International Journal of Medical Informatics 119 (2018) 103–108

Contents lists available at ScienceDirect

International Journal of Medical Informatics

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

Connecting healthcare and clinical research: Workflow optimizations through seamless integration of EHR, pseudonymization services and EDC systems

Philipp Bruland*, Justin Doods, Tobias Brix, Martin Dugas, Michael Storck

Institute of Medical Informatics, University of Münster, Münster, Germany

ARTICLE INFO

Keywords:
Data management
Electronic data capture
Health information systems
Pseudonymization
Workflow optimization

ABSTRACT

Objective: In the last years, several projects promote the secondary use of routine healthcare data based on electronic health record (EHR) data. In multicenter studies, dedicated pseudonymization services are applied for unified pseudonym handling. Healthcare, clinical research and pseudonymization systems are generally disconnected. Hence, the aim of this research work is to integrate these applications and to evaluate the workflow of clinical research.

Methods: We analyzed and identified technical solutions for legislation compliant automatic pseudonym generation and for the integration into EHR as well as electronic data capture (EDC) systems. The Mainzelliste was used as pseudonymization service, which is available as open source solution and compliant with the data privacy concept in Germany. Subject of the integration was the local EHR and an in-house developed EDC system. A time and motion study was conducted to evaluate the effects on the workflow.

Results: Integration of EHR, pseudonymization service and EDC systems is technically feasible and leads to a less fragmented usage of all applications. Generated pseudonyms are obtained from the service hosted at a trusted third party and can now be used in the EDC as well as in the EHR system for direct access and re-identification. The evaluation of 90 registration iterations shows that the time for documentation has been significantly reduced in average by 39.6 s (56.3%) from 71 ± 8 s to 31 ± 5 s per registered study patient.

Conclusions: By incorporating EHR, EDC and pseudonymization systems, it is now feasible to support multicenter studies and registers out of an integrated system landscape within a hospital. Optimizing the workflow of patient registration for clinical research allows reduction of double data entry and transcription errors as well as a seamless transition from clinical routine to research data collection.

1. Introduction

Adoption rates of electronic documentation systems, in particular electronic health records (EHR) and electronic data capture (EDC) systems, are steadily increasing in the recent years [1]. Several advantages such as availability of information, better communication, higher readability, support in different tasks, decrease of documentation errors, etc. are reported to accompany with the use of such systems [2,3]. Studies have shown however, that manually transferred information from one system to another contains a large source of errors [4]. Electronic documentation, reusing of those and sharing of data between stakeholders in healthcare has even been deemed so important by politics that legislations such as the Health Information Technology for Economic and Clinical Health Act (HITECH) in the United States in 2009 [5] or the eHealth-law in Germany [6] were decided.

Introduction of new applications harbors the potential risk of workflow interruptions frequent task switching causing stress and dissatisfaction [7,8]. In addition to that, clinicians and general practitioners spend 25–50% of their daily work for EHR-related tasks such as documentation or administrative issues and the time for research documentation comes on top of the daily routine work [9–11].

In this regard, the re-use of routinely collected medical data, e.g. for research or quality assurance purposes, offers several incentives, for instance the reduction of redundant documentation or increase of data quality [12]. Clinical trials usually consist of several stages in which routinely collected data is eligible to support e.g. trial feasibility, patient recruitment or the execution of trial documentation. Recent research has shown that data elements required for clinical trials overlap

* Corresponding author at: Albert-Schweitzer-Campus 1 – A11, Institute of Medical Informatics, University of Münster, 48149, Münster, Germany.
E-mail addresses: philipp.bruland@uni-muenster.de (P. Bruland), justin.doods@uni-muenster.de (J. Doods), tobias.brix@uni-muenster.de (T. Brix), dugas@uni-muenster.de (M. Dugas), Michael.storck@uni-muenster.de (M. Storck).

https://doi.org/10.1016/j.ijmedinf.2018.09.007
Received 21 March 2018; Received in revised form 2 July 2018; Accepted 6 September 2018

with available ones in the EHR [13–17]. Therefore, different projects have developed technical infrastructures that address single and multicenter approaches for a broad legislation compliant, syntactical and semantical interoperable solution to accompany the heterogeneous landscape of EHR systems and organizational burdens [18–20]. In Münster, we developed the x4T-architecture (exchange for Trials) that allows the automatic pre-population of study eCRFs (electronic Case Report Forms designed to capture all of the protocol required information to be reported to the sponsor on each study subject) with validated routine data of the EHR [21]. Technical limitations of common EHR systems resulted in a middleware component called "Clinical Interface" that handles the secure pre-population data transfer with EDC systems.

In clinical research, de-identification of subjects i.e. study patients plays an essential role preserving and respecting patient's privacy. Pseudonymization services provide secure and legislation compliant solutions to ensure (1) that subjects can be identified only by the person in charge of identifying data and (2) that patients who appear in multiple institutions obtain the same pseudonym and are not handled as two different persons. Aamot et al. analyzed and compared different methods and applications for pseudonymization services [22]. In all pseudonymization methodologies a trusted third party (TTP) is placed to trustworthy handle pseudonyms. Examples for pseudonymization services are the generic Pseudonym Administration Service (gPAS) [23] and the "Mainzelliste" [24]. Both services offer web services for managing and obtaining pseudonyms.

