

Using Real-time Data to Drive Better Decisions, Faster

Paula Brown Stafford, BS, MPH

President, Clinical Development, Quintiles, Durham, North Carolina

Andrew Garrett, PhD

Vice President, Global Biostatistics and Medical Writing, Quintiles, Reading, United Kingdom

The new health landscape requires biopharmaceutical companies to conduct smarter clinical trials and produce better outcomes faster. Electronic data capture and electronic data review tools enable drug developers to review data shortly after it is captured, creating opportunities for improving the process and outcomes of clinical trials, and proactively managing quality, patient safety, and risk. We examine the extraordinary opportunity for the biopharmaceutical industry to use real-time data to drive better decisions faster. We identify the advantages of access to real-time data during a clinical trial, while recommending guidelines for

mitigating associated risks. Topics discussed include the following:

- Technologies that make access to real-time clinical trial data widely available
- Regulatory implications of reviewing real-time data throughout clinical trials
- Clinical trial roles and the rationale for granting real-time data access
- Guidelines and controls for use of real-time data during clinical trials
- How real-time data enhance patient safety, strengthen quality, and accelerate timelines

Key Words

Electronic data capture;
Electronic data review;
Real-time data access;
Clinical trials

Correspondence Address

Paula Brown Stafford,
Quintiles, Durham, NC
(email: paulabrown.stafford@quintiles.com).

INTRODUCTION

Advances in technology have enhanced data capture capabilities, making more timely and accurate information available to drug developers. Specifically, electronic data capture (EDC) systems electronically collect patient data directly from an investigative site, and data integration tools combine electronic data from various sources in real time. These advances, alongside data standards initiatives brought forward by the Clinical Data Interchange Standards Consortium (1), create new opportunities to streamline the clinical research process. Drug developers can now access more complete and higher-quality data faster.

Over the past several years, EDC has enabled the use of adaptive trial designs, or multistage trials during which drug developers review accumulating data at predetermined intervals to decide how to modify aspects of a study after it has started. These trials and the technology that support them reduce the gap between accruing clinical data and decision making, offering drug developers an alternative to lengthy and increasingly expensive traditional studies.

The same technology supporting adaptive clinical trials makes possible ongoing review of data throughout the course of a clinical trial. If

used appropriately, real-time data open the door for drug developers to improve drug development in the following ways:

1. Enhancing patient safety
2. Strengthening quality
3. Accelerating timelines

With opportunity, however, comes risk. Technology has made information that once required statistical expertise to assemble available to a wider group of people. As with adaptive clinical trials, more formal and clearly established protocols describing how ongoing data review will be conducted and documented and plans for implementing an audit trail enabling regulators to easily follow a drug developer's logic and method are required. With this, the use of real-time data can help to make significant strides in clinical research.

REAL-TIME DATA ACCESS: KEY CONSIDERATIONS

The notion of allowing drug developers open access to data during clinical trials is somewhat controversial, with the concern that even aggregate, blinded data in the wrong hands can be used to bias a trial. Still, with access to technology and a desire to enhance drug safety,

there is now more pressure than ever to utilize novel safety review tools to explore associations, event combinations, coding disparities, safety signals, and more by reviewing accruing data throughout the course of a clinical trial.

The US Food and Drug Administration (FDA) has signaled both concern and encouragement for the use of real-time data and the novel tools that enable it. In its guidance “Establishment and Operation of Clinical Trial Data Monitoring Committees,” the FDA stated, “Even aggregate data on safety and efficacy may be informative; these data . . . are best limited to those who cannot otherwise carry out their trial management responsibilities” (2).

Yet, during a presentation at the 19th annual Euromeeting in March 2007, the FDA concluded, “The Critical Path Initiative and Opportunities list and the FDA response to the IOM (Institute of Medicine) report on safety highlight the need for new tools and processes for signal detection and clinical trial review” (3). However, the FDA has also made clear that protecting study blinding is “particularly important to avoid the introduction of bias in the study conduct and to maintain confidence in the validity of the study’s result” (4).

