

Health Data Systems at a Crossroads: Lessons Learned From 25 Years of Hospital Discharge Data Reporting Programs

Funded by the Agency for Healthcare Research and Quality

December 2012

Denise Love, R.N. MBA
Barbara Rudolph, PhD

THE NATIONAL ASSOCIATION OF
HEALTH DATA ORGANIZATIONS



Health Data Systems at a Crossroads: Lessons Learned From 25 Years of Hospital Discharge Data Reporting Programs

Background

The health care industry is undergoing massive legal and structural transformation. States, the federal government, industry, and consumers need comprehensive information on disease incidence, treatment costs, and health outcomes more than ever, but the information is not readily available. The “American Recovery and Reinvestment Act of 2009” (ARRA) provides for massive investments into health information technology. The Health Information Technology for Economic and Clinical Health Act (HITECH) provisions will make available incentive payments totaling up to \$27 billion over 10 years to achieve liftoff of a national system of EHR’s.¹ As of the end of August \$6.9 billion in Medicare and Medicaid incentive payments have been made to providers and hospitals to adopt the “Meaningful Use Stage 1” of certified health IT products.² While this investment in clinical information systems is important, it would be wise to also invest in the established health data systems to take advantage of the new data opportunities that will emerge with EMRs and HIEs. This investment could leverage the existing infrastructure which includes to collect and aggregate data, staff resources, and expertise on how best to report publicly on results.

As states continue to face recession-based budget deficits, growing numbers of uninsured, and are working to reform their health care systems, legislators and others are seeking information about populations and markets to inform payment and reform policies. EMRs are not yet “shovel ready” for cross-system information exchanges ; thus, data systems established over the course of decades will continue to supply much of the information for market and policy decisions during the transition to clinically-based information systems.

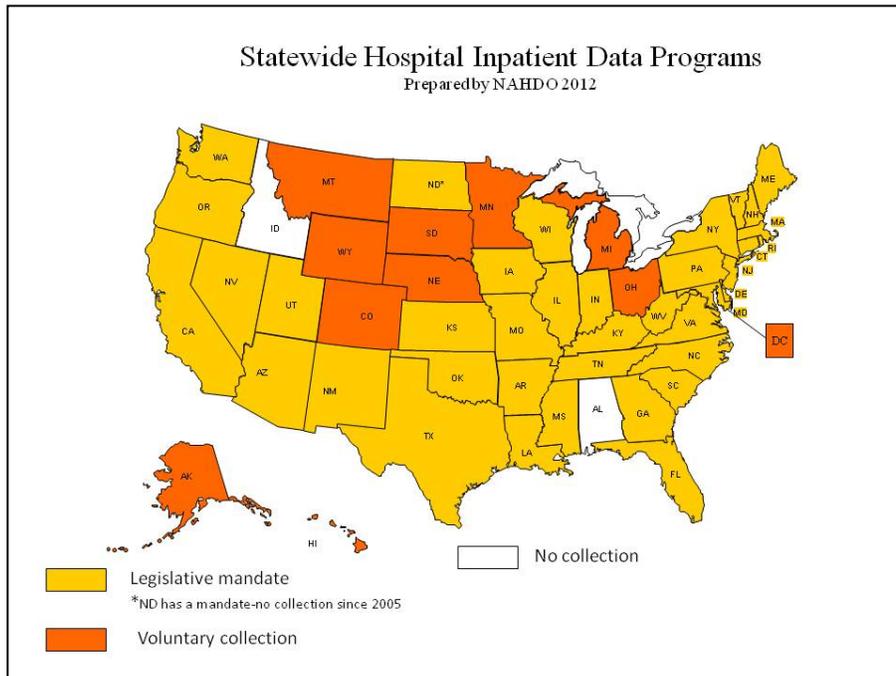
In the context of the EMR, it is important to understand how existing data systems differ from newly emerging systems, and how existing systems were able to overcome many of the challenges facing those who will be engaged in EMR aggregation efforts. This paper, written by the National Association of Health Data Organizations (NAHDO) for the Agency for Healthcare Research and Quality (AHRQ), will provide an overview of statewide hospital discharge data systems, and the implementation lessons learned over the past 25 years, as well as a call for hybrid data systems combining EMR data with discharge data.

Overview of hospital discharge data reporting

¹ Blumenthal, D., and Tavenner, M. “The Meaningful Use Regulation for Electronic Health Records.” *New England Journal of Medicine*, August 5th, 2010.

