Options for a Statewide Health Data Reporting System in Mississippi

Prepared by the National Association of Health Data Organization for the Center for Mississippi Health Policy

November 2007

THE NATIONAL ASSOCIATION OF HEALTH DATA ORGANIZATIONS

NAHDO

Table of Contents

Table of Contents i
List of Figures & Tables iii
Executive Summary iv
Summary of Recommendations vii
Introduction 1
The Benefits of a Patient-level Statewide Health Data Reporting System 12
Public safety and injury surveillance and prevention 12
Public health, disease surveillance and disease registries 13
Public health planning and community assessment 13
Public reporting for informed purchasing and comparative reports 13
Quality assessment and performance improvement 14
Health services and health policy research 14
Private sector and commercial applications 14
Informing policy deliberations and legislation 14
Employer and Purchaser Use of Administrative Data 14
An Integrated State Data System in South Carolina 15
The Value of Emergency Department Data Value 17
The Value of Ambulatory Surgery Data 18
Planning Considerations 19
Scope of Data Collection 19
Data Format 20
Data Elements 22
Governance and Funding 28
Determination of data steward (data ownership and control) 33
Funding Mechanism 34
Planning-The Advance Work 37
Analysis of Expected Costs to Implement and Operate a Health Data Program 37
Implementation Requirements 44
Data Management 46
Data Access and Dissemination Policies 55
Data Modification 56
Strategies for Adding Value 63
Data Linkage 64
Appendix 1 67
Appendix 2 68
Appendix 3 72

List of Figures & Tables

Figure 1: State Data Collection Approaches vi
Figure 2: Number of State-Level Data Sets by First Year of Collection 2
Table 1. Strengths and Weaknesses of Hospital Discharge Data 3
Figure 3: Maine Claims Data Flow 6
Table 2. Data Sets Integrated in South Carolina 16
Table 2. Core UB-04 Billing data elements 23
Table 3. Collection of Race and Ethnicity Data by Collection Method (mandatory versus voluntary) 25
Figure 4: State Data Collection Approaches 28
Table 4. Strengths and Weaknesses of Mandatory and Voluntary Reporting Initiatives 29
Table 5. Advantages and Disadvantages of Funding Approaches 35
Table 6. Cost Considerations for a State Health Data Program (Year 1 is in 2007 dollars) 38
Figure 5: Data Submission Process 43
Table 7. Proposed Data Elements for an Emergency Department Data Collection System 44
Figure 6. Data Storage & Dissemination: General Practices 55
Table 8. Data Elements Available for Analysis and Release Physician Administrative Data Collection System Bureau of Health Information Draft, for purposes of discussion (Wisconsin) 5/4/2000 56
Table 9. 57
Table 10. Public Use Data Elements (example) 58
Table 11. Restricted Data Elements 58
Table 12: Marginal Cost-Effectiveness of Supplementing Administrative Data with Increasingly Costly Data to Improve Risk Adjustment of IQIs and PSIs 63
Figure 7: Data sets linked by respondents of NAHDO Survey on Linkage 66

Executive Summary

Mississippi is one of the few states or jurisdictions without a state health data reporting program, leaving important gaps in information to guide policy and market decisions. In 2007, 46 states and the District of Columbia maintain hospital inpatient discharge databases[1]. Because these data typically include detailed information on all patients discharged from the acute care hospitals in the jurisdiction, they are a complete, ongoing source of health care information that has proven useful in health care cost, quality, access, and research applications. These systems are the foundation for policy, research, and market information in those states and a growing source of health care data for national programs.

Statewide inpatient discharge data are routinely used to study and monitor issues of public safety, including injury surveillance and prevention.[2] Hospital Discharge Data are used in a remarkably wide range of applications.

- Public safety and injury surveillance and prevention;
- Public health, disease surveillance and disease registries;
- Public health planning and community assessment;
- Public reporting for informed purchasing and comparative reports;
- Quality assessment and performance improvement;
- Health services and health policy research ;
- Private sector and commercial applications;
- Informing policy deliberations and legislation; and
- Employee & purchaser use.

This report is intended to facilitate planning and implementation decisions in Mississippi and lay out a framework for a establishing a statewide health data program in Mississippi. This report can be used as a tool to solicit input across potential data suppliers and expected data users. The recommendations in this report are based on experience in implementing similar systems in other states, and lays out the considerations and the options for organizing and operating a health data program. The establishment of a healthcare data program involves a series of decisions which are systematically addressed in this report, including the following issues:

- Type of healthcare data to be collected
- Defining the scope of data collection initiatives

[1] For the purposes of this paper, inpatient data refer to ***all-payer*** data (including self and uninsured) for all patients admitted to ***acute care (non-federal) hospitals*** in the state for a fiscal/calendar year/or quarterly periods, and collected into an annual data base. Records are collected ***by hospitalization***, not by individual, and are represented at the ***discharge level*** rather than as aggregated statistics.

[2] Schoenman, et. al.

- Governance and funding options to promote stability
- Analysis of expected costs to implement and operative a statewide health data program
- Cost considerations for the reporting burden to providers of the data
- Consideration of data standards and reporting requirements to identify the most appropriate and efficient formats for submitting data across various data providers
- Analysis of data access and data dissemination requirements, including recommendations on how the data should be stored and released., including access and confidentiality provisions

Over the past 30 years, the number of state health data programs has expanded from 7 in 1981 to 47 in 2007. States have established inpatient data systems which have served as the foundation for non-inpatient reporting systems. Today, 35 states have added ambulatory surgery data to inpatient reporting requirements and 27 states have also added ED data. Six states have implemented all payer all claims data reporting from commercial and public payers in those states and several other states are in the planning stages. Gathering data from payers poses a set of different challenges than hospital discharge data reporting. Payer-supplied data promote understanding about the cost and efficiency of care; however, because these data have important gaps, such as the uninsured and self-pay encounters and complete diagnostic codes, such as External Cause of Injury codes, the data may not fully support injury surveillance and public health assessment.

A typical health data program establishes and maintains an infrastructure to collect, analyze, and disseminate comparable health care data for multiple users and uses, while guarding the confidentiality of the data. States have adopted various approaches to statewide hospital discharge data reporting. As depicted in Figure 1, three basic organizational models for state health data programs have emerged:

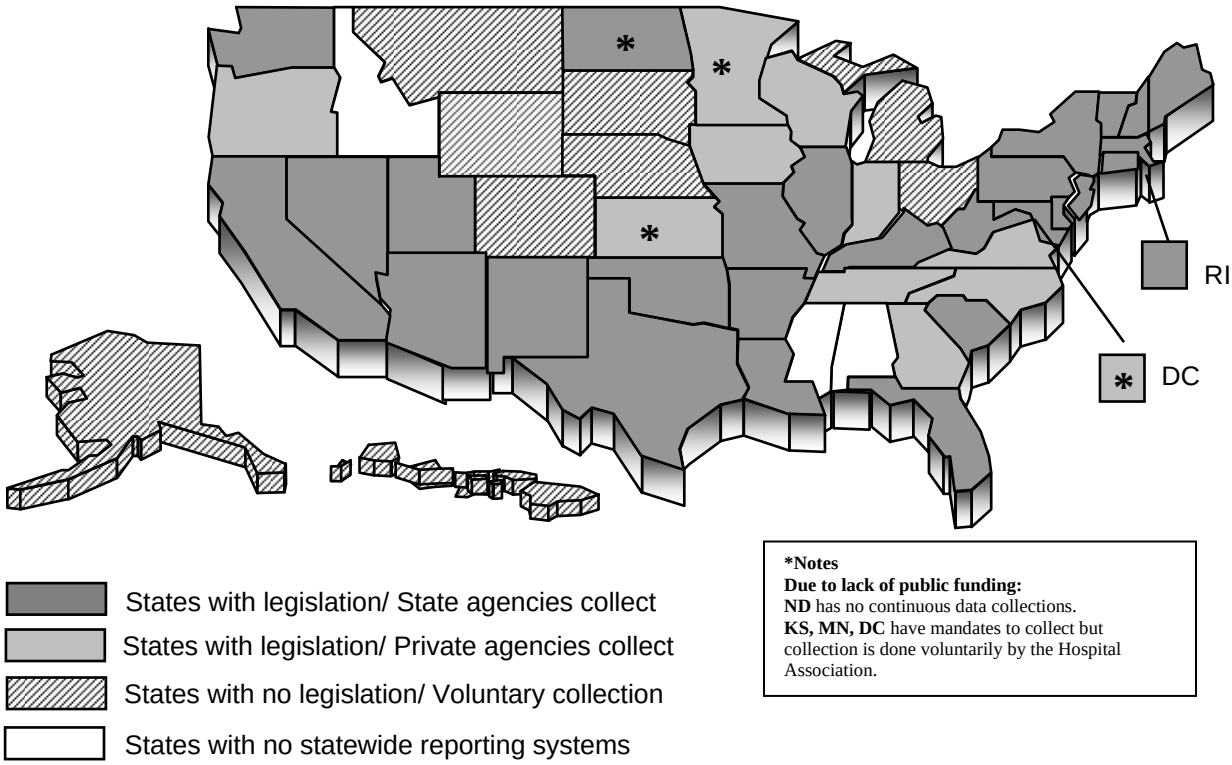
1) a public agency collects the data under a legislative mandate (28 States);

2) a private agency collects the data as the delegated authority to the state (11 States)

3) a private agency collects data voluntarily (9 States).

Each approach to data reporting has strengths and weaknesses, which this report will discuss. While a mandate does not assure compliance to reporting requirements, or a source of sustainable funding, it does provide a useful tool for planning, reporting, and consensus-building across all stakeholders.

Figure 1: State Data Collection Approaches



An ongoing, stable source of funding is essential to continuous data collection. A handful of states have enacted legislative mandates but have either not funded data collection or have phased out funding. Other states or jurisdictions have enacted mandates that were not funded. In those cases (Kansas, Minnesota, and the District of Columbia) the public agency has negotiated an arrangement with the hospital association in that state/jurisdiction to provide inpatient data. In these states/jurisdictions, the public agency is restricted in how they can use the data, thus limiting its utility for many stakeholders.

Which of the healthcare databases to develop and in which order is an important policy decision. With the exception of one state, all have started with inpatient data collection, expanding on that to include other types of healthcare data:

- 35 states have added ambulatory surgery data to inpatient reporting requirements; and
- 27 states have added ED data;
- Six states have added all payer all claims reporting from public and commercial payers

As Mississippi considers its options for a statewide health data program, looking to its public and commercial payers may be an alternative to ‘an all patient all provider’ (hospital discharge) data system. As an emerging system, all payer all claims data model, reported by public and private

payers, has already been implemented by four states (Maine, Massachusetts, New Hampshire, and Vermont), while two states are in the planning stages using Maine's model.

Users of healthcare data are generally diverse, including various government agencies, provider associations and individual health care providers, consumer organizations and individual patients, health care insurers and other health care purchasers (e.g., large employers), policymakers, researchers, and private-sector interests such as database vendors and consultants.

Where to locate a public program depends on the local environment and the opportunities and challenges the local environment presents in terms of leveraging existing Information Technology (IT) and staffing resources. There are advantages and disadvantages to either approach. An independent agency may be more flexible and responsive to stakeholder needs than a larger agency with layers of bureaucracy. However, the costs may be higher if a separate infrastructure must be established.

Analysis of Costs

Many decisions will impact the costs of a health data program in Mississippi. States have structured their health data programs in various ways, often reflecting the types of funding. Regardless of where a statewide health data program is housed, either as an independent state agency or within a larger umbrella organization, a core information technology and workforce infrastructure must be funded and at a sufficient level to establish and maintain the essential functions of planning, data collection, data management, analysis, and dissemination. Some states have built stand-alone or independent systems that may or may not leverage the existing IT environments of a larger agency. Others have relied on vendors and/or shared analytic resources with other entities in their state.

In this report we discuss the various cost considerations to guide funding decisions. Costs estimates for the planning, establishment, and ongoing operations of a statewide health data program will vary, depending on various factors. As a starting place for planning, estimates range from $277,406 for year-1 (planning), $387,894 for year-2 (implementation), and $401,295 for year 3 (full reporting and ongoing operations). These estimates were based on an inpatient volume of 407,000 per year and 94 acute care facilities reporting.

Summary of Recommendations

To assure comprehensive reporting and sustainability, a legislative mandate and general appropriations are the recommended authority and funding options, with balanced stakeholder governance. Mississippi should adopt national standards for its reporting requirements from acute care facilities and capture patient demographic data elements that can support population-based studies and measure and monitor health disparities. The de-facto national standard for state reporting systems is the UB data maintained by the National Uniform Billing Committee (NUBC). The UB-04 is the current version, which replaced the UB-92 in May, 2007. Data protection and release policies should be adopted that balance public availability of information with the protection of patient confidentiality. Finally, the Mississippi health data program should promote public use of data and generate research products, along with aggregate reports, that add value to the data and stimulate community and health improvements.

Specific recommendations are summarized in the section that follows. The following recommendations are based on the experience of statewide health data program implementation in other states and the options more likely to yield results in a cost-effective manner.

Data Collection

Mississippi should implement a hospital data reporting system to be accomplished in two stages, inpatient then outpatient, to be followed by other providers at a later time.

Inpatient data systems serve as the foundation and platform for outpatient data reporting in states. While ambulatory surgery and Emergency Department (ED) data are growing in importance as care shifts to outpatient settings, no state has implemented outpatient data reporting without first establishing an inpatient reporting system.

NAHDO recommends that Mississippi include race/ethnicity as a part of the mandatory reporting requirements. This recommendation is consistent with the National Research Council of the National Academies 2003 report, "Eliminating Health Disparities: Data and Measurement Issues". Mandatory reporting is more likely to result in reporting compliance. However, due to the sensitive nature of these variables, and due to limitations related to data reporting practices, NAHDO recommends these variables not be released publicly until sufficient evaluation and data quality assessment is completed in future years.

Measuring and monitoring health disparities is an important national and state priority. Capturing the appropriate data poses challenges, but industry support and collection tools are helping to overcome data collection challenges.

NAHDO recommends the inclusion of key patient demographic data elements as a part of both inpatient and outpatient discharge data requirements in Mississippi. These data elements should include, at the minimum, patient social security number (SSN), hospital medical record number, date of birth, and gender. If feasible, patient name and address should be collected to facilitate the assignment of a unique number, but also to facilitate geo-coding and public health applications.

The collection of patient demographic information improves the quality and utility of hospital discharge data. Over 28 states use patient SSN, which is usually encrypted to protect the identification of a particular patient.

Unique patient identifiers, such as patient SSN, name, address, and date of birth should be excluded from all public reports in their raw form and the release of identifiable information be carefully regulated for authorized research and public health applications.

Sensitive and potentially identifiable information should not be released to the public. Data aggregation, recoding, suppression, and data review and release policies are strategies states use to protect patient confidentiality.

Since Mississippi hospitals will be reporting Present on Admission (POA) indicator to Medicare beginning October 1, 2007, Mississippi's inpatient health data reporting requirements should include POA as a required core data element. For reasons of provider reporting cost and burden, NAHDO recommends that Mississippi NOT release this data element in public use files until the quality of the data element is evaluated in several years. NAHDO recommends that Mississippi NOT include clinical or laboratory data elements with

initial inpatient reporting requirements. These elements could be assessed for inclusion at a later time as automation and standards evolve.

The POA is a cost-effective way to improve the utility of inpatient data for future applications such as quality improvement and reporting.

Data Standards

To reduce provider reporting burden, the data formats should be aligned with the Uniform Bill (UB) standard maintained by the National Uniform Billing Committee (NUBC).

The adoption of national standards will reduce provider reporting burden and improve the comparability of the information across providers and states.

The legislation in Mississippi should reference "national standards" for reporting requirements, not specifically referencing UB-04, as standards change. Administrative rules should specify the UB-04 core data elements as required data elements, to align with national standards and reduce provider reporting burden.

The legislation should provide broad authority and provisions, to permit flexibility as standards and information needs evolve.

Governance and Funding

NAHDO recommends that Mississippi should statutorily mandate a publicly-controlled health data reporting program, thus joining 39 other states that have legislatively mandated such reporting.

Legislation is more likely to result in:

- Comprehensive reporting across all providers
- Expansion to non-inpatient reporting by providers
- Public availability and public health access to health information
- Transparency in methods of data collection and reporting.

The legislation should establish stakeholder representation in the form of a data commission or committee with rulemaking authority.

Broad stakeholder input and a fair decision-making process is essential to building a community data system that meets the diverse needs of the users for multiple uses.

Mississippi should consider the health data program data steward structure that is most likely to meet the objectives of public availability, sustainability of funding, equitable access, and independence and neutrality. An analysis of options should be undertaken in order to leverage an existing infrastructure and IT capacity as well as to assure the authority or ability to eventually expand to outpatient (non-hospital) settings in the future.

The selection of data ownership and control is important to the success and sustainability of a health data program and is a decision that is best made after careful deliberation by local stakeholders.

Mississippi should appropriate legislative funds to fully support the start-up and maintenance of a statewide health data reporting program. Funding should be in the form of general appropriations sufficient to fund the core infrastructure and staffing to support the implementation and ongoing operations of a health data program. A fee assessment, on all providers (including hospitals, nursing homes, free-standing ambulatory facilities) or on health plans, can be a source of core funding, if the political environment is favorable.

Core funding will provide revenues for the planning and implementation of statewide reporting in Mississippi. NAHDO recommends legislative general appropriations, as an investment in measuring, monitoring, and improving the health of the population. If feasible, an industry fee assessment may be considered, as providers and payers are the highest volume users of the data in most states.

Data Access and Dissemination

During the planning phases, the Mississippi health data program should develop a data release plan and establish policies that support the plan. The plan should provide for the release of de-identified data in the form of a public use data set, controlled by the use of a data use agreement; the release of a research-oriented data set for bona-fide research and federal programs, such as the AHRQ Healthcare Cost and Utilization Project (HCUP), either with IRB approval or data policy board review.

Mississippi could model the data release plan after other state health data program policies, which have a proven track record in providing useful information while protecting the confidentiality of the data.

States have established comprehensive data dissemination policies which could serve as a model for Mississippi. In over 30 years of discharge data reporting, there is no known breach of patient confidentiality by a health data program. These programs have effectively balanced the public good of the information with the need to protect patient confidentiality.

Mississippi should price these data products in a manner that balances the need for data sales revenues with data access for authorized uses by legitimate users.

Revenue from data product sales is an important component of most state health data program budgets. However, no data program relies solely on sales to support data collection activities. Other revenue streams from legislative appropriations, membership, or fee assessments offset the cost of data to users.

The Mississippi data dissemination and use plan, recommended earlier, should provide for data access by the authorized public health programs and support data linkage and data integration. This is an important mechanism to fill important data gaps and reduce the burden on providers to report multiple registry and surveillance data elements.

A large number of states provide inpatient and other data sets to their public health authority and programs to fulfill a variety of information needs, from injury and chronic disease surveillance (asthma, diabetes, cardiovascular care) to maternal-child health program purposes. The data are useful to public health to identify population (and community) health priorities and needs.

Mississippi should not adopt a web-based data query system until at least year three of operation, after the release of standard data products has been accomplished. At that time, Mississippi should research state-developed query systems and adopt one of these at no to low cost, to generate aggregate statistics and reports.

Web-based data query systems (WDQS) are a useful tool for disseminating hospitalization statistics to the public. A WDQS will reduce the number of requests for data sets and reports, reducing the burden to produce and the risk of circulating patient-level data. State systems are available to Mississippi in the future, after the data system is established and stable.

About NAHDO

This report is prepared by the National Association of Health Data Organizations (NAHDO) and draws from its work with state health data programs. Since 1986, NAHDO has promoted the uniformity, comparability, and public availability of statewide health care data. Using state-to-state transfer of technical assistance and lessons learned, NAHDO has built a national network or community of practice around the collection and use of hospital discharge data. The lessons learned in states with mature programs benefit those, like Mississippi, that are in the beginning stages of planning and implementing statewide reporting

Introduction

Mississippi is one of the few states or jurisdictions without a state health data reporting program. In 2007, 46 states and the District of Columbia maintain hospital inpatient discharge databases[3]. Because these data typically include detailed information on all patients discharged from the acute care hospitals in the jurisdiction, they are a complete, ongoing source of health care information that has proven useful in health care cost, quality, access, and research applications. These systems are the foundation for policy, research, and market information in those states and a growing source of health care data for national programs.

States are not alone in the quest to engage consumers and improve quality. In August of 2006, President Bush signed an Executive Order to promote efforts to improve transparency and quality in health care for all health programs administered or sponsored by the federal government. The Centers for Medicare and Medicaid Services (CMS) is shifting rapidly from a passive payer of services to an active purchaser and is implementing a series of provisions outlined in the Medicare Modernization Act and the Deficit Reduction Amendment. And state policymakers see the public reporting of pricing and quality information as a way to engage the consumer and reform the market. At the national and local level, health care providers are also becoming much more engaged in the measurement of quality and efficiency, establishing a variety of "alliances" that are providing input into the measurement of healthcare quality. Some medical specialty societies are developing new measures. These measurement activities require a reliable and valid source of data—hospital discharge systems are a less costly source of data for these initiatives. Many states are now using these endorsed measures as part of their measurement programs. What is clear from all this activity is that administrative data (particularly discharge data that has been expanded to include a limited number of clinical data elements) is in high demand.

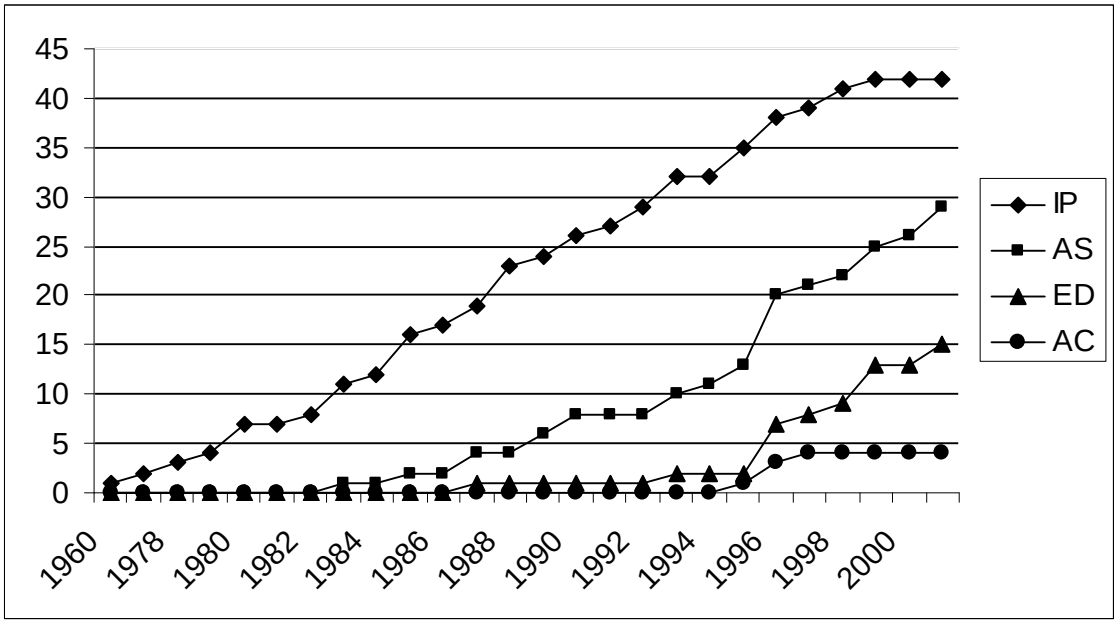
Many believe that at least a partial solution to the health care crisis is the publication of information to understand the variation in cost, quality, and access as a starting place to inform policy and market decisions and enable consumers to be actively involved in decisions. Many state and local data initiatives have proven successful, in part, because the scale is manageable and the information, collected locally, is understandable and actionable.[4]

Over the past 30 years, the number of state health data programs has expanded from 7 in 1981 to 47 in 2007. States have established inpatient data systems which have served as the foundation for non-inpatient reporting systems. Today, 35 states have added ambulatory surgery data to inpatient reporting requirements and 27 states have also added ED data. Figure 2 depicts the growth of different databases over time.

