Paper 171-2012

Applying Business Analytics to Optimize Clinical Research Operations

Dave Handelsman, SAS Institute Inc., Cary, NC

ABSTRACT

SAS® is widely accepted as the gold standard for determining safety and efficacy for clinical trials, and it provides the primary mechanism for preparing data for traditional clinical research analysis activities. However, most SAS users in the biopharmaceutical industry are unaware of the broad range of SAS analytics that are widely applied in other industries. This paper discusses and describes how SAS business and advanced analytics can be used to design better trials, forecast patient-based activities, and optimize other operational processes.

Applying business and advanced analytics to clinical trial operations represents a new and improved approach to reducing the cost and time associated with managing clinical research projects. As a result, the roles of SAS experts in the biopharmaceutical industry are expanded.

INTRODUCTION

Bringing a new drug to market is an expensive proposition—the cost is often estimated at $1.2 billion. This investment includes the entirety of work, from discovery in the lab to approval by the national regulatory agencies. The investment varies, based on therapeutic areas, geographies, and a myriad of other factors. What does not vary, however, is that the expense is extraordinarily high, and ongoing investments of this magnitude are not sustainable.

The biopharmaceutical industry is undergoing a widespread revolution. The era of the blockbuster drug is rapidly fading. Many high-revenue-producing drugs are about to have expired patents, and the pipeline to replace these drugs and their associated revenue is not encouraging.

To compound the issues facing the biopharmaceutical industry, private and government health insurers are carefully evaluating their reimbursement policies to ensure that their payments are being well spent. Health plans are defining reimbursement strategies that pay only for drugs with differentiated proven effectiveness over existing and frequently less expensive treatments. Additionally, the US government has allocated over $1 billion to fund comparative effectiveness research to provide evidence about which therapies perform better.

All of these factors are contributing to the clear message that “the good old days of the pharmaceutical industry are gone forever.” Companies must identify ways to work efficiently and effectively to ensure that their investment dollars are well spent.

Of the $1.2 billion estimated for each research program, several hundred million dollars are allocated just for clinical trials. The cost for each clinical trial can be as high as tens of millions of dollars, sometimes even higher. With research programs this size, it is easy to see why project execution must be efficient. However, most manufacturers apply only basic tools to management.

SAS, with its rich set of advanced analytics tools for business analytics, provides an ideal means to bring rigor to the decision-making processes and management of clinical research projects. Although SAS has been widely accepted as the gold standard for providing statistical capabilities to determine the safety and efficacy of individual and integrated clinical trials, and although SAS is frequently the tool of choice in clinical trial data quality and transformation activities, only a few SAS analytical tools have historically been applied during the clinical trials execution process. A recent search of the SAS Web site yielded fewer than 1,000 hits for variations of the term *biostatistics*, and it yielded almost 20,000 hits for *business analytics*. There is an enormous opportunity for applying SAS business analytic capabilities to the execution of clinical trials.

As shown in Figure 1, there are several key areas where SAS analytical capabilities can bring significant efficiencies to the clinical research process. Each of the key areas in Figure 1 describes business processes that are associated with clinical research operational activities. The operational activities typically occur much earlier in the process than classical biostatistics activities, and operational activities use limited, if any, analytics-based decision-making efforts. Instead, decisions about operational activities are most typically made using unmanaged and manually populated spreadsheets, and they are guided more by past experiences than hard data.

