Integrating Clinical Research to the Healthcare Enterprise: from the RE-USE project to the EHR4CR platform

AbdenNaji El Fadly¹, Bastien Rance²*, Noël Lucas², Charles Mead³, Gilles Chatellier², Pierre-Yves Lastic⁴,⁵, Marie-Christine Jaulent¹, Christel Daniel¹,²,⁶

¹ INSERM, UMR_S 872, Eq. 20, Université Paris Descartes, Paris
²AP-HP, Clinical Research Unit, Hôpital George Pompidou, Paris, France
³National Cancer Institute, 3rd Millenium Consulting, Center for Biomedical Informatics and Information Technology, 2115 East Jefferson St., 6th floor Rockville, MD 20852 USA
⁴CDISC Board of Directors and Sanofi-aventis R&D, Chilly-Mazarin, F-91385, France
⁵National Cancer Institute, 3rd Millenium Consulting, Data Privacy and Healthcare Interoperability Standards, USA
⁶AP-HP, Medical Informatics Department, Hôpital George Pompidou, Paris, France

* Current affiliation: National Library of Medicine, NIH, USA
bastien.rance@nih.gov
Most of the slides of this presentation are borrowed from Christel Daniel
Reusing EHRs in Clinical Research

• During the design phase
  – Clinical trial feasibility

• During the instruction phase
  – Patient recruitment
    • Only 7% of eligible patients enroll in a clinical trial
    • 86% of all trials fail to enroll on time

• During the implementation phase
  – Spontaneous reporting of drug adverse events
  – Data capture
    • 30 to 50% of redundant data (between EHRs and EDC system)

Draft version 0.1, March 3, 2006; The eClinical Forum and PhaRMA EDC. The Future Vision of Electronic Health Records as eSource for Clinical Research
Why is it so difficult

• Different **workflows**

• Different **ethical and regulatory contexts**
  – FDA Guidelines Computerized Systems Used in Clinical Investigations (CSUCI), 21 CFR Part 11, HIPAA
  – …

• Different information technology **standards**
  – Health care: HL7 v3 (e.g. CDA r2), LOINC, SNOMED CT, ICD-10…
  – Clinical Research: CDISC ODM, CDASH, MedDRA…
EHR/CDMS approaches to integrating data

• **Single-source Concept**
  – All data are entered into an electronic source document (interface to the EHR or to the EDC system) and then flow – after appropriate contextual parsing - into the EHR and the EDC system respectively.
    • IHE integration profile “Retrieve Form for Data Capture” (RFD) + content profile “Clinical Research Data Capture”

• **Extraction and Investigator Verification**
  – All data are first entered into the EHR and then extracted, verified and transferred to the EDC system
    – RE-USE project approach

• **Direct Extraction from EHR**

*CDISC Electronic Source Data Interchange (eSDI) Group, Leveraging the CDISC Standards to Facilitate the use of Electronic Source Data within Clinical Trials Version, Nov 20. 2006, http://www.cdisc.org/stuff/contentmgr/files/0/2f6eca8f0df7caac5bbd4fadfd76d575/misc/docs/esdi.pdf*
Same need of standard-based integration profiles

- **Integrating the Healthcare Enterprise**
  - “Integration profiles”: guidelines for implementing "transactions" between different components of a distributed health information system, using established standards such as HL7 or DICOM.

- **“Connectathons”**
  - Rigorous testing process

> 200 participants, > 100 systems, > 50 vendors
Same need of semantic Interoperability services

- The semantic of the data collected by a given EHR or EDC system are almost always system-specific
  - Different **Models of Use** associated to different “interface terminologies”

- **Model of Meaning**
  - Serves as a **standardized representation** to which multiple Models of Use can be mapped
  - Used by both EHR and EDC system vendors as the basis for enabling cross-vendor semantic interoperability.
Need of semantic Interoperability services

• **Model of meaning**
  – Library of **agreed clinical data structure definitions** (e.g. HL7 templates)
    • Based on generic reference models for representing clinical data (e.g. ISO/HL7 RIM) and on standard data types (ISO 21090)
    • Explicitly bound to reference terminologies/ontologies (e.g. LOINC, SNOMED-CT, ICD-10) through value sets

• **Semantic interoperability framework** of the RE-USE project
  – Library of data elements bound to SNOMED 3.5 VF concepts
  – Pivot representation supporting the mapping process between data elements of the eCRFs & data elements already existing in the library of the EHR
RE-USE project

• AP-HP: Most important French University Hospital Organization with
  – 38 hospitals:
    • 1,000,000 hospitalized patients
    • 23,000 beds
    • 1500 day care
    • 850 home care capacity
  – 90,000 employees including 15,300 physicians
  – Georges Pompidou University Hospital (HEGP) (853 beds)
    • EHR: DxCare® (Medasys©)
• AP-HP: First research center about human beings in France
  – 354 active research projects
  – 35,000 enrolled patients
  – Sponsors: AP-HP, pharma industry, public institutions
  – EDC system: MARVIN® (XClinical ©)
RE-USE: use case

