Statewide Hospital Discharge Data: Collection, Use, Limitations, and Improvements

Roxanne M. Andrews

Objectives. To provide an overview of statewide hospital discharge databases (HDD), including their uses in health services research and limitations, and to describe Agency for Healthcare Research and Quality (AHRQ) Enhanced State Data grants to address clinical and race–ethnicity data limitations.

Principal Findings. Almost all states have statewide HDD collected by public or private data organizations. Statewide HDD, based on the hospital claim with state variations, contain useful core variables and require minimal collection burden. AHRQ's Healthcare Cost and Utilization Project builds uniform state and national research files using statewide HDD. States, hospitals, and researchers use statewide HDD for many purposes. Illustrating researchers' use, during 2012–2014, *HSR* published 26 HDD-based articles on health policy, access, quality, clinical aspects of care, race–ethnicity and insurance impacts, economics, financing, and research methods. HDD have limitations affecting their use. Five AHRQ grants focused on enhancing clinical data and three grants aimed at improving race–ethnicity data.

Conclusion. ICD-10 implementation will significantly affect the HDD. The AHRQ grants, information technology advances, payment policy changes, and the need for outpatient information may stimulate other statewide HDD changes. To remain a mainstay of health services research, statewide HDD need to keep pace with changing user needs while minimizing collection burdens.

Key Words. Administrative data, hospital, race, ethnicity, clinical data

Almost all states have data organizations collecting administrative data (discharge abstract or claims records) for all hospitalizations in non-Federal facilities in their state. The data collection efforts cover inpatient stays and, increasingly, emergency department, ambulatory surgery, hospital observation services, and other hospital outpatient services. Although these statewide discharge databases were created for state governments, local communities, and the hospital industry, they have been widely used by health services researchers to examine policy, care delivery, and clinical issues. Nonetheless,

as with all data sources, there are limits to their uses related to data content. To foster successful approaches to improving statewide data, the Agency for Healthcare Research and Quality (AHRQ) awarded eight 3-year grants to improve the clinical content and race–ethnicity information in these databases. The primary goals were to improve the statewide data for local uses and for AHRQ's Healthcare Cost and Utilization Project (HCUP).

The purpose of the article is to provide background information on statewide hospital discharge data and the context for the other articles in this special issue of *HSR* that focus on the products and lessons learned by the Enhanced State Data grantees. This article provides an overview of statewide hospital discharge data, including content and coverage, and its evolution and improvement over time. It describes the HCUP state and national datasets derived from the statewide data. The article highlights the use of these data in health services research and includes a description of state, local, and hospital industry uses. Limitations of the data are also noted. The article concludes with a discussion of the need to improve the clinical detail and race–ethnicity data and a description of the AHRQ-sponsored Enhanced State Data Grant program to improve the data in these areas.

DATA COLLECTION—OVERVIEW

Over the last 40 years, hospital discharge data have become one of the mainstay data sources for health services research. From innovative collection in a handful of states in the early 1970s, these data now encompass almost all states in the United States. There are several reasons for their wide availability and use. Because existing claims data are their foundation, the resources to create the datasets are modest when compared to primary data collection such as surveys or medical record abstraction. Using the hospital claims, standard format (Uniform Bill) minimizes the burden on hospitals to report data, on states to process incoming files, and on analysts to use the datasets. The datasets contain core data elements that are valuable for many types of analyses and applications. Analyses can be done at various levels of aggregation, including discharge/visit, patient (in some states with encrypted identifiers), hospital, physician (in some states), and geographic areas (zip code, county,

Address correspondence to Roxanne Andrews, Ph.D., Center for Delivery, Organization, and Markets, Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850; e-mail: roxanne.andrews@comcast.net.

state, region, nation). The datasets include records for all-payers, including the uninsured, and generally include all non-Federal acute care hospitals in a state. AHRQ's HCUP transforms the statewide data into uniform research files to facilitate multistate analyses. HCUP also develops national datasets from these data for national estimates (AHRQ 2015). Numerous software and standard methodologies are available to facilitate analysis, including clinical groupers, risk adjustment/severity of illness methodologies, quality of care measures, and economic measures. Online query systems provide for easy access to statistics generated from the data. Linkages to other datasets (through hospital, geographic, patient, or physician identifiers) expand the analytic capacity of statewide data.

Statewide discharge data collection practices are similar across states, but many variations exist (Love, Rudolph, and Shah 2008). State law often mandates collection, but in some states collection is voluntary. The statewide data organization (SDO) is generally a state government agency, but it may also be a state hospital association or other private data organization. The data are a summary record for each hospital stay with such information as patient demographics (age, sex, zip code of residence), clinical information about the hospitalization (e.g., diagnoses, procedures performed, length of stay), billing information (e.g., expected payer category, amount charged), and the hospital identity. The foundation for the typical statewide discharge database is the national hospital claims standard—the Uniform Bill (currently the UB-04).

States differ in UB data elements included in their statewide discharge data, based on local needs. Many, but not all, states include physician identifiers, patient identifiers (deidentified in released data), and data elements from the line item detail on the UB (e.g., use and charges for individual services such as intensive care units). A few states (such as California) have legacy systems that have some commonalities with the UB, but they have many variations from it. Because the purpose of the UB is for payment, and not public health and research, all SDOs add data elements not contained in the UB, in particular expected payer category. Most states add race–ethnicity and a few states add other data such as the patient's primary language, do-not-resuscitate indicators, mother's medical record number (on newborn record), and newborn birthweight. One state (Pennsylvania) has collected additional detailed clinical information, or specific laboratory findings, to support quality reporting for nearly 20 years. Generally hospitals are required to submit all the data elements the SDO collects, but in some states hospital submission of specific data elements is voluntary (e.g., race–ethnicity).

Beginning in the early 1990s, HCUP began a voluntary collaboration with SDOs to leverage their data collection efforts to build uniformly formatted national and state hospital encounter-level datasets for research (AHRQ 2015). HCUP greatly expands the availability and use of the statewide discharge data. The AHRQ statewide discharge and visit-level databases include the State Inpatient Databases, the State Emergency Department Databases, and the State Ambulatory Surgery and Services Databases. After receiving the data from SDOs, HCUP conducts standard data quality checks, creates uniformly formatted files, and, to facilitate research, adds data elements derived from the statewide data and linkages to other data (e.g., AHA Annual Survey). These added data elements include clinical classifications based on the ICD-9-CM (International Classification of Diseases, 9th Edition, Clinical Modification) diagnosis and procedures codes (e.g., Clinical Classification Software and comorbidity measures) and sociodemographic indicators based on patient zip code (e.g., urban–rural measures and median income of the patient's zip code, in quartiles). HCUP also provides linkable files for some states, such as hospital characteristics (from linkage to the American Hospital Annual Survey), cost-to-charge ratios (for each hospital using information derived from Medicare Cost Reports), hospital market structure files (for studies on competition and market forces), and revisit variables (to examine readmissions). HCUP samples the SID and SEDD to create three databases for national estimates—National (Nationwide) Inpatient Sample (NIS) for estimates of inpatient care, Kid's Inpatient Database (KID) for estimates of inpatient care for children, and Nationwide Emergency Department Sample for estimates of ED visits.[1]

DATA COLLECTION—IMPROVEMENT OVER TIME

A key to the continued, and expanded, use by researchers is that the datasets have evolved and improved over time. The number and types of datasets have increased substantially as the value of the data became evident. Initially, a few pioneering states began collecting discharge data in the 1970s and 1980s, primarily to support health planning and hospital rate setting. In the mid-1980s, additional states began collecting the data and the motivation shifted to providing public information in a competitive environment, primarily on costs, but occasionally including quality (Epstein 1992; Epstein and Kurtzig 1995). During the mid-1980s, the Medicare reimbursement system moved to prospective payment for hospitalizations based on diagnosis-related groups

(DRGs), which rely on ICD-9-CM diagnosis and procedure codes. Consequently, hospitals and the government closely monitored coding quality, resulting in improved accuracy (Fisher et al. 1992), though DRG reimbursement incentives may lead to upcoding (Hsia et al. 1992). During the 1990s, the emphasis on public reporting on quality of care grew, almost all remaining states created statewide collection of inpatient data, and many of the established systems responded to the need for information on outpatient services by expanding collection to ambulatory surgery and emergency department visits.

