

Integrating Research Data Capture into the Electronic Health Record Workflow: Real-World Experience to Advance Innovation

by Marsha Laird-Maddox; Susan B. Mitchell, MSN, RN; and Mark Hoffman, PhD

Abstract

As the adoption of electronic health records (EHRs) increases, more opportunities are available for leveraging the system and the data to facilitate research. Historically, for patients enrolled in clinical research trials or studies, data have been documented in the medical record, and then study-related data are manually reentered into an electronic case report form in a research system. By utilizing data collected in the EHR to prepopulate electronic case report forms, manual transcription is reduced, data quality is improved, and the workflow for capturing research data is streamlined. Past efforts to integrate EHRs and research systems for the purposes of data capture have demonstrated that interoperability is possible. This article highlights how Cerner Corporation and Florida Hospital collaborated to extend an existing standard to implement a workflow called Integrated Data Capture.

Introduction and Background

The use of electronic health record (EHR) systems by healthcare organizations has accelerated since the passage of the Health Information Technology for Economic and Clinical Health Act in 2009, which included incentives for the meaningful use of health information technology.¹ With the increasing amount of patient information captured in EHRs, more opportunities are available to leverage the system and data to facilitate research. While comparable in their ability to store clinical data, EHRs and research systems differ in their workflow and regulatory compliance needs. The primary use of an EHR system is to facilitate clinical care while improving the quality of healthcare delivery and enhancing the safety of patients.² Emphasis is placed on workflows that support the provision of care. Research systems, often known as electronic data capture systems, focus primarily on electronic documentation, collection, and management of data captured by clinical research sites participating in a given study. Priority is placed on workflows that enable verification of data integrity and validity. Although similar data, such as medications, test results, and problems, are collected in both types of systems, fundamental differences in the use of the systems have driven the need for separate purpose-built systems.

Historically, for patients enrolled in clinical research trials, data have been documented in the medical record, and then study-related data are manually reentered into an electronic data capture system.³ Greatly reducing or even completely eliminating redundant data entry by utilizing EHR data to prepopulate a research database can increase research data collection efficiency, minimize transcription errors, and expedite database lock (finalization of data to be studied).

While not all study data are likely to be found in the EHR's clinical documentation, significant overlap is possible. The RE-USE (Retrieving EHR Useful data for Secondary Exploitation) project leveraged a semantic mapping process to match EHR data to elements of the electronic case report form for research. This work found that 13.4 percent of the data needed in the electronic case report forms could be directly mapped from data in the EHR.^{4,5} A study conducted by Siemens and the Frauenklinik of the Technical University of Munich found that between 48 percent and 69 percent of the electronic case report form data could be prepopulated using their integrated EHR–electronic data capture solution.⁶ Murphy et al demonstrated an EHR system that included custom-built screens for capturing research-related data, which were then extracted from the EHR database.⁷

A number of organizations have worked to address the challenge of EHR and research system integration by proposing data, process, and technology standards. Integrating the Healthcare Enterprise (IHE) and the Clinical Data Interchange Standards Consortium (CDISC) worked with multiple EHR and electronic data capture vendors and pharmaceutical companies to develop the Retrieve Form for Data Capture profile in 2007.⁸ Retrieve Form for Data Capture is a method for gathering data within a user's current application context to support the prepopulation of forms retrieved from an external source such as an electronic data capture system.⁹ CDISC and IHE have hosted interoperability demonstrations leveraging Retrieve Form for Data Capture since 2007.¹⁰ Cerner Corporation, Greenway Health, Allscripts, and other EHR vendors have participated and demonstrated that data captured in the EHR system can be electronically transmitted to data capture systems for research, prepopulating relevant data elements.