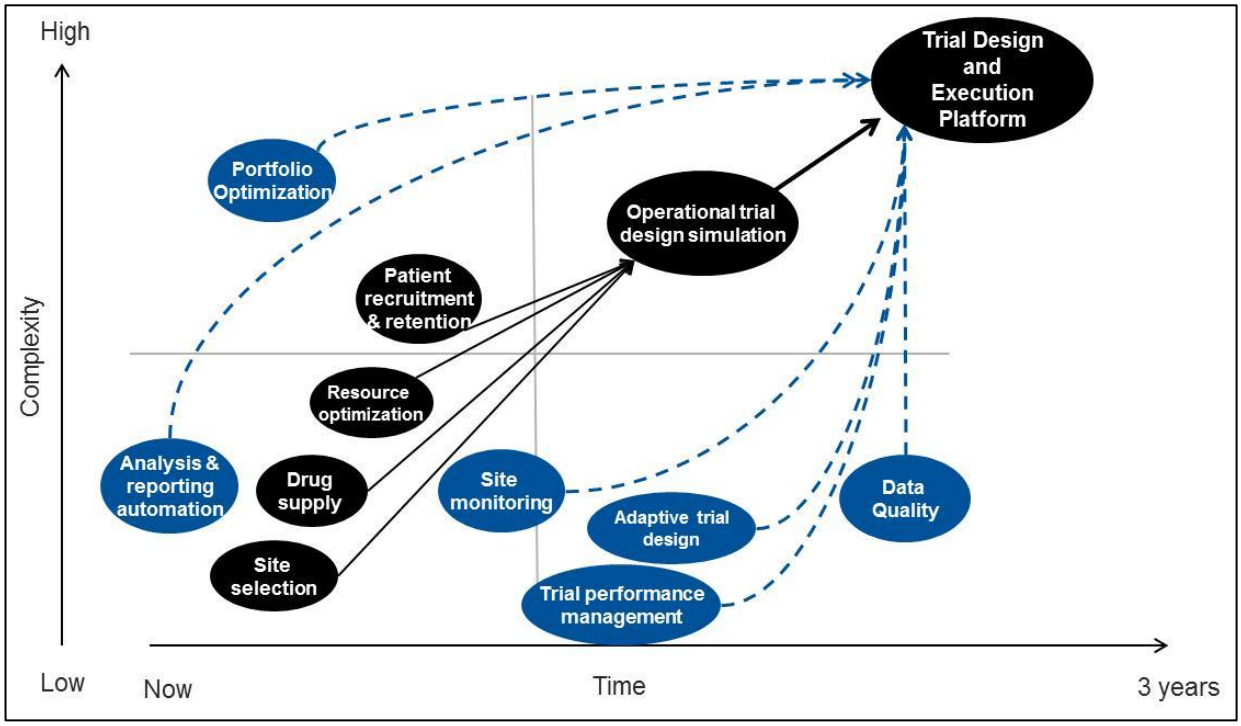

Figure 1. Key Areas for SAS Business Analytics in Clinical Trials Execution and Optimization

In many cases, there are direct parallels between specific clinical trial processes and business processes in other industries where SAS is used to create informed and optimized decisions. This paper addresses each key area in Figure 1, describes the business process involved, identifies how analytics can help, and uses examples from other industries where appropriate.

However, analytics will not be successful when executed in a vacuum. Organizations must invest in data preparation so that analytics can be successfully applied. They must invest in personnel who can lead business process initiatives. As with all change, the keys to success are people, process, and technology. This paper provides the foundation to successfully change the process and to enable the biopharmaceutical industry to develop new therapies with optimized operational activities.

OPERATIONAL TRIAL DESIGN SIMULATION

For years, clinical trials have typically been designed using traditional biostatistical techniques. Although experimentation, by its very nature, means that not all expected results will be achieved, the business needs of clinical trials necessitate a more robust approach. A poorly designed trial has many ramifications beyond failing to prove the desired endpoints. It can bring entire research programs to a halt because expected safety or efficacy is lacking. There are the additional considerations of cost, ethical treatment of patients, and wasted resources.

Trials can be more successfully designed by applying analytics to historical data to assess the likelihood of the trial's success in terms of operational and statistical outcomes. In some cases, a trial might be designed optimally for demonstrating safety or efficacy. But, this same trial is impossible to execute because of restrictive inclusion and exclusion criteria. On the other hand, a trial might be easy to execute, but impossible to meet its scientific objective.

