Throwing Data over the Wall: The Next Big Threat to Data Quality

Documents Defining Data Rules Don’t Stay with the Data

Premise
- Each group defines “quality” differently
- Groups don’t understand others’ data rules
- Data are “tossed over the wall” to the next group
- Data rules are lost as the data travel among groups
- Don’t know how the data can be used, posing a threat to quality

Possible Solutions
- Centrally-recorded in-depth cross-silo exchange in plain language to define who does what to/with the data
- Build quality into study processes so data issues can’t be repeated
- Sometimes data rules must differ, so rules must travel with the data
- CDISC’s machine-readable protocol could help to record rules (e.g., when to start collecting AEs/SAEs)
- ICH E3 Clinical Study Reports Part 9.6 Data QA could be used to describe these rules

Groups see only inside their silos

Data handling rules are lost

Examples

1. Hmm... Q3 on Visit 2
  Hamilton Depression survey is blank—can’t calculate a total score. Better verify

2. Biostat says Q3 on Visit 2 Ham-D is blank
   OK, I’ll ask the site

3. Q3 on Visit 2 Ham-D is blank
   Subject’s gone—I can’t retrieve it

4. Site says it can’t be retrieved after the visit, so don’t send queries
   OK. We’ll stop

5. Sorry—it can’t be fixed
   Now I know not to query subject surveys

6. Darn. That’s 1 less subject for the efficacy pool
   OK, I’ll check it & retrain the site

7. Site 8 seems to have a lot of blanks in the Ham-D forms
   Hmm... That means 1 less efficacy subject

8. Data Repository
   Time Started: Informed Consent
   Time Started: Day 1

Data handling is a threat to quality
- Data Rep for S1 shows both AEs and SAEs starting with informed consent
- Data Rep for S2 shows only AEs starting with informed consent
- Data Rep for S3 shows only SAEs starting with informed consent

Without checking every protocol, there is no way to know if an absence of AEs before Day 1 is because none occurred or because AEs weren’t captured.

Author: Kit Howard, MS, CCDM, CRCP
Founder & Owner, Kestrel Consultants, Inc.
410 Rose Dr, Ann Arbor MI 48103
Kit@KestrelConsultants.com +1 734-576-3031
www.KestrelConsultants.com

FDA
- Focus: Individual subject & aggregate data across entire submission
- Quality is: Same as sponsor silos plus consistency across submission
- Note: May have little data handling information

Site
- Focus: individual subject &/data points
- Quality is: clear protocol, simple CRFs, rational queries, regulatory compliance

Clinical Data Manager
- Focus: within-subject consistency
- Quality is: complete CRFs that meet edit check specs

Sponsor / Monitor
- Focus: individual subject
- Quality is: complete CRFs that reflect source, regulatory compliance at site

Stats Analysis Plan
- Focus: cross-subject & cross-site patterns and consistency
- Quality is: data that conform to analysis assumptions, all sites similar

Journals
- Focus: Manuscript
- Quality is: Unknown. Get summary of analysis plan, rarely protocol & never data rules
- Result: Publishing based on very limited info but may greatly affect treatments

Data Repositories
- Focus: Many
- Quality is: Difficult to assess. Data design, handling and analysis rules aren’t available
- Result: Data may be used inappropriately and will not know

Documents Defining Data Rules Don’t Stay with the Data