Integrating Electronic Health Records and Electronic Case Report Forms: The RE-USE Project

Christel Daniel¹,², Naji El Fadly¹,³, Bastien Rance², Pierre-Yves Lastic⁴, Noël Lucas²

¹ INSERM UMRS 872, team 20, Paris Descartes University, Paris, France
² European George Pompidou Hospital, Paris, France
³ Medasys, France, ⁴ Sanofi-Aventis & CDISC Board of Directors, France

Presented by P-Y Lastic, PhD, (pierre-yves.lastic@sanofi-aventis.com)
Senior Director, Data Privacy & Healthcare Interoperability Standards at Sanofi-Aventis R&D and Chairman, European CDISC Coordinating Committee, at the CDISC Interchange Japan in Tokyo, July 21, 2010
Most important French University Hospital Organization with
- 38 hospitals: 1,000,000 hospitalized patients, 23,000 beds, 1500 day care and 850 home care capacity
- 69,000 employees including 15,300 physicians

- G. Pompidou University Hospital
  (HEGP: Hôpital Européen Georges Pompidou) (853 beds)

- EHR: DxCare® (Medasys©)
  -> Orbis (AGFA Healthcare©)
**AP-HP clinical research**

- First research center about human beings in Europe
  - 354 active research projects and 35,000 enrolled patients
  - Sponsors: AP-HP, pharma industry, public institutions
- Research structures
  - 18 federated research institutes, 7 Clinical Investigation Centers, 12 bio-medical research centers, 8 Clinical Research Unit
  - 112 INSERM (French National Institute for Health and Medical Research) teams, 30 CNRS (French National Institute for Scientific Research)

http://www.aphp.fr/site/recherche_innovations/presentation1.htm
http://www.aphp.fr/site/connaitre/chiffres_recherche.htm
Overview

- Why it is so hard to integrate patient care & clinical research
- Facing semantic interoperability issues
  - e.g IHE “Retrieve Form for Data capture” (RFD) integration profile
- The REUSE project at AP-HP
  - Single source concept: “Retrieve and Integrate Form for Data capture”
  - Mapping patient care & clinical research data elements
- Key messages
Integrating Patient Care & Clinical Research

- Patient recruitment
  - Only 7% of eligible patients enroll in a clinical trial
  - 86% of all trials fail to enroll on time
  - Women, minorities, children and special populations are under-represented

- Enhance spontaneous reporting of drug adverse events
  - Spontaneous Reporting System (SRS)

- Enhance data capture
  - EHR contains 30% to 50% of items of research forms


*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 hard?

- Different workflows
- Different ethical and juridical contexts (e.g. patient consent)
- Different information system quality requirements
- Different IT standards
  - Patient care
    - HL7 v3 information models (e.g. Clinical Document Architecture (CDA r2))
    - Reference terminologies (LOINC, SNOMED CT, ICD-10, etc)
  - Clinical research
    - Operational Data Model (ODM)
    - Clinical Data Acquisition Standards Harmonization (CDASH)
    - MedDRA (adverse events)

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

- Electronic Healthcare Record for Clinical Research (EHRCR)
  - US & Europe joint initiative (eClinical Forum, EuroRec, CDISC, HL7)
  - Profile passed HL7 Ballot in 2009
  - http://www.EHRCR.org

- EHR as source data system: Clinical research regulatory requirements for data collection, management, extraction, security
  - FDA mandated CDISC Electronic Source Data Interchange (eSDI) guidance published in 2005
  - European Medicines Agency guidance to be issued in 2010

- Biomedical Research Integrated Domain Group (BRIDG) Model
  - Information Model for Clinical Research (co-chair CDISC, HL7, NCI & FDA)
  - Used to set up the new NCI information System Architecture
  - Used to ensure consistency of CDISC standards development
  - Used as a Domain Analysis Model by the HL7 RCRIM Work Group
EHR for Clinical Research
Where are we?

From the “one hospital - one sponsor” proof of concept...

