



Mats Sundgren,
PhD
Industry Science
Director, i-HD



Gynet Santiago
Clinical
Research
Coordinator,
Memorial
Sloan Kettering
Cancer Center
(MSK)



**Joseph
Lengfellner**
Sr. Director,
Memorial
Sloan Kettering
Cancer Center
(MSK)

Streamlining Clinical Trials with eSource: Insights from MSK

Memorial Sloan Kettering use case explores the potential of EHR-to-EDC.

Today, clinical trial data collection faces increased complexity due to data duplication between research systems and electronic health records (EHRs), particularly in oncology studies. Manual processes, such as data entry, consume significant time and resources. Over 50% of clinical trial data is duplicated between research systems and hospital EHRs, with around 20% of total study costs typically allocated to data duplication and verification.¹

The number of data points collected in oncology trials has surged dramatically over the past decade, driven by advances in genomic technologies, digital health tools, and personalized medicine. While this increase in data volume enhances our ability to understand and treat cancer more effectively, it also presents significant challenges in data management and analysis. In typical Phase III oncology trials, approximately 10,000 data points are collected per patient. However, in trials incorporating genomic and digital health data, this number can multiply ten to a hundredfold.^{2,3}

As a result, study sites, particularly clinical research coordinators (CRCs), face increased workloads, complex data management tasks, and the need for specialized training, leading to higher personnel costs. Sponsor companies incur additional expenses due to investments in advanced data management tools, extended trial timelines, and source data verification (SDV). To mitigate these challenges, sites and sponsors must explore innovative solutions, collaborate with stakeholders and vendors, and ensure regulatory compliance to improve efficiency, reduce costs, and enhance cancer treatment outcomes.

EHR to electronic data capture (EHR-to-EDC), or eSource technology and processes, offers a promising solution, delivering significant time and cost savings by streamlining data transfer.^{4,5} The European EIT Health-supported EHR2EDC project estimated a total time of three minutes per data point in oncology trials, demonstrating substantial efficiency gains.^{6,7,8}

This paper presents a use case from Memorial Sloan Kettering Cancer Center (MSK), emphasizing the critical role of CRCs. It highlights the burden-

some nature of current double-entry processes and supports the three-minute-per-data-point estimate. Additionally, it revisits the value generated for investigational sites and sponsors by introducing this technology, exploring the transformative potential of eSource for future clinical trials.

Methodology

The study employs a qualitative structured interview guide to understand the data entry process during clinical trials. An in-depth, semi-structured interview was conducted with an experienced CRC to gather insights into workflows, challenges, and perceptions regarding data entry tasks. Open-ended questions allowed for rich, nuanced responses, capturing the complexity and variability of the process. The validity and generalizability of the interview outcomes were assessed and found to correspond well with CRC practices at MSK.

CRCs: Key players in trial success

Gynet Santiago is a seasoned CRC at MSK, with 16 years of experience and involvement in over 50 clinical studies. She summarizes her role as, “I am responsible for collecting, extracting, and entering data for research projects, databases, and clinical trial protocols. This includes reviewing patient charts, existing databases, and other sources within a specific timeframe.”

The CRC is pivotal in executing clinical trials, with responsibilities spanning various trial phases to ensure compliance with regulatory requirements and ethical standards, while prioritizing patient safety and data integrity. During trial execution, CRCs manage participant recruitment, screening, and enrollment, ensuring informed consent is obtained. They accurately collect and document clinical data, monitor participants for adverse events, and ensure adherence to the trial protocol. CRCs act as liaisons between participants, investigators, and sponsors, maintaining communication and addressing issues. They also ensure regulatory compliance by maintaining accurate documentation and assisting

with audits. These responsibilities are crucial for the smooth operation of trials, data integrity, and participant safety. Additionally, CRCs often manage multiple studies simultaneously, requiring efficient context switching throughout the week.

Navigating the complexities of EDC data entry

In clinical trials, EDC data entry is a multi-faceted process involving several crucial steps performed by a CRC (see Figure 1). This process, while systematic, is often challenging due to its time-consuming nature and the meticulous attention to detail required. The following is an overview of how a CRC manages the process of retrieving and manually entering data into an EDC system.

