The Importance of Data Quality

by Julie A. Dooling, RHIA, CHDA

The quality of healthcare data impacts every decision made along the patient care continuum. The demand for accurate and reliable data has never been more important.

The sharing of today’s healthcare data is far reaching. Healthcare data is created at the source by providers such as physician groups, hospitals, skilled nursing facilities and pharmacies during the normal course of business. The data is then transformed into secondary data where it is shared with entities like payers, clearinghouses, third party vendors, public health agencies, government agencies, health information organizations, and the consumer. Each entity may have different purposes for collecting and using the data. Focus could range from the clinical, administrative, or financial aspects of data.

Data is collected in many ways and formats, using different technologies and standards often dependent upon the care setting. Technologies and data standards will vary from pharmacies to laboratories to acute care facilities; each likely to have a different technology vendor and complying with different data standards such as RxNorm, LOINC, or the C-CDA.

Healthcare organizations collect data that may never be used. Estimates indicate that healthcare data consists of 50 petabytes.1 To put this into perspective, one petabyte is enough to store the DNA of the entire population of the US and then clone them, twice.2 Data should be collected and aggregated with a purpose in mind; turning the data into meaningful information where patient care decisions are positively and responsibly influenced.

Ensuring the reliability and integrity of healthcare data begins with the accuracy and the completeness of the data captured in the patient’s health record (source data). At a high level, this process includes elements like information governance where information is recognized as an asset; patient matching where error and duplicate rates are monitored and remediated; identification and authentication of all authors contributing to an entry; a process for amendments and record corrections; and adequate audit trail functionality. It is important to note that established policies, procedures, and staff education must exist for these processes to achieve and maintain a high level of compliance.

With the pressure of payment reform and lowering healthcare costs on every healthcare organization’s agenda, performing data analysis and analytics to show value in data is a priority. To remain competitive and viable, healthcare organizations must be able use data to positively affect the quality of care, contain costs, and manage patient populations.

Measuring quality healthcare data comes in different forms such as face validity checks or more sophisticated methods such as descriptive, predictive, and prescriptive analytics.

Quality runs through the health information management (HIM) information lifecycle of capturing, processing, storing, and disposing of patient information. The HIM professional will continue to focus on collecting and maintaining quality data which will, in turn, deliver Reliable Information and Responsible Care.

Did you know that AHIMA has the only data analysis credential in healthcare? To learn more about the Certified Health Data Analyst (CHDA), visit ahima.org/certification/chda.

Notes:
AHIMA. “Assessing and Improving EHR Data Quality (Updated).” Journal of AHIMA 84, no. 2 (March 2013): 48–53 [expanded online version].

References:
1. healthcareitconnect.com/infographic-big-data-is-a-big-deal/
2. computerweekly.com/feature/What-does-a-petabyte-look-like

Quiz

1. Healthcare organizations collect data that may never be used.
   a. True
   b. False

2. The four Data Quality Management Model Domains are:
   a. Creating, Applying, Warehousing, and Analysis
   b. Collection, Application, Warehousing, and Analysis
   c. Collection, Application, Warehousing, and Deletion
   d. None of the above

3. The demand for accurate and reliable data was more important in the 1960s.
   a. True
   b. False

4. “Data Consistency” is defined as:
   a. When data is not reliable
   b. The extent to which the healthcare data are reliable and the same across applications
   c. This definition does not apply to healthcare data
   d. None of the above

5. Quality runs through the HIM information lifecycle of capturing, processing, storing, and disposing of patient information.
   a. True
   b. False