Although all trials are designed with goals of scientific and operational rigor, life sciences research organizations are beginning to more broadly apply simulations to the design process. With simulations, different designs are defined, and then a series of trials are simulated with advanced analytical software. These simulations identify parameters that will result in the most optimally designed trial. A prototype of a clinical trial design simulation tool is shown in Figure 2. Figure 2 shows the outcome of running multiple simulations, and then selecting the most highly rated simulation. In this case, the results include the expected mean trial length and cost and the standard deviation associated with these metrics.

Figure 2. Clinical Trial Design Simulation Prototype

Rigorous trial design at a macro level can be further refined by developing more accurate modeling inputs to the design process. These additional enabling models include patient retention and recruitment, resource optimization, site selection, and drug supply. Importantly, these methodologies can also function independently as part of the overall trial execution process, even if data-driven designs are not being used.

PATIENT RECRUITMENT AND RETENTION

Patient recruitment is widely recognized as the single biggest bottleneck of the clinical trials process. In a CenterWatch 2007 survey, fewer than 7% of US sites reported meeting their enrollment timelines. As patient enrollment timelines slide, additional sites might be initiated. This creates increased and unexpected costs. Final delivery milestones are frequently missed and cost overruns are common.

Confounding the problem of delayed patient recruitment is over-recruitment to address patient-retention concerns. For a clinical trial, an assessment is made to determine the number of patients required to meet the trial's statistical endpoints. Because the trial manager expects some patients to fail to complete the trial (ideally, this number is a quantitative assessment based on previous trials), the overall patient recruitment goal is inflated. In effect, research companies are compounding the problem of missed patient recruitment goals by adding patients to these existing goals, which all but ensures delays in determining the outcomes of the trials.

Applying patient-retention processes results in more patients completing the trial. In addition, it reduces the number of patients that need to be enrolled for the trial outcomes to be statistically valid. Retention processes are widely practiced in other industries, especially where the focus is on customer retention. In other industries, companies look for trends in their existing customers' behaviors (for example, fewer clicks on retail Web sites or a diminished use of a long-distance calling plan). Then, these companies intervene to retain the customers' business (for example, offer a targeted discount). In clinical trials, intervention might take the form of increased contact from the investigator or another similar activity to retain the patient's participation in the clinical trial.

The parallels between customer retention and patient retention are clear. It costs much less to retain a customer or patient than it costs to recruit one. By applying patient-retention processes using existing and common tools available to other industries, patient recruitment can be managed more effectively. For example, Earthlink (http://www.sas.com/success/earthlink.html) implemented customer-retention capabilities based in SAS "to identify customers who might be thinking of leaving so it can […] convince customers to stay 'Linked.'" Similarly, strategies should be deployed to identify patients so that their trial participation can be continued.

RESOURCE OPTIMIZATION

At its core, resource optimization is having sufficient resources to complete a measurable task. In all organizations, but especially in larger organizations, the ability to assess, understand, and implement resource allocation is largely left to spreadsheets and whiteboards. Excess capacity in one part of the organization goes unused. At the same time, external contractors are brought in to complete a task that could readily be completed by existing resources. For clinical trials with complex project schedules, frequent and unexpected delays, and nontrivial employee turnover, spreadsheets and whiteboards are simply insufficient tools. Instead, a fluid, data-driven approach must take into account the planned project schedule and other competing projects. It must accommodate the frequent changes that occur during the clinical-execution process.

For resource management, the most compelling area to look at is clinical research associate (CRA) staffing and site-monitoring visits. Most research organizations plan their staffing needs around a linear work timeline. This timeline includes tasks such as site selection, site initiation, patient recruitment, and so on. However, work rarely occurs in a linear fashion, and its forecasted completion is directly related to site selection and patient recruitment.

If CRA staffing for site-monitoring visits is based on anticipated workload, it becomes very important for that workload to be accurately forecast. Instead, organizations typically rely on an approach of *x* patients per month for recruitment, although historical models almost always indicate that patient enrollment is not nearly linear. Based on this false assumption, it becomes impossible to optimally staff and train the CRAs because the anticipated enrollment, which is the ultimate driver for site-monitoring visits, lags the linearly forecasted enrollment. If you look at a basic example as shown in Figure 3, the number of idle CRAs initially assigned to a project continues to grow as enrollment lags. These idle CRAs represent wasted resources who could have been allocated to other projects. Although it is possible to reassign the CRAs, there are significant costs involved because CRAs must be fully trained on any new project in accordance with organizational standard operating procedures. A more analytics-based approach to this resource problem includes the use of accurate forecasting analytical tools to better anticipate model-based patient recruitment.