Finding and allocating time. One of the first tasks a CRC faces is finding and allocating time for EDC data entry. This often involves scheduling data entry activities around other demanding clinic responsibilities. For instance, a CRC might set aside Friday afternoons every two weeks for data entry once all other clinic activities have been addressed. Prioritizing EDC data entry among other tasks can be challenging, as the coordinator must balance immediate clinical duties with the time-consuming process of data entry. The irregularity and unpredictability of clinical activities can make it difficult to allocate consistent time slots for this purpose. Or as Santiago puts it, “Each day is based on what the priority is for that day, that’s how I know how to allocate my time. So, for example, if there’s a data lock coming up or an audit coming up for a particular study, that’s always going to be a top priority for me for that day or for that week.”

Patient information retrieval and manual searching for patient information. Once time is allocated, the CRC manually searches for patient information within EHRs. Unlike records specifically designed for clinical trials, these records are complex and primarily geared toward billing, with patient information spread across various locations, pages, and fields, including free-text entries. Confidence in the data source can vary, and locating the correct data points requires a detailed, page-by-page review, making this step particularly time-consuming.

The complexity of EHR systems presents additional challenges. Patient information is not always recorded in a standardized manner, which means that the CRC must have a thorough understanding of how different healthcare providers document data.

EDC Entry Elements

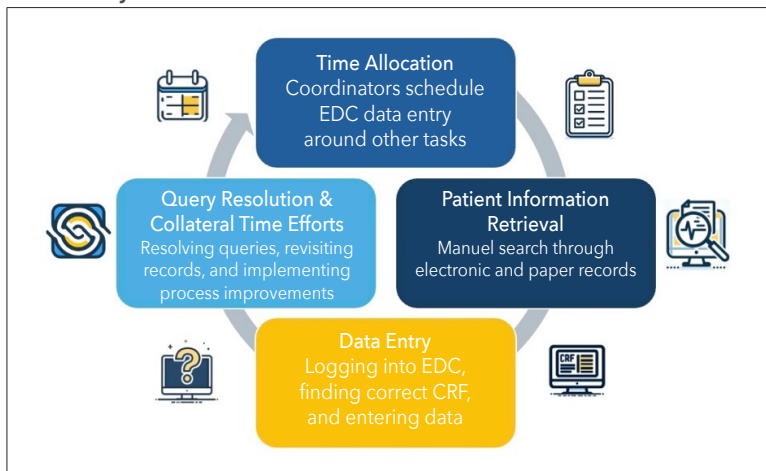


FIGURE 1. The key steps for clinical research coordinators in EDC data entry.

SOURCE: Sundgren et al.

This lack of standardization can lead to difficulties in finding specific data points, as information might be spread across multiple pages or hidden within lengthy notes. Santiago explains: “The way that EHR is set up is, once you open the records, it displays different categories of information in a particular folder, and it can be a challenge to locate the right data point because of the large amount of information that you must read through to find that correct data point.”

Data entry. After retrieving the necessary patient information, the CRC logs into the sponsor’s EDC system using the appropriate URL and credentials. The next step involves navigating the EDC system to find the correct case report form (CRF) for data entry. This can be challenging due to the system’s complexity and the need to ensure that the information entered matches the patient’s records accurately. The CRC must be familiar with the EDC system’s layout and functionality to efficiently locate the correct CRF pages. It is a demanding process as Santiago explains: “It really is familiarizing yourself with the EDC. So EHR is one aspect of it where you’re able to know where it is. And now you must go to the EDC and confirm, okay, I’m entering this information in the proper place that it needs to be for the database.” And sometimes errors can multiply—“There’s challenges when you’re entering data into the EDC and inaccurate system queries require information that’s not always available,” adds Santiago.

Challenges and post-data entry review. After entering the data, the CRC reviews it for

A Cost and Time Saver?