² *Government Health IT*, 9/6.

Hospital discharge data are a summary of patient and provider information from a hospital stay. The hospital discharge record includes information on: patient demographics; payer type; charge for the care delivered; coded patient diagnoses; length of stay; procedures performed during the stay; admission source; discharge status; physician providers; and hospital identifiers. Hospital discharge data sets are collected by public or private health data organizations; the health data organization serves as the data aggregator or data steward. With national standards for format and data specifications maintained by the National Uniform Billing Committee (NUBC), the data collected by the health data organizations are fairly uniform in format and data specifications across states and providers. In 2012 48 states and the District of Columbia collect inpatient hospital discharge data records for each patient discharged from an acute care hospital licensed in that state or jurisdiction. Figure 1 below shows state collection practices.



Impact/value of widely available health care/hospital data sets

Since the advent of statewide hospital discharge data reporting programs in the 1970's, a wide variety of uses for the data have arisen, including: program and policy evaluations using data for benchmarking and impact studies; market share analyses; provider performance studies on quality and efficiency of care; utilization studies for hospital service planning; motor vehicle crash outcomes studies, environmental tracking, and a variety of other uses of the data. In addition, entire allied professional areas of study have developed, e.g., the research and development of risk adjustment systems, disease groupers, and efficiency software, for use with discharge data.

What is lost without comparative, statewide health care data

Without the hospital data currently supplied by states, it is hard to imagine a reliable source for understanding state trends in hospital quality and cost. Statewide databases are the principal source for conducting trend analysis on quality of care received and the concomitant cost of care received. The statewide data are also aggregated to examine similar trends at the federal level, through the Healthcare Cost and Utilization Program (HCUP), maintained by AHRQ. Over forty states contribute data for health services research and policy information. In addition, an entire system of measuring and understanding quality and patient safety has been formalized in the AHRQ Quality Indicators and the Patient Safety Indicators. Many of these indicators have now been endorsed through the National Quality Forum for public reporting. These indicators provide specific hospital comparative information on outcomes and processes of care, which can be examined year-to-year.

The release to the public of quality of care results from comparative, statewide health care performance data, started slowly when New York, Pennsylvania and California began to push out public reports on cardiac mortality in hospitals. These early and limited transparency initiatives shook the complacency of consumers, providers and purchasers of care, around the quality of surgical care. The information provided by these reports fueled other public reporting efforts in the public and private sectors and continues to be extended as purchasers of care demand greater information. While other data sources are being used for some very specific purposes, hospital discharge data still provides the largest proportion of public information available to consumers for making informed choices about providers.. New analytic techniques for risk adjustment, grouping, and linkage to other databases have kept this data useful as a primary source of healthcare information.

The public health community also relies on hospital discharge data for a wide array of uses—from environmental tracking, epidemiologic and cost information for specific tracking of traumatic brain injury, cancer, and other conditions, identifying prevalence and hospital costs associated with influenza and other outbreaks, and for basic denominators in a wide variety of conditions like hospital acquired infections. While condition-specific registries have provided the numerators and significant specificity (on tumors and other conditions) for public health work, the discharge databases have been the main source for determining the denominators and cost of care in large populations.

The health services research community has focused on geographic variation in care delivery, quality of care disparities, cost, efficiency and comparative effectiveness research, also rely heavily on hospital discharge data. In 2009, AHRQ reported over 200 articles published in research journals using data from HCUP. In addition, states which also release discharge data directly to researchers, such as California, New York, Florida, and others, are likely to have at least that many research users. Frequently this data is linked to death certificate data, birth data, motor vehicle crash data, hospital descriptive information, and condition-specific registries.

A common community data set

Today, discharge data provides the full community of users with information that is relatively current, has provider identifiers, and is cost efficient. The information is used in: public displays, such as websites, dynamic web query systems, and in traditional reports. It provides a broad array of information not found in individual registries and is more cost efficient to collect than other sources. It can also be de-identified to allow broader use, than is possible with other clinical data sources.

Because they are widely available and broadly used, hospital discharge data could serve as the backbone for a hybrid EMR/discharge “package” of information. Statewide discharge data combined with clinical data in an EMR can supply both the numerator and denominator for examining outcomes of care and cost effectiveness of treatments. Without discharge data, it is difficult to attain denominators across institutions, which are critical to any analysis of care delivery. In addition, the common structure and relative uniformity of hospital discharge data across providers and states, allows for regional and national comparisons.