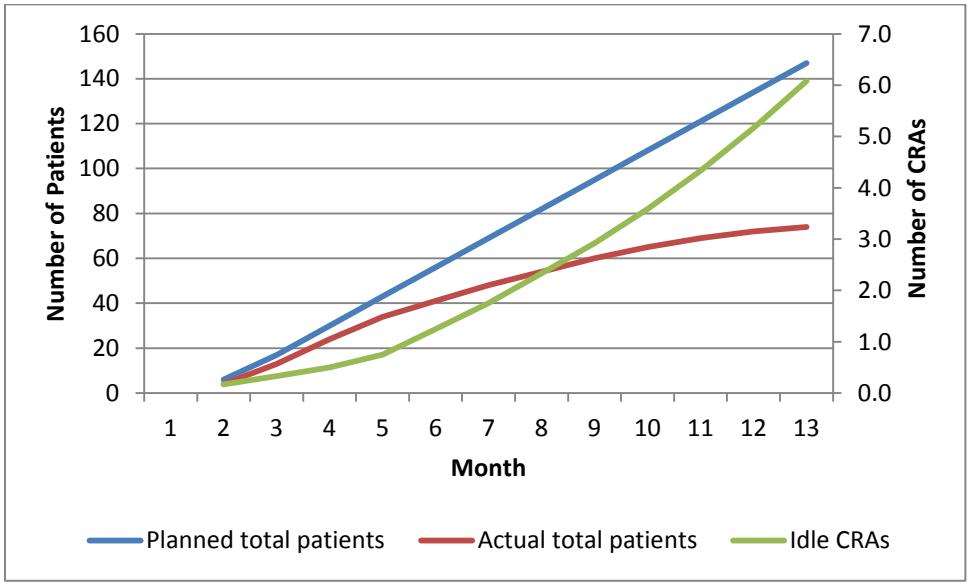

Figure 3. Planned versus Actual Patient Enrollment

In many ways, this problem is similar to the supply chain issues associated with how manufacturers manage their supply inventories against sales forecasts. Here, the goal is to have just enough inventory to meet expected demand, with some calculated additional inventory to meet unexpected demand. In the case of AmBev (http://www.sas.com/success/ambev.html), this problem manifests itself in the task of delivering beverages for retailers to sell. There are multiple distribution sites, multiple destinations, and an ongoing manufacturing and delivery process. By applying business analytics, AmBev is able to provide the right inventory to the right destination at the right time. A similar strategy can be deployed for CRA or other departmental resources.

DRUG SUPPLY

As more analytically driven business processes are built and, specifically, as the industry deploys adaptive trials at an increasing pace, one of the key issues that must be addressed is clinical supply management. For every clinical trial, the dosing material must be manufactured, packaged, blinded appropriately, and distributed. This process is very important and ensures that designated treatments are available for sites to provide to patients. In addition, it ensures that treatments are available in the correct distribution to accommodate the expected distribution of patients at each site.

For traditional clinical trials, this process is initially straightforward because a drug is shipped based on spreadsheet-driven enrollment projections. However, these projections vary in their accuracy. More of the drug must be distributed to account for fluctuations in the projection, sites must return unused drugs for redistribution to other sites, or sites must dispose of unused drugs. More accurate enrollment projections enable more accurate clinical supply forecasting. In addition, it prevents over-distributing study medication and avoids resupply and waste issues. For longer-term trials, just-in-time distributions that are closely aligned with forecasted enrollment and site-monitoring visits provide an excellent way to optimize drug distribution.