Florida Hospital (Orlando, Florida), one of the country's largest not-for-profit hospitals with a widespread installation of the Cerner Millennium EHR, agreed to collaborate with Cerner to implement the Retrieve Form for Data Capture workflow in a real-world environment. Cerner partnered with the Translational Research Institute for Metabolism and Diabetes (TRI) at Florida Hospital to create a streamlined system that integrates research data capture into a standard care workflow. A workflow based on Retrieve Form for Data Capture was implemented to electronically transmit relevant participant data captured in the Millennium EHR to the Cerner *Discovere* research data capture system. The TRI selected an investigator-initiated, noninterventional diabetes study for the first Integrated Data Capture implementation. Although the study is not subject to the same regulatory requirements as an interventional study, the TRI aimed to design the system to meet the requirements of regulated research. This article describes the implementation of research data capture integration at the TRI, features of the integration, and how the system was designed and implemented to enhance the research process while maintaining regulatory compliance.

The Systems

Millennium is the EHR system used by the Florida Hospital system and the TRI. In addition to the core EHR capabilities, Millennium supports several workflows related to clinical research: study management, enrollment tracking, trial screening, and recruitment. Clinicians can use Millennium to easily identify patients who are in a study and to view relevant, research-related information such as study documents and contact information for key study personnel. Millennium can promote protocol compliance through the use of predefined order sets, facilitate research billing by delineating standard-of-care versus research charges, and alert key study personnel to activity of study participants, such as a participant's being admitted to the hospital or being prescribed a contraindicated medication.

Discovere is the Cerner system designed for research data capture and is a separate, web-based platform that can be used independently of Millennium. *Discovere* supports traditional electronic case report form data capture, data management, participant surveys, patient-reported outcomes, and study reporting.

Integrated Data Capture is the process that enables the electronic transmission of relevant data from the Millennium EHR to *Discovere* (See Figure 1). Integrated Data Capture is an extension of the Retrieve Form for Data Capture workflow.

Initial Retrieve Form for Data Capture–based Implementation

Cerner’s interoperability demonstrations using Retrieve Form for Data Capture prior to the work with the TRI successfully demonstrated retrieving a form from a research data capture system that would be populated with data captured in the Millennium EHR. From the Millennium system, a Continuity of Care Document is generated from the EHR and contains the most recently populated values for the relevant data elements. A Continuity of Care Document is an XML-based HL7 standard used in the exchange of clinical data between healthcare providers.¹¹ A script of code transforms the Continuity of Care Document into a format that can be used by the research system. The electronic case report form is displayed within a new window, and the values from the EHR are prepopulated in the appropriate fields. At that point, the user can enter additional research-specific data, modify values from the EHR, and save the form. The user can complete this process with minimal interruption to the current EHR session. Data can flow from the EHR to an electronic case report form without manual reentry.

Improving Integrated Data Capture

In preparation for the use of Integrated Data Capture at the TRI, the team assessed the Retrieve Form for Data Capture–based workflow described above and identified areas of improvement that were critical to successful usage in a real-world setting. The assessment included identifying the type of data that would need to be captured from the EHR for the study, identifying the circumstances in which the data would need to be modified, and evaluating what metadata would need to be captured and presented for audit purposes. Collaboration between Florida Hospital and Cerner resulted in enhancements that extended the Retrieve Form for Data Capture–based workflow (See Figure 2). These areas for improvement and subsequent enhancements are described below.

Additional Data Categories

Enhancement was needed to correct the fact that some data elements required by the study were available in the EHR but not represented by the Continuity of Care Document. In past demonstration projects, the content that the Millennium EHR could push to a research system was limited to demographics, vital signs, adverse events (problems, diagnoses, and allergies), and medications. The TRI case report forms captured additional structured information that was in the EHR but not in this list. Examples of these elements are family and surgical history, laboratory results, task completion information, and data captured within custom forms in the EHR.

To maximize the benefit of Integrated Data Capture, we extended the categories of data available for prepopulating the electronic case report form. This task was accomplished using embedded database queries and aliases to match data from one system to the other. The aliasing was achieved by mapping the unique identifier of the field that displays the discrete EHR data to clinicians to the unique identifier of the data capture field in the electronic case report form.

