International Journal of Medical Informatics 103 (2017) 89–94

Contents lists available at ScienceDirect

International Journal of Medical Informatics

journal homepage: www.elsevier.com/locate/ijmedinf

A comparative effectiveness study of eSource used for data capture for a clinical research registry

Amy Harris Nordo[a,*], Eric L. Eisenstein[b], Jeffrey Hawley[a], Sai Vadakkeveedu[a], Melissa Pressley[a], Jennifer Pennock[a], Iain Sanderson[a]

[a] Duke University School of Medicine- Office of Research Informatics, 2424 Erwin Road, Durham, NC 27701, USA
[b] Duke University Clinical Research Institute, 2400 Pratt Street, Terrace Level 0311, Durham, NC 27705, USA

ARTICLE INFO

Keywords:
Data accuracy
Data collection
Data retrieval
Information extraction
Registries
Time-motion analysis

ABSTRACT

Objective: This pilot study compared eSource-enabled versus traditional manual data transcription (non-eSource methods) for the collection of clinical registry information. The primary study objective was to compare the time spent completing registry forms using eSource versus non-eSource methods The secondary objectives were to compare data quality associated with these two data capture methods and the flexibility of the workflows. This study directly addressed fundamental questions relating to eSource adoption: what time-savings can be realized, and to what extent does eSource improve data quality.

Materials and methods: The study used time and motion methods to compare eSource versus non-eSource data capture workflows for a single center OB/GYN registry. Direct observation by industrial engineers using specialized computer software captured keystrokes, mouse clicks and video recordings of the study team in their normal work environment completing real-time data collection.

Results: The overall average data capture time was reduced with eSource versus non-eSource methods (difference, 151 s per case; eSource, 1603 s; non-eSource, 1754 s; $p = 0.051$). The average data capture time for the demographic data was reduced (difference, 79 s per case; eSource, 133 s; non-eSource, 213 s; $p < 0.001$). This represents a 37% time reduction (95% confidence interval 27% to 47%). eSourced data field transcription errors were also reduced (eSource, 0%; non-eSource, 9%).

Conclusion: The use of eSource versus traditional data transcription was associated with a significant reduction in data entry time and data quality errors. Further studies in other settings are needed to validate these results.

1. Introduction

For more than 60 years, the cost of conducting biomedical research has increased exponentially while net productivity has declined [1,2]. The greatest cost increases have occurred in late phase clinical trials where > 65% of total costs are site-related (for site management and site trial work) and the complexity of protocol-mandated activities has escalated [3,4]. Several initiatives are investigating ways to reduce clinical trial costs without compromising their scientific validity. These efforts have focused on reducing costs by monitoring only high risk tasks and studies (risk-based monitoring) and the secondary use of existing registry and billing data [5–8]. However, none of these initiatives address an area of high cost and great inaccuracy in clinical research studies, namely the collection and transcription of study data from the patient's health record into the clinical study's electronic case report form (eCRF).

2. Background and significance

There is growing interest in utilizing EHR data for clinical research. According to the American Medical Informatics Association (AMIA), "*Secondary use of health data can enhance healthcare experiences for individuals, expand knowledge about disease and appropriate treatments, strengthen understanding about the effectiveness and efficiency of our healthcare systems, support public health and security goals, and aid businesses in meeting the needs of their customers*" [9]. Typically, EHR data is manually abstracted and then entered into electronic clinical research forms (eCRF). The next step in making EHR data available for clinical research is to directly link the EHR and eCRF systems. The Federal Drug Administration (FDA) coined the term "eSource" to represent the secondary use of EHR data for completing eCRFs using interoperability standards [10,11].

The Retrieve Form Data Capture (RFD) standard provides eSource

* Corresponding author at: Duke Office of Research Informatics, 303 Village Loop Drive, Apex, NC 27502, USA.
E-mail addresses: amy.nordo@duke.edu, amy.nordo@gmail.com (A.H. Nordo).
URL: http://mailto:amy.nordo@duke.edu (A.H. Nordo).

http://dx.doi.org/10.1016/j.ijmedinf.2017.04.015
Received 11 January 2017; Received in revised form 10 April 2017; Accepted 23 April 2017

capability. RFD is an Integrating the Healthcare Enterprise (IHE) standard that allows for secure interoperability between systems by providing a window into the EHR so that the eCRF form can be auto-populated using previously mapped EHR data elements [12]. RFD also enables other study-specific data elements to be entered directly into the eCRF at the point of care and from within the EHR, posting the eCRF data into the study database and not the EHR itself.

This study tests the hypothesis that eSource data management reduces time and transcription errors, with the implication of savings in study costs, particularly if widely adopted in the most expensive late phase clinical trials.