[3] For the purposes of this paper, inpatient data refer to all-payer data (including self and uninsured) for all patients admitted to acute care (non-federal) hospitals in the state for a fiscal/calendar year/or quarterly periods, and collected into an annual data base. Records are collected by hospitalization, not by individual, and are represented at the discharge level rather than as aggregated statistics.

[4] Colmers, J. Public Reporting and Transparency. The Commonwealth Fund, February 6, 2007.

Figure 2: Number of State-Level Data Sets by First Year of Collection



Source: 2003 Nationwide Health Data Inventory, AHRQ/HCUP Planning Document

In four states, the health data program is being used as a platform to expand to 'all payer all claims systems. In the states of Maine, New Hampshire, Vermont, and Massachusetts, health care claims databases are capturing data from all public and private payers to collect data from paid medical, dental and pharmacy claims files. The sources of the data are licensed commercial carriers and third party administrators, CMS, and state Medicaid programs. These data do not replace, but augment, the existing facility-based reporting in those states.[5]

Hospital discharge databases are a proven and cost-effective source of data for market, policy, and research applications. Because they contain detailed, record level information on all inpatient (or ambulatory/ED) encounters, these population-based data sets support a large range of uses and serve diverse audiences. When properly collected, validated, and disseminated, these data support a robust set of applications, discussed in the next section of this report. The data, like any data set, have important strengths and weaknesses. The strengths include the relative uniformity of the data across providers and states. Because they are derived from national billing standards, most providers are able to report the data without undue burden. The data are widely used in utilization, quality, public health and research.

Perhaps the greatest limitation of hospital discharge data is that they lack clinical detail, because they are designed for billing and administrative functions, not clinical decision making. Coding practices vary across providers and, because the data are coded after the patient is discharged, coded data are not as timely as some would prefer. Table 1 summarizes these strengths and weaknesses.

[5] Prysunka, A. Maine Health Data Organization, Health Care Claims Data: The Fuel for the Gauges, Presented at the NAHDO 21st Annual Meeting , Washington D.C., December 5, 2006.

Table 1. Strengths and Weaknesses of Hospital Discharge Data

Strengths	Weaknesses
Full census—all discharges are available in the file.	Clinical data, such as laboratory results and pathophysiologic information, are not included
National standards increase consistency of data across hospitals	Discharge data lack timeliness due to coding, reporting, and validation of data
Large number of cases or observations	Coding practices may vary across providers
Cost-effective to collect when compared to surveys and medical records abstraction	
Can be linked to other data sets or enhanced with clinical data to augment information	
Provides baseline and trend information on health care cost, quality, and access	

For the purposes of this report, a state health data program is a public or private initiative that collects patient-level data and includes all patients and all payers from generally all acute care facilities in a state. All health data programs have established inpatient data reporting systems initially, using inpatient data systems as a building block for future expansion. After successfully establishing an inpatient data program, many states have progressed to other facility-based data reporting. In 2007, a growing number of states are augmenting facility reporting systems with all payer all claims data from payers. The various data sets that states maintain are described below:

Hospital Inpatient Discharge Data: Inpatient care is the provision of surgical and non-surgical health care services to individuals admitted to non-Federal acute care hospitals. Records are collected by hospitalization, not by individual, and are represented at the discharge level rather than as aggregated statistics. Inpatient data include all-payer data (including self and uninsured) for all patients admitted to an acute care hospital in the state for a fiscal or calendar year period, or by periods that can be collected into an annual database. Inpatient data generally contain a complete collection of demographic, clinical and billing data.

Ambulatory Surgery Data: Ambulatory surgery care is the provision of surgical care performed on an outpatient basis (i.e., care that falls into one of the surgical DRG categories) and includes data from hospital-based ambulatory surgery centers in non-Federal acute care hospitals and, depending on state collection practices, free-standing ambulatory surgery centers. This includes surgical procedures that could have been alternately performed on an inpatient basis (e.g. cataract surgery and hernia repair), but excludes minor procedures (e.g., toenail removal or skin biopsy). Ambulatory surgery care has also been defined as patients identified in Hospital Separation Abstracts as outpatients, who underwent surgery (based on Diagnostic Related Groups (DRGs)) in an operating room.

Ambulatory surgery data are defined as all-payer data (including self and uninsured) collected for a fiscal or calendar year period, or by periods that can be collected into an annual database. Ambulatory surgery data are collected by visit, not by individual, and are represented at the encounter level rather than as aggregated statistics. Ambulatory surgery data generally contain a complete collection of demographic, clinical and billing data.

Emergency Department Data: Emergency department care is the provision of surgical or non-surgical health services to individuals admitted to or registered in a non-Federal acute care hospital emergency department. Emergency department data are defined as all-payer data (including self and uninsured) for all patients admitted to the emergency department in non-Federal acute care hospitals, for a fiscal or calendar year period or by periods that can be collected into an annual database. Emergency department data ideally contain a complete collection of demographic, clinical and billing data.

Pursuing Quality Data: Using Statewide Health Data as the Source of Quality Indicators,

"Building a database is the single most important thing states can do. (Including) an individual identifier (to track) people through time…. And (to improve quality the states should) identify hospitals doing better jobs with coding and hospitals not coding well, and work with them to improve the quality of the database".

Leslie Roos, Ph.D., Professor, Drake Center for Management Studies and Dept of Medicine, University of Manitoba, November 11, 1987

Observation Data: Outpatient Observation care is provision of stays for patients not admitted to the hospital. Observation stays include the evaluation and treatment of patients expected to be stabilized and released within 24 hours[6]. Services are for the reasonable and necessary evaluation and treatment services furnished on a hospital's premises to determine the need for admission to the hospital. Monitoring by hospital staff and use of a bed are included within these services. Examples of observation care include extended recovery following a complication of an outpatient procedure (e.g. poor pain control, intractable vomiting, delayed recovery from anesthesia). Outpatient observation is not a substitute for an inpatient admission, for medically stable patients needing diagnostic testing or outpatient procedures, patients needing therapeutic procedures (e.g. chemotherapy, dialysis) that are routinely provided in an outpatient setting, or for patients awaiting nursing home placement, or as a convenience for the patient, family, or physician[7]. Stays for patients admitted to the hospital are included in a state's Inpatient Database.

Ambulatory Care Data: Ambulatory care is the provision of non-surgical health services to individuals outside their home by health-care professionals. These services may take place in hospital-based or freestanding clinics, or office settings. Examples of such settings include but are not limited to community health centers, group and private medical practices, outpatient clinics, student health services, primary care clinics, and specialty diagnostic centers. Ambulatory care data ideally contain a complete collection of demographic, clinical and billing data.

Regardless of the structure and approach to collecting statewide health data, there are common characteristics:

- Most states base their data collection requirements on the national billing standard (UB-04) because of the robust nature of this standard. (It should be noted for that the basic UB-04 standard includes such data as: diagnosis codes, inpatient and outpatient

[6] Massachusetts Health Data Consortium. Data and Research: Outpatient Observation Stays in Massachusetts. Available at http://www.mahealthdata.org/data/observation/index.html. Accessed on December 11, 2007

[7] Baer RK. Hospital Guidelines for Outpatient Observation Services. MPRO. Available at http://www.mpro.org/hospital/pdf/CMS_observation_services.pdf. Accessed on December 11, 2007

procedure codes, external cause of injury codes, patient demographic data, detailed service and associated charge data, and payer identification.)

- All health data programs maintain, at the minimum, inpatient encounters from all acute care hospitals and all payers, including uninsured and self-insured. Over half of the statewide health data programs have also expanded reporting to include ambulatory surgery and Emergency Department encounters.
- States typically exclude federal hospitals and payers, such as Veterans Administration Medical Centers (VAMC), military hospitals, and Indian Health Service facilities. However, some states have successfully included these facilities to statewide reporting, with voluntary submissions.

An Emerging State Data System: All Payer All Claims Payer Data

An important consideration for a statewide health data program is to determine the source of data. Providers and payers are the logical sources because they are licensed and regulated by the state and the number of potential data suppliers is manageable. Physician information can be obtained indirectly through providers and payers. Direct physician reporting poses huge reporting challenges for many reasons, including the large number of physicians (data sources) and the varying capacity of physician offices to report uniform data in an electronic form.

As Mississippi considers its options for a statewide health data program, looking to its public and commercial payers may be an alternative to an all patient all provider (hospital discharge) data system. All payer all claims data reported by public and private payers is an emerging source of data in states. Four states, Maine, Massachusetts, New Hampshire, and Vermont have implemented an all payer all claims system, adopting the "Maine" system for reporting. Two states, Hawaii and Utah are in the planning stages and are seriously considering the Maine model. Kansas and Maryland collect claims data but are incompletely different formats.

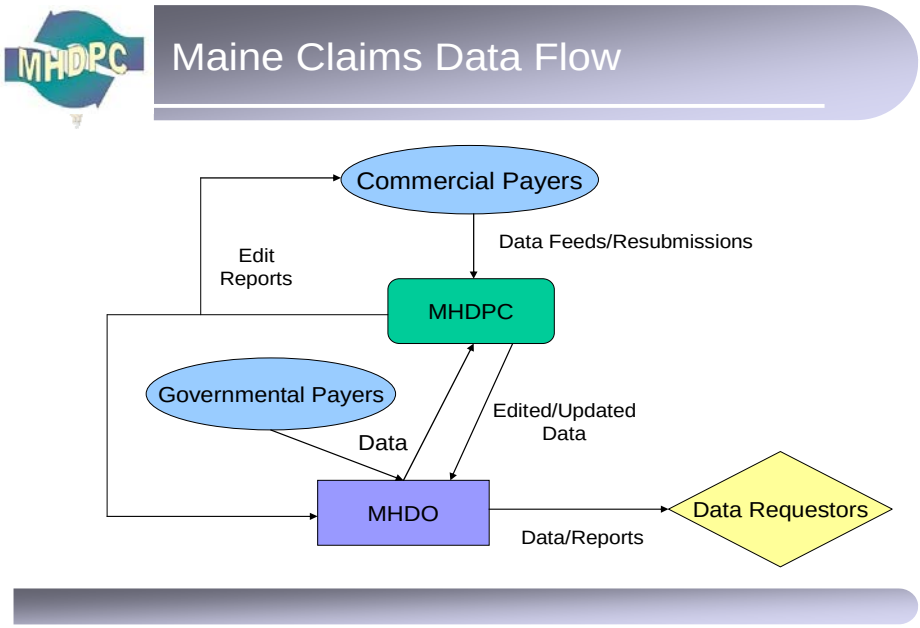
The Maine Health Data Organization (MHDO) and the Maine Health Data Processing Center have developed a platform for collecting all claims data, which paves the way for other states and has the potential to reduce technical challenges associated with this system. The Maine database structure includes the following:

- Paid medical, dental, pharmacy claims files for all covered services rendered to privately-insured residents;
- Eligibility and membership files;
- Health care service provider files; and
- Home-grown procedure and taxonomy code files[8].

[8] Prysunka, A. Maine Health Data Organization, "All Provider/All Payer Claims Database Status Report", NAHDO's 22nd Annual Meeting, October 2007, San Diego, CA

Figure 3 below illustrates the claims data flow in Maine:

Figure 3: Maine Claims Data Flow



The claims database in Maine includes the following data elements:

Encrypted social security numbers	Type of product (HMO, POS, Indemnity)
Type of contract (individual, family)	Patient demographics (DOB, gender, residence, relationship to subscriber)
Diagnoses codes (including E-codes)	Procedure codes (ICD, CPT, HCPS, CDT)
NDC code/generic indicator	Revenue codes
Service dates	Service provider (name, tax ID, payer ID, specialty code, city, state, zip)
Prescribing physician	Plan payments
Member payment responsibility (co-pay, co-insurance, deductible)	Date paid
Type of bill	Facility type

What is excluded in the Maine system are services provided to the uninsured, denied claims, workers' compensation claims, services to out-of-state residents, Diagnosis-related Group assignment (DRG), and premium information, along with capitation and administrative fees and clinical data, such as test results. Maine has acquired Medicare patient-level data and merged these data with commercial claims, including Part D pharmacy data, which is in a commercial claims data format from the payers.

With respect to financing the data collection, the MHDO has legislative authority to equally assess fees to health care providers and payers, which are divided as follows:

- 38.5 percent hospitals (based on net patient service revenue)
- 11.5 percent non-hospital providers (based upon fixed categorical assessments)
- 38.5 percent carriers (based upon premiums written and as reported to the Insurance Commissioner)
- 11.5 percent Third Party Administrators (TPAs) (based upon claims paid for plan sponsors).

According to MHDO sources, because the fees are distributed across the system and are based on revenues, with smaller payers and providers paying less than their larger counterparts, there have been few complaints from payers and providers. Further, these fees are placed in a dedicated revenue account, not the general fund, so unexpended revenue carries over to the next fiscal year and the assessments are adjusted downward accordingly.

If Mississippi's goals for a statewide health data system are primarily public health assessment, improvement, and prevention applications, then the all payer all claims option may not be the best starting place for Mississippi, for several reasons:

- Data completeness and quality may vary more across payers than providers. Payers do not retain data for all fields submitted by providers in their data warehouses, which may limit important public health studies, such as injury surveillance applications, which rely on complete External Cause of Injury coding (E-codes).
- Global claims aggregate multiple services into one bill, masking many services.
- A small percentage of claims are still submitted manually, resulting in missing data.
- The claims data are not a complete census of the states' population, like hospital discharge data, which include self-pay and uninsured patients.
- Medicare data acquisition policies by states are under review by the Centers for Medicare and Medicaid Services (CMS).

On the other hand, if Mississippi's goals are to promote understanding about the cost and efficiency of care, then the all payer all claims data approach may have merit. Because Mississippi has fewer payers to manage, this option should not be ruled out, especially if discharge data reporting cannot be implemented, or as an augmentation to facility reporting at a future date. When combined with facility data, the all payer all claims system provides information about true costs of health care as well as episodes of care.

Maine has successfully addressed many of these issues and the Northeast states (New Hampshire, Vermont, Maine, Massachusetts, Rhode Island) are working together to resolve other technical and political issues associated with payer reporting, including the harmonization of data collection and release rules and the sharing of technology and provision of support to other states developing similar systems.

This report is prepared by the National Association of Health Data Organizations (NAHDO) and draws from its work with state health data programs. Since 1986, NAHDO has promoted the uniformity, comparability, and public availability of statewide health care data. Using state-to-

state transfer of technical assistance and lessons learned, NAHDO has built a national network or community of practice around the collection and use of hospital discharge data. The lessons learned in states with mature programs benefit those, like Mississippi, that are in the beginning stages of planning and implementing statewide reporting.

This report will lay out an implementation framework for a statewide health data program in Mississippi, based on the experience in other states. This report is intended to inform the planning and formation of data policies in the State of Mississippi and can be used as a tool to solicit input across potential data suppliers and expected data users. The report will layout options and recommendations for the following implementation steps usually associated with the implementation of a statewide health data program:

- Defining the scope of data collection initiatives;
- Governance and funding options to promote stability;
- Analysis of expected costs to implement and operative a statewide health data program;
- Consideration of data standards and reporting requirements to identify the most appropriate and efficient formats for submitting data across various data providers;
- Cost considerations for establishing and operating a health data program and the reporting burden to providers of the data.
- Analysis of data access and data dissemination requirements, including recommendations on how the data should be stored and released., including access and confidentiality provisions;

Data reporting burden and costs are important considerations for any data system. Other states have leveraged existing national standards and data systems to reduce these burdens. The planning and implementation sections will discuss the various approaches and trade-offs states must make when establishing a system.

Discharge data are distinct and different than data that many Health Information Exchanges (HIEs) or Regional Health Information Organizations (RHIOs) propose to collect in the future. The HIEs are forming to exchange clinical data across providers at the point of care. No HIE or RHIO has developed a common, shared data repository such as a statewide hospital discharge data base. A discharge data program is complementary to such efforts, and in some states, serves as a facilitator to data exchange and sharing.

Information Gaps in Mississippi

With the exception of four states – Alabama, Idaho, Mississippi, and North Dakota -- all states currently collect hospital discharge data in some form. With a near universal collection of hospital discharge data cross states, the non-collecting states face a definite set-back in the information age of increasing consumer demands for hospital care information and transparency in healthcare delivery.

An all-patient, all-payer discharge data system will fill critical information gaps in Mississippi and is consistent with the recommendation in the American Health Planning Association

(AHPA) October 2006 report commissioned by the Mississippi Department of Health. The AHPA report analysis of the State Health Plan and the Certificate of Need program was limited due to "the lack of patient-level hospital discharge and outpatient use data". The report went on to say that "patient origin and medical market data are limited" and that an "all payer patient-level hospital discharge data system is needed to permit the geographic and service-specific planning that is needed to determine how best to meet the needs of growing communities. The current state health plan does not address these questions directly. Reliable, comprehensive data are needed to permit future editions to address such questions"[9].

But health planning is not the only gap that exists in Mississippi. Based on common uses and practices documented in other states, statewide hospital data will fill important information gaps that exist in Mississippi:

- Prevalence of disease and morbidity across the population;
- Hospital and ED utilization for the state as a whole or for local subdivisions;
- Surveillance data of rare conditions where registries and surveys are impractical;
- Cost of care for specific individuals and populations and payers;
- Quality of care and access to care by different groups in the state;

In this section, we compare the health demographics of Mississippi with Arkansas and when appropriate, the average for the United States. Mississippi and Arkansas are similar in population size (2.9 million, 2.8 million respectively) and have a similar proportion of dependent populations (elderly and the young). Also similar are the numbers of Medicare and Medicaid recipients. Other similarities include educational achievement, percent of the population with disabilities, disparity between blacks and whites in terms of early prenatal care, infant mortality, low-birth weight babies, and age-adjusted death rates per 100,000 population. (Tables of population profiles of Mississippi, Arkansas, and Louisiana are in Appendix 1).

Differences are found in the availability of hospital beds—community hospital beds per thousand residents are significantly higher in Mississippi (4.5 beds per thousand residents versus 3.5 in Arkansas, versus an average of 2.8 beds in the United States). The occupancy rate is slightly lower (58% versus 59% Arkansas and 67% in the United States), but the average length of a hospital stay is longer in Mississippi. Most striking is the absence of Medicaid recipients enrolled in health maintenance organizations (0.1% versus 6.4% Arkansas, and 23.4% for the United States) (Appendix 1).

While we have these statistics available for Mississippi, it was not possible to look at other important health data about the population, except for a very small number of measures available on the CMS Hospital Care website. For example, we can only find very limited information on the CMS website; it is focused on a small number of processes of care and mortality for heart

[9] Health Services Planning Certificate of Need Regulation in Mississippi, American Health Planning Association, October 2006, (page 56). Available at http://www.msdh.state.ms.us/msdhsite/_static/resources/1845.pdf . Accessed Nov, 2007

failure, heart attack and pneumonia, but nothing on deliveries, pediatric conditions, or other costly surgeries.

In Arkansas, the state discharge data easily allows analysis of length of stay, charges, diagnoses, etc. They report on hospital stays for:

- heart transplants,
- burns,
- bone marrow transplants,
- bowel resections,
- skin grafts,
- bariatric procedures,
- bypass grafts,
- infectious diseases,
- kidney and urinary tract diseases,
- respiratory distress syndrome,
- cardiac dysrhythmias,
- spinal cord injuries,
- and, many other conditions and procedures.

For each of these types of discharges, Arkansas can analyze by:

- type of payer (Medicaid, Medicare, Private Insurance, Self-pay, Charity care);
- patient characteristics (age, gender, race, zip code of residence);
- charges for the stay;
- stays by region or city within the state; and
- the miles patients travel to get to a hospital.

Linked with other information, such as Emergency Department data, states can use the combined data to have a better understanding of how to prevent use of EDs for routine care, and whether there are areas of the state where access to ambulatory care is lacking, or to understand which injuries are most common and most costly in their state. Other linkages, such as linkages to the birth data, can provide a wide array of information about mothers and their prenatal care, characteristics of the mothers and babies, type of birth, weight of the baby, smoking and alcohol use during pregnancy, etc.

Mississippi may be able to acquire some of this information through public health registry information but not the comprehensive range of information that is available in a discharge data system. National data sources are insufficient for state-based studies, especially for special population groups, rural areas, and rare conditions. Primary data collection, in the form of surveys and registries are expensive and impose tremendous burden on the providers supplying the data. The large volume of observations in statewide discharge data are more accurate for small geographic areas than survey data, including national surveys such as the National Hospital Discharge Survey (NHDS) conducted by the National Center for Health Statistics of the Centers for Disease Control and Prevention (CDC).

The remainder of this report will lay out a framework for implementing hospital discharge data reporting in Mississippi, including the benefits or value of statewide discharge data and the planning and implementation considerations. States have adopted several approaches to implementation, each with advantages and disadvantages. What the best health data programs in the country have in common are the following:

- There is a clear sense of purpose for the data collection;
- All providers in a class report their data, either voluntarily or under a mandate;
- There is a source of stable funding to assure consistency in data and analytic workforce capacity;
- There is involvement of multiple stakeholders in decision-making and governance;
- Whether the data are reported voluntarily or mandatory, the data are not proprietary, but are considered a public or community resource;
- Data collection uses standardized methods and validation;
- There are proper policies in place to protect patient confidentiality and to prevent inappropriate use;
- Public use and research data are available for multiple users at a reasonable cost; and
- The cost of reporting is balanced with the value and benefit to stakeholders.

The Benefits of a Patient-level Statewide Health Data Reporting System

A health data program strengthens collaboration and partnerships in a state. First, key stakeholders must work together to solve the political and technical issues related to data collection and reporting. Next, the data are a tool to bring communities together to study and understand priority health care issues and problems and work together solve these problems. States use hospital discharge data in a remarkably wide range of applications. The database users are similarly diverse, including various government agencies, provider associations and individual health care providers, consumer organizations and individual patients, health care insurers and other health care purchasers (e.g., large employers), policymakers, researchers, and private-sector interests such as database vendors and consultants. In 2005, a report funded by the Agency for Healthcare Research and Quality (AHRQ), was produced by the National Opinion Research Center (NORC) and NAHDO; it was titled,, "The Value of Hospital Discharge Databases"[10]. In this report the prevalent uses of discharge data were identified and included:

- Public safety and injury surveillance and prevention;
- Public health, disease surveillance and disease registries;
- Public health planning and community assessment;
- Public reporting for informed purchasing and comparative reports;
- Quality assessment and performance improvement;
- Health services and health policy research;
- Private sector and commercial applications;
- Informing policy deliberations and legislation; and
- Employer and Purchaser Use of Administrative Data

Public safety and injury surveillance and prevention

Inpatient discharge data are routinely used to study and monitor issues of public safety, including injury surveillance and prevention.[11] Statewide data permit the targeting of communities and populations suffering the burden of preventable injuries and can be used to inform and evaluate policies. The National Highway Traffic Safety Administration (NHTSA) has awarded participating states competitive cooperative agreements, the Crash Outcomes Data and Evaluation System (CODES). By linking existing data sets, such as hospital discharge, accident reports, death certificates, and ambulance information, states can fulfill expanded data needs without the additional expense and delay of new data collection. The linkage improves data

[10] Schoenman, J A, et al. The Value of Hospital Discharge Databases. NORC at University of Chicago and National Association of Health Data Organizations (NAHDO); 2005.
[11] Schoenman, et. Al.

quality and states benefit from state-specific injury and financial outcome information about motor vehicle crashes.[12] States that have added Emergency Department (ED) data reporting have an even more robust capacity to track rates of injury and costs associated with injury, including valuable information about non-fatal injuries that result in hospitalization.