Utilizing real-time data, therefore, requires careful management of the risks associated with introducing bias into a study, and mitigating them by ensuring that those individuals most vested in whether a study is successful on safety or efficacy criteria are most remote from the data access.

A potential temptation to be avoided is the case whereby a drug developer observes a higher incidence of adverse events in a placebo-controlled trial at a particular investigative site. The temptation in such an instance might be to make an educated guess that the active group is the one experiencing more adverse events, potentially leading a drug developer to limit enrollment at that site, while encouraging enrollment at another site with fewer adverse events. Similarly, a graphic of aggregated data for a quantitative efficacy outcome might reveal a bimodal distribution, suggestive of a difference between the treatments. Although such shifts

are seldom observed early in a trial, since limited data are available, a small risk remains.

Therefore, although the sponsor organization is charged to act on early safety signals—and must be supplied with the data to do so—data relating to efficacy could be more strictly controlled to prevent data mining aimed at determining whether the study is proceeding successfully.

The functioning of a typical data monitoring committee (DMC) for interim analyses provides a useful paradigm in the context of ongoing data access. In a closed session, the DMC acts as a third party, reviewing certain unblinded data prepared by an independent biostatistics center. This independence ensures that those running the trial are not influenced by knowledge of the treatment effect. However, the DMC also operates an open session, attended by the sponsor, where aggregated blinded data are reviewed, such as patient enrollment, protocol adherence, descriptions of the patient population, and descriptions of safety and perhaps efficacy. To prohibit information from biasing a trial, the information provided for both open and closed sessions must be detailed in the DMC charter, updated as the trial progresses and well documented.

By specifying in study protocols or in separate data charters exactly how information will be used, who will have access to what type of data and when, and establishing appropriate firewalls and a clear audit trail as per FDA guidance (5), controls over the use of real-time data can be established and enforced, thereby minimizing risk. Additionally, sponsors can allay regulator concerns about the introduction of bias and avoid a compound being rejected for perceived mishandling of trial data. Ongoing review of real-time data is generally accepted by regulators when study sponsors set out a plan and follow it carefully.

CONTROLLING DATA ACCESS

To establish the safe and effective use of real-time data, it is important to grant access to professionals on the basis of the role they play within the clinical trial, and in accordance with

international good clinical practice guidance. For example, while a clinical research associate (CRA) may engage in a range of activities, the primary responsibility of this role is to monitor and ensure the scientific integrity of the data being collected at investigator sites and the well-being of trial patients. To carry out their responsibilities, CRAs require access to the patient data being collected at investigator sites; however, these individuals should not be granted access to aggregated on-treatment data. Instead, aggregated data is best placed in the hands of the DMC to perform activities such as sample size reestimation in an adaptive trial design.

Keeping in mind that specific job roles and responsibilities may differ from sponsor to sponsor, Tables 1 and 2 provide guidance for granting access to real-time data based on professional roles.

Care must always be taken to ensure that real-time access to on-treatment data that may inadvertently unblind patients and groups of patients, such as pharmacokinetic or biomarker samples, is not provided. Indeed, even open-label comparative studies can be reviewed in a blinded manner simply by restricting access to specific data—for instance, device information for medical device trials. However, access to pretreatment data across the range of data captured in a clinical trial is broadly acceptable since no knowledge of the treatments can be imparted or inferred. From a system perspective, it is also important to balance the risks of providing real-time data access with the administrative burden of managing an overly complex security protocol. In fact, in Table 2, access levels have been deliberately simplified to balance such competing demands. Practical implementation along these lines necessitates defining role groups, distinguishing between pretreatment and on-treatment data by visit (or domain for event data), restricting access according to investigator site ID, and restricting reporting to one patient at a time for some roles. However, a security protocol that is designed to classify each variable on an individual study basis (eg, is blood pressure a safe-

ty or efficacy variable) is likely to be overly burdensome.

