Electronic Data Capture and DICOM Data Management in Multi-Center Clinical Trials

Daniel Haak*, Charles-E. Page, Thomas M. Deserno
Department of Medical Informatics, RWTH Aachen University, Pauwelsstr. 30, 52057 Aachen

ABSTRACT

Providing eligibility, efficacy and security evaluation by quantitative and qualitative disease findings, medical imaging has become increasingly important in clinical trials. Here, subject's data is today captured in electronic case reports forms (eCRFs), which are offered by electronic data capture (EDC) systems. However, integration of subject's medical image data into eCRFs is insufficiently supported. Neither integration of subject's digital imaging and communications in medicine (DICOM) data, nor communication with picture archiving and communication systems (PACS), is possible. This aggravates the workflow of the study personnel, in special regarding studies with distributed data capture in multiple sites. Hence, in this work, a system architecture is presented, which connects an EDC system, a PACS and a DICOM viewer via the web access to DICOM objects (WADO) protocol. The architecture is implemented using the open source tools OpenClinica, DCM4CHEE and Weasis. The eCRF forms the primary endpoint for the study personnel, where subject's image data is stored and retrieved. Background communication with the PACS is completely hidden for the users. Data privacy and consistency is ensured by automatic de-identification and re-labelling of DICOM data with context information (e.g. study and subject identifiers), respectively. The system is exemplarily demonstrated in a clinical trial, where computer tomography (CT) data is de-centrally captured from the subjects and centrally read by a chief radiologists to decide on inclusion of the subjects in the trial. Errors, latency and costs in the EDC workflow are reduced, while, a research database is implicitly built up in the background.

Keywords: Workflow, clinical trials, integration, DICOM, PACS, open source

1. INTRODUCTION

Providing eligibility, efficacy, and security evaluation by quantitative and qualitative disease findings, medical imaging is looming large in clinical trials today [1]. Here, the traditional paper-based case report forms (CRFs) are today substituted by electronic CRFs (eCRFs), which are provided by electronic data capture (EDC) systems [2]. Automatic evaluation of data directly during entry, modelling of specific views for various roles participating in the trial (e.g. research nurse), and web-based designs for de-central data capture for multi-centered studies, is supported. This reduces errors, latency and costs in the study workflow [3].

However, available EDC systems lack in support of subject's medical image data. Neither integration of subject's digital imaging and communications in medicine (DICOM) data into eCRFs, nor interfaces for communication with picture archiving and communication systems (PACS), is supported. Hence, image data captured in trials has to be manually mapped to the correct eCRF afterwards. This produces additional steps for the study personnel in the data capture workflow, neutralizing the advantages of eCRFs. In special regarding multi-centered studies with de-central data capture on multiple sites, error-proneness, latency and costs increase.

Hence, a first approach to connect an EDC system with a PACS has been presented by van Herk [4]. Here, the PACS Conquest is integrated into the popular EDC system OpenClinica. Conquest is queried from inside the eCRF for subject's medical images. The DICOM objects are de-identified and connected to the eCRF via the web access to DICOM persistent objects (WADO) protocol [5]. After this, the subjects' eCRF and DICOM data is transferred to a research server. In this approach, DICOM objects have to be available in the PACS first and advanced DICOM viewing functionality is not available.

*Corresponding author: Daniel Haak, Department of Medical Informatics, RWTH Aachen University, Pauwelsstr. 30, D - 52057 Aachen, Germany, email: dhaak@mi.rwth-aachen.de; phone: +49 241 80 85174, fax: +49 241 80 82426.

Medical Imaging 2016: PACS and Imaging Informatics: Next Generation and Innovations, edited by Jianguo Zhang, Tessa S. Cook, Proc. of SPIE Vol. 9789, 97890M · © 2016 SPIE · CCC code: 1605-7422/16/$18 · doi: 10.1117/12.2218257

A similar approach has been presented by Skripcak et al. [6], also connecting Conquest and OpenClinica scripts. In this approach, subjects' image data is integrated into the eCRF via a standalone client. After this, the data is forwarded to the PACS and WADO references inserted in the eCRF. For retrieval of the DICOM data, the DICOM viewers Weasis and DVW have been integrated. However, in this architecture, the DICOM object cannot be integrated via web. A special software has to be installed on all client systems.

In our previous work, a comprehensive evaluation of 28 DICOM viewer tools has been performed [7]. Including open source, free and commercial tools the viewer have been investigated concerning 26 criteria. The criteria include (i) platform, (ii) interface, (iii) support, (iv) two-dimensional (2D), and (v) three-dimensional (3D) viewing functionality. The criteria results have been analyzed regarding the use cases (i) central viewing in hospitals, (ii) de-central viewing in web-based systems and (iii) advanced viewing with rich 3D functionality. The viewer Oviyam, Weasis, and Xero have been suggested as optimal for the de-central viewing use case, which is particular important for multi-centered clinical trials.

