



Clinical Researcher

The Authority in Ethical, Responsible Clinical Research

April 2023 (Volume 37, Issue 2)

PEER REVIEWED

Electronic Source Reduces Protocol Deviations Compared to Paper Source in Clinical Trials

Ryan Chen, BA; Calvin Nash, BA; Nicole Mastacouris, MS; Garrett Atkins, MPH; Daniel Dolman; Jonathan Andrus, MS; Raymond Nomizu, JD



Web-based clinical research software solutions that let sites configure data templates to specific trial protocols and capture data against those templates in a clinical setting are known as electronic source (eSource) solutions. eSource leverages the use of edit checks and other technology innovations to reduce the incidence of protocol deviations in clinical trials compared to paper source (pSource). While many sites have not yet transitioned away from the traditional pSource model, eSource may represent an efficient systematic alternative to this traditional method. To investigate the effect of eSource at the site level, we report the study of protocol deviations using a commercial eSource solution vs. site-generated pSource across three independent study sites within a large network.

Background

Source data were first defined in section 1.51 of the International Council for Harmonization's Good Clinical Practice (ICH-GCP) guide as the original data in records and certified copies of original records of clinical findings, observations, or other activities used for the evaluation of a clinical trial. Source documents are defined in section 1.52 of the ICH-GCP as the original documents or records that store source data. Source data are fundamental to trials as they are analyzed to establish study outcomes. Attempts have been made to provide guidance on source document generation and completion since source data ultimately become core study data.

In compliance with federal and state law, this documentation should be attributable, legible, contemporaneous, original, accurate, enduring, available and accessible, complete, consistent, credible, and corroborated.{1} These standards were created to protect human subjects and safeguard study integrity. Despite industry standards, some sites struggle to meet all source criteria; in fact, source documents are the most commonly cited document type in findings during monitoring inspections and audits.{2} Therefore, additional consideration is needed to improve initial source document accuracy and reliability.

The State of the Industry for Paper vs. Electronic Source

Like source, the case report form (CRF) is an electronic or paper tool used to collect and store clinical information associated with clinical research protocols. Unlike source, CRFs are typically a secondary data reservoir designed to standardize data collection across multiple sites as specified by the sponsor.{1,3} CRFs are designed by the study stakeholders to capture required endpoints while source is the responsibility of the investigators. Typically, sites transcribe data from site-captured source templates into the CRFs, and sponsors then perform source data verification (SDV) to confirm the accuracy of the CRF transcriptions against the underlying source before locking the CRF data and extracting for statistical analysis.

Presently, electronic CRFs (eCRFs) are auditable digital records of clinical investigation data that can be systematically captured, reviewed, managed, stored, and analyzed.{3} The flexibility of the eCRF and its potential integrability into other medical technological platforms—like the electronic medical record (EMR)—eliminates transcriptional errors during eCRF data entry.{4}

The pharmaceutical industry has largely migrated from paper CRFs (pCRFs) to eCRFs in response to issues of data quality and cost efficiency.^{5} Multiple studies have demonstrated a substantial improvement in data quality via a reduction in data entry errors when eCRFs were implemented over pCRFs.^{6–8} eCRFs allow sponsors and sites to perform point-of-entry logic checks and auto-query generation, which result in faster turnaround times for query resolution.^{8–10}

A more efficient query resolution process translates into shorter post-recruitment study times and reduced study costs due to financial mitigation in data management and site monitoring processes. eCRFs offer several additional advantages, such as instantaneous data submission, ease of handling, and overall efficiency.^{7}

Making the Case for eSource

Use of eSource may offer significant advantages over pSource models. Electronic documentation offers greater accuracy, accessibility, security, and efficiency, and may help reduce the deviation burden of clinical research sites. Benefits of the transition from paper to electronic documentation can be modeled by the industry move from traditional paper models to digital systems.^{3,6–8} This can result in improved workflows, reduced cost, and increased productivity, making electronic documentation a preferable option for many organizations. The transition to eSource may yield similar benefits and should be explored.

Despite these benefits, the adoption of electronic documentation can pose a challenge for facilities that have streamlined their workflow based on pSource models. The implementation of new electronic systems may require significant changes in established practices, potentially causing disruptions and resistance among personnel accustomed to traditional methods. Moreover, facilities that have adopted the eSource model may face integration challenges when attempting to incorporate their records into eCRFs required by study sponsors. This may be exacerbated by the fact that few eSource models have permission for direct transcription, making it difficult for facilities to transfer their data seamlessly into eCRFs. These challenges may be particularly severe for facilities that have limited experience with electronic systems.