... to a cross boarder architecture for multi sites clinical trial from different sponsors
Different approaches …

● “Source data extraction”
  ● Some source data (patient care) of the EHR are exported to the clinical trial database (CDMS)
    ● IHE integration profile “Retrieve Form for Data Capture” (RFD) + content profile “Clinical Research data Capture”

● “Single source”
  ● All source data (patient care + research) are captured into the EHR and exported to the CDMS
    ● REUSE project

● “Single source used as CDMS”
  ● All source data (patient care + research) are captured into the EHR=CDMS
... the same need of semantic interoperability

Patient care (HL7)

Proprietary information models & “interface terminologies”

EHR

Clinical Research (CDISC)

CDMS

HL7
CDA r2
Reference terminologies (LOINC, SNOMED CT, ICD-10, etc)

CDISC
ODM
CDASH

HL7 RCRIM
IHE Quality, Research, and Public Health Profiles
Key Aspects of the HL7 CDA documents

- Derive their meaning from the HL7 v3 Reference Information Model (RIM) and use HL7 v3 Data Types
- Consist of a header and a body
  - Header provides information on patient, provider, encounter, and authentication
  - Body contains
    - narrative text / multimedia content (level 1)
    - optionally augmented by coded equivalents (levels 2 & 3) (Release 2 (2005))
- CDA documents are encoded in Extensible Markup Language (XML)
- Implementation Guides
  - ASTM/HL7 Continuity of Care Document (CCD)
  - IHE Patient Care Coordination Templates
Structure of CDA documents

Main Components of CDA

```xml
<ClinicalDocument>
  ...
  <StructuredBody>
    <Section>
      <text>...</text>
      <Observation>
        ...
        </Observation>
        <Observation>
        <Reference>
          <ExternalObservation>
            ...
            </ExternalObservation>
          </Reference>
        </Observation>
      </Section>
      <Section>...</Section>
    </StructuredBody>
  </ClinicalDocument>
```

© René Spronk
Semantic interoperability

From CDA r2 (CCD) to ODM

HL7 CDA r2

CDISC ODM-CDASH
ODM-CDASH example

Vital signs

- **Item definition including CDASH normalized data structures**

  ```xml
  <ItemDef DataType="integer" Length="3" Name="PA systolique" OID="ID.PA_SYSTOLIQUE">
    <Question>
      <TranslatedText xml:lang="fr">Systolique la plus elevee</TranslatedText>
    </Question>
    <MeasurementUnitRef MeasurementUnitOID="MU.MMHG" />
    <Alias Context="CDASH" Name="VSORRES"/>
  </ItemDef>
  ```

- **CDASH - Vital Signs – VS (Findings)**

  - VSTEST used to capture the test name
  - VSORRES, is used to capture the result.

- **Item group definition & item references**

  ```xml
  <ItemGroupDef Name="Bilan commun" OID="IG.BILAN_COMMUN" Repeating="No">
    <Description>
      <TranslatedText xml:lang="fr">Bilan commun</TranslatedText>
    </Description>
    <ItemRef ItemOID="ID.POIDS" Mandatory="Yes" />
    <ItemRef ItemOID="ID.TAILLE" Mandatory="Yes" />
    <ItemRef ItemOID="ID.PA_SYSTOLIQUE" Mandatory="Yes" />
  </ItemGroupDef>
  ```
Vital signs section contains coded measurement of a patient’s vital signs

```xml
<organizer classCode='CLUSTER' moodCode='EVN'>
  <templateld root='2.16.840.1.113883.10.20.1.32'/>
  <code code='46680005' displayName='Vital signs'
    codeSystem='2.16.840.1.113883.6.96' codeSystemName='SNOMED CT'/>
</organizer>

<observation classCode='OBS' moodCode='EVN'>
  <templateId root='1.3.6.1.4.1.19376.1.5.3.1.4.13.2'/>
</observation>
```

*observation* elements are using the Vital Signs Observation template.

```xml
<observation classCode='OBS' moodCode='EVN'>
  <templateId root='1.3.6.1.4.1.19376.1.5.3.1.4.13'/>
  <code code='8480-6' displayName='Intravascular systolic'
    codeSystem='2.16.840.1.113883.6.1' codeSystemName='LOINC'/>
  <effectiveTime value= ''/>
  <value xsi:type='PQ' value= '' unit=' mm[Hg] '/>
</observation>
```
## IHE PCC (CDA r2) example

