Query Wording
By Rick Brogden

In between the reporting of clinical data and the statistical analysis of that data, there are data queries, and lots of them. Data queries ask questions like, “Please verify that the result you recorded on the Case Report Form (CRF) is the result you intended to record.” They might explain that a respiratory rate is “out of range.” Queries sometimes do not say enough for the recipient to understand what is being asked, or conversely, say so much that the recipient can change the data to resolve the issue without even checking the source documents.

Most of the swarm of queries comes from Data Management. Is this barrage of queries really necessary? To answer this question, it is helpful to review Data Management’s role.

Take Me to the Bridge

Data Management is the bridge between the clinical and statistical sides of a clinical trial. This bridge has to be sturdy, stringent and clean enough to ensure that the reported data can be analyzed. It has to make sure that what is submitted for analysis is accurate, adheres to good clinical practice, and meets regulatory requirements and guidelines. The Data Management bridge ensures that the protocol is followed by all involved and, in cases where there are protocol violations — be they major or minor — communicates to the sites that what they reported is not what was expected.

Clinical programmers program the study database to generate queries when there are missing, inconsistent or illogical data points. Data Management and Clinical personnel may write other queries, based on the completed CRFs, the protocol, the regulatory rules, and the sponsor guidelines.

In paper CRF studies, auto-generated queries are usually reviewed by Data Management to detect patterns and minimize unnecessary queries before being sent to the sites. With Electronic Data Capture (EDC) systems, queries appear almost immediately after the data is entered and saved. The site thus has to do its best to understand what is being asked without the help of Data Management “pre-review” it may have previously received in paper CRF studies.

Site monitors are writing more and more of the queries. Monitors may see EDC data that does not match the source documents or find values that require investigation. In the past, these queries took the form of “yellow stickies” on the paper CRF pages. In these cases, as with comparisons of Adverse Events (AEs) to Concomitant Medications or AEs to Medical History, monitors write manual queries to resolve or verify issues.

Query Wording Guidelines

Since the existence of a query means there is some confusion, it is helpful when a query communicates clearly, without adding to the confusion. Considering the importance of queries to statistical analysis and the various roles of the people that may be writing and reading queries, it is useful to have some guidelines for dealing with queries on a day-to-day basis.

A successful query is one that will be understood. When writing queries, ensure the person reading the query understands what the query is asking.
A query is sent to a site to address an issue with a specific data point or set of data points, perhaps out of thousands of data points that were reported. Guideline one for writing proper queries is thus:

**Guideline 1.** Tell the site what it reported.

Begin queries by explaining there are data points the site needs to review, where those data points are, and what is recorded as the values for those data points.

- Systolic Blood Pressure is reported as 76.
- Date of Birth is reported as 12Dec1902.

In these examples, the site is directed to exactly what needs to be reviewed.

Once pointed in the right direction, the query recipient should now be directed to what the issue is with the data point(s) in question. Explain why the query was sent. So guideline two must be:

**Guideline 2.** Tell the site what is wrong with what it reported.

Continuing with the previous two examples, our first two query writing guidelines could now read something like:

- Systolic Blood Pressure is reported as 76; however, this is out of the acceptable range.
- Date of Birth is reported as 12Dec1902; however, per Inclusion Criteria, subjects must be between the ages of 18 and 65 to be included in this study.

These examples show what was reported and state the issues with the reported value.

The next step is to explain how to address the issue. Addressing the issue could be as simple as verifying the data is correct as reported, or it could require changing what is reported to adhere to study and/or regulatory guidelines. So guideline three must be:

**Guideline 3.** Ask the site to correct or verify what it reported.

The completed queries may now read:

- Systolic Blood Pressure is reported as 76; however, this is out of the acceptable range. Please correct or verify the value as reported.

(This query indicates what was reported (SBP = 76), what the issue is (out of range), and asks the site to review and correct or verify the value.)

- Date of Birth is reported as 12Dec1902; however, per Inclusion Criteria, subjects must be between the ages of 18 and 65 to be included in this study. Please revise the Date of Birth or provide a reason for the out-of-range value.

