*MEDINFO 2010*
*C. Safran et al. (Eds.)*
*IOS Press, 2010*

*doi:10.3233/978-1-60750-588-4-1324*

# The REUSE project: EHR as single datasource for biomedical research

**AbdenNaji El Fadly[a], Noël Lucas[b], Bastien Rance[b], Philippe Verplancke[d], Pierre-Yves Lastic[c], Christel Daniel[ab]**

[a] *INSERM, UMR_S 872, Eq. 20, Paris, F-75006 France; Université René Descartes, Paris, F-75006 France*
[b] *AP-HP, Hôpital George Pompidou, Département d'Informatique Hospitalière, Paris, F-75015 France*
[c] *CDISC Board of Directors and Sanofi-aventis R&D, Bagneux, F-92220, France*
[d] *XClinical GmbH, Munich, Allemagne;*

**Abstract**

*Integrating biomedical research and patient care is a challenging issue requiring interoperability solutions. During a clinical trial, clinical data are captured twice, first in the Electronic Health Record (EHR) and then in the Clinical trials Data Management System (CDMS). The aim of REUSE (Retrieving EHR Useful data for Secondary Exploitation) project is to provide a single source solution for electronic data capture to the investigators of a university hospitals involved in a multi-centric clinical trial. We first investigated the differences between the workflows of patient care and biomedical research to specify the use of EHR for clinical trials. Then we defined a semantic interoperability framework in order to enable the reuse of EHR clinical data and implemented a mediator that transforms CDISC Operational Data Model (ODM) XML into proprietary XML document templates of different EHR solutions and vice-versa. Implementing electronic data capture for biomedical research within EHR eliminates redundant data entry, thus improving data quality and processing speed. Moreover, unlike other initiatives such as IHE integration profile "Retrieve Form for Data Capture" (RFD), the REUSE approach ensures that all clinical data is kept in the EHR whatever the context of data capture is.*

***Keywords:***

HL7, CDISC, ODM, IHE profile, RFD, Electronic health record, Clinical data management system, Biomedical research, Clinical trial, Cardiovascular, Single source.

## Introduction

The role that EHR could play in a biomedical research context has been investigated in many recent works. Clinical data in EHR may be useful during the design phase providing a better understanding of real patient population; during the instruction phase, by improving patient recruitment in clinical trials, as well as during the implementation phase by optimizing clinical data entry or adverse event reporting [1-4].

Current information systems for patient care on one hand, and for biomedical research on the other hand, are completely disconnected, even in university hospitals which have the most successful implementations of Electronic Health Records (EHR).

Since, there is evidence for a relevant overlap of routine and research oriented documentation [5,6], interoperability between Electronic Healthcare Record (EHR) and Clinical trials Data Management System (CDMS) for research would be of major advantage. Factors limiting EHR/CDMS interoperability are related to organizational, regulatory and technological issues.

**From an organizational point of view**, patient care and clinical research processes are very different. Furthermore, the functionalities expected from EHR and CDMS differ significantly in these two contexts. Recent efforts have been made to expand the functionality of EHRs for their use in clinical research.

As an example, the HL7 EHR System Functional Model [7] provides a reference list of over 160 functions that may be present in EHRs. Its functionality has been recently expanded by the global EHR/CR Functional Profile Project [8], a collaborative effort to expand and adapt the functionalities of EHR and associated systems, networks, and processes to support clinical research.

**From a regulatory point of view**, health data processing is subject to different constraints in patient care and biomedical research. Concerning data security and system validation, requirements for CDMS are also subject to higher constraints than for EHR. In the USA for instance, a system used to manage clinical trials data needs to comply with the FDA [9] Guidance "Computerized Systems Used in Clinical Investigations" (CSUCI) [10].

**Finally, from a technical point of view**, the standards for archiving and transmitting health data are defined by different organizations. On one hand, in the domain of biomedical research, the CDISC (Clinical Data Interchange Standards Consortium) organization [11] was created in 1997 by a group of individuals from pharmaceutical companies, clinical research organizations and software providers, working closely with the US FDA in order to develop platform-independent standards that support the electronic acquisition, exchange, regulatory submission and archiving of health data.