Over time, SDOs modify their discharge data in response to new priorities and revisions to national data standards. The earliest collections in the 1970s and 1980s often used the Uniform Hospital Discharge Data Set (UHDDS) as the standard. The UHDDS was a minimal basic dataset aimed at meeting a variety of analytic needs (National Center for Health Statistics 1980), though it did not include charge or payment information. Many SDOs added charge information to their UHDDS reporting requirements to fill this gap. During the late 1980s and early 1990s, most states transitioned to the UB standard, which they considered readily available from hospitals, and which included charge and other cost-related information that was valuable for the increasing focus on cost containment (Larks 1986). SDOs generally change their submission requirements when the National Uniform Billing Committee (NUBC) updates the UB. The UB updates occur as often as several times a year and include two major revisions since the early 1980s (from the UB-82 to UB-92, and to UB-04). The diagnosis and procedure code sets used in the data —ICD-9-CM and CPT® (Current Procedural Terminology, for procedures on some outpatient records)—change regularly (annually or sooner). Beginning in October 2015, hospitals will transition from the ICD-9-CM coding system to ICD-10-CM (International Classification of Diseases, 10th Edition, Clinical Modification) for diagnoses and ICD-10-PCS (Procedure Coding System) for procedures.

When SDOs innovate by collecting data elements not on the UB, their innovations serve as tests of the feasibility and usefulness of the novel data. Other SDOs may implement the successful innovations and national standards organizations may adopt them. For example, in 1991 just 14 SDOs (of the 31 existing SDOs) collected race data. By 2008, this had increased to 43 states (of 47 SDOs) (Andrews 2011). Similarly, in the mid-1990s two pioneering states—California and New York—collected present on admission indicator (POA) for each diagnosis on the discharge record to identify complications arising during the stay. Ten years later (2007) the

NUBC added it to the UB (largely driven by Medicare's need to identify hospital acquired conditions for payment restriction, required by Deficit Reduction Act of 2005) (Iezzoni 2007), and by 2009 at least 38 states collected it on some or all of their discharge records (Kassed et al. 2011).

USE OF THE DATA

States, researchers, the public health community, and the hospital industry use statewide hospital inpatient and outpatient data for a variety of purposes. Through their review of the literature, interviews with key informants and web searches, Schoenman and colleagues (Schoenman et al. 2005, 2007) identified the following types of uses of the data (Table 1): public safety, injury surveillance and prevention, public health, disease surveillance, disease registries, health planning, community assessments, public reporting for purchasing and comparative reports, quality assessment, performance improvement, commercial applications, and health services and policy research.

Health service researchers have recognized the value of statewide hospital discharge data since their inception. Indeed, the seminal small area variations study by Wennberg and Gittelsohn (1973) used 1969 data from an early SDO (Cooperative Health Information Center of Vermont). Researchers' use has grown so substantially that by 2014 an average of 50 journal articles each month used HCUP data[2] (the average would be higher if articles using SDOs' data were included).

To provide illustrations of current types of health services research using statewide discharge databases, Table 2 identifies 26 articles published in *HSR* over a recent 3-year period (2012–2014). Of these, 8 used data provided from the SDO and 19 used HCUP data (one study used both SDO and HCUP data). Of the 19 HCUP studies, 10 used the SID (occasionally including the SASD and SEDD), 7 used the NIS, and 2 used the KID. The studies often augmented the discharge data with data from other sources such as the AHA Annual Survey, Cost Reports from CMS or the states, cancer registries, Census Bureau data, the Area Health Resource File, and special datasets such as Medicaid enrollment files, AMA Master file, and data on health information technology implementation. These other data sources, which expand the analytic potential of discharge data, link to discharge data through hospital identifiers, physician identifiers, geographic area identifiers (e.g., county or zip

Table 1:   Types of Uses of Statewide Hospital Discharge Data

Type of Use	*Description*
Public Safety and Injury Surveillance and Prevention	States use the data to routinely track injury such as those from accidents, falls, and weapons.
Public Health, Disease Surveillance, and Disease Registries	States use the data for monitoring diseases that are generally treated in hospitals, such as severe trauma. In some cases, the data are used to supplement disease registry data, such as cancer registries. States use the data to report health care utilization related to Federal programs such as maternal and child health block grants.
Public Health Planning and Community Assessments	States use the data to monitor progress in meeting Healthy People goals, potential impacts from hospital mergers, closures and conversions from nonprofit to for-profit status. States and local communities use the data to obtain hospital utilization rates for specific conditions (e.g., avoidable hospitalizations) and use this with other information about community health, resources, and population.
Public Reporting for Informed Purchasing and Comparative Reports	States use the data to develop all-payer public reports on hospitals. In some cases, this is very basic information such as caseloads for specific conditions and surgeries, length of stay, and charges. In some states, the reports are more sophisticated, involving complex analyses to provide risk-adjusted outcomes, such as mortality and readmissions.
Quality Assessment and Performance Improvement	Beyond the public comparative reports, hospitals use the information for quality assessments and internal improvements. The statewide data provide hospitals with information on their own performance as well as benchmarks on their peers.
Private-Sector and Commercial Applications	Hospitals, their consultants, and other private entities use the data to provide hospitals and health plans information on the business and financial side of hospital services. Hospitals are interested in patient flow and market share analyses—both their own and other hospitals in their market area, and efficiency assessments such as case mix- or severity-adjusted length of stay.
Health Services and Health Policy Research Applications	The data are used for research on diverse health services and health policy topics. These include studies on the effects of financing and delivery systems on hospital service use and outcomes, hospitalizations for ambulatory care sensitive conditions (that provide a window on access to high-quality outpatient care), racial differences in use and outcomes, geographic variations, effects of government policies, and comparative effectiveness of different clinical practices and interventions.

Source: Schoenman et al. (2007).

Table 2:   Articles Using Discharge Datasets Published in *Health Services Research*, 2012–2014

| | Discharge Dataset | | | | | | | Additional Datasets | | | | |
| | HCUP Dataset | | | | | | | | | | | |
Author	SID	SASD	SEDD	NIS	KID	State	AHA File	Cost Reports	Census Bureau	AHRF	Cancer Registry	Other Data Sources
Bazzoli et al. (2012)	X						X		X	X		
Bradley et al. (2012)						X	X				X	
Braithwaite et al. (2013)				X			X		X			
Conti (2013)				X								
Davies et al. (2013)						X						Medicaid enrollment files
Dawes et al. (2014)						X					X	
Deily et al. (2013)						X			X			HiMSS Database; State Managed Care Reports
Downey et al. (2012)				X								
Drösler et al. (2012)				X								National data from other countries
Ghaffarzadegan, Epstein, and Martin (2013)						X						AMA Physician Masterfile
He and Mellor (2013)						X		X	X			State publications; Bureau of Labor Statistics
Hernandez-Boussard et al. (2012)				X								
Hockenberry et al. (2014)	X	X	X			X	X					
Holmes, Freburger, and Ku (2012)	X						X	X		X		CMS Provider of Services
Howard and Shen (2014)	X	X										

