How Do We Define Clinical Trial Data Quality if No Guidelines Exist?

By Moe Alsumidaie and Artem Andrianov
May 2015

Many of us speak about the importance of clinical trial data quality and integrity, yet the lack of data quality standards and definitions introduces subjectivity risk in clinical trials. For example:

- A **clinical trial monitor/ CRA** interprets data quality as verifying source documents and visit procedures.
- To a **data manager** and **bio-statistician**, data quality can mean deviations in the data.
- To a **quality manager**, data quality could translate to not following SOPs.
- To a **study director** overseeing global clinical trials, data quality can mean fraud, patient safety and subsequent delays in study completion.

Variances in data quality reduce the statistical power of the sample size, which results in needing to enroll more patients, and delaying clinical trial completion timelines. In this article, we will go through common data quality definitions and offer recommendations on addressing the varying data quality facets.

Data Quality: "Fit for Purpose"

When it comes to data quality, fit for purpose models depend on the data strategy of the study. Fit for purpose methodologies suggest that data quality improves as data collection strategies become more targeted towards the study's objectives (i.e., critical data points). To elaborate, incorporating additional data points that have nothing to do with a protocol's endpoints (non-critical data points) not only introduces risk during endpoint analysis [1, 2], but also exhausts project management resources on verifying the quality of non-critical data. According to the EMA, quality ultimately depends on the measured variable, strong statistical power, acceptable errors, and clarity for clinical effects [3].

Albeit 'fit for purpose' study design may advocate that leaner protocols exhibit characteristics of improved quality, it is important to emphasize that study teams must also engage the commercial and health economics groups to incorporate and collect data that can be used to support payer submissions, and commercial activities.
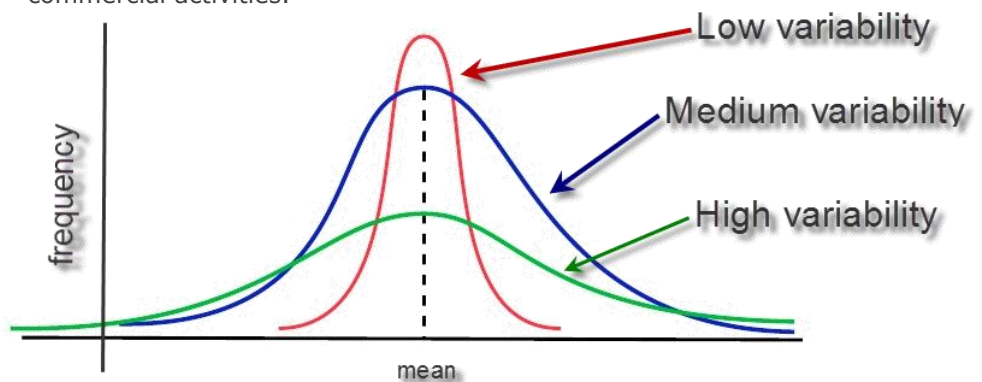

Figure 1: Impact of data variability on statistical power

Data Quality = Lower Variability

It is widely known that variances in data quality reduce the statistical power of the sample size, which results in enrolling more patients, and delaying clinical trial completion timelines. Unluckily, real-life data is dirty; it is inconsistent, duplicated, inaccurate, and out of date. All these data defects contribute towards data variability, which lowers statistical power. Figure 1 illustrates data variability and its impact on statistical power.

Figure 1 demonstrates three scenarios involved with variability as it relates to data quality. Lower data quality exhibit higher variability and lower confidence intervals compared to higher data quality, which yields higher confidence intervals. It is important to emphasize that study teams need to focus on improving data quality from all aspects; this can include patient nonadherence, variability with coordinator mandated measurements, source data quality, and data collection methodologies. The ultimate benefit of improving data quality involves the notion that study teams will need to enroll less patients because of sufficient statistical power.

Data Quality = Consistency + Integrity

Data consistency refers to the validity of the data source in which it is stored, and data integrity attributes towards the accuracy of the data that is stored within the data source. The traditional source model is a good example that demonstrates how a process can introduce risks towards data consistency and integrity.