To date, no studies have compared the effect of pSource to eSource on protocol deviation rates at the site level. To understand the benefits of eSource, we looked at a group of research sites that were transitioning from pSource to an eSource model and studied the impact of each methodology on the rate of protocol deviations observed with each method.

Methodology

We focused on a commercially available eSource solution (CRIO) being used by a study site network, Benchmark Research, that performs third-party data collection for clinical trials in the public and private sector. As of 2020, this network had helped its clients conduct more than 1,000 clinical trials and interface with more than 40,000 participants, contributing to the development of new vaccines and medical therapies worldwide.^{11} The network's central office services encompass performing study visits and monitoring data entry to sponsors' eCRFs.

We analyzed the anonymized data gathered by the network from three of its large clinical trial sites (in New Orleans, La. and Austin and Fort Worth, Texas) from January 1 through March 22, 2022. During this period, all three sites were in the process of transitioning from pSource templates to eSource templates. Because the sites preferred not to change source data methodology mid-study, each site during this period had a mix of studies using each data collection methodology. The network configured its own study templates using the aforementioned eSource product's built-in configuration modules, which allows sites to build their own schedule of events and configure questions, instructions, and edit checks based on sponsor protocols. No sitewide hardware system upgrades were necessary to facilitate the transition to eSource, as the product is accessible in web-based browsers.

All clinical trials being conducted at the sites during this period were reviewed. Only in-person visits were considered. During this period, the network's central office had in place numerous mature processes governing the identification and classification of protocol deviations. All protocol deviations were identified in a running study-specific protocol deviation log. The network's quality assurance (QA) team then reviewed each deviation and classified them as either site-generated or other-generated. Site-generated deviations are those whose root cause was in site performance, and which were not previously approved by the sponsor. Since the

network already reviews source completion as part of its QA protocol for sponsors, no additional clearance was necessary to access eSource.

In our study, we reviewed the protocol deviation logs across all trials. Trials implementing the eSource solution served as the experimental group and trials using pSource served as the control group. The number of site-generated protocol deviations in both models were compared. The number of protocol deviations per completed study visit was the endpoint. Completed study visits served to calibrate across the relative volume of work performed.

The rate of protocol deviation (RPD) was calculated by dividing the total amount of site-generated protocol deviations by the total number of visits performed. Percent reduction of RPD was calculated as the percent difference (see Table 1).

Table 1: Deviation Rates in pSource vs. eSource

	Source method	Total Visits	Total Deviations	Deviation rate (as percentage)	Reduction
Site 1	Paper	695	10	1.44%	
	CRIO	2394	16	0.67%	54%
Site 2	Paper	1022	8	0.78%	
	CRIO	1766	8	0.45%	42%
Site 3	Paper	1197	11	0.92%	
	CRIO	1312	10	0.76%	17%
Total (Weighted Avgs)	Paper	2914	29	1.00%	
	CRIO	5472	34	0.62%	38%

A one-tailed, two proportion Z test was utilized to determine if using eSource significantly decreased the RPD as compared to pSource. All samples are independent and simply random. The null hypothesis states that the use of eSource does not reduce the RPD. The alternative hypothesis states that eSource decreases the RPD. A significance level of 0.05 was used ($\alpha = 0.05$). Figure 1 demonstrates the formula used to calculate the Z-score for the two proportion Z test. \hat{P}_1 represents proportions from sample 1 and \hat{P}_2 represents proportions from sample 2. \hat{p} represents the total pooled proportion from both samples. n_1 represents the sample size of sample 1, and n_2 represents the sample size of sample 2.

Figure 1: Two Proportion Z Test Formula

$$Z = \frac{\hat{P}_1 - \hat{P}_2}{\sqrt{\hat{p}(1 - \hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Results

The use of eSource decreased the RPD across all sites. At site 1, the RPD using pSource was 1.44% while the RPD with eSource was 0.67%; there was a 54% reduction in deviation ($Z=1.9574$, $\alpha=0.05$, $P<0.025$). At site 2, the RPD was 0.78% and 0.45% with pSource and eSource, respectively; there was a 42% reduction in deviation ($Z=1.1108$, $\alpha=0.05$, $P=0.1335$). At site 3, the RPD was 0.92% with pSource and 0.76% with eSource; there was a 17% reduction in deviation ($Z=0.4305$, $\alpha=0.05$, $P=0.3336$).

The weighted RPD with pSource across all three sites was 1%, while the weighted deviation rate of eSource was 0.62%. There was a 38% reduction in deviation rate when eSource was used as opposed to traditional pSource ($Z=1.8879$, $\alpha=0.05$, $P=0.02938$).