Source Preservation

Values in the research system that were captured through Integrated Data Capture should not be overwritten. In the first iteration of Integrated Data Capture, values displayed in the electronic case report form could be overwritten via manual entry. This ability could reintroduce the risk of data transcription error and put the electronic case report form out of sync with the source document (the EHR). The Discovere system was enhanced so that values captured via Integrated Data Capture can only be edited via a menu option that contains other values from the EHR (discussed in the next section).

Flexibility

A range of data must be available to be saved for research purposes. In the first iteration of Integrated Data Capture, the most recently documented value for a field that mapped to an electronic case report form field would be used. The most recently documented value for a particular item in the patient chart is not necessarily the appropriate value to be saved in the electronic case report form. If multiple clinical values, for example, multiple blood glucose values, are captured, the value captured specifically for research purposes needs to be saved in the electronic case report form. Additionally, at times a set of data

is collected twice in the electronic case report form; for example, a set of vital signs may be repeated. To ensure capture of the most appropriate value, the system was enhanced to receive a range of data from the EHR. Once the electronic case report form is displayed, the user can filter the data available to a specific date range. Although the most recent values for that range are populated in the mapped fields, the user can select an earlier result to be saved to the research database. This option is crucial because manual entry is not allowed.

The name of the user who documented the result in the EHR as well as the date and time of documentation is displayed next to each result value, to aid in its selection. This feature can be especially useful in an inpatient setting where, for example, vital signs are taken frequently and by multiple clinicians. In these situations, the contextual metadata for the results help the researcher locate the result most relevant to the research study, for example, the result captured by a research nurse.

Indication of Data Collection Mode

Discovere can be used in two modes—manual data entry mode without active connection to the EHR, and Integrated Data Capture mode in which Discovere is invoked from an EHR session. Users requested a visual indicator in the application so they would know which mode the system is in.

Audit Information with EHR Data

In traditional data capture, an item history displays the date, time, and creator of data in an electronic case report form field, as well as a record of subsequent changes with an accompanying reason for change. The code scripts that send the data to the research system and the item history in Discovere were enhanced to support the display of additional audit information for data in the electronic case report form that originated in the EHR. To show a clear sequence of events, the item history now reflects that the source of the value in the electronic case report form field is the EHR; the date, time, and user who documented the data in the EHR; and all of the previously mentioned audit information specific to the electronic case report form.

Support of Repeating Data

Certain pieces of participant information need to be displayed in a list or table format because they are repeating instances of the same case report form fields. Examples include medications, adverse events, and medical history. In the initial implementation of Integrated Data Capture, related items such as these could not be grouped into a list. Discovere and the script that drives EHR data gathering were updated to support grouping of multiple repeating pieces of information. For medications, the name, dosage, route, and start date, for example, can all be displayed in one row per medication. The Integrated Data Capture implementation also supports configurable rules that define when the system will consider a set of data as unique and create a new row.

Grouping of Related Data Elements

In the first iteration of Integrated Data Capture, a single value was sent to the electronic case report form independent of any other single value. For example, systolic and diastolic blood pressure values would be gathered and populated independently. The Discovere system was enhanced to support grouping of related values. Values captured together in the EHR can be populated together in the electronic case report form; for example, blood pressure data are selected and displayed as a group. Other examples of values that could require grouping include laboratory results, such as the components of a laboratory panel.

Benefits of Integrated Data Capture

The collaborative development and implementation decisions made by Florida Hospital and Cerner have made the following benefits possible.

Improved User Experience

Integrated Data Capture streamlines the research data capture process by allowing a user to bypass the typical steps required in completing electronic case report forms (See Figure 3).

Improved Data Quality

Integrated Data Capture reduces the risk of a transcription error because data are entered directly into the electronic case report form, as opposed to manually.

Support of Auditing and Monitoring

The electronic case report form data contain details about the capture of the information in the EHR. These values are kept synchronized with the EHR, and the details of changes are recorded, providing a complete, clear sequence of events.

Earlier Database Lock

While working in a study participant's EHR record, a care provider can submit the data to the electronic case report form in real time. This capability prevents a time lag between EHR capture and transcription of data, allowing for participants' data to be captured and finalized more quickly. With improved data quality, data cleansing takes less time.