Because market competition is central to the U.S. health care delivery system, the healthcare industry will continue to rely on hospital and other administrative data sets to compare utilization patterns and patient migration for care. In addition, many of the new Accountable Care Organizations (ACOs) are being founded by hospitals in combination with physician practices, and these ACO’s will need to incorporate discharge data with the clinical record for assessing the efficiency and quality of care delivery.

Discharge databases: A creative response to political barriers

Data collection and aggregation across providers is fraught with political and technical challenges. Hospital discharge data bases are no exception, but what is notable is that the agencies implementing these systems had to become adept at working with their stakeholders, especially the provider community, to overcome resistance to reporting.

Health data agencies have established a roadmap for putting together complex data reporting initiatives that could inform other initiatives, including an EMR reporting system. The long history of collection, cleaning and use of hospital discharge data for secondary use provides a number of lessons for EMR clinical data aggregation, validation and use. Many of the same political and technical challenges faced in the collection and use of discharge data shall arise with EMR data when it is aggregated and adopted for *secondary use* in quality reporting. The lessons learned in discharge data system development and management could help provide solutions to the most difficult issues faced by EMRs (Table 1).

Table 1. Lessons learned that might be relevant to the EMR/clinical data aggregation

Common Challenges	Proven Solutions
Provider reporting burden to respond to an unfunded mandate	Adoption of industry standards minimizes the provider reporting burden. Hospitals reporting to states rely on standardized code sets and data formats such as the UB-04 and the HIPAA Electronic Transactions.
Provider resistance to initial aggregation of data	To stimulate market decisions, compliance to reporting requirements in lieu of rate-setting was the strategy in some states. All state health data organizations work to build value to providers as a major consumer of the data sets. Aggregators of clinical information should also consider the clinician as a primary customer.
Patient privacy and confidentiality concerns	Health data agencies have adopted encryption, cell suppression, field aggregation and statistical techniques to de-identify the data. Carefully-crafted release policies and practices have resulted in no known breach of hospital data despite 40 years of releases. Unfortunately, this is not true of the provider community. Those who will aggregate large volumes of hybrid information will need to assure security of the data. Data agencies can assist aggregators in assuring a secure environment.
Responding to and overseeing numerous requests for data with limited staff	Standard public use files (de-identified); research files with IRB-approval, and web-query systems meet most frequently-requested statistics/reports. Clinical data aggregators will need sophisticated tools and processes to respond to data requesters.
Limited funding for the data system and agency operations	Strategic pricing of data sets and products to generate revenue and control access are key to maintenance of data systems. Several states have mandated industry (provider or payer) fee assessments, based on provider and/or payer market share, to fund their data system. While new systems will have initial federal support, continuing into the future will require new funds.
Chronic under-funding for hospital discharge data implementation and operations	State-to-state transfer of best practices, tools, knowledge through the NAHDO network and the HCUP Partners network. Clinical aggregators and reporters will need to work together to share best practices in order to assure survival.
State-specific or local codes	NAHDO and states have actively worked with standards organizations to incorporate state reporting needs into national standards. Present on Admission and race/ethnicity are examples of this work. The new ACOs and data aggregators will need to develop expertise in standard setting to assure that data are comparable across regions and states.*
Overcoming data limitations	Because the discharge data bases are based on billing standards, they lack clinical data such as laboratory results. Coding practices vary across providers, and the data capture facility-based care, not outpatient or ambulatory care. Health data agencies have adapted to these limitations. A growing number of states are enhancing their administrative data with clinical data elements, such as Present on Admission indicator for each inpatient diagnosis and linking the data to laboratory and other public health registries. Consistent feedback of the data to hospitals and broad use of the data by various users has shown to

