ƷS MEDIDATA

Solving the EHR-to-EDC Challenge: A Scalable-First Approach

Table of Contents

Introduction

As most readers are likely aware, Electronic Data Capture (EDC) systems have been used in clinical trials for many years to collect, clean, transfer, and process data, primarily through the use of electronic case report forms (eCRFs). However, despite the significant progress and widespread adoption of Electronic Health Records (EHRs) following the enactment of the Health Information Technology for Economic and Clinical Health (HITECH) Act of 2009, their integration into clinical research has been comparatively slow. EHRs are real-time, patient-focused records that incorporate information from various healthcare providers involved in a patient's care. The widespread adoption of EHR systems, with over 95% of U.S. hospitals implementing them, has resulted in improved efficiencies, increased reimbursements, and enhanced patient care (ONC, 2019).

EHRs contain a rich collection of information that can provide several benefits when leveraged for clinical research purposes, including (McCord, 2019; Nordo, 2019):

Reduced trial costs

Faster trial completion

Increased generalizability of results

Enhanced recruitment

Feasibility assessment of research protocols

Improved outcome analysis

Expanded scope of research

Earlier identification of safety events

EHRs also facilitate the connectivity and interoperability required for the exchange of electronic health information with other systems, such as in the case of increasing engagement between regional- and national-level Health Information Exchanges (HIEs). HIEs are secure electronic central repositories of patient data that allow relevant stakeholders to safely access and seamlessly transfer a patient's medical information, with some even allowing users to tap into EHRs without point-to-point integrations (ONC, 2023).

Why Solve the EHR-to-EDC Challenge?

While clinical researchers have long sought to repurpose EHR data at scale for conducting clinical studies, several hurdles have hindered this progress, including poor interoperability between EHRs and other systems, poor data quality, and the intensive effort required by sites to manually transcribe electronic data from EHRs (and other systems) to EDC systems (Collen, 1990; Garza, 2020). These challenges are exacerbated by the fact that the sheer volume of data being collected is constantly increasing, with Phase III trials collecting 3.6 million data points on average—a sevenfold increase in volume from 20 years ago (Tufts CSDD, 2021). Additionally, healthcare systems often store data in systems outside EHRs, with Johnson et al. (2018) reporting that more than half of the hospitals in their sample used four or more electronic methods to transfer summaries of their care records to organizations outside their hospital system.
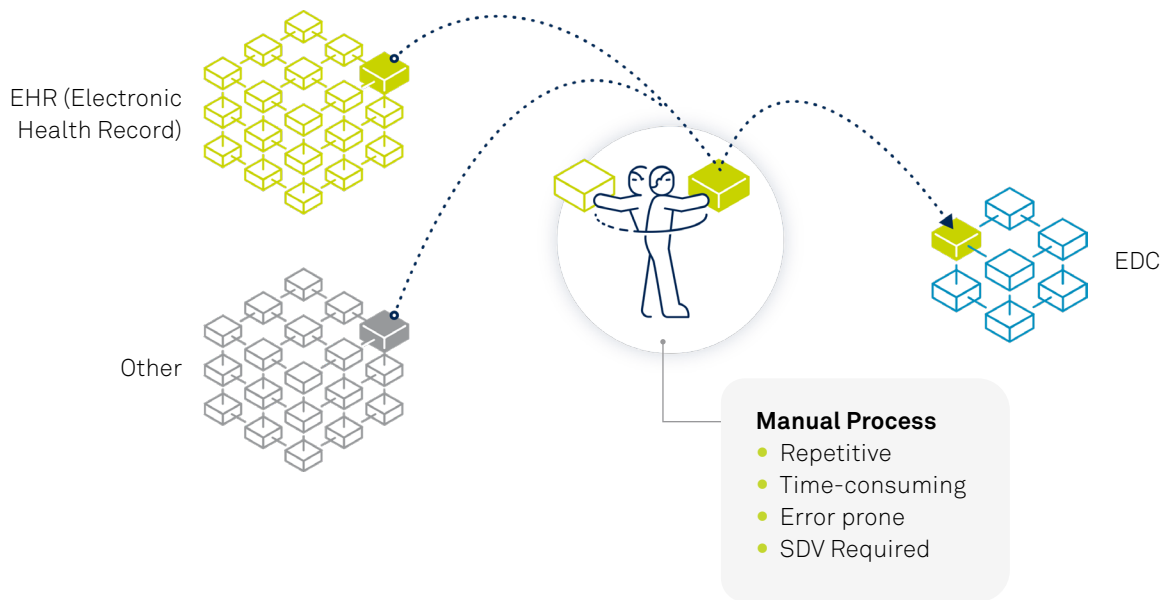
Overall, a pain point that is commonly highlighted by site research coordinators is:

"

Why am I manually entering data that is already available somewhere else?"

