnicity because neither is necessary for adjudicating claims. Race and ethnicity data should be collected once, directly from the beneficiary, rather than indirectly through observation by providers at each health care encounter. The data should be available in separate enrollment databases. However, for statewide data systems, linking enrollment and claims data for numerous health plans is costly, and states usually forgo the cost and the information. Other states require hospitals to collect race/ethnicity data and submit it to the state with UB data.

Without a standard, statewide data systems vary considerably in their collection of race and ethnicity data. Nine of the twelve states collect race data (Arizona, Illinois, and Washington do not). Despite the rapid growth of the Hispanic population in the United States, only five of the twelve states (Florida, New Jersey, New York, Pennsylvania, and Wisconsin) collect data on Hispanic ethnicity.

Notwithstanding the legitimate controversies, race and ethnicity should be included in every statewide data system. Some mechanism must be devised to combine claim and enrollment data elements needed for policy analysis. Failing to collect such information leaves us unable to detect intergroup differences in treatments or incidence of disease or to identify strategies to care for underserved populations. We recommend adoption of the standard established by the U.S. Office of Management and Budget. This standard for race has been used for almost two decades throughout the federal government in the decennial census, in other surveys, in data collection meeting statutory requirements, and in other program monitoring.

Compliance. Some UB data elements are unambiguous in definition and straightforward in coding. Yet some statewide systems diverge from the standard. We list examples of noncompliance to illustrate areas in which statewide data organizations might improve the compatibility of their data with that from other systems.

Diagnoses and procedures. Diagnoses and procedures summarize the clinical content of a patient’s medical record, including the reason for admission, other conditions identified during the hospital stay, and treatments provided. This is usually the only clinical information in administrative databases upon which payments are based. In addition, this information is often used to classify hospital stays for
comparisons of the cost and quality of care among hospitals and, increasingly, among physicians. Whether clinical information from the medical record is coded accurately cannot be addressed here, but how these data elements are handled in statewide databases poses other threats to their accuracy.

Diagnoses and procedures are always recorded in statewide inpatient data systems as alphanumeric values according to the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM). However, handling is not consistent. Eleven of the twelve statewide data systems record diagnosis and procedure codes with implicit decimals (for example, “25000” represents “250.00”); in one system the decimal is explicit. Two data systems pad blank positions on the right with zeros (for example, zero-padding makes “410_ _” become “41000,” an invalid code).

Although zero-padding can be identified and undone, it requires a costly iterative process of truncating and verifying until a valid code is achieved and always runs the risk of creating a valid code from a data-entry error. To avoid these problems, we recommend standard instructions that specify left-justified alphanumeric values, implicit decimals, and no zero-padding.

Discharge status. Discharge status, or “disposition of patient,” describes the arrangement or event ending a hospital stay. It is used to record death during the hospital stay or whether the discharge is routine (to home) or to another facility. From this information inpatient mortality rates can be calculated; the effects of health policies and programs on referral patterns among acute, subacute, and long-term care facilities can be assessed; and increases in the use of home health care following hospitalization can be monitored.

Most of the twelve state data systems we examined included the UB discharge status categories—for example, discharged to home (routine), to short-term general hospital, or to skilled nursing facility. In addition to ten UB categories, states add six other categories, such as discharged to another hospital for specific types of care. The values used to represent the UB categories also differ among the states.

We recommend two reassessments of this data element. First, the NCVHS should determine whether its categories are consistent with how health care is organized today. Second, statewide data organizations should advise the NCVHS on state-specific needs that should be reflected in the national standard.

**Consensus.** A few data elements derived from the UB-92 are so important that they also should be given standard definitions. First are patients’ age and length-of-stay, elements used to replace birth, admission, and discharge dates to protect patients’ privacy in publicly released data sets. Second are type of payer and charge
categories, derived elements essential for making cost and quality comparisons.

Patients' age. Patients' age is essential for identifying and studying subsets of the hospitalized population (newborns, children, and the elderly) and for basic age/sex risk adjustment for differences among patients. States derive patients' age in two ways: as the date of birth to the date of hospital admission; or as the date of birth to the date of hospital discharge. Fortunately, most hospitals stay are short, so the reference point for calculating age is usually unimportant. However, consistent rules and four-digit years for dates would avoid errors. Also, given privacy issues that restrict release of dates, data organizations should agree to define age using a consistent reference point.

Length-of-stay. Length-of-stay is the standard used by every statewide data system for comparing resource use among hospitalized patients. Like patients' age, length-of-stay is derived by at least two different methods. This is most problematic when admission and discharge fall on the same day. In eight states length-of-stay for a “same-day stay” is one day; in four states length-of-stay for such a stay is zero days. In addition, administrative leave days (when recorded) are not handled consistently; sometimes they are counted in the length-of-stay, and sometimes they are not.

