

Appendix A: Data Quality Management Model Domains and Characteristics

Characteristic	Application	Collection	Warehousing	Analysis
<p>Data Accuracy</p> <p>The extent to which the data are free of identifiable errors.</p>	<p>To facilitate accuracy, determine the application's purpose, the question to be answered, or the aim for collecting the data element.</p> <p>Standard acceptable values should be used where available.</p> <p>Where possible value flags and constraints should be implemented.</p> <p>For example, data entry of height into EHRs should flag or highlight very small (less than 12 inches) or very tall (over seven feet) heights.</p>	<p>Ensuring accuracy involves appropriate education and training along with timely and appropriate communication of data definitions to those who collect data. The applications should constrain entry to allowable values where possible.</p> <p>For example, data accuracy will help ensure that a patient height cannot be entered erroneously as five inches when it is in fact 50 inches. In addition to a primary data error, this would impact any calculated fields such as Body Mass Index (BMI).</p>	<p>To warehouse data, appropriate edits should be in place to ensure accuracy, such as basic field length checks.</p> <p>Also, error reports are generated related to transfers to and from the warehouse.</p> <p>All warehouses should have a correction and change management policy to track any changes.</p>	<p>To accurately analyze data, ensure that the algorithms, formulas, programming, and translation systems are correct.</p> <p>For example, ensure that the encoder assigns correct codes and that the appropriate DRG is assigned for the codes entered.</p> <p>Continual data validation is important to ensure that each record or entry within the database is correct.</p>

Characteristic	Application	Collection	Warehousing	Analysis
<p>Data Accessibility</p> <p>Data items that are easily obtainable and legal to access with strong protections and controls built into the process.</p>	<p>The application and legal, financial, process, and other boundaries determine which data to collect. Ensure that collected data are legal to collect for the application. For example, recording the date of birth and race in the EHR is appropriate and should only occur once with verification. Then the values should roll forward.</p>	<p>When developing the data collection instrument, explore methods to access needed data and ensure that the best, least costly method is selected. The amount of accessible data may be increased through system interfaces and integration of systems. For example, the best and easiest method to obtain demographic information may be to obtain it from an existing system. Another method may be to assign data collection by the expertise of each team member. For example, the admission staff collects demographic data, the nursing staff collects symptoms, and the HIM staff assigns codes.</p> <p>Data entry should undergo a cost- benefit analysis process to determine which method provides the best data most efficiently.</p>	<p>Technology and hardware impact accessibility. Establish data ownership and guidelines for who may access or modify data and/or systems. Inventory data to facilitate access.</p> <p>In the EHR it may be advisable to establish data ownership or governance at the data element level, especially data which are reused. For example, allergies are recorded by many different clinicians and come in many forms. Who defines what an allergy is? How does this impact the use of allergies in the EHR, especially for clinical decision support?</p>	<p>Access to complete, current data will better ensure accurate analysis and data mining. Otherwise results and conclusions may be inaccurate or inappropriate.</p> <p>For example, use of the Medicare case mix index (CMI) alone does not accurately reflect total hospital CMI. Consequently, strategic planning based solely on Medicare CMI may not be appropriate.</p>

Characteristic	Application	Collection	Warehousing	Analysis
<p>Data Comprehensiveness</p> <p>All required data items are included. Ensures that the entire scope of the data is collected with intentional limitations documented.</p>	<p>Clarify how the data will be used and identify end users to ensure complete data are collected for the application. Include a problem statement and cost-benefit or impact study when collected data are increased.</p> <p>For example, in addition to outcome it may be important to gather data that impact outcomes.</p>	<p>Cost-effective comprehensive data collection may be achieved via interface to or download from other automated systems.</p> <p>Data definition and data precision impact comprehensive data collection (see these characteristics below).</p>	<p>Warehousing includes managing relationships of data owners, data collectors, and data end- users to ensure that all are aware of the available data in the inventory and accessible systems. This also helps to reduce redundant data collection.</p>	<p>Ensure that all pertinent data impacting the application are analyzed in concert.</p> <p>This is especially important when EHR clinical decision support is utilized. Incomplete data can result in underreporting a numerator or denominator.</p>