Given that approximately 70% of the data entered into EDC systems are duplicated from EHRs and other source systems, this reaction is not surprising (Sundgren, 2021). Today's industry standards require research coordinators to manually transcribe large volumes of data already existing in EHRs for the purpose of entering them into EDC systems. This process involves identifying and reviewing specific patient and visit records in the EHR (as well as in systems outside the EHR) and then determining the data that need to be transferred to the EDC from specific reports. This is carried out by toggling between the EHR (or another system outside the EHR) and EDC systems to identify and capture accurate information, which often involves copying EHR data onto a sheet of paper before manually entering them into an EDC system (Figure 1). Unsurprisingly, this manual re-entry of EHR data into EDC systems is an enormous challenge for sites. According to the results of a survey conducted by the Society of Clinical Research Sites, 98% of sites indicated that they manually re-enter data into EDC systems, while nearly 70% of them stated that more than 50% of their EDC data is re-entered using existing EHR data (Society for Clinical Research Sites, 2022). In addition to the huge amount of time that this process consumes, it also introduces transcription errors, resulting in additional time and effort spent on raising and resolving queries—a laborious undertaking, given that the average number of queries per Phase III study (n = 20) has been found to be 96,980 (Stokman, 2021).

Figure 1. Today's industry standard involves manual data (re)entry.

EHR (Electronic Health Record)

Other

EDC

Manual Process
- Repetitive
- Time-consuming
- Error prone
- SDV Required

The EHR-to-EDC challenge, and indeed the overall data entry challenge, continues to plague almost all personnel involved in the clinical trial ecosystem. According to Medidata's research, solving this challenge at scale can address specific concerns raised by sites, sponsors, and partners, including the following:

SITE CONCERNS: Data already existing in the EHR, requirement for excessive data entry, high frequency of toggling.

SPONSOR CONCERNS: High monitoring costs, too many data entry errors, and desire to keep sites satisfied.

PARTNER CONCERNS: Industry demand for better technologies to improve processes, desire to keep sites satisfied.

While several one-off solutions have been developed in an attempt to allow for data extraction from EHR to EDC, they are unfortunately not scalable for a variety of reasons (Table 1). One such reason is the requirement for a significant level of effort at the site since the solutions have been limited to single-EHR and single-institution implementations. Furthermore, these limited solutions require an IT infrastructure to be developed, systems to be installed, data to be mapped, and many data agreements to be negotiated between the site, the vendor, and the sponsor for each specific trial. After completing all these initial steps, the system can finally be connected to the EDC system.

Table 1: Challenges associated with scaling traditional EHR-to-EDC solutions.

POINT TO POINT SITE CONNECTIONS

Need dedicated IT resources from sites

Individual data sharing agreements

Time spent on configuration, testing, reconfiguration

Often involves EHR vendor resources and additional costs to sites

DATA MAPPING

Healthcare data is inherently complex

Not effective at scale given time and effort required

Inability to control for all variables in data, human verification still needed

SLOW ADOPTION

Still no universal and widely adopted standard

Ambiguity in current standards

EHR to EHR configuration variations

While many of these solutions are EDC-agnostic, they still require significant amounts of time and effort for installation and configuration, before even beginning a study. This typically takes multiple months to accomplish, and implementing these solutions is generally a heavy burden on sites.

Furthermore, IT resources, data sharing agreements, and the time spent testing and configuring the solution often involve an EHR vendor. While the industry is dominated by major enterprise vendors, such as Epic, Cerner, Centricity, and Allscripts, there are many small and specialty vendors as well, with one estimate suggesting that there are more than 1,000 EHR systems available on the market (Fouque, 2022).

In addition, since many enterprise EHRs are site-hosted and site-configured, variability among even single implementations of the same EHR is often observed. Overall, a primary challenge in the sharing of EHR data for clinical research is the different and inconsistent data standards followed by organizations using different EHRs and EHR vendors, and between clinical care and research sites. With such enormous variability between, and within, healthcare and clinical research systems, a major challenge has been to identify a suitable method to handle this variability in the context of standards that lack flexibility and are not readily adaptable to the diverse needs of different stakeholders across its many potential use cases.