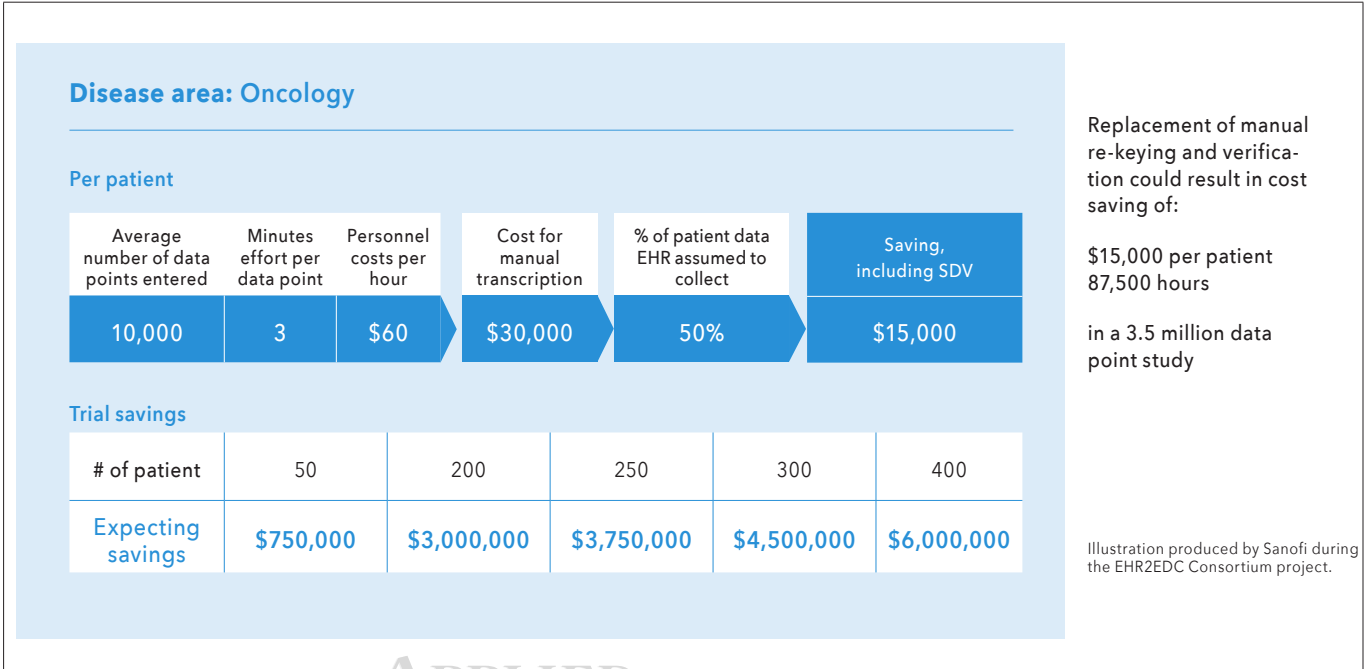


FIGURE 2. A retrospective analysis to estimate the average total time spent per data point per patient in a Phase II-III oncology study.

SOURCE: Sundgren et al.

accuracy and resolves any discrepancies. This review process might involve understanding the data’s meaning, consulting with other study team members, and performing self-quality assurance checks to ensure data integrity. Ensuring data accuracy is crucial, as any errors can lead to significant issues in the trial’s results and regulatory compliance. Therefore, the CRC must be diligent in verifying that the data entered the EDC system accurately reflects the information in the patient’s medical records. Santiago highlights one common issue: “Sometimes the same data is being entered into different places in the EDC. So, I’m always pushing back and saying, why are we entering the same data in multiple places?”

Query resolution and collateral time efforts. Queries from the sponsor or contract research organization (CRO) can arise weeks or months after initial data entry, requiring the CRC to spend additional time for resolution. This involves revisiting medical records to locate specific data points, often repeating original data retrieval steps. If significant time has passed, the CRC may need to refamiliarize themselves with the patient’s records and data points.

Managing queries is time-consuming and demanding, often due to double data entry. As Santiago explains, “Query resolution is part of the role. I would say 50% is data entry, and 50% can be query resolution, especially if a study is active.” Queries can be frequent and complex. Adds Santiago: “Queries can be raised weekly or even daily, and sponsors usually require them to be resolved within one or two weeks.” This process involves thorough research, confirming data accuracy. During this process, the CRC may also need to check or verify with other study personnel to address the query. Monitoring visits by sponsors or CROs are crucial for ensuring data quality and protocol adherence. These visits involve reviewing data entered into the EDC system and verifying its accuracy against source documents. The CRC may also implement process improvements based on queries and share these with the broader study team.

The interview highlights the multifaceted role of a CRC, including data entry, patient information retrieval, and managing queries. The CRC must navigate systems such as EDCs, clinical trial management systems, and EHRs, often facing challenges such as locating specific data points and handling system queries. Effective query resolution is

crucial, often consuming as much time as initial data entry, underscoring the need for efficient processes and tools to streamline these tasks.