In 2001, CDISC published the first version of its Operational Data Model (ODM) [12], that defines the different types of information related to clinical trial: clinical trial metadata (definition of protocol and used items), baseline data of the clinical trial (e.g. the normal values of laboratory tests), user's administrative data (access rights and user profiles), patient clinical data and audit trail data (tracking data changes).

On the other hand, for patient care, the efforts of structuring health information and communication systems, are carried out within the European (CEN TC 251) [13] and international (HL7, Health Level 7) [14] committees for standardization. With regards to the EHR, in 2003 the Reference Information Model (RIM) proposed by HL7 became an ISO standard [15], from which are derived the models of both messages and documents (such as Clinical Document Architecture (CDA)) [16]. In order to facilitate the integration of patient care and biomedical research information systems, efforts have been initiated between HL7 and CDISC within the RCRIM (Regulated Clinical Research Information Management) technical committee of HL7, in which NCI (National Cancer Institute) [17] and FDA are actively involved. As part of these efforts, the BRIDG project aims to create a Domain Analysis Model covering the entire field of biomedical research and to develop common semantics for the different actors (hospitals, clinical trials sponsors, health authorities, etc.) [18-20].

We identified some initiatives that propose solutions for integrating information systems for patient care and biomedical research [21,22].

According to the CDISC Electronic Source Data Interchange Initiative [23], we can distinguish three different solutions for EHR/CDMS integration. First in the "*source data extraction and verification"*, part of the source data are extracted from the EHR, the investigator has to verify that the extracted clinical research data from the EHR reflect the source data required for the clinical trial before they are to a separate clinical trial database (CDMS)). In the "*single source"*, source data are all captured into the EHR and the clinical trial data are exported to the CDMS. In "*EHR used as CDMS"*, the source data are all captured into the EHR which is the clinical trial database (CDMS). Many pilot projects have implemented one of these different types of EHR/CDMS integration solutions allowing the reuse of patient care data, but with rare exceptions most examples consist of institution-specific approaches that do not use standards and thus lack the broad interoperability required for multicenter trials

Among "*source data extraction and verification"* solutions, Integrating the Healthcare Enterprise (IHE) published the Retrieve Form for Data-capture (RFD) "integration profile" [24]. Combined with the Clinical Research Content Profile, this integration profile provides a method of data capture within a used application (e.g. an EHR) that meets the requirements of an external system (e.g. a CDMS). An IHE "integration profile" is a real situation describing exchanges of information, called "transactions" of different components of a distributed health information system, called "actors". IHE provides guidelines for implementing these "transactions", using established computer standards such as HL7 or DICOM.

The RFD profile specifies a solution to integrate EHR and CDMS. If an external organization needs to collect clinical data as part of a clinical trial, the RFD profile will allow the EHR user (study investigator, for example) to access to the electronic Case Report Form (eCRF) of the clinical trial without leaving the EHR, and enter the required information so that they are collected by the external agency. However this solution still constrains the clinician to provide and manage two distinct data sources, duplicating data entry from one to another except for only several items described in the Clinical Research Content Profile.

In this context, the objective of REUSE (Retrieve data in EHR Useful for Secondary Exploitation) project is to provide a "single source" integration solution between EHR and CDMS using international IT standards (CDSIC, HL7) and to evaluate it in the hospital context of Assistance Publique-Hôpitaux de Paris (AP-HP).

In this paper, we first describe how we modeled the biomedical research process in a university hospital environment in order to specify the use of EHR in this context. Then we present the design and implementation of the "single source" solution between EHR and CDMS that takes into account the regulatory constraints of biomedical research and addresses semantic interoperability issues. At last we present the experimentation of the REUSE architecture in the context of a multi-centric clinical trial in the cardiovascular domain.