Table 2: *Continued*

| | Discharge Dataset | | | | | | Additional Datasets | | | | | |
| | HCUP Dataset | | | | | | | | | | | |
Author	SID	SASD	SEDD	NIS	KID	State	AHA File	Cost Reports	Census Bureau	AHRF	Cancer Registry	Other Data Sources
Levit, Friedman, and Wong (2013)	X							X				State hospital financial data; Medicare area wage index
Maeda, Raetzman, and Friedman (2012)	X							X				
Mark et al. (2013a)	X						X	X				State hospital disclosure reports
Martin et al. (2013)						X						
Maxwell et al. (2014)				X								
Morriss (2013)					X							
Reiter, Jiang, and Wang (2014)	X						X	X		X		Interstudy
Romley et al. (2014)					X							
Seymour et al. (2012)	X											
Smith et al. (2012)				X								
White (2014)	X							X	X			Provider Specific File, Impact File, Federal Regulations

AHA, American Hospital Association; AHRF, Area Health Resources File; AMA, American Medical Association; CMS, Centers for Medicare and Medicaid Services; HCUP, Healthcare Cost And Utilization Project; HIMSS, Healthcare Information and Management Systems Society; KID, Kid's Inpatient Database; NIS, National (Nationwide) Inpatient Sample; SASD, State Ambulatory Surgery and Services Databases; SEDD, State Emergency Department Databases; SID, State Inpatient Databases.

code), patient identifiers (accessible on some SDO data after a special application process).

The *HSR* article topics were wide-ranging covering health policy, access and utilization of care, impact of race–ethnicity and insurance, clinical and quality aspects of care, economics and financing, and research methods development. As examples of health policy studies, Mark et al. (2013a) examined the impact of nurse staffing legislation on patient outcomes. White (2014) and He and Mellor (2013) examined the impact of changing Medicare pricing policy. Studies investigating the impact of race or insurance coverage include Bazzoli et al.'s (2012) research on the impact of safety-net hospital closures and conversions on racial–ethnic minorities' access (travel times) and Morriss's (2013) examination of uninsured neonates' mortality. Seymour et al. (2012) studied hospital variation in intensive care unit use, an example of utilization studies. Several studies focused on clinical practice or quality of care, including Howard and Shen's (2014) assessment of practice changes following a clinical trial involving percutaneous coronary intervention. Of the many studies on complications and patient safety, Hernandez-Boussard et al. (2012) examined the relationship between patient safety and surgical volume. Smith et al.'s (2012) article crosses several topical areas (patient safety, insurance coverage, and economics) by studying Medicaid, hospital financial stress, and adverse medical events for children. Several additional articles examined the economics and financing of care, such as Reiter, Jiang, and Wang's (2014) investigation of how safety-net hospitals fared financially over the recession. The methods-focused articles included one germane to measurement of quality for public reporting and pay-for-performance (Davies et al. 2013), one on estimating hospital prices (Levit, Friedman, and Wong 2013), and one on international comparisons of patient safety given disparate national data systems (Drösler et al. 2012).

In addition to the topic areas described above, the studies have other notable features. Three studies used inpatient data to examine the effects of outpatient programs or reimbursement (Conti 2013; Deily et al. 2013; He and Mellor 2013). Two studies used discharge data with other data for simulations (Braithwaite et al. 2013; Ghaffarzadegan, Epstein, and Martin 2013). Most studies were about adults, but some studies focused on elderly adults (Davies et al. 2013) or children (Smith et al. 2012; Morriss 2013; Romley et al. 2014). Most were patient-level studies, but some aggregated the discharge-level data for hospital-level analyses (Braithwaite et al. 2013;

Davies et al. 2013; He and Mellor 2013; Levit, Friedman, and Wong 2013; Mark et al. 2013a; Reiter, Jiang, and Wang 2014) and one aggregated to metropolitan statistical area (White 2014). While many studies used broad patient types, some concentrated on specific types of conditions or procedures, including cancer (Bradley et al. 2012; Dawes et al. 2014); cardiac (Howard and Shen 2014; Maxwell et al. 2014; Romley et al. 2014); influenza (Braithwaite et al. 2013); pregnancy, delivery, or neonatal care (Deily et al. 2013; Ghaffarzadegan, Epstein, and Martin 2013; Morriss 2013); and a specific category of surgeries (Ghaffarzadegan, Epstein, and Martin 2013; He and Mellor 2013; Martin et al. 2013; Howard and Shen 2014; Maxwell et al. 2014; Romley et al. 2014).

LIMITATIONS OF THE DATA

Statewide discharge data have limitations that affect their usefulness and accuracy for some analyses. Schoenman et al. (2007) described the limitations as falling into three types: quality of data elements, missing data elements, and excluded populations. One data quality problem they identified concerns the accuracy of some ICD-9-CM-coded diagnoses and procedures, including miscoding and omission of comorbidities. O'Malley et al. (2005) noted a number of factors that may affect the data quality of the ICD-9-CM diagnoses, including the quality of information in the medical record, coder training and experience, facility quality control, and unintentional and intentional coding errors. Incentives to maximize reimbursement and to code only clinical information that affects reimbursement may affect the discharge data. Another shortcoming is that the maximum number of diagnosis and procedure fields is limited on some state's data. For example in 2012, some SDOs had a maximum of 9 or 10 diagnoses, while most had a maximum of 20 or more (Coffey et al. 2015). Thus, some states may have incomplete information for the more complex cases, affecting analyses (Iezzoni et al. 1992; Romano and Mark 1994). Another data quality concern mentioned by SDOs (Barrett et al. 2014) concerns expected payer, particularly that Medicaid enrollees in managed care may be miscoded as privately insured (Chattopadhyay and Bindman 2005). Data quality also suffers for multistate analyses when states collect data elements differently (Coffey et al. 1997), such as collecting different categories for expected payer categories (Barrett et al. 2014) or for race–ethnicity (Geppert et al. 2004; Andrews 2011).

Missing data elements limit the usefulness of the statewide data. For example, the statewide discharge data include hospital charges but do not include the hospital's costs to provide the services or the reimbursed amounts from payments by health plans or patient copays (Riley 2009). Many states do not include patient identifiers which would be useful for examining readmissions or for linkages to outside datasets. Similarly, physician identifiers are not collected in all states, limiting analyses such as studies of variations in physician practice patterns and outcomes. The data do not include information concerning events outside the hospital, such as out-of-hospital deaths. As discussed further in a later section, the statewide discharge data include some clinically related data elements such as ICD-9-CM-coded diagnoses and procedures, but they generally do not include detailed clinical data such as physiological measures, laboratory results, and functional status.

Another type of limitation is that the statewide data generally do not cover Federal hospitals, such as Veteran's Administration and Indian Health Service. Most states also do not have records for residents who use hospitals in another state (border crossing), although some SDOs enter into agreements to share such information, and analysts using data from multiple states may be able to identify border crossing through the patient zip code data element when available.

Researchers may attenuate some limitations through statistical approaches, data manipulations, or linkages to outside datasets. For example, statistical imputation methods may address missing data (Houchens 2015). In the absence of patient identifiers to identify readmissions, one study created an algorithm to develop patient-level records for leukemia and lymphoma patients using sex, age, race, insurance status, and zip code (Mitchell et al. 1997). Linkages expand the uses of the data (Riley 2009; Bradley et al. 2010). For example, linking the discharge data to death certificate files expands the mortality information to include out-of-hospital deaths (Herrchen, Gould, and Nesbitt 1997; Zingmond et al. 2004; Mark et al. 2013b). By using hospital cost reports (Medicare and or state financial records) with the discharge data, researchers can estimate the hospital cost to produce the care (Friedman et al. 2002; Riley 2009) or price (health plan payments) of the stay (Levit, Friedman, and Wong 2013). Because the data do not include information on a patient's income, researchers often employ a proxy measure using the median income of the patient's zip code (Krieger et al. 2002).