Based on these results, and in contrast to the other approaches, we demonstrated a system that allows to integrate subject's image data directly via the eCRF [8]. The architecture utilizes the open source components OpenClinica as EDC system, DCM4CHEE as PACS and Weasis as DICOM viewer software. Communication between the components is performed via WADO.

In this paper, a final system architecture connecting OpenClinica, DCM4CHEE and Weasis is presented allowing decentral storage and retrieval of DICOM objects in multi-centered studies. The system is demonstrated by a clinical trial, in which subject's computer tomography (CT) data is de-centrally captured, but centrally read by a chief radiologists to decide on inclusion of the subjects in the trial or exclusion as screening failure.

2. MATERIALS AND METHODS

2.1 OpenClinica

OpenClinica (OC) is a popular EDC and clinical data management system [2,3]. The web application offers rich functionality for storage and management of subject's data in clinical trials. OC follows industry standards and has been approved by the Food and Drug Administration (FDA). The Community Edition (Version 3.5) is licensed as open source. Hence, OC is, in special, attractive for low-budgeted multi-centered studies. OC is one of the most popular EDC systems and supported by a large user community continuously developing extensions for the system such as OC-Big.

2.2 OC-Big

OC-Big[1] (Version 1.1.3) is an add-on for OC, which allows to integrate binary large object (BLOB) data into OC's eCRFs [9]. The extension is embedded into the eCRF and substitutes OC's native file upload component. OC-Big ensures stable transfer of the data (e.g. by automatic splitting into smaller parts) and considers the context of the data (e.g. subject, study, event identifier) by re-labelling of the data after successful transfer.

2.3 DCM4CHEE

DCM4CHEE[2] (Version 2.17.2) is a PACS and part of DCM4CHE, a community-driven software collection of open source utilities for the healthcare enterprise. DCM4CHEE is open source and includes a web frontend. The PACS has a modular structure and offers several interfaces, e.g. WADO to serve DICOM objects via the web.

2.4 Weasis

Weasis[3] (Version 2.0.2) is a web-based DICOM viewer, which is also offered by the DCM4CHE community. The open source tool is developed in Java. An additional component, the Weasis PACS Connector (WPC), allows Weasis to gather DICOM data from PACS such as DCM4CHEE. In addition, Weasis can be instantiated in web browser via the Java network launching protocol (JNLP).

[1] http://idmteam.github.io/oc-big/
[2] https://dcm4che.atlassian.net/wiki/display/ee2/Home/
[3] https://dcm4che.atlassian.net/wiki/display/WEA/Home/

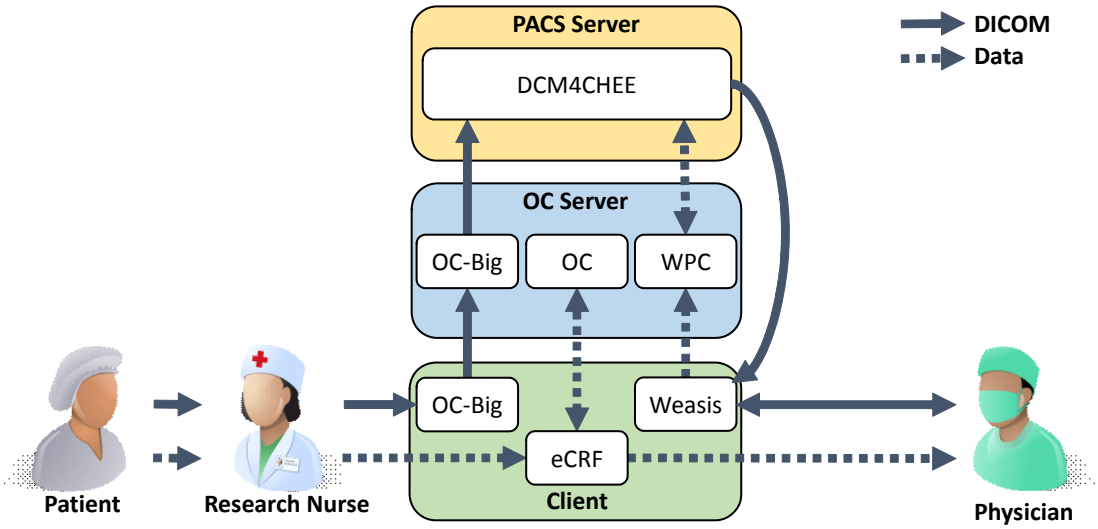

Figure 1: System architecture of EDC and PACS integration.