Public health, disease surveillance and disease registries

Hospital discharge data support a variety of disease surveillance systems, such as chronic disease prevention and control, to measure the burden of chronic diseases on a population and support planning and evaluation of disease programs and interventions. In states with disease registries that are independent of discharge data reporting, the discharge data provide a reference data base to assess compliance to specific disease registry reporting. Discharge data also are often used to estimate the financial burden of specific illnesses or conditions, an important component to garnering support for community assessment and improvement interventions across stakeholder groups.

Public health planning and community assessment

Inpatient discharge data are also an important source of information for public health planning and community needs assessments. The detailed data can supplement existing data bases, such as Certificate of Need and hospital financial data, to monitor and measure the performance of the health systems and assess potential impact of hospital mergers, closures and plan for future needs based on population projections. A baseline and trend of hospital utilization patterns is important to such projects. Community assessment in many states, relies on the discharge data in a given geographic area to evaluate inpatient and ED admission rates for special populations by diagnosis, diagnosis severity, and causes of admission or injury. Further, these data can identify services that are lacking in a community and can supply information for plans for future allocation of resources.

Public reporting for informed purchasing and comparative reports

Transparency and quality reporting laws have been enacted in over 30 states[13], with the intent to stimulate consumer decision making and market forces to improve the cost and quality of health care. At least 16 states produce a state-sponsored hospital quality consumer report or website (NAHDO) and a growing number of hospital associations and private health care organizations are publishing price and quality reports for consumer audiences.

Pursuing Quality Data: Using Statewide Health Data as the Source of Quality Indicators, November 11, 1987

"Hospitals can request of the state agencies, which are gathering and analyzing data, what they want and need in the data analyses. This is very important".

Leslie Roos, Ph.D., Professor, Drake Center for Management Studies and Dept of Medicine, University of Manitoba, November 1987.

[12] Marta Benavente1, Michael A. Knodler1, Heather Rothenberg .Case Study Assessment of Crash Data Challenges: Linking Databases for Analysis of Injury Specifics and Crash Compatibility Issues. Transportation Research Record. 2006; 1953:180-86.

[13] National Conference of State Legislatures, State Legislation Relating to Disclosure of Hospital and Health Charges, November 2007 (http://www.ncsl.org/programs/health/Transparency.htm)

Quality assessment and performance improvement

Providers use statewide measures of quality and utilization for internal quality improvement programs. The data may or may not be publicly reported. Some state health data programs provide hospital-specific comparative reports or statistics for the providers' internal use. Often these reports, derived from statewide data, will include the hospital's statistics alongside peer group and state rates.

Health services and health policy research

Hospital discharge data are used extensively in basic health services research, where the goal is to establish a foundation of knowledge in health services or health policy. Numerous studies have used discharge data to examine how systems for organizing, financing, and delivering health services affect hospital utilization, costs, and outcomes (NORC, pg 27). The presence of a statewide health data program also benefits research institutions in that state, by providing a means to produce preliminary studies and statistics to define the scope of a local problem, which makes that institution more competitive when applying for national grants. Reduction in health disparities is a national priority, and Mississippi will benefit from studies targeted to study the relationship between racial or ethnic status and health care utilization and outcomes.

Private sector and commercial applications

Discharge data are used for economic and market applications, ranging from hospital strategic planning to the development of proprietary tools that generate information for purchasers, providers, and consumers. In states that make public data sets available, the potential revenue from sales of these data to consulting firms, health care providers, or health information management vendors can be significant, while providing the private sector with valuable information, such as market share reports and patient origin reports so important to hospital strategic planning activities. Competitive bidding and enrollment, comparisons of utilization and outcomes statistics for peer hospitals are also possible with statewide data. In states where the discharge data are publicly available to the private market, there have been demonstrated benefits, including stronger price competition, community-wide and hospital-specific quality improvement initiatives, and more aggressive purchaser negotiations.

Informing policy deliberations and legislation

A source of census data on hospitalized patients has the potential to guide policy decisions and evaluate legislative initiatives, including health care reform. By establishing a baseline and trend of hospital utilization, the impact of legislation can be assessed on specific populations and geographic areas; in addition, the scope of a particular problem can be more readily defined.

Employer and Purchaser Use of Administrative Data

Healthcare purchasers have supported state data collections and have shown that the data are useful for understanding the range and the variation across hospitals in utilization patterns, length of stay, re-admissions to hospitals, and pricing. Some employer purchasers have also produced information for employees to assist them in their choice of provider; although many plans have limited choices for the consumer.

Healthcare purchasers, such as the Alliance Healthcare Cooperative in Madison WI, have also initiated quality reporting and pay for performance programs with hospitals. They have used administrative data to supply the necessary information. Other business coalitions utilizing administrative data include the Pacific Business Group on Health and the Niagara Health Quality Coalition. Large state purchasers, including Medicaid programs and state employee purchasing groups have utilized state data to examine differences in utilization across their state. In states with all payer all claims reporting from payers, comparative information about the cost of episodes of care and negotiated discounts can be studied.

An Integrated State Data System in South Carolina

The Office of Research and Statistics of the South Carolina Budget and Control Board (ORS) maintains the South Carolina inpatient hospital billing system, the outpatient surgery data system, the emergency department visits data system, the home health data system, health manpower, health education and facilities data systems, the Master File data system of clients of State Health and Human Services agencies, and is the State Data Center responsible for Census products and analysis. ORS is a leader in establishing a statistical data warehouse, linking and integrating data across agencies and sectors, that supports program evaluation and policy and program decisions in South Carolina.

> Today, the South Carolina data warehouse has been used to evaluate the effectiveness of program. Examples abound and include an evaluation of the Communi-Care program, which provides free prescription drugs to the uninsured. By reducing hospital inpatient and Emergency Department utilization, Communi-Care's impact could be quantified. Similar savings have been documented for senior receiving Meals on Wheels and for mental health case management populations.

Beginning with inpatient data in 1986, ORS has expanded data collection to include Ambulatory Surgery and Emergency Department data, which, along with Medicaid claims data, serve as the foundation for the integrated state system, which today encompasses social services and health information. South Carolina has created tools and methods for disseminating data and encrypting or de-identifying data and makes these available to other states at no or low cost.

Ideally, an integrated data system attempts to capture the full range of health and human services experiences of the populations being served, so a range of information from administrative data systems can be relevant and yield insight into many areas of interest[14].

Information on hospitalizations, emergency department utilization, outpatient or ambulatory surgeries, office visits, home health, and nursing services provide a comprehensive picture of health services utilization. By integrating hospital data with other sources of data from Government, private-sector, and not-for-profit programs, the prevalence of many health conditions and can more fully assessed through a population's health services utilization.

To illustrate the types of available administrative datasets that can potentially be integrated, Table 2 lists datasets that are routinely integrated in South Carolina.

---

[14] Tools for Monitoring the Health Care Safety Net, "Integrated State Data Systems", 2003, Walter P. "Pete" Bailey, M.P.H., *http://www.ahrq.gov/data/safetynet/bailey.htm*

Table 2. Data Sets Integrated in South Carolina[15]

Agency or Program	Information Available
Census Data	Decennial and estimates/projections
SC First Steps	Needs assessment data for children age 5 and
SC Department of Disabilities & Special Needs	Client Database
SC Vocational Rehabilitation	Client Database
SC Department of Mental Health	Client Database
SC Labor, Licensure, and Regulation	Licensed Physicians Database
SC Department of Health & Environmental Control	Vital Records, Emergency Medical Services, Ambulance, BabyNet, Children's Rehabilitative Services, various Maternal & Child Health files
SC Department of Public Safety	Motor Vehicle Crashes
SC Department of Juvenile Justice	Juvenile Justice Referral Database
SC Private Healthcare Providers	Inpatient hospitalizations, emergency department visits, outpatient surgeries, home health visits
Free Clinics	Client Database
SC Department of Education	Student demographics, Palmetto Achievement Challenge Test (PACT, a standardized test) and Exit Exams, 1st grade readiness
SC Department of Social Services	Temporary Assistance to Needy Families (TANF), Wage Match and Work Support, Food Stamps, Foster Care Tracking, Child Protective Services, Adult Protective Services, Child Support Services
SC State Law Enforcement Division	Criminal History File, Crime Incidents
SC Department of Health & Human Services	Medicaid claims data, Child Care Voucher System, Community Long Term Care, Division on Aging
SC State Health Plan	Medical claims data for State employees
SC Department of Alcohol & Other Drug Abuse Services	Client service files

[15] Bailey, 2003.

As South Carolina has demonstrated, the integration of data from multiple programs and sources can take the policymaker far beyond a single data source. Integrated data promote an understanding the underlying problems of the populations receiving services and facilitate the evaluation of the impact of the services that the program provides. South Carolina has used integrated data to look across the system, build partnerships, and improve quality of life and foster independence on those who rely on the safety net.

The Value of Emergency Department Data Value

The AHRQ report, the Value of Hospital Discharge Databases, studied the uses of inpatient data for a variety of market and policy applications. Through its work with state health data programs to expand beyond inpatient reporting, NAHDO has identified the value of ED and ambulatory surgery data.

As ED data collection grows across the country, states are realizing the tremendous utility of this data set. Emergency Department data provide important information in a growing number of states. The ED data set is being used for policy and planning. Because it includes data on all patients and all payers in a state, the ED data provide a unique window into the performance of the health system as a whole, as well as patient subgroups and communities within a state. States with ED data are able to research the following issues:

- Identify patterns of care including, but not limited to, injury or disease classes (asthma, heart attack and stroke).
- Identify patients and high-risk groups receiving emergency department services distributed by age, race/ethnicity (if collected), gender, and payment source.
- Identify seasonal deviations and other patterns of change over time for emergency department utilization. This would also include disease or injury specific groupings.
- Identify high-risk groups and neighborhoods that have high emergency department visit rates
- Identify the co-morbid conditions that impact care outcomes including, but not limited to, injury or disease classes (asthma, heart attack and stroke).
- Identify the distribution and cost of services provided to emergency department patients. This would include medication administration, ancillary services and procedures that were administered or performed during the visit.
- Track patterns of care for emergency department visits distributed across geographic regions of the state over time and by hospital type.
- Identify outcomes by source of admission for emergency department visits.
- Track the location of injury episode and exposure resulting in an emergency department visit as well as the source of admission for emergency department services.

- Track emergency department visits that lead to an inpatient admission or subsequent re-admissions for emergency or inpatient services.
- Identify other contributing factors (e.g., severity, secondary complications, specialist referral, etc.) impacting the cost of an emergency department visit.
- Analyze the composition and cost of resources consumed for emergency department visits for primary-care sensitive conditions better treated in other settings.
- Identify the distribution of patients in payer groups including the uninsured, and homeless to detect differences in usage and costs for emergency care.

The Value of Ambulatory Surgery Data

Many procedures that had routinely been performed in hospitals are now also performed in ambulatory surgery centers. While some have completely moved to outpatient settings others are performed in both—costs for the procedures may vary widely between hospital and ambulatory surgery. Analysis of the differences in charges requires one to look at patient severity while assessing the differences in charges. In some areas of the state use of an ambulatory surgery center (ASC) may strictly be a result of location, in other areas the ASC's are competing directly with hospitals for patients.

Often the procedures performed in outpatient settings are considered elective rather than emergent. That is, they are not entering in an emergency, but are scheduled. One way to distinguish severity of cases treated would be to assess whether the patient was non-emergent versus entering through the emergency department. Alternatively, you could also look at the number of co-morbid conditions that were present on admission for patients in the hospital or surgery center. Thus, one would find out if the cases served by hospitals needed emergency surgery or had more co-morbidities and therefore, were inappropriate for surgery in an ASC.

Outcomes of care are also important to analyze especially as more complex procedures are done at ASC's—are patients being admitted to the hospital from ambulatory surgery centers due to complications of care? Alternatively, are hospital patients with the same condition/procedure being re-admitted for problems with care delivery? These issues may trigger licensing or regulatory initiatives—without data patients may be harmed, or alternatively ASC's may save considerable dollars for payers and citizens by providing safe and less expensive care.

Planning Considerations

The establishment of a health data program involves a series of decisions. Each decision influences the reporting burden and costs, system sustainability, and how the information is eventually disseminated and used. Fortunately, the experience of successful state programs informs the decisions of emerging programs, such as Mississippi, about the following:

- The scope of the data collection
  - Data formats
  - Data elements
- Governance and Funding
- Data ownership and control
- Data access and dissemination policies

This section will discuss the various options and make recommendations that will guide guide the planning considerations and discussions in Mississippi. This report is written with the assumption that a statewide health data program will be established in Mississippi. Because of its unique demographics and population health challenges that are well-documented, Mississippi would benefit tremendously from a state health data program. Statewide reporting, beginning with hospital inpatient and Emergency Department data would provide a window into health care use across all providers and payers, for all hospital encounters in Mississippi. Not only would it establish a baseline of utilization of the most costly health care in Mississippi, the ongoing collection of such data will provide important trend data and information. With the capture of key data, such as a unique patient identifier and patient race and ethnicity with the discharge data, Mississippi will be well on the way to establishing a source of community data to stimulate a community assessment and improvement collaborative, guide policy development, and promote research. A comprehensive reporting program in Mississippi also has the potential to fill important data gaps nationally, including health utilization information for the state in the AHRQ National Healthcare Quality and Disparities Reports, mandated annually by the U.S. Congress[16].

Scope of Data Collection

Recommendation: Mississippi should implement a hospital data reporting system to be accomplished in two stages, inpatient then outpatient, to be followed by other providers at a later time.

To reduce provider reporting burden, the data formats should be aligned with the Uniform Bill (UB) standard maintained by the National Uniform Billing Committee (NUBC)[17].

[16] The National Healthcare Quality Report is a national comprehensive effort to measure the quality of health care in the U.S. Developed by the Agency for Healthcare Research and Quality, the report includes utilization and quality information from the Healthcare Cost and Utilization Project, HCUP. States like MS that do not supply data to the HCUP do not have state-level data included in this report to Congress. www.ahrq.gov/QUAL/nhqr03/nhqrsum03.htm.

[17] The National Uniform Billing Committee (NUBC) was brought together by the American Hospital Association (AHA) in 1975 and it includes the participation of all the major national provider and payer organizations. The NUBC was formed to develop a single billing form and standard data set that could be used nationwide by institutional providers and payers for handling health care claims.

This recommendation is consistent with the recommendation in the American Health Planning Association (AHPA) October 2006 report commissioned by the Mississippi Department of Health. Inpatient data is relatively comparable across hospitals and there are a wide range of data management and analytic tools available.

Lessons learned in other states:

The New York State experience implementing their emergency department data collection system is a text book example of the value of standards. The legislature had been convinced by an emergency department doctor in New York City that emergency department data was necessary to address the financial and health issues with a rapidly growing incidence of juvenile diabetes. The legislature unanimously passed the legislation, but provided no additional funding for the system design and development.

ED data collection, in conjunction with the inpatient data, provides valuable information about health care access and utilization for the population. Ambulatory surgery data, from hospitals and freestanding ambulatory care centers, provide important information about the cost and quality of care that is shifting from inpatient to outpatient settings. However, because of the unique issues[18] surrounding the collection and use of both ED and ambulatory surgery data sets, it is advisable for Mississippi to direct its initial data development efforts to inpatient data. There are solutions for the challenges related to ED and ambulatory surgery data, but these are best solved once the health data program has been established and has experience in working with discharge data.

Data Format

Health data organizations that maintain statewide health data programs have proven that leveraging existing data flows from hospital automated billing systems is a cost-effective way to gather detailed data on hospital utilization. The de-facto national standard for state reporting systems is the UB data maintained by the National Uniform Billing Committee (NUBC). The UB-04 is the current version, which replaced the UB-92 in May, 2007.

As of July 2007, a growing number of states were collecting the data in UB-04 format and were also including the Present on Admission indicator, following the lead of CMS and its requirement that providers report POA for Medicare reimbursement. It is important to note that historically the UB has been used for multiple purposes, most notably for institutional claims and state reporting. Over time, the addition of diagnosis and procedure codes, external cause of injury codes, and most recently (in the UB-04) the present on admission indicator, were data elements recommended by the National Committee on Vital and Health Statistics as necessary for state discharge systems. Gradually these elements have been added to the billing form and adopted for billing purposes.

With the implementation of the UB-04 additional data elements, such as Race and Ethnicity and Do Not Resuscitate Order, have been added to the UB for the sole purpose of meeting state reporting requirements. It is important to note that the UB-04 now officially supports data necessary for institutional claims as well as state reporting purposes. The electronic format

[18] States that collect ED data have grappled with the following issues: defining an ED visit, procedural coding variation across providers. States that collect ambulatory surgery data must address the scope (hospital-based and/or freestanding, defining an ambulatory surgery visit, procedural coding variation across the system, and grouping of the data for analysis.

designed to transmit the UB content is the ANSI ASC X12 837 institutional implementation guide. This standard implementation guide is mandated by HIPAA for institutional claiming purposes. This is significant because the ANSI ASC X12 837 institutional implementation guide was approved as a national standard by the ANSI consensus process as well as by the federal rule making process. For state reporting there is a companion ANSI ASC X12 approved implementation guide, the Health Care Service Data Reporting Guide. It is important to note that there has been much effort to align the common data elements from the Institutional HIPAA Claim and the Health Care Service Data Reporting implementation guides.

The value of implementing standard solutions was highlighted at the 2007 NAHDO annual meeting. During the Present On Admission panel discussion, the vendor representative made it clear that implementing 50 different state solutions added considerably to the cost and the complexity of the vendor solutions, and parenthetically to what states would pay for vendor solutions.

With the rapidly escalating costs of health care along with large questions about the quality of care being delivered, there is an increasing need to measure the quality of the health care delivered. As a result there has been much work to develop process and outcome quality measures to assess the strengths and weaknesses of the health care delivery system in the United States. It is clear that to compare the quality of care across the country standards become increasingly important. In addition, to better assess the quality of care additional clinical data is needed. The most significant evidence of this is the Medicare requirement to collect the Present on Admission Indicator to help determine payment algorithms for some healthcare associated conditions (pressure ulcers, infections, falls) later in 2008. The organization most responsible for developing standards to report clinical data is Health Level 7. As these clinical standards evolve, the state reporting uses of this same data would be best advised to use the lessons learned from implementing the billing standards for reporting purposes. The data should be collected once and used many times. It takes persistence and patience to prove the utility of using data for multiple purposes. States need to leverage existing data systems and information flows to collect the necessary information while minimizing the burden on the collection source.

Economies of scale can occur when single collection solutions can be applied to a multiple of uses. The integrated use of clinical and administrative data standards offers that possibility.

While some have suggested that the Electronic Health Records (EHRs) will be the principal source of data for quality measurement, recent evidence indicates that these systems are not ready for that role and won't be for some time (See GAO Report, 2007[19]). Most hospitals have multiple legacy systems for various components of the medical record, for example, x-rays, laboratory information, pharmacy may all be in separate electronic systems, making it very complex to retrieve—each of these separate systems also have different standards for the data. What is more likely to occur (and is already happening in some states) is that a select stream of electronic data elements from EHR's will enhance the administrative data for submission to states. See Chapter 5: Strategies for Adding Value for more detail.

[19] US Government Accounting Office. Hospital Quality Data: HHS Should Specify Steps and Time Frame for Using Information Technology to Collect and Submit Data. Report to the Committee on Finance, U.S. Senate, GAO 07-320, April 2007.

Hospital's UB claims processing can be effectively used to collect inpatient discharge data;

The UB-04 has become public health reporting friendly. During the analysis leading up to the adoption of UB-04 each form locator on its predecessor, the UB-92, was reviewed for its continued utility. The UB-92 dedicated several form locators for local use. The analysis revealed that many of these local use form locators were used to support state inpatient and outpatient discharge data reporting requirements. This led to great variation on how data, such as race and ethnicity, was reported across the country.

In an effort to reduce this variation it was decided that the local use form locators would be eliminated in the UB-04. Since many states were using these local use form locators to support their state discharge data reporting requirements, it was decided that the UB-04 needed to also support state reporting requirements in lieu of providing form locators for local use. The result of that decision is that the UB-04 now supports the reporting of a national Race and Ethnicity code set (maintained by CDC and used in the 2000 Census), a national code for Do Not Resuscitate, additional external cause of injury codes, the Present on Admission Indicator, amongst other things. These additions, which NAHDO and the public health community facilitated, are important to state health data programs.

Prior to the development of the UB-04 getting a data element approved by the NUBC for state reporting purposes only was always a challenge. When the NUBC decided to remove local form locators to reduce regional variability in reporting UB data, the committee realized this could not be done unless state discharge data reporting needs were also accommodated. There were just too many states that used the UB to meet their state reporting needs. That included the ubiquitous use of local use form locators to satisfy state specific requirements. That realization by NUBC led to significant changes to the UB-04 resulting in the public health discharge data friendly form it is today.

Recommendation: Mississippi should collect the UB-04 core data elements. Collection of these national standards will assure the comparability and utility of the information and will also minimize the provider reporting burden.

Data Elements

In choosing the data elements to support the particular use cases for which the data is collected states use one of two approaches.

- Use standards, typically the UB billing standard as maintained by the National Uniform Billing Committee (NUBC). It is these same standards that are the core data content of the HIPAA institutional claim format (ANSI X12 837 Institutional).
- Use of state defined proprietary idiosyncratic data elements designed for specific state purposes.

It should be noted that practically all states do abstract additional data for disparity assessments, readmission rates, and patient safety use cases. Examples of data elements abstracted for these purposes would be Race and Ethnicity, Personal Identifiers, and Present on Admission Indicators amongst other things.

Some of the advantages of using such national standards are the comparability—the same data definitions are common across the country, the availability of "out of the box" vendor solutions based on national standards, and the lessened reporting burden for the provider community using the same standards as required by HIPAA for claims and state reporting purposes.