When dates are unavailable, consistency in the length-of-stay calculation may be important—for example, for policies directed at short-stay hospitalizations such as deliveries. We recommend adoption of a standard that records same-day stays as zero days to distinguish them from one-day stays that span two calendar days. We also recommend a consistent method for handling administrative leave days—identify but do not subtract them from length-of-stay.

Type of payer. Payer type is important for tracking use and costs of health services under different insurance arrangements, monitoring the impact of changes in government programs on various insured groups, and, with increasing importance, assessing the impact of managed care. Despite its importance for policy analyses, expected payer is among the least-standardized data elements.

HIPAA requires the adoption of a standard unique identifier for every health plan. Policy analysts need a method to classify similar insurance arrangements across thousands of health plan identifiers.
All twelve statewide data organizations have created such categories, but they vary greatly (Exhibit 1). Most states identify separate categories for Medicare, Medicaid, commercial insurance, Workers’ Compensation, and self-pay. Because managed care plans and other government payers are more difficult to categorize, they have spawned numerous variations. Whereas no one state has more than twelve payer categories, the twelve states collectively have forty-two different categories.

We recommend that the NCVHS study options for creating health insurance categories that are useful for policy analysis. As part of this, the NCVHS should give in-depth consideration to a classification for managed care plans. Several issues should be considered. First, the organization of managed care should be considered because different models can have different impacts on patients, provider incentives, and the delivery of health care services. Second, financial arrangements must be considered because the same plan or group sponsor can have different arrangements with different providers for the same group of enrollees. Third, there is a growing need for a method to distinguish public from private managed care patients. When public coverage is provided under contract to private insurers, the insurance cards for public and private beneficiaries are sometimes indistinguishable. Fourth, an update mechanism is essential because the rate of change in third-party arrangements will continue to thwart efforts to keep any payer classification up to date.

Charge categories. Charge categories are also crucial to in-depth assessments of how health care dollars are spent and what types of services comprise a hospitalization. Uncompensated care, diagnosis-based reimbursement, negotiated discounts, and capitation have made charges an imprecise measure of payments to hospitals. However, charges remain one of the only ways to approximate the costs of hospital care and to measure the relative costs of medical treatments. Standard charge categories would be a valuable source of information for the National Health Accounts and for approximating cost of care using departmental cost-to-charge ratios.

All twelve statewide data systems collect total charges, but only seven collect more detailed charges. Of these, nearly every state has a different scheme. Two use the 300+ UB-92 revenue codes that identify specific accommodations and ancillary services. The other five report aggregated categories such as room and board, pharmacy, and laboratory. The number of categories range from five (Colorado) to forty-two (Massachusetts), and there is little consistency in what they contain. For example, a pharmacy category is available from all five states, but there are three combinations of UB-92 revenue codes.
that define the pharmacy content. In short, cross-state comparisons are impossible with the charge categories in use today.

We recommend that the NCVHS consider a standard classification for charges (or payments) that could be used to describe how health care dollars are spent at state and national levels. States that want to make comparisons with other states would benefit from such standardization. The states should be consulted about the classification that makes sense for their uses, such as for health planning and cost containment.

Identifiers. Identifiers have a multitude of uses in the health care system that go beyond simplifying the electronic exchange of information. Any application that requires linking information across multiple sources or across time within one source will benefit from a system of unique identifiers. HIPAA directs the HHS secretary to take into account multiple uses for identifiers and to specify...
the purposes for which a unique health identifier may be used (sec. 1173[b]). Congress recognized that use of individually identifiable information may be appropriate for certain purposes, including health care–related research, which would not compromise individual privacy. Indeed, studies of health outcomes, referral patterns, disease surveillance, and cost of illness may involve multiple admissions, hospitals, health plans, and/or states and require unique national identifiers. With unique identifiers, physician/patient encounters or claims could be linked to physician characteristics, to insurer information, or to records of prior encounters for evaluation of complete episodes of care. Such linkages would reduce the cost of evaluating health care services. The challenge facing the secretary in adopting standard health identifiers and appropriate protections is illustrated by the variations in patient, hospital, and physician identifiers found in statewide data systems.

Patient identifiers. Patient identifiers in the UB-92 include the patient’s name, address, control number, medical record number, and the insured’s identification number assigned by the payer. Statewide databases generally include one type of patient identifier, but the types vary considerably. The identifier may be unique to the admission, to the hospital, to the health plan, or to the state. Two states collect Social Security numbers, which are unique to the individual, nationwide.

