Understanding Publicly Available Healthcare Data (2016) - Retired

Save to myBoK

This update supersedes the September 2013 practice brief, Understanding Publicly Available Healthcare Data, which replaced the January 2010 practice brief with the same title.

Healthcare organizations increasingly collect and utilize data for internal and external reporting on patient safety, quality measurement, patient satisfaction initiatives, reimbursement, and general business functions. Federal regulations such as the Health Insurance Portability and Accountability Act (HIPAA), the Health Information Technology for Economic and Clinical Health (HITECH) Act, including the EHR incentive program, and the Patient Protection and Affordable Care Act (ACA) require quality healthcare data. As a result, more data are publicly available to healthcare organizations, government and business entities, and individual healthcare consumers for benchmarking and analysis. Members of the public have an increased awareness and interest in reliable data regarding patient safety and quality of care, the incidence of disease outbreaks (e.g. Ebola, Zika viruses), as well as syndromic surveillance (e.g. diabetes, obesity, Alzheimer's).

A major source of public data is claims data, which has a variety of secondary uses. Analysis of claims data can produce information about:

- Population health and access to care
- Disease trending
- State and national health care expenditures and costs
- Drug utilization and health care coverage by payer type

Present on admission (POA) values, readmission rates, and hospital-acquired conditions (HAC) affect reimbursement as well as drive patient safety reporting and the Affordable Care Act (ACA) risk-adjustment methodology. Claims data support benchmark comparisons among facilities, quality of care reporting, and risk analysis. Healthcare providers and payers monitor the accuracy of billing practices to detect lost revenue opportunities and billing noncompliance. The use of data is a key resource for strategic planning, measuring the healthcare environment and providing business intelligence. Healthcare organizations must ensure they gather and report quality, accurate healthcare data to derive confidence and value from its reuse.

This practice brief provides background on the different types of publicly available healthcare data and the factors that determine if the data serve to meet the user's needs.

Access Requirements for Publicly Reported Data

Publicly reported data are typically aggregated, focusing on healthcare services of facilities and providers, and do not contain protected health information (PHI) or personally identifiable information (PII). PHI or PII data users must meet the Health Insurance Portability and Accountability Act (HIPAA) privacy requirements for access to and release of the data.

HIPAA defines PHI as information that "is created or received by a health care provider, health plan, public health authority, employer, life insurer, school or university, or health care clearinghouse" and "relates to the past, present, or future physical or mental health or condition of an individual; the provision of health care to an individual; or the past, present, or future payment for the provision of health care to an individual."[1]

PII, on the other hand, refers to information that an individual can use to uniquely identify, contact, or locate a single person when used alone or in combination with other sources. PII is not necessarily restricted to health information.

The implementation of health IT systems has created an expansion in the amount and complexity of data produced by healthcare services. The format and location of the data vary and range from state and national health databases to facility-based health and financial information systems. As a result, the amount and types of publicly available healthcare data continue to grow as more healthcare organizations transition to electronic health records and the healthcare industry pushes for transparency.

When evaluating available data for analysis, it is important to remember that each data source may have special requirements for access and use. These requirements may include privacy release approvals, pre-arranged financial agreements, or source-specific software.

In some cases, an organization may require specific software to download and analyze its data. The organization housing the data generally makes the source-specific software available. The Centers for Medicare and Medicaid Services (CMS) has a search tool called CMS Data Navigator that allows access to a wide range of Medicare and Medicaid data such as disproportionate share hospital rates, vital statistics data, and health outcomes survey data. A list of CMS publicly available files (subject to privacy release approvals and computing resources) is available at https://www.cms.gov/Research-Statistics-Data-and-Systems/Files-for-Order/FilesForOrderGenInfo/index.html. The CMS easy-to-use filtering tools provide an interactive analytical experience utilizing the available CMS site data sets. Data users may access CMS data sets at Data.CMS.gov.

Types of Publicly Reported Data

Publicly reported healthcare data vary from quality and patient safety measurement data to patient satisfaction results (e.g., Hospital Consumer Assessment of Healthcare Providers and Systems, prevention of medication errors, etc.). The aggregated data range from local to national perspectives, such as state-specific public health conditions to national morbidity and mortality statistics.

Agencies use data visualizations such as star ratings on different quality measure reports or to enable a healthcare provider to benchmark its own reports against other providers. Current publicly reported data typically fall into the following general categories:

- Quality or performance measure data
- Utilization data
- Clinical condition data

Quality or Performance Measure Data

Demonstrating quality healthcare remains an industry focus, and as such, the types of quality and performance measures data available to the public and providers continue to grow.