Industry Changes Have Finally Enabled the Development of a Scalable Solution

Fundamental regulatory and data standard changes have provided significant opportunities for organizations to integrate EHRs and EDCs and finally develop scalable solutions to solve the EHR-to-EDC challenge.

The 21st Century Cares Act (the "Cures Act"), passed in 2016, sought to reduce the regulatory burden associated with the use of EHR systems through provisions focused on advancing interoperability and requiring developers to not engage in information blocking or preventing the access, exchange, or use of electronic health information. This legislation also includes the provision that every healthcare organization within the country must share data with entities that have rightful access to it, and this sharing must be digitally conducted in a format that adheres to specific structures and standards.

The Office of the National Coordinator for Health Information Technology (ONC) was tasked with implementing the key provisions of the Cures Act designed to advance interoperability, including supporting the access, exchange, and use of electronic health information and addressing occurrences of information blocking (Federal Register, 2020). Furthermore, the ONC has adopted API-enabled "read" services and recommended using the HL7® Fast Healthcare Interoperability Resources (FHIR®) standard, which lays down the established technical requirements for software developers while also offering a set of accepted implementation specifications (ONC, 2020). The HL7 FHIR standard, often referred to as the "next generation" standards framework (Garza, 2020), is a simple yet highly versatile benchmark that allows for the accommodation of diverse healthcare processes, be it as a stand-alone standard or in tandem with other widely used standards. HL7 FHIR is increasingly being used to support healthcare data exchange using a variety of applications, with all U.S. healthcare organizations supporting some level of exchange utilizing HL7 FHIR.

Overall, HL7 FHIR has enabled the health applications market to leverage data from any EHR in a standard format, while also allowing for a standardized interface that functions seamlessly with EHR operating systems by feeding information directly into the provider workflow to facilitate document-based exchange.

While the industry has been trying for many years to find the best methodology, practice, workflow, and approach for extracting data from EHRs and feeding them into EDC systems, addressing this complex industry-wide challenge requires a novel and multi-pronged approach because mapping existing EHR data to the information required by EDC systems is an extremely difficult task, with even the most advanced solutions not able to map anywhere close to 100% of the necessary data.

For instance, Ganza et al. (2020) mapped clinical study data elements from three therapeutically diverse studies (labeled "A," "B," and "C") to HL7 FHIR resources. All exchangeable contents were defined as "resource," representing a small collection of clinical and administrative information that could be captured or shared, such as "Observation," "Patient," and "Specimen." With regard to FHIR coverage, this investigation revealed considerable variability in the percentage of data that was mappable across the three studies, which was 45%, 55%, and 79% for studies "B," "A," and "C." The results of this study offer valuable insights into the coverage and utility of the FHIR standard, further underscoring the need for multi-pronged approaches that do not rely solely on FHIR-based mapping (but rather, leverage FHIR as one of many integration standards).

The enactment of the Cures Act paved the way for an efficient EHR-to-EDC integration solution by codifying data exchange standards, which can be used by industry, providers, payers, and technology vendors to guide the design of their health IT systems for seamless information flow. This important regulatory tailwind has been the fundamental driver of organizations becoming more standards-compliant and achieving ubiquitous interoperability. The fact that the use of HL7 FHIR and other standards-based interoperability mechanisms is now federally mandated for all healthcare organizations across the U.S. has been a game-changer for retrieving data from a diverse set of clinical sources.

Additionally, over the past decade, there has been an increasing number of healthcare providers participating in HIEs, as they offer a variety of benefits, such as helping facilitate coordinated patient care, reducing duplicative treatments and testing, and increasing efficiency by eliminating unnecessary paperwork and handling.

Lastly, regulatory agencies have long recognized the benefits of EHR data in clinical research and have encouraged the development of solutions that advance their use in such research. While the FDA has published guidance on the use of EHR data in FDA-regulated clinical studies (FDA, 2018), the European Commission has adopted a recommendation on a European electronic health record exchange format to facilitate the cross-border interoperability of EHRs in the European Union (European Commission, 2022). Furthermore, multiple large-scale efforts have been underway for several years, to specify patient clinical information standards in a manner that enables computable interoperability among diverse systems across the world (Nordo, 2019).