## Material and methods

### Material and context: AP-HP clinical trial "Arcadia"

The AP-HP is the most important French University Hospital Organization (38 hospitals with about 23,000 beds , 1400 one-day care, 850 capacities of care at home and 1,000,000 hospitalized patients per year; 90,000 employees including 19,000 physicians).

The aim of the New Information System (NIS) project of the AP-HP, is to implement the EHR Orbis ® (AGFA Healthcare ©). However some hospitals such as the Georges Pompidou European Hospital (HEGP) already have an EHR e.g. DxCare® (Medasys©) [25].

In AP-HP, biomedical Research is carried out within 18 Research Institutes; 8 Clinical Investigation Centers (phase I & IIa, Pharmacokinetic/Pharmacodynamic (PK/PD), etc.), 10 Clinical Research Units and 100 INSERM teams. The AP-HP direction of the biomedical research promotes electronic data capture and some experiments are conducted in different hospitals. Within the HEGP Clinical Research Unit several Clinical Data Management Systems (CDMS) are being evaluated such as OpenClinica® (Akaza Research©) or Marvin® (XClinical©) [26].

Arcadia is a biomedical research study conducted at HEGP Clinical Research Unit in the cardiovascular domain that was classified in the "common care" field of French law. The use of EHR as data entry solution was subjected to the favorable opinion of a Committee for the Persons Protection (CPP) [27], which are the French Institutional Review Board (IRB).

### Business analysis and design of REUSE

We conducted a series of meetings with the AP-HP managers (from AP-HP direction of the biomedical research), clinicians and researchers (from the Clinical Research Unit (CRU) of HEGP)) in order to carry out the business process models describing the clinical trial implementation activity in a context of institutional sponsorship in France. Then we defined the different stages composing the process which could be simplified and automated thanks to a use of the EHR.

We integrated into the process models elements from the BRIDG dynamic model comprising the use cases and the diagrams activity of clinical trials. We also took into account the CDMS specifications defined in the FDA guidelines "CSUCI" and in the EHR Functional Model, extended to biomedical research.

### REUSE implementation

The REUSE architecture implements the “single source” concept so that all data, whether entered in the patient care context or in the biomedical research context, are stored and accessible in the EHR. We developed modules that allow: 1) to import within the EHR a SNOMED encoded electronic case report form (eCRF) of a clinical trial, 2) to align existing clinical data with the eCRF data elements and 3) to export the clinical trial data towards a CDMS.

We used JAVA programming language for its stability, its large selection of library and because it is platform-independent, particularly JDOM Application Program Interface (API) [28] for its simplicity compared to SAX and DOM parser [29] and for the richness of its functionalities.

## Results

### Process models and functional perimeter of REUSE

We modeled the biomedical research activities of institutional sponsorship in a university hospital and specified the functional perimeter of the EHR within this context.

The activity diagrams made it possible to better analyze the main phases of a clinical trial: the submission phase where the principal investigator and the CRU edit the research protocol and submit it to the sponsor; the instruction phase where the CRU prepares all the tools for the data management and completes the regulatory and financial tasks and the implementation phase during which the patients are included (information of the actors, inclusion lists, patients consents) and the data captured before monitoring and statistic analysis. This analysis made it possible to better identify the role of the main actors at the time of each activity of the clinical trial implementation process (physician-investigator, methodologist, Structure Coordinator, Clinical Trial Coordinator (CTC), statistician, Data-Manager, Clinical trial technician (CTT), Clinical Research assistant (CRA), Administrative and Financial Manager (AFM) and sponsor). The REUSE business process diagram has a hierarchical structure: each modeled phase is made up of several structured activities, themselves made up of sub-activities. In total, it is composed of 15 activity diagrams from which we identified 7 activities and tasks likely to be simplified and automated, thanks to the use of the EHR: "To develop a clinical trial database", "To propose a data-management plan", "To develop eCRF", "To carry over existing data in the EHR", "To modify administrative patients data", "To modify data of clinical research", "To Monitor data» (Figure 1).