Conclusion

This article describes an initial Retrieve Form for Data Capture–inspired implementation of Integrated Data Capture, followed by enhancements based on lessons learned. We deployed capabilities such as increasing the data available for prepopulation, securing the prepopulated fields so that they could only be modified with other EHR-supplied values, providing the user with notification that EHR data are available, adding audit information, supporting repeated elements such as medications, and enabling data to be populated as a group.

The success of Integrated Data Capture will be evaluated as the TRI and other research institutes at Florida Hospital leverage the technology and process during future studies. Implementations such as the one at the TRI can serve to inform and provide lessons learned for future iterations of standards and regulatory guidance. For workflows such as Integrated Data Capture to gain broad acceptance, continued collaboration is needed among researchers, industry, and regulatory organizations.

Marsha Laird-Maddox is a senior engagement leader in Population Health Consulting at Cerner Corporation in Kansas City, MO.

Susan B. Mitchell, MSN, RN, is a senior manager of Research Information Systems at Florida Hospital in Orlando, FL.

Mark Hoffman, PhD, is the director of the Center for Health Insights at the University of Missouri–Kansas City in Kansas City, MO.

Acknowledgements

We would like to thank Jane Griffin, RPh, Lisa Kaspin, PhD, and Ginger Nedblake for their contributions to this manuscript.

