Understanding Publicly Available Healthcare Data (2013 update) - Retired

Save to myBoK

This practice brief has been updated. See the latest version here. This version is made available for historical purposes only. This update supersedes the January 2010 practice brief "Understanding Publicly Available Healthcare Data."

Healthcare organizations increasingly are required to collect and externally report more data, whether for patient safety, quality measurement, or patient satisfaction initiatives. In recent years, federal regulations such as ICD-10-CM/PCS, the "meaningful use" EHR Incentive Program, and accountable care organizations (ACOs) have supported the need for quality healthcare data. As a result, more data are publicly available to healthcare organizations and the general public for benchmarking and analysis.

For example, certain Present on Admission values affect reimbursement; however, they also drive patient safety reporting and risk-adjusted methodology. Healthcare data may also affect life or health insurance coverage. Employers may decide which life or health insurance coverage options are available to employees based on reported data. Because of this, healthcare organizations must ensure that they gather and report quality, accurate healthcare data.

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

Access Requirements for Publicly Reported Data

Publicly reported data typically contain aggregate data that is neither protected health information (PHI) nor personally identifiable information (PII) and is focused on healthcare services of facilities and providers. If the data are PHI or PII, then the privacy requirements must be met for release.

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 can be used to uniquely identify, contact, or locate a single person or can be used with other sources to uniquely identify a single individual. PII is not necessarily restricted to health information.

Increasing amounts of healthcare data have become available for analysis as the industry implements more health IT systems. The format and location of the data vary and range from state and national health databases to facility-based health 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.

Source-specific software is necessary to download and analyze data from the organization housing the data. For example, the Centers for Medicare and Medicaid Services (CMS) 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. The list of publicly available files (subject to privacy release approvals and computing resources) is available at www.cms.hhs.gov/FilesForOrderGenInfo.

The Agency for Healthcare Research and Quality (AHRQ) also provides access to different types of data that are primarily used for quality and utilization management purposes. For example, 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.

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 a local to national perspective, such as state-specific public health conditions to national morbidity and mortality statistics.

Publicly reported data may be presented for consumer use such as star ratings on different quality measure reports or for a healthcare provider's use to benchmark its own reports against surrounding 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 measure 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 externally reports the data to a third party, which presents the data in its own unique way, typically known as a report card.

Some of the most common quality and performance measure organizations are detailed below.

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. They were quickly endorsed when the National Quality Forum was formed and tasked with evaluating and approving quality and performance measures. The measures began in the 1980's with HMOs, and have grown to over 90 percent of plans voluntarily gathering and reporting data. The National Committee for Quality Assurance (NCQA) credentials an independent third party, which then certifies the results before they are reported.

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, with more detailed reports available for purchase. HealthGrades also offers educational information to providers on how it determines the ratings.

The Joint Commission is an independent, not-for-profit organization that accredits and certifies healthcare organizations on the quality of care they provide to individuals. It provides certification and accreditation information to patients and healthcare providers. The information is based on information provided from other Joint Commission-accredited facilities. Accreditation is voluntary.

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 in regards to the median Medicare payment and number of Medicare patients treated for specific conditions.

AHRQ provides information regarding quality and performance measures as well as a searchable compendium of healthcare report cards. The reports in the compendium are designed for consumers. The compendium includes comparative data on quality for healthcare organizations and provider type.[4]

Utilization Data

Public performance and quality measurement organizations typically report on the use of healthcare services through patient volume and charges for selected diagnoses and procedures. AHRQ provides multiple databases that contain this information through HCUP. HCUP also contains software tools to facilitate use of the data. The databases combine state, federal, and industry data, which can be used for research 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 US. Two major components of the data are the household, which focuses on individuals and their providers, and insurance. The data can be used for private planning and to help policy makers better understand the nation's healthcare needs and how to best meet them.

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 contains claims data from hospital outpatient encounters. A complete list of CMS data files regarding research, statistics, data, and systems is available at www.cms.hhs.gov/home/rsds.asp.

State-specific information may also be available, depending on the individual state. For example, in California, the Office of Statewide Health Planning and Development (OSHPD) was developed to provide 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.

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 a 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).

CDC has numerous programs that record disease statistics, management, treatment, and control. This information can be aggregated by specific populations such as workplace, diseases, and healthcare setting. It routinely works with state public health agencies such as the Paul Coverdell National Acute Stroke Registry to measure, track, and improve the quality of care. It is working to implement state-based registries to achieve these goals.

"Examples of Publicly Reported Data," [below], 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 be already analyzed and displayed in percentages or graphic displays.

Analytic Considerations

To ensure quality of the information that is shared, organizations must take into account the following fundamental considerations when analyzing data, whether the data are internal or available to the public.

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 these data independent of the company? If a specialty society or academic group with no ties to the pharmaceutical industry conducted the study, would the results differ?

