Today's Broken EDC Model and How the CRIO Model Fixes It

A PRIMER

Author: Raymond Nomizu, CEO at CRIO

sales@clinicalresearch.io
www.clinicalresearch.io
(617) 302-9845

02

A NEW MODEL FOR CONDUCTING CLINICAL TRIALS

This primer describes CRIO's integrated eSource-EDC model, in which data are captured once by the site against electronic source templates built by the sponsor, and then surfaced into a sponsor-facing application called Reviewer, where the sponsor can review, query, lock, code and extract the data. This model eliminates the vast majority of the need for secondary data entry into an EDC, onsite monitoring by CRAs, source data verification and redundant data management workflows, thus driving substantial gains in quality, speed and cost savings.

To truly understand the benefits of CRIO's model, it is important to describe the current model in clinical trials. To understand this model, we need to document the specific steps sites currently perform, and what sponsors do as a result.

03

The Current Site Workflow is Highly Manual with Multiple Failure Points

Translate the protocol into eSource template

Collect data against the templates

Enter into EDC

In the current model, sponsors provide the protocol in PDF format to clinical research sites, who are responsible for reading, interpreting and then executing the protocol requirements for all aspects of patient recruitment and data collection. Typically, protocols require that the investigative sites collect critical data such as subject eligibility, clinical endpoints, medical history, medications, adverse events, drug accountability, and patient compliance.

To comply with the protocol, sites construct paper source templates to guide their workflows and collect the appropriate data. Oftentimes, a single study coordinator will spend hours reviewing the 100+ pages in the protocol and translating the requirements into a step-by-step workflow to ensure the site collects data properly. This process is done manually, often with little guidance from the sponsor or CRO, and with no formal quality control and review by the sponsor or CRO.

Here's an example of what the protocol might say regarding vitals:

- Vital signs include blood pressure (resting more than 5 minutes in the sitting position) and pulse rate.
- The investigator will assess any vital sign outside the range of [X-Y] for clinical significance.
- Blood pressure should be determined using the same arm, and the same position throughout the study.

04

And here's how one site might write the source template:

VITAL SIGN MEASUREMENTS	
☐ Check here if vitals not done. Reason:	
Date of Visit	__ __ / __ __ __ / __ __ __ __
Time of Position	___ : ___ (24 hr clock)
Position	☐ Sitting ☐ Supine (deviation – explain below:) ☐ Standing (deviation – explain below:)
Time of Vitals *(needs to be 5 minutes+)*	___ : ___ (24 hr clock)
Reference Arm *(use same arm as v1)*	
Arm Used	
Pulse Rate	
Systolic BP	
Diastolic BP	
Investigator's Interpretation	☐ Within Normal Limits ☐ Abnormal, not clinically significant ☐ Abnormal, clinically significant
Investigator Initials and Date	

Figure 1: Sample source template

On the day of the patient visit, the study team will print these paper templates to guide their workflows and record data. Then, after the visit is complete, the sites will manually transcribe these data into the electronic case report forms (eCRFs) housed in an Electronic Data Collection (EDC) system provisioned by the sponsor.

The eCRF is usually a condensed version of the procedure, generally focused on the endpoint data, not necessarily all the procedural elements:

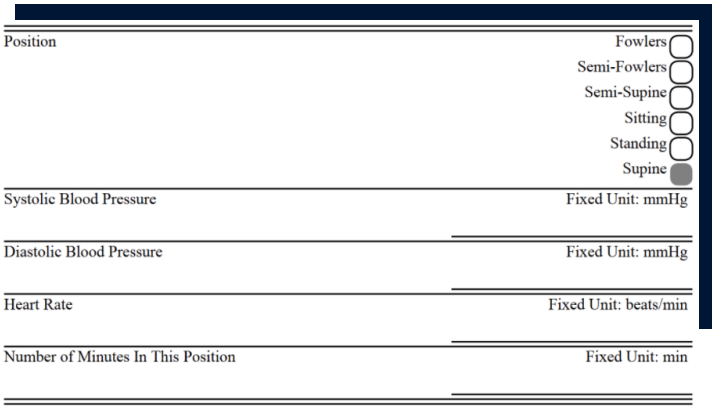


Figure 2: Sample eCRF

Because of this highly manual process, there is no assurance that the data in the eCRF were captured per protocol. This lack of assurance stems from multiple failure points in this process:

Sites design incorrect source templates

Because sites write their templates based on their interpretation of the protocol, different sites may end up writing completely different workflows. Inevitably, some sites will miss critical elements of the protocol when writing their templates, or, worse yet, codify protocol deviations because they incorrectly designed the templates.