Both public and private organizations collect and provide quality and performance measures data to consumers. Facility and provider participation in these initiatives is voluntary unless mandated by legislation. Organizations may choose to participate in voluntary initiatives for reasons such as financial incentives or public perception. Each facility or provider reports the data to a third party, which presents the data in its own way, typically known as a report card.

The following list provides examples of some of the most common quality and performance measure organizations.

The Healthcare Effectiveness and Data Information Set (HEDIS) are measures used for reporting on insurance companies.[2] The HEDIS measures are some of the most long-standing, well-vetted measures available. The National Quality Forum, tasked with approving quality and performance measures, endorsed the HEDIS measures. The measures began in the 1980s with Health Maintenance Organizations (HMOs), and have grown to where over 90 percent of plans voluntarily gather and report data. The National Committee for Quality Assurance (NCQA) credentials an independent third party, which then certifies the outcome before reporting the results.

The Leapfrog Group was started by a group of large employers who wanted to influence the quality and affordability of healthcare based on their healthcare service purchases.[3] The organization provides a consumer and a hospital view of its data, allowing for ease in comparing and selecting healthcare providers. Participation in reporting information is voluntary.

HealthGrades is an independent organization that offers consumers ratings of physicians, hospitals, and nursing homes based on data from the most current three-year data set available from CMS, several individual states (where available), publicly available directories, and telephone surveys. The information is free. However, HealthGrades makes more detailed reports available for purchase. HealthGrades also offers educational information to providers on how it determines the ratings.

The Joint Commission is one of several healthcare accrediting organizations in the United States. This independent, not-for-profit organization accredits and certifies healthcare organizations on the quality of care they provide to individuals. The Joint Commission sets benchmarks utilizing information from other Joint Commission accredited facilities.[4] Accreditation is voluntary. Another accrediting body is DNV GL Healthcare, which integrates CMS Conditions of Participation requirements with the ISO 9001 Quality Management Program. It provides certification and accreditation information to patients and healthcare providers.[5] State licensing requires inspections of hospital facilities and provides similar data. The results of these quality surveys provide valuable data for comparison of facilities.

Hospital Compare is a CMS-sponsored site that provides consumers with information on hospital performance and recommended patient treatments. The site offers Medicare beneficiaries' perceptions of the care they received at a facility. It also provides information on the utilization of facilities regarding the median Medicare payment and number of Medicare patients treated for specific conditions. Similar comparisons are available for nursing homes, home health, and Medicare Part D drug plans with a simplified comparison using a star rating system.

The Agency for Healthcare Research and Quality (AHRQ) provides a variety of data and information to healthcare professionals, policymakers and consumers. Its stated mission is "to produce evidence to make healthcare safer, higher quality, more accessible, equitable and affordable, and to work within the U.S. Department of Health and Human Services and with other partners to make sure that the evidence is understood and used."[6] AHRQ sponsors the Healthcare Cost and Utilization Project (HCUP), which provides access to a family of databases containing both public and private hospital care data. However, accessing the data set requires an agreement that outlines limitations and provisions of data use. Organizations must also cite AHRQ if the data are used in reports.

Healthcare Utilization Data

Public performance and quality measurement organizations typically report on the use of healthcare services through patient volume and fees for selected diagnoses and procedures, which are presented in. AHRQ's HCUP databases. HCUP also contains source-specific software tools to facilitate use of the data. The databases combine state, federal, and industry data, which researchers can use for data on broad healthcare topics.

AHRQ also supports the Medical Expenditure Panel Survey, which provides data on the cost and use of healthcare services and health insurance across the U.S. Two major components of the data are household, which focuses on individuals and their providers, and insurance. Healthcare organizations can use the data for private planning. Policy makers can use the data to understand the nation's healthcare needs.

CMS offers numerous data files regarding the cost of healthcare services. While the titles vary, many of the data elements are claims data from different sources.

The Medicare Provider Analysis and Review (MedPAR) file contains claims data for acute care inpatient and skilled nursing facilities; the Hospital Outpatient Prospective Payment System (HOPPS) contains claims data from hospital outpatient encounters.