To elaborate, from a data integrity standpoint, a coordinator may misinterpret and incorrectly record medical measurements on a paper CRF; without automated validations, a coordinator will not know they made an error until they input the data into EDC. From a consistency perspective, paper source introduces all kinds of risks; if the coordinator, for example, loses the paper source, there is no validity towards the data in EDC, and paper source can be modified without validated tracking systems (i.e., erasing or rewording measurements and reproducing paper CRFs from memory). eSource is known to significantly improve both data consistency and integrity.

The Polaris Perspective on Addressing Data Quality

Stan Woollen, FDA's Bioresearch Monitoring (BIMO) Program Coordinator, offered a data quality validation process - ALCOA, which stands for **Attributable, Legible, Contemporaneous, Original,** and **Accurate**. These characteristics are universal for detecting data defects.

Lauren Kelley, the Associate Director of GCP compliance with Polaris Compliance Consultants is an experienced auditor and study

monitor. To illustrate the concept of data quality, she likes to share the following example.

On a routine monitoring visit, Lauren once discovered that the source data for a subject's study visit was entirely missing. Naturally, she reported this to the site. On a subsequent visit, Lauren noted that all the missing visit data appeared in the subject's chart. Unfortunately, she also learned that the sub-investigator completed the page weeks after the subject visit – with data that included vital signs – from memory.

To disguise the omission, the sub-investigator had also backdated the record to reflect the date of the subject visit instead of the actual date of data completion. That she did all this **L**egibly was her only homage to ALCOA. The backdated entry violated one of the tenets of proper **A**ttribution. The intervening weeks between the subject visit and the data recording meant that the source was not **C**ontemporaneous. Was the data she entered at least **O**riginal? In the sense that she probably invented it, perhaps, but its lack of originality in the *regulatory* sense was clear when she explained, "I don't know where it is. I know I recorded it, but it is probably in somebody else's records by mistake." And what are the chances that the data were **A**ccurate?

When Lauren is asked to talk to study teams about mitigating data quality risks and to site personnel about minimizing inspection findings, she advises people to stick with what is fundamental: ALCOA.

How to build a "data quality culture" within an organization?

There is no regulation on data quality; each company has to define its own standards. We imply that improving data quality can not only generate better results, but, also minimize the amount of subjects needed in a clinical trial (therefore, costs) and reduce

timeline slippage. Below are six simple steps to improve data quality strategy within organization:

Step 1. Design leaner protocols that minimize unnecessary data collection.

Step 2. Define the data quality strategy, and data governance. Develop an implementation plan.

Step 3. Assign the roles and responsibilities for data governance.

Step 4. Develop policies, data quality metrics, ongoing monitoring and reporting through existing Risk-based Monitoring (RbM) technologies. For this purpose, a number of emerging technologies are appearing on the market (e.g., Cyntegrity).

Step 5. Using data quality tools and technology to monitor the trends.

Step 6. Build data standards program applicable for all clinical trials and apply standard metrics.

Data Quality Needs to be Defined and Managed

The subjectivity of data quality introduces many risks and variability towards clinical trials, especially when it comes to improving clinical trial outcomes, minimizing the amount of subjects needed to maintain statistical validity, and sustaining data quality and integrity. Clearly defining and categorizing data quality is the first step towards better managing clinical trial outcomes. Correspondingly, study teams can better leverage RbM technologies to not only better manage data quality, but also, reduce clinical trial costs, and enhance clinical trial results.

References:

[1] EMA, "Reflection paper on risk based quality management in clinical trials," 18-Nov-2013. [Online]. Available: http://www.ema.europa.eu/docs/en_GB /document_library/Scientific_guideline/2 011/08/WC500110059.pdf.

[2] L. M. Friedman, C. D. Furberg, and D. DeMets, Fundamentals of Clinical Trials. Springer Science & Business Media, 2010. pp 212

[3] Institute of Medicine (IOM), Davis Jr, Nolan VP, Woodcock J, Estabrook RW.eds. Assuring Data Quality and Validity in Clinical Trials for Regulatory Decision Making. Workshop Report. Roundtable on Research and Development of Drugs, Biologics, and Medical Devices, Division of Health Sciences Policy, Institute of Medicine. Washington DC: National Academy Press, 1999

Contact Details

Cyntegrity Germany GmbH
Altenhöferallee 3
60438 Frankfurt am Main

Tel: +49 6192 - 470 - 113 - 50
Fax: +49 6192 - 470 - 113 - 59

Artem Andrianov, PhD
Managing Director

Email: post@cyntegrity.com
Web: www.cyntegrity.com