The lack of standardization across protocol design is a high risk for sponsors. As a scientific protocol, the data methodology should be standardized across sites to ensure that confounding factors are eliminated and patient safety is protected. However, the current process does not have any standardization to it - instead, sites are translating the sponsor's specification into their own site-specific specification. If there are 20 sites on a trial, the protocol effectively ends up being specified 20 different ways by 20 different people.

Sites make mistakes against their own templates

Because they default to using paper charts, sites make mistakes against their own templates. For instance, in the example source template above, the site might mis-record the date of the procedure, record the wrong arm, leave a field blank, forget to initial and date, or record a signature date inconsistent with the vitals date. Because the data collection process is so error prone, sites commit protocol deviations that could easily have been avoided if they were using a real-time electronic system. These deviations may go undetected for months, until a monitor visits the sites and reviews the source.

Sites do not enter complete data into the EDC in a timely manner

Entry into the EDC is often delayed, usually by days and sometimes weeks. Sometimes, sites enter parts of the eCRF record but not all. As a result, when an eCRF form is blank, it's impossible for a sponsor to know whether it's blank because the data wasn't collected, or the site hasn't entered the data yet. This lack of transparency creates an overhanging question on the accuracy and completeness of the data in the EDC, and is one of the reasons necessitating review of source data.

06

Because the EDC is a Secondary Entry System, Sponsors Must Review the Source Data, Not Just the eCRF

The FDA holds the sponsor liable for all aspects of the clinical trial, including protocol compliance, patient safety, Principal Investigator oversight, and investigator compliance with ICH-GCP principles. [21 CFR 312.60 and ICH E6 (r2)] Sponsors cannot fulfill this obligation from reviewing the eCRF data alone. They can only fulfill this obligation by reviewing the source data itself. Here are the deficiencies of relying on the eCRF data alone:

The eCRF is usually not a complete manifestation of protocol compliance

The eCRF usually contains a condensed version of the required endpoint data, but not all the data required to demonstrate that these endpoint data were collected per protocol-prescribed methodology. For instance, the eCRF in the above example captures the position and required time period, but not the actual sitting time and vitals time, which could demonstrate that the requisite resting time was not met. It also does not provide evidence that the same arm was used throughout.[1]

The eCRF does not contain evidence of PI oversight

While the PI signature on the eCRF is usually obtained prior to data lock, the only way to ensure actual PI oversight on the trial is to review the source data. The source data will contain evidence that the PI was the one who actually performed the medical procedures, or that the PI reviewed the data collected by others in a timely manner. The source data will contain the Investigators' progress notes, which are text entries that amplify or clarify the PI's clinical reasoning in making eligibility or safety decisions. For instance, a PI may document his or her reasoning as to why a given Adverse Event does not rise to Serious status, or why it does not merit discontinuing the protocol-prescribed IMP dose. If the PI's reasoning is flawed, or improperly documented, an auditor could create a finding of inadequate PI oversight.

1 Although there may be an aspiration for the eCRF to demonstrate protocol compliance, the reality is that protocols are extremely complex, and most eCRF templates are incomplete in capturing all the components specified in the protocol. CRIO's experience is that source contains 2x the data points as the eCRF.

07

The eCRF does not demonstrate that data were collected per ICH-GCP principles or by qualified personnel

ICH-GCP requires that source data be captured per ALCOA+ principles - Accurate, Legible, Contemporaneous, Original, Attributable, Complete, Consistent, Enduring, and Available. Only direct review of the source data can confirm that this was done. In particular, the "attribution" component is of paramount importance because ICH-GCP requires that only appropriately qualified personnel specifically delegated by the PI can perform study-related tasks. Only the source data contains evidence of this. While the EDC may have audit trails, it only has audit trails of who entered the data, not who actually collected the data. [2]

The Process of Reviewing Site Data is Inefficient and Costly, With Multiple Redundancies