Key CRC task during trial execution

Three minutes and 5,000 hours: A significant investment of time

During the European EIT Health-sponsored EHR2EDC project from 2018 to 2020, AstraZeneca and Sanofi collaborated on a retrospective analysis to estimate the average total time spent per data point per patient in a Phase II-III oncology study. The combined analysis determined that local study personnel spend at least three minutes per data point. This time encompasses all the tasks we covered in the interview, including data searches, manual entry, checking, reviewing, and resolving queries.

To put this in context, for an oncology study involving 10 patients, each with 10,000 data points, this equates to about 5,000 hours of work. Scaling up to an ordinary Phase III study with over 200 patients, the time and associated costs increase exponentially (see Figure 2 on facing page).

Based on the MSK use case, we can conclude that straightforward data points can be entered within three minutes total time if the data is normal. However, complex data points, such as those involving adverse events or abnormal lab values, can take significantly longer. Santiago explained that query resolution is a significant part of her workload, often consuming 50% of her time. Queries can range from simple confirmations to complex requests requiring additional information from treating physicians or further investigation. She noted that resolving queries can be particularly challenging and time-consuming when dealing with historical data. She agreed that the estimation of three to five minutes per data point is reasonable for straightforward cases but emphasized that overall time varies. Santiago also highlighted the potential for eSource technology to streamline data entry and reduce manual processes and query resolution, potentially saving more than 50% of the current time required. This underscores the variability and complexity of data management in clinical research.

What is eSource technology and what can it do?

EDC-to-EDC, or eSource, is a streamlined process specifically designed to transfer routinely collected EHR data relevant to clinical research, particularly

Moving Beyond Manual

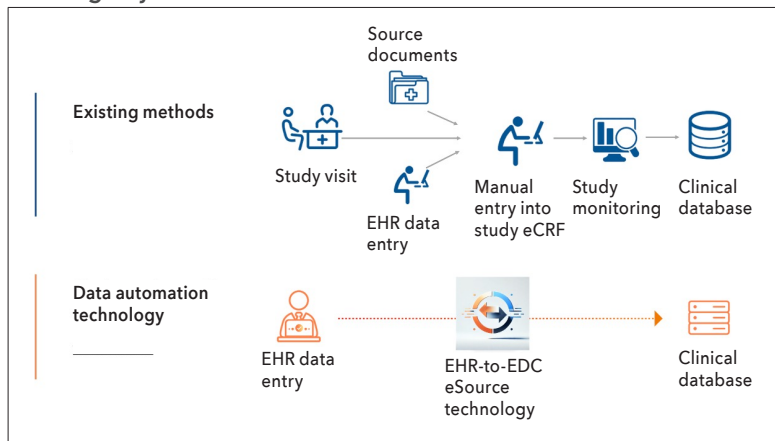


FIGURE 3. The eSource approach can enhance overall clinical trial efficiency.

SOURCE: Sundgren et al.

data collected at investigator/hospital sites during clinical trials (including randomized controlled trials). This transfer occurs with patient consent, meeting regulatory eSource requirements, into either an EDC system or the sponsor's study database. It's important to emphasize that EHR as eSource does not involve replacing staff or vendors for data entry at sites, duplicative data capture, or extraction of datasets from the EHR.

In the case of the “eSource data flow” and the transfer of mapped study data from EHR to study data base, it will always be in the control of the CRC, who will have oversight and final say of data transfer. In this perspective, the eSource technology should be viewed as a “virtual research assistant.”

Contrasting traditional EDC trial with eSource-enabled trial

The eSource approach improves clinical trial conduct. By automating data transfer from existing EHR directly to EDC systems, it eliminates manual entry, reducing investigator site burden and enhancing data quality (see Figure 3). This automation minimizes transcription errors, improving data accuracy and streamlining and sometimes eliminating SDV, leading to cost savings and faster trial timelines. Furthermore, eSource fosters collaboration by integrating EHR and EDC systems. Sponsors gain valuable insights into site data collection processes, enabling the optimization of electronic CRFs to align with specific EHR capabilities. This collaborative approach ultimately enhances overall clinical trial efficiency.

Ongoing studies at MSK utilizing data from past clinical trials to simulate feasibility and outcomes with eSource technology have shown promising results. These findings strongly support the value generation as outlined in this report.