State-specific information may also be available, depending on the individual state. A good resource for finding state specific information is the state department of public Hhealth. For example, in California, the Office of Statewide Health Planning and Development (OSHPD) provides the state with an enhanced understanding of the structure and function of its healthcare delivery systems. Various healthcare data originating from a multitude of different types of settings are available in an effort to publish valuable information about healthcare outcomes and promote quality care and healthcare accessibility. OSHPD also provides public access to the "Healthcare Atlas," which is a powerful interactive geographic information system (GIS) tool. Healthcare Atlas utilizes "hospital and other healthcare facility data including facility locations, financial and utilization reports, and patient level summaries". Furthermore, users of Healthcare Atlas can create "patient origin and market share maps" utilizing data from over 68 million patient records.[7] The Statewide Planning and Research Cooperative System (SPARCS) data for the state of New York is another source for open de-identified healthcare data sets for download or for online query.

Clinical Condition Data

Although the focus of publicly reported data may be utilization or quality specific, clinical conditions may be included in the data, such as quality measures reported for myocardial infarction. In other instances, the clinical condition may be the primary focus of the data, as with many state public health sites and the Centers for Disease Control and Prevention (CDC).

The CDC has numerous programs that record disease statistics, management, treatment, and control. CDC data have attributes that allow the user to aggregate by specific populations, such as workplace, disease, or healthcare setting. The CDC routinely works with state public health agencies to implement state-based registries such as the Paul Coverdell National Acute Stroke Registry to measure, track, and improve the quality of care. Table 1, "Examples of Publicly Reported Data," provides an additional overview of common publicly available data. Some data are available in downloadable files, allowing organizations to analyze and compare them against their own data. Other data may already be analyzed and displayed in percentages, reports, or graphic displays; or, data may be sorted and filtered for analysis using the tools provided by the website.

Analytic Considerations

Organizations must take into account the following fundamental considerations when analyzing internally held or publicly available data to ensure the quality of the information they share.

The context of the data. It is important to consider who collected the data and for what purpose. Take for example a pharmaceutical company that conducts a study on the side effects of a drug it manufactures. How reliable are the data? If a specialty society or academic researcher with no ties to the pharmaceutical industry conducted the study, would the results differ?

Organizations should consider using internal benchmarking reports to compare diagnoses, procedures, complications, length of stay, or other measures for conducting accuracy analysis of their own data. If an organization only uses HCUP data (discussed earlier), results would represent only theinpatient Medicare population. Therefore, validating the data source to ensure it reflects the population desired is imperative to ensuring quality information.

The method used to gather the data. Public health organizations collect data in two ways. In the first method, agencies gather individual health-related events or services, often by geographic region, over a specific period, for a specific health-related procedure or diagnosis, or via some other parameter. This method is limited to those who seek services and may not accurately represent the entire population. The second method involves using sampling techniques to gather survey data from a defined subset of the population. Results obtained using this method may contain bias due to different methodologies, data collection periods, data dictionary definitions, or tabulation methods. Inconsistent measurements, calculations, and conditional rules can also introduce unreliable results.

Whether the data were peer reviewed. Organizations should always ask if data were peer reviewed and if not, why not. To be peer reviewed, a group of qualified experts in a given field must have performed an impartial review of the data.

During the review process, these experts evaluate factors that might negatively affect the reliability of the data. Reviewers consider whether the data are reproducible, whether the sample size is statistically adequate, whether the organization collected the data over a clearly defined time interval, and so forth.

The scope of the data must match the need. It is important to ensure that the data can correctly address the needs of the project or study. Important issues include the completeness of the data, design, confidentiality, availability, accessibility, usability, cost and the size of the database.

The Future of Publicly Available Data

HIM professionals are facing many challenges, including the extraordinarily extensive volume of data collected, and the way organizations collect, report, store, and manage healthcare data to ensure data integrity and availability. HIM professionals must be cognizant of data standards and definitions when analyzing data. The implementation of ICD-10-CM/PCS and the movement toward widespread health IT use create the potential for even more data collection and reporting. The adoption of the ICD-10 classification system has increased the granularity of data collected and the inclusion of data from clinical sources. The widespread adoption of electronic health records systems presents a bounty of data that has previously been underutilized or not easily obtained. Consumer health data from personal health monitoring devices presents another source of data adding to the mix. The harnessing of "Big Data" accumulated from many sources presents opportunities for data-driven knowledge discovery.