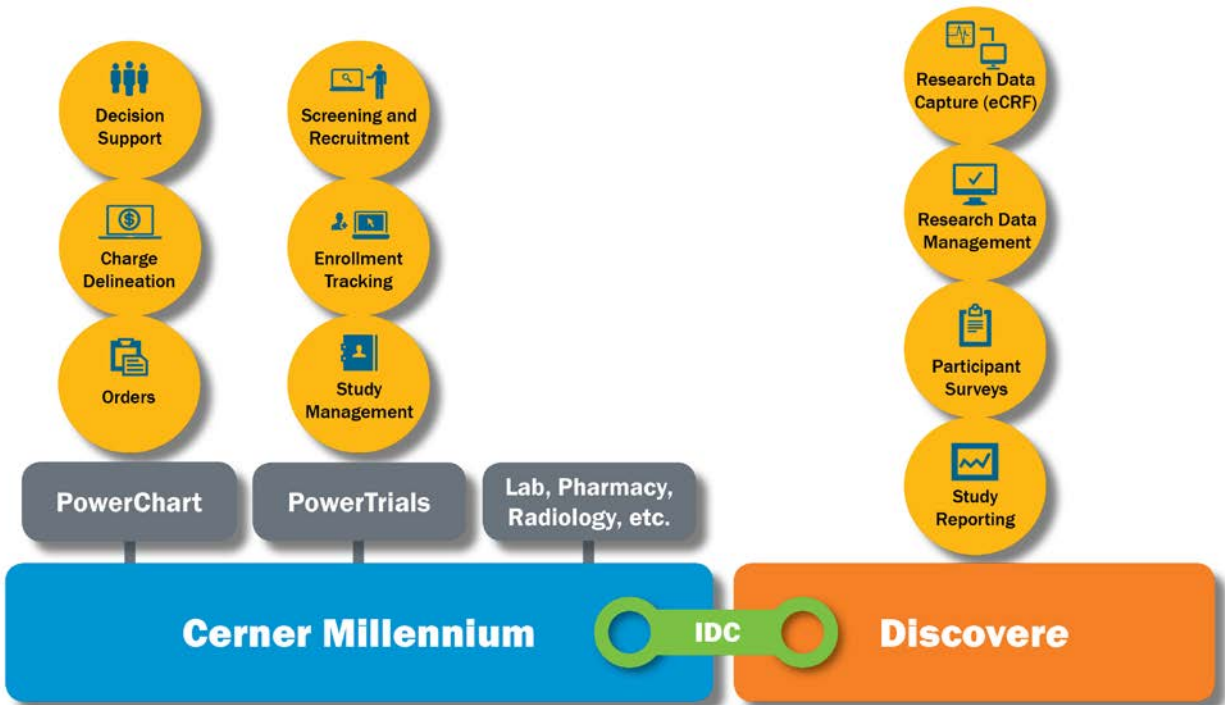
Notes

1. Jha, Ashish K., Matthew F. Burke, Catherine M. DesRoches, Maulik S. Joshi, Peter D. Kralovec, Eric G. Campbell, and Melinda B. Buntin. "Progress toward Meaningful Use: Hospitals' Adoption of Electronic Health Records." *Journal of Managed Care* 17, no. 12 (spec. no.) (2011): SP117–SP124.
2. Bartlett, Michael, Suzanne Bishop, Catherine Celingant, Gary Drucker, Tricia Gregory, Linda King, Susan Klimek, John Mestler, Brad Michel, Richard Perkins, Sharon Powell, Christian Reich, and Selina Sibbald. *The Future Vision of Electronic Health Records as eSource for Clinical Research*. eClinical Forum/PhRMA EDC/eSource Taskforce. September 14, 2006. Available at [http://www.w3.org/wiki/images/2/21/HCLSS\\$ClinicalObservationsInteroperability\\$FutureEHR.pdf](http://www.w3.org/wiki/images/2/21/HCLSS$ClinicalObservationsInteroperability$FutureEHR.pdf).
3. Ibid.
4. El Fadly, AbdenNaji, Bastien Rance, Noël Lucas, Charles Mead, Gilles Chatellier, Pierre-Yves Lastic, Marie-Christine Jaulent, and Christel Daniel. "Integrating Clinical Research with the Healthcare Enterprise: From the RE-USE project to the EHR4CR Platform." *Journal of Biomedical Informatics* 44, suppl. 1 (2011): S94–S102.
5. El Fadly, AbdenNaji, Noël Lucas, Bastien Rance, Philippe Verplancke, Pierre-Yves Lastic, and Christel Daniel. "The REUSE Project: EHR as Single Datasource for Biomedical Research." *Studies in Health Technology and Informatics* 160, pt. 2 (2010): 1324–28.
6. Zahlmann, Gudrun, Nicole Harzendorf, Ulrike Schwarz-Boeger, Stefan Paepke, Markus Schmidt, Nadia Harbeck, and Marion Kiechle. "EHR and EDC Integration in Reality." *Applied Clinical Trials Online*. November 16, 2009. Available at <http://www.appliedclinicaltrials.com/appliedclinicaltrials/article/articleDetail.jsp?id=641682>.
7. Murphy, Elizabeth C., Frederick L. Ferris III, and William R. O'Donnell. "An Electronic Medical Records System for Clinical Research and the EMR-EDC Interface." *Investigative Ophthalmology & Visual Science* 48, no. 10 (2007): 4383–89.
8. Clinical Data Interchange Standards Consortium. "Healthcare Link Initiative." 2012. Available at <http://www.cdisc.org/healthcare-link>.
9. IHE International, Inc. "IHE IT Infrastructure Technical Framework Supplement: Retrieve Form for Data Capture (RFD) Trial Implementation." August 19, 2011. Available at http://www.ihe.net/Technical_Framework/upload/IHE_ITI_Suppl_RFD_Rev2-2_TI_2011-08-19.pdf.
10. Clinical Data Interchange Standards Consortium. "Healthcare Link Initiative."

11. Health Level Seven International. “HL7/ASTM Implementation Guide for CDA R2—Continuity of Care Document (CCD) Release 1.” Available at http://www.hl7.org/implement/standards/product_brief.cfm?product_id=6.