These changes in the industry landscape have enabled the development of a scalable multi-disciplinary approach to solving the EHR-to-EDC challenge that is less focused on mapping data and more focused on presenting data to users. This solution is further described in the following section.

Medidata's Multi-Pronged Approach to Overcome the EHR-to-EDC Challenge

Although previous attempts at EHR-to-EDC integration were too complex to solve the problem practically, it is still vital to enhance data quality and release some of the burden on sites by making data entry simpler and faster, while also reaping the previously discussed benefits of using EHR data in clinical research.

Medidata's priority, in this context, is "scalability first." Consistent with this priority, it adopted a multi-pronged approach when building its EHR-to-EDC solution—**Rave Companion** (enabled with **Medidata Health Record Connect**) Rave Companion, a simple-to-use browser plugin, reduces data entry efforts for clinical trial sites by making it easier and faster to

populate Rave EDC forms with source data from other systems (e.g., EHR, CTMS, eSource) and documents (e.g., lab values in a spreadsheet). Medidata Health Record Connect is a healthcare data interoperability engine for securely and compliantly acquiring, transforming, and exchanging EHR data. Combined, this solution* accelerates data entry, eliminates transcription errors, and reduces queries, enabling sites to spend more time on higher-value research activities and with their patients.

*Note: For interventional trials only. All users must be employed by the institution where the data is accessed, under the supervision of a physician or PI (Principal Investigator) at all times, and be involved in monitoring safety events and study outcomes for patients enrolled in research.

Rave Companion

- ✓ Replicates the Rave EDC form in a floating 'always on top' window

- ✓ Eliminates bouncing between applications

- ✓ **90% faster** completion of Rave EDC forms*
  *using Health Record Connect

- ✓ Fewer errors and queries

ABC Pharma XYZ Patient Linked ●
002 EMR Linked ●
Patient Details

[All] [Date]
10/30/2022 - 30 + Day(s)

Lab Results

Albumin 6
3.9

Sodium 6
135
Autofilled from:09/27/2022 12:30 pm

Potassium 4
4.4
Autofilled from:09/27/2022 12:30 pm

Presents patient EHR data using Medidata Health Record Connect

Just click or click and drag to capture any data from any system or document

Or simply type, check boxes, or select from pull-downs to manually complete fields

When a Clinical Research Coordinator (CRC) at the clinical trial site opens a Rave EDC form, Rave Companion automatically pops up a compact replica of the form. If the site is connected to Medidata Health Record Connect, they can link the patient in Rave EDC to their healthcare record by providing one-time patient identification information. Once linked, Companion queries Health Record Connect to find and present matching data for the form, enabling the form to be completed over 90% faster than with manual data entry.

Unlike other EHR-to-EDC solutions that can take months to implement on each study, Medidata Health Record Connect doesn't require site-by-site EHR system integrations and data transfer agreements or complex study- and system-specific EHR-to-protocol mappings, so it's up and running immediately.

If the site isn't connected to EHR data or there's data that's in another system or document, the CRC can still use Rave Companion's always-on-top, floating form to take it alongside the other window and complete the form without having to switch back and forth. And the CRC always has eyes on the data. Nothing is saved until the CRC clicks 'Save' on the form in Rave EDC. All Rave Companion actions are audited so that traceability is maintained.

Rave Companion has been designed with the user in mind. It is highly focused on presenting data to the user. Notably, Companion directly writes data into Rave EDC forms but provides users with the opportunity to review and make edits to the data before saving it. This is important since our initial research found that users did not prefer a direct data flow from the EHR into the EDC system. In this way, users can first review the data to ensure it reflects the right value, the right visit, the right date and time, etc., thus ultimately facilitating the completion of the form with no requirement for manual data entry. Furthermore, provenance is maintained in the audit trail for every field in the eCRF to easily determine whether the data were manually entered into Rave EDC, or populated via Rave Companion, either with EHR data provided by Medidata Health Record Connect or with Rave Companion's 'click to capture' feature.