Alongside real-time data access, data may be summarized at certain cutoff times for specific purposes—for instance, to enable the DMC to perform a safety evaluation or to undertake a dry run on the database to prepare for the final analysis. In the case of dry runs, outputs are produced using a dummy randomization so that their format and layout can be finalized. Data derivation rules are also tested and refined to accommodate unexpected data combinations while highlighting data issues. In both cases there are clear advantages to having the most up-to-date data to summarize. Yet, although processes and documentation are most likely to be in place for DMC dissemination, there has been less control over the dissemination of dry-run outputs. Although the risk of influencing trial conduct may be viewed as remote, since most dry runs are performed toward the end of a trial or after the recruitment process is complete, it remains good practice to plan, control, and document which professionals have access to what data, when they have access, and exactly why access was granted. It is recommended that this information is also detailed in the protocol defining real-time data access.

ENHANCING PATIENT SAFETY

Identifying patient safety concerns during clinical trials is of paramount importance, and the earlier issues associated with developmental treatments are flagged and addressed, the better. The ongoing review of patient data throughout a trial plays a critical role in a drug developer's ability to monitor patient safety events, identify early safety signals, build a safety profile, and check protocol compliance. The faster this information is made available, the quicker sponsors and health authorities will be empowered to stop a study that is found to put patients at risk, or to develop additional safety measures for their protection.

The utility of real-time data with regard to evaluating safety is illustrated with the application of Hy's law (6), which is used during drug development to determine whether a drug

TABLE 1

Description of Clinical Trial Roles and Rationale for Granting Real-time Data Access		
Job Role	Role Description	Rationale for Granting Real-time Data Access
Lead CRA/CRA	Ensure the well-being of patients enrolled in trial. Ensure trial conduct is compliant with the approved protocol amendment(s). Monitor and ensure the scientific integrity of study data—in terms of accuracy, completeness, and verifiability—collected at investigator sites.	To ensure that enrolled patients meet the inclusion/exclusion criteria. To monitor the screening and recruitment numbers. To check the accuracy and completeness of data entries and that missed visits and assessments, withdrawals, and AEs are reported within required time frames.
Clinical manager	Coordinate cross-functional teams for a specific study to ensure quality deliverables on time, within budget, and in accordance with SOPs, policies, and procedures.	To monitor patient enrollment and completion to ensure cross-functional tasks are planned and timelines are met. To investigate issues of protocol and SOP compliance—including issues raised by others.
Global clinical lead	Coordinate cross-functional teams for a specific compound or indication to ensure quality deliverables on time, within budget, and in accordance with SOPs, policies, and procedures.	To monitor patient enrollment and completion across a program of studies to ensure regulatory submission tasks are planned and timelines are met. To investigate issues of patient eligibility, protocol compliance, and patient withdrawal in relation to the drug program as a whole.
Safety physician	Undertake regular surveillance activities directed at a specific developmental compound, and prepare safety evaluation documents for review in cases where safety signals are identified or suspected.	To identify potential safety issues as early as possible during a trial by exploring associations, event combinations, and temporal dependencies that may require action leading to protocol modification or termination, and that may impact other ongoing or planned studies.
Project statistician/statistical programmer	Undertake statistical analysis and interpretation for a specific study. Write and test programs to perform data manipulation and derivation, and produce tables, listings, and figures.	To identify potential data errors, inconsistencies, patterns, and outliers that may impact the results or analysis methods. To implement statistical methods to identify poor practices or potential fraud at investigator sites. To develop programs using real data to accelerate delivery times post-database lock. To support the ongoing evaluation of a study by undertaking programming tasks to facilitate safety surveillance activities and provide information for DMCs (via an independent, unblinded statistical center).
Program statistician/statistical programmer	Undertake statistical planning, analysis, and interpretation for a program of studies directed toward regulatory submission. Design integrated databases, write and test programs to combine and summarize data across studies.	To develop analysis plans and programs to integrate and report data across studies to characterize thoroughly a developmental compound and to accelerate submission delivery times. To support the ongoing evaluation of a developmental compound by undertaking programming tasks across studies to facilitate safety surveillance activities.
Lead data manager	Ensure the accuracy, completeness, and consistency of data through data capture, edit checks, and data clarification for a specific study.	To check data accuracy, completeness, and consistency. To raise and resolve data queries.
Medical writer	Create documents for a specific study that describe research results accurately and clearly, and that comply with regulatory guidance in terms of content, format, and structure.	To prepare patient narratives on an ongoing basis and highlight potential data issues to the project team.

