



One version of true: Improving Performance Metrics in Clinical Trials

by Mark Donaghy, 2010

Providing a high level of transparency and penetration into study performance as measured using defined and consistently applied metrics enables the different participants in the study project management framework to grasp a quick view of the study's progress, get an understanding of the most critical performance issues, and to take action to isolate and eliminate the sources of poor performance quickly and efficiently.

Industry experts report that phase I-IV clinical research accounts for over 35% of a drugs development costs and is the source of some of many significant delays.¹ In 2006 Applied Clinical Trials reported that 86% of trials experience delays², and 94% of delayed studies were found to have delays of greater than one month³. In its 2010 report the Manhattan Institute stated that the monetary value that patients assigned to the benefits of access to the three drugs studied one year earlier would have equalled \$27.3 billion, while increasing producer profits by \$4.9 billion.⁴ In our experience in 25 years of clinical trials, many delays in a trial can be avoided through early detection and resolution of issues before they influence the majority of the trial.

Increasingly in the past three years sponsor companies have been turning to performance metrics in their drug development programs as a way to identify and reduce delays and operational issues. Doing so offers several benefits: project managers get a glimpse of the programs progress well before a final verdict is pronounced and the soundness of project plan has become moot. Trial sites can receive better information on the specific actions needed to achieve strategic objectives, and investors have a better sense of the programs overall performance, since operational indicators usually reflect realms of tangible and intangible value.

The reality is that few companies realize the benefits of operational performance measurement. Why? Because they fail to set the correct objectives and to identify, measure, analyze and act on the right operational performance metrics. In many cases measurements are made, but the cause and effect link between actions and improvements has not been demonstrated. In other cases action does not occur when required. Finally, it is frequently observed that the very process of measuring and analyzing

operational metrics is inconsistently executed and the resulting actions are unwarranted or even damaging to the study.

The purpose of this article is to review some of the common failings in developing and implementing effective operational metrics programs. Future articles will discuss deciding how to pick which metrics provide the greatest insight into your development program, and the use of metrics to identify and resolve issues before they become critical problems.

Building in metrics during study design: Plan - Do - Check - Act

Performance metrics are the project management version of the data and analysis in a longitudinal clinical study. Measurements of key indicators of project 'health' are taken at predetermined intervals, and conclusions drawn on actions required to maintain the 'health' of the study.

Experienced clinical trial designers recognize the importance of building a complete clinical trial design prior to starting any major activities within the trial. This should include the designing of the metrics to monitor the program.

Performance metrics programs produce the best results when established in the design phase of a clinical study. Application of the Deming Cycle (Plan – Do – Check - Act) encourages a planned, systematic, and explicit alignment of study endpoints with objectivity in performance data collection and analysis. At the beginning of the study additional benefits can be realized such as early alignment of specific roles to how the results will be measured with the metrics, alignment and analysis of performance standards to other similar projects to determine what produced successful results and what produced subtle results which were contrary to the intent of the performance metric program. In addition, with all stakeholders signed off in advance, the results of the metric program are less subject to agency issues and interpretive art.

Focus on what matters

Focus is critical to any metrics program. Without focus a metrics program will rapidly begin to measure many areas which are not on the critical path supporting the end-points selected as objectives. Clear sight of the goal is lost, and the metrics program itself becomes obfuscated and inefficient.

In measuring performance, the project manager is trying to gather information to help make management decisions to affect change that, hopefully, will improve that performance. For example, project performance measures may be used to identify specific CRF fields that create an abnormal number of queries and use substantial amounts of monitor and CDM time.

What are we to measure? Why is it significant? The metrics must be appropriate to the organizational level that can immediately effect change based on information it learns in order to control the performance of the project and to allow managers to make critical decisions to bring the project to closure successfully. A metrics program needs to have clear end-points that provide the purpose of the

process. These measures must be collected frequently, perhaps even weekly, depending on the duration of the project.

Many companies select performance metrics on the basis of the cause and effect relationships which exist between project performance objectives, and the activities mapped to those objectives. Common milestones in clinical trial project management include:

<ul style="list-style-type: none"> • Project progress. <p>Is the purpose of the metrics to</p> <ol style="list-style-type: none"> a. measure the efficiency of specific activities, or b. is it to monitor overall progress to provide early warning signs that the project plan is moving off track? 	<ul style="list-style-type: none"> • General indicator of whether the project will end as scheduled, and an indicator of the number of months of general overhead which may be associated with delays.
<ul style="list-style-type: none"> • Data quality metrics and data completeness? 	<ul style="list-style-type: none"> • Issues with useable data often result in larger numbers of subjects being required, longer database lock periods and issues in data analysis
<ul style="list-style-type: none"> • Resource efficiency (drug stockpiles, etc) 	<ul style="list-style-type: none"> • Inefficient use of resources may result in drug spoilage or waste at the end of a study

Similar to any study design, clear endpoints makes the selection of the variables (metrics) to be measured simpler and more effective. A common breakdown point in the performance metrics program is to rely on preconceptions of what activities have the cause and effect relationship to the endpoint without testing them. In an unstructured metrics program it has been the habit to select metrics based on the informal understanding of clinical trials. While inherent knowledge is useful, the additional value that a thoughtful and process driven approach adds to a performance metric program is considerable if only to assess the total metric environment before the final group of metrics optimized.

Most sub-processes have more than one measurable point that will provide data on how that sub-process is proceeding. The different points of measurement may support the different objectives for the metrics program. Usually, however, there are critical points in the process that provide a more powerful indication of how the process is proceeding.

