Lesson 5: Data Quality Control and Assurance

Data contamination
Process or phenomenon, other than the one of interest that affects a value and results in erroneous values.

Quality assurance & quality control
Activities that prevent errors from entering or remaining in a data set. These activities ensure data quality before it is collected, entered, or analyzed, and monitor and maintain the quality of the data throughout the study’s data life cycle.

Types of errors

<table>
<thead>
<tr>
<th>Errors of commission</th>
<th>Errors of omission</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incorrect/inaccurate data is entered</td>
<td>Data or metadata is not recorded</td>
</tr>
<tr>
<td>Examples: malfunctioning instrument or mistyped data</td>
<td>Examples: inadequate documentation, GPS unit runs out of power, forgotten measurements</td>
</tr>
</tbody>
</table>

Main steps for QA/QC activities in data life cycle

QA/QC Best Practices: During data collection
- Consider double entry, where data is keyed in by two different people, agreement checked by computer, and mistakes corrected
- Transcribe the data by listening to a recording of yourself reading the data or by using a text-speech program
- Design data storage well:
  - minimize how many times a piece of data must be entered
  - use consistent terminology
  - atomize data: enter one piece of data per cell
- Document changes to data to avoid duplicate error checking and to allow revision if necessary

QA/QC Best Practices: After data collection
- Make sure data line up in proper columns
- Make sure there are no missing, impossible, or anomalous values
- Perform statistical summaries, like means and standard errors
- Look for outliers, or extreme values, to identify possible data contamination.
  - This can be done with graphs, maps, or statistical methods.

QA/QC Best Practices: Before data collection
Define standards to be used during the project, including:
- data formats
- code
- measurement units
- metadata to be captured

Assign responsibility for data quality

Local contact information

View all Education Modules at https://www.dataone.org/education-modules

QA/QC Best Practices: Before data collection