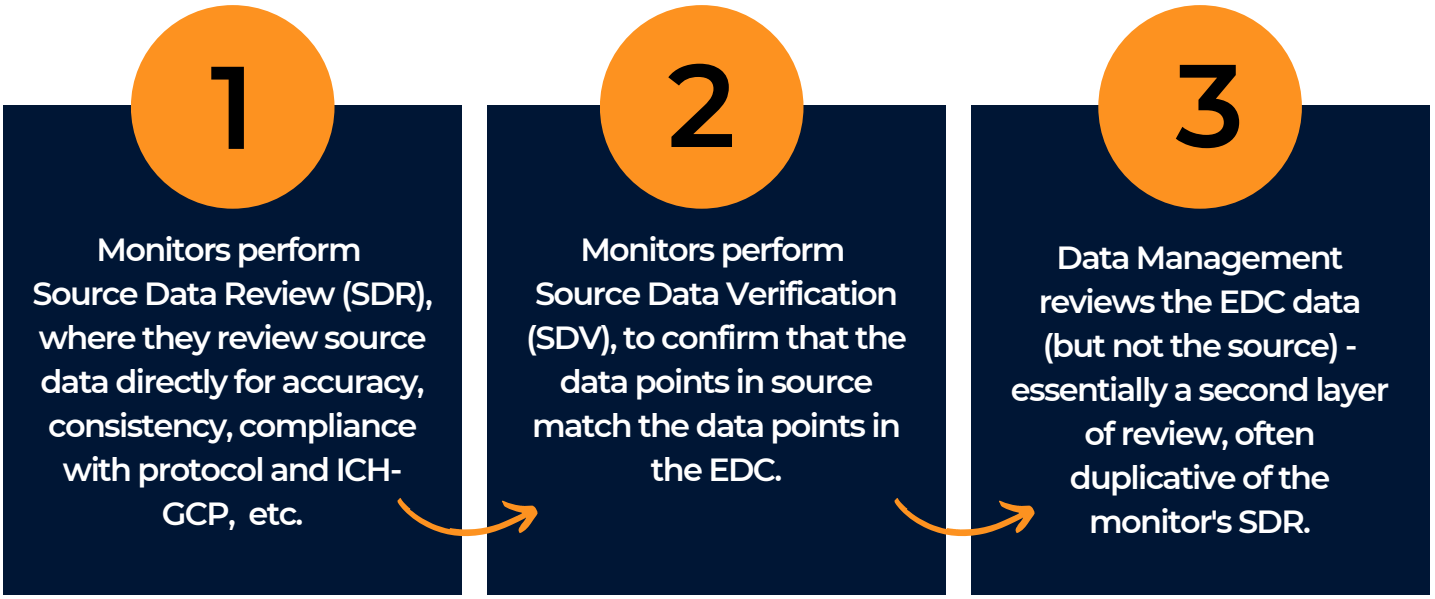
CRA visits site to perform Source Data Review

And Source Data Verification, after which

DM reviews data in the EDC

2 Indeed, "EDC data entry" usually has its own duty on the delegation log. Thus, a study team may specifically bifurcate the responsibility for performing a procedure from the responsibility of entering the procedure's endpoint data into the eCRF.

Because of the failure points described above, Sponsors and CROs have built a complex, multi-step process for reviewing and cleaning site data. There are in fact three separate review cycles of site data:



Let's start with Monitoring. Because sponsors need to review the source data to confirm protocol compliance and PI oversight, and to confirm the data in the eCRF, Clinical Research Associates (CRAs) perform onsite monitoring visits. That means they schedule the monitoring visits with sites weeks or months in advance, and then travel onsite to review the completed paper charts. On these visits, they do two separate processes, usually simultaneously, and in roughly equal time allotment.[3]

First, monitors do SDR. Monitors independently review the source data to ensure it was collected per protocol, that the documentation adheres to the ALCOA+ principles, that the PI has provided oversight (as evidenced by progress notes, signatures, and/or direct performance of medical procedures), and that the data are internally consistent and complete. In other words, monitors review the source data in totality to ensure that the data, on a stand-alone basis (i.e., without reference to the eCRF), sufficiently documents protocol compliance and contains the endpoint data required to assess the efficacy and safety of the investigational product.

3 Per a CRIO time study conducted of CRAs, monitors spend on average 31% of their time on Source Data Review and 28% of their time on Source Data Verification, and the vast majority do both at the same time when reviewing source charts.