The main advantage of using the proprietary idiosyncratic data elements is the flexibility of being in control of the data elements to meet specific state needs. This flexibility will create additional burden on healthcare providers, however, and NAHDO does not recommend this approach. The use of national standards, currently UB-04 provide the following advantages:

- Comparability of information across providers and states;
- Reduced provider reporting burden, as most providers have the capacity to produce a uniform bill;

Table 2. Core UB-04 Billing data elements

• Patient Control Number • Medical Record Number • Type of Bill • Provider Identifiers (including National Provider Identifier) • Statement Covers From and Through Dates • Patient Identifier • Patient Name • Patient Address (including city, state, and zip code) • Patient Birth Date • Patient Sex • Admission / Start of Care Date • Priority (Type) of Visit • Point of Origin for Admission or Visit (Replaced Source of Admission) • Discharge Hour • Patient Discharge Status • Condition Codes (Includes amongst other things – Do Not Resuscitate Indicator) • Accident State • Occurrence Codes (Includes amongst other things – Accident dates) • Occurrence Span Codes (Includes amongst other things – SNF level of care dates) • Value Codes (Includes amongst other things – Newborn birth weight) • Revenue Codes • HCPCS / Accommodation Rates / HIPPS Rate Codes • Service Date	• Health Plan Identification Number • Release of Information Certification Indicator • Assignment of Benefits Certification Indicator • Prior Payments – Payer • Insured's Name • Patient's Relationship to Insured • Insured's Unique Identifier • Insured's Group Name • Insured's Group Number • Treatment Authorization Code • Employer Name (of the Insured) • Principal Diagnosis Code and associated Present on Admission Indicator • Other Diagnosis Codes and associated Present on Admission Indicator • Admitting Diagnosis Code • Patient's Reason for Visit • Prospective Payment System (PPS) Code • External Cause of Injury and associated Present on Admission Indicator • Principal Procedure Code and Date • Other Procedure Codes and Dates • Attending Provider Name and Identifiers • Operating Physician Name and Identifiers • Other Provider Name and Identifiers (includes Other Operating Physician, Referring Provider, and Rendering Provider) • Service Units • Total Charges • Non-covered Charges

For an example of data submittal requirements, go to the New York SPARCS Submission Manual at http://www.health.state.ny.us/statistics/sparcs/index.htm

Recommendation: The legislation in Mississippi should reference "national standards" for reporting requirements, not specifically referencing UB-04, as standards change. Administrative rules should specify the UB-04 core data elements as required data elements, to align with national standards and reduce provider reporting burden.

Developmental Data Elements for Public Health and Quality Measurement

States are adding tremendous value to their hospital discharge data by collecting data elements that support analyses, such as race and ethnicity and key clinical data. A unique patient number, also is valuable for purposes of data quality, longitudinal analysis, and data linkage. State programs that did not include these data elements with initial reporting are revising their collection policies to permit the addition of these key variables, once they recognized how the utility of the entire data set improves with these additions.

Race and Ethnicity

The Institute of Medicine's Report, "Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care" illuminated what had been apparent to many--that health outcomes vary by race and ethnicity. *Unequal Treatment* established that disparities occur due to differential treatment within clinical settings of members of different races and ethnicities. In so doing, the report marked a turning point for the study of health care disparities.[20]

A growing number of states require race and ethnicity as a part of their hospital discharge data requirements. These data elements have been controversial in the past, due in part to the variation in coding across hospitals and sensitivities surrounding race and ethnicity issues. But this is changing. With the promulgation of the Office of Management (OMB) Directive 15 standards into the national ANSI X12N billing standards and the American Hospital Association's Health Research Educational Trust's (HRET) health disparities project, the barriers have been diminishing. The Health Research and Educational Trust, with support from the Commonwealth Fund, launched the online HRET disparities toolkit in February 2005 to assist hospitals and other healthcare providers in collecting information on race, ethnicity, and primary language from patients.[21]

Both mandatory and voluntary health data programs are capturing the data in response to national, state, and industry initiatives directed to eliminating health disparities in health care. States do vary in their approach to race/ethnicity data collection: 1) states with mandatory reporting include race/ethnicity as part of the data reporting requirements; 2) states with mandatory reporting do not require or request the voluntary submission of race/ethnicity with hospital discharge data; or 3) stats without mandatory reporting include race/ethnicity as part of the voluntary reporting requirements. Regardless of approach, all states recognize the limitations and sensitivities surrounding these data elements and use them cautiously in their analyses. One of the greatest challenges states have is the issue of "observed" race/ethnicity as coded by the hospital admitting clerk versus "self-reported" by the patient.

The National Research Council of the National Academies report, "Eliminating Health Disparities: Measurement and Data Needs" recommended the collection of data on race,

[20] Institute of Medicine. Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care; 2003
[21] Health Research Educational Trust, 2005, Available at http://www.hretdisparities.org/. Accessed Nov. 2007.

ethnicity, socioeconomic position, and when feasible, acculturation and language use".[22] In order to meet the data collection and analysis needs for measuring and reducing disparities, both the public and private sector will need to work together to develop state data and design interventions.[23]

Because of the very real concerns about race and ethnicity data, many states exclude these data elements from their public use files.

Mississippi's demographics are such that the capture of race/ethnicity with discharge data has the potential to yield important information for research.

Recommendation: NAHDO recommends that Mississippi include race/ethnicity as a part of the mandatory reporting requirements. This recommendation is consistent with the report, Eliminating Health Disparities: Data and Measurement Issues CITE)[24] Mandatory reporting is more likely to result in reporting compliance.

However, due to the sensitive nature of these variables, and due to limitations related to data reporting practices, NAHDO recommends these variables not be released publicly until sufficient evaluation and data quality assessment is completed in future years.

States that require the reporting of race/ethnicity as a part of the discharge data submission show higher rates of compliance than states that collect the data voluntarily or that do not require resubmission of the data if it is missing or invalid (96 percent compliance for mandated reporting versus 83 percent compliance for voluntary submission of this field) (Table 3).

Table 3. Collection of Race and Ethnicity Data by Collection Method (mandatory versus voluntary)[25]

Mandatory Reporting			Voluntary Reporting		
Range	% missing	% compliance	Range	% missing	% compliance
Average	3.2%	96.8%	Average	17.07%	82.9%
Minimum	.076%	92.6%	Minimum	0%	23.4%
Maximum	7.45%	99.9%	Maximum	76%	100%

Unique Patient Identifier

The purpose of most state health data programs is for public disclosure of information, but without violating patient confidentiality in any way. State health data programs have balanced these seemingly opposing functions through a combination of statistical, methods and policies in such a way that there is no breach of patient identity by any state health data program on record.

[22] National Research Council of the National Academies. Eliminating Health Disparities: Measurement and Data Needs;2003, page 95.
[23] National Research Council of the National Academies, Eliminating Health Disparities: Measurement and Data Needs;2003, page 99.
[24] National Research Council of the National Academies. Eliminating Health Disparities: Measurement and Data Needs;2003
[25] Data Sources*: 1998 Data Inventory for HCUP Partners (1999 collection), 1997 HCUP statistics*, Telephone interviews with states, January 2000

Most state health data programs capture a unique patient identifier in the form of the patient's social security number (SSN), or a derivative of the SSN, which is encrypted (de-identified) into an agency-assigned number. For example, a derivative number may be a synthetic number generated from the patient's date of birth, part of the SSN, gender, and (in states that collect it, patient initials or name). States without the means to collect and encrypt this information are not able to fully utilize the data they collect because they are unable to identify readmissions, follow patients across hospitals and settings, or link hospitalization data to other public health data sets, such as mortality or birth files to analyze outcomes. The de-duplication of data is hindered, so data quality also suffers.

Health data programs that collect the patient's SSN treat this data element as a "strictly confidential" data element. In other words, whether prohibited by legislation or data policies, it is never released in its raw or identifiable form. It must be noted that patient SSN is not without problems. An estimated 20 percent of the patients do not have a SSN (newborns and some immigrants). The patient may not know their SSN at the time of admission (or be conscious to supply the number) and no provider can deny treatment and admission to a patient who does not supply the SSN. It would be preferable to have an alternative patient identifier, but one does not exist. The federal government, under the Administrative Simplification provisions of the Health Insurance Portability and Accountability Act of 1996, was to promulgate standards for the development of a unique national patient number. For reasons, both political and economic, this has not occurred. Therefore, it is left to states and private entities that collect and manage health care data, to address the issue of patient number assignment and protection. Few states have the technical infrastructure to assign their own master patient index (MPI) or number across providers and systems, nor the funding to support such an undertaking.

Therefore, the SSN is the logical solution for many data initiatives. And, despite concerns, the SSN is used widely for purposes other than social security. The federal government, state and local governments, and private businesses all widely use SSNs. No federal law imposes broad restrictions on businesses' and state and local governments' use of SSNs when that use is unrelated to a specific federal requirement[26]. Currently, governments and businesses frequently use SSNs to identify and organize individuals records and to exchange information in order to verify information on file, to coordinate benefits or services, or to ensure compliance with certain federal laws. Private health care organizations, historically, have always asked patients for an SSN, but they do not deny services if a patient refuses to provide the number.[27]. With the emerging threat of identity theft, many insurers are now assigning proprietary subscriber numbers.

The unique identifier combined with the patients date of birth, zip code, and gender are important for improving data quality (de-duplication) and data utility (the linkage of data with other data sets such as ED or other public health data sets (such as birth certificate data for birth outcomes). Because discharge data have limitations, which include the lack of detailed clinical data, and because the cost of clinical abstraction is high,, many health data programs (or their public health partners) conduct data linkage to fill data gaps.

---

[26] United States General Accounting Office, Testimony Before the Subcommittee on Government Management, Information, and Technology, Committee on Government Reform, House of Representatives "Social Security: Government and Other Uses of the Social Security Number are Widespread", GAO/T-HEHS-00-120, May 2000. http://www.gao.gov/archive/2000/he00120t.pdf

[27] GAO, May 2000.

The discussion under “Data Policies and Dissemination” and “Filling Data Gaps” sections describe how states protect the patient’s privacy and link other data to add value to the discharge data.

Recommendation: NAHDO recommends the inclusion of key patient demographic data elements as a part of both inpatient and outpatient discharge data requirements in Mississippi. These data elements should include, at the minimum, patient SSN, hospital medical record number, date of birth, and gender. If feasible, patient name and address should be collected to facilitate the assignment of a unique number, but also to facilitate geocoding and public health applications.

Unique patient identifiers, such as patient SSN, name, address, and date of birth should be excluded from all public reports in their raw form and the release of identifiable information be carefully regulated for authorized research and public health applications

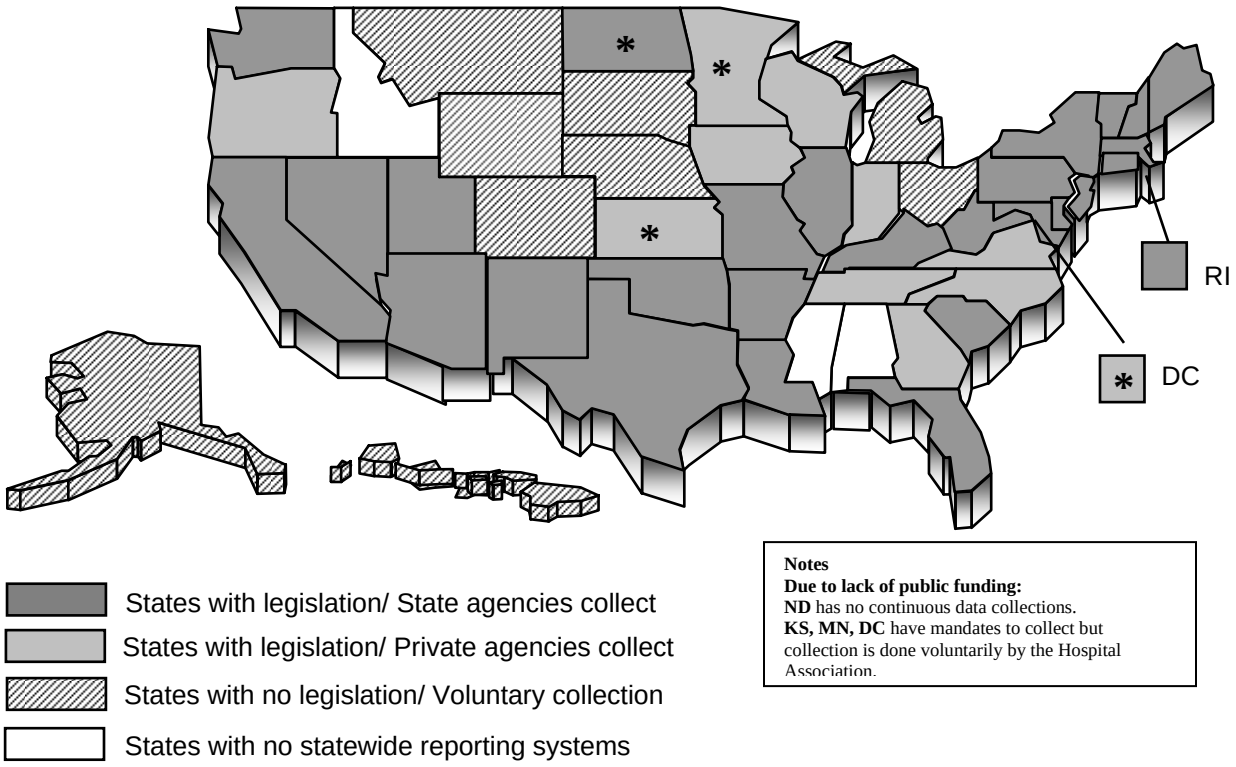
Governance and Funding

Once the information priorities and data needs are determined, there are a series of decisions that must be made about how the data collection and reporting will be organized and funded, and how it will operate. Perhaps the most important decision influencing a state health data program is how it will be governed. This decision drives almost all of the others, including the core issue of data ownership and control. States have taken several approaches to program implementation and each state must also select the structural and operational approach that is the best fit with their local environment.

The first issue is whether there is political support for enacting mandatory reporting. Thirty-nine states have legislative mandates that guide the policies for data collection and disclosure. In states where legislation that governs the structure and funding of a health data program has not been enacted, voluntary initiatives have been established. These voluntary initiatives are filling critical data gaps, but the issues of data access, funding, and ownership can be uneven, even contentious.

Figure 4 illustrates the various approaches states have taken to inpatient data collection across states.

Figure 4: State Data Collection Approaches



An ongoing, stable source of funding is essential to continuous data collection. A handful of states have enacted legislative mandates but have either not funded data collection or have phased out funding. In North Dakota, the state collected data from public and private payers for several years and used the data to study episodes of care and patient utilization patterns, but data collection terminated when funding was eliminated. Other states or jurisdictions have enacted mandates that were not funded. In those cases (Kansas, Minnesota, and the District of Columbia) the public agency has negotiated an arrangement with the hospital association in that state/jurisdiction to provide inpatient data. In these states/jurisdictions, the public agency is restricted in how they can use the data, thus limiting its utility for many stakeholders.

Each approach to data reporting has strengths and weaknesses, summarized below in Table 4. While a mandate does not assure compliance to reporting requirements, or even funding to support the health data program, legislation is a useful tool to "level the playing field" and require that all providers participate. A mandate also presumes transparency in system implementation, report methodologies, and disclosure practices, as the policies for each of these activities is spelled out through a public process. However, legislation may not provide the flexibility that a voluntary initiative may have, especially if the legislation is too specific or restrictive. Appendix 2 contains a table summarizing state and private health data programs.

Table 4. Strengths and Weaknesses of Mandatory and Voluntary Reporting Initiatives

Type of governance	Strengths	Weaknesses
Mandatory reporting	Assigns authority All providers are required to report Uniformity and timeliness in reporting can be compelled Provisions for non-compliance States the framework for data reporting and access State law exempts the health data program from HIPAA Privacy Rule regulations, in most cases	Legislation may be too specific and therefore inflexible Changes in scope and data elements may require changing state law Program funding is subject to political pressure and data funding competes with other state programs May be enacted as an unfunded legislative mandate, but this is rare
Voluntary reporting	Flexibility may be more responsive to market information needs Funding is not dependent on the legislative process May meet with less resistance from the provider community	Community-based participation is difficult to sustain over time The process may not be transparent in data collection and analytic methods Private initiatives are less likely to reveal hospital-level information to the public (or restrict or suppress this information) Providers that refuse to participate cannot be compelled to do so, which threatens the utility and sustainability of the data initiative If hospital association-based, the collection of data from freestanding and outpatient facilities is less likely

Type of governance	Strengths	Weaknesses
		Data access policies may be uneven and restrictive for certain users or uses Funding is not subject to the public process and the major funder could influence the scope and process Data reporting and quality cannot be required.

Both the mandatory and the voluntary programs provide useful information for local and national stakeholders, when properly designed. Other commonalities include:

- Most collect uniform billing data (currently Uniform Bill 04 (UB-04)
- Most collect inpatient data from acute care providers/hospitals statewide
- All provide the data back to providers/hospitals for market and quality improvement practices
- Most provide the data to public health programs for surveillance activities, such as injury or chronic disease monitoring and interventions
- Many state health data programs eventually provide data for national initiatives, such as the AHRQ Healthcare Cost and Utilization Project (HCUP), the National Highway Traffic Safety Administration's Crash Outcomes Data and Evaluation System (CODES), and/or the Center for Disease Control and Prevention's Environmental Public Health Tracking Network.

Appendix 3 includes the citations of the various states' discharge data reporting legislation. Typically, legislation in states with successful reporting programs will contain the following provisions:

- Policy making authority and oversight by a commission or committee that is comprised of key stakeholder representatives;
- Compliance provisions and penalties for non-compliance;
- Flexibility in the scope and content of data reporting requirements;
- Reference to a national standard (e.g. uniform bill) rather than specific data element listing;
- A review and validation process for data suppliers to verify their data; and,
- Exemption or extension request process for special circumstances precluding a data supplier from reporting.

Recommendation: NAHDO recommends that Mississippi statutorily mandate a publicly-controlled health data reporting program, joining 39 other states that have legislatively mandated such reporting.

Legislation is more likely to result in:

- Comprehensive reporting across all providers
- Expansion to non-inpatient reporting by providers
- Public availability and public health access to health information
- Transparency in methods of data collection and reporting.

Recommendation: The legislation should establish stakeholder representation in the form of a data commission or committee with rulemaking authority.

Over one-half of state health data programs (53%) are governed by a statutory committee, and the remaining (47%) are governed by a board or data council. In general, a statutory committee has rulemaking authority and the make-up of the committee or commission is established in law (see Utah Health Data Committee text box).

Successful state health data programs have established trust with their data suppliers and all of the stakeholders in the community. This is done in several ways. In some states, a statutory committee or commission is authorized to make the data policies. In states such as Pennsylvania and Utah, purchasers, employers, payers, public health, consumer, provider, physician, and nursing representatives are appointed. The statutory committee has rule-making authority and makes all of the decisions in a public process. In some states, a data advisory body may be appointed to shape data policy, but this entity does not have rule-making authority. In private initiatives, often it is a data council or board that shapes the collection and release policies.

Utah modeled its legislation after the Pennsylvania Health Care Cost Containment Council (PHC4), which has been a leader in quality and outcomes reporting for decades. PHC4's 25-member Council consists of individuals who represent health care purchasers, providers, insurers, consumers and state government officials. (http://www.phc4.org/council/nr072505.htm).

Like Utah's Health Data Committee, the PHC4 representation is embedded in legislation, defining the make-up of its members:

The Utah Health Data Committee

The Utah Health Data Authority Act (26-33a) establishes the Utah Health Data Committee and defines its purpose "to direct a statewide effort to collect, analyze, and distribute health care data to facilitate the promotion and accessibility of quality and cost-effective health care and also to facilitate interaction among those with concern for health care issues." The committee "shall be composed of 13 members appointed by the governor with the consent of the Senate. No more than seven members of the committee may be members of the same political party. The membership of the committee shall include:

- one person employed by or otherwise associated with a hospital
- one physician, who spends the majority of his time in the practice of medicine
- one registered nurse
- three persons employed by or otherwise associated with a business that supplies health care insurance to its employees, at least one of whom represents an employer employing 50 or fewer employees
- one person employed by or associated with a third-party payor that is not a Health Maintenance Organizations and Limited Health Plans;
- two consumer representatives from organized consumer or employee associations
- one person broadly representative of the public interest
- one person employed by or associated with Health Maintenance Organizations or Limited Health Plans;
- two people representing public health.

- The Secretary of Health
- The Secretary of Public Welfare.
- The Insurance Commissioner.
- Six representatives of the business community, at least one of whom represents small business, who are purchasers of health care as defined in section 3, none of which is primarily involved in the provision of health care or health insurance, three of which shall be appointed by the President pro tempore of the Senate and three of which shall be appointed by the Speaker of the House of Representatives from a list of twelve qualified persons recommended by the Pennsylvania Chamber of Business and Industry.
- Three nominees shall be representatives of small business.
- Six representatives of organized labor, three of which shall be appointed by the President pro tempore of the Senate and three of which shall be appointed by the Speaker of the House of Representatives from a list of twelve qualified persons recommended by the Pennsylvania AFL-CIO.
- One representative of consumers who is not primarily involved in the provision of health care or health care insurance, appointed by the Governor from a list of three qualified persons recommended jointly by the Speaker of the House of Representatives and the President pro tempore of the Senate.
- Two representatives of hospitals, appointed by the Governor from a list of five qualified hospital representatives recommended by the Hospital and Health System Association of Pennsylvania one of whom shall be a representative of rural hospitals.
- Two representatives of physicians, appointed by the Governor from a list of five qualified physician representatives recommended jointly by the Pennsylvania Medical Society and the Pennsylvania Osteopathic Medical Society.
- An individual appointed by the Governor who has expertise in the application of continuous quality improvement methods in hospitals.

Pennsylvania Health Care Cost Containment Council (PHC4)

The Council is an independent state agency formed under Pennsylvania statute (Act 89,as amended by Act 14) in order to address rapidly growing health care costs. The Council's strategy to contain costs is to stimulate competition in the health care market by:

A) giving comparative information about the most efficient and effective health care providers to individual consumers and group purchasers of health services; and

B) giving information to health care providers that they can use to identify opportunities to contain costs and improve the quality of care they deliver.

Act 89, as amended by Act 14, specifically assigns the Council three primary responsibilities:

1) to collect, analyze and make available to the public data about the cost and quality of health care in Pennsylvania;

2) to study, upon request, the issue of access to care for those Pennsylvanians who are uninsured;

3) to review and make recommendations about proposed or existing mandated health insurance benefits upon request of the legislative or executive branches of the Commonwealth

The Council is funded through the Pennsylvania state budget. In addition, the Council receives revenue through the sale of its data to health care stakeholders around the state, the nation, and the world.

- One representative of nurses, appointed by the Governor from a list of three qualified representatives recommended by the Pennsylvania State Nurses Association.
- One representative of the Blue Cross and Blue Shield plans in Pennsylvania, appointed by the Governor from a list of three qualified persons recommended jointly by the Blue Cross and Blue Shield plans of Pennsylvania.
- One representative of commercial insurance carriers, appointed by the Governor from a list of three qualified persons recommended by the Insurance Federation of Pennsylvania, Inc.
- One representative of health maintenance organizations, appointed by the Governor from a list of three qualified persons recommended by the Managed Care Association of Pennsylvania

Other governance models include advisory and other types of policy boards. Another model of governance is the Virginia Health Information (VHI), a private, non-profit organization that collects and disseminates hospital and health plan data in the state of Virginia. VHI is a "delegated authority" to the Virginia Department of Health, which means that, under the state's mandate, the Department of Health contracts with a non-governmental entity, such as VHI, to oversee the implementation of data system development. VHI has an independent board and relies on a combination of public and private funds, especially data sales, for revenues[28].

Recommendation: Whichever governance model is selected for the Mississippi health data program, NAHDO recommends that all stakeholders be represented.