Each activity included in the functional perimeter of the EHR is described by a sequence diagram specifying the actors’ interactions using EHR within the REUSE solution.

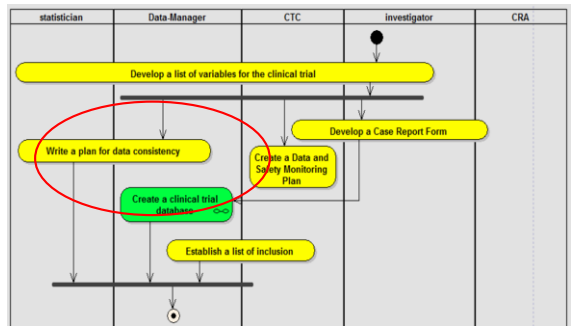


*Figure 1- Example of process activities of clinical trial implementation which can benefit from the use of the EHR: "To create a clinical trial database" and "To propose a data-management plan"*

### REUSE implementation

We defined the “Retrieve & Integrate Form for Data capture” (RIFD) integration profile in order to implement the “single source” concept for the REUSE project (Figure 2).

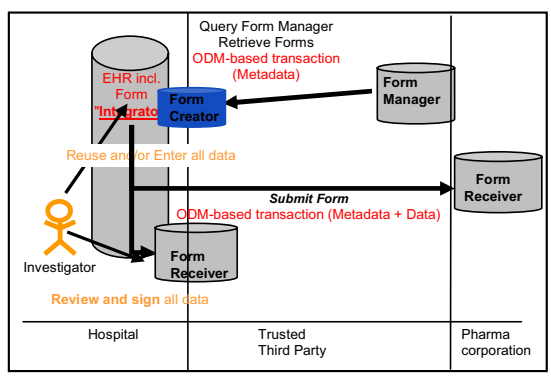


*Figure 2- Actors and transactions involved in the "Retrieve & Integrate Form for Data Capture" (RIFD) integration profile proposed*

The “Retrieve & Integrate Form for Data capture” (RIFD) integration profile consists in three component or “actors”: “Form Manager”, “Form Creator” and “Form Receiver” and two transactions “Query-Retrieve Form” and “Submit Form”.

***The transaction "Query-Retrieve Form"***

This transaction consists in importing within the EHR an eCRF of clinical trial provided by the Contract Research Organization (CRO). It relies on an ODM message to transfer the metadata of the eCRF from the "Form Manager" to the "Form Creator". Then, the eCRF is encoded using SNOMED v3.5 VF codes as much as possible and integrated into the EHR. The pre-filling process of the eCRF by pre-existing clinical data in the EHR is achieved through mechanisms specific to the EHR. These mechanisms are based on mappings between clinical items of different forms (used in a patient care as well as in biomedical research context) and SNOMED v3.5 VF used as pivot terminology.

***The transaction "Submit Form"***

This transaction consists in exporting the data of the clinical trial captured in the EHR towards the CDMS. It relies on an ODM-based transaction between the "Form Filler" and the "Form Receiver". Data validation is done in the CDMS, i.e. in an application which the management depends neither on the investigator nor on the sponsor.

The "Retrieve & Integrate Form for Data capture" (RIFD) integration profile was implemented by the EHR DxCare® (Medasys©) and the CDMS Marvin® (XClinical©).

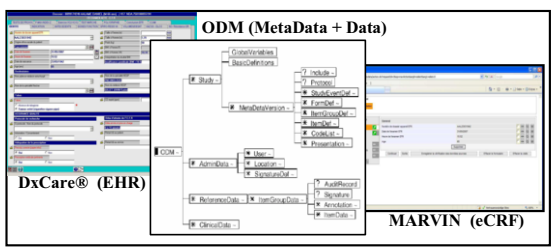


*Figure 3- Export of clinical data and meta-data of a clinical trial from the EHR (DxCare, Medasys) to CDMS (MARVIN, XClinical) using the REUSE CDISC ODM mediator.*

In HEGP, in the context of the "Arcadia" multi-centric study, figure 3 shows how the CDISC ODM mediator allows transferring data entered into the EHR to the CDMS.