Consumers are becoming more actively involved in their healthcare, looking to the Internet for guidance on providers or facilities of choice. In PricewaterhouseCoopers' Health Research Institute 2016 release of its annual forecast of top health care trends, 73% of consumers surveyed would be willing to share their healthcare data to aid in the diagnosing and treatment of others. Also 83% of consumers surveyed would be willing to share their healthcare data to aid in diagnosing and treating themselves.[8] This consumer support welcomes new opportunities for publicly reported data and health data exchange. Additionally, payers continue to call for performance-based results on which to base reimbursement and utilization decisions, as they do not want to pay for HACs or unnecessary readmissions. The ability to share information across systems and healthcare entities will continue to push the need for clinical data standards.

The Quality Reporting Document Architecture (QRDA) is one initiative developing an electronic data standard for exchanging patient-level quality measurement information between health information systems. This enables the use of existing standards for clinical data exchange to report current quality measures.[9] This initiative is one example of how improved data collection and reporting will lead to improved quality of care.

Public reporting can also improve quality initiatives such as value-based purchasing (VBP), further helping Medicare beneficiaries choose distinct and merited healthcare services. Already a key component, VBP programs could enhance reporting to:

- Address the needs of multiple stakeholder audiences
- Employ display methods or decision supports that facilitate fair and accurate decision making
- Ensure consumer understanding of performance data displays[10]

Data that can inform consumer decision making could be part of the VBP program and publicly reported on the CMS Hospital Compare Web site.

The introduction of Accountable Care Organizations (ACOs) affects the future of publicly-available data. An ACO is a group of providers of healthcare services and suppliers who form a legal entity to be accountable for the quality, cost, and overall care of a population of patients. According to CMS, ACOs will provide information about the quality of care their providers deliver on more than 60 different clinical quality measures.[11]

Despite the efforts mentioned above, some consumers fear the increase in data collection and reporting will adversely affect their health or life insurance policies. Some consumers question whether insurance companies will deny coverage based on available patient data.

Changes in clinical classification systems and new technology will allow the industry to capture, report, share, analyze, and access more detailed patient health information, which in turn will improve the quality of care. This increased granularity in data, coupled with the ability to combine the varied sources and components of health data, yields a bounty of data sets for organizations and consumers to explore using the power of computer-based analysis tools and techniques. AHIMA's Information Governance Principles for Healthcare (IGPHC) ™ is an important factor for all organizations to consider, as organizations must be committed to ethical and responsible handling of healthcare information.[12] Organizations should implement information governance practices that include assigning ownership of publicly-reported data. With a continued focus on value-based care and pay for performance (P4P) initiatives, the volume and diversity of healthcare data becomes a resource for value analysis and outcomes. The reuse of large stores of available data fuels a multitude of healthcare initiatives and purposes ranging from monitoring patterns of care, patient safety, cost trends, gaps in access to care, clinical research, business intelligence, and more. Table 1, provides examples of publicly reported data. It is not an all-inclusive list.

Table 1: Examples of Publicly Reported Data

Organization	**Examples of Available Data**
Agency for Healthcare Research and Quality (AHRQ)	Medical Expenditure Panel Survey Consumer Assessment of Healthcare Providers and Systems Healthcare Cost and Utilization Project State Ambulatory Surgery Database (34 states) Healthcare Cost and Utilization Project State Emergency Department Database (32 states) Healthcare Cost and Utilization Project State Inpatient Database (48 states)
Centers for Disease Control and Prevention (CDC)	United States Cancer Statistics National Center for Health Statistics (NCHS) data files CDC Wonder Online Databases Behavioral Risk Factor Surveillance System (BRFSS) Survey data files
Centers for Medicare and Medicaid Services (CMS)	Medicaid Statistical Information Files Medicaid Analytic eXtract Files Hospital Compare Nursing Home Compare Home Health Compare
HealthGrades	Quality ratings per organization's defined methodology
National Committee for Quality Assurance (NCQA)	HEDIS measures allowing comparison of healthcare plan performance
Partners in Information Access for the Public Health Workforce	Links to numerous Health Data Tools and Statistics from government and public health sources, behavioral surveys, and organizational initiatives.
The Joint Commission (TJC)	Quality Check directory
The Leapfrog Group	Overall patient safety and selected procedures per organization's defined methodology

Notes

[1] Health Insurance Portability and Accountability Act of 1996. Public Law 104-191. 45 CFR § 160.103.

