Data Quality:
New methods for minimizing data queries, finding discrepancies, and correcting them efficiently

MAGI
Clinical Trial Agreements, Budgets & Regulatory Conference – West
Las Vegas, 14 October 2008

Kit Howard, MS
Principal & Owner, Kestrel Consultants
How can sites and sponsors minimize data queries?
Easy!!!

The Sites Will Produce Perfect Data
The End
Reduce the number of queries

Why are the queries sent?

To improve data quality

What’s the matter with the data?

It’s not high quality

What does “high quality” mean?

We need to define “quality”
Defining Quality

FDA
ALCOA plus 3 C’s

Deming
Conforming to the customer’s expectations.
Nothing else matters.

Sales
Giving the customer not only what they want, but what they didn’t know they wanted

Juran
“Fit for their intended uses in operations, decision making and planning.”
“Quality” is not an absolute attribute

It depends on the needs and perceptions of the customer and the degree to which they are met
Who is the Customer?

Subject/Patient

IRB

Regulation

Site

Submission/Publication

Post-marketing surveillance

Registry/Company Warehouse/Janus

DATA

CSR Writer

Sponsor Clinician

CRA

CD

Statistician

Programmer
What Do They Want?

- Fast enrolment of eligible subjects
- Access to accurate and complete data ASAP
- Full timely info about potential safety issues

Requirements are in the Protocol
What Do They Want?

- Complete well-organized source data & docs
- Accurate and timely CRF completion
- Requirements are in the Monitoring Guidelines
• Data delivered in timely fashion
• Internally consistent data on CRFs completed in accordance with instructions
• Data that show protocol compliance & conform to assumptions & expectations

Requirements are in **Data Management Plan**, including **CRF Completion Instructions**
What Do They Want?

- Complete data that conform to their programming assumptions

Requirements are in the User Requirements
What Do They Want?

- Complete data that conform to protocol requirements and their analysis assumptions
- No outliers, unusual values or unexpected events
- Data that are sufficiently consistent between sites to produce a trustable analytical result

Requirements are in the Statistical Analysis Plan
Two Types of Requirements

More Obvious: Subject-Specific

Typical Errors:
- Blank fields
- Transcription errors

Result in
- Increased data variability
- Potentially fewer evaluable subjects
- May mask safety issues

Less Obvious: Cross-Subject

Typical Issues:
- Values different at one site
- Data at one site are more or less variable

Result in
- Skewing or biasing the data
- False study conclusions
- May never be detected
Differing Variability of Scores
Contributing Factors: Sponsors

- Multiple sources for quality requirements
- Different study designs
- Inconsistent operational requirements
- Use of multiple vendors in a given study
- Overutilized resources (e.g., CRAs)
- Poor recognition that site handles many studies
- Lack of personal interaction with site results in site = number, not site = people
Contributing Factors: Sites

- Multiple studies running at one time
- No access to many quality requirements
- Multiple data capture systems within and across studies
- High staff turnover where knowledge transfer can be challenging
- Regulatory requirements for publication research very different from industry research
- Defined quality systems are uncommon
Where does that leave the site?
Three Primary Drivers for Queries

• Different and sometimes conflicting requirements for the data
• Lack of understanding between the various customers re the different requirements
• No definition of “how good is good enough,” i.e., no assessment of the risk of various data errors
How Good is Good Enough?

• The cleaner data have to be, the more queries are generated
• Sponsors tend to clean all data completely
• Institutes of Medicine report in 1999 stated that data are good enough if they “support the same interpretations and conclusions as those derived from error-free data”
• It does not say that quality data are error-free
Apply Risk-Adjusted Approach

<table>
<thead>
<tr>
<th>Domain Type</th>
<th>Likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
</tr>
<tr>
<td>Primary (high)</td>
<td>Medium*</td>
</tr>
<tr>
<td>Secondary (medium)</td>
<td>Low</td>
</tr>
<tr>
<td>Tertiary (low)</td>
<td>Low</td>
</tr>
</tbody>
</table>

* Risk Level

- Risk: Likelihood and severity of impact on study conclusions or subject safety
- Primary Domain: primary efficacy and safety (primary endpoints, AEs)
- Secondary: summarized domains (e.g., con meds)
- Tertiary: non-summarized domains (e.g., physical exam)
Site
Quality Management System
Quality Management System

- Defined by ISO 9000:2000 as “a systematic approach to managing the organization's processes so that they consistently turn out product that satisfies customers' expectations.”
- “Systematic approach” is defined documented processes (e.g., SOPs)
- “Process” is everything you do as part of a clinical trial
- “Product” is the data
- “Customer” is the various parts of the sponsor
- “Expectations,” if not met, lead to queries
Creating and Improving Quality

• Plan / Do / Check / Act
  – Plan: define the process and decide what quality looks like
  – Do: perform the process
  – Check: check how the results met the expectations
  – Act: Change the process to improve