3. Objective

The primary study objective is to compare the time spent completing the eCRF using traditional (non-eSource) and eSource-enabled workflows. The secondary objective is to compare data quality associated with these data capture methods. This study directly addresses fundamental questions relating to eSource adoption: what time-savings can be realized, and to what extent does eSource improve data quality.

4. Materials and methods

4.1. Product design

Duke University Office of Research Informatics developed middleware, called RADaptor, that uses the RFD standard to electronically call a study eCRF from a REDCap [13] database into Duke's Epic (Epic Inc, Verona, WI) EHR. This use of the RFD standard is part of Epic's model research functionality, and configuration for a study to use RADaptor can be accomplished within Epic's normal research configuration in the research (RSH) record. Study team members using RADaptor open the EHR, and "call" the eCRF. Once security is authorized, the eCRF will appear within the EHR window. Data points that have been previously mapped will auto-populate in the eCRF based upon EHR data availability. For the present study, data points mapped with RADaptor included those contained within the EHR's continuity of care document (CCD) [14]. Data elements that are not mapped appear as unanswered (study-specific) eCRF sections. Data can be edited or study-specific data can be entered directly into the eCRF form. Edited or study-specific data in the eCRF will not be stored in the EHR, being posted directly to the study's REDCap database.

4.2. Study design

This is a single site, observational comparative effectiveness study that examines the impact of Duke's RADaptor on eCRF completion workflows. The study protocol employs time-motion methods to compare eSource and non-eSource workflows for a single center OB/GYN study.

Study data collected included: process time, motion, mouse clicks, and keystrokes. Typically mouse click and keystroke research data are collected manually in a laboratory environment using pseudo patients. This is a tedious, error prone process. To address this problem, we augmented the direct observations of the industrial engineers with software that documented keystrokes, mouse clicks and took video recordings in the actual work setting. The Duke University School of Medicine Institutional Review Board (IRB) approved the study protocol on December 1, 2015 (PRO00068189).

4.3. Information security

At the recommendation of the Duke Information Security Office (ISO), two desktop computers and two laptop computers were acquired and configured with keylogging and study software. These desktop computers were set up and enabled in the study team members' normal workspace at the start of each observation period. At the end of each observation period, study desktops were removed, the study teams' passwords reset and the original work computers replaced.

4.4. Study participants

The study included a convenience sample of Duke University clinical research staff working on a Prematurity Prevention registry. Subjects had access to the institution's EHR, Epic, and were eligible to access the primary study's eCRF, REDCap. Study data were collected and managed using Research Electronic Data Capture (REDCap) electronic data capture tools hosted at Duke University. REDCap is a secure, web-based application designed to support data capture for research studies, providing 1) an intuitive interface for validated data entry; 2) audit trails for tracking data manipulation and export procedures; 3) automated export procedures for seamless data downloads to common statistical packages; and 4) procedures for importing data from external sources [15]. This software tool was developed and supported by the US National Institute of Health's National Center for Advancing Translational Sciences, is intended to support smaller investigator-initiated studies, and is considered the premier software tool in this niche.

The clinical study team is proficient in the use of REDCap and has completed multiple studies using this system and requested the Prematurity Prevention Registry to be built in REDCap. The eCRF was built by the Duke Office of Clinical Research (DOCR) and is considered to be a typical registry eCRFs for principal investigator initiated studies in terms of complexity.

The eCRF was created by the Duke Office of Clinical Research (DOCR) and its complexity is typical of investigator initiated registry eCRFs at our institution. This eCRF contains 401 data elements relating to a mother's pregnancy and birth of the infant; however, most of these data elements are not entered for each case. The present study is limited to the eCRF's demographic section. The eSource methodology allowed for 7 of 14 demographic data elements to be auto-populated into the demographics form. Abstracting in the present study was limited to the copying of information from patient electronic medical records to the eCRF demographics section.

4.5. Observation

Three study participants were observed completing two different workflows in their Duke work environment.

4.5.1. Non-eSource observation

Fig. 1 shows the non-eSource workflow performed by a Clinical Research Coordinator (CRC) and a Data Entry Technician completing registry data collection. A CRC initializes the workflow by opening a patient's record in the Epic EHR, then transcribes the data from the EHR onto the paper case report form (CRF). The CRC gives the completed paper CRF to a data entry technician to transfer the information from the paper CRF to the eCRF.

4.6. ESource observation

The eSource workflow requires one CRC to complete data collection (Fig. 2). A CRC initializes the file by opening the patient's medical records and the eCRF, all within Epic's hyperspace (initialization phase). The CRC then verifies the eSourced variables, (demographic information), that were pre-populated with data extracted from the EHR utilizing the RFD standard, and simply clicks the "Save" button (eSourced phase). The CRC then finds the necessary information in the EHR to manually complete the remainder of the fields in the supplemental form (supplemental phase) which appears within the EHR; appearing to the end user as if it is a part of the EHR.