Once a unique identifier is chosen, it must be protected, especially in publicly available databases commonly used by state and federal policymakers. All states limit access to individual identifiers, encrypted or not, to those with a need to know them. Some states use encryption to protect individual identifiers while allowing their use for legitimate purposes. However, when a state encrypts an individual identifier, the encrypted identifier becomes unique to an individual only within the state. Thus, encryption methods should be uniform across data systems but impervious to unauthorized decryption. HIPAA directs the HHS secretary to adopt security standards to ensure data integrity and confidentiality and to protect against unauthorized uses or disclosures. We recommend that consideration be given to state-of-the-art protections that preserve the uniqueness of individual identifiers for legitimate policy and research uses, in keeping with congressional intent.

Hospital identifiers. Hospital identifiers vary across statewide data systems. Ten of the twelve systems record a state hospital identifier; for eight, this is the only hospital identifier. New York retains several hospital identifiers: Medicare provider number, Medicaid provider number, insurer’s provider number, and federal tax number. Other states pick and choose from among these.
Hospital identifiers that are specific to the state impede data linkages. For example, to link a state’s hospital discharge data to the Medicare database of hospital providers, one must establish a translation between state identifiers and the Medicare provider number. This is time-consuming and difficult because the state and Medicare often view the institution in slightly different configurations, and because hospitals change—they merge, split, or otherwise alter their identities—which makes the painstaking translations quickly obsolete. The National Provider Identifier (NPI), discussed below, should solve this problem.

**Physician identifiers.** Physician identifiers required by the UB-92 differ by type of payer. Medicare requires the uniform physician identification number (UPIN) and physician name; other payers have their own identifiers or require names only. Among the twelve statewide data systems, four use the UPIN; seven use state medical license numbers; one uses hospital-defined numbers for physicians; and two do not collect physician identifiers at all.

Identifiers for physicians have problems similar to those for hospital identifiers. A physician who practices in multiple states may be identified differently by each state. A single identifier may be used to submit claims for a group of physicians practicing together. HIPAA specifically directs the secretary to consider multiple locations and specialty classifications of physician providers in adopting unique identifiers (sec. 1173[b][2]).

- **The National Provider Identifier.** When fully implemented, the NPI could apply to every individual, group, and organization involved in health care services or supplies. If, after public comment, the HHS secretary chooses the NPI as the standard, then NPIs would be issued to all providers within two years, as required by HIPAA.

**Lessons For Policymakers And Data Organizations**

The story of the fire hoses teaches us that standards alone do not ensure their universal adoption. This paper illustrates real-world deviations from an existing standard and makes recommendations for the national standard. Here we suggest lessons for policymakers charged with developing national standards and ways for statewide data organizations to benefit from them.

- **For policymakers.** Lesson 1: A compelling reason must exist for standards to be developed. HIPAA provides the compelling force to establish national standards for health data. It was enacted with broad industry support. The federal mandate preempts state and payer-specific variability, confers on the HHS secretary the authority previously lacking to resolve the multiple interests and objectives for setting standards for health care data, and specifies dead-
Lesson 2: Strong incentives are the key to offsetting the high costs of change, especially of altering information systems to fit a national standard. HIPAA provides for sanctions on organizations that do not comply with the standards. However, the stronger incentives from HIPAA are economic. Electronic transactions using national standards should reduce the cost of each transaction to the health plan and to the provider, yielding savings that can be passed along as lower charges and lower premiums. Providers and plans that fail to adopt the standards would be at a competitive disadvantage.

Lesson 3: Input from a broad array of data users can lead to more durable standards. HIPAA requires the HHS secretary to consult with private standard-setting organizations and the NCVHS in establishing standards. Open communication will be essential to accommodate local data requirements, to temper the urge to customize, and to comply with the implementation schedule.

Lesson 4: Strategies for updating and expanding national standards are important to their continued relevance and acceptance, especially in a rapidly changing health care system. Payer categories illustrate the confusion that results when standards do not keep pace with the times. Judicious but timely revisions will be important for maintaining confidence in the system of standards.

For statewide data organizations. HIPAA provides the federal mandate for data standards for the purpose of electronic data exchange. Statewide data organizations that assemble data for other purposes can benefit from the requirements of HIPAA, if they assume responsibility for adopting and preserving standardization in their derived data systems.

Recommendation 1: Adopt the national standards, by increments if necessary. Adoption of the national standards will lead to enhanced capabilities for combining data across geographic boundaries for many purposes. How mature statewide systems respond to national standards remains to be seen. This paper illustrates several instances in which real progress toward standards can be achieved by small changes in data processing procedures. HCUP-3 participants who recognize the value of such data and the cost of noncompliance should take the lead in this endeavor.

Recommendation 2: Resist the urge to customize. If customization is necessary, preserve a map back to the standard. This paper provides ample evidence of customization of the UB-92, which occurred because there was no national standard. With HIPAA, both mature and developing statewide systems will be able to benefit by adopting national standards. Of course, local adaptation will be necessary sometimes to meet local needs for information and policy...
development. However, customization need not destroy compatibility with other systems. If finer categories are needed, be certain that these new categories can be reassembled into the standard.

