The Importance of Remote Data Access and Analysis in Clinical Trials

The migration from paper-based data collection to electronic data capture (EDC) systems makes it possible to analyse trial data from any location as soon as it is collected. If leveraged properly, advanced data access and analysis methods can speed up the trial, make it safer, and even unlock substantial cost and time savings. The FDA has encouraged companies to implement these advanced methods in order to rely less on monitoring at clinical sites and more on centralised data analysis. This article will discuss the potential considerations that must be addressed when choosing remote data access and analysis platforms—including accessibility, security, ease of use, speed, data freshness, comprehensiveness, and correctness—and the specific advantages that can be gained when modern remote data analysis techniques are applied to a study. This article will also briefly discuss how modern data access and analysis can facilitate advanced adaptive trials.

Accessibility and Security
Accessibility considerations relate to who can access study data and by what means they access the data. The goal of accessibility is to make sure that data is available to the people who need it, when they need it, and in a manner that helps them do their job faster and better. The flip side of accessibility is security, or making sure that unauthorised users are not able to access study data. A well-designed remote data platform must find the optimal balance between removing friction points between clinicians and their data and preserving security.

There are two main sources of friction from an accessibility standpoint: the data access point and the login structure. The first friction point is typically the relative flexibility of the data access point, the tool the user actually uses to see the data. For example, can the data be accessed from any computer, or only specific machines? Through other devices, such as an iPad or smartphone? Over any internet connection or only proprietary networks? Using common computer programs, like a web browser, or only through specific software that must be installed on all computers of all users? Although increasing the flexibility of the data access raises some security concerns, today’s modern data security methods allow for extremely flexible data access points, with the ideal data access option being any web-enabled device. This would allow, for example, a study sponsor to access preliminary results from her hotel room while traveling, or for a study administrator to receive real-time information about serious adverse events (SAEs) on his smartphone when he’s out of the office.

The other primary point of friction is the number of times a user must log in to gain access to all of the data and tools they need. Clearly strong user authentication is necessary to secure sensitive trial data, but ideally such authentication should follow a “single sign-on” model, where, for example, a user can log in to an electronic portal once, and then the permissions embedded in the portal automatically provide the user with access to all of his authorised data and tools. Many EDC vendors have begun to recognise the importance of single sign-on, particularly where multiple systems are involved, and are developing integrated systems that allow users to work seamlessly across multiple systems after a single authentication. This has the advantage of both creating less interruption in a user’s workflow, and also reducing the risk of security breaches from human error, such as users writing down their passwords in order to keep track of multiple logins.

A key element of security that does not create accessibility friction is audit trails. It is essential that any remote data access platform keep an accurate audit trail of every action performed on the platform, and also respect and propagate the audit trail requirements of any underlying system. A comprehensive audit trail allows administrators to verify that logins, permissions and other security structures are functioning as designed, and is also imperative to have in the extreme case of an actual security breach. In this respect, an integrated EDC and remote data platform may be considered more secure than a manual data collection and on-site review system, because every step of the EDC system can be electronically recorded and reviewed, which is not true for all steps of manual data capture.

When a remote data access platform is optimised for accessibility, it can reduce study bottlenecks, help clinicians and other users perform their jobs more efficiently, and ultimately decrease study duration. For example, suppose the study sponsor, reviewing preliminary results in her hotel room mentioned above, is looking at the results of patient recruitment screeners administered that day. If the sponsor notices that a certain aspect of the protocol is causing unnecessary screen failures, she could immediately start working on a protocol amendment to avoid the problem. If the data was less accessible, she may not have been aware of the problem until later, raising the number and cost of unnecessary screen failures.

As noted above, any increase in accessibility must be counterbalanced by an assessment of the related security risks. Even small increases in accessibility can have a large impact on the efficiency of study administration.

Ease of Use and Speed
Once a user has access to the remote data platform, the next criterion to assess is how quickly and easily the user can obtain the information he needs to know about the data. The key components of this assessment are whether computer programming knowledge is required to interact with the data, how well the platform highlights and elevates key points of interest and, relatedly, how quickly a user can drill down on a point of interest, how closely different...
systems are integrated, and whether the audit trail supports efficient data interaction. Each of these components is discussed below.

Not every clinician is a programmer, nor should they need to be a programmer in order to make the most of their tools. Clinicians are domain experts, and their tools should complement their skills instead of getting in their way. This has not always been the case. One of the most basic workflows for reviewing clinical data, and one that is still often used today, requires a reviewer to determine what question they want to answer about the data and to then relay that question to an SAS programmer or statistician who then converts the question into a database query and returns the results to the reviewer. This process is both error-prone and time-consuming, not to mention frustrating for a reviewer, particularly if, for example, the reviewer realises very quickly upon seeing the results that she needs to ask a different question or that the database query didn’t match the question she had in mind. Smarter data platform tools can translate a clinician’s input into the appropriate database query, allowing for ad hoc reporting and instant results. The best of these tools have intuitive interfaces that a clinician can quickly understand and leverage the metadata associated with standardised EDC data, so that the platform can handle even complex clinician questions without the need for programmer involvement.

Another way that better remote data platforms help users review data more easily is by automatically highlighting important information. One feature that does this is high-level dashboard overviews that are configurable by the user, enabling the user to see the metrics that he most cares about. For example, a clinical monitor may want to see a dashboard of how many patients are currently enrolled at each study site when she logs in to the platform, while a data manager may want to see how many data discrepancies have been reported in the data cleaning process and how they were resolved. All of these dashboards are possible if the appropriate databases are connected to the platform, and ideally the platform should allow the user to easily reconfigure the dashboard as their informational needs change.