AHRQ ENHANCED STATE DATA GRANTS: IMPROVED CLINICAL CONTENT AND RACE–ETHNICITY DATA

To foster the improvements of clinical content and patient race and ethnicity data in statewide hospital discharge data, the AHRQ awarded eight 3-year grants in 2010 (AHRQ 2014a). These grants were part of the American Recovery and Reinvestment Act funding to improve the data infrastructure for comparative effectiveness research (AHRQ 2010; O'Day et al. 2014).

Improving Clinical Content

The clinical information in statewide discharge data focuses on diagnoses and procedures coded in ICD-9-CM (and for certain types of records, CPT®) codes. It also includes other data elements concerning the patient's clinical status such as discharge status and admission type (e.g., emergency, urgent, elective, trauma). As noted earlier, in recent years most states have begun to collect a "present on admission" indicator associated with each diagnosis to distinguish complications that began during the hospital stay from conditions that were present at hospital admission. While these are valuable data, they do not include more detailed information on the clinical status of the patient such as physiological measures (Tabak, Johannes, and Silber 2007; Escobar et al. 2008; Hayward 2008) or functional status (Iezzoni and Greenberg 2003) that are needed for some research. The growing availability of clinical data in electronic medical records, coupled with increasingly sophisticated health information technology, creates new opportunities to enhance the clinical content of hospital discharge data. Additional clinical detail would be useful for more precise measurement of patient severity and risk adjustment, as outcome measures or to identify a specific clinical study sample. Thus, clinically enhancing these data would expand their capacity to support studies concerning comparative effectiveness, quality improvement, efficiency, and health policy. The enhanced data may also improve the accuracy of reports on provider quality (Pine et al. 2007; Tabak, Johannes, and Silber 2007).

Table 3 provides information about the five Enhanced State Data Grant projects that focused on improving the clinical content of statewide hospital discharge data (AHRQ 2014a). These projects were aimed at improving the data in their state by linking in electronic clinical laboratory data (Hawaii, Minnesota, New York grants), hospital pharmacy data (Minnesota grant), pre-hospital emergency care data (New Jersey grant), and/or vital record birth and death certificate data (Florida, Minnesota, New Jersey grants).

Table 3:   State Data Enhancement Grants to Improve Clinical Content

State	Grantee and Source of Statewide Discharge Data	Enhancement of Statewide Hospital Discharge Data
Florida	*Grantee* University of South Florida *Source of Statewide Discharge Data* Florida Agency for Health Care Administration	Linked inpatient, emergency department, and ambulatory surgery data to vital statistics (birth and death) and hospital financial data to create a multiyear enhanced statewide maternal child dataset. Strategies for linkage of the maternal and child records were created to overcome such challenges as multiple births and records with missing patient identifiers.
Hawaii	*Grantee* Queen's Medical Center *Source of Statewide Discharge Data* Hawaii Health Information Corporation	Linked statewide data to laboratory results data; further developed master patient identifier to link and track patients across hospitals throughout the state.
Minnesota	*Grantee and Source of Discharge Data* Minnesota Hospital Association	Increased number of hospitals submitting laboratory data for linkage to discharge data, added inpatient pharmacy linkage, improved linkage of patients across hospitals and with death certificates. Focused on treatment of acute decompensated heart failure.
New Jersey	*Grantee* Rutgers University *Source of Statewide Discharge Data* New Jersey Department of Health	Linked inpatient and emergency department data to prehospital emergency medical services data and death certificates. Focused on therapeutic hypothermia on survivors of cardiac arrest.
New York	*Grantee and Source of Discharge Data* New York Department of Health	Linked discharge data to laboratory results data. Focused on coronary artery bypass graft surgery and elective percutaneous coronary interventions.

Source: AHRQ (2014a).

Using information and materials developed by the grantees, AHRQ developed the Clinical Content Enhancement Toolkit (AHRQ 2014b) for other SDOs and interested parties to learn from the grantees' experiences. The toolkit focuses on supplementing statewide discharge data with laboratory, pharmacy, vital statistics, and emergency medical services data and covers:

- Information to make the case for improving the clinical content of statewide discharge data to improve the accuracy for quality measurement and research.

- Materials to recruit and train hospitals.
- Information on relevant electronic data standards.
- Information and templates related to the process of data collection and linkages.

Improving Race–Ethnicity Data

Despite the substantial improvement in the number of states collecting race–ethnicity, two types of problems remain: (1) the most current data standards and recommended coding for race and ethnicity data have not been implemented widely and (2) the quality of race/ethnicity data remains suspect. Good quality, standardized race/ethnicity data are important to support studies on racial and ethnicity disparities (National Research Council 2004; National Committee on Vital and Health Statistics 2005; Institute of Medicine 2009). In addition, even when disparities are not the focus of a study, many studies use race and ethnicity data as one of the several patient characteristics to examine in relation to the main study question and/or as control variables. In addition, states and local communities need good quality discharge data to support their disparity monitoring and reduction efforts (Hanlon and Raetzman 2010).

SDOs vary in the race–ethnicity information they collect (Geppert et al. 2004; Andrews 2011). This inconsistency causes problems in multistate analyses and developing national estimates. About half the states collect the data using the current 1997 Office of Management and Budget (OMB) categories (OMB 1997), a Federal standard that is a useful minimal standard (Institute of Medicine 2009). About half of the states use the older 1977 OMB standard. Among states collecting race data, information about Hispanic ethnicity varies—a few states collect no information to identify Hispanic patients; about two-thirds have a separate data element for ethnicity (the preferred approach), while the remaining states have a combined race–ethnicity coding (this may meet the 1997 OMB standard, but it is not the preferred method). A few states collect more granular data than the 1997 OMB standards (with the ability to roll-up to OMB categories). Hawaii, for example, collects detailed Asian and Pacific Islander categories (Chinese, Filipino, Japanese, Native Hawaiian, and other Pacific Islanders) that coincide with the demographics of their state. A few states collect information about multiracial status—either "multirace" as a separate category or multiple fields for separate reporting of the multiple race and ethnicity categories.

This variability across states largely stems from the absence of race and ethnicity on the national standard for the hospital claim, unlike most data elements collected in statewide discharge data. In 2007, the NUBC added race and ethnicity to the UB as options for "public health reporting" (e.g., to report to SDOs), but they remain unacceptable on claims to health plans. The UB optional fields include separate reporting of race and ethnicity and the capability of multiracial reporting. The UB uses the Centers for Disease Control Race and Ethnicity Code Set, which has a hierarchical structure that provides for detailed reporting of race and ethnicity but also for "rolled up" codes that are compatible with the current OMB standard. Several states have migrated to this UB optional standard in recent years.

In addition to the standardization problems, there are ongoing concerns about the accuracy and completeness of the race and ethnicity data. One concern is that hospitals may use observation to collect the information (Gomez et al. 2003; Hasnain-Wynia, Pittman, and Pierce 2004) rather than the preferred method of asking the patient. The research on accuracy of race–ethnicity coding in hospitals suggests that it is particularly problematic for American Indians and Alaska Natives, but good for non-Hispanic whites and blacks (Blustein 1994; Korenbrot, Ehlers, and Crouch 2003; Fiscella and Meldrum 2008). Another problem area is the level of missing data in some states and hospitals, particularly when the reporting is voluntary and not mandated (Geppert et al. 2004).

Table 4 provides information about the three Enhanced State Data Grant projects that focused on improving the quality of race and ethnicity data in statewide hospital discharge. These projects were aimed at improving the data by (1) education and training of hospital staff concerning the importance of collecting accurate race–ethnicity data and methods to accurately collect self-reported information from patients (California and New Mexico), (2) assessing the quality of race–ethnicity reporting through a variety of quantitative approaches (California, New Mexico, Oregon/Washington), and (3) changing state regulations to standardize and expand data collection (New Mexico).