While it is possible to apply standard company performance metrics, project managers are aware that the concept of one-size-fits-all works about as well in performance metrics as its does in clinical studies. A more customized set of performance metrics and trigger points may provide better information.

Each metric needs to map to an end-point. It needs to be key information in the longitudinal study of the health of your drug development program. The value of an individual performance measure is

limited, however, but combined, the measures provide a robust assessment of the quality of project management for individual projects and programs. If applied consistently over time and used for internal and external benchmarking, the measures will provide the information needed for day-to-day management and long-term process improvement.

Consistency

Consistency of metrics within a study is critical if trend analysis is to be effective, and there is an effect on absolute value analysis. In both cases, low adherence will result in a higher number of false positives and negatives in the analysis of the metrics. The associated risks are that the call-to-action may not occur when it is needed due to a false negative resulting in a late response to issues, or may occur when not required resulting in the misuse of scarce project resources and some relationship damage.

In its most basic form, consistency in metrics refers to data values measured at a point in time being consistent with values measured at other points in time. A more complicated notion of consistency includes consistency between studies or trial sites, between different cohorts of trial sites with different start dates, and with sets of predefined constraints such as interval of metric measurement or event based measurement triggers.

More formal consistency constraints can be formed as a set of rules that support the validity and reliability relationships between performance metric values, either across an episode metric value set, or along all values of a single metric. However, be careful not to confuse consistency with accuracy or correctness.

Consistency may be defined within different contexts:

- Between one set of attribute values and another attribute set within the same episode metric set (record level consistency) For example, within the CRF pages from one trial visit.
- Between one set of metric values and another similar metric value in different records (cross-record consistency) Between CRF pages from more than one subject, but for the same point in the visit schedule.
- Between one set of metric values and the same metric values within the same record at different points in time (temporal consistency) Same subject, at different points in the same treatment schedule.
- Consistency may also take into account the concept of “reasonableness,” in which some range of acceptability is imposed on the values of a set of metrics.

Consistency in performance metric programs are best developed by establishing an analysis framework early in the clinical study with sufficient detail on how it will be executed to address common questions on when and how each metric will be calculated in a process very similar to defining how variables will

be captured when preparing CRFs. As a result eDC platforms are very strong tools to collect a significant amount of performance data, but at the cost of setting up the eDC to collect and report it.

Comparability

For performance measures to have meaning and provide useful information, it is necessary to establish comparisons. The comparisons may evaluate progress in achieving given goals or targets, assess trends in performance over time, or weigh the performance of one organization against another.

In some industries metrics can be designed without the need to be directly comparable to any other point in the metric scale. These metrics are used by highly experienced people in a relatively homogeneous industry who have implemented metric systems for many studies and have developed for a sense of when the metric is beginning to indicate trouble for the underlying project. In general, however, even highly experienced people have developed ways to make sure that their metrics are comparable to some reference point so as to be useful in building a common framework and supporting their discussions with other stakeholders and to persuade them that action is required.

Reference points for comparability are found in two forms: absolute reference points and trending reference points. An absolute reference point is one where a line has been drawn in the sand, and if the measured metric crosses that line it calls for an action. Trending reference points float and may be, for example triggered by a secondary attribute of the measured metric. Rate of change between two absolute reference points accelerating or decelerating may be the trigger.

Assessment

Part of the process of performance metric management for proactive monitoring deals with establishing the relationship between recognized data flaws and business impacts. In order to do this, one must first be able to distinguish between “good” and “bad” data. The attempt to qualify performance metrics is a process of analysis and discovery involving an objective review of the measured data values through quantitative measures and analyst review and establishing thresholds that signal an event has occurred. For example, if recruitment at a particular site is behind by a specified amount, the project manager should make the proper notifications and start planning for how the deficiency will be recovered.

It is typical for metrics to be designed and measured without taking the key step of defining at what points the metrics trips a threshold. While a data analyst may not necessarily be able to pinpoint all the thresholds, the ability to document situations where metric data values look like they warrant attention provides a means to communicate these instances with subject matter experts whose business knowledge can confirm the existences of problems.

Setting the trigger point for a call to action requires an understanding of the underlying activities.

Call to action

A very common failing in performance metric systems is to fail to specify when a threshold has been breached, and what actions need to occur in the event a performance metric breaches a threshold. Who is to be informed, and what activities need to happen. How urgently do those activities need to be executed?

Conclusion

Similar to clinical studies, the success or failure of a performance metrics program is often predetermined by the design of the program at the outset of the study. In fact, the success of an underlying study may be inhibited by a badly designed set of performance metrics that encourages counter-productive behaviours.

The success of performance metrics has been clearly demonstrated in other industries. There, comparative analysis of competing processes has allowed project managers to select in or out activities that advance the whole organization towards success, and to control their processes more effectively.

Life sciences and biotechnology are currently under a tremendous resource strain, and investors are looking for new ways to decide which companies they should support. While the potential of the underlying drug is of critical importance, so is the ability to demonstrate that the development team can effectively bring the development project to a close efficiently and performance metrics should be part of that story.

¹ Crowner, Robert, "Drug Pricing vs. Drug Development," Action Institute for the Study of Religion and Liberty, May 2002.

² Canavan, Carl, "Integrating Recruitment into eHealth Patient Records," Applied Clinical Trials, June 2006.

³ Weinberg, Myrl, "Engaging the Public in Clinical Research," Regulatory Affairs Focus, July 2006.

⁴ Philipson, Tomas, and Sun, Eric, "Cost of Caution: The Impact On Patients of Delayed Drug Approvals," Manhattan Institute, June 2010.