Also consider internal benchmarking reports that could be used to compare procedures, complications, or other measures. If an organization only used the HCUP data (discussed earlier), their results would be representative of the Inpatient Medicare population only. 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 data can be collected in two ways. The first is by gathering individual health-related events or services, which can be collected by geographic region, over a specific time 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, time frames, data dictionary definitions, or tabulation methods. Inconsistent measurements, calculations, and conditional rules can also introduce unreliable results.

It is necessary to understand potential flaws in the methods of data collection and assumptions of causality reported, whatever method was chosen.

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 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. Consideration is given during this process to whether the data are reproducible, whether the sample size is statistically adequate, whether the time interval over which the data was collected is clearly defined, and so forth.

The scope of the data. 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, data design, data confidentiality, size of the database, data availability, accessibility, date range, and cost of the data.

The Future of Publicly Available Data

HIM professionals are facing many changes, including the way organizations collect, report, and manage healthcare data. With ICD-10-CM/PCS on the horizon and the movement toward widespread health IT use, there is the potential for even more data to be collected and reported.

Consumers are becoming more actively involved in their healthcare, looking to the Internet for guidance on providers or facilities of choice. In PricewaterhouseCooper's Health Research Institute release of its annual forecast of top health care trends in 2012, six in ten people surveyed said they would be comfortable having their personal health information shared among health organizations—including hospitals, doctors, insurance companies and pharmaceutical companies—if it would help to improve the coordination of their care. [5] 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 don't want to pay for Hospital-Acquired Conditions (HACs) or unnecessary readmissions. The ability to share information across systems and healthcare entities will continue to push towards the need for clinical data standards.

According to the Health Story Project, an alliance of healthcare vendors, providers, and associations working to develop data standards for HIE:

Approximately 1.2 billion clinical documents are produced in the United States each year. Dictated and transcribed documents make up around 60 percent of all clinical notes. These documents contain the majority of physician-attested information and are used as the primary source of information for reimbursement and proof of service. This tremendous source of clinical information is underutilized in current computer-based record systems. [6]

With the advancement of HIE, there is opportunity for more of this clinical narrative to be linked to the structured data of the electronic health record, advancing the use of data collection and reporting.

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 means that existing standards for clinical data exchange can be used to report current quality measures.[7] The initiative is one example of how improved data collection and reporting will lead to improved quality of care in the future.

Public reporting can also improve quality initiatives such as Value-Based Purchasing, further helping Medicare beneficiaries choose distinct, merited healthcare services. As a key component of the VBP program, public reporting could be enhanced 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[8]

Data that can inform consumer decision making could be part of the VBP program and publicly reported on the CMS Hospital Compare Web site. Data displays also could be modified so that beneficiaries could more easily interpret results.

The introduction of accountable care organizations (ACOs) will also impact the future of publically 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.[9]

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 question whether insurance companies will deny coverage based on available patient data.

For example, could certain body mass index levels that put patients at risk for a heart attack or other health concerns jeopardize insurance coverage? Will social histories including smoking and alcohol play into the equation for coverage? Similarly, could this increase in data impact consumer life insurance policies? There are no answers to these questions yet, but the purpose for looking at the data could change as more data are collected.

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.

Examples of Publicly Reported Data

Organization	**Available Data**
Agency for Healthcare Research and Quality	Medical Expenditure Panel Survey Consumer Assessment of Healthcare Providers and Systems Healthcare Cost and Utilization Project State Ambulatory Surgery Database (28 states) Healthcare Cost and Utilization Project State Emergency Department Database (27 states) Healthcare Cost and Utilization Project State Inpatient Database (40 states)
Centers for Disease Control and Prevention	United States Cancer Statistics
Centers for Medicare and Medicaid Services	Medicaid Statistical Information Files Medicaid Analytic eXtract Files

	Hospital Compare Nursing Home Compare Home Health Compare
HealthGrades	Quality ratings per organization's defined methodology
NCQA	HEDIS measures allowing comparison of healthcare plan performance
The Joint Commission	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. "How and Why Leapfrog Started." Online at www.leapfroggroup.org/about_us/how_and_why.
4. Agency for Healthcare Research and Quality. "Health Care Report Card Compendium." www.talkingquality.gov/compendium/index.asp?mode=about#information.
5. Health Data Management. "Top Healthcare Rends for 2012." http://www.healthdatamanagement.com/resource-center/?id=43623
6. Health Story Project. "About the Health Story Project." Available online at www.healthstory.com/about/about.htm.
7. Alschuler, Liora, et al. "Quality Reporting Document Architecture (QRDA) Initiative Phase I Final Report." December 13, 2007. Available online at www.hl7.org.
8. 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.
9. AHIMA. "Accountable Care: Implications for Managing Health Information."

Updated by

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

Acknowledgements

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. "Understanding Publicly Available Healthcare Data (2013 update) - Retired" (AHIMA Practice Brief, September 2013)