Characteristic	Application	Collection	Warehousing	Analysis
<p>Data Consistency</p> <p>The extent to which the healthcare data are reliable and the same across applications.</p>	<p>Data are consistent when the value of the data is the same across applications and systems, such as the patient's medical record number. In addition, related data items should agree.</p> <p>For example, data are inconsistent when it is documented that a male patient has had a hysterectomy.</p>	<p>The use of data definitions, extensive training, standardized data collection (procedures, rules, edits, and process) and integrated/interfaced systems facilitate consistency.</p> <p>Static data should be moved between users. For example, once date of birth has been definitively established, age at the time of treatment should be calculated, not entered by a user who might make an error.</p>	<p>Warehousing employs edits or conversion tables to ensure consistency. Coordinate edits and tables with data definition changes or data definition differences across systems. Document edits and tables.</p> <p>When new data are loaded it should be checked against existing data for consistency.</p> <p>For example, is someone reporting a different race for a patient?</p>	<p>Analyze data under reproducible circumstances by using standard formulas, scientific equations, programming, variance calculations, and other methods. Compare "apples to apples."</p> <p>Any manipulation of data, aggregating or otherwise, should be documented thoroughly. For example, how is BMI calculated and has the formula been checked?</p>

Characteristic	Application	Collection	Warehousing	Analysis
<p>Data Currency</p> <p>The extent to which data are up-to-date; a datum value is up-to-date if it is current for a specific point in time. It is outdated if it was current at a preceding time yet incorrect at a later time.</p>	<p>The appropriateness or value of an application changes over time.</p> <p>In EHRs it is imperative that the guidelines and algorithms be up-to-date. For example, acceptable blood pressure ranges have lowered, as have target HbA1C levels.</p>	<p>Data definitions change or are modified over time. These should be documented so that current and future users know what the data mean. These changes should be made in accordance with data governance policies and practices. Further, they must be communicated in a timely manner to those collecting data and to the end users.</p>	<p>To ensure current data are available, warehousing involves continually validating systems, tables, and databases. The dates of warehousing events should be documented.</p>	<p>The availability of current data impacts the analysis of data.</p> <p>For example, analyzing the long-term incidence or prevalence of disease requires data in a different timeframe than when trying to track a disease outbreak for biosurveillance purposes.</p> <p>Validating data from various fiscal and calendar years should also be considered.</p>
<p>Data Definition</p> <p>The specific meaning of a healthcare related data element.</p>	<p>The application's purpose, the question to be answered, or the aim for collecting the data element must be clarified to ensure appropriate and complete data definitions.</p> <p>Does the system use the Office of Management and Budget (OMB) standard for race and ethnicity? If not, what are the definitions and acceptable values?</p>	<p>Clear, concise data definitions facilitate accurate data collection.</p> <p>For example, the definition of patient disposition may be "the patient's anticipated location or status following release or discharge." Acceptable values for this data element should also be defined. The instrument of collection should include data definitions and ensure that the application limits data collection to the allowed values.</p>	<p>Warehousing includes archiving documentation and data. Consequently, data ownership documentation and definitions should be maintained over time. Inventory maintenance activities (purging, updates, and others), purpose for collecting data, collection policies, information management policies, and data sources should be maintained over time also.</p>	<p>For appropriate analysis, display data needs to reflect the purpose for which the data were collected. Appropriate comparisons, relationships, and linkages need to be shown.</p>