Within the EHR, the ARCADIA study consists of seven forms: Inclusion, initial evaluation, initial abdominal imaging, initial neurovascular imaging, sample, adverse events, and serious adverse events. In using the CDISC ODM mediator, the forms and responses to questions of each patient participating in the clinical trial, have been respectively transformed into Study Event and Subject Data and been generated in the ODM output file with ensuring that the meaning of the data is appropriately preserved, so that the receiving system (ie, researchers) could appropriately use the data.

## Discussion and Conclusion

We have modeled the implementation process of clinical trials of institutional sponsorship in a university hospital in order to design, develop and test an integration profile between Electronic Health Record (EHR) and Clinical Data Management System (CDMS) and exploit the EHR clinical data for biomedical research.

The benefits for the health care professionals consist on the economy of a double entry, which is error-prone and time consuming. The benefits for the CRU are to facilitate the data monitoring quality in the electronic Case Report Form. The automatic transfer of data reduces the risk of error, monitoring focuses on data that are not pre-filled from the EHR. Last but not least, due to the "single source" approach that we have adopted, the benefits for the patient are related to the storage in the EHR of all clinical data collected either in a patient care context or a biomedical research context. We believe that considering the EHR as the single source of data may improve patient safety insofar as some of the data collected in a biomedical research setting, such as adverse events of treatment delivered in the clinical trial context, are likely to be useful to support diagnosis and treatment of the patient.

Unlike the RFD IHE profile the "Retrieve & Integrate Form for Data capture" (RIFD) integration profile, that we have implemented in REUSE, not only allows displaying an eCRF within an EHR, but also importing and integrating it in the EHR. This increases the possibilities of pre-filling the electronic Case Report Form (eCRF) compared to what is proposed in the RFD IHE profile. Indeed, even when the RFD profile is completed by the Clinical Research Data capture (CRD) profile [30], the pre-filling of the eCRF is limited to some predefined clinical items. These items are selected from the content modules of IHE Continuity of Care Document (CCD) [31] because they are common to patient care and biomedical research. This set of items, defined generically for every clinical trial, will be by nature limited and therefore, from our point of view, will only slightly address the double entry issue.

An interoperability framework and the use of data standards (HL7 or CDISC "templates" and reference terminologies) are essential for any successful implementation of a data collection system designed to reuse patient data in the context of multi-centric clinical trials. Multicenter trials require data from different sites to be submitted to a central data center, with whom the site's relationship may exist for only a single trial. For reuse of patient care data to be feasible, data collection methods must be easy to implement and use, and must minimize disruption at the clinical site.

Furthermore, successful adoption of IT systems has always been a challenge involving numerous social and organisational factors. The understanding of end user needs and workflow is a key element of the REUSE project. Indeed, these factors must be considered even more carefully when considering a fundamental re-engineering of an enterprise like clinical research enterprise.

Standards have the potential to make this possible by allowing investigational sites to use existing systems without the burden of data transformation.

The main drawback of the "single source" approach is that a new "actor" – "Form Creator" - has to be implemented by the EHR in order to integrate eCRF designed using the CDISC ODM format.The experience gained during the project REUSE should allow us to build an interoperability framework using HL7 and CDISC standards, by developing and implementing the content standards such as a clinical note input templates and eCRF templates derived from protocol specification.

This will facilitate the implementation of Single Source solutions in a European multi-centric context including different vendors, different CROs and different sponsors.

**Acknowledgments**

We would like to express our gratitude to the staff of HEGP, particularly the staff of URC and the computer department, Mr Nicolas DE SAINT JORRE and Prof. Joël MENARD.

## References

[1] Powell J, Buchan I. Electronic Health Records Should Support Clinical Research. J Med Internet Res. 2005; 7(1): e4.

[2] Pakhomov S, Weston SA. Electronic Medical Records for Clinical Research: Application to the Identification of Heart Failure. Am J Manag Care. 2007; 13(part 1):281-288.