• Clinical study **Arcadia:**
  – Treatment modalities for **hypertension** caused by fibro-muscular dysplasia

• Approved by the HEGP Institutional Review Board for the use of the EHR as a source for clinical research data (mandatory in French law)

• **Integration** into DxCare® (EHR) of the 7 e-CRFs of the Arcadia study
  – Inclusion, initial evaluation, initial abdominal imaging, initial neurovascular imaging, biological tests, adverse events, and serious adverse events
RE-USE Architecture

• Implementation on an integration profile
• Synchronization process between EHR and eCRF
  – ISO21090 data types and terminology mapping service
• Integration profile to enable the sharing and exchange of clinical research data between EHR and CDMS
RE-USE design and implementation

Hospital

Template consumer

Clinical Research

Template source

Distribute Document Set

Document source

Document consumer

Provide & Register Document Set

Document Registry/repository

Query/Retrieve Document

XTS (Cross Enterprise Template Sharing)

XDM (X Document Media Interchange)

XDS (Cross Enterprise Document Sharing)

10/20/2011
(1) Cross enterprise template sharing

Patient care (HL7)

Hospital n 1

Template Consumer

Clinical Research (CDISC)

CRO n 1

Template source

(1) Meta data
CDISC ODM

Integrating Clinical Research to the Healthcare Enterprise: from the RE-USE project to the EHR4CR platform,
caBIG VCDE Virtual Journal Club
DxSync uses the semantic interoperability framework to synchronize the data elements of the eCRF with data elements already existing in the library of the EHR.
(3) Data capture & transfer to CDMS

Patient care (HL7)

Hospital n 1

(2)

Template Consumer

Document Source

Clinical Research (CDISC)

CRO n 1

(1)

Template source

Document Repository

(3) Clinical data (meta data + data)

CDISC ODM

Integrating Clinical Research to the Healthcare Enterprise: from the RE-USE project to the EHR4CR platform, caBIG VCDE Virtual Journal Club
Mapping of Interface Terminologies

• Method (in the study)
## Mappings

<table>
<thead>
<tr>
<th></th>
<th>Automatic mapping (%) (TP+FP)/n</th>
<th>Expert validated mapping (%) (TP+FN)/n</th>
<th>Precision (%) TP/(TP + FP)</th>
<th>Recall (%) TP/(TP+FN)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arcadia-IT / PatientCare-IT</td>
<td>15.9</td>
<td>13.4</td>
<td>70</td>
<td>84</td>
</tr>
<tr>
<td>Arcadia-IT/SNOMED 3.5</td>
<td>19.4</td>
<td>86.2</td>
<td>84</td>
<td>55</td>
</tr>
<tr>
<td>PatientCare-IT/SNOMED 3.5</td>
<td>32.4</td>
<td>65.3</td>
<td>62</td>
<td>31</td>
</tr>
</tbody>
</table>

Size of the sets: Arcadia-IT 232, PatientCare-IT 419
Prepopulation of the eCRF with EHR values

Current weight (78 kg)

Last known weight (78kg)
Limitations

• The results presented in this study are a **baseline**
  – **Limited linguistic resources** (e.g. no resource comparable to the UMLS in French)
    • Presence of acronyms, abbreviations...
  – Mapping based on **edit-distance** (distance between Strings)
    • Some features could be optimized marginally
  – **SNOMED 3.5** (SNOMED International very different from SNOMED CT)
Limitations

• Mapping process identifies **candidates** for pre-population of variables
  
  **Data quality**
  
  • Health Care: Clinician entering the data, Data Management
  • Clinical Research: Clinician entering the data, Research Technician control the data, Data Management, (if any problem is detected), Research Technician control again...

  **Data-type**
  
  • Clinical Research: binary question
  • Health Care: multiple choice question

  **Context** (precise semantic)
  
  • Heart pulse
  • Heart pulse 6 hours after receiving the treatment
From RE-USE to EHR4CR

From the “one hospital - one sponsor” proof of concept...
...
... to a cross boarder architecture for multi sites clinical trials with different sponsors = EHR4CR
Future (current) work

• The **EHR4CR project** ([http://www.ehr4cr.eu/](http://www.ehr4cr.eu/)) 2011-2014
  – **Platform and business model** for re-using EHR data for supporting medical research in Europe

• **3 different scenarios** across different therapeutic areas
  – Protocol feasibility
  – Patient recruitment
  – Data capture & adverse event reporting

• **32 participants**
  – 10 Pharmaceutical Companies (members of EFPIA)
  – 22 Public Partners (Academia, Hospitals and SMEs)

• **11 pilots** validating the solutions across 5 European countries (under different legal frameworks)

• **6 millions Euro** & in-kind participation of Pharma