2.5 System Architecture

Utilizing the presented components, our system architecture (Fig. 1) is composed of (i) OC's eCRF, (ii) OC-Big's web front-end, and (iii) Weasis, on the client side. A middleware component is interconnected, which includes (iv) OC's and (v) OC-Big's backend, as well as (vi) the WPC. In addition, (vii) DCM4CHEE is connected as PACS.

While the usual subject's study data is captured by the research nurse via the eCRF, DICOM data is transferred by OC-Big via the hypertext transfer protocol (HTTP) to the middleware component. Here, the data is extracted, de-identified, and context information in the DICOM header re-labelled. After this, the DICOM data is stored in DCM4CHEE. During this a WADO reference is generated and inserted into the eCRF.

A mouse click on this reference by the physician, triggers the invocation of Weasis. Weasis queries DICOM objects corresponding to the subject identifier from DCM4CHEE via the WPC. The resulting image objects are retrieved by Weasis and visualized in the browser.

2.6 De-Identification and Record Linkage

Usually, DICOM data of clinical trials is not neatly de-identified and subject related data is still stored in the DICOM header. Hence, the record has to be de-identified and connected to the context of the subject in the study. In the first step the DICOM header is de-identified by erasure of a set of 22 DICOM tags, which are designated to store subject's identification data such as the SubjectName (0010,0010) field. After this, the DICOM fields SubjectID (0010,0020), ClinicalTrialSubjectID (0012,0040) and ClinicalTrialSiteID (0012,0030) are re-labelled with the subject, study and event ID, which are extracted from the eCRF.

2.7 Evaluation Study

The workflow of the system architecture is exemplarily demonstrated in the VitaVasK study (ClinicalTrails.gov #NCT01742273), where the effect of vitamin K1 on the human body is analyzed. In this study, CT data is de-centrally captured in multiple sites and centrally viewed by an expert, who analyzes the images to decide on inclusion of the subject in the trial or exclusion as screening failure. The CT data is collected in the sites Aachen, Düsseldorf, Erlangen Coburg (Germany) and Stockholm (Sweden). So far 40 of 300 planned subjects have been enrolled.

3. RESULTS

The resulting EDC workflow of subject's CT images in the VitaVasK study is as follows (Fig. 2): On the study's sites, the research nurses invokes OC-Big from inside the eCRF and selects subject's DICOM data from the file system (compressed in a zip archive here) (Step 1). The data is transferred by OC-Big to the PACS and references to the

DICOM objects are stored in the eCRF (Step 2). For viewing, the radiologist on the central site clicks on a reference. This triggers the invocation of Weasis and the subject's CT images are immeditaley visualized. Rich functionality for image interaction (e.g. measurements) is now available (Step 3). So far 29 CT data sets of the 40 subjects enrolled in the VitaVasK study have been successfully integrated, and these subjects have been included or excluded based on this data.