A direct integration of pseudonymization services into EDC systems can be found in a few research data capture solutions. In addition to that, Schrimpf et al. reported about the linkage of a randomization service into the open-source EDC system OpenClinica [25]. In terms of patient safety and regulatory requirements, it is essential to be aware of whether a patient participates in a clinical trial to report current conditions as serious adverse events. Nevertheless, to our knowledge the seamless and overarching integration of a pseudonymization service into the direct context of EHR and EDC systems has not been performed so far.

Since EHR, EDC and pseudonymization applications are generally disconnected the aim of this research is to design, implement and evaluate an IT architecture that allows the direct pseudonymization of patient within the EHR by integrating the systems and workflows of EHR, pseudonymization and EDC system.

2. Material and methods

System integrations require a careful analysis of the clinical workflow in which they are introduced. Agfa ORBIS is used as main EHR system at the University Hospital Münster (UKM). Based on experiences and interviews with study nurses, we analyzed the workflow of patient pseudonymization, adding of subjects within the EDC system and the manual data transfer process. Technical capabilities of the EHR ORBIS, the pseudonymization service "Mainzelliste" and the x4T-EDC system [26] were examined to propose an integrated IT architecture.

Since we already gained experiences with the Mainzelliste, we chose this solution for our implementation. The Mainzelliste is a web-based pseudonymization service offering a user interface and a also RESTful services to allow secure creation of pseudonyms and the interaction with other applications. It is in line with the German data protection legislation for research projects. Per default, patient administrative data such as name*, surname*, date of birth*, birth name, zip code and city (* = mandatory) are used as identifying parameters to generate a pseudonym. Additional arbitrary parameters can be configured for the de-identification procedure in the Mainzelliste. An integrated customizable algorithm discovers unsure cases of similar entries to prevent spelling errors.

Currently, x4T's "Clinical Interface" tackles the communication between the EHR system and EDC solutions based on web services and international standards such as CDISCs Operational Data Model (ODM) [27]. Depending on available EHR interfaces, customary communication servers can also be integrated. The Clinical Interface requests routine medical data from the EHR system, converts it into ODM and transfers it to the EDC system.

To propose a generalizable and reusable infrastructure we reviewed available interfaces of all IT components and analyzed possibilities to extract data from and store information in the EHR. × 4T's "Clinical Interface" was enhanced to handle the secure connection with external web service-based applications, namely the interface towards the Mainzelliste and a single sign-on mechanism towards the EDC system. Furthermore, the interface was also capable of receiving pre-population data from a primary source to pre-fill study documentation forms with clinical values including provenance data. We developed an EHR form that allows two registration options straight from the EHR system: 1) enter a pseudonym manually or 2) generate a pseudonym using a pseudonymization service.

Transferring and handling of medical data always requires respecting data privacy and protection regulations. Therefore, a formal description (procedure directory) of involved IT systems and processes was jointly developed with and approved by the local data protection officer.

2.1. Workflow evaluation

Context of the evaluation was a multi-center cross-sectoral registry for patients with traumatic brain injury (TBI), which was funded by the Ministry of Health, Equalities, Care and Ageing. Partners from acute care, early rehabilitation, long-term rehabilitation and coordinators of follow-up care participated in this project.

The evaluation was performed before and after implementation of the pseudonymization service into the EHR system and interconnecting all systems. Due to the circumstances that physicians usually perform the task of patient registration and documentation infrequently spread over the workday, we chose a laboratory evaluation setting. Three independent observers were instructed in the workflow and time measurement. Three study nurses, who were used to the old workflow since they worked with the systems, were observed. All study nurses were trained after introduction of the new system infrastructure. Evaluation of the new workflow took place four months later. For both observations we obtained 30 eligible patient names from the study's principal investigator to be included in the TBI register. These names and date of births were included in an Excel-spreadsheet for the evaluation.

Outcome measure of the evaluation was the time between opening a patient's electronic medical record within the EHR as a starting point for transferring research data and the successfully registered patient within the EDC system. In-between the x4T-EDC was accessed, the Mainzelliste used for pseudonym generation and the spreadsheet managed to insert the corresponding pseudonym. We measured the execution time before and after the implementation on six days including a block of five patients each day. Each participant paused for half an hour time between before and after the observation. The task order was switched every day to prevent bias in terms of being familiar with already known patient names from the prior round. Each study nurse got an own instance of the Excel-spreadsheet to compare the manually generated pseudonyms afterwards in the previous workflow. Automatically generated pseudonyms in the new workflow were compared with the human created ones of the old workflow to examine whether the integrated infrastructure operates correct. In order to discover typing errors, generated pseudonyms were compared among each study nurse.

IBM SPSS version 25 was used for descriptive statistic and independent *t*-test calculations.