Figure 2: With Rave Companion (enabled with Medidata Health Record Connect), sites are presented with potential matches from the patient's EHR, from which the site reviews and selects the right data point for the eCRF. Once a data point is selected, all other data points within the eCRF can be auto-filled from the EHR based on the date range selected.
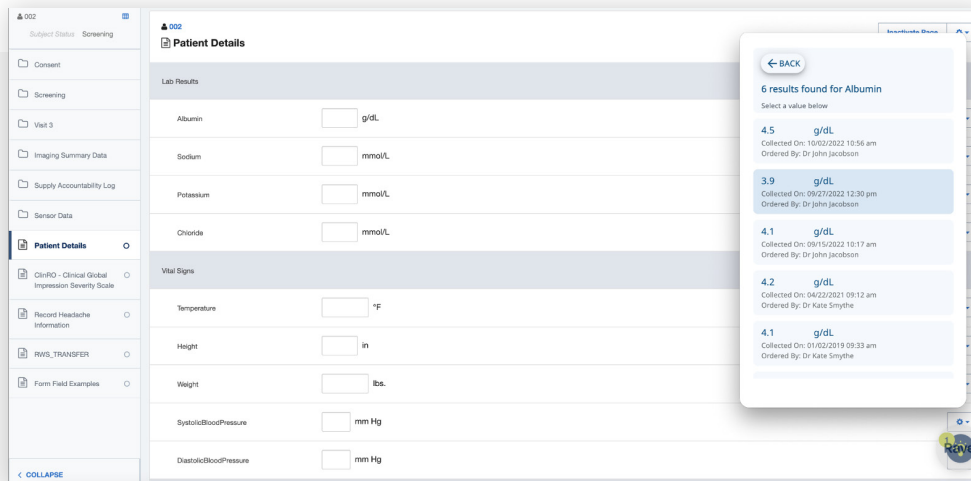
Medidata Health Record Connect is a uniquely scalable solution for accessing patient healthcare information with out-of-the-box connectivity to tens of thousands of healthcare organizations and physician practices through HIEs, dramatically reducing the need to create individual connections to EHR systems at each institution. Health Record Connect leverages multiple data format and standards to access healthcare data, including HL7, CCDA, and FHIR.

Figure 3: A Multi-Pronged, Scalable Approach to Accessing Healthcare Data
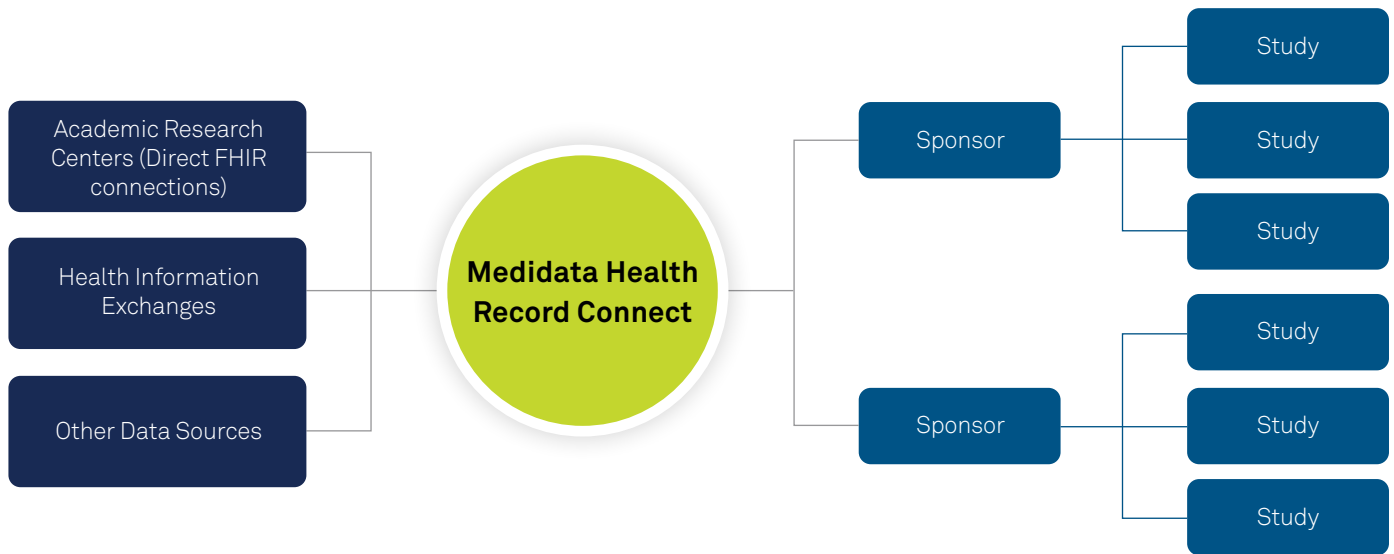
Summary

EHRs provide a comprehensive collection of information that can provide several bona fide benefits to clinical trials, such as faster trial completion, enhanced recruitment, and earlier identification of safety events. Furthermore, they facilitate the connectivity and interoperability required for the exchange of electronic health information with other systems.