	improve coding practices by the hospitals over time. EMR systems will also have data limitations and will require users to improve existing structures.
Timeliness and availability of information	Agencies maintaining hospital discharge data systems operate under different legal and political environments, which influence data availability. Quarterly data reporting and lengthy review and validation periods affect the timeliness of data release by the agency. The complexity of aggregating EMR data will also slow down the timeliness of information and the availability of reporting—these problems create tension between custodians of data and end-users. Establishing accepted electronic editing and validation processes will address timeliness; policies for release must be clear and have acceptable timeframes.
Lack of national and regional benchmarks and analytic tools	States established the vision of multi-state aggregation of hospital data which later became the Healthcare Cost and Utilization Project (HCUP) managed by AHRQ. States contribute data and receive technical support, tools, and benchmarks in return. The new EMRs/HIEs will also need regional and national benchmarks and would benefit from a similar structure.
Resistance to public reporting, especially comparative quality reports created demand for advanced data tools	In an effort to overcome resistance and instill fairness, state agencies have been leaders in pioneering risk adjustment methods and enhancing their hospital data sets with clinical information. Further work will be needed as databases are combined, existing tools won't work without revisions. Building or re-building these tools is expensive, and coming to agreement across data integrators on the most appropriate tools is time-consuming.
Data ownership and control issues arise when combining data across stewards	Establishing clear rules for data integration and continued data maintenance of combined systems requires legal and political negotiation. Lack of trust can kill data integration. Legal restrictions on the sale of EMR data may limit release.

*Mayo Clinic is engaged in an ONC-funded research project examining the very early standardization of EMR data; the project is titled Mayo Clinic's Strategic Healthcare IT Advanced Research Project (SHARP). Key elements are standardizing clinical data normalization, (i.e., agreeing how diabetes is defined in the EHR); merging and standardizing patient data from non-electronic forms with the EHR; identifying patient phenotypes (absorption and processing of nutrients); creating efficient methods for doing the first three; and identifying incorrect or inconsistent data.

Implications for EMRs

EMRs will need to address these similar issues, but they are faced with significantly increased complexity. State hospital discharge systems may have as few as 40 variables to collect and manage—while EMR's could have 10 or more times that number of variables. The number of items alone will create significant technical challenges—combine that with clinical knowledge and measurement on hundreds of areas of care—it becomes extremely complex. Today there are still many areas of medical care where there is strong disagreement on diagnosis, labeling, and appropriate treatment—coming to consensus on these areas for standardization will be difficult and costly. Can agreement be reached and complexity reduced enough to create at least a standard minimum data set? EMRs have very limited standards for structure, and the information collected may be too broad. There is still a significant amount of information in free text fields—yet forcing limited choices for documentation may harm treatment. Diagnosing and treating clinical conditions requires numerous types of inputs—laboratory, radiology, observations, paper inventories, surveys, etc—all of this information must become part of the EMR—providing valuable data for clinicians. Yet, turning all of this information into a single database to

be used for quality measures, payment, and decision-support requires standardization. The standard setting bodies are working on various areas—but still have a significant way to go before EMR interoperability for performance measurement is standardized and implemented. The rules and standards will need to undergo rigorous and potentially contentious review by those who stand to lose market share with standardization in this now lucrative industry.

Historically, physician culture and environments of care have been based on “rugged individualism”, a guild system, and more recently on corporate (for profit) philosophy where creating value/profits have dominated over public and societal good. Will it be possible for physician culture to change toward a more managed and monitored approach to care giving? Or will physicians resist external monitoring, payment reforms, and social medicine? State hospital discharge systems have been based on differing social constructs—the public good through public health, transparency, and collaborative efforts to improve data quality and public reporting.

At present, the established EMR systems have only been used on a very limited basis to provide quality reporting. Many of today’s endorsed quality measures cannot be populated from stand-alone automated systems for a variety of reasons. At present, most systems would require a new defined “report” for each quality measure—a report would require programming, and after programming, a run of the system to get the necessary data. The lack of standardization of EMR systems would require additional work to meet the definitions within most performance measures. This multi-step approach will be costly, and is probably avoidable if systems were required to have specific data elements with standard definitions and formats. The pulling of information out of these systems was designed for clinical care, not for performance measurement. Performance measures require consistent measurement across providers, conditions, and settings.

At present, the National Quality Forum has built a “framework for standardizing a set of quality measures for EMRs”—called the Quality Data Set (QDS), and has developed a minimal set of measures that could be employed by EHRs.³ And, these measures have been approved by the Healthcare Information Technology Standards Panel (HITSP) and put into updates of the Quality Interoperability Specifications. However, implementation and use across entities is still far on the horizon. There are still many other components which need to be put in place for this to work, including agreement on risk-adjustment systems, modification of EHRs to collect many non-clinical indicators necessary for risk adjustment, and systems for acquiring denominators across entities. In Stage 1 of Meaningful Use, one type of communication of structured or unstructured information (e.g., diagnostic test results, problem list, medication list, or medication allergy list) was to be exchanged between two EHR’s, but they were allowed to fail and still get credit. While standards for operability will be increased in future stages, time periods for implementation are often subject to industry changes.