Determination of data steward (data ownership and control)

There are three main approaches state programs have adopted for data stewardship and control and these models reflect the legislative and political environment in a state. Three organizational models for state health data programs have emerged.

1) a public agency collects the data under a legislative mandate (28 states);

2) a private agency collects the data as the delegated authority to the state (11 states)

3) a private agency collects the data voluntarily (9 states).

A legislatively mandated reporting initiative generally determines the type of entity that will collect, manage, and release the data. Voluntary initiatives are always privately managed, either by a hospital association or a non-profit entity. In states with legislative mandates, the data steward can be either a state agency or a "delegated authority" for the state (for example, a hospital association or a non-profit entity).

States that have health data programs without a legislative mandate have established statewide reporting voluntarily. These states have recognized the benefits of a common health care database and have worked with hospitals to secure voluntary compliance. Often, the voluntary health data program is managed by the hospital association, which maintains the information for

[28] Virginia Health Information. About us. At http://www.vhi.org/about_home.asp . Accessed Nov. 2007

the benefit of its members, and in many cases, also shares data with the community. In Hawaii, the Hawaii Health Information Corporation (HHIC) is a membership organization that collects the data voluntarily from hospitals. HHIC is governed by a Board of Directors and provides a source of data for public health, market studies, quality improvement, and community assessment.

This data steward decision is an important one. The data steward is responsible for maintaining the database and controls who has access to the database. Successful management of these responsibilities is difficult, yet key to success for any data program. Ideally, a data steward will be seen as a neutral or independent entity that is credible and trusted by all of the stakeholders. However, local environments vary and factors may preclude achieving the "ideal", so health data programs must choose the approach that is most practical.

Recommendation: Mississippi should consider the health data program data steward structure that is most likely to meet the objectives of public availability, sustainability of funding, equitable access, and independence and neutrality. An analysis of options should be undertaken in order to leverage an existing infrastructure and IT capacity as well as to assure the authority or ability to eventually expand to outpatient (non-hospital) settings in the future.

Where to locate a public program depends on the local environment and the opportunities and challenges the local environment presents in terms of leveraging existing Information Technology (IT) and staffing resources. Some states have placed the health data program within the Department of Health, which facilitates data integration and linkage with other public health data sets and leverages the analytic and IT infrastructure. In some states, a separate state agency is established, with its own policy board or commission. There are advantages and disadvantages to either approach. An independent agency may be more flexible and responsive to stakeholder needs than a larger agency with layers of bureaucracy. However, the costs may be higher if a separate infrastructure must be established.

Funding Mechanism

The key to a stable statewide health data program is a stable source of ongoing funding. Unlike other state data systems, such as vital statistics, state health data programs receive no federal funding. In states that have mandated hospital reporting, the states and providers bear the costs, through general fund appropriations or fee assessments on providers (and/or other facilities and insurers). Voluntary health data reporting programs are usually funded through membership fees and/or grants and contracts. All programs, mandatory and voluntary, rely on revenue from data product sales and data programs must balance the need for data sales with affordable data access.

> *In the words of one state health data program director, "Data systems are not like light switches. You can't flip them on one year and then flip them off the next and expect to see progress in data collection and use. States must have ongoing support to keep the pipeline open".*
>
> *NORC-NAHDO Report, 2005.*

Because a publicly-governed health data program should benefit the public and all stakeholders, NAHDO recommends public funding to establish and sustain statewide health data reporting in Mississippi. This core funding, when combined with future data product revenues and future contracts and grants, will provide a diversified revenue stream to support core activities.

Core funding will create the financial stability that is essential to building and maintaining discharge data systems. State health data programs take years to establish. These programs must rely on relationships and trust, and stable data flows. Most data programs, both public and private, have diversified their revenues, capitalizing on the growing demand and value of the statewide data. Revenues for health data programs come from the following sources:

- General appropriations;
- Fee assessments on facilities and/or health plans;
- Funding from membership dues and fees; and/or
- Data sales.

States with public funding, in the form of fee assessments or general appropriations tend to have the largest and most mature data reporting systems. These publicly-funded programs support access and use to all stakeholders. Private, voluntary reporting initiatives generally rely on membership dues for core funding. The public funding could be in the form of a fee assessment on all providers and health plans or through general appropriations, or both. In fact, ideally, the start-up of a health data program would be funded by a general appropriation that could be adjusted downward as fee assessments on the industry are implemented.

All health data programs, public and private, eventually supplement their core revenues with data sales revenues, but these revenues don't accrue for one or two years, after the data system is established. Data sales alone are not generally sufficient to support the core infrastructure of a health data program, but do provide important supplemental revenue to maintain and update the system.

Because start-up is so critical, legislative appropriations are recommended to assure that staffing and technical infrastructure are supported to complete planning, policy development, and data collection activities. Based on the experience in other states, there are advantages and disadvantages to the various funding approaches (Table 5).

Table 5. Advantages and Disadvantages of Funding Approaches

Funding Mechanism	Advantages	Disadvantages
General Appropriations	Legislative accountability Assures program continuity through mandate Assures hospitals will submit data/support of the state behind the program Providers cannot use fee assessment to pressure legislators to kill program	Competes with other budget priorities Rarely sufficient to solely fund and maintain a data system Often have travel and training moratoriums--difficult to maintain skills of staff without travel to workshops, training and conferences
Fee assessment on health systems, providers	Shared expense and accountability across the industry More resources than generally found in state budgets--allows up-to-date software, hardware Providers tend to be more engaged with data and products	Subject to political pressure Demands for special private tools to analyze data may be greater

Funding Mechanism	Advantages	Disadvantages
Private funding (membership dues)	Not subject to political influence The data program must remain relevant to the funders to survive	Proprietary interest may override the public interest Public access may be limited or restricted Sustainability is dependent on member funding/value

Recommendation: Mississippi should appropriate legislative funds to fully support the start-up and maintenance of a statewide health data reporting program. Funding should be in the form of general appropriations sufficient to fund the core infrastructure and staffing to support the implementation and ongoing operations of a health data program. A fee assessment, on all providers (including hospitals, nursing homes, free-standing ambulatory facilities) or on health plans, can be a source of core funding, if the political environment is favorable.

The next section discusses the cost considerations for health data program development and maintenance.

Planning-The Advance Work

Analysis of Expected Costs to Implement and Operate a Health Data Program

Cost Considerations

Many decisions impact the costs associated with the development and maintenance of a statewide data collection system. In this section we provide information on a base system of data collection and limited reports. For purposes of clarity, we assume in the costs shown below that the healthcare data program will be an independent unit, which may be housed in a larger public health agency. We make this assumption because this is the most common form for health data organizations. We recommend that program decisions are independent of other public health agency programs—there are often competing priorities in other units, either for financing or program goals. For example, if placed within a public health surveillance unit, that unit may seek more clinical information than is feasible for the system as a whole. Or, an infectious disease unit may attempt to drive the agenda towards closed reports for providers. Thus, independence is needed in order to meet the mission of the health data program.

Within this section, we do provide some information on what costs (shown with an asterisk*in this section) could be saved by housing the data collection system within a larger public health agency. While savings may occur; it is also possible that other units would assess fees. It is likely that the unique requirements of this data collection would not be met by existing infrastructure. These infrastructure requirements will be discussed primarily in the sections on "IT Infrastructure and Implementation." Other assumptions for the inpatient data collection include:

- Data collection is a full census of all acute care discharges;
- Submission of data to the system would be done via an electronic system, whether web-based or other electronic process;
- Hospitals are required to correct data with errors and re-submit;
- A series of data files will be produced for different users; and,
- A limited number of reports would be produced within funding constraints.

Other considerations that were taken into account include: the number of discharges (407,000); and the number of acute care facilities (94) in Mississippi[29] that would be submitting information and their associated discharges. The major driver for ongoing costs is the volume of data to be processed (WA Ambulatory Surgery Report, p. 32)[30]. Therefore, when ED data and ambulatory surgery data are added, volume increases and therefore costs will increase over what is shown below.

[29] Health Forum LLC, American Hospital Association, 2005 AHA Annual Survey 2006; special data request, March 2007. Available at http://www.ahaonlinestore.com/ProductDisplay.asp?ProductID=637]

[30] Washington State Department of Health, Final Report to the Legislature: Ambulatory/Outpatient Data System Feasibility Study, July 1998.

Program Costs

Critical to the success of any new data system is adequate planning for staffing and IT infrastructure, along with a plan for ongoing maintenance activities for the system. These will be discussed in greater detail below. In this section, we will focus on a summary of what it will take to bring a system up—allowing for the capture of standardized inpatient data from Mississippi hospitals. Planning and system development (following the legislation) is likely to take six months to a year before data actually roll into the system. If all software and tools is vendor supplied and installed, less upfront time may be needed (although that would depend on the vendor selection, bidding and contracting processes in the state), but costs may be higher than shown below. System development costs are summarized in the following table. Areas that are shaded indicate potential variation in cost due to selections made by the program. More details follow Table 6.

Table 6. Cost Considerations for a State Health Data Program (Year 1 is in 2007 dollars)

						YR1	YR2	YR3
Staffing	#						*est 2.5% increase.*	*est 2.5% increase*
Admin/Policy Analyst	1	100.00%	$93,800	$70,000	$19,600	$89,600	$91,840	$94,080
IT Director	1	60.00%	$76,800	$46,080	$12,902	$58,982	$60,457	$61,932
Sr Analyst Researcher	1	100.00%	$60,000	$60,000	$16,800		$78,720	$80,640
Analyst	1	80.00%	$44,800	$35,000	$9,800		$45,920	$47,040
DBA 6 months then IS support	1	40.00%	$76,800	$30,720	$8,602	$39,322	$40,305	$41,288
Admin Asst - budget payroll	1	10.00%	$34,000	$3,400	$1,156	$4,556	$4,670	$4,784
sub-total	6			$245,200	$68,860	$192,460	$321,912	$329,763
Hardware / software	**#**		**Per setup**					
Hard/Software seats	4		$3,000			$12,000		
LAN server & nodes						$5,000	$1,500	$1,500
Oracle / SQL server License - or SAS licenses						$20,000	$20,000	$20,000
Secure Server						$10,000	$10,000	$10,000
Other analytic software				$5,000		$5,000		
Severity Adjustment modeling software						$10,000	$11,000	$11,000
sub-total						$62,000	$42,500	$47,500

Supplies	#	Each seat					
Office	4	$750			$3,000	$3,075	$3,152
Printing / Pub		$4,000			$4,000	$4,100	$4,203
Travel / Board Meetings					$2,500	$2,563	$2,627
Phone + Long distance					$1,500	$1,500	$1,500
sub-total					$11,000	$11,238	$11,481
Rent					$11,946	$12,245	$12,551
sub total					$11,946	$12,245	$12,551
TOTAL					$277,406	$387,894	$401,295
Savings if integrated (IT + 1/4 Analyst)					$39,322	$51,785	$53,048
Additional Vendor Services					$86,322	$71,805	$72,788
	annual $ /sq. ft	Sq ft per office	# of offices				
Rent calculations	$22.97	130	5	650			

Staffing

It is essential to have experienced and skilled staff in electronic healthcare data collection—the tasks described below for each of the employees reflect the need for skilled employees. When a healthcare data processing system has manual editing and other activities, lesser skilled individuals can be used, but more of them are required than is shown. The salaries are estimates, some states' pay structures may be higher other states' pay structures may be lower. An electronic version of the budget spreadsheet is provided, this allows the user to auto update with specific Mississippi information related to the salary and benefits, as well as other figures shown.

Staffing assumptions were based on estimated number of hours needed for specific components of the position. As shown in the table, for two of the positions there are possibilities for substitutions of employees from other program areas. However, it should be noted that additions to the data collection, such as Emergency Department or Ambulatory Surgery data collection , would likely utilize the time that was freed by using employees from other program areas.

It is anticipated that the planning phase will require hiring of the Program Director and IT Director. As planning concludes and implementation begins, the remaining two staff persons should be hired.

Year 1 Program Director (Administrator/Senior Level Health Policy Analyst) Tasks for this position will include hiring and supervising staff; working with the appointed Board to determine policies for data submission and data dissemination; planning and implementing the physical structure/environment for the unit; establishing connections with key stakeholder groups; identifying key resources within the state; developing a dissemination strategy based on policies decided by the Board; working with the press and other public entities requiring information; promulgating administrative rules if required. The Program Director will oversee the development of the IT infrastructure and will, with the IT Director develop a hospital submission manual. It is likely that any training that hospitals require will also be done by the Program Director and IT Director.

Year 1 IT Director (Programmer/ Database Administration/Network Administration*/Data Security*/Website*). This IT Director position will be needed whether or not some services are provided by vendors. See bottom of table for estimated vendor costs for assisting the IT Director in designing and implementing the hardware/software system for data submission, editing, and file preparation. This position requires knowledge and experience with a range of IT services, including web systems, networked environment, secure communications, structured query language, and other programming languages. The IT Director will also be installing and maintaining analytic software, and building web-based information sites. The IT Director will also install all workstations with hardware and software, and support services to staff. Some IT services, if part of an integrated system, can be provided by others. For example, network administration, security standards, etc., could be handled by other agency personnel in IT. However, it is important to note that one person must be responsible for production system “uptime”; electronic submissions by hospitals will occur 24 hours per day and 7 days a week. Many hospitals process data across the 24 hours, and the IT infrastructure should be available at all times, except for planned ongoing maintenance (if there are planned outages, hospitals must be notified in advance). [These systems are in a production environment and differ from common email and web services.]

Year 2 Analyst/Writer. In Year 2, new staff for processing incoming data and developing reports will need to be added to the two positions described above. They will conduct analyses for both in-house monitoring of systems, and for public reports, and custom data requests. The Analyst/Writer will need a good understanding of the health policy environment, public health issues, and some programming skills.

*Year 2 Analyst Support** In year 2, additional support will be needed to address follow-up issues with hospitals and data customers. In addition, this individual will assist with editing and preparation of tables and other documents for public reports. Ongoing communication with hospitals regarding data issues is required. It would be helpful if this person was familiar with claims coding in a hospital environment.

IT Infrastructure (Hardware/Software licenses/Internet Access/ programming)

Hardware The following types of hardware will be needed: WEB server, Storage and Processing Server, Backup Device, 3-4 desktops and 1 laptop with network or wireless cards.

Software Licensing: The following activities will require software licenses: Internet hosting, database storage, editing software (in-house and hospital edits); secure communications, network administration, severity adjustment tools and analytic software. Costs for software are highly variable depending upon types of operating system and software selected—e.g., some states may use Linux on a UNIX box for their database—costs are low for software, but there are less available supporting software programs than when using other types of data management software, and it may be more difficult to hire a Unix DBA and programmer. ORACLE and SAS licenses cost more, but it is easier to find programmers. SAS is used by many states for data management and analysis. SAS has many components, but likely the base component and desktop statistical analysis tools would be needed (licensing varies but could be between $15,000 and $30,000)*. Adding more SAS components such as business intelligence or SAS web products could bring licensing costs up significantly. One of the most expensive upfront costs of the system will be the automated data editing—some states have designed their own, others use vendor systems. Internal development of an in-house and hospital editing tool can take substantial time and dollars—we recommend either adopting another state's system or purchasing editing tools through a vendor. The actual edits can be programmed to run from other state's software—but you need: both a sophisticated programmed system for accepting files, editing files and returning files to hospitals, a hospital editing tool and file return system for edited files to be added to the database. The entire process must be tracked and secure. Auditing software is also needed and audits must be reviewed to assure security and safety of the data. A secure and reliable back-up system must also be designed, implemented and maintained.

Training Costs (both for staff and hospitals)—Any migration away from use of standard claim data elements will require some training for hospitals. Depending on the file structure that is selected (hierarchical or flat file) the amount of training needed may vary. Complex hierarchical file structures will require more training than a flat file. Training will require services of both the Program Director and IT Director.

Facility costs shown in Table 6 include a server room with controlled temperature and a security system or a steel cage system in a larger air conditioned space). Costs vary depending upon location. Rent is estimated at $22.97 per square foot for a 'cabled environment withT1 lines' for high speed internet connections.

Ongoing Costs

In the table above, three years of cost, beginning with 2007 costs, are displayed. Clearly, much could change in terms of the electronic environment and also demands for information. However, most costs will remain and there will be some increases, given ongoing maintenance and updates for the system (software and hardware). Revenues from data sales have not been included as an offset, given that those revenues will depend on marketing and the cost of the products. The value of the products will be determined by timeliness, data accuracy, and ease of access. If the process for acquiring data is too difficult or time-consuming, or the data processing is too slow, there will be fewer users of the data.

Ongoing costs should include training costs for staff to maintain and update skills in programming, data processing, IT environment, health services research, etc.

New costs will also accrue with the addition of ED data and ambulatory surgery data, but will be significantly reduced by the presence of a functioning inpatient system.

Hospital Burden

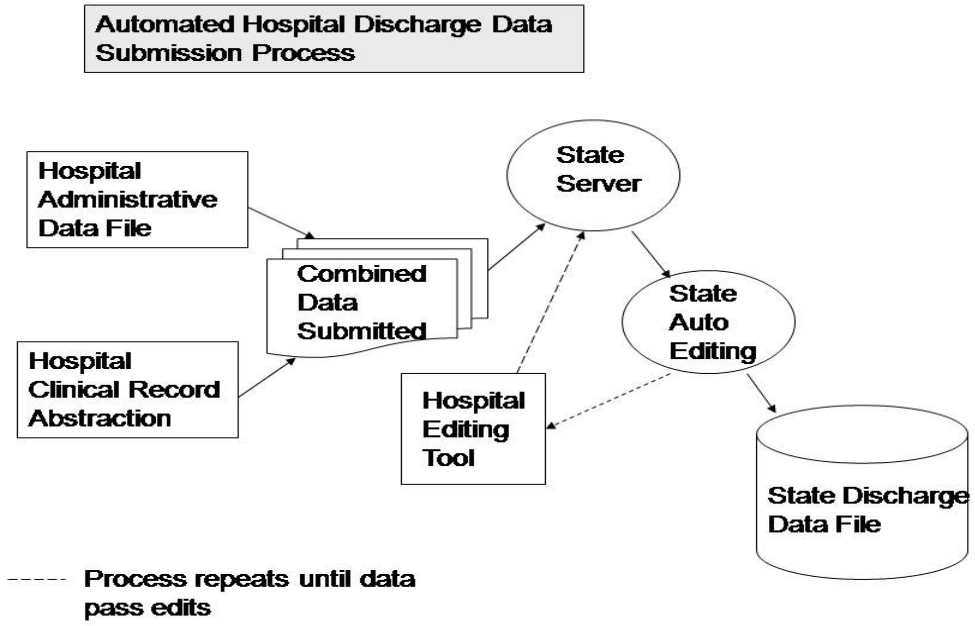
In terms of hospital burden, you can expect hospitals will have some upfront costs associated with submitting data via a secure server to the data agency. Much of the upfront costs will be incurred for programming and are somewhat dependent upon the complexity of the hospitals IT infrastructure. If there are significant changes to the data elements (that is, they vary from the standard claim), in-house programming costs for designing the system and hospital abstraction and file loading costs will increase. Abstraction of clinical data elements may require either a medical records professional or RN—depending on the level of clinical detail requested. The least costly method is to require submission of standard claim data elements; however as discussed above, it will be important to add several additional clinical and demographic data elements to meet the needs of public health and policy stakeholders. Changes in the data elements submitted over time will also require re-programming, therefore, we strongly recommend that the initial requirements can stand for at least 2-3 years before making changes to the types of data elements within the inpatient system.

There have been several efforts to document upfront and ongoing hospital burden related to data submission; many of those are based on some abstraction of data elements, not just submission of the standard claim form. For example, according to the report by Booz, Allen, and Hamilton for the Hospital Quality Alliance[31], quarterly submissions for CMS hospital quality measures related to Acute Myocardial Infarction ranged from $10 per record abstraction to a high of $132 per record. The hospitals' data source in the sample ranged from 40% administrative claims data to 100% claims data. There are no comparative figures for cost of submission related to standardized claims data.

According to Booz, Allen and Hamilton, primary cost drivers for hospitals include: ongoing hospital or contract staff time for monitoring submission of the data on a quarterly basis, as well as editing of the data as needed; and the application vendor fees. When new data elements are added hospitals will likely need vendor modification to their IT infrastructure. Figure 5 below is an example of the data submission process for hospitals when both standardized data and some clinical data elements are submitted by hospitals. The specific data elements are discussed in the implementation section of this report, immediately following the schematic.

[31] Booz-Allen-Hamilton, Hospital Quality Reporting in the United States: A Cost Analysis for the Hospital Quality Alliance. December 2006. Available at http://www.hospitalqualityalliance.org/hospitalqualityalliance/files/BAH1206.pdf. Accessed Nov. 2007

Figure 5: Data Submission Process



Recommendations: Costs are dependent upon a variety of decisions: governance, location of the organization (free-standing or within a larger organization), funding availability, and whether the system is predominantly built in house or whether it is predominantly done with vendors. In either case, some vendor services will be necessary. Other cost determinants include the structure of the data file, whether hierarchical or flat file; the submission mode (fully electronic including editing o r electronic submission with manual processing). Staffing costs are also variable given state position salary and benefit structure.

Lastly, the costs for a data collection system, as shown in Table 6, are linked to Figure 5 Infrastructure design and the earlier assumptions made in this section about:

- State governance and mandatory submission;
- Combination of state and vendor IT system development;
- Hierarchical file for database, flat files for analysis and distribution;
- Submission mode is fully electronic and includes hospital electronic editing tool;
- Salary structure dependent upon state guidelines;
- Inclusion of specific dollars for training of staff and hospitals in budget; and
- An assessment of needs of stakeholders for a limited number of reports.

Implementation Requirements

Data Elements

Below, Table 7 lists proposed data elements to be used for an emergency department data collection in Wisconsin—the table also contains brief definitions and entities that have proposed use of these elements. The data elements that were proposed for collection by the health data organization (Bureau of Health Information) are indicated in the Source column with the initials BHI. Note that the majority of the BHI data elements listed also represent the needs for a typical inpatient discharge data system. These BHI ED data elements, like the inpatient standard data elements, are derived from national standards. The table below demonstrates that, once inpatient data reporting is established in Mississippi, adding ED reporting should not be too burdensome, as the data elements are similar and are aligned with billing standards hospitals already utilize. Additionally, the table reflects data elements that could be linked to the ED collection from other available sources of information.