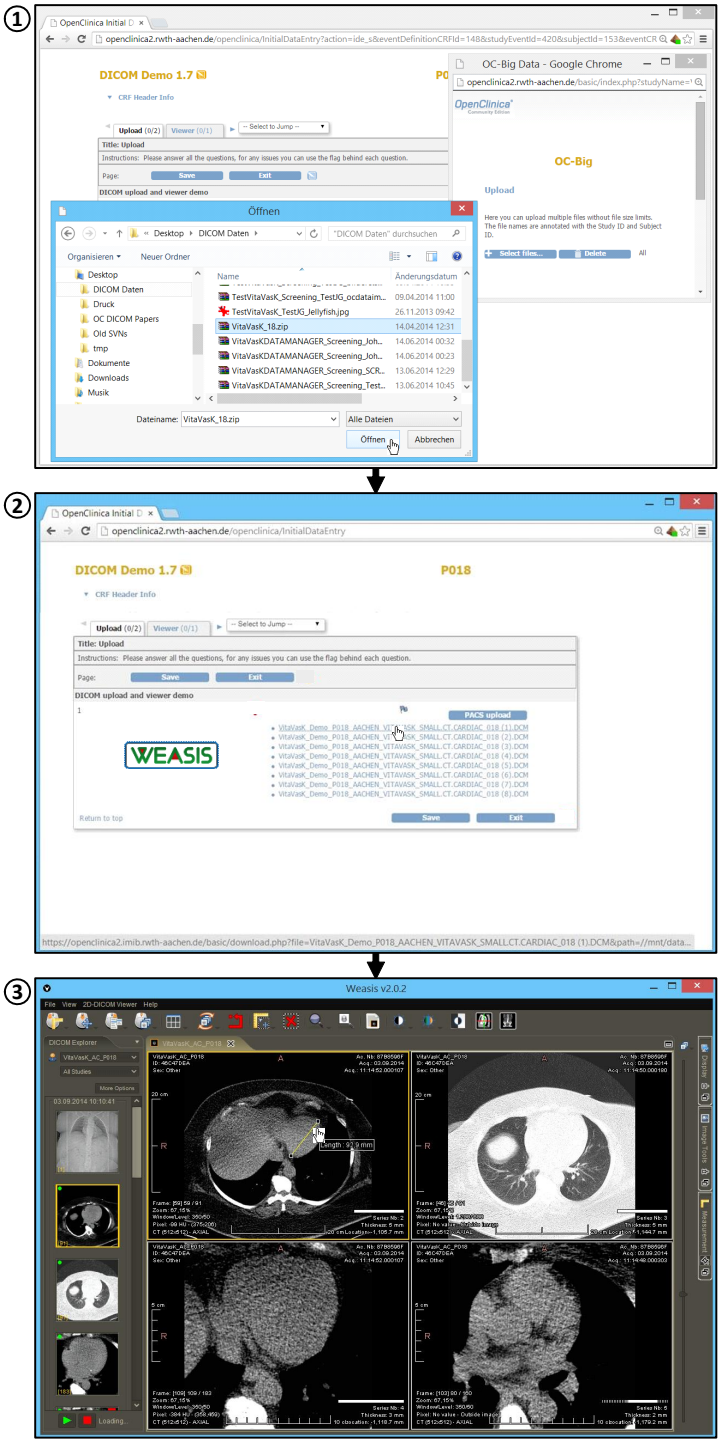

Figure 2: Final EDC workflow in the eCRF with DICOM & PACS integration.

4. DISCUSSION

Focusing on EDC in multi-center studies a completely web-based system has been developed, consisting of the open source components OpenClinica, DCM4CHEE and Weasis. The architecture allows integration and retrieval of subject's DICOM data directly in the eCRF. The data is automatically de-identified and connected to the context of the subject in the study, ensuring data privacy and consistency, respectively. In addition, a research database is implicitly built-up in the background, collecting de-identified DICOM data of several clinical trials, which can be maybe re-used in future for currently not known research questions.

With Weasis, a powerful DICOM viewer is integrated. As evaluated in [7], Weasis achieved the highest rank regarding optimal workflow in multi-center clinical trials and rich functionality for two- (2D) and three-dimensional (3D) image viewing (e.g. windowing, pseudo-coloring) and interaction (e.g. measurements, annotations) are available. Hence, Weasis includes rich functionality to support the physicians during the diagnosis.

In our architecture the eCRF is the primary endpoint for the research nurse. While the system communication is hidden from the user in the background, the interaction completely takes place in the eCRF. This is in contrast to the approaches of van Herk and Skripcak et al., in which the data has to be stored in the PACS or integrated via standalone client first, respectively.

However, integration and retrieval of DICOM objects directly in the eCRF, is in particular important in multi-centered trials, where subject's DICOM data is collected de-centrally. As illustrated in Figure 3 (top), the workflow of the study personnel suffers, if subjects' DICOM data cannot be integrated into the eCRF. To transfer the DICOM data on traditional ways, (i) the material for data acquisition (e.g., patient identifier labels on memory cards) has to be prepared and labelled; (ii) the material has to be bundled and shipped to the imaging site via mail; (iii) the image or biosignal data has to be recorded; (iv) the data has to be copied to a transferable memory card or compact disc; (v) the data has to be shipped to an expert; (vi) the analysis results of the expert have to be sent back via facisimile and, finally, (vii) the results