Discussion

The use of eSource platform consistently decreased site-wide deviation rates compared to when conventional pSource was used. Using data from three participating sites, the calculated weighted deviation rate significantly decreased when eSource was used compared to when traditional pSource was used between a set period. The statistically significant decrease in weighted deviation rate demonstrates that eSource use leads to an overall decrease in RPD.

While the decrease in weighted deviation rates was statistically significant using a one-tailed two proportion Z test, this was not always the case at each independent site. At site 1, the reduction in deviation rate was statistically significant at 54%. The reduction in deviation rate at site 2 was not statistically significant at 42%. The reduction in deviation rate at site 3 was also not statistically significant at 17%. While the RPD from eSource use may not be statistically significant at sites 2 and 3, the baseline RPD rates with pSource use are already so low that the

observed site-level improvements with eSource implementation may be difficult to quantify as significant. Site 1 also had nearly double the number of visits implementing eSource compared to sites 2 and 3; this may demonstrate that the efficacy of eSource is tied to the degree of experience a site has working with it.

A nearly 40% reduction in the rate of protocol deviations due to the use of eSource could benefit clinical trials at the sponsor and patient level by improving data quality and patient safety. The improvement in data quality is consistent with the eSource solution's features, such as real-time edit checks and visit window calculations which are designed to automate and flag error-prone processes. The site network's QA team attributes the drop in deviations to these automation features.

Furthermore, eSource improves patient safety by protecting patient privacy. Recommendations for best practices in anonymizing patient-level clinical trial data include the ability to standardize processes across data holders, practically delivering large volumes of data, and cost efficiency.^{12} eSource addresses all of these criteria due to its practicality, which allows a cost-efficient way for different parties to access large quantities of data simultaneously in a secure manner. eSource can further improve cost efficiency by decreasing the downstream effects of protocol deviations, such as prolonged data-cleaning times and associated costs for corrective action plans. The resources spent on these tasks could be reinvested by the sponsor and study stakeholders into other aspects of clinical research.

In addition to decreasing the rate of site deviations, using eSource offers many other advantages over pSource. The site network's QA team reported several advantages from using eSource over pSource charts. When pSource is used, quality control (QC) often requires having site staff scan paper charts and e-mail them to QC personnel, who in return annotate those scanned documents and then return the annotated scans to the site. After the sites correct the errors, the sites must then scan and e-mail back the corrected paper templates for review. Multiple cycles of this process may occur, which can exponentially increase the time spent on these clerical tasks. Furthermore, there is often a delay between data capture and data QC due to the administrative time required to scan and send in paper documents.

By contrast, QA teams can review the eSource platform directly to review the source data shortly after they are collected, or even in real time during an ongoing study visit, and then issue electronic queries for the site to resolve. The site network's system, for example, has built-in change management, so that changed source is highlighted for successive review. Since all data are electronic, multiple stakeholders can view the data at the same time, across geographies, enabling more efficient global workflows.

Further Considerations

Several logistical and administrative challenges may limit the transition to an eSource model. One of the key challenges is the burden associated with initial implementation, as it requires the necessary hardware, software licenses, and network resources. Additionally, sites must train their research staff on the new systems to ensure that legal, regulatory, and quality standards are met, especially during the transitional stage. Sites with smaller volume may find the initial costs of eSource system licensing and personnel training to be uneconomical.

It is also important for study sites to consider the cost efficiency of adopting eSource systems in their circumstances; large-volume sites, such as those that generate a high volume of documents, may benefit more from the transition to digital documentation models than smaller sites. This is because the costs associated with managing paper-based documentation, such as printing, quality assurance, storage, or transportation, can be significant for organizations with large volumes of data. As a result, sites with greater recruitment volume are most likely to benefit from the eSource model, while smaller sites are less likely to experience a significant mitigation of data management costs.

It is important to note that, currently, there is no standardization in the utilization of eSource systems for clinical trials. A study site's choice in software may or may not easily interface with a sponsor's computer systems. While this lack of standardization and potential lack of compatibility may introduce challenges, it may also allow study sites to select software that best suits their needs. As the standardization of eSource systems evolves, it may help simplify the process of choosing eSource systems and improve clinical trial workflow.

Limitations

While this study provides valuable insight, several limitations should be noted. The level of protocol complexity for each study was not assessed and used as an independent variable for analysis. Greater protocol complexity scores could potentially lead to more deviations compared to studies with lower protocol complexity scores. Additionally, we did not consider the number of patients per study, which may be associated with an increased rate of protocol deviations due to a larger volume of study visits. Diagnostic and laboratory data values were not considered as metrics for protocol deviations, as these types of data are usually procured from external contract research organizations and automatically transmitted to the sponsor's eCRF.