[3] Murphy EC, Frederick L. An Electronic Medical Records System for Clinical Research and the EMR–EDC Interface. National Institutes of Health. 2007; 48(10): 4383–4389.

[4] The Future Vision of Electronic Health Records as eSource for Clinical Research (Version 1.0). Anonymous eds. eClinical Forum / PhRMA EDC/eSource Taskforce. 2006.

[5] Williams JG, Cheung WY, Cohen DR, Hutchings HA, Longo MF, Russell IT. Can randomised trials rely on existing electronic data? A feasibility study to explore the value of routine data in health technology assessment. Health Technology Assessment 2003; 7(26):1–117

[6] Kahn MG, Kaplan D. Configuration Challenges: Implementing Translational Research Policies in Electronic Medical Records. Academic Medicine. 2007; 82(7):661-9.

[7] The HL7 EHR System Functional Model : http://www.hl7.org/ehr/downloads/index_2007.asp

[8] EHR/CR Functional Profile Project : http://www.ehrcr.org

[9] Food and Drug Administration: http://www.fda.gov

[10] FDA Guidance for Industry "CSUCI": http://www.fda.gov/cder/guidance/7359fnl.htm

[11] Clinical Data Interchange Standards Consortium: http://www.cdisc.org

[12] Operational Data Model: http://www.cdisc.org/models/odm/v1.3/index.html

[13] Comité européen de normalization, Technical Committee 251: http://cen.iso.org

[14] Health Level Seven: http://www.hl7.org

[15] HL7 RIM: http://www.hl7.org/Library/data-model/RIM/modelpage_mem.htm

[16] Klein A, Prokosch HU, Müller M, Ganslandt T. Experiences with an Interoperable Data Acquisition Platform for Multi-centric Research Networks Based on HL7 CDA. Methods of Information in Medicine. 2007; 46 5: 580-5.

[17] National Cancer Institute: http://www.nci.nih.gov

[18]Biomedical Research Integrated Domain Group (BRIDG) model: http://www.bridgmodel.org

[19] Souza T, Kush R, Evans J. Global Clinical Data Interchange standards are here. Drug Discov Today. 2007; 12(3-4):174-81.

[20] Fridsma DB, Evans J, Hastak S, Mead CN. The BRIDG Project: A Technical Report. J Am Med Inform Assoc. 2008; 15: 130-7.

[21] Integrating the Healthcare Enterprise: http://www.ihe.net.

[22] Healthcare Information and Management Systems Society: http://www.himss.org

[23] CDISC Electronic Source Data Interchange Initiative: http://www.cdisc.org/stuff/contentmgr/files/0/2f6eca8f0df7caac5bbd4fadfd76d575/miscdocs/esdi.pdf

[24] RFD Supplement. Anonymous eds. IHE IT Infrastructure Technical Framework 2006.

[25] MedaSys: http://www.medasys.com

[26] XClinical: http://www.XClinical.com

[27] Picard S, Pellet J, Brulet JF, Trombert B. Les aspects juridiques et éthiques de la protection des données issues du dossier médical informatisé et utilisées en épidémiologie : un point de la situation. Santé Publique. 2006; Vol 18 :107-17.

[28] JDOM: http://www.jdom.org

[29] Java API for XML Processing: http://java.sun.com/j2se/1.4.2/docs/api/

[30] CRD Supplement. Anonymous eds. IHE Quality, Research and Public Health (QRPH) Technical Framework 2008.

[31] Dolin RH, Giannone G, Schadow G. Enabling Joint Commission Medication Reconciliation Objectives with the HL7 / ASTM Continuity of Care Document Standard. AMIA. 2007; 186-190.

**Address for correspondence**

Naji El Fadly, Christel Daniel
INSERM UMR_S 872, Eq 20
Faculté de Médecine Paris V
15, rue de l'École de Médecine 75006 Paris – France
Mail: nelfadly@yahoo.fr; christel.daniel@crc.jussieu.fr