Outlook

Consider an example: An industry-sponsored oncology trial includes three investigational sites, each with 10 patients. All sites meet the minimum eSource HCO requirements for early adopters, including EHRs with structured data and coding standards for data domains such as laboratory results, vital signs, and medications. They have adopted the FHIR® (Fast Healthcare Interoperability Resources) standard to ensure smooth information flow across platforms and organizations.

In this example, all sites are eSource-enabled, encompassing 300,000 data points (10,000 per patient) across the mentioned domains, accounting for over 50% of the entire study data. Implementing eSource technology in this study is expected to significantly reduce site burden, cutting site staff time by 50%, equivalent to more than 2.5 full-time equivalent months per site. Most issues from double entry and query resolution will be eliminated, leading to faster trial completion. There will be a significant reduction in SDV from the sponsor's side and a notable decrease in study costs, estimated at \$150,000 per site.

In the coming years, hybrid eSource-enabled trials will include more eSource-capable sites among traditional “double-entry EDC” sites. This hybrid approach may not immediately reduce the study database lock time, but it will enhance efficiency and reduce site burden at eSource sites. A transformational shift will occur when trials become 100% eSource-enabled. These trials will likely have fewer but larger super-sites, transferring over 80% of all study data as eSource. This shift will enhance value generation for investigational sites by reducing the total time per data point by over 90%. This improvement in efficiency and resource utilization will almost completely eradicate SDV and other time-consuming processes. Fully eSource-enabled studies are expected to reduce total study costs by over 20%, achieving database lock significantly faster.

This MSK use case highlights the significant benefits of eSource technology and corroborate time metrics in supporting CRC and site personnel during trial execution. The widespread adoption of eSource technology represents a paradigm shift in hospital

clinical trials, optimizing processes, saving time, and reducing double entry. If innovations such as eSource, or EHR-to-EDC, were fully implemented, they could also enable research that might otherwise not happen, paving the way for the development of medicines that otherwise would never reach patients. 🌟

References

1. Kalankesh, L.R.; Monaghesh, E. Utilization of EHRs for Clinical Trials: A Systematic Review. *BMC Med. Res. Methodol.* 2024. 24 (70). <https://bmcmedresmethodol.biomedcentral.com/articles/10.1186/s12874-024-02177-7>
2. Passut, C. Oncology Trials Outpacing Rest of the Field in Complexity and Duration. CenterWatch. May 10, 2021. <https://www.centerwatch.com/articles/25599-oncology-trials-outpacing-rest-of-the-field-in-complexity-and-duration-study-shows>
3. Monira, S.; Toral, M. Possibilities and Challenges for Digital Medicine in Oncology Clinical Trials. *Clinical Researcher.* August 15, 2023. <https://acrpnnet.org/2023/08/possibilities-and-challenges-for-digital-medicine-in-oncology-clinical-trials/>
4. Parab, A.A.; Mehta, P.; Vattikola, A.; et al. Accelerating the Adoption of eSource in Clinical Research: A TransCelerate Point of View. *Ther Innov Regul Sci.* 2020. 54 (5), 1141-1151. <https://pubmed.ncbi.nlm.nih.gov/32128701/>
5. Garza, M.; Myneni, S.; Nordo, A.; et al. eSource for Standardized Health Information Exchange in Clinical Research: A Systematic Review. *Stud Health Technol Inform.* 2019. 257, 115-124. <https://pubmed.ncbi.nlm.nih.gov/30741183/>
6. Sundgren, M.; Keong Wong, W. eSource Interoperability Between EHR and EDC. *Applied Clinical Trials.* 2023. 32 (8), 21-23. <https://www.appliedclinicaltrials.com/view/esource-interoperability-between-ehr-and-edc>
7. Ammour, N.; Griffon, N.; Djadi-Prat, J.; et al. TransFAIR Study: A European Multicenter Experimental Comparison of EHR2EDC Technology to the Usual Manual Method for eCRF Data Collection. *BMJ HCL.* 2023. 30 (1). <https://informatics.bmj.com/content/30/1/e100602>
8. Beresniak, A.; Schmidt, A.; Proeve, J.; et al. Cost-Benefit Assessment of Using Electronic Health Records Data for Clinical Research Versus Current Practices: Contribution of the Electronic Health Records for Clinical Research (EHR4CR) European Project. 2016. *Contemp Clin Trials.* 46 (1), 85-91. <https://pubmed.ncbi.nlm.nih.gov/26600286/>