• Provides a context for CAPA
Plan / Do / Check / Act

- Define and map the process with the controls
- Decide what “good quality” results are
- Determine how to measure and document results

We’ll use CRF completion as an example
Plan / Do / Check / Act

• Define and map the process with the controls
  – Who does it and with what?
    • Sponsor’s CRF completion guidelines
      – Are they complete and do you fully understand them?
      – If you follow them precisely will you avoid all queries?
  – How is it done?
    • When should they be completed?
    • How do you know they are done?
    • What happens if some data are not available yet?
  – What about QC?
    • Who checks the CRFs after completion and for what and when?
    • How and by whom are errors corrected?
Plan / Do / Check / Act

- Determine how to measure quality and document results
  - How define an error? Count them?
  - Where does this get documented? By whom?
  - What is done with the results?

- What triggers a system fix?
  - What type of error should trigger a process change?
  - Who analyzes the process to determine root cause?
  - How do you ensure that the fix is not a band-aid?
  - How are the process changes implemented?
Plan / Do / Check / Act

• The rest is basically executing the plans
  – Do: perform the processes (i.e., complete CRFs)
  – Check: apply quality checks and assess for acceptability
  – Act: institute improvements
Practical To-Dos: General

• Get certification and other general training
• Develop formal training on your processes
  – Document who attends
  – Include quiz or other learning measure
  – Train new staff (ensure continuity if CRC leaves)
• Ensure sufficient resources to do the study
• Build study activities into the daily routine
• Try to standardize your practices for all trials
Cross-Functional Communication

• Site collaborate with sponsor’s study team
  – Understand the operational aspects of the protocol
    • Do a “dummy” set-up
    • Do this early, preferably before accepting the study
  – Participate in the CRF review and especially in review of CRF completion guidelines
  – Get copies of the Monitoring Guidelines and the Data Management Plans
  – Ensure that all the items you’ll be checked on are somewhere in your study guidelines
Cross-Functional Communication

- Site collaborate with sponsor’s study team
  - Meet with monitors and data manager to ensure common understanding (don’t assume monitors understand CDM)
  - Establish consistent process and tools to ensure you apply the right processes to the right study
Cross-Functional Communication

• Site collaborate with sponsor’s study team
  – Site can communicate with other sites via email, telecons
    • Sites can share information and tips (e.g., subject recruitment, CRF confusion, test administration)
    • This helps to ensure common practices, interpretations, etc. and reduce cross-site variability
    • Include monitor and CDM occasionally for common questions (e.g., why asking a given query)
Some Thoughts

• Do you know your current performance? E.g.,
  – #s of eligible subjects enrolled per time frame?
  – #s of queries generated from the CRF data? How many is too many?
  – # of findings on a study monitor’s report?

• Do you know the accuracy of:
  – Your source data records?
  – Your transcription from source data to CRFs?

• What is your target performance?
You’ve done all this and STILL get a query!
Resolving queries efficiently

- Really understand what the query is asking and why the data are in question
  - Don’t let the query lead you to a solution
  - Reference DMP, monitoring guidelines, CRF completion guidelines and internal guidelines
  - If the query is comparing more than one data point check them both
  - Check against source doc for transcription error
  - Apply the “sniff test” – do the data smell right based on the rest of the subject’s experience?
Resolving queries efficiently

• Assuming there is an error, can it be corrected?
  – If ‘no’, then document that on the query form and in the source data
  – If yes, acknowledge it and correct it
    • Is the query targeting the right data point?
    • Will changing the data point will affect other data points?
    • Make the change thoughtfully, correctly and completely
    • Document changes fully in source data, the CRF(s) and the query form
    • This is more time consuming but will help prevent follow-up queries.
Resolving queries efficiently

• Apply your quality plan criteria for documenting errors

• If necessary, determine why the error occurred and how to prevent the error from occurring again

• Always update data correctly – e.g., single crossout, enter change, initial and date; or the appropriate electronic update method
Benefits of a QMS Approach

• Your data will be much higher quality
• Each new study will be easier to set up
• Your costs will be reduced
• The burden on personnel will be reduced
• Transitioning to new staff will be much easier
• You’ll be a much more attractive investigative site for sponsors
• Vastly reduce the chances of 483s, warning letters, being disbarred and going to jail
Oh, and by the way, you’ll get fewer queries.
References & Links


• *The Detection of Suspicious Data*, presentation, Alec Vardy, CV Therapeutics, Oct 2003; special thanks for the data trend examples


• Data Quality Research Institute project “Data Trend Evaluation for Increasing Data Quality”, work in progress.


• Stan Woollen as quoted on FirstClinical.com in Q&A files on GCP

• International Standards Organization, ISO 9000 [www.iso.org](http://www.iso.org)

• How to review a protocol quickly: [http://www.researchadvocates.org/article014.htm](http://www.researchadvocates.org/article014.htm)