Using the products and materials developed by the grantees, AHRQ developed the Race and Ethnicity Data Improvement Toolkit (AHRQ 2014c) as an online resource for SDOs, hospitals, and researchers interested in improving hospital race and ethnicity data. The toolkit includes:

Table 4:   State Data Enhancement Grants to Improve Race and Ethnicity Data

State(s)	Grantee and Source of Statewide Discharge Data	Enhancement of Statewide Hospital Discharge Data
California	**Grantee** University of California, Los Angeles **Source of Statewide Discharge Data** California Office of Statewide Health Planning and Development	Created and disseminated materials for hospitals to consistently collect self-reported race, ethnicity, language data (R/E/L). Conducted educational webinars for hospitals. Performed a baseline assessment of hospitals' R/E/L data collection, reporting, and accuracy. Developed candidate audit measures and tested them by linking the discharge data to birth certificates and cancer registries. Compared California's data quality to six other states using the auditing approach.
New Mexico	**Grantee and Source of Discharge Data** New Mexico Department of Health	Revised state regulations to mandate hospital reporting of race, ethnicity, and tribal identifier data and to align coding with 1997 OMB standards. Formed a focus group of AI/AN persons to guide implementation of the collection of tribal identifiers. Conducted hospital key informant survey to understand hospital practices and target methods to improve R/E data collection. Developed training material and conducted in-person hospital trainings concerning the need to improve, and methods to collect, self-reported R/E data. Evaluated educational intervention concerning: (1) accuracy of R/E data compared to patient survey reporting; (2) change in hospital staff's knowledge about R/E data collection; (3) change in completeness of reported R/E data.
Oregon and Washington	**Grantee** Northwest Portland Area Indian Health Board/ Northwest Tribal Epidemiology Center **Source of Statewide Discharge Data** Oregon Health Authority and Washington Department of Health	Conducted record linkages and data analyses to identify racial misclassification of AI/AN in discharge data using the most complete roster of Northwest AI/AN people. Improved collaboration between the Tribal Epidemiology Center and state data organizations concerning race data quality and the need for accurate data for state surveillance systems for minority health.

Source: AHRQ (2014a).

- Information to make the case for improving race, ethnicity, and language data, including the benefits of improving the data collection and examples of using the data to track and reduce disparities.
- Education and training materials for hospital staff, including scripts for data collection, example questionnaires, and training evaluation materials.
- Quantitative approaches to examine the accuracy of the data collected (e.g., through linkages to outside databases or follow-back surveys).

Conclusion

Statewide discharge data support a wide range of health services research, as well as state, local, and hospital information needs. Though widely used, they have limitations. The AHRQ-funded Enhanced State Data Grants addressed limitations in the clinical content and race–ethnicity data. Pine et al. (2015) discuss the challenges the grantees faced in conducting their projects and their plans for sustaining their efforts following the grant funding. The goal of the AHRQ web-based toolkits, this AHRQ-sponsored *HSR* special issue, and other dissemination avenues is to demonstrate that improvements are possible and to provide technical resources to foster widespread enhancements based on the grantees' experiences.

In addition to possible future improvements stimulated by the AHRQ grants, the statewide discharge data are likely to change in other ways. The data have evolved over the last four decades in response to user needs, policy priorities, identified limitations, and national data standards. Going forward, we can expect these influences to prompt modifications—some foreseeable, but many unknown.

The most significant and challenging change is the transition from ICD-9-CM to ICD-10-CM/PCS in October 2015. The ICD-10-CM/PCS is substantially different from ICD-9-CM, with greater specificity (with many times more codes), and a different coding structure and format (Utter et al. 2013). Hospitals, health information coders, SDOs, and the Federal government have been working diligently to prepare their data systems for the new code sets. For example, AHRQ developed ICD-10 versions of its HCUP clinical grouper tools, is converting Quality Indicators software to ICD-10-CM/PCS (AHRQ 2014d), and has other online resources for SDOs (AHRQ

2014e). The transition presents challenges for researchers, who must understand the structure and format of the new code sets to use them appropriately. Some researchers are concerned about the quality of data during the initial transition (Krive et al. 2015), as well as potential ongoing coding problems (Utter et al. 2013). Longer term, because the two coding systems are so different, analyses using multiyear data with the two coding systems will be problematic. Researchers will be essential in the transition by evaluating coding accuracy, identifying appropriate and inappropriate uses of the code sets, developing methods that exploit the code sets' strengths, and creating nuanced bridging methods for multiyear studies involving new and old code sets.

Statewide health care data may be available for a broader range of services in the future. Many states have extended hospital data to ED and ambulatory surgery encounters, and at least a dozen SDOs are collecting information on hospital observation services (Hockenberry et al. 2014). In the last decade, several states began collecting inpatient, outpatient, and pharmacy claims data from insurance plans (Porter et al. 2015). These All-Payer Claims Databases (APCD) aim for claims from all-payers, except the uninsured, but may exclude some payers such as Medicaid and Medicare. With a dozen existing systems varying in service and plan coverage, APCDs are at a promising developmental stage, similar to where statewide discharge data were in the 1980s. They have the potential to support studies that complement the population-based statewide discharge data, for example, on health care payments, ambulatory services, and episodes or continuum of care (Love and Steiner 2011).

Advances in computers and health information technology could influence the collection and distribution of the statewide discharge data, possibly resulting in more timely data access. The electronic health record could provide a wealth of clinical information to the discharge abstract, as well as additional patient demographics (e.g., preferred language). EHRs are designed for clinical care, and optimal uses for research are emerging (Holve and Calonge 2013; Gardner et al. 2014; Randhawa 2014; Tai-Seale et al. 2014). The Enhanced State Data grantees improved clinical content through linkages to other data systems, such as electronic laboratory, pharmacy, and emergency medical services data. Alternatively, the UB could add key clinical information from the EHR if useful for payers (e.g., for pay-for-performance measures), similar to the POA indicator addition. If future payment policy does not require the UB, SDOs would need a new standardized reporting foundation, which could include EHR-based clinical data in a consensus-based core dataset (e.g., an updated UHDDS).

Government policy changes could affect the content and limitations of statewide discharge data. For example, prompted by the Patient Protection and Affordable Care Act (ACA), at least one state (Arizona) is planning to modify its expected source of payment field to track the impact of HIE coverage (Barrett et al. 2014). State variability in Medicaid expansion under the ACA may challenge data collection and analyses of payer codes. In particular, some states are proposing to use Medicaid funds for premium support for insurance through health insurance exchanges (Kaiser Family Foundation 2015). When submitting data to SDOs, hospitals will likely have difficulty identifying such patients as Medicaid-funded and instead code the payer as private insurance (similar to their current difficulty identifying Medicaid managed care patients). Changes in payment policy to bundled payments could limit the information about individual hospital stays and visits if payers no longer require a separate claim for each encounter. More dramatically, if claims were not required under a future payment system, SDOs would need to find another data standard to replace the UB.

Use of statewide discharge data grew over the last 40 years because the data continually met two key requirements—they are useful and require minimal burden to create and use. To maintain usefulness, they will need to adapt to changing needs of major stakeholders for population-based data—states, hospitals, policy makers, and researchers. To minimize resources in creating useful data, they need to continue leveraging government and payment reporting requirements, use data standards, and stay current with computer and health IT advances. Although SDOs and the hospital sector will be the central players, future advancements will require collaborations with researchers, the Federal government, and private foundations to demonstrate value, develop analytic methods, identify needed improvements, and support and test innovations. Through such collaborations over the last 40 years, statewide hospital discharge databases have become an important component of the data infrastructure for health services research.

ACKNOWLEDGMENTS

Disclosures: None.
Disclaimers: None.