[2] National Center for Quality Assurance. "HEDIS & Performance Measurement." http://www.ncqa.org/HEDISQualityMeasurement.aspx.

[3] The Leapfrog Group.About Leapfrog group: Available online at: http://www.leapfroggroup.org/about_leapfrog

[4] The Joint Commission information available at https://www.jointcommission.org/

[5] DNV GL information available at https://www.dnvgl.com/assurance/index.html

[6] Agency for Healthcare Research and Quality. About AHRQ http://www.ahrq.gov/cpi/about/profile/index.html

[7] OSHPD. “OSHPD Healthcare Atlas.” http://www.oshpd.ca.gov/general_info/Healthcare_Atlas_Details.html

[8] Pricewaterhouse Coopers Health Research Institute information available at http://www.pwc.com/us/en/health-industries/top-health-industry-issues/high-tech-database.html

[9] Alschuler, Liora, et al. "Quality Reporting Document Architecture (QRDA) Initiative Phase I Final Report." December 13, 2007. Available online at www.hl7.org.

[10] Department of Health and Human Services. "Report to Congress: Plan to Implement a Medicare Hospital Value-Based Purchasing Program." November 21, 2007. www.cms.hhs.gov/AcuteInpatientPPS/downloads/HospitalVBPPlanRTCFINALSUBMITTED2007.pdf.

[11] AHIMA. "Accountable Care: Implications for Managing Health Information."

[12] AHIMA. “Information Governance Principles for Healthcare(IGPHC)”

Prepared by

Patty Buttner, RHIA, CDIP, CHDA, CCS
Shannon H Houser, PhD, MPH, RHIA, FAHIMA
Jill Flanigan, MLS, MS, RHIT
Raymound Mikaelain, RHIA
Barbara Ryznar RPh, MSHI, CPHIMS, CHDA, CAPM
Clarice Smith, RHIA, CHP

Acknowledgements

Dee Banet, RN, MSN, CCDS, CDIP
Maria Barbetta
Aurea Beidler, MHA, RHIA, CHC, CHPS
Angie Comfort, RHIA, CDIP, CCS
Becky Dicus
Julie Dooling, MSHI,RHIA,CHDA
Melanie Endicott, MBA/HCM, RHIA, CCS, CCS-P, FAHIMA
Karen Gibson, MSA, RHIA, FAHIMA
Karen Gallagher Grant, RHIA, CHP
Vickie Griffin, RHIT,CCS
Linda K. Howard, CCS
Sheila A. Johnson-Siegler, RHIT, CCS
Lesley Kadlec, MA, RHIA, CHDA
Laurie Miller,RHIT, CCS-P
Andrea Romero, RHIT, CCS, CPC
Angela Rose, MHA, RHIA, CHPS, FAHIMA
Donna Rugg, RHIT, CCS, CDIP

Kathleen Scott, RHIA
Tracey Stanely, RHIT
Nicole Van Andel, MS, RHIA, CHPS
Lou Ann Wiedemann, MS, RHIA, CDIP, CHDA, FAHIMA Donna Wilson
Kelly Wilson, MBA, RHIA, CHIP, FAHIMA, MCRM
Holly Woemmet, MA, RHIA,CHPS

Prepared by (2013 Update)

Jill S. Clark, MBA, RHIA, CHDA, FAHIMA

Acknowledgements (2013 Update)

Sue Bowman, MJ, RHIA, CCS, FAHIMA
Lisa Brooks-Taylor, RHIA
Katherine Downing, MA, RHIA, CHP, PMP
Melanie Endicott, MBA/HCM, RHIA, CDIP, CCS, CCS-P, FAHIMA
Theresa Rihanek, MHA, RHIA, CCS
Angela Dinh Rose, MHA, RHIA, CHPS, FAHIMA
Diana Warner, MS, RHIA, CHPS, FAHIMA

Prepared by (Original)

June Bronnert, RHIA, CCS, CCS-P
Jill Clark, MBA, RHIA
Jane Cook, CPC
Rachael G. D'Andre, CPHQ, RHIA
Laurine Johnson, MS, CPC-H, RHIA
Crystal Kallem, RHIA

Acknowledgments (Original)

Margaret Foley, PhD, RHIA, CCS
Lou Ann Schraffenberger, MBA, RHIA, CCS, CCS-P

Article citation:
AHIMA Practice Brief. "Understanding Publicly Available Healthcare Data (2016) - Retired" (Updated June 2016)

Driving the Power of Knowledge