Conclusion

eSource platforms may offer a cost-efficient, reliable alternative to traditional pSource documentation models in clinical trials. Besides improving data quality and patient safety, eSource platforms simplify other processes in clinical trials such as recruitment, the QA process, and stipend management for patients. It is important for independent study sites to consider the cost efficiency of implementing eSource systems; recruitment rates, data management cost analysis, software licensing, and hardware requirements may limit access to eSource platforms depending on site volume. As the landscape of eSource systems evolves, more research is needed to identify other independent variables that may influence the effectiveness of such solutions.

Disclosures

Jonathan Andrus and Raymond Nomizu are currently affiliated with Clinical Research IO (CRIO). Garrett Atkins and Daniel Dolman are currently affiliated with Benchmark Research. Benchmark Research is a third-party firm which provided data to CRIO for this study, and has no financial gain from the research being conducted.

References

1. European Medicines Agency. 2002. ICH Topic E6(R1). Guideline for Good Clinical Practice. Note for Guidance on Good Clinical Practice. https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e6-r1-guideline-good-clinical-practice_en.pdf
2. Bargaje C. 2011. Good documentation practice in clinical research. *Perspect Clin Res* 2(2):59–63. doi:10.4103/2229-3485.80368
3. U.S. Department of Health and Human Services, Food and Drug Administration. 2013. Guidance for Industry—Electronic Source Data in Clinical Investigations. <https://www.fda.gov/media/85183/download>
4. Mitchel JT, Kim YJ, Choi J, et al. 2011. Evaluation of Data Entry Errors and Data Changes to an Electronic Data Capture Clinical Trial Database. *Drug Inf J* 45(4):421–30. doi:10.1177/009286151104500404
5. Fleischmann R, Decker AM, Kraft A, et al. 2017. Mobile electronic versus paper case report forms in clinical trials: a randomized controlled trial. *BMC Med Res Methodol* 17(1):153. <https://doi.org/10.1186/s12874-017-0429-y>
6. Malik I, Burnett S, Webster-Smith M, et al. 2015. Benefits and challenges of electronic data capture (EDC) systems versus paper case report forms. *Trials* 16(Suppl 2):P37. doi:10.1186/1745-6215-16-S2-P37
7. Zeleke AA, Worku AG, Demissie A, et al. 2019. Evaluation of Electronic and Paper-Pen Data Capturing Tools for Data Quality in a Public Health Survey in a Health and Demographic Surveillance Site, Ethiopia: Randomized Controlled Crossover Health Care Information Technology Evaluation. *JMIR Mhealth Uhealth* 7(2):e10995. doi:10.2196/10995
8. Pavlović I, Kern T, Miklavcic D. 2009. Comparison of paper-based and electronic data collection process in clinical trials: costs simulation study. *Contemp Clin Trials* 30(4):300–16. doi:10.1016/j.cct.2009.03.008
9. Le Jeannic A, Quelen C, Alberti C, Durand-Zaleski I; CompaRec Investigators. 2014. Comparison of two data collection processes in clinical studies: electronic and paper case report forms. *BMC Med Res Methodol* 14(7). doi:10.1186/1471-2288-14-7
10. Rorie DA, Flynn RWV, Grieve K, et al. 2017. Electronic case report forms and electronic data capture within clinical trials and pharmacoepidemiology. *Br J Clin Pharmacol* 83(9):1880–95. doi:10.1111/bcp.13285
11. <https://www.benchmarkresearch.net/>
12. Tucker K, et al. 2016. Protecting patient privacy when sharing patient-level data from clinical trials. *BMC Med Res Methodol* 16(1):5–14.

Ryan Chen, BA, (Ryan.Chen5@umassmed.edu) is a medical student at the University of Massachusetts Chan Medical School. He earned his bachelor's degree in biochemistry and molecular biology at Boston University.

Kalvin Nash, BA, is a medical student at The Johns Hopkins University School of Medicine.

Nicole Mastacouris, MS, is a medical student at the Loyola University Chicago Stritch School of Medicine.

Garrett Atkins, MPH, works at Benchmark Research.

Daniel Dolman works at Benchmark Research.

Jonathan Andrus, MS, is President and COO of Clinical Research IO and Owner of Good Compliance Services, LLC.

Raymond Nomizu, JD, is Co-Founder and CEO of Clinical Research IO.