Nonsensical Data Input Produces Undesired Output for Data Analysis

Save to myBoK

By Lou Ann Wiedemann, MS, RHIA, CHDA, CDIP, FAHIMA

There is a growing hunger in healthcare today for attaining meaningful data. It is no longer an option to sift through large amounts of data stored in a variety of systems, spending valuable man hours to produce information that can be utilized to make good business decisions. Allowing for nonsensical data input most often results in undesired outputs—in other words, "garbage in, garbage out."

As more emphasis is being placed on quality, outcomes, and new payment methodologies, the healthcare sector is looking inwards toward their own data to inform decision making. As new data analyst roles emerge, health information management (HIM) professionals have a new arena to showcase their skills and knowledge.

HIM professionals have been painting the picture of patient care for years, interpreting the record and stories into meaningful data through code assignments, morbidity and mortality, case mix index, and basic statistics. As the amount of data has grown, and the data itself has become more complex, it has become essential to analyze healthcare data and use it for multiple secondary purposes.

Data basically goes through three processes or stages before it is considered meaningful:

- Data capture
- Data provisioning
- Data analysis

Stage 1: Data Capture

Data capture involves understanding how data gets into a healthcare system. This step can include multiple mechanisms such as direct data entry, use of pick lists, free text, or integrating multiple source systems. More than any other factor, understanding how data capture occurs has the greatest impact on data analysis.

Data capture, done correctly, ensures that the right information is collected, in the right format, so that it can be used accurately in analysis. Most data analysts do not directly enter information, so understanding the impact of good data capture and what source systems are used is imperative. Moreover data analysts are often not a part of the system development process, so the data that are accessible to analysts for analysis is often dictated by the vendor or health IT implementation team.

Stage 2: Data Provisioning

Data provisioning is the need to pull information from multiple systems. For example, a data analyst is asked to provide a report on quality improvement efforts related to adverse drug reactions. The analyst will need to know that the following systems must be queried:

- Electronic health record system for basic demographic information
- Encoder, if the organization uses coded data to track adverse drug reactions
- Pharmacy system to identify national drug code information on both the drug the patient reacted to as well as the drug administered for the reaction
- Billing system to identify a potential increase in costs or charges

Pulling this data together, in a common format, and communicating it through common identifiers is incredibly time consuming. Accomplishing that without an error is a major feat. Most organizations choose to use an enterprise data warehouse that aggregates, organizes, and catalogues structured data queries.

Stage 3: Data Analysis

Data analysis is the piece of the puzzle that occurs once data capture and provisioning has taken place. Good data analysis will include data quality checks, data discovery, data interpretation, and finally data presentation.

Data quality checks include all of those steps the analyst uses to ensure that the data is accurate. It may also feature steps that include understanding the audience to which the data will be presented and tailoring the data to that group. This also assists the audience in understanding the size, complexity, and accuracy of the data.

Data discovery encompasses an initial search of the data for extreme measures, outliers, or trends that may be meaningful. Sometimes the face value of the data generates a validation of its own. Does the data represent what the analyst thinks that it should? Does it cover all of the pieces needed?

Data interpretation is most often associated with analysis. This step in the process would include identifying variables for analysis. This is where analysts would apply statistical procedures or formulas to the data to ensure that the hypothesis is correct.

Data Presentation Skills Equally Important

The final step, which is critical, is presenting the data. An analyst can spend hours or days reviewing data, applying statistical algorithms, and verifying data-backed findings. But without appropriate presentation skills the full meaning of the data will not be conveyed. The analyst must tell a story in a way that is easily consumable for the intended audience. The need for accurate and timely data analysis is great. The creation of precise data is a critical component of healthcare today. To ensure that organizations do not fall into garbage in, garbage out scenarios, high-quality analysis must occur in healthcare organizations.

References

Taylor, Lisa Brooks. “Show Me the Data.” *Journal of AHIMA* 84, no. 9 (September 2013): 60-61.

Eramo, Lisa A. “Healthcare’s Data Revolution: How Data is Changing the Industry and Reshaping HIM’s Roles.” *Journal of AHIMA* 84, no. 9 (September 2013): 26-32.

Staheli, Russ. “4 Best Practices You Should Know for Analyzing Healthcare Data.” *Health Catalyst*. November 4, 2013. www.healthcatalyst.com/healthcare-analytics-best-practices.

Lou Ann Wiedemann (lou-ann@wiedemann@ahima.org) is a senior director of HIM practice excellence at AHIMA.

Article citation:
Wiedemann, Lou Ann. "Nonsensical Data Input Produces Undesired Output for Data Analysis" *Journal of AHIMA* 85, no.6 (June 2014): 50-51.