TABLE 2

Guidelines for Granting Access to Blinded Real-time Trial Data				
Job Role	Data Type/Access Level ^a			
	Administrative data (eg, CRF completion; protocol/patient compliance; enrollment and visit attendance data, etc)	Pretreatment data (ie, any trial data collected up until the first exposure to study drug, including lab, vital signs, pretreatment adverse events, pretreatment efficacy, pretreatment safety, demographic/patient accounting data, etc)	On-treatment and follow-up efficacy data (eg, NINCDS-ADRDA scores for an Alzheimer study)	On-treatment and follow-up safety data (ie, after treatment lab, vital signs, adverse events, ECG, concomitant medications, etc)
Lead CRA/CRA	Single study access	Single study access	Single study access (by-patient only)	Single study access (by-patient only)
Clinical manager	Single study access	Single study access	No access	No access
Global clinical lead	Multistudy access	Multistudy access	No access	No access
Safety physician	Multistudy access	Multistudy access	Multistudy access	Multistudy access
Project statistician/statistical programmer	Single study access	Single study access	Single study access	Single study access
Program statistician/statistical programmer	Multistudy access	Multistudy access	Multistudy access	Multistudy access
Lead data manager	Single study access	Single study access	Single study access	Single study access
Medical writer	Single study access	Single study access	Single study access	Single study access

^aBoth aggregated and by-patient data unless otherwise specified.

could cause fatal liver injury or sufficient damage to require liver transplant. Most drugs causing severe liver injury do so infrequently, and the incidence of damage is generally not picked up during the course of a trial. Access to accruing laboratory data enables faster detection of markers (eg, elevated serum aminotransferases accompanied by increased serum total bilirubin) that cannot be explained by other causes, therefore, playing a critical role in assessing the potential toxicity of a drug.

It is important, therefore, in a real-time way, to obtain data generated from patients' medical history, vital signs, laboratory, concomitant medications and diseases, and adverse events. Early detection relies not only on accessing data in near real time, but on the ability to integrate information that has traditionally been housed in disparate locations so that meaningful trending reports can be created. Depend-

able reports made available shortly after data capture are critical to the flow of useful information, and are necessary for all levels of study management to take action when safety concerns arise. Graphical displays are particularly helpful to identify patterns and to evaluate temporal relationships for individual patients. More advanced graphics include motion-enabled bubble plots that plot data in three dimensions through time. Review tools depend upon real-time data to allow drug developers to identify possible safety concerns as they materialize, rather than months after the fact. These interactive tools enhance DMC meetings by enabling reviews to be performed more effectively and thoroughly, providing the capability to drill down into specific patients and link information efficiently—a common frustration for committee members who have a limited window to undertake such tasks.

As an additional protection, accumulating safety information can be combined and compared with safety data from previously completed studies, essentially forming a dynamic, integrated summary of safety. In this way, drug developers are more quickly able to identify and monitor for safety flags and adverse events, making the most current trial more informative with regard to a compound's relative risk/benefit. As information accrues, researchers are aware of the totality of safety information available for the compound being studied by incorporating it with previously acquired data. Adverse events in excess of previous studies might be an indicator for an independent safety reviewer, such as a DMC or project safety physician, to further scrutinize the data.