For information that needs to reach the user even faster than a dashboard can provide, a remote data platform should allow users to configure alerts that are triggered in real-time, as soon as certain events occur. These alerts can be configured to automatically send an email or text message, so the notified users do not need to wait until they log in to a computer to know about the event. The most common application of alerts is for serious adverse events (SAEs), but stakeholders may want to be alerted to other events, for example, if a patient misses a scheduled visit or drops out of the study. By receiving alerts, stakeholders can address potential issues quickly, before they become bigger problems that may delay or derail the study.

On the other end of the spectrum from features that highlight information are features that allow users to quickly and easily drill down into the data and interact with it on a granular level. A well-designed data platform should bridge the two main barriers to a user being able to drill down on data quickly: the need to form database queries and the need to integrate multiple data systems. As discussed above, a data platform should do the work of creating database queries so that a user without computer programming training can obtain query results through a simple interface. A data platform should also do the work of integrating multiple data sources. For example, if data collected at sites are stored in an Oracle database, but patient demographic information is collected in an Excel file, the data platform should have the ability to run a query involving both sets of information without moving the data from its source. The alternative method, to manually combine the data sources into one database, which is what many study administrators continue to do, tends to be extremely costly and introduces an unnecessary opportunity for error.

Finally, a data platform can allow for quicker and easier data review by leveraging the audit trail, which, as noted above, is a necessary feature for security reasons. For example, if a reviewer can easily review their audit trail and annotate it with notes describing the review process, the reviewer can easily go back to the beginning of a review path that turns out to be unhelpful and start a new branch, or another reviewer can use the audit trail as a template for looking at a different set of data, speeding up the process of data exploration.

The advantages of a fast, easy to use data platform that empowers non-programmers to review study data are tremendous. By eliminating the number of steps between a reviewer determining a question and the reviewer receiving the data that answer that question, advanced data platforms can substantially decrease costs, remove common sources of error, help identify potential issues quickly, before they become costly problems, and greatly increase the speed at which clinical data can be reviewed.
Data Quality (Freshness and Comprehensiveness)

Many of the benefits of a well-designed remote data platform that are referenced above are amplified by the quality of the underlying data. For purposes of this discussion, the quality of the data is measured by how fresh it is and how comprehensive it is, as described in more detail in the following paragraphs.

Freshness describes the length of time since data were collected; the more recently data are collected, the fresher they are. In a paper-based CRF system, it may take weeks between when the data are collected and when they are accessible in the data platform, whereas top-of-the-line EDC systems can provide real-time, instant access to data. The value of freshness depends on the purpose of the data. Where the goal is to identify potential issues in the study, such as SAEs, patient recruitment failures or cost overruns, fresh data are crucial. Freshness is also important for study sponsors planning to implement the FDA’s recommendation to use centralised analysis for ongoing studies. One purpose of centralising the analysis is to highlight sites that are not progressing appropriately compared to the other sites, and make appropriate adjustments while the study is ongoing. The fresher the data, the more valuable such adjustments are likely to be.

Comprehensiveness describes the extent to which all relevant data sources are integrated into the remote data platform. For example, connecting IVR and recruitment data can help the sponsor keep track of how many screen fails, enrolments and randomisations are occurring. Connecting relevant clinical trial management systems (CTMS) and other clinical systems to an inventory system can help the sponsor track supplies and make sure each site has the right inventory for its scheduled patients, while connecting the budget and expense system can help the sponsor track the costs of its study and manage the financial aspects of the study. A relevant data source may also be the results of another study. The ability to compare results across studies can be extremely important in identifying adverse effects or confirming drug effectiveness.

The cutting-edge features of advanced data platforms are limited in value if the underlying data are of low quality. A data platform should facilitate the rapid input of fresh data and the integration of multiple data sources by, for example, connecting directly to EDC sources and incorporating software components that can run queries across different data sources without moving the data from the sources.

Correctness

It is important to note that, although advanced data platforms offer many exciting new features and the potential for valuable innovations in clinical trial design and process, they still must meet the most fundamental need of any software used in clinical data review: the software must be correct. In particular, any data used in part for electronic submissions to the FDA need to be Title 21 CFR Part 11 compliant. These regulations can create some obstacles for data platform design, and in evaluating any remote data platform, an essential question is whether its correctness has been validated through industry-accepted methods, and whether the platform provider is committed to responding quickly and thoroughly to any bugs that users may encounter.

Remote Data Platforms and Adaptive Trial Design

While advanced computer and data analytics technology have been transforming industries from finance to retail to national defence, the clinical trial industry has lagged in taking advantage of these new technologies. That is likely to change in light of recent recommendations by the FDA for study sponsors to adopt new processes, including the centralised data analysis described above, and adaptive trial design, which involves building decision points into a study protocol where clinicians can adapt the protocol based on the initial results of the trial. In order to implement adaptive trial design, clinicians need a data platform that supports quick, easy, and centralised review of the initial trial results. In selecting this platform, study sponsors should consider the criteria described in this article in order to maximise the benefit of both the trial methods recommended by the FDA and the substantial cost and time savings offered by the newest data platform and analytics technologies.

References


Rick Morrison is co-founder and CEO of Comprehend Systems. Previously, Rick spent over a decade writing software for clinical trials, including tools that are now used by the FDA and top pharma. Email: rmorrison@comprehend.com