Figure 1

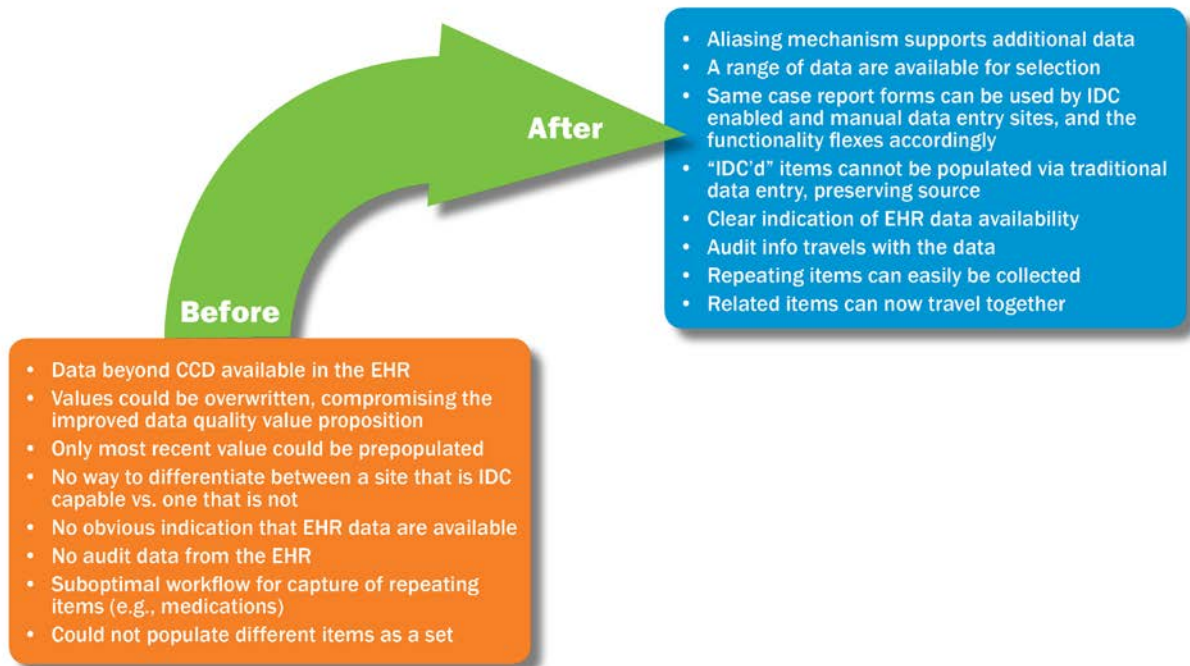
Connection between the Electronic Health Record and the Research System



eCRF, electronic case report form; IDC, Integrated Data Capture.

Figure 2

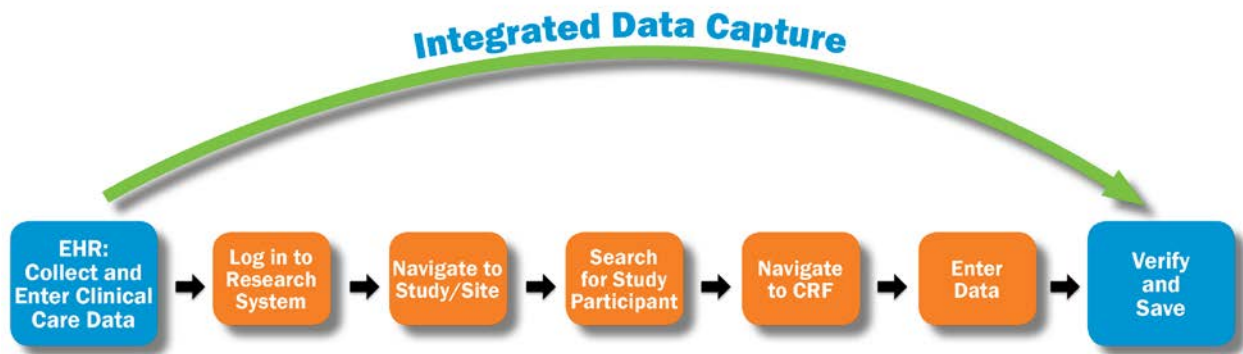
Enhancements to Extend Initial Retrieve Form for Data Capture–based Implementation



CCD, Continuity of Care Document; EHR, electronic health record; IDC, Integrated Data Capture.

Figure 3

Improved Workflow



CRF, case report form; EHR, electronic health record.