STRENGTHENING QUALITY

The quality and integrity of a clinical study depends on the ability to maintain protocol compliance from the outset. With greater regulatory scrutiny over protocol deviators and violators, and a record number of FDA warning letters, drug developers must ensure that patients not only meet inclusion and exclusion criteria, but also adhere to the protocol throughout the course of a clinical study. Ensuring protocol compliance is becoming increasingly difficult as protocols increase in complexity; however, real-time data play a significant role in helping drug developers quickly and easily detect protocol deviators and violators, thereby managing the ultimate quality of a clinical trial. Protocol adherence can also be strengthened by investigating the capture of all scheduled data at each visit, by monitoring visit attendance according to protocol schedules, and by reviewing reasons for patient withdrawal.

The quality of a study is also dependent upon the highest level of data accuracy. For example, if errors at any trial site go undetected, they can compromise the quality and integrity of the entire study. Real-time data play a critical role in helping researchers achieve data accuracy from the outset, spotting missing data, identifying data errors, or highlighting poten-

tially fraudulent data early. For example, if a laboratory assessment occurs without the completion of a corresponding clinical visit, then a missing page report can be generated in real time.

Furthermore, real-time data allow for signal detection of quality issues long before issues become real problems. By using data to spot trends or early warning signs, drug developers can proactively investigate and correct potential issues early—ensuring both corrective and preventive actions are taken. Quality signals include edit checks firing more often than the targeted value, prompting additional training or coaching to the problematic investigator site. On the other hand, edit checks firing too infrequently could suggest that a review of the edit check code is required.

The use of data to detect potential fraud is well documented (7), and statistical techniques have been developed that may point to further investigation of specific investigator sites. Methods include: end digit preference where the last digit of a set of assessments has an unusual pattern; and detecting inliers (rather than outliers) through the investigation of multivariate data structures where fraudulent data often exhibit markedly less variability relative to that observed elsewhere. Fraud remains an uncommon problem, but real-time data access enables checking programs to be put in place that encourage proactive trial management in this area.

The use of real-time data allows for ongoing review of a site's past site performance, the number of subjects and the rate of site recruitment, staff feedback on protocol compliance, site contact and record keeping, information received from data management, inaccurate or repetitive data, and other variables. Real-time access to each of these and other data points enable more timely site review triggers, which help to prevent the quality of the study from being threatened.

The ability to demonstrate that real-time access to data is being utilized for the purposes of improving the conduct of the trial is likely to be

well received by regulators. Real-time data access can also support and inform a sponsor's quality plan.

ACCELERATING TIMELINES

Technology designed to capture and integrate accruing patient data has enabled the increasingly widespread use of adaptive clinical trials. Adaptive designs, which use a set of design rules to define a priori which modifications can be incorporated into the trial design, reduce the length of time required to complete a clinical study and have potential to accelerate the development of promising therapies. Use of adaptive designs allows stopping decisions at the earliest possible time point, avoiding subjecting patients to noneffective or unsafe therapies.

Adaptive designs involve interim analyses, which are planned at the outset and well controlled. It is understood that results must be handled carefully, with firewalls put in place to ensure that those with access to the data are not a part of the study team.

Ongoing review of real-time data differs from adaptive trials in that access to data is more open and, to date, its use has been less well defined. Furthermore, adaptive trials typically require access to treatment assignments whereas real-time data access does not. It is important to note, however, that one commonly used sample size reestimation method for adaptive trials only utilizes the pooled estimate of the variance (or the probability of events) from the blinded aggregated data. But like adaptive trials, ongoing review of data can play an important role in the acceleration of a study's timeline.

For example, open access to blinded, real-time information makes faster data cleanup possible. Undetected errors in data capture can delay database lock or, worse, compromise the overall study. Errors are much more quickly detected with the use of real-time data than has traditionally been the case. This helps to ensure accuracy and encourages standardization of training and processes across trial sites, as well as quicker remediation when a problem is identified.