Table 7. Proposed Data Elements for an Emergency Department Data Collection System

Element	Definition	Source(s)
Facility ID	Facility where patient seeks or receives outpatient emergency care	BHI, STAC, DEEDS, Panel
*Facility Type	Code to identify hospital EDs, free standing urgent care clinics, urgent care clinics connected to hospitals, etc.	Panel
Date of service / ED arrival date	Date (month, quarter) of ED service	BHI, STAC, DEEDS, Panel
Encrypted case ID / unique identifier	Identifier used by the facility to identify the patient at admission (medical record number, etc.)	BHI, STAC, WEMSIS, DEEDS
*ED Admission type	Code indicating the priority of admission: emergency, urgent, elective, newborn Alternative from Ky. ED Triage routine, urgent, serious, critical	BHI, KY
ED Admission source	Code indicating the source of admission: physician referral, clinic referral, HMO referral, transfer from hospital, transfer from skilled nursing facility, transfer from another health care facility, emergency room, court/law enforcement, unknown	BHI
*Time in / arrival time	Time documented in patient's record for the ED visit	STAC, DEEDS, Panel
Diagnosis codes	Principal and up to 8 other diagnosis codes describing the condition established, after study, to be chiefly responsible for causing a patient's admission	BHI, STAC
E-code	ICD-9-CM code describing the external cause of an injury, poisoning, or adverse effect	BHI, STAC, WEMSIS, Panel
Procedure codes	Principal and up to 5 other codes describing a procedure performed for definitive treatment or that was necessary to treat a complication rather than for diagnostic, exploratory, or therapeutic purposes	BHI, STAC
*Chief complaint	Patient's reason for seeking care or attention, expressed in words as close as possible to those used by the patient or responsible informant, entered as code with associated text description or as text description alone	DEEDS, Panel
*Patient severity	Classification of patient's severity: requires immediate evaluation or treatment, requires prompt evaluation or treatment, time to evaluation or treatment not critical, or unknown. Alternative classification: critical, emergency, urgent, observation. (see ED Triage above)	DEEDS, Panel
Discharge diagnosis	Encoded description of ED disposition diagnosis	DEEDS, Panel

Element	Definition	Source(s)
ED Discharge status	Code for the arrangement or event ending a patient's ED visit: discharged to home or self care, discharged to another short-term general hospital, discharged/transferred to a skilled nursing facility, discharged/transferred to an intermediate care facility, discharged/transferred to another type of institution, discharged/transferred to home under care of organized home health service organization, left against medical advice, transferred or discharged to a home intravenous provider, expired, discharged to hospice-home, discharged to hospice-medical facility	BHI
*Time out / departure time	Time when patient leaves ED	STAC, DEEDS, Panel
*Mode of transport to ED	Patient's mode of transport to ED: ground ambulance, helicopter, police, walk-in (following private transport, public transport, law enforcement transport, not specified), other, unknown	STAC, WEMSIS, DEEDS, Panel
*EMS response unit ID	Identifier for EMS unit that transported patient to ED	STAC, WEMSIS, DEEDS
*EMS agency ID	Identifier for EMS agency that transported patient to ED	DEEDS
Referral source to ED	Individual or group that determined patient should seek care in ED: self-referral, EMS transport, practitioner or health care facility referral, internal facility referral or transfer, law enforcement, acute care hospital transfer, other health care facility transfer, other, unknown	DEEDS, Panel
*Incident site type (place of injury, location emergency occurred)	Type of place where patient's injury occurred, entered as home, residential institution, school or other institution and public administrative area, sports and athletic area, street highway, trade and service area, industrial and construction area, farm, other, or unspecified	STAC, WEMSIS, DEEDS, Panel
ED disposition	Patient's anticipated location or status following ED visit, entered as discharged to home or self-care; transferred/discharged (to another short-term general hospital, skilled nursing facility, intermediate care facility, another type of institution, home under care of home intravenous drug therapy provider, or home under care of certified home provider/program); left (without receiving medical advice against leaving or with receiving medical advice against leaving); placed in designated observation unit; admitted (to hospital floor bed, intermediate care/telemetry unit, ICU, or OR); died; other; or unknown	DEEDS, STAC, Panel
*Receiving facility ID	Identifier for facility to which patient is transferred or discharged at conclusion of ED visit	STAC, DEEDS, Panel
*Mode of transport for transfers following ED visit		STAC
*Transport personnel Inpatient admission codes	EMT, RN, MD, etc.	STAC Panel
Attending/Admitting provider ID	Identifier for ED provider responsible for the patient's care during the ED visit	BHI, STAC, DEEDS, Panel
Attending provider type/specialty		DEEDS, Panel
Consulting provider ID	Identifier for consultant provider who participates in patient's care during the ED visit	DEEDS
Consulting provider type/specialty		DEEDS, Panel
Performing provider ID	Identifier of provider who performs ED procedure; designated as Other Physician ID at BHI	BHI, DEEDS
Performing provider type/specialty		DEEDS, Panel

Element	Definition	Source(s)
*Inpatient provider ID	Identifier of provider whose inpatient service patient is admitted to	DEEDS
*Inpatient provider type/specialty		DEEDS, Panel
Patient zip code	Zip code of patient residence	BHI
Patient county of residence	County of patient residence	BHI, STAC
Patient sex	Sex of patient	BHI, STAC, DEEDS
Patient date of birth / age	Patient's date of birth (integer age)	BHI, STAC, WEMSIS, DEEDS
*Patient race	Race of patient	BHI, STAC, WEMSIS, DEEDS, Panel
*Patient ethnicity	Ethnicity of patient	BHI, DEEDS, Panel
Primary and Secondary payer ID	Medicare, Medicaid, WPS, CHAMPUS or CHAMPVA, Non-Medicaid Blue Cross and Blue Shield, Other	BHI, STAC, DEEDS, Panel
Primary and Secondary payer type	FFS, Alternative Health Care Insurance Plans (HMO, PPO, PPA), workers' compensation, general relief, self pay, etc.	BHI, DEEDS, Panel
Total charges		BHI, STAC, DEEDS, Panel

Note: DEEDS (Data Elements for Emergency Department System) is proposed by CDC and others, STAC (State Trauma Advisory Committee) WEMSIS (Wisconsin Emergency Medical Services Information System) BHI (Bureau of Health Information) Panel (Technical Advisory Panel convened November 30 2000 by BHI) KY(Kentucky Project on Emergency Data Collection; offered by a panel member)

**Indicates those elements that do not appear to be routinely collected by Emergency Departments in the state at current time, although potentially available for collection. These might be suggested for a second phase of implementation.*

For an example of data submittal requirements, go to the New York SPARCS Submission Manual at http://www.health.state.ny.us/statistics/sparcs/index.htm

Data Management

Edits

The goal of all state reporting systems is to be accurate and timely. The enemy is coding variations. The national information technology agenda it based on the principle of interoperability. The American Health Information Community (AHIC) through naming "break through" use cases has established the priority areas for interoperable solutions. The Health Information Technology Standards Panel (HITSP) is charged with naming the relevant standards for each of these use cases along with the task of harmonizing any gaps amongst the standards.

System edits are the data steward's weapon against coding variations. Existing state reporting systems use a variety of editing strategies. Some utilize extensive data element by data element validation. Some utilize a post auditing strategy to assess trends across data submissions. Some states utilize both techniques. The level of sophistication of editing strategies is directly proportional to the complexity of the data system. The amount of editing will also affect the timeliness of data availability.

Systems that use extensive editing techniques to improve the accuracy of the data will also need enough resources to monitor and follow up on data that does not meet the editing criteria. Data left uncorrected or re-submitted late can adversely affect the timeliness of the data.

The use of data standards is of particular value for decisions about editing the data. States that have a long history of collecting state discharge data have learned through experience the right level to apply edit criteria to each collected data element as well as the right trending to expect. Edits on standard data elements in states with established systems are easily shared with newly developing systems using those same data element standards. This would not be the case when a new system is developed using state specific idiosyncratic data elements.

Derived Data Elements

Most existing state reporting systems produce a file or relational data base that contains additional data elements that were derived from submitted data that is necessary to enhance analysis capabilities. An example of this would be the derivation of DRG's for severity adjustment based on the information submitted by the provider. In the case of derived DRG's being added to an analysis friendly output file, this would require a state system to purchase the necessary grouping software as well as adding the addition of these derived fields to the programming requirements specification document. Other derived fields would be specific to individual state needs. The programming and cost implications would be dependent on the nature of the derived fields being added.

File Maintenance

For a variety of reasons all state reporting systems need a process to correct or delete submitted data. How to make such corrections or deletions is an important system design feature. It is important to note that this design feature implies a bi-directional communication strategy with the provider community.

Severity Adjustment

Data adjustment, such as case mix or severity adjustment is a method used to account for differences in patient characteristics (e.g., age, income, and type and severity of illness needing treatment) likely to affect the outcomes of care (e.g., death, physical functioning, resource utilization and/or cost) independent of the actual medical treatment given. Data adjustments are only needed when there is interest in ***making comparisons***. When health care data are used to understand health care processes or other characteristics of the health care system, there is no need to "adjust" the data. Adjustment is used to ***increase the validity of comparisons*** between providers, or groups of providers, by accounting for inherent differences in the patients served by each provider or group. Adjustment is used only to control for factors that are ***outside of the control of the providers*** to be compared.

There are two main reasons for adjusting health care data:

- To *predict* the amount and cost of care that an individual (or population) will use, i.e., the propensity of an individual (or group) to consume health care resources in a specified time period) or
- To *compare* outcomes after medical intervention.

The term "risk adjustment" is often used to refer to the type of adjustment needed to predict amount and cost of care. The most common application of risk adjustment is in the context of identifying appropriate capitation payments for health plans or providers. Plans' costs depend on at least four factors: patient characteristics, efficiency of providers, intensity of services or

treatments provided, and unpredictable events (e.g., chance or random factors, such as accidents). If plans (or providers) receive the same unadjusted premium (payment) for each subscriber, plans (providers) with healthier members reap an unearned windfall and plans (providers) with sicker populations face unfair losses.

The term "severity adjustment" is often used to refer to the type of adjustment needed to compare outcomes after medical intervention although actually there are various degrees of adjustment can be applied for this purpose, e.g., adjusting for the age-sex of patients, adjusting for differences in case-mix , and/or adjusting for differences in severity. The need for severity adjustment in comparing outcomes was first highlighted in the mid 1980s when the Health Care Financing Administration first publicly released hospital-level mortality figures. This public release followed the implementation of DRGs for paying hospitals for the care of Medicare patients and the resulting concerns that hospitals were discharging patients "quicker and sicker." It turned out that the facility with the most aberrant death rate in this initial release (an observed rate of 87.6% compared with a predicted rate of 22.5%) was a hospice caring for terminally ill patients. The model had failed to account for differences in patients' risk of death. This led to the development of numerous off-the-shelf severity measures designed specifically for comparing hospital death rates.

Key point: Adjustments in rates by age, gender, and severity of illness is only needed when there is interest in ***making comparisons***. When health care data are used to understand health care processes or other characteristics of the health care system, there is no need to "adjust" the data. Adjustment is used to ***increase the validity of comparisons*** between providers by accounting for inherent differences in the patients served by each provider.

The first step in identifying an appropriate adjustment strategy is to identify the purpose of adjustment in rates and other statistical measures:.

- Utilization of services
- Practice variation
- Community health
- Ambulatory sensitive conditions
- Chronic care conditions
- Service volume within certain disease populations
- Procedures performed for certain diseases
- Resources needed to treat certain diseases
- Determining the best way to treat certain diseases
- Improving the healthcare of the population
- Intervention patterns.

There are a number of open source and proprietary risk adjustment versions; on average the proprietary versions are more sophisticated, but because of their proprietary nature they are a "black box." Increasingly, healthcare providers are calling for the use of open-source products, so they can determine the validity of the underlying methodology.

Appropriate adjustment of data will serve to avoid arguments such as "my patients are sicker so they cost more to treat." Adjusting charges to account for patient differences will mean that any remaining differences in charges are the result of other factors. The intent is to develop a method for adjusting for factors that are outside of a provider's control.

Strategies to Improve Comparability

The health data program should involve all of its stakeholders to define and document the methodologies that will be used to prepare the data and reports. The methods should be transparent and publicly available. State health data programs apply a series of tools and methods to improve the comparability of the data across hospitals. Using various grouping methodologies, peer group and stratification, adjusting for case mix and outliers are all used to prepare the data for public use files and reports.

The following section summarizes how Utah addressed issues of data comparability and public reporting. Peer grouping and adjustment were important issues of concern. Utah formed several technical advisory groups to help create the methods and peer group strategies below. Because there are so many rural hospitals in Utah, it was important to differentiate between the rural and urban hospitals in reports and adjust for the actual utilization hospitals experienced.

Utah Health Data Committee's Standard Report 1, 1993

Diagnosis Related Group (DRG)

The DRGs were developed for CMS (formerly the Health Care Financing Administration) as a patient classification scheme which provides a means of relating the type of patients a hospital treats (i.e., its case mix) to the costs incurred by the hospital. While all patients are unique, groups of patients have common demographic, diagnostic and therapeutic attributes that determine their resource needs. All patient classification schemes capitalize on these commonalities and utilize the same principle of grouping patients by common characteristics.

The use of DRGs as the basic unit of payment for Medicare patients represents a recognition of the fundamental role a hospital's "sicker" patients play in determining resource usage and costs, at least on average. "The DRGs, as they are now defined, form a manageable, clinically coherent set of patient classes that relate a hospital's case mix to the resource demands and associated costs experienced by the hospital." (Diagnosis Related Groups, Seventh Rev., Definitions Manual, page 15.)

Each discharge in the Utah Health Discharge Database (UHDDB) was assigned into a DRG based on the principal diagnosis, secondary diagnoses, surgical procedures, age, sex, and discharge status of the patient. This report includes 62 selected DRGs which covered about 60% of all discharges that occurred in 1996, when this report was first published.

All-patient Refined (APR) DRG

The APR-DRGs are a patient classification scheme developed by 3M Health Information Systems that follows the basic DRG methodology of classifying patients into disease categories, but further subdivides each disease category into severity of illness classifications. With a few exceptions, a patient in each disease category (called consolidated DRG) is assigned into one of four levels of severity: no/minor complication or co-morbidity (CC), moderate CC, major CC, and extreme CC. Some of the exceptions to the four-level classification are newborns and neonates which are assigned to APRDRGs formed with the severity of condition already built-in (e.g., APRDRG 590: Neonate,

birthweight <750g with major procedure). A significant revision of the APR-DRG classifications occurred April 1, 1998. This 2000 report attempts to match as closely as possible the APR-DRGs used in earlier reports. APRDRG categories were used to define charge and length of stay outliers and calculate the Case Mix Index. This report includes 25 selected APR-DRGs which covered about 60% of all discharges that occurred in 1996, when this report was first published.

Outlier Cases

Some patients have exceptionally low or high lengths of stay (LOS) or total facility charges. A hospital's charges can be affected by just a few unusually long (or short) or expensive (or inexpensive) cases. These high or low values could be a result of coding or data submittal errors, particularly in length of stay, total charges, or data elements that affect DRG assignments. Other reasons for exceptionally low LOS or charges could be due to death or transfer to another facility. Exceptionally high LOS or charges could be due to a catastrophic condition. Whatever the reason, these values, referred to as "outliers," distort the averages and were excluded from calculations. LOS or facility charge high outliers are defined in this and succeeding reports as values above 2.5 standard deviations from the mean. Means and standard deviations are APR-DRG specific and calculated on a statewide basis. The low outliers were defined as a non-newborn or non- normal delivery discharge with less than a $300 charge. However, the calculations in this report do not exclude low outliers. A preliminary analysis showed that of the discharges that met this definition, a high proportion are in the DRG, "Other factors influencing health status," for which it was difficult to determine whether they were true outliers.

Case-Mix Index

An important source of variation among hospitals in summary measures of outcome - such as length of stay, total charges, and severity of illness - is the differences in the complexity of the patients they treat. To allow for a meaningful comparison of outcome measures among hospitals, an adjustment factor based on patient complexity should be applied. For this reason, four case-mix indices (all-patient, acute, obstetric, and pediatric) have been calculated for each hospital and are shown on the tables in this report. For example, a hospital's case-mix index of 1.15 means that the overall case mix of a hospital requires 15 percent greater intensity of resource use relative to the state as a whole.

Hospital Peer Groups

Comparing summary outcome measures (length of stay, total charges, readmission rates, mortality rate) among hospitals has always been a controversial issue because of the difficulty of defining what makes hospitals "comparable." As discussed previously, summary outcome measures vary among hospitals depending on various factors such as location, bed size, ownership, affiliation, and teaching status. If all these factors were to be considered in defining peer groups, each hospital might end up in a group by itself. The question then is why define peer groups at all? The answer is that given hospital-level data, users tend to compare hospitals. Without peer groupings to refer to, readers would compare a hospital with either the state level data or to another arbitrarily chosen hospital. Therefore, it was decided that this report would contain summary statistics for a hospital's peer group as well as for the hospital and the state. Having decided this, the next issue was the basis for the grouping, which is discussed next.

Among various factors which affect a hospital's average charges, location and case mix indicators play important roles in determining the complexity of patients treated in the hospital. Therefore, the bases for the 1993 hospital grouping are location (urban/rural) and the all-patient case-mix index, except for psychiatric and substance abuse hospitals and non-comparable hospitals.

The hospitals are assigned to peer groups according to 1996 UHDDB acute case-mix index (CMI). The 1996 UHDDB acute CMI is shown below.

Group 1: Acute Care, Urban, High CMI	LDS Hospital 1.4475
	University of Utah Hosp & Clinics 1.3357
Group 2: Acute Care, Urban, Upper Medium CMI	St. Mark's Hospital 1.1155
	McKay-Dee Hospital Cntr 1.0659
	Salt Lake Regional Medical Center 0.9747
	Utah Valley Medical Center 1.0635
Group 3: Acute Care, Urban, Lower Medium CMI	Cottonwood Hospital Med Center 0.7741
	Davis Hospital and Medical Center 0.7130
	Lakeview Hospital 0.9416

	Mountain View Hospital 0.8149
	Ogden Regional Medical Center 0.8883
	Pioneer Valley Hospital 0.8875
Group 4: Acute Care, Urban, Low CMI	Alta View Hospital 0.6263
	American Fork Hospital 0.5710
	PHC Hospital 0.8017
	Jordan Valley Hospital 0.5665
	Orem Community Hospital 0.4337
Group 5: Acute Care, Rural, High CMI	Ashley Valley Medical Center 0.6986
	Brigham City Community Hospital 0.8371
	Castleview Hospital 0.9461
	Dixie Medical Center 0.8308
	Logan Regional Hospital 0.6477
	Valley View Medical Center 0.6812
Group 6: Acute Care, Rural, Low CMI	Allen Memorial Hospital 0.5736
	Bear River Valley Hospital 0.5888
	Beaver Valley Hospital 0.5328
	Central Valley Med Center 0.6238
	Delta Community Medical Center 0.5608
	Fillmore Community Med Center 0.5371
	Garfield Memorial Hospital 0.5426
	Gunnison Valley Hospital 0.5147
	Kane County Hospital 0.5698
	Milford Valley Memorial Hospital 0.4786
	San Juan County Hospital 0.5427
	Sanpete Valley Hospital 0.5698
	Sevier Valley Hospital 0.6262
	Tooele Valley Regional Med Center NA
	Uintah Basin Med Center 0.6286
	Wasatch County Hospital 0.5749
Group 7: Psychiatric & Substance Abuse & Specialty Hospitals	Benchmark Regional North
	Benchmark Regional South
	Highland Ridge Hospital
	Olympus View Hospital
	Rivendell Psychiatric Center
	University Neuropsychiatric Institute
	Special Hospitals (not comparable)
	Bonneville Health and Rehabilitation
	HEALTHSOUTH Rehab Hosp of Utah
	Primary Children’s Med Center 1.4717
	South Davis Community Hospital
	The Orthopedic Specialty Hospital
	Utah State Hospital
	Veterans Hospital

Some industry experts contend that hospital comparisons are meaningful only when confined to a specific treatment, service or procedure. Thus, in analyzing total charges for Coronary Artery Bypass Graft (CABG), one would only compare among hospitals that perform non-urgent CABGs. Among these hospitals, a logical peer grouping would be based on the volume and severity mix of discharges associated with this procedure. While this is obviously ideal, it may not be possible to provide this in a summary report for consumers. There are many examples of state reports available to guide Mississippi's public reporting.

System Maintenance

When developing a new system the one constant is that there will be changes in the future. The questions that need answering when the system is first conceived and developed will evolve as the system becomes more established. The technologies available to implement the system will change. The level of sophistication of the data collection and data use staff will change. The combination of these factors will ensure that this brand new system will need to be changed in the future.

Factoring in the need for ongoing resources to maintain the state reporting system and planning for future growth is an important design feature.

IT Infrastructure

Legacy state reporting systems, which are still prevalent today, require a magnetic media (such as tape, CD, or diskette) as the principal input mechanism. These magnetic media would typically be sent from the provider to the organization responsible for the collection on a periodic basis (monthly, quarterly, bi-annually).

More recent state reporting systems, which are progressively becoming more common, transmit the data electronically using a secured internet site or a dedicated connection. This technology supports the same monthly, quarterly, or bi-annual data submission as well as providing the capability of real time submissions.

The other IT consideration is the storage of data in legacy "flat" files or in a relational data base. The level of sophistication of the data storage methodology is directly proportional to the complexity of the data system.

The hardware platform used by these state reporting systems is largely dependent on the volume of inpatient and outpatient discharges in each state as well as the access requirements to that data. The range of hardware solutions is a PC platform, a server based configuration, and a large scale mainframe.

Similarly, software needs to support submission and analytic uses of the data. The extent of software needed would (as with the criteria for selecting the hardware platform) be dependent on the volume of inpatient and outpatient discharges in the state as well as the access requirements to that data.

Developing a system using electronically transmitted data that is maintained in relational data bases would require a higher level of technical expertise than would be necessary for data sent via magnetic media and stored on "flat" files. The less technical solution, though, would require significantly more clerical staff to process the incoming data.

Compliance Issues

State data systems fall into two categories, voluntary or mandated collection. In states with mandated collection the rules and / or regulations usually include the penalties for non-compliance.

There are different levels of compliance to be considered when developing a new system. The first level would be a simple tracking mechanism. Such tracking would provide utilization information about the reporting providers. The next level of sophistication would use the data to assess the quality of care based on reported outcomes. In this case compliance would include the completeness of reported data. (e.g., reporting of Other Diagnosis codes to more fully document complications). The level of compliance checking will depend on the level of sophistication of the state reporting system.

It has been the experience of several states (e.g., Pennsylvania and New York) that publicly publishing compliance statistics has a bigger impact on achieving high levels of compliance on all levels. This would apply equally to voluntary and mandated data collection models.

An even better vehicle, though, for improving provider compliance submitting data is to use the data in meaningful ways. For example in New York State, the governor commissioned a report to address inefficiencies in the health delivery system for hospital and nursing services. Based on this report, recommendations were made to close or merge hospitals or nursing homes based on the data reported to the state. It is not coincidental that there was a flurry of activity to submit or correct the data being used prior to the publication of the report.

(www.nyhealthcarecommission.org/final_report.htm)

The Pennsylvania Cost Containment Council (www.phc4.org) publishes a Hospital Performance Report along with hospital comments. This is a high profile example of how using the data provides greater transparency on the state of health care using state data. Reports, such as this, continue to be the best vehicle for achieving compliance with state reporting requirements.

Cross-cutting Implementation Lessons Learned

NAHDO has been actively involved with state health data program development since 1986, facilitating state-to-state interaction and transfer of technologies and lessons learned. States seeking to implement and expand discharge data reporting systems are eager to seek help from their peers across the states and states with experience in implementation are willing to share what they have learned. Over the years, NAHDO has identified the following cross-cutting issues to statewide health data reporting:

- Data programs must strike a balance between data completeness and quality and the timeliness of data release and reports. Data quality improves over time and with use. Conclusions and reports must be developed with caution in the first years of data reporting.
- Data reporting requirements that are aligned with national standards reduce the reporting burden on the providers. Data elements and definitions that are not standard increase data reporting costs and may not be comparable across providers and states.