On the other hand, adaptive trials present a different set of challenges. These trials typically define different dosing levels in the trial. They might require investigators to switch to a new dosing level based on analytical results described in the protocol. Because these dosing levels are determined only when the trial is in process, it requires a different approach to clinical supply. The clinical supply team must be able to model and forecast which doses are needed at which sites within some level of confidence to optimize the availability of the right dose at each site. The only other alternatives are to ship multiple dose packages to each site in anticipation of the various dosing possibilities, or to react when dosing regimens are changed. Neither of these alternatives is appealing. The former option requires significant expense to oversupply each site, and the latter option risks delays in supply distribution.

SITE SELECTION

There are many critical aspects of a clinical trial, but perhaps the most critical is site selection. (This is in terms of successful trial execution and operations.) Site selection provides the foundation to meet the project's timeline for patient recruitment. Ultimately, site selection enables the trial study report to meet its final delivery dates. In short, sites that are not able to recruit sufficient patients seriously impact the ability to complete the trial on time. These sites are sinkholes for trial budgets. The average sponsor cost to open and close an investigator site, regardless of whether any patients are enrolled, is estimated at $50,000. And, 80% of these sites enrolled one or fewer patients in Phase II and III in 2008 and 2009. The effects are compounded when additional sites need to be opened at $50,000 each to make up for the patient-recruitment shortfall from the previously selected sites.

Beyond the direct relationship between successful sites and patient-recruitment milestones, sites indirectly affect the trial's goals through cost and quality issues. Sites that collect poor-quality data (which requires additional site-monitoring visits and disproportionate numbers of data queries) increase the expense of the trial through additional labor costs. In addition, these sites potentially interfere with the ability to accurately calculate the outcome of the trial because of poor-quality data issues.

Sites are frequently selected because their physicians are identified as key opinion leaders. Key opinion leaders are considered important to the pharmaceutical company sponsoring the trial. These key opinion leaders are perceived to influence how other physicians would treat similar patients. This is the first step in marketing a drug under investigation.

Despite the many ways a research program can be affected (time, cost, and quality), very little effort is made to identity and select optimal research sites. Instead, pharmaceutical companies primarily rely on past recruitment performance. (For example, how many patients did the site recruit previously?) Site questionnaire responses that indicate an estimate (by the site) of its ability to recruit appropriate patients are frequently used as well. These basic measures, the latter of which is frequently a biased guess by the site, are limited in their ability to identify the *best* site because *best* is not typically quantified as part of the selection process.

Instead, the *best* or optimal site is determined by assessing data points (recruitment history, queries, requeries, calculated thought leadership, and so on), weighting these data points, and quantitatively determining which sites are, indeed, best. Clearly, a low-cost site that cannot recruit patients has little value. Similarly, a site that recruits many patients, but collects low-quality data at a high cost, is limited in value.

SAS has been applying analytics to solve business problems such as these for years, primarily with its solutions for supplier intelligence, which enable comprehensive sourcing capabilities. Just as it is not enough to select the lowest-cost supplier, it is not enough to select the best recruiter (historically). A comprehensive solution enables research organizations to apply analytics-based decisions when selecting the best research sites.

SITE MONITORING

Monitoring data and processes at the site is a critical aspect of clinical trials. In many cases, it is the costliest aspect of the trial. Historically, site monitoring has been managed on a calendar basis, with CRAs visiting sites without regard to the workload that is associated with that visit. Frequently, a site is visited when it has a comparatively low number of patients, or when little has changed since the last visit, or simply because a visit is due based on the calendar. This simplistic approach creates a situation that has all of the expenses of a site visit, with little of the value. And, in fact, it wastes resources that would be better used ensuring quality elsewhere within the clinical trial.

Two trends for site monitoring in clinical trials are emerging. The first trend is applying resourcing techniques to accurately forecast patient enrollment. Beyond the need to have the right number of CRAs resourced to a trial on an ongoing basis, there is additional value in having the CRA conduct site-monitoring visits when there is sufficient data at the site to justify the expense of the visit. Through accurate patient-enrollment forecasting and accurate patient-visit-schedule forecasting, it becomes straightforward to calculate the monitoring workload at each site. With accurate forecasting, CRAs can schedule site visits to coincide with a full-day's work of monitoring. By aligning the workload with the visits, efficient monitoring visits are the result. Each site visit is maximized in value.