Characteristic	Application	Collection	Warehousing	Analysis
<p>Data Granularity</p> <p>The level of detail at which the attributes and values of healthcare data are defined.</p>	<p>A single application may require varying levels of detail or granularity.</p> <p>For example, census statistics may be utilized daily, weekly, or monthly depending upon the application. Census is needed daily to ensure adequate staffing and food service. However, the monthly trend is needed for long-range planning. Similarly, lab test results may be trended at various levels of detail.</p>	<p>Collect data at the appropriate level of detail or granularity.</p> <p>For example, the temperature of 100° may be recorded. The granularity for recording outdoor temperatures is different from recording patient temperatures. If patient Jane Doe's temperature is 100°, does that mean 99.6° or 100.4°?</p> <p>Appropriate granularity for this application dictates that the data need to be recorded to the first decimal point while appropriate granularity for recording outdoor temperatures may not require it.</p>	<p>Warehouse data at the appropriate level of detail or granularity.</p> <p>For example, exception or error reports reflect granularity based on the application. A spike (exception) in the daily census may show little or no impact on the month-to-date or monthly reports.</p>	<p>Appropriate analysis reflects the level of detail or granularity of the data collected.</p> <p>For example, a spike (exception) in the daily census resulting in immediate action to ensure adequate food service and staffing may have had no impact on analysis of the census for long-range planning. Of particular note for analysis is the impact of any rounding which might be done for numerical data.</p>

Characteristic	Application	Collection	Warehousing	Analysis
<p>Data Precision</p> <p>Data values should be strictly stated to support the purpose.</p>	<p>The application's purpose, the question to be answered, or the aim for collecting the data element must be clarified to ensure data precision.</p> <p>What level of detail is needed for the data collection purpose? Are age ranges or four U.S. regions sufficient?</p>	<p>To collect data precise enough for the application, define acceptable values or value ranges for each data item.</p> <p>For example, limit values for gender to male, female, and unknown; or collect information by age ranges or allow more detailed collection to fully meet the needs.</p>	<p>Are warehouses receiving and storing all data elements being transferred from the source system?</p>	<p>If the precision of the data has been altered in the analysis is the process understood and well documented?</p>
<p>Data Relevancy</p> <p>The extent to which healthcare-related data are useful for the purposes for which they were collected.</p>	<p>The applications purpose, the question to be answered, or the aim for collecting the data element must be clarified to ensure relevant data.</p>	<p>To better ensure relevancy, complete a pilot of the data collection instrument to validate its use. A "parallel" test may also be appropriate, completing the new or revised instrument and the current process simultaneously.</p> <p>Communicate results to those collecting data and to the end users. Facilitate or negotiate changes as needed across disciplines or users.</p>	<p>Establish appropriate retention schedules to ensure availability of relevant data.</p> <p>Relevancy is defined by the application.</p> <p>It may be appropriate for warehouses to subset data related to its relevancy for certain uses.</p>	<p>For appropriate analysis, display data to reflect the purpose for which the data were collected. This is defined by the application. Show appropriate comparisons, relationships, and linkages.</p>

Characteristic	Application	Collection	Warehousing	Analysis
<p>Data Timeliness</p> <p>Concept of data quality that involves whether the data is up-to-date and available within a useful time frame. Timeliness is determined by how the data are being used and their context.</p>	<p>Timeliness is defined by the application.</p> <p>For example, patient census is needed daily to provide sufficient day- to-day operations staffing, such as nursing and food service. However, annual or monthly patient census data are needed for the organization's strategic planning.</p> <p>In the EHR, vitals may be taken once per visit for ambulatory care patients, but every 15 minutes or more often for critically ill patients.</p>	<p>Timely data collection is a function of the process and collection instrument.</p> <p>In the EHR, system performance plays an important role in data timeliness. Data display should be sub-second and data entry should occur instantaneously.</p>	<p>Warehousing ensures that data are available per information management policy and retention schedules.</p> <p>For EHR or clinical data warehouses, is the data updated concurrently or does it occur in a batch process?</p>	<p>Timely data analysis allows for the initiation of action to avoid adverse impacts. For some applications, such as allergy-drug or drug-drug interactions, timely may be seconds. For others, such as the prevalence of a disease over time, it may be years.</p>

Article citation:
 AHIMA. "Data Quality Management Model. Appendix A: Data Quality Management Model Domains and Characteristics." *Journal of AHIMA* 83, no.7 (July 2012): 68-71.