In addition, drug developers are able to conduct a blinded dry run of the output process at various points throughout a study, which enhances data cleaning and aids in finalizing data derivation rules and validating programming code. The manner in which data are organized in tables (of which there can be hundreds) can be reviewed at preplanned intervals by the global clinical lead overseeing the trial and other members of the sponsor organization months in advance of the trial's completion, thereby expediting data output at a study's conclusion.

It is imperative with dry runs to put in place protections against the misuse of data and to establish at the outset exactly who has access to what information, when. By ensuring that the data under review are aggregated but ungrouped, or grouped randomly using dummy assignments, those most concerned with the efficacy of the trial are only able to review table formatting and not mine the data in any way that would jeopardize the integrity of the trial.

Finally, quotas for enrollment, which nearly 80% of trials fail to meet, are better managed with real-time data review (8). Access to real-time data offers researchers a timely look at the state of enrollment for the trial as a whole, and an opportunity to put a corrective action plan in place when the data indicate enrollment is slower than predicted.

CONCLUSION

As technology and tools better enable the capture, integration, and evaluation of data in real time, drug developers will have new opportunities to enhance the process and outcomes of clinical trials in the new health landscape. For drug developers interested in leveraging real-time data access, however, it is imperative to recognize that it must be done in a planned, controlled, and documented way to avoid risks associated with biasing a trial.

It is important to consider formulating clear review policies, as well as the use of appropriate data review tools that limit user access to data. The intent to review real-time data throughout the course of a trial should also be

clearly established at the outset within a clinical trial protocol or separate real-time data access protocol. Finally, a clear audit trail must be established so that regulators can easily determine how and when data were reviewed and by whom.

The advantages of reviewing real-time data far outweigh the dangers and allow for earlier detection of patient safety issues, faster timelines, and higher-quality clinical trials.

Acknowledgments—The authors would like to thank the following colleagues at Quintiles who reviewed and provided valuable comment on earlier drafts of the manuscript: Vladimir Dragalin, Thomas Grundstrom, Michael Fiola, David McGowan, Michael O’Kelly, and Lindsay Singler. The authors would also like to express thanks to Gary Koch for his valuable contribution.

REFERENCES

1. Clinical Data Interchange Standards Consortium. <http://www.cdisc.org/mission-and-principles>.
2. Food and Drug Administration. Guidance for clinical trial sponsors: establishment and operation of clinical trial data monitoring committees, section 4.2.2. March 2006. <http://www.fda.gov/OHRMS/DOCKETS/98fr/01d-0489-gdl0003.pdf>.
3. O’Neill R. Signal detection in clinical trials: some perspectives on new tools and processes—a critical path update. 19th DIA Annual Eurometing, Vienna, Austria, March 26–28, 2007.
4. Food and Drug Administration. Guidance for industry: adaptive design clinical trials for drugs and biologics. February 2010. <http://www.fda.gov/downloads/Drugs/guidancecomplianceregulatoryinformation/guidances/ucm201790.pdf>.
5. Food and Drug Administration. Guidance for industry: part II: electronic records; electronic signatures—scope and application. August 2003. <http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm070295.pdf>.
6. Food and Drug Administration. Guidance for industry: drug-induced liver injury: premarketing clinical evaluation. July 2009. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM174090.pdf>.
7. O’Kelly M. Using statistical techniques to detect fraud: a test case. *Pharm Stat*. 2004;3(4):237–246. <http://onlinelibrary.wiley.com/doi/10.1002/pst.137/abstract>.
8. Hess J, Litalien S. Web-based patient recruitment: best opportunity to accelerate clinical trials. *Cutting Edge Information*, 2005.

Paula Brown Stafford has disclosed that she is an employee of Quintiles, that she received honoraria from the UNC CTSA External Advisory Board, and that as an employee of Quintiles she currently undertakes, and in the past has undertaken, extensive work for the biopharmaceutical industry. Andrew Garrett has disclosed that he is an employee of Quintiles and that in this capacity he currently undertakes, and in the past has undertaken, extensive work for the biopharmaceutical industry.