The second trend is risk-based monitoring. With this approach, CRAs focus on sites with the greatest risks. CRAs provide less monitoring of sites with minimal risks. In this case, *greatest risk* refers to the likelihood that data quality issues at the site will impede, interfere, or impact the successful conclusion of the trial and potentially the entire research program. For sites that have historically good track records of performance and quality, monitoring visits are reduced. Sites with a higher risk profile receive commensurately increased monitoring visits.

In August 2011, the FDA issued draft guidance entitled *Oversight of Clinical Investigations — A Risk-Based Approach to Monitoring.* At the time this paper was written, that guidance had not yet reached final status, but it indicates that FDA recognizes the need to apply a risk-based methodology to this critical business process. Specifically, the guidance "recommends that each sponsor design a monitoring plan that is tailored to the specific human subject protection and data integrity risks of the trial"

By reviewing the past performance at a site, and building an analytical model that indicates the characteristics of sites that are most at risk for data quality issues, it becomes straightforward to implement a risk-based monitoring approach. Furthermore, it becomes practical to deploy risk-based monitoring so that the efforts to target monitoring are fully optimized. As a result, a predetermined quantitative risk profile for the trial as a whole is met.

ADAPTIVE TRIAL DESIGN

Adaptive trials have the potential to fundamentally change how clinical trials are executed. But, they face an uphill battle in overcoming perceived difficulties in the real world. In reality, many adaptive trial designs have been well proven over the years, and they are broadly accepted by industry and regulatory agencies. The main obstacle is concerns about the rigor of an adaptive trial.

The expected rigor, however, is no different from what would be expected when a traditional trial design is executed. Furthermore, the potential increase in overall trial efficiency is significant.

By designing trials that allow predefined operational changes based on real-world accumulated data, individual trials provide a significantly increased return on investment. Several trials can be combined into one, which dramatically reduces the total number of trials necessary for submission and approval. Executing fewer trials reduces the need for multiple rounds of enrollment, recruitment, data management, and analysis. It eliminates the white space between trials where iterative planning cycles are spent.

Adaptive trials continue to gain momentum in the industry and with the FDA, but they are not without risk. By applying a robust simulation process, you can determine the best approach in moving a research project forward, and you can reduce the risks of adaptive trials.

TRIAL PERFORMANCE MANAGEMENT

Various systems are deployed throughout the industry to measure and monitor clinical trial performance. Typically, data is captured and managed through clinical trial management systems, with Excel being the most frequent tool of choice for ongoing data capture and maintenance. The data associated with trial performance management is simply the means to an end. The goal is to provide actionable intelligence about the execution of clinical trials.

All too often, trial performance management systems provide only the most basic reporting information about product status, such as the number of sites initiated, the number of patients enrolled, and so on. These systems provide a reasonable accounting of what has happened to date. But, they do very little in terms of forecasting metrics about the completion of the trial. Advanced analytics provide a way to integrate available data from an ongoing trial with data from completed trials. The likely operational outcome of the trial is reported through advanced analytics. These likely outcomes include more sophisticated approaches than just straight-line projections for meeting budget and timeline metrics and can create confidence intervals and alerts.

DATA QUALITY

Data quality is a critical component to trial execution. But, more often than not, data quality is thought of only in the contexts of clinical data queries and corrections. In fact, in a clinical trial, there is a disproportionate amount of effort correcting each and every data point that is questionable. The reality is that data inherently has inconsistencies, and correct processes and methodologies must monitor and address these inconsistencies. It is not necessary to correct every inconsistency.

The issue of data quality is not and should not be limited to only clinical trial research data. Significant time and expense is spent manually reconciling trial operations data from multiple systems. This task is almost always manual, and it involves simple activities such as ensuring that investigator names and site numbers are consistent and accurate across different systems. An automated process would be a better approach. More importantly, common

operational data should be organized so that a single version of the truth is accessible to all downstream systems and users.