### Vital signs

<table>
<thead>
<tr>
<th>LOINC</th>
<th>Description</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>9279-1</td>
<td>RESPIRATION RATE</td>
<td>/min</td>
</tr>
<tr>
<td>8867-4</td>
<td>HEART BEAT</td>
<td></td>
</tr>
<tr>
<td>2710-2</td>
<td>OXYGEN SATURATION</td>
<td>%</td>
</tr>
<tr>
<td>8480-6</td>
<td>INTRAVASCULAR SYSTOLIC</td>
<td>mm[Hg]</td>
</tr>
<tr>
<td>8462-4</td>
<td>INTRAVASCULAR DIASTOLIC</td>
<td></td>
</tr>
<tr>
<td>8310-5</td>
<td>BODY TEMPERATURE</td>
<td>Cel or [degF]</td>
</tr>
<tr>
<td>8302-2</td>
<td>BODY HEIGHT (MEASURED)</td>
<td></td>
</tr>
<tr>
<td>8306-3</td>
<td>BODY HEIGHT^LYING</td>
<td>m, cm,[in_us] or [in_uk]</td>
</tr>
<tr>
<td>8287-5</td>
<td>CIRCUMFERENCE.OCCIPITAL-FRONTAL (TAPE MEASURE)</td>
<td></td>
</tr>
<tr>
<td>3141-9</td>
<td>BODY WEIGHT (MEASURED)</td>
<td>kg, g, [lb_av] or [oz_av]</td>
</tr>
</tbody>
</table>
“Retrieve Form for Data Capture (RFD)” - Content profile “Clinical Research data Capture”

- Actor Form Filler
- Form Filler
- Form Manager
- CDMS
- EHR
- Reference Implementation CCD Option
- Form Manager
- CRD CCD
- XSLT
- ODM/CDASH
<table>
<thead>
<tr>
<th>CDASH Domain</th>
<th>Clinical Database Variable Name</th>
<th>Optionality</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demography</td>
<td>BRTHYR</td>
<td>R</td>
<td>Year of subject’s birth.</td>
</tr>
<tr>
<td></td>
<td>BRTHMO</td>
<td>R</td>
<td>Month of subject’s birth.</td>
</tr>
<tr>
<td></td>
<td>SEX</td>
<td>R</td>
<td>The assemblage of physical properties or qualities by which male is distinguished from female; the physical difference between male and female; the distinguishing peculiarity of male or female. (NCI – CDISC Definition).</td>
</tr>
<tr>
<td>Medical History</td>
<td>MHTERM</td>
<td>R</td>
<td>Verbatim or preprinted CRF term for the medical condition or event.</td>
</tr>
<tr>
<td></td>
<td>MHONG</td>
<td>R</td>
<td>Identifies the end of the event as being ONGOING or RESOLVED.</td>
</tr>
<tr>
<td></td>
<td>CMTRT</td>
<td>R</td>
<td>Verbatim drug name that is either pre-printed or collected on a CRF.</td>
</tr>
<tr>
<td></td>
<td>CMSTDTC</td>
<td>R</td>
<td>Date when the medication was first taken.</td>
</tr>
<tr>
<td>Substance Use</td>
<td>SUTRT</td>
<td>R</td>
<td>The type of substance (e.g., TOBACCO, ALCOHOL, CAFFEINE, etc. Or CIGARETTES, CIGARS, COFFEE, etc.).</td>
</tr>
<tr>
<td></td>
<td>VSORRES</td>
<td>R</td>
<td>Result of the vital signs measurement as originally received or collected.</td>
</tr>
<tr>
<td></td>
<td>VSORRESU</td>
<td>R</td>
<td>Original units in which the data were collected.</td>
</tr>
<tr>
<td>Lab Test Results</td>
<td>LBDTC</td>
<td>R</td>
<td>Date of sample collection.</td>
</tr>
<tr>
<td></td>
<td>LBORRES</td>
<td>R</td>
<td>Result of the measurement or finding as originally received or collected.</td>
</tr>
<tr>
<td></td>
<td>LBORRESU</td>
<td>R</td>
<td>Original units in which the data were collected.</td>
</tr>
<tr>
<td></td>
<td>LBTEST</td>
<td>R</td>
<td>Verbatim name of the test or examination used to obtain the measurement or finding. Note: any test normally performed by a clinical laboratory is considered a lab test.</td>
</tr>
<tr>
<td>ECG Test Results</td>
<td>LBDTC</td>
<td>R</td>
<td>Date of sample collection.</td>
</tr>
<tr>
<td></td>
<td>EGTESTCD And/or EGTEST</td>
<td>R</td>
<td>Verbatim name of the test or examination used to obtain the measurement or finding.</td>
</tr>
<tr>
<td></td>
<td>EGORRES</td>
<td>R</td>
<td>Result of the measurement or finding as originally received or collected.</td>
</tr>
<tr>
<td></td>
<td>EGTEST</td>
<td>R</td>
<td>Verbatim name of the test or examination used to obtain the measurement or finding.</td>
</tr>
</tbody>
</table>
## Option CCD - Required CCD data elements