(This query describes what was reported (Date of Birth = 12Dec1902), what the issue is, and what to review and correct or clarify.)

In certain cases, additional information could be helpful, for example:

- Respiratory Rate is reported as 22 at Week 4; however, this is out of the specified acceptable range (12-20). Please verify the Respiratory Rate or provide the correct value.

Here the site was to guided to Week 4, where the out-of-range value was reported. If working on a paper study, the page number could be useful information. If there are dates, as in the case of Adverse Events, Concomitant Medications, Medical History, and the like, these dates could be helpful in the query text as well. Give the query recipient what he or
she needs, without giving too little or too much information. Keep it concise, guide the
query recipient to what needs review, state what the issue is, and then ask that the data
point(s) be verified or corrected as appropriate.

There is a fine line between asking the site to resolve an issue by correcting or verifying
what was reported and telling them how to do it. Jean Toth-Allen, Office of Good Clinical
Practice, Office of the Commissioner, states that “it would be inappropriate for a monitor to
provide a ‘leading’ suggestion when identifying a questionable value on a case report form
or in the source documents. Such a suggestion could be construed as telling the site to
substitute a specific value, even if it is not factual.”

On the other hand, not suggesting the nature of the problem can baffle the site. (Of course,
onece the site understands the issue, it can still arbitrarily change the data to make the
query go away.) The best approach is to inform the site of the nature of the problem, but
not to tell the site too much about how to resolve it. Guideline 4 is thus:

**Guideline 4.** Do not suggest to the site what to report.

The following queries are thus inappropriate:

- Outcome for Headache is blank and should be “Resolved.” Please correct.
- Number of tablets returned does not match source. Please change current value of 9
to 7.

By changing these queries a bit, we can make them fall within the guidelines:

- Outcome for Headache is blank. Please enter a value.
- Number of tablets returned is reported as 9. This value does not match source docs.
  Please revise or explain the value.

**Other Factors to Consider**

Unfortunately, query wording is not always this simple. For example, an Adverse Event was
once reported that was really two symptoms reported as one event. The symptoms were
marked as being serious. Coding could not be done until the two terms were split. So a
query was sent asking the site to split the symptoms into two events, and the site agreed.
Now, there were two serious AEs in the database. When the data was reconciled with the
safety data, Data Management had two terms, Safety had one, and nothing matched —
except the dates.

Everything in query writing is not going to be simple, and this instance is a good example of
why. Other things that may need to be considered could even include the English language.
In global studies, English is usually the language of the study but usually is not the first
language of everyone involved in the study. Understanding what is written in English can be
challenging, for example:

- The bandage was wound around the wound.
- His attempt to pierce the vein was in vain.

In the case of one global study, Data Management was trying to help sites whose primary
language was not English, while still following the guidelines for query writing. Data
Management pointed out the various data points that needed to be reviewed, stated what
the issue was, and then concluded the query with something like, “Since point A is this and
point B is that, please advise if…”

Data Management found that some sites were not responding to the queries. Why? These
sites were confusing the word “since” with the word “sense.” Once the problem was
identified, the word “because” was substituted for the word “since” and sites began to “sense” what issue(s) were being addressed in the queries.

The English language is funny. There are homonyms, homophones, acronyms and an ever-increasing melting pot of words. The global nature of modern clinical studies when accompanied by the English language means not only thinking about the issue being addressed in the query, but also about how words may be translated or “heard.”

**Conclusion**

Writing a helpful query is not always easy, but the simple guidelines addressed here, along with thinking about the nature of the query recipient(s), will help in most instances.

Queries are a means of communicating issues that could interfere with the statistical analysis of the data being collected. If the wording of queries is not effective, then communication will not be effective either. By using the guidelines set forth in this article, communication will improve, as will the ability to assist in providing better health options to the people who need them.

**Author**

Rick Brogden is a Clinical Data Associate III and former Training Specialist at INC Research. Contact him at 1.919.334.3633 or rbrogden@incresearch.com.