Still to be determined is whether there will be an “end-product” of the aggregated EMR data—will physicians approve use beyond that mandated by CMS? That is, will there be public use files—as there

³ National Quality Forum (NQF). *Health Information Technology Automation of Quality Measurement: Quality Data Set and Data Flow*. Washington, DC: NQF; 2009.

is with hospital discharge data? Will the aggregated data be available to purchasers, researchers, policy makers? State discharge systems are made public under public law—will there be intervention by states to assure the public good?

A vision for the future: Hybrid data

Hospital discharge data is not your mother's administrative data anymore—several key data element enhancements have already been made to discharge data, and states are engaged in pilot testing additional clinical data enhancements to discharge data systems. More than 25 states have added Present on Admission to their state discharge systems--this element has allowed states to report on conditions, such as infections, that are hospital-acquired. Anticipated enhancements include the addition of laboratory data to discharge systems, along with additions from public health registries (cancer, vital records, PHIN, NHSN) and data which have the necessary backbone allowing data linkage with administrative/financial data. A number of states are now pilot testing through AHRQ funding, the addition of key laboratory findings. AHRQ funded research has shown these data elements are an effective and efficient approach for enriching administrative systems with clinical information.^{4 5} The CDC is also funding a project to link hospital laboratories with public health agencies so that the labs can electronically transmit reportable test results. Electronic laboratory data can be used in conjunction with hospital data to conduct disease surveillance electronically to meet meaningful use requirements.⁶

This new enriched data source will be further enhanced by ICD-10 coding (adding much greater detail to diagnosis coding); it is scheduled by CMS to be in place by October 1, 2014. It is anticipated that EMR data systems in the future will also be able to provide standardized data for certain common chronic clinical conditions. It will be critical to link this information with the administrative data to assess efficiency of care. Over time it is expected that a hybrid system, combining other key clinical data elements with the administrative data, will be developed and will provide purchasers, payers, public health and consumers with much greater information than currently exists. This type of hybrid data source will broadly encompass all stakeholders—rather than EMR or registry data which is principally focused on clinicians.

In order for hybrid systems to be developed, there are a number of actions needed:

- Concerted federal-state-industry education effort for consumers and clinicians about the value of health information, the protections in place for sensitive data and the “common good.”

⁴ “Adding Clinical Data Elements to Administrative Data for Hospital-Level Reporting.” Final Report AHRQ Contract #233-02-0088, Task Order 13, Vol. 1, July 3, 2006

⁵ M.Pine, B. Jones, Y-B. Lou “Laboratory values improve predictions of hospital mortality.” *Int J Qual Health Care* (1998) 10 (6): 491-501. doi: 10.1093/intqhc/10.6.491

⁶ M.Mosquera, CDC award boosts hospital labs' ability to report result. *Government Health IT*, Monday, February 21, 2011. Accessed online 2/22/2011 at <http://www.govhealthit.com/newsitem.aspx?nid=76365>

- Following the educational effort, a strategy for development of patient identifiers, or a strategy for more standardized and widely available linkage methods and tools must be developed.
- Collaboration among stewards of various data systems is also critical, and can be accomplished through federal and state initiatives and pilots.
- Data standards committees should focus on and consider key clinical elements for standardization. Many specialized features are part of the approved HL7 EMR functional specification yet they are not currently *required* for certification by the Certification Commission for Health Information Technology (CCHIT). This means that as providers are incented to purchase *certified EMR* systems they will find that those systems do *not* include the features needed to support key public health functions, including registries, while a developed hybrid data system could.
- Some funding for these strategic efforts and for development and implementation within state discharge systems is also needed. Potentially, some incentives might be provided to states to add standardized clinical data elements to their discharge abstracts.

Conclusions

Despite chronic underfunding, hospital discharge data reporting programs have survived over the years by adopting practical solutions to complex challenges. Today they generate much of the health care data in use by health services researchers, public health, and industry for comparative studies, surveillance, and quality improvement. NAHDO suggests that policy makers assure that hospital discharge data systems are not swept away by proprietary clinical data systems, but instead hospital discharge systems should be the backbone for integrated healthcare information in hybrid systems in the community domain.