PORTFOLIO OPTIMIZATION

For biopharmaceutical organizations, selecting a portfolio has typically been performed using a lens that shows all promising therapies having a path to market approval. When financials are strong and investment is high, this view can certainly be accurate. However, as the biopharmaceutical industry continues to evolve, top-line revenue and bottom-line profitability are constantly threatened.

The critical idea behind portfolio optimization is that decisions must be made based on a risk and benefit approach. If there is limited funding, the right decisions must be made in terms of not only which treatments are truly likely to reach an approved state, but which treatments, once approved, support the revenue strategies of the organization.

Portfolio optimization can be thought of as protocol-selection optimization. The ability to select the best protocol design and the best protocol targets (in terms of limited time and funding) is crucial. If a trial has a low likelihood of meeting corporate objectives in terms of time, cost, or hypothesis, it should be shelved indefinitely.

AUTOMATED ANALYSIS AND REPORTING

For many years, SAS has been the analytics tool of choice to determine clinical trial outcomes. The manual on-site process is remarkably consistent among biopharmaceutical organizations. Each of these organizations should aspire to create an automated analysis and reporting environment for a metadata-driven and powerful SAS analytic solution.

To revamp the manual process, the organization should begin with a structured protocol that defines the trial at a granular level. The trial metadata is then used to acquire and apply the necessary analytical components. For example, the protocol would specify the data to be collected, the analytical techniques to be applied, the design of the trial, primary and secondary endpoints, and protocol violation rules. Trial metadata definitions would be machine-readable. They would automate the creation of the batch analytical programs to answer the research questions in the protocol. For example, if the protocol is amended to include a new protocol violation rule, the *new* process would automatically cascade the rule throughout the system.

There are several critical components that must be in place in order for this process to be automated. First, a machine-readable, granular-level protocol must be developed. Trial metadata needs to be merged with metadata from other systems. These other systems include analysis and reporting tools and available and expected data models. Validation algorithms need to be constructed to ensure that the true result of applying the metadata to the data makes logical and analytical sense.

Deploying an automated analysis and reporting system creates the ability to run a blinded or unblinded analysis at the proverbial push of a button. The analyses are designed and available at the earliest segment of the trial. Analyses can be executed as soon as sufficient data is available.

CONCLUSION

Ultimately, the individual components would be integrated into a comprehensive trial design and execution platform. This platform does not currently exist. Most biopharmaceutical organizations have implemented only a few components. In many cases, shared and migratory spreadsheets are the only tools.

As the life sciences research organizations wrestle with the evolving business and scientific landscape, analytics provide a way to gain control of their business processes. There is little doubt that the clinical trial process can be executed more efficiently. It is possible to streamline the process at multiple points.

The greatest potential for leveraging analytics is in applying simulations to trial design, and then executing the trial based on that optimal trial design. Many promising therapies are subjected to unsuccessful trials, and they result in wasted resources and delayed approvals. Trials must be designed to have the highest likelihood of scientific and economic success. When simulation indicates that an adaptive design should be chosen, the value of analytics is further multiplied as the adaptive trial frequently results in fewer trials, reduced costs, and accelerated time to approval.

At the execution level, analytics enable research companies to focus their efforts where they matter the most—accurate forecasts for planning and risk-based leveraging of resources. It cannot be business as usual if the life sciences research organizations want to continue to add to their long history of success.

REFERENCES

http://csdd.tufts.edu/research/research_milestones

Hunt, Vivian, Nigel Manson, and Paul Morgan. December 2011. "A wake-up call for Big Pharma." McKinsey Quarterly.

http://blogs.sas.com/content/hls/2011/08/11/all-hail-the-end-of-adaptive-clinical-trials

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Dave Handelsman
SAS Center for Health Analytics and Insights
SAS Campus Drive
Cary, NC 27513
Work Phone: 919-531-0303
E-mail: david.handelsman@sas.com
Web: http://www.sas.com/industry/life-sciences