Recommendation 3: Get involved in standard setting. Statewide data organizations can contribute to the successful development of national standards by joining the standard-setting discussions. They can explain the inadequacies of standards when applied to their own uses. If the HHS secretary views derived data elements, such as payer and charge categories, as beyond HIPAA and her (or his) authority, these data organizations may want to form a group to deal with these issues. Statewide data organizations could improve the uniformity of their systems even if existing standard-setting organizations do not address all of their issues.

In conclusion, it would be foolhardy to conclude that HIPAA’s federal mandate is the magic wand for transforming health data to consistency and compatibility. Great strides toward standardized data characterized the three decades before HIPAA became law. At the same time, technologies that could have facilitated standardization have actually resulted in the proliferation, decentralization, and customization of health data systems. HIPAA sets the course for standardizing administrative data for health care. Serious effort will be required to make standardization a reality. The secretary of HHS faces a challenge not only in establishing the standards but also in motivating their use, acceptance, and evolution.

This paper was completed while the authors were at the Agency for Health Care Policy and Research (AHCPR). This paper does not represent the policy of either AHCPR or the U.S. Department of Health and Human Services (DHHS). The views expressed herein are those of the authors and no official endorsement by AHCPR or the DHHS is intended or should be inferred. The authors gratefully acknowledge the stimulus and support of the twelve states and hospital associations represented in this paper and the persons in these organizations who reacted to an earlier version of this paper: Henry Brown (WA), Michael Boyson (CO), Debra Gressel (FL), Kathy McCaffrey (CA), John Noak (IL), and Judith Nugent (WI). The authors thank many others who also shared their knowledge and insights: Jim Blakley, William Buczko, Mark Epstein, J. Michael Fitzmaurice, Marjorie S. Greenberg, Barbara Kurtzig, Ann Meadows, William Sobaski, Barbara Starfield, and several anonymous reviewers.

NOTES
2. For other standard-setting groups and reviews of the use of core data elements for health care encounters, see Report of the National Committee on Vital and Health Statistics: Core Health Data Elements (Hyattsville, Md.: National Center for Health
3. Systems of inpatient hospital data include one record for each completed inpatient hospital stay. Because discharge from the hospital is the event that triggers the inclusion of a record, the data are often referred to simply as discharge data. Because billing is an administrative function for a hospital and discharge data are often derived from billing records, these data are sometimes called administrative data as well.


5. D.N. Mendelson and E.M. Salinsky, “Health Information Systems and the Role of State Government,” Health Affairs (May/June 1997): 106–119; and P. Starr, “Smart Technology, Stunted Policy: Developing Health Information Networks,” Health Affairs (May/June 1997): 91–105. HCUP-3 produces two databases that provide uniform data on hospital inpatient stays: the Nationwide Inpatient Sample (NIS) and the State Inpatient Databases (SID). The NIS is available from the National Technical Information Service at 703-487-4650. The SID contains data on all inpatient stays in seventeen states and is available from the data sources. For more information about the NIS, call 301-594-1406 or send e-mail to hcupnis@ahrpr.gov.

6. Arizona Department of Health Services (30 June 1993); California Office of Statewide Health Planning and Development (August 1994); Colorado Hospital Association (10 February 1993); Florida Agency for Health Care Administration, Center for Health Statistics (12 October 1993); Illinois Health Care Cost Containment Council (First Quarter 1993); Iowa Hospital Association (20 May 1993); Massachusetts Rate Setting Commission (8 July 1994); New Jersey Department of Health, Health Care Planning, Financing, and Information Services (23 August 1994); New York Department of Health, Statewide Planning and Research Cooperative System (SPARCS) (May 1993); Pennsylvania Health Care Cost Containment Council (1 January 1993); Washington State Department of Health, Office of Hospital and Patient Data Systems, Comprehensive Hospital Abstract Reporting System (CHARS) (February 1993); and Wisconsin Office of Health Care Information (14 March 1994).

7. Summary tables are available from the authors upon request. Contact Rosanna Coffey, The MEDSTAT Group, Suite 400, Box 11, 4401 Connecticut Avenue, NW, Washington, DC 20008.

8. U.S. Office of Management and Budget (OMB) Statistical Policy Directive no. 15, “Race and Ethnic Standards for Federal Statistics and Administrative Reporting,” Federal Register (12 May 1977). Current categories for race are white, black, Native American/Alaskan Native, Asian/Pacific Islander, and other; ethnicity categories are Hispanic and non-Hispanic. Criticisms of these categories, particularly for classifying persons of mixed race, have led the OMB to initiate a thorough review of the directive with a final decision on changes expected in 1997 (Federal Register, 60FR 44674, 1995).


10. Some consumer advocacy groups have seriously opposed the use of the Social Security number for a health identifier because of its widespread use for a variety of purposes and its easy availability.