While traditional approaches to EHR-to-EDC have existed for years, changes in the industry landscape, the adoption of new technologies, and openness to interoperability are all opportunities to create solutions that scale.

ⓘ **To learn more about Rave Companion visit:**
https://www.medidata.com/en/clinical-trial-products/clinical-data-management/rave-companion/

References

Collen MF. Clinical research databases--a historical review. J Med Syst. 1990;14(6):323-344.

European Commission. Exchange of electronic health records across the EU. 2022. Available at: **https://digital-strategy.ec.europa.eu/en/policies/electronic-health-records**. Accessed May 2023.

FDA. Guidance Document: Use of electronic health record data in clinical investigations, Guidance for industry. 2018. Available at: **https://www.fda.gov/regulatory-information/search-fda-guidance-documents/use-electronic-health-record-data-clinical-investigations-guidance-industry**. Accessed May 2023.

Federal Register. 21st Century Cures Act: Interoperability, Information Blocking, and the ONC Health IT Certification Program. Vol. 85, No. 85. Friday, May 1, 2020.

Fouque, J. 2022. The EHR Landscape is a Mess: One Way to Gain Control. Available at: **https://fdiinc.com/ehr-landscape-gain-control/**. Accessed May 2023.

Garza MY, Rutherford M, Myneni S, Fenton S, Walden A, Topaloglu U, Eisenstein E, Kumar KR, Zimmerman KO, Rocca M, Gordon GS, Hume S, Wang Z, Zozus M. Evaluating the Coverage of the HL7® FHIR® Standard to Support eSource Data Exchange Implementations for use in Multi-Site Clinical Research Studies. AMIA Annu Symp Proc. 2021 Jan 25;2020:472-481.

Johnson, C., Pylypchuk, Y., Patel, V. 2018. Methods Used to Enable Interoperability among U.S. Non-Federal Acute Care Hospitals in 2017. ONC Data Brief. No. 43.

McCord KA, Hemkens LG. Using electronic health records for clinical trials: Where do we stand and where can we go? CMAJ. 2019 Feb 4;191(5):E128-E133.

Society for Clinical Research Sites. 2022. A Survey of 120+ clinical trial site representatives in Oct/Nov 2022.

Nordo AH, Levaux HP, Becnel LB, Galvez J, Rao P, Stem K, Prakash E, Kush RD. Use of EHRs data for clinical research: Historical progress and current applications. Learn Health Syst. 2019 Jan 16;3(1):e10076.

ONC. Hospitals' Use of Electronic Health Records Data, 2015-2017. ONC Data Brief. No. 46. April 2019.

ONC. CURES ACT FINAL RULE: Standards-based Application Programming Interface (API) Certification Criterion. 2020. Available at: **https://www.healthit.gov/sites/default/files/page2/2020-03/APICertificationCriterion.pdf**. Accessed May 2023.

ONC. Health Information Exchange. 2023. Available at: **https://www.healthit.gov/topic/health-it-and-health-information-exchange-basics/health-information-exchange**. Accessed May 2023.

Rave Companion Factsheet. Capture EHR Data into Rave EDC in a Few Clicks. 2023. Available at: **https://www.medidata.com/wp-content/uploads/2023/02/Rave-Companion-Factsheet-Jan-23.pdf**. Accessed May 2023.

Stokman, P.G., Ensign, L., Langeneckhardt, D., Mörsch, M., Nuyens, K., Herrera, D., Hochgräber, G., Cassan, V., Beineke, P., Kwock, R. and Voortman, A. 2021. Risk-based Quality Management in CDM An inquiry into the value of generalized query-based data cleaning. Journal of the Society for Clinical Data Management, 1(1).

Sundgren, M., Ammour, N., Hydes, D., Kalra, D., Yeatman, R. 2021. Innovations in Data Capture Transforming Trial Delivery. Applied Clinical Trials. Volume 30, Issue 7/8. Available at: **https://www.appliedclinicaltrialsonline.com/view/innovations-in-data-capture-transforming-trial-delivery**. Accessed May 2023.

Tufts Center for the Study of Drug Development. 2021. January/February Tufts CSDD Impact Report: Rising Protocol Design Complexity is Driving Rapid Growth in Clinical Trial Data Volume. January/February Vol. 23 No. 1.