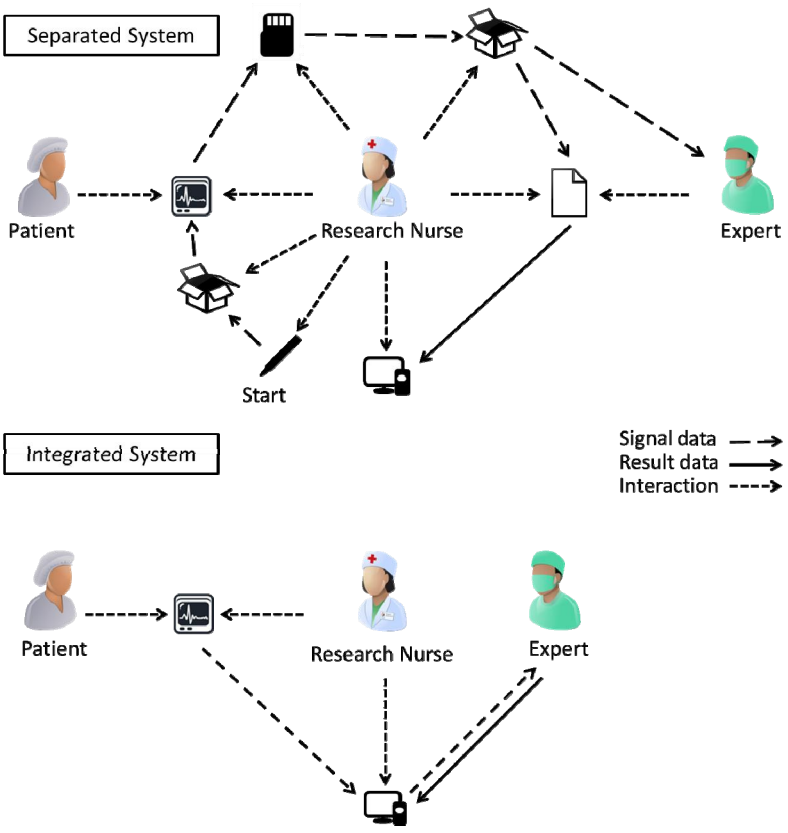

Figure 3: Workflow improvement by integration of systems.

have to be manually entered into the EDC system. In contrast, using an integrated systems the study personnel is able to transfer the eCRF and DICOM data directly via the central EDC system and de-central access to subjects' data. This massively reduces the number of error-prone manual steps.

However, in our architecture, the image data is currently transferred via WADO to the DICOM viewer. Since WADO does not support any storage of results data back into the PACS, annotations (e.g. DICOM-SR) made by the physicians are not stored. In future, this could be solved by, for instance, by introducing DICOM Store over the Web by Representations State Transfer (STOW-RS) [10] communication.

CONCLUSION

We presented a system architecture, which connects the EDC system OpenClinica with the DCM4CHEE PACS and the DICOM viewer Weasis. Using such a system, the workflow for the study personnel in the multi-centered trials is simplified and errors, latency and costs are reduced. As exemplarily demonstrated in the VitaVasK study, a remote decision of subject's inclusion in a trial can be made immeditaley at a central side, although the DICOM data of the subject has been de-centrally captured in multiple sites. Integration of advanced DICOM functionality in the system (e.g. for storage of annotations) will be investigated in future.

REFERENCES

[1] Miller CG, Krasnow J, Schwartz LH. Medical imaging in clinical trials. London: Springer London; 2014

[2] Pavlović I, Kern T, Miklavčič D. Comparison of paper-based and electronic data collection process in clinical trials: Costs simulation study. Contemporary Clinical Trials 2009; 30(4):300–16.

[3] Shah J, Rajgor D, Pradhan S, McCready M, Zaveri A, Pietrobon R. Electronic data capture for registries and clinical trials in orthopaedic surgery: open source versus commercial systems. Clin Orthop Relat Res 2010; 468(10):2664-71.

[4] van Herk M. Integration of a clinical trial database with a PACS. JPCS. 2014;489:12099

[5] Koutelakis GV, Lymperopoulos DK. PACS through web compatible with DICOM standard and WADO service: advantages and implementation. In: International Conference of the IEEE Engineering in Medicine and Biology Society; 2006. p. 2601–5

[6] Skripcak T. Lessions learned from integrating OpenClinica with other IT systems [Internet]. [Cited 28 August 2015]. Available from: https://community.openclinica.com/sites/fileuploads/akaza/cms-community/Tomas%20Skripcak%20-%20Lessons%20learned.pdf

[7] Haak D, Page C, Deserno TM. A Survey of DICOM Viewer Software to Integrate Clinical Research and Medical Imaging. J Digit Imaging. Epub ahead of print.

[8] Haak D, Page CE, Reinartz S, Krüger T, Deserno TM. DICOM for clinical research: PACS-integrated electronic data capture in multi-center trials. J Digit Imaging 2015; 28: epub ehead of print.

[9] Haak D, Samsel C, Gehlen J, Jonas S, Deserno TM. Simplifying electronic data capture in clinical trials: workflow embedded image and biosignal file integration and analysis via web services. J Digit Imaging 2014;27(5):571–80.

[10] Le Maitre A, Fernando J, Morvan Y, Mevel G, Cordonnier E. Comparative performance investigation of DICOM C-STORE and DICOM HTTP-based requests. In: 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). p. 1350–3