For instance, in the above Vitals example (see Figure 2), the CRA/Monitor might ask:

1. Does the source evidence completion of Vitals per protocol?
2. Are the data accurate and internally consistent?
3. Did the PI review any out of range values?
4. Were the data collected in a manner consistent with ALCOA+?
5. Was the person collecting the data appropriately delegated and trained?

Second, the CRA does SDV, where the CRA compares the data points in source against the data points in the EDC, and marks off in the EDC that the data points have been source data verified. This ensures that when Data Management looks at the eCRF data, they are looking at the correct dataset.

After the CRA review, Data Management now looks at the eCRF and performs their own separate data review which is almost always a subset of the same SDR that the CRA previously performed.

For instance, in the above Vitals example, the Data Manager might ask these exact same questions:

1. Does the eCRF evidence completion of Vitals per protocol?
2. Are the data accurate and internally consistent?
3. Did the PI review any out of range values?

Data Management cannot, by definition, perform all of the reviews done by the CRA. For instance, Data Management cannot confirm that the person collecting the data was appropriately delegated and trained, or that the documentation of the vitals was done contemporaneously with the procedure. However, the data checks that DM does are almost always what the CRA is already doing as part of SDR.

In summary, the clinical research industry has effectively created two parallel data sets with two parallel data reviews. First, the sites collect data in source, then manually re-key a subset of the data into the eCRF. Then, the CRA's perform review of the source data, while DM performs review of the eCRF data. What holds this together is SDV, which ensures that the two data sets are the same.

The CRIO Model Streamlines All of This

The CRIO model streamlines all of these inefficiencies. With CRIO, there is only one data template - an eSource template that codifies the protocol and acts as the “eCRF”. There is only one point of data entry: the real-time, contemporaneous entry at source. And there is only one data set. There is thus only one review cycle, and this review cycle can be done remotely, centrally and continuously.

In the CRIO model, CRIO’s study design team will review the protocol and design the study template. This ensures there is one uniform data collection process, not one created uniquely by each site. By ensuring that all sites operate off the same workflow, sponsors build standardization into the data methodology.

In designing the eSource template, CRIO will sequence the procedures in the protocol specified order, enrich them with detailed instructions, and incorporate alerts, branching logic and other automation to guide the site through the correct process. Because these alerts operate in real-time, at point of capture, they ensure protocol compliance, leading to cleaner data from the outset.

Here is an example of how CRIO might design the same vitals procedure referenced above:

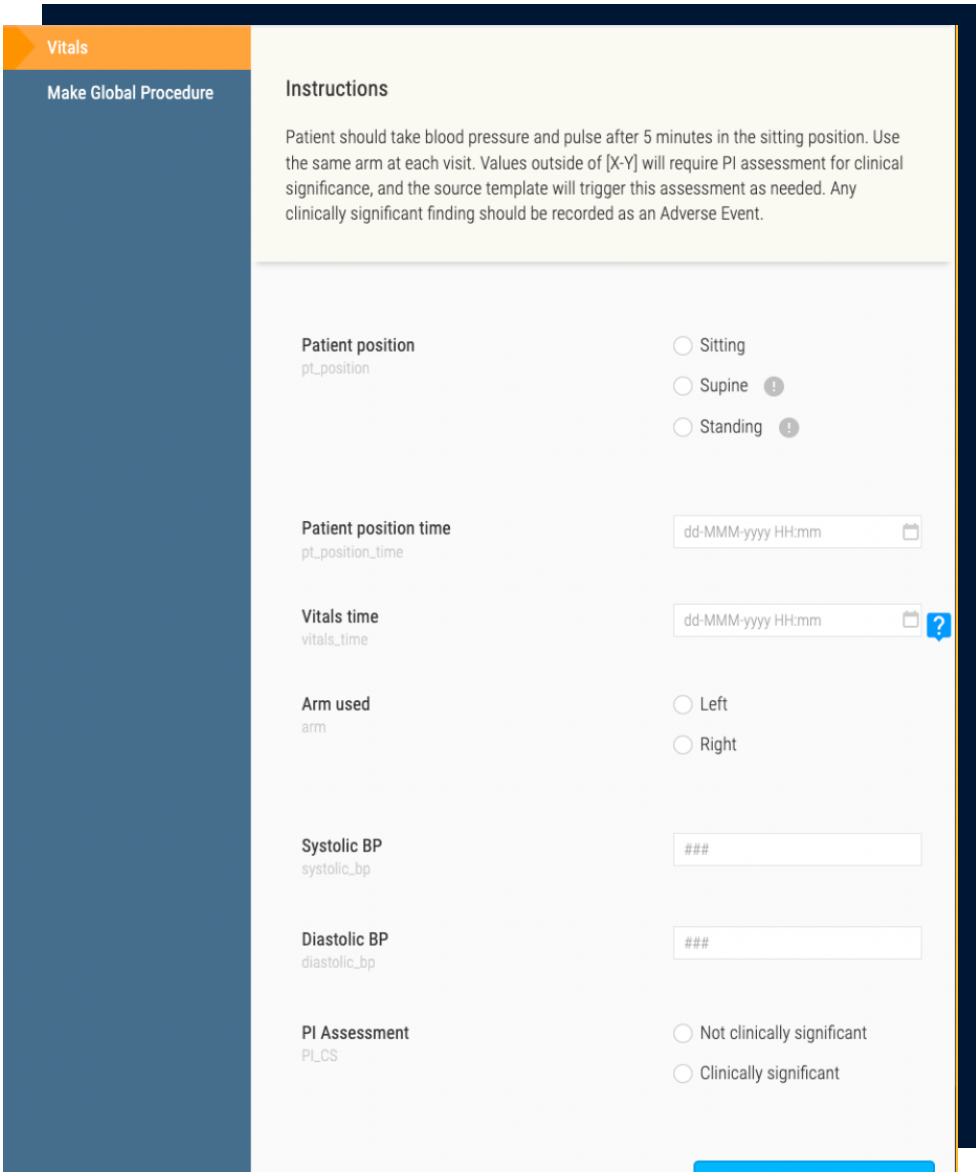


Figure 3

In the procedure above, CRIO would have automation logic to drive compliance, such as:

- An alert appears if a position other than Sitting is selected
- An alert appears if the vitals and position time is less than 5 minutes
- PI assessment is triggered based on the values entered

The site user can also use a number of built-in features such as:

- The ability to display the history of past values, thus ascertaining the previous position and arm used, or prior blood pressure readings for comparison.
- The ability to mark the entire procedure, or a specific value, as “Not Done”, whereupon the system requires the user to leave a reason.
- The ability to leave a progress note.
- The ability to tag the Investigator in an internal comment after the visit for clarification, or to call attention to a data point for PI review.

In addition, within CRIO, the study designer may designate “Vitals Time”, “Systolic” and “Diastolic” as “Core” variables - Core meaning that they represent the subset that the statisticians need to perform their analyses. This subset is what should be extracted and provided to the statisticians at the end of the study.

The quality gains from this process are significant and unequivocal - in the order of 40-70%. In one large scale published study a site network with an already advanced QA process experienced a 38% reduction in protocol deviations from moving off paper to CRIO for source data collection. In CRIO’s review of FDA audits of sites utilizing the CRIO system, we observe that sites using CRIO have a 70% more favorable outcome from the audits than the overall industry average.

Once the sites save the electronic data, CRIO’s system will send that saved data on an anonymized basis directly into CRIO Reviewer, a Sponsor facing application that lets the Sponsor or CRO review, query, lock, code and extract the data. This process ensures that the Sponsor or CRO receives the data immediately - not when the site gets around to entering the data - and without any lingering questions of accuracy and completeness that taint the current eCRF data review process.