- A data review and validation process is an opportunity for providers to identify problems with the data before it is released and instills fairness and trust.
- Timeliness of coded data will always be a challenge for discharge data reporting, due to the nature of the reporting (at least 45 days after the close of a quarter). Data programs must balance the timeliness with accuracy.
- Data providers benefit from feedback to highlight reporting problems and coding inconsistencies.
- Providers are the most frequent users of the data. As key stakeholders, the guidance and input from the provider community is invaluable to system implementation.
- Budget and workforce constraints are common across most health data programs. Partnering with other entities, like public health or academic health centers, to analyze the data and borrowing of programs and tools from other states may help offset these constraints.
- Healthcare data systems are political; it is prudent to maintain a strong stakeholder base. This means producing products that are useful to the stakeholders, and continuously seeking input regarding new products.

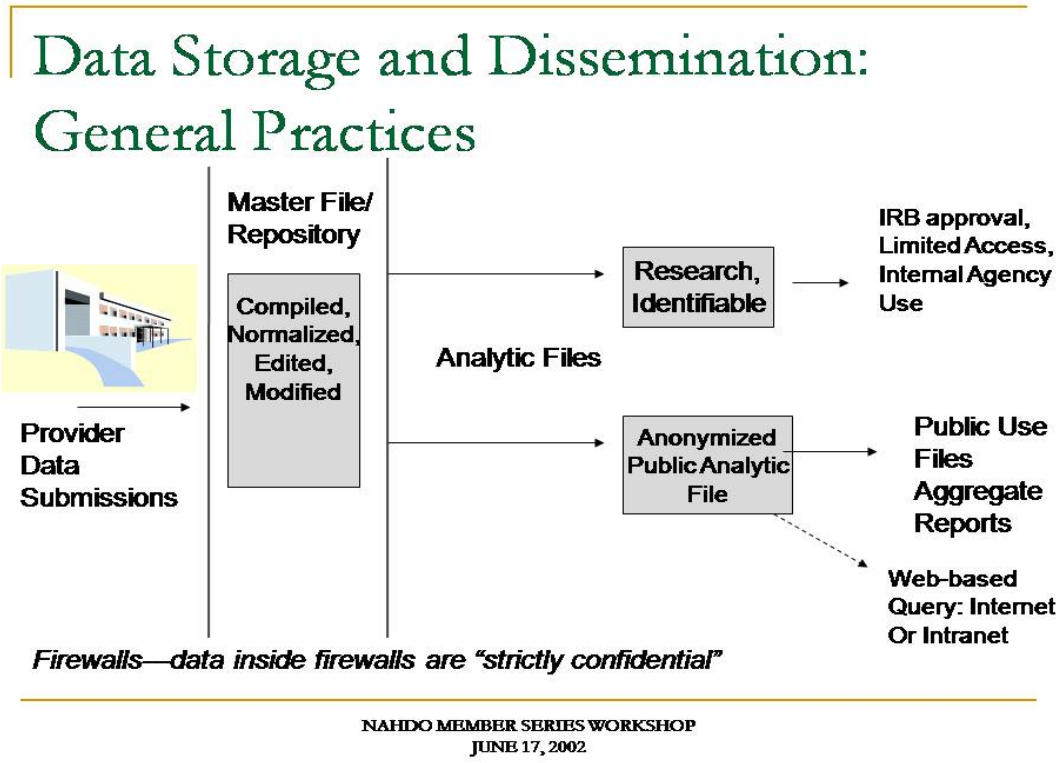
Data Access and Dissemination Policies

Most health data programs have carefully designed their data release policies to balance the protection of patient confidentiality and identification with appropriate use that benefits the public good. Both state and private health data programs, including hospital associations and other organizations, collect personal health information from providers and disclose the information to others, such as health service researchers, other providers, state health departments, and businesses. The disclosures are for purposes that improve the efficiency of the health system and public health policies, both of which are accepted and valued by the providers, but are made without individual patient authorization.

State health data programs perform this aggregation and dissemination function with a good track record over time of maintaining the confidentiality of the information through internal controls and through data use agreements that place restrictions on the use of the data by the receivers. Figure 1 illustrates the general health data program framework for receiving and disseminating data. Note that the master file, or repository, is generally classified by legislation or data policies as "strictly confidential", limiting the access to identifiable data to non-authorized program staff and, protected from discovery or requests for public records.

Figure 6 below illustrates the methods used to store and protect the data.

Figure 6. Data Storage & Dissemination: General Practices



The health data program uses a combination of methods and policies to add layers of protection to the data.

Data Modification

Behind the firewalls, the health data program, in a secure Information Technology (IT) environment, modifies or changes the raw data to make the data more uniform and also less individually-identifiable. Data modification techniques change the identifiable/raw data into a statistical abstract that anonymizes the individual. Methods include:

- Recoding and aggregation (example: date of birth is aggregated into age categories and dates of admission and discharge are calculated into length of stay and quarter of discharge). This is done in order to mask personal identity.
- Suppression: sensitive diagnoses or conditions (mental health, drug/alcohol, HIV)) or other conditions with less than a certain number, such as 5, are suppressed or hidden. This is to reduce the risk of re-identification for rare events or conditions protected from public disclosure by law. All patient identifying information is suppressed.
- Encryption: In states that collect patient a unique patient number, that number is encrypted, using an irreversible and stable algorithm that the health data program applies to change the number into an agency-assigned, de-identified number.

Additional methods, such as smoothing and data swapping, may also be applied to further mask the data and protect against re-identification.

Below is a draft table of data elements and how they were proposed to be released for the Wisconsin Physician Data Collection program, and a second section describing the availability of information within products. This type of pre-collection information is useful when explaining how information is protected to various stakeholders.

Table 8. Data Elements Available for Analysis and Release
Physician Administrative Data Collection System
Bureau of Health Information Draft, for purposes of discussion (Wisconsin) 5/4/2000

Data Elements:	Data Elements Approved by Statute for Standard Files	Data Elements Requiring Independent Review Board Approval	Confidential Data Elements (Patient-Identifiable)	Other Data Elements
Patient Information				
City, town, or village			✓	
Zip code		✓		
County of residence	✓			
Gender	✓			
Date of birth			✓	
Age		✓		
Birth month/year		✓		
Five-year age category	✓			
Payer Information				

Primary payer category	✓			
Secondary payer category	✓			
Provider Information				
Billing physician ID		✓		
Performing physician ID		✓		
Referring physician ID		✓		
Physician specialty				✓
Health care facility location	✓			
Practice type				✓
Medical group or corporate affiliation				✓
For Each Visit				
Date of visit			✓	
Month or quarter of visit	✓			
Procedure codes	✓			
Diagnosis codes	✓			
Charge(s)	✓			
Adjustment factor(s)				✓

Variables Available in Bureau of Health Information Products
Physician Administrative Data Collection System
Bureau of Health Information
Draft, for purposes of discussion 5/4/2000

Table 9.

Availability in BHI Products:	Available as basis for BHI Standard Reports (Aggregate Information Products)	Available in Custom Data Files and Custom Reports	Available in Standard Public-Use Data Files
Data Elements Approved by Statute for Standard Data Files	Yes	Yes	Yes
Data Element Requiring Independent Review Board Approval	Yes	With case-by-case IRB approval	No
Confidential Data Element (Patient- Identifiable)	Yes	No	No
Other Data Element	Yes	With one-time IRB approval	With one-time IRB approval

Regulatory Policies

Even prior to the Health Insurance Portability and Accountability Act of 1996 (HIPAA) and the HIPAA Privacy Regulations which defines provisions for the use and protection of personal health information,, many health data programs had policies in place to govern the release and protection of data. In states with legislation, these policies were guided by law, which in most

cases, is more restrictive (and therefore exempt) from the HIPAA Privacy Rule provisions. Policies in place across health data programs may be included in the authorizing legislation. These policies generally include the following actions:

1. Classification of data fields as public, restricted, and confidential with policies to govern each. The data commission, oversight committee, or board generally adopts detailed data release policies. Examples of such classification are as follows:

Table 10. Public Use Data Elements (example)

Day of week of admission	Day procedure performed
Day of week of discharge	E-code(s)
Month of discharge	Expected Payer
Patient age or age category on admission	Admission source
Discharge Status	Patient count of residence and/or zip
Primary and secondary diagnoses/procedures	Length of Stay (calculated)
Discharge Status	Charge and/or revenue codes
Provider identifier	

Most states collect, but never release direct patient identifiers or other patient identifiable information protected from release by federal or state law. Most states have prohibitions within the statute for any attempt to identify or contact the patient from the discharge data systems. However, this does not apply to the actual provider of care, who may contact their patient or be able to identify their patient within the data. The healthcare provide clearly has other records that can assist them in identifying individuals within the larger data set as being patients in their facility.

a) Processes for applying for the use of restricted or protected data fields for research purposes. Restricted data elements may require a data board review process as well as a data use agreement. In some states, an Institutional Review Board (IRB) approval is required, either by law or as an added layer of protection. Examples of data elements that states restrict are included in Table 11:

Table 11. Restricted Data Elements

Dates (admit, discharge, patient DOB)	Sensitive diagnoses/conditions
Medical record number	Unique patient number (encrypted)
Exact charges	

Note: Approval by an Institutional Review Board (IRB) for research projects beyond that required by their university or college or public health authority applications that use healthcare data, can impose an onerous burden and hinder research. Most uses and users can use the de-identified data set that has few restrictions. HIPAA provides for use by public health programs,

such as trauma and birth defects registries, under the public health exemptions (Section 164.512) and some state health data programs have worked out data release and use agreements for public health uses.

b) Public use file Data Use Agreements that govern the terms and conditions for using the data. The use agreements specifically prohibit the attempt to re-identify patients or link the data to other data sets, with penalties defined. Many states are defining various levels of public and restricted use files, depending on the type of users. For example, the Massachusetts Division of Health Care Finance and Policy prepares a limited use data file for six different levels of detail, with an application process that requires justification for the level of request.

State health data programs, both public and private, were compelled to review their state laws and their data release and access policies with the enactment of the HIPAA Privacy Rule provisions. In the case of mandated programs, even where the state laws were more restrictive and thus exempt from HIPAA, data modification and regulatory procedures were reviewed and aligned with HIPAA provisions. In the case of private, voluntary programs that are subject to HIPAA provisions, business use agreements and limited (de-identified data sets) are the mechanism by which access and release is governed.

Note: The governance of a health data program may influence the availability of both public and research data, so therefore this is an important decision for the State of Mississippi. NAHDO recommends legislation to govern data disclosure and protections that are reasonable and equitable across users and uses. Private, voluntary systems are more likely to vary in their release and access policies, thus hindering public health and research in those states without a publicly controlled system.

Pricing of Products: Most organizations charged with statewide collection of hospitalization data generate some revenues through data sales, in order to partially offset their data collection expenditures. The current data sales practices across states show no standard pricing formula. The pricing structure is, however, generally a function of the volume of discharges, the intended use of the data, the level of details required in the data file (e.g. public as opposed to research data), the availability of general budget for data collection, and a host of other factors.

The prices of annual hospital discharge data files vary dramatically across states. Florida recently reduced their price from $600 to $100 this year, to promote access and use of the public use files. In contrast to Florida prices, the Pennsylvania data which are enhanced with additional clinical data, are priced at $9,000 for non-profit use; and $33,000 for commercial vendors.

Data stewards follow no particular data pricing structure. Some charge by number of record -- e.g., the South Carolina Office of Research and Statistics charges $1.25/1,000 records. Others price their data per release and data product. Massachusetts sells their state inpatient data for $1,000 to $7,000 for an annual file, depending upon the purpose of the use.

In general, states with a larger number of discharges generate greater amounts of revenues but there are exceptions. NAHDO data indicated that *"for smaller states the range is from 10K to 105K, larger states range from $113K to 3.1million in annual data sales. "*

Types of Products

Data stewards make large number pf products available to their users, which typically include[32]:

- Public data set;
- Research restricted data set;
- HCUP data files;
- Custom Data requests;
- Standard reports;
- Standard aggregate utilization profile reports;
- Special reports or analyses on a specific topic, such as Coronary Artery Bypass Graft outcomes, or healthcare acquired infections;
- Data briefs (one page information sheets on a single topic);
- Web-based query systems.

Public use data sets are commonly designed to provide general health care information to a wide spectrum of users after stripping off data elements containing sensitive or patient identifiable information. From these data, individual patients cannot be identified. Patient's age range, diagnosis codes, physician's name, and payer codes are typically included, but other variables such as county of residence, and severity of illness are also included by some states.

Standard research restricted data may include all of the variables collected in the hospital discharge abstract if needed for the specific research project. Only those data that are deemed necessary for the project are released and these data are only made available to researchers that have fulfilled requirements and meet standards outlined in the data sharing policies, including Institutional Review Board approval and potentially, other policy or privacy committee approval for bona fide research.

The national use of state data bases are growing. Some states (currently 38) participate in the Healthcare Cost and Utilization Project (HCUP) sponsored by the Agency for Healthcare Research and Quality (AHRQ). Over 30 states provide data to the National Highway Traffic Safety Council (NHTSA-funded) Crash Outcomes Data and Evaluation System (CODES) which links motor vehicle crash reports with trauma, hospitalization, drivers license, and mortality data to evaluate the outcomes of accidents and shape prevention and safety policies. The Center for Disease Control and Prevention's (CDC) Environmental Public Health Tracking Network is standardizing environmental health indicators derived from inpatient and ED data to evaluate environmental health outcomes across states.

Data stewards also respond to custom data requests, which are usually filled based on staff resources. Requesting agencies are generally asked to pay for staff time and access to records (usually by number of variables accessed, and by number of quarters or years) to accommodate special requests.

[32] National Association of Health Data Organizations. NAHDO Membership Survey. 2006.

Web-based Data Query Systems

Many state agencies have to respond to increasingly complex and numerous requests for health statistics; in order to save staff resources they develop and use Web-based Data Query Systems (WDQSs). A WDQS is an Internet-based application enabling dynamic database query, facilitating information retrieval by the direct manipulation of query variables and rapid retrieval of query results[33]. During the last decade, there has been a proliferation of WDQSs implementation by state and local public health agencies from 11 known systems in 2001, to 43 web query systems in 2005[34]. Examples of state-developed query systems include: The Missouri Department of Health and Human Services' Missouri Information for Community Assessment (MICA), the Massachusetts Department of Public Health's MassCHIP system, the Utah Department of Health's Indicator-based Information System for Public Health (IBIS-PH), and Washington's King County Health Department's Vista system. Both Missouri's MICA and Utah's IBIS query systems have been adopted across multiple states at no charge for the technology to the adopting states, except for the adopting states' IT, training, and analytic infrastructure costs.

Forecasting the budgetary impact of a WDQS is important, yet challenging. But it is well established that additional funding is essential to implement a WDQS[35]; [36]. Building a WDQS now is, however, considerably less complicated and less resource intensive , because states are willing to share their systems with other states. States are actively exporting their technologies to peer organizations, and several guidelines are available for states implementing these systems. For instance, the CDC funded a tool for states, consisting of two products: a report titled "Web-Based Systems for the Dissemination of Health-Related Data: A Guide for Public Health Agencies Developing, Adopting, or Purchasing Interactive Web-based Data Dissemination Systems" and a model prototype web system, a static system replicating the kind of system a state may want to develop, adopt, or purchase[37].

WDQSs require upfront resources dedicated to building the system, but once in place they offer many benefits as they help streamline the data distribution process with considerably limited ongoing resource commitment. WDQSs allow users to immediately see the result of their query and allow policy makers, health professionals, and the public to submit questions (requests for data) and receive answers (tables) over the Internet[38]. The online distribution of information through WDQS reduces printing costs. WDQSs also provide common data sources and indicators for community health assessments, connecting data suppliers and users in ways never before possible.

---

[33] Ahlberg, C., & Shneiderman, B. Visual information seeking: Tight coupling of dynamic query filters with starfield displays. In R. M. Baeker, G. Grudin, W. A. S. Buxton, & S. Greenberg (Eds). Readings in human-computer interaction: Toward the year 2000 (pp. 450-456). San Francisco: Morgan Kaufmann; 1993

[34] Friedman, DJ, Parrish RG. Characteristics and desired functionalities and data sets of state Web-based Data Query Systems. Journal of Public Health Management & Practice. 2006; 12(2):196-200

[35] National Association of Health Data Organizations (NAHDO) and National Association of Public Health Statistics and Information Systems (NAPHSIS) Inventory of State Using WDQS, 2001

[36] National Association of Health Data Organizations (NAHDO) and National Association of Public Health Statistics and Information Systems (NAPHSIS) Inventory of State Using WDQS, 2004.

[37] CDC, Division of Public Health Surveillance and Informatics. Web-Based Systems for Dissemination of Health-Related Data: A Guide for Public Health Agencies Developing, Adopting, or Purchasing Interactive Web-Based Data Dissemination Systems online http://www.cdc.gov/epo/dphsi/asb/orcmacro.htm retrieved Sept 29, 2005.

[38] Wisconsin DOH. Wisconsin Interactive Statistics on Health (WISH) available at: http://dhfs.wisconsin.gov/wish/

State health data programs with the least resources and thin analytic infrastructure, like Mississippi, may benefit the most from a WDQS, yet they may not have the resources to fully implement such a system. For example, when Arkansas' health data program was ready to disseminate hospital data via the web, Arkansas adopted an early version of Utah's query system. Utah provided technical assistance and training to Arkansas to permit implementation at the lowest cost to Arkansas possible. States with query systems are willing to provide their systems to states seeking to adopt such programs, thus reducing the costs and maximizing the utility of the data.

Recommendation: During the planning phases, the Mississippi health data program should develop a data release plan and establish policies that support the plan. The plan should provide for the release of de-identified data in the form of a public use data set, controlled by the use of a data use agreement; the release of a research-oriented data set for bona-fide research and federal programs, such as the AHRQ Healthcare Cost and Utilization Project (HCUP), either with IRB approval or data policy board review.

Recommendation: Mississippi should price these data products in a manner that balances the need for data sales revenues with data access for authorized uses by legitimate users.

Strategies for Adding Value

States, following the lead of national standard-setting entities and CMS, are enhancing their discharge data to more fully support quality reporting and measurement. Over time, the addition of diagnoses and procedure codes, external cause of injury codes were implemented based on recommendations by the National Committee on Vital and Health Statistics (NCVHS) which identified these data elements as essential for state discharge data systems. Over time, these data elements have been added to the billing form and adopted by the NUBC. With the implementation of the UB-04, additional data elements, such as present on admission (POA) indicator, race and ethnicity, and Do Not Resuscitate (DNR) order indicator have been added to the UB for the purpose of state reporting. The UB-04 now officially supports data necessary for institutional claims submission as well as state reporting purposes.

As of July of 2007, many states have adopted the new data collection format UB-04. Some states have added additional detail or data elements to the core UB format, such as the patient's race and ethnicity, unique patient number, additional diagnosis indicators, and present on admission, etc.

A number of states are engaged in pilot programs to enhance their existing UB-04 data using either present on admission codes or the addition of a limited set of numeric laboratory data. These additions were shown in a paper by Michael Pine, et.al.[39], to be cost-effective additions to state administrative data. The additional data elements strengthen the state data by allowing for more accurate severity adjustment; in this case for the AHRQ IQIs. [In addition, the Present on Admission data element is critical to understanding whether a complication or infection was present at the time of admission versus acquired during the hospital stay.] Not only are the added data elements useful, they are less expensive to collect than data abstraction of information from the clinical record. See the table below for differences in cost regarding data collection. When using the table please keep in mind that these are marginal costs, over and above the cost of other standard UB-04 data elements.

Table 12: Marginal Cost-Effectiveness of Supplementing Administrative Data with Increasingly Costly Data to Improve Risk Adjustment of IQIs and PSIs

Best Case Scenario			
Added Data	Increase in Cost per Abstracted Record	Increase in Effectiveness1	Cost-Effectiveness2
Present on Admission Code	$0.33	8.44%	$0.04
Numerical Laboratory Data	$0.67	16.68%	$0.04

[39] M.Pine, H.S. Jordan, A. Elixhauser, et.al., Enhancement of Claims Data to Improve Risk Adjustment of Hospital Mortality. JAMA, 297(1), January 3, 2007.

Vital Signs; Other Laboratory Data	$0.91	8.55%	$0.11
Key Clinical Findings; Aggregate Scores	$4.80	8.72%	$0.55
Worst Case Scenario			
Added Data	Increase in Cost per Abstracted Record	Increase in Effectiveness1	Cost-Effectiveness2
Present on Admission Code	$0.33	8.44%	$0.04
Numerical Laboratory Data	$3.33	16.68%	$0.20
Vital Signs; Other Laboratory Data	$9.86	8.55%	$1.15
Key Clinical Findings; Aggregate Scores	$24.91	8.72%	$2.86

[1] Effectiveness = percentage of hospitals with less than 0.5 standard deviations of data-related bias.
[2] Cost-Effectiveness = cost per one percent increase in effectiveness.

For an example of state implementation experiences, the following link contains a paper describing the evolution of an enhanced data collection system in New York State, including its authorization, development, testing and implementation. (Author: Bob Davis, Consultant, Health Data Standards, LLC) http://www.phdsc.org/implcasestudies/case4EC_x12.htm

Recommendation: Since Mississippi hospitals will be reporting Present on Admission (POA) indicator to Medicare beginning October 1, 2007, Mississippi's inpatient health data reporting requirements should include POA as a required core data element. For reasons of provider reporting cost and burden, NAHDO recommends that Mississippi NOT include clinical or laboratory data elements with initial inpatient reporting requirements. These elements could be assessed for inclusion at a later time as automation and standards evolve.

Data Linkage

Hospital discharge data can be combined with other data sources to fill important information gaps. States with hospital data systems frequently augment and enhance their data by using a technique called record level data linkage. Record linkage is the task of deciding whether two or more records belong to the same entity (individuals, hospitals, geographic region, families, or households) from one or more data sources. Linking the healthcare and public health data at an individual level requires presence of a unique identifier in all databases being linked. Since such unique identifiers, generally the Social Security Number (SSN), are often not available for some or all of the records, accurate record linkage becomes difficult.

An example of a common data linkage is linking inpatient hospital discharge data with birth certificate data; this combines the billing utilization data with maternal and newborn clinical data for robust outcomes studies. The Crash Outcome Data and Evaluation System (CODES) is an example of a national data linkage project initiated by the National Highway Transportation and Safety Agency (NHTSA) in collaboration with states, in which hospital discharge data,

emergency department data, ambulance records, post-acute care data are linked to drivers license, crash report, and death certificate data to address a variety of research questions related to motor vehicle accidents and formulate policies relative to safety.

Hospital discharge data can be combined with other data sets to study healthcare use and outcomes for special populations or for targeted conditions. Hospital discharge data are used in community health assessment projects to assess variation in admissions for preventable conditions such as asthma, injuries, or chronic diseases. Many of these conditions may be preventable with the proper access to outpatient care or through community outreach and educational initiatives.

Some states link hospital discharge data with other databases routinely as well as on special requests.

Some of these applications are:

a) De-duplication: Annual hospital discharge data file or ED data file may contain several duplicates because it is completed through a batch process with weekly, monthly, or quarterly data, and it goes through several rounds of edits. Individual discharge record level linkage allows states to identify and remove duplicates.

b) For augmenting hospital discharge data: States augment hospital discharge data by borrowing information from other sources (e.g. linking ED visits data, hospital discharge data, Medicaid Enrollee data and death certificate data to investigate the burden of heart disease) for more sophisticated analyses, and to reduce data collection burden (e.g. link hospital characteristics file like AHA files with hospital discharge data).