NOTES

1. The development and contents of the HCUP databases are documented in the HCUP User Support website available at http://www.hcup-us.ahrq.gov
2. Based on a search for journal articles published in 2014 that used at least one of the HCUP databases. The search was made with the HCUP Publications Search feature at http://www.hcup-us.ahrq.gov/reports/pubsearch/pubsearch.jsp

REFERENCES

Agency for Healthcare Research and Quality [AHRQ]. 2010. "Fact Sheet: American Recovery and Reinvestment Act Investments in Comparative Effectiveness Research for Data Infrastructure" [accessed on June 22, 2015]. Available at http://archive.ahrq.gov/funding/arra/factsheets/osfsinfra.html

Agency for Healthcare Research and Quality [AHRQ]. 2014a. "Data Innovations. Healthcare Cost and Utilization Project (HCUP)" [accessed on May 25, 2015]. Available at www.hcup-us.ahrq.gov/datainnovations/grants.jsp

Agency for Healthcare Research and Quality [AHRQ]. 2014b. "Clinical Content Enhancement Toolkit. Healthcare Cost and Utilization Project (HCUP)" [accessed on June 22, 2015]. Available at www.hcup-us.ahrq.gov/datainnovations/clinicalcontentenhancementtoolkit/home_toolkits.jsp

Agency for Healthcare Research and Quality [AHRQ]. 2014c. "Race and Ethnicity Data Improvement Toolkit. Healthcare Cost and Utilization Project (HCUP)." [accessed on June 22, 2015]. Available at www.hcup-us.ahrq.gov/datainnovations/raceethnicitytoolkit/home_race.jsp

Agency for Healthcare Research and Quality [AHRQ]. 2014d. *Data Innovations—ICD-10-CM/PCS Resources: Tools. Healthcare Cost and Utilization Project (HCUP)*. Rockville, MD: Agency for Healthcare Research and Quality [accessed on June 22, 2015]. Available at http://www.hcup-us.ahrq.gov/datainnovations/icd10_tools.jsp

Agency for Healthcare Research and Quality [AHRQ]. 2014e. "Data Innovations—ICD-10-CM/PCS Resources. Healthcare Cost and Utilization Project (HCUP)" [accessed on June 22, 2015]. Available at http://www.hcup-us.ahrq.gov/datainnovations/icd10_resources.jsp

Agency for Healthcare Research and Quality [AHRQ]. 2015. "Healthcare Cost and Utilization Project" [accessed on June 22, 2015]. Available at www.hcup-us.ahrq.gov

Andrews, R. M. 2011. "Race and Ethnicity Reporting in Statewide Hospital Data: Progress and Future Challenges in a Key Resource for Local and State Monitoring

of Health Disparities." *Journal of Public Health Management and Practice* 17 (2): 167–73.

Barrett, M.,L. Lopez-Gonzalez, A. Hines, R. Andrews, and J. Jiang. 2014. "An Examination of Expected Payer Coding in HCUP Databases." HCUP Methods Series Report # 2014-03. Agency for Healthcare Research and Quality [accessed on June 22, 2015]. Available at http://www.hcup-us.ahrq.gov/reports/methods/methods.jsp

Bazzoli, G. J., W. Lee, H. M. Hsieh, and L. R. Mobley. 2012. "The Effects of Safety Net Hospital Closures and Conversions on Patient Travel Distance to Hospital Services." *Health Services Research* 47 (1 Pt 1): 129–50.

Blustein, J. 1994. "The Reliability of Racial Classifications in Hospital Discharge Abstract Data." *American Journal of Public Health* 84 (6): 1018–21.

Bradley, C. J., L. Penberthy, K. J. Devers, and D. J. Holden. 2010. "Health Services Research and Data Linkages: Issues, Methods, and Directions for the Future." *Health Services Research* 45 (5 Pt 2): 1468–88.

Bradley, C. J., B. Dahman, L. M. Shickle, and W. Lee. 2012. "Surgery Wait Times and Specialty Services for Insured and Uninsured Breast Cancer Patients: Does Hospital Safety Net Status Matter?" *Health Services Research* 47 (2): 677–97.

Braithwaite, S., B. Friedman, R. Mutter, and M. Handrigan. 2013. "Microsimulation of Financial Impact of Demand Surge on Hospitals: The H1N1 Influenza Pandemic of Fall 2009." *Health Services Research* 48 (2 Pt 2): 735–52.

Chattopadhyay, A., and A. B. Bindman. 2005. "Accuracy of Medicaid Payer Coding in Hospital Patient Discharge Data: Implications for Medicaid Policy Evaluation." *Medical Care* 43 (6): 586–91.

Coffey, R. M., J. K. Ball, M. Johantgen, A. Elixhauser, P. Purcell, and R. Andrews. 1997. "The Case for National Health Data Standards." *Health Affairs (Millwood)* 16 (5): 58–72.

Coffey, R., M. Barrett, R. Houchens, E. Moy, R. Andrews, E. Moles, and N. Coenen. 2015. *Methods Applying AHRQ Quality Indicators to Healthcare Cost and Utilization Project (HCUP) Data for the 2014 National Healthcare Quality and Disparities Report (QDR)*. HCUP Methods Series Report # 2015-02. Agency for Healthcare Research and Quality [accessed on June 22, 2015]. Available at http://www.hcup-us.ahrq.gov/reports/methods/methods.jsp

Conti, M. S. 2013. "Effect of Medicaid Disease Management Programs on Emergency Admissions and Inpatient Costs"." *Health Services Research* 48 (4): 1359–74.

Davies, S., O. Saynina, E. Schultz, K. M. McDonald, and L. C. Baker. 2013. "Implications of Metric Choice for Common Applications of Readmission Metrics." *Health Services Research* 48 (6 Pt 1): 1978–95.

Dawes, A. J., R. Louie, D. K. Nguyen, M. Maggard-Gibbons, P. Parikh, S. L. Ettner, C. Y. Ko, and D. S. Zingmond. 2014. "The Impact of Continuous Medicaid Enrollment on Diagnosis, Treatment, and Survival in Six Surgical Cancers." *Health Services Research* 49 (6): 1787–811.

Deily, M. E., T. Hu, S. Terrizzi, S. Y. Chou, and C. D. Meyerhoefer. 2013. "The Impact of Health Information Technology Adoption by Outpatient Facilities on Pregnancy Outcomes." *Health Services Research* 48 (1): 70–94.

Downey, J. R., T. Hernandez-Boussard, G. Banka, and J. M. Morton. 2012. "Is Patient Safety Improving? National Trends in Patient Safety Indicators: 1998-2007." *Health Services Research* 47 (1 Pt 2): 414–30.

Drösler, S. E., P. S. Romano, D. J. Tancredi, and N. S. Klazinga. 2012. "International Comparability of Patient Safety Indicators in 15 OECD Member Countries: A Methodological Approach of Adjustment by Secondary Diagnoses." *Health Services Research* 47 (1 Pt 1): 275–92.

Epstein, M. H. 1992. "Uses of State-Level Hospital Discharge Databases." *Journal of AHIMA* 63 (4): 32–9.

Epstein, M. H., and B. S. Kurtzig. 1995. "Statewide Health Information: A Tool for Improving Hospital Accountability." In *Accountability and Quality in Health Care: The New Responsibility*, edited by L. E. Markson, and D. B. Nash, pp. 51–71. Terrace, IL: Joint Commission on Accreditation of Healthcare Organizations.

Escobar, G. J., J. D. Greene, P. Scheirer, M. N. Gardner, D. Draper, and P. Kipnis. 2008. "Risk-Adjusting Hospital Inpatient Mortality Using Automated Inpatient, Outpatient, and Laboratory Databases." *Medical Care* 46 (3): 232–9.

Fiscella, K., and S. Meldrum. 2008. "Race and Ethnicity Coding Agreement between Hospitals and between Hospital and Death Data." *Medical Science Monitor* 14 (3): SR9–13.

Fisher, E. S., F. S. Whaley, W. M. Krushat, D. J. Malenka, C. Fleming, J. A. Baron, and D. C. Hsia. 1992. "The Accuracy of Medicare's Hospital Claims Data: Progress Has Been Made, But Problems Remain." *American Journal of Public Health* 82 (2): 243–8.