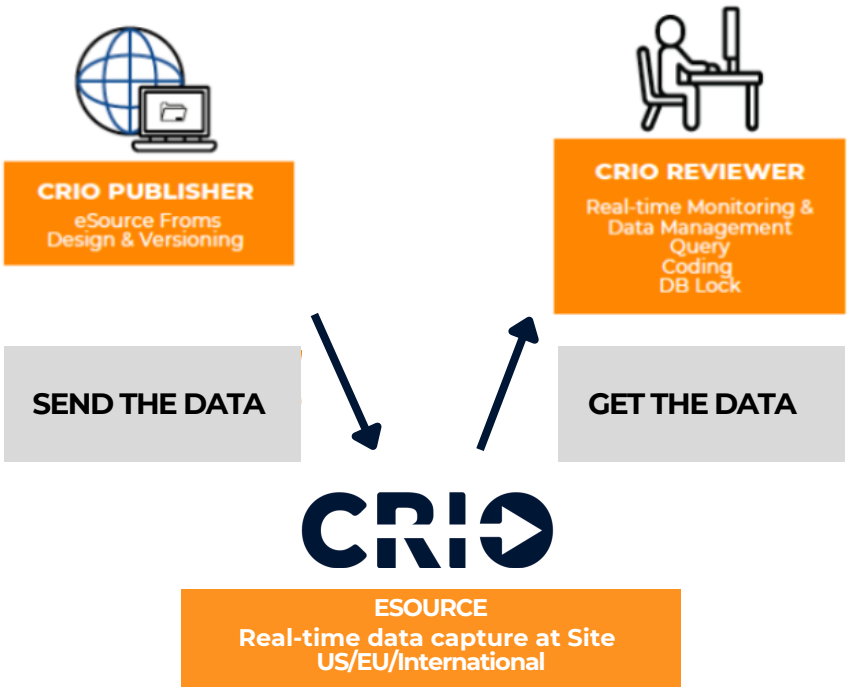
Because there is now only one data set, sponsors can collapse three separate workflows (Source Data Review, Source Data Verification and DM Review) into one review workflow. In this workflow, which we call “Clinical Data Monitoring”, a CRA can review the data within days of the sites saving them. This means that monitors no longer have to coordinate, schedule and perform onsite travel just to review source data. It also means that with faster review cycles, monitors can identify actual or potential deviations earlier, thus preventing further deviations. **Based on a time study analysis, CRIO estimates that at least half of monitoring time can now be eliminated.** From a total cost perspective, this enables savings on 50%+ of monitoring time, most of the travel expenses, and virtually all of the DM time spent on form-by-form review.

Furthermore, the CRIO Model enables ongoing locking of data. This is because CRIO eSource secures the PI sign-off of source data as the visits are completed since it leverages the traditional role of PI's in reviewing source data contemporaneously as part of their oversight. Once the clinical team has reviewed the data, all queries are resolved, and PI sign-off is secured, CRIO Reviewer enables the clinical team to lock the visit. With continuous data locking, sponsors and CROs can accelerate database lock, and therefore improve time to market

Why Is CRIO's Software Uniquely Suited to do this?

Many EDC vendors make the claim that their system can be used as a Direct Data Capture tool, thus claiming similar benefits. **These systems cannot replicate the benefits of CRIO and that's because CRIO's system is founded on a groundbreaking architecture.**

CRIO's system consists of three separate components:



1 First, CRIO Publisher allows sponsors to create, version and publish source templates to sites, thus standardizing the template. Critically, and unique among vendors, the receiving sites can add their own customized visits and procedures to the template - for example, a site could add a pre-screening visit, or a unique site-specific procedure they wish to incorporate as part of their SOPs. This makes source creation a truly collaborative process that reflects site needs. Without the ability to accommodate site-specific workflows, a sponsor-mandated "DDC" system could end up straight jacketing sites, hampering their successful adoption of the system.

2 Second, once the templates are published, sites use CRIO eSource, a site-native workflow tool, to populate the templates with contemporaneous data collection. The data collection is part of a broader workflow designed to streamline site operations and facilitate compliance. For example, CRIO offers integrated patient recruitment, eConsent, EMR integration, patient scheduling/payments, and built-in phone and text communications between the site and patient. Critically, CRIO's eSource system has Protected Health Information (PHI), which sites need to run their trials, but this PHI is stored on CRIO's regionally dispersed servers.[4] In this architecture, CRIO's eSource database remains local to the geography, thus giving regulators assurance that their citizens' PHI does not leave their geographic boundary.

3 Third, when sites save data in CRIO eSource, the data are transmitted to CRIO Reviewer, a separate, sponsor-facing application. Only study-level, anonymized data are transmitted. This database is separate from the site database and ensures that sponsors do not have direct access to PHI. Sponsors can use Reviewer to review the visits as they are completed, and query, lock, medically code, and extract the data.

This architecture gives the site and sponsor each their own specific application and database, allowing CRIO to optimize the user experience for each user type. Publisher and Reviewer work across geographies, which lets sponsors manage a study globally, while the eSource application is operated locally within the sites' regions. Both the site and sponsor retain direct ownership over their respective data sets, unlike a unitary EDC system, where there is only one data set. This dual database structure embodies the principle of ICH-GCP in that the investigators retain control of their source data.