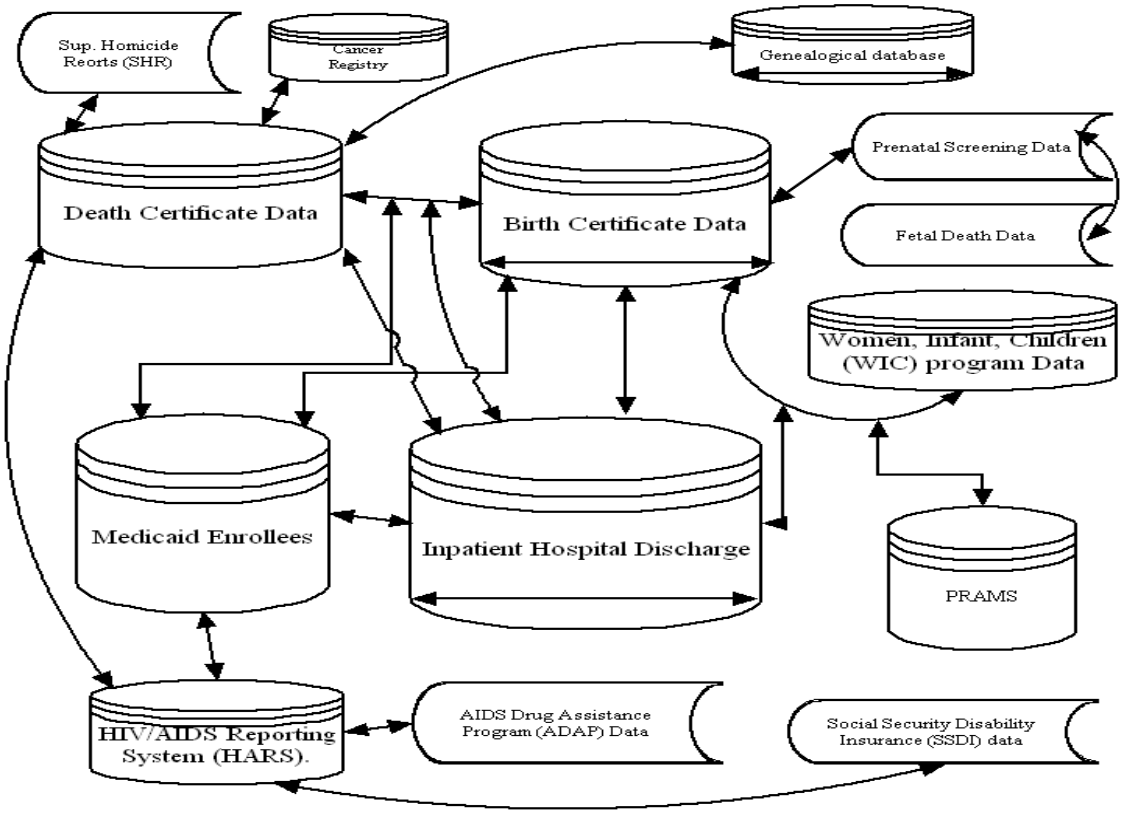
c) Hospital Readmissions Indicator: State also use record level linkage of hospital discharge data file for tracking hospital readmissions and make this indicator available for research on quality of care.

States without patient identifiers can make use of record linkage software that uses a combination of existing data elements to link data using a probabilistic model. Some of the probabilistic linkage software is available to states at no cost; these include: Link King and Link Plus. State to state sharing of best practices and lessons learned through NAHDO helps states in decisions about record linkage.

Inhibitors of record linkage include non availability of semi-unique and unique identifiers, needed for linkage. Further, there can be privacy provisions which forbid linkage because linkage can create a greater risk of exposure to sensitive information and invasion of individual privacy.

States regularly link Hospital Discharge data with Birth Certificate, Death Certificate, and Medicaid Enrollee Data, as shown in the Figure 7 below.

Figure 7: Data sets linked by respondents of NAHDO Survey on Linkage



Note: The arrow inside a database indicates linkage of data files within the database.

Given the advantages of application of record linkage reviewed in this section, NAHDO recommends that Mississippi collects SSN and other variables required for record linkage. Standard variables required for discharge record linkage are name, birth date, gender, and zip code. States collecting other information such as SSN can use that as well.

Recommendation: The Mississippi data dissemination and use plan, recommended earlier, should provide for data access by the authorized public health programs and support data linkage and data integration. This is an important mechanism to fill important data gaps and reduce the burden on providers to report multiple registry and surveillance data elements.

Appendix 1

Population profile of Mississippi compared with neighboring states, and U.S.

Population Measures	Mississippi	Arkansas	Louisiana	USA
Population, 2006 estimate	2,910,540	2,810,872	4,287,768	299,398,484
Population, percent change, April 1, 2000 to July 1, 2006	2.30%	5.10%	-4.10%	6.40%
Persons under 5 years old, percent, 2005	7.30%	6.70%	7.10%	6.80%
Persons under 18 years old, percent, 2005	25.60%	24.30%	25.40%	24.80%
Persons 65 years old and over, percent, 2005	12.30%	13.80%	11.80%	12.40%
Female persons, percent, 2005	51.40%	51.00%	51.40%	50.70%
White persons, percent, 2005 (a)	61.20%	81.30%	64.10%	80.20%
Black persons, percent, 2005 (a)	36.90%	15.70%	33.10%	12.80%
American Indian and Alaska Native, percent, 2005 (a)	0.40%	0.70%	0.60%	1.00%
Asian persons, percent, 2005 (a)	0.70%	1.00%	1.40%	4.30%
Native Hawaiian; Other Pacific Islander, percent, 2005 (a)	0.00%	0.10%	0.00%	0.20%
Persons reporting two or more races, percent, 2005	0.60%	1.20%	0.80%	1.50%
Persons of Hispanic or Latino origin, percent, 2005 (b)	1.70%	4.70%	2.80%	14.40%
White persons not Hispanic, percent, 2005	59.70%	77.00%	61.60%	66.90%
Living in same house in 1995 and 2000, pct 5 yrs & over	58.50%	53.30%	59.00%	54.10%
Foreign born persons, percent, 2000	1.40%	2.80%	2.60%	11.10%
High school graduates, percent of persons age 25+, 2000	72.90%	75.30%	74.80%	80.40%
Bachelor's degree or higher, pct of persons age 25+, 2000	16.90%	16.70%	18.70%	24.40%
Percent of Persons with a disability, age 5+, 2000	20.9%	20.5%	20.5%	16.6%
Mean travel time to work (minutes), workers age 16+, 2000	24.6	21.9	25.7	25.5
Homeownership rate, 2000	72.30%	69.40%	67.90%	66.20%
Housing units in multi-unit structures, percent, 2000	13.30%	13.90%	18.70%	26.40%
Median value of owner-occupied housing units, 2000	$71,400	$72,800	$85,000	$119,600
Households, 2000	1,046,434	1,042,696	1,656,053	105,480,101
Persons per household, 2000	2.63	2.49	2.62	2.59
Median household income, 2004	$34,278	$35,295	$35,216	$44,334
Per capita money income, 1999	$15,853	$16,904	$16,912	$21,587
Persons below poverty, percent, 2004	19.30%	15.60%	19.20%	12.70%
Source U.S. Census Bureau: State and County QuickFacts. At http://quickfacts.census.gov/qfd/				

Selected Health Status Indicators by race in Mississippi, compared with neighboring states, and U.S

Indicators	Mississippi	Arkansas	Louisiana	USA
Early prenatal care by race and Hispanic origin of mother, 2002-04				
All races	84.4	81.1	84.5	83.8
White	90.8	84.6	90.8	88.9
African American/Black	77.2	73.3	75.9	76.1
Low-birthweight live births (Percent of live births weighing less than 2,500 grams), by race and Hispanic origin of mother, 2002-04				
All races	11.4	8.92	10.69	7.94
White	8.5	7.78	7.76	7.05
African American/Black	15.31	14.54	14.91	13.56
Infant mortality rates (Infant deaths per 1,000 live births), by race and Hispanic origin of mother -- 2001-03				
All races	10.5	8.5	9.8	6.9
White	7.1	7.6	7	5.7
African American/Black	14.7	13.1	13.9	13.6
Age-adjusted death rates per 100,000 population, by race, Hispanic origin, 2002-04				
All races	1,010.60	939.9	994.5	826.5
White	949.6	915.2	925.4	811
African American/Black	1,166.30	1,160.10	1,201.50	1,059.70
Source: National Center for Health Statistics. Health, United States, 2006: With Chartbook on Trends in the Health of Americans. Hyattsville, MD: 2006				

Healthcare access and utilization, Mississippi compared with neighboring states, and U.S

Indicators	Mississippi	Arkansas	Louisiana	USA
Number of active physicians per 10,000 civilian population, 2004	18.4	20.5	25.3	26.3
Community Hospital Beds per 1,000 resident population, 2004	4.5	3.5	3.8	2.8
Occupancy rates (%) in community hospitals, 2004	58	59	61	67
State mental health agency per capita expenditures for mental health services, 2003	$93	$30	$51	$92
Medicare enrollees, enrollees in managed care, payment per enrollee, and short-stay hospital utilization, 2003				
Enrollment in thousands	438	443	621	40,203
Percent of enrollees in managed care	0.6	0.5	11.3	13.1
Payment per fee-for-service enrollee	6,794	5,680	7,826	6,618
Discharges per 1,000 enrollees	446	403	461	369
Average length of stay in days	6.2	5.8	6.1	6
Medicaid recipients, recipients in managed care, and payments per recipient, 2003				
Recipients in thousands	717	702	995	51,971
Percent of recipients in managed care	45	67	59	59
Payment per recipient	3,582	3,151	3,632	$4,487
Percent of Persons enrolled in health maintenance organizations (HMOs), 2005	0.1	6.4	10.7	23.4
Percent of population without health insurance coverage, 2002-04	17.2	16.7	18.8	15.5
Source: National Center for Health Statistics. Health, United States, 2006: With Chartbook on Trends in the Health of Americans. Hyattsville, MD: 2006				

Appendix 2

State	State Mandate	Organization Name	Organization Type	Inpatient Data Collected in 2006	Ambulatory Surgery	Emergency Department	Quality Reports	URL
Alabama		No Data Collection Reported						
Alaska	no	Alaska State Hospital and Nursing Home Association	Hospital Association	Voluntary				www.ashnha.com/
Arizona	yes	Arizona Department of Health Services	Public Agency	Mandatory	Mandatory	Mandatory		www.azdhs.gov/
Arkansas	yes	Arkansas Department of Health	Public Agency	Mandatory				www.healthyarkansas.com/
California	yes	Office of Statewide Health Planning & Development	Public Agency	Mandatory	Mandatory	Mandatory	✓	www.oshpd.state.ca.us
Colorado	no	Colorado Health & Hospital Association	Hospital Association	Voluntary	Voluntary			www.cha.com
Connecticut (3)	yes	Chime, Inc.	Hospital Association	Voluntary	Voluntary	Voluntary		www.chime.org
		Connecticut Office of Health Care Access	Public Agency	Mandatory				www.ohca.state.ct.us/
Delaware	yes	Delaware Health Statistics Center	Public Agency	Mandatory				www.dhss.delaware.gov/dhss/dph/hp/healthstats.html
District of Columbia	yes (1)	District of Columbia Hospital Association	Hospital Association	Voluntary				http://dchealth.dc.gov/doh/site/default.asp
Florida	yes	Florida Agency for Health Care Administration	Public Agency	Mandatory	Mandatory	Mandatory	✓	www.fdhc.state.fl.us/SCHS/index.shtml
Georgia	yes	GHA: An Association of Hospitals & Health Systems	Hospital Association	Mandatory	Mandatory	Mandatory		www.gha.org
Hawaii	no	Hawaii Health Information Corporation	Private Agency	Voluntary	Voluntary	Voluntary		www.hhic.org
Idaho		No Data Collection Reported						
Illinois	yes	Illinois Department of Public Health	Public Agency	Mandatory	Mandatory			www.idph.state.il.us/
Indiana	yes	Indiana Hospital & Health Association	Hospital Association	Mandatory	Mandatory	Mandatory		www.inhha.org/
Iowa	yes	Iowa Hospital Association	Hospital Association	Mandatory	Mandatory	Mandatory		www.ihaonline.org
Kansas	yes (1)	Kansas Hospital Association	Hospital Association	Voluntary	Voluntary	Voluntary		www.kha-net.org

Options for a Statewide Health Data Reporting System in Mississippi

State	State Mandate	Organization Name	Organization Type	Inpatient Data Collected in 2006	Ambulatory Surgery	Emergency Department	Quality Reports	URL
Kentucky	yes	Kentucky Department for Public Health	Public Agency	Mandatory	Mandatory		✓	www.chs.state.ky.us/publichealth
Louisiana	yes	Louisiana Department of Health and Hospitals	Public Agency	Mandatory				www.dhh.state.la.us/
Maine	yes	Maine Health Data Organization	Public Agency	Mandatory	Mandatory	Mandatory	✓	www.healthweb.state.me.us/start.asp
Maryland	yes	Health Services Cost Review Commission	Public Agency	Mandatory	Mandatory	Mandatory	✓	www.hscrc.state.md.us
Massachusetts	yes	Division of Health Care Finance and Policy	Public Agency	Mandatory		Mandatory	✓	www.state.ma.us/dhcfp
Michigan	no	Michigan Health & Hospital Association	Hospital Association	Voluntary	Voluntary			www.mha.org/mha_app/index.jsp
Minnesota	yes (1)	Minnesota Hospital Association	Hospital Association	Voluntary	Voluntary	Voluntary		www.mnhospitals.org
Montana	no	MHA - An Association of Montana Health Care Providers	Hospital Association	Voluntary				www.mtha.org/index.htm
Mississippi		No Data Collection Reported						
Missouri (3)	yes	Missouri Department of Health and Senior Services	Public Agency	Mandatory		Mandatory	✓	www.dhss.mo.gov/
		Hospital Industry Data Institute	Hospital Association	Voluntary	Voluntary	Voluntary		http://web.mhanet.com/
Nebraska	no	Nebraska Hospital Association	Hospital Association	Voluntary	Voluntary	Voluntary		www.nhanet.org
Nevada	yes	Center for Health Information Analysis at University of Nevada at Las Vegas	Public Agency	Mandatory	Mandatory			www.unlv.edu/Research_Centers/chia/
New Hampshire	yes	New Hampshire Department of Health & Human Services	Public Agency	Mandatory	Mandatory	Mandatory	✓	www.dhhs.state.nh.us/DHHS/DHHS_SITE/default.htm
New Jersey	yes	New Jersey Department of Health & Senior Services	Public Agency	Mandatory	Mandatory	Mandatory	✓	www.state.nj.us/health/hcsa/index.html
New Mexico	yes	New Mexico Health Policy Commission	Public Agency	Mandatory				http://hpc.state.nm.us/
New York	yes	New York State Department of Health	Public Agency	Mandatory	Mandatory	Mandatory	✓	www.health.state.ny.us/nysdoh/sparcs/sparcs.htm
North Carolina	yes	North Carolina Hospital Association	Hospital Association	Mandatory	Mandatory			http://facility-services.state.nc.us
North Dakota	yes (2)	North Dakota Department of Health	Public Agency					www.health.state.nd.us/
Ohio	no	OHA: Ohio Hospital Association	Hospital Association	Voluntary	Voluntary	Voluntary		www.ohanet.org/

Options for a Statewide Health Data Reporting System in Mississippi

State	State Mandate	Organization Name	Organization Type	Inpatient Data Collected in 2006	Ambulatory Surgery	Emergency Department	Quality Reports	URL
Oklahoma	yes	Oklahoma State Department of Health	Public Agency	Mandatory	Mandatory			www.health.state.ok.us/
Oregon	yes	Oregon Association of Hospitals & Health Systems	Hospital Association	Mandatory	Mandatory			www.oahhs.org
		Office for Oregon Health Policy and Research	Public Agency				✓	www.ohpr.state.or.us
Pennsylvania	yes	Pennsylvania Health Care Cost Containment Council (PHC4)	Public Agency	Mandatory	Mandatory		✓	www.phc4.org
Rhode Island	yes	Rhode Island Department of Health	Public Agency	Mandatory		Mandatory	✓	www.health.state.ri.us/
South Carolina	yes	South Carolina State Budget & Control Board	Public Agency	Mandatory	Mandatory	Mandatory		www.ors.state.sc.us
South Dakota	no	South Dakota Association of Healthcare Organizations	Hospital Association	Voluntary	Voluntary	Voluntary		www.sdaho.org/
Tennessee	yes	Tennessee Hospital Association	Hospital Association	Mandatory	Mandatory	Mandatory		www.tha.com
Texas	yes	Texas Health Care Information Council	Public Agency	Mandatory	Mandatory	Mandatory	✓	www.dshs.state.tx.us/thcic/
Utah	yes	Office of Health Statistics, Utah Department of Health	Public Agency	Mandatory	Mandatory	Mandatory	✓	http://health.utah.gov/hda/
Vermont (3)	yes	Vermont Association of Hospitals and Health Systems	Hospital Association	Voluntary	Voluntary	Voluntary	✓	www.vahhs.com/
		Division of Health Care Administration	Public Agency	Mandatory			✓	www.bishca.state.vt.us/HcaDiv/hcadefault.htm
Virginia	yes	Virginia Health Information	Private Agency	Mandatory	Mandatory		✓	www.vhi.org
Washington	yes	Washington State Department of Health	Public Agency	Mandatory				www.doh.wa.gov/data/data.htm
West Virginia	yes	West Virginia Health Care Authority	Public Agency	Mandatory				www.hcawv.org
Wisconsin	yes	Wisconsin Hospital Association	Hospital Association	Mandatory	Mandatory	Mandatory		www.dhfs.state.wi.us/healthcareinfo

Options for a Statewide Health Data Reporting System in Mississippi

State	State Mandate	Organization Name	Organization Type	Inpatient Data Collected in 2006	Ambulatory Surgery	Emergency Department	Quality Reports	URL
Wyoming	no	Wyoming Hospital Association	Hospital Association	Voluntary				www.wyohospitals.com/
Totals	**39(4) States with a Mandate to Collect**			**47 States With Systems**	**35 States With Systems**	**27 States With Systems**	**18 States With Reports**	

notes:

1. District of Columbia, Kansas, Minnesota: The Department of Health has a mandate to collect data, but the hospital association collects the data voluntarily.
2. North Dakota: The Department of Health has a reporting mandate but has not collected data since 2005. Number in mandate totals represents only mandated states who collect data and does not include ND.
3. Conneticut, Vermont & Missouri have dual systems where data is collected by both the Health Department and the Hospital Association.
4. Only 38 of the 39 states with mandates actually collect data (see note #2)

Appendix 3

State	Data Type	Statute or Regulation Authorizing Collection
Arkansas	IP	Arkansas Code Annotated 20-7-301 et seq.
California	IP	California Code of Regulations, Title 22, Division 7, Ch. 10-Health Facility Data, Article 8. Discharge Data Reporting.
California	AC	California Health and Safety Code
Colorado	IP	None
Colorado	AS	None
Connecticut	IP	
Delaware	IP	Delaware Code/Title 16 Health & Safety; Part II Regulatory Provisions Concerning Public Health; Chapter 20 Uniform Health Data
Florida	IP	Florida Statutes, Chapter 408.061 (11)
Florida	AS	Florida Statutes, Chapter 408.061 (11)
Georgia	IP	GHA is not named in law, however the law stipulates that if another entity was collecting, they would have to use GHA.
Illinois	IP	Collection, Disclosure, and Confidentiality of Health Statistics (77 Ill. Administrative Code 1005); PA 92-0597; 20 ILCS 2215/2-1
Illinois	AS	PA 92-0597; 20 ILCS 2215/2-1
Indiana	IP	IC 16-21-6-6
Indiana	AS	
Iowa	IP	641 Iowa Administrative Code 177.3
Iowa	AS	641 Iowa Administrative Code 177.3
Kansas	IP	KS 65-6801
Kentucky	IP	Administrative Regulation 902 KAR 17:040
Kentucky	AS	Administrative Regulation 902 KAR 20:008
Maine	IP	Chapter 1683
Maine	AS	Chapter 1683
Maine	ED	Chapter 1683
Maine	AC	Chapter 1683
Maryland	IP	COMAR 10.37.06
Maryland	AS	COMAR10.37.07
Maryland	ED	COMAR 10.37.04
Maryland	AC	COMAR 10.37.04
Massachusetts	IP	114.1 CMR 17.00 Requirement for the Submission of Hospital Case Mix and Charge Data Administration Bulletin 02-06
Michigan	IP	NO STATUTUE. Data Acquisition and Access Policies and Procedures Michigan Health & Hospital Association Access and Oversight Committee (Revised March 1996)
Missouri	IP	Rules of DOH 19 CSR 10-33.010
Missouri	AS	Rules of DOH 19 CSR 10-33.010
Missouri	ED	Rules of DOH 19 CSR 10-33.010
New Hampshire	IP	RSA 126:25 (Health Care Data)
New Hampshire	AS	RSA 126:25 (Health Care Data)
New Hampshire	ED	RSA 126:25 (Health Care Data)
New Jersey	IP	N.J.A.C. 8:31B-2
New Mexico	IP	Health Information Systems Act, Section 24-14A-1 et seq. NMSA 1978
New York	IP	Title 10 Section 400.18
New York	AS	Title 10 Section 755.10
North Carolina	IP	Article II of Chapter 131E OF North Carolina General Statutes

North Carolina	AS	Article II of Chapter 131E of North Carolina General Statutes
North Dakota		Chapter 23-01.1: Health Care Data Committee
Oklahoma	IP	OK Hlth Care Info Sys. Act. Title 63 Sect. 1-701
Oregon	IP	ORS 442.120
Oregon	AS	ORS 442.120
Pennsylvania	IP	Chapter 911 Data Submission & Collection
Pennsylvania	AS	Chapter 911 Data Submission & Collection
South Carolina	IP	Code of Laws of South Carolina 1976, section 44-6-170
South Carolina	AS	Code of Laws of South Carolina 1976, section 44-6-170
South Carolina	ED	Code of Laws of South Carolina 1976, section 44-6-170
Tennessee	IP	Chapter 1200-7-3 (Regs); TCA 68-1-108 (Law)
Tennessee	AS	Chapter 1200-7-3 (Regs); TCA 68-1-108 (Law)
Tennessee	ED	Chapter 1200-7-3 (Regs); TCA 68-1-108 (Law)
Texas	IP	Chapter 108. Texas Health and Safety Code
Utah	IP	R428-10
Utah	AS	R428-11
Utah	ED	R426-1-7(I)
Vermont	IP	18 USA Section 9410, 9453, 9454.
Vermont	AS	18 USA Section 9410, 9453, 9454.
Vermont	ED	18 USA Section 9410, 9453, 9454.
Virginia	IP	House Bill 1307, Chapter 7.2 Health Care Data Reporting
Virginia	AS	House Bill 2763 Chapter 0341
Washington	IP	Washington Administrative Code (WAC) 246-455
West Virginia	IP	WV 65 CSR13 4.3.2
Wisconsin	IP	Chapter 153, Wisconsin Stats. And Chapter HFS 120, Wisconsin Administrative Code
Wisconsin	AS	Chapter 153, Wisconsin Stats. And Chapter HFS 120, Wisconsin Administrative Code

BOLD = Leaders in public release of hospital level data