Friedman, B., J. De La Mare, R. Andrews, and D. H. McKenzie. 2002. "Practical Options for Estimating Cost of Hospital Inpatient Stays." *Journal of Health Care Finance* 29 (1): 1–13.

Gardner, W., S. Morton, S. C. Byron, A. Tinoco, B. D. Canan, K. Leonhart, V. Kong, and S. H. Scholle. 2014. "Using Computer-Extracted Data from Electronic Health Records to Measure the Quality of Adolescent Well-Care." *Health Services Research* 49 (4): 1226–48.

Geppert, J. J., S. J. Singer, J. Buechner, L. Ranbom, W. Suarez, and W. Xu. 2004. "State Collection of Racial and Ethnic Data." In *Eliminating Health Disparities: Measurement and Data Needs*, edited by M. Ver Ploeg, and E. Perrin, pp. 232–48. Washington, DC: National Academies Press.

Ghaffarzadegan, N., A. J. Epstein, and E. G. Martin. 2013. "Practice Variation, Bias, and Experiential Learning in Cesarean Delivery: A Data-Based System Dynamics Approach." *Health Services Research* 48 (2 Pt 2): 713–34.

Gomez, S. L., G. M. Le, D. W. West, W. A. Satariano, and L. O'Connor. 2003. "Hospital Policy and Practice Regarding the Collection of Data on Race, Ethnicity, and Birthplace." *American Journal of Public Health* 93 (10): 1685–8.

Hanlon, C., and S. Raetzman. 2010. *State Uses of Hospital Discharge Databases to Reduce Racial and Ethnic Disparities*. Agency for Healthcare Research and Quality. Contract No. HHSA-290-2006-00009-C [accessed on June 22, 2015]. Available at http://www.hcup-us.ahrq.gov/reports/race/C18RECaseStudyReportforWEBfinal99.pdf

Hasnain-Wynia, R., M. Pittman, and D. Pierce. 2004. *Who, When and How: The Current State of Race, Ethnicity, and Primary Language Data Collection in Hospitals.* New York: The Commonwealth Fund.

Hayward, R. A. 2008. "Access to Clinically-Detailed Patient Information: A Fundamental Element for Improving the Efficiency and Quality of Healthcare." *Medical Care* 46 (3): 229–31.

He, D., and J. D. Mellor. 2013. "Do Changes in Hospital Outpatient Payments Affect the Setting of Care?" *Health Services Research* 48 (5): 1593–616.

Hernandez-Boussard, T., J. R. Downey, K. McDonald, and J. M. Morton. 2012. "Relationship between Patient Safety and Hospital Surgical Volume." *Health Services Research* 47 (2): 756–69.

Herrchen, B., J. B. Gould, and T. S. Nesbitt. 1997. "Vital Statistics Linked Birth/Infant Death and Hospital Discharge Record Linkage for Epidemiological Studies." *Computers and Biomedical Research* 30 (4): 290–305.

Hockenberry, J. M., R. Mutter, M. Barrett, J. Parlato, and M. A. Ross. 2014. "Factors Associated with Prolonged Observation Services Stays and the Impact of Long Stays on Patient Cost." *Health Services Research* 49 (3): 893–909.

Holmes, G. M., J. K. Freburger, and L. J. Ku. 2012. "Decomposing Racial and Ethnic Disparities in the Use of Postacute Rehabilitation Care." *Health Services Research* 47 (3 Pt 1): 1158–78.

Holve, E., and N. Calonge. 2013. "Lessons from the Electronic Data Methods Forum: Collaboration at the Frontier of Comparative Effectiveness Research, Patient-Centered Outcomes Research, and Quality Improvement." *Medical Care* 51 (8 Suppl 3): S1–3.

Houchens, R. 2015. *Missing Data Methods for the NIS and the SID.* HCUP Methods Series Report # 2015-01. Agency for Healthcare Research and Quality [accessed on June 22, 2015]. Available at http://www.hcup-us.ahrq.gov/reports/methods/methods.jsp

Howard, D. H., and Y. C. Shen. 2014. "Trends in PCI Volume after Negative Results from the Courage Trial." *Health Services Research* 49 (1): 153–70.

Hsia, D. C., C. A. Ahern, B. P. Ritchie, L. M. Moscoe, and W. M. Krushat. 1992. "Medicare Reimbursement Accuracy under the Prospective Payment System, 1985 to 1988." *Journal of the American Medical Association* 268 (7): 896–9.

Iezzoni, L. I. 2007. "Finally Present on Admission But Needs Attention." *Medical Care* 45 (4): 280–2.

Iezzoni, L. I., and M. S. Greenberg. 2003. "Capturing and Classifying Functional Status Information in Administrative Databases." *Health Care Financing Review* 24 (3): 61–76.

Iezzoni, L. I., S. M. Foley, J. Daley, J. Hughes, E. S. Fisher, and T. Heeren. 1992. "Comorbidities, Complications, and Coding Bias. Does the Number of Diagnosis Codes Matter in Predicting In-Hospital Mortality?" *Journal of the American Medical Association* 267 (16): 2197–203.

Institute of Medicine. 2009. *Race, Ethnicity and Language Data: Standardization for Health Care Quality Improvement.* Washington, DC: National Academy Press.

Kaiser Family Foundation. 2015. "Medicaid Expansion in Arkansas" [accessed on June 22, 2015]. Available at http://kff.org/medicaid/fact-sheet/medicaid-expansion-in-arkansas/

Kassed, C., N. Kowlessar, A. Pfunter, J. Parlato, and R. M. Andrews. 2011. *The Case for the POA Indicator: Update 2011.* HCUP Methods Series Report # 2011-05. Agency for Healthcare Research and Quality [accessed on June 23, 2015]. Available at http://www.hcup-us.ahrq.gov/reports/methods/methods.jsp

Korenbrot, C. C., S. Ehlers, and J. A. Crouch. 2003. "Disparities in Hospitalizations of Rural American Indians." *Medical Care* 41 (5): 626–36.

Krieger, N., P. Waterman, J. T. Chen, M.-J. Soobader, S. V. Subramanian, and R. Carson. 2002. "Zip Code Caveat: Bias Due to Spatiotemporal Mismatches between Zip Codes and US Census-Defined Geographic Areas—The Public Health Disparities Geocoding Project." *American Journal of Public Health* 92 (7): 1100–2.

Krive, J., M. Patel, L. Gehm, M. Mackey, E. Kulstad, J. J. Li, Y. A. Lussier, and A. D. Boyd. 2015. "The Complexity and Challenges of the International Classification of Diseases, Ninth Revision, Clinical Modification to International Classification of Diseases, 10th Revision, Clinical Modification Transition in EDs." *American Journal of Emergency Medicine* 33 (5): 713–8.

Larks, M. 1986. "State Systems Gather More Health Care Data But Lack Uniformity." *Business and Health* 3 (5): 49–50.

Levit, K. R., B. Friedman, and H. S. Wong. 2013. "Estimating Inpatient Hospital Prices from State Administrative Data and Hospital Financial Reports." *Health Services Research* 48 (5): 1779–97.

Love, D., B. Rudolph, and G. H. Shah. 2008. "Lessons Learned in Using Hospital Discharge Data for State and National Public Health Surveillance: Implications for Centers for Disease Control and Prevention Tracking Program." *Journal of Public Health Management and Practice* 14 (6): 533–42.

Love, D., and C. Steiner. 2011. *Fact Sheet: Key State Health Care Databases for Improving Health Care Delivery.* All-Payer Claims Databases Council [accessed on May 26, 2015]. Available at http://apcdcouncil.org/sites/apcdcouncil.org/files/HDD_APCD_Fact_Sheet_021411.pdf

Maeda, J. L., S. O. Raetzman, and B. S. Friedman. 2012. "What Hospital Inpatient Services Contributed the Most to the 2001-2006 Growth in the Cost Per Case?" *Health Services Research* 47 (5): 1814–35.