4 Currently the U.S., Canada, EU, and Australia.

Traditional EDC vendors have only one application and one database. Thus, their system cannot house PHI, which makes it difficult to deliver a truly site-native workflow. This means it is not possible for sites to use a traditional EDC system for eConsent, appointment scheduling, text reminders, or rapid medical records retrieval - just to name a few of the workflows capabilities within the CRIO system. All of these workflows collectively and in total drive protocol compliance at the site level. And, furthermore, regulators in EU and elsewhere have voiced skepticism of a system where PIs do not have direct control over their source data.

Most Critically, CRIO's eSource is Already Used by Sites Around the World

Because the CRIO eSource system is inherently site-facing, **over 2,000 sites worldwide in 20+ countries have adopted CRIO's eSource by licensing directly from CRIO**. These sites are using it as their platform of choice across their studies. In total, CRIO is being deployed on nearly 5,000 unique protocols for almost 1,000 sponsors, across every major therapeutic area.

On some trials, upwards of 30% of U.S. sites are using CRIO. For instance, on the Pfizer and Moderna COVID-19 vaccine trials, 27% of the U.S. sites are current CRIO clients. Over the coming years, this percentage will only increase as the site industry matures, with larger networks consolidating locations and adopting eSource to drive standardization.

CRIO site clients are high performing: 40-70% better data quality, greater openness to technology (by definition), 40% higher enrollment[5] and 2x the patient diversity compared to the industry average.

To be clear, Sponsors and CROs can deploy CRIO to non-CRIO sites. CRIO will set these sites up with their own CRIO accounts and provide training. Because CRIO has a proven track record of site adoption, these new-to-CRIO sites will find the system intuitive and easy to use.

5 Based on an analysis comparing CRIO site enrollment to the trial average, as derived from CT.gov data. This most likely reflects self-selection (i.e., sites that invest in CRIO are higher performing to begin with) as well as the fact the sites have more time to focus on enrollment. This would be akin to saying people who attend gyms regularly are healthier - one could envision both correlation and causation factors at play.

What Are the Prerequisites For the CRIO Model?

CRIO recommends three prerequisites.

1. As a site selection criterion, sponsors should make it a condition of the trial that the site utilize CRIO eSource. In CRIO's experience, close to 98% of non-native CRIO sites would accept a trial on these terms.

   Additionally, CRIO will make its worldwide client network available for site selection through the CRIO Connect platform. In most major therapeutic areas, CRIO can offer over 100 active investigators who are eager to take on a study that leverages their chosen platform and relieves them of the burden of EDC entry.

2. Sponsors and CROs need to entrust CRIO to design the source template. Source templates utilize different design principles than traditional eCRF forms. CRIO has constructed Study Design Guidelines that codify best practices and specifically articulate differences in approach between eSource and eCRF. CRIO will build templates to these Guidelines. Doing so maximizes site adoption and thus data quality. Over time, CRIO can train its partners to build eSource templates themselves.

3. Sponsors and CROs should adopt the new Clinical Data Monitoring process for maximum efficiency. CRIO can supply a draft Monitoring Plan. This process leverages the real-time delivery of data by adopting continual, remote and centralized review of site source data, combined with risk analysis from reports that CRIO can furnish using its business intelligence tool. This Monitoring Plan could also incorporate some of the same data checks frequently used by Data Management in the traditional model. CRIO will provide suggested data checks based on the new system and process.

It's Time To Move Forward

The traditional EDC system had its time and place. Through no fault of their own, no EDC vendor could lay claim to the integrity, completeness and accuracy of the endpoint data they house compared to the data collected within CRIO eSource. With CRIO's unique sponsor-facing solution, sponsors can realize these benefits across the entire trial, not just the part of the dataset collected by the forward-thinking sites that happen to already be CRIO clients.

In short, sponsors now have a much simpler, faster and cheaper path to their destination: reliable, protocol-driven endpoint data.

If you have questions about this paradigm, please contact CRIO for a discussion or product demo. We can also introduce you to a certified CRO partner who is aligned with your goals of better, faster data at a reasonable price point.

Explore CRIO at www.clinicalresearch.io.