# Tools for Data Analysis: New Toolkit Provides Resources for Health Data Analysts



By Jill Clark, MBA, RHIA

Healthcare entities collect a vast amount of administrative, clinical, and financial data. Health data analysts play a critical role in helping translate these data into meaningful information that can be used to identify improvement opportunities, recognize trends, and assist in decision making.

Within the next five years, HIM professionals will be faced with implementing and maintaining electronic health records, transitioning to ICD-10-CM/PCS, managing a plethora of government audits, and reporting quality measures. Healthcare entities will have an increased interest in analyzing and interpreting their data to better prepare for and manage these operational and financial challenges.

## The Right Tools for the Job

HIM professionals working with data reporting and analysis will be required to know the various sources for data within their organization, how data are structured, what questions to ask when developing a report, and the appropriate method to display data.

Asking the right questions ensures a report's outcome supports the user's request. Questions to consider before developing a data report might include:

- What patients are to be studied?
- What does the requester want to know about these patients?
- In what format and structure does the requester want to see the information?
- What is the requester's timeline or turnaround requirements?
- What is the requester's price range or budget, if known?

Tools to facilitate data analysis and reporting come in many forms, such as a data dictionary. A data dictionary is a descriptive list of names, definitions, and attributes of data elements to be collected in an information system or database. Its purpose is to standardize definitions and ensure consistent use. It facilitates a common understanding of an organization's data quality when developing reports and analyzing information.

AHIMA's newly published "Health Data Analyst Toolkit" includes an example of a data dictionary, excerpted [below]. The full sample data dictionary includes other fields such as the source system, the date first entered, and the logic for including the item. HIM professionals can customize their data dictionaries according to their needs.

## About the Toolkit

The "Health Data Analyst Toolkit" was developed to aid HIM professionals working with healthcare data in a variety of settings across the country. The toolkit begins by discussing data acquisition, specifically focusing on the purpose of the data dictionary. It then discusses how information is requested and validated. The toolkit also includes examples of good and poor data displays.

The second half of the kit includes a list of common formulas and statistics, a glossary of terms, and an annotated bibliography. Finally, the toolkit concludes with a sample case study illustrating how data are collected, analyzed, and transformed into information for reporting.

The toolkit is available online at www.ahima.org/resources/infocenter/datamanagement.aspx.

## Data Dictionary Sample (Excerpt)

| Data field | Name | Definition | Data Type | Format | Field size | values |
|---|---|---|---|---|---|---|
| Admission Date | ADMIT_DATE | The date the patient is admitted to the facility as an inpatient | Date | MMDDYY | 8 | Admission date cannot precede birth date or 2007. No hyphens or slashes. |
| Census | CENSUS | The number of inpatients present in the facility at any given time | Numeric | X to XX | 3 | Any whole number from 0 to 999 |
| Ethnicity | PT_ETHNIC | The patient's ethnicity. Must be reported according to official<br><br>Office of Management and Budget (OMB) categories. | Alpha-numeric | EX; letter must be uppercase | 2 | E1 = Hispanic or Latino ethnicity; E2 = non-Hispanic or Latino ethnicity |
| Infant Patient | INFANT_PT | A patient that has not reached 1 year of age at the time of discharge | Alpha-numeric | Age in months = XD to XXD OR XM to XXM | 3 | Must be > 0 AND < 1 year |
| Inpatient Daily Census | ID_PAY_CENSUS | The number of inpatients present at census-taking time each day plus any inpatients who were both admitted and discharged after the previous day's census-taking time | Numeric | X to XX | 3 | Any whole number from 0 to 999 |
| Medical Record Number | MR_NUM | The unique number assigned to a patient's medical record. The medical record is filed under this number. | Alpha-numeric | Xxxxxx; requires leading zeroes | 6 | 000001 to 999999 |
| Patient Age | PT_AGE | Age of patient calculated using most recent birthday attained prior to or on same day as discharge | Numeric or alpha-numeric | Age in days = xD to xxD OR Age in months = xM to xxM OR Age in years = x to xxx | 3 | Age must be > 0, and < OR = 124 years; children less than 1 year must be > 0M AND < 1 year |
| Patient Sex | PT_SEX | Patient sex | Alpha-numeric | Letter; must be upper case | 1 | M = Male<br><br>F = Female<br><br>U = Unknown |
| Patient Zip Code | PT_ZIP_CODE | Zip code of patient's residence | Alpha-numeric | XXXXX-XXXX | 11 | 00000 to 99999; 00000 = Unknown<br><br>99999 = Foreign |

| | | | | | | |
|---|---|---|---|---|---|---|
| Pediatric Patient | PED_PT | A patient that has not reached 18 years of age at the time of discharge | Numeric | Age in days = XD to XXD OR age in months = XM to XXM OR age in years = X to XXX | 3 | Age must be > 0 AND < 18 years; children less than 1 year must be > OM AND < 1 year |

Jill Clark (jill.clark@ahima.org) is a professional practice manager at AHIMA.

**Article citation**:
Clark, Jill S. "Tools for Data Analysis: New Toolkit Provides Resources for Health Data Analysts" *Journal of AHIMA* 82, no.2 (February 2011): 40-41.