Mark, B. A., D. W. Harless, J. Spetz, K. L. Reiter, and G. H. Pink. 2013a. "California's Minimum Nurse Staffing Legislation: Results from a Natural Experiment." *Health Services Research* 48 (2 Pt 1): 435–54.

Mark, T. L., W. Lawrence, R. M. Coffey, T. Kenney, B. C. Chu, E. R. Mohler 3rd, and C. Steiner. 2013b. "The Value of Linking Hospital Discharge and Mortality Data for Comparative Effectiveness Research." *Journal of Comparative Effectiveness Research* 2 (2): 175–84.

Martin, B. I., S. K. Mirza, G. M. Franklin, J. D. Lurie, T. A. MacKenzie, and R. A. Deyo. 2013. "Hospital and Surgeon Variation in Complications and Repeat Sur-

gery Following Incident Lumbar Fusion for Common Degenerative Diagnoses." *Health Services Research* 48 (1): 1–25.

Maxwell, B. G., J. K. Wong, D. C. Miller, and R. L. Lobato. 2014. "Temporal Changes in Survival after Cardiac Surgery Are Associated with the Thirty-Day Mortality Benchmark." *Health Services Research* 49 (5): 1659–69.

Mitchell, J. M., K. R. Meehan, J. Kong, and K. A. Schulman. 1997. "Access to Bone Marrow Transplantation for Leukemia and Lymphoma: The Role of Sociodemographic Factors." *Journal of Clinical Oncology* 15 (7): 2644–51.

Morriss, F. H. Jr 2013. "Increased Risk of Death among Uninsured Neonates." *Health Services Research* 48 (4): 1232–55.

National Center for Health Statistics. 1980. *Uniform Hospital Discharge Data: Minimum Data Set.* Report of the National Committee on Vital and Health Statistics. DHEW Pub No. 80-1157. Washington, DC: Government Printing Office.

National Committee on Vital and Health Statistics. 2005. *Eliminating Health Disparities. Strengthening Data on Race, Ethnicity, and Primary Language in the United States.* Hyattsville, MD: Department of Health and Human Services.

National Research Council. 2004. *Eliminating Health Disparities: Measurement and Data Needs* edited by M. VerPloeg, and E. Perrin. Washington, DC: The National Academies Press.

O'Day, B., T. Kieffer, S. Forrestal, and D. Esposito. 2014. "American Recovery and Reinvestment Act Investments in Data Infrastructure." *Journal of Comparative Effectiveness Research* 3 (6): 591–600.

Office of Management and Budget [OMB]. 1997. "Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity." Federal Register: 62: No. 210 [accessed on June 22, 2015]. Available at http://www.whitehouse.gov/omb/fedreg/1997standards.html

O'Malley, K. J., K. F. Cook, M. D. Price, K. R. Wildes, J. F. Hurdle, and C. M. Ashton. 2005. "Measuring Diagnoses: ICD Code Accuracy." *Health Services Research* 40 (5 Pt 2): 1620–39.

Pine, M., H. S. Jordan, A. Elixhauser, D. E. Fry, D. C. Hoaglin, B. Jones, R. Meimban, D. Warner, and J. Gonzales. 2007. "Enhancement of Claims Data to Improve Risk Adjustment of Hospital Mortality." *Journal of the American Medical Association* 297 (1): 71–6.

Pine, M., N. M. Kowlessar, J. L. Salemi, J. Miyamura, D. S. Zingmond, N. E. Katz, and J. Schindler. "Enhancing Clinical Content and Race/Ethnicity Data in Statewide Hospital Administrative Databases: Obstacles Encountered, Strategies Adopted, and Lessons Learned." *Health Services Research* 50 (4 Pt 2): 1300–21.

Porter, J., D. Love, A. Costello, A. Peters, and B. Rudolph. 2015. *All-Payer Claims Database Development Manual: Establishing a Foundation for Health Care Transparency and Informed Decision Making.* All-Payers Claims Database Council and West Health Policy Center [accessed on May 26, 2015]. Available at http://apcdcouncil.org/all-payer-claims-database-development-manual

Randhawa, G. S. 2014. "Building Electronic Data Infrastructure for Comparative Effectiveness Research: Accomplishments, Lessons Learned and Future Steps." *Journal of Comparative Effectiveness Research* 3 (6): 567–72.

Reiter, K. L., H. J. Jiang, and J. Wang. 2014. "Facing the Recession: How Did Safety-Net Hospitals Fare Financially Compared with Their Peers?" *Health Services Research* 49 (6): 1747–66.

Riley, G. F. 2009. "Administrative and Claims Records as Sources of Health Care Cost Data." *Medical Care* 47 (7 Suppl 1): S51–5.

Romano, P. S., and D. H. Mark. 1994. "Bias in the Coding of Hospital Discharge Data and Its Implications for Quality Assessment." *Medical Care* 32 (1): 81–90.

Romley, J. A., A. Y. Chen, D. P. Goldman, and R. Williams. 2014. "Hospital Costs and Inpatient Mortality among Children Undergoing Surgery for Congenital Heart Disease." *Health Services Research* 49 (2): 588–608.

Schoenman, J. A., J. P. Sutton, S. Kintala, D. Love, and R. Maw. 2005. *The Value of Hospital Discharge Databases.* Final report submitted to the Agency for Healthcare Research and Quality under contract number 282-98-0024. Rockville, MD: Agency for Healthcare Research and Quality [accessed on June 22, 2015]. Available at http://www.hcup-us.ahrq.gov/reports/final_report.pdf

Schoenman, J. A., J. P. Sutton, A. Elixhauser, and D. Love. 2007. "Understanding and Enhancing the Value of Hospital Discharge Data." *Medical Care Research and Review* 64 (4): 449–68.

Seymour, C. W., T. J. Iwashyna, W. J. Ehlenbach, H. Wunsch, and C. R. Cooke. 2012. "Hospital-Level Variation in the Use of Intensive Care." *Health Services Research* 47 (5): 2060–80.

Smith, R. B., L. Dynan, G. Fairbrother, G. Chabi, and L. Simpson. 2012. "Medicaid, Hospital Financial Stress, and the Incidence of Adverse Medical Events for Children." *Health Services Research* 47 (4): 1621–41.

Tabak, Y. P., R. S. Johannes, and J. H. Silber. 2007. "Using Automated Clinical Data for Risk Adjustment: Development and Validation of Six Disease-Specific Mortality Predictive Models for Pay-for-Performance." *Medical Care* 45 (8): 789–805.

Tai-Seale, M., C. J. Wilson, L. Panattoni, N. Kohli, A. Stone, D. Y. Hung, and S. Chung. 2014. "Leveraging Electronic Health Records to Develop Measurements for Processes of Care." *Health Services Research* 49 (2): 628–44.

Utter, G. H., G. L. Cox, P. L. Owens, and P. S. Romano. 2013. "Challenges and Opportunities with ICD-10-CM/PCS: Implications for Surgical Research Involving Administrative Data." *Journal of the American College of Surgeons* 217 (3): 516–26.

Wennberg, J., and A. Gittelsohn. 1973. "Small Area Variations in Health Care Delivery." *Science* 182 (4117): 1102–8.

White, C. 2014. "Cutting Medicare Hospital Prices Leads to a Spillover Reduction in Hospital Discharges for the Nonelderly." *Health Services Research* 49 (5): 1578–95.

Zingmond, D. S., Z. Ye, S. L. Ettner, and H. Liu. 2004. "Linking Hospital Discharge and Death Records-Accuracy and Sources of Bias." *Journal of Clinical Epidemiology* 57 (1): 21–9.