<table>
<thead>
<tr>
<th>Data Element</th>
<th>Optionality</th>
<th>Template ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date of Birth</td>
<td>R</td>
<td>patientRole/patient/birthTime</td>
</tr>
<tr>
<td>Gender</td>
<td>R</td>
<td>patientRole/patient/administrativeGenderCode</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>O</td>
<td>patientRole/patient/ethnicGroupCode</td>
</tr>
<tr>
<td>Race</td>
<td>R2</td>
<td>patientRole/patient/raceCode</td>
</tr>
<tr>
<td>Active Problems</td>
<td>R</td>
<td>1.3.6.1.4.1.19376.1.5.3.1.3.6</td>
</tr>
<tr>
<td>Past Medical History</td>
<td>R2</td>
<td>1.3.6.1.4.1.19376.1.5.3.1.3.8</td>
</tr>
<tr>
<td>Procedures and Interventions</td>
<td>R2</td>
<td>1.3.6.1.4.1.19376.1.5.3.1.1.13.2.11</td>
</tr>
<tr>
<td>Social History</td>
<td>R2</td>
<td>1.3.6.1.4.1.19376.1.5.3.1.3.16</td>
</tr>
<tr>
<td>Current Medications</td>
<td>R</td>
<td>1.3.6.1.4.1.19376.1.5.3.1.3.19</td>
</tr>
<tr>
<td>Vital Signs</td>
<td>R2</td>
<td>1.3.6.1.4.1.19376.1.5.3.1.1.15.3.2</td>
</tr>
<tr>
<td>Physical Exam</td>
<td>R2</td>
<td>1.3.6.1.4.1.19376.1.5.3.1.1.9.15</td>
</tr>
<tr>
<td>Allergies and Other Adverse Reactions</td>
<td>R</td>
<td>1.3.6.1.4.1.19376.1.5.3.1.3.13</td>
</tr>
<tr>
<td>Coded Results</td>
<td>R2</td>
<td>1.3.6.1.4.1.19376.1.5.3.1.3.28</td>
</tr>
</tbody>
</table>
# Mapping CCD -> CDASH

<table>
<thead>
<tr>
<th>Clinical Research Data Capture Data Elements</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CDASH Domains</strong></td>
<td><strong>CCD Reference</strong></td>
</tr>
<tr>
<td>Demography</td>
<td>CCD Header Information</td>
</tr>
<tr>
<td>Medical History</td>
<td>Active Problems, Past Medical History, and Procedures and Interventions</td>
</tr>
<tr>
<td>Concomitant Medication</td>
<td>Current Medications</td>
</tr>
<tr>
<td>Substance Use</td>
<td>Social History</td>
</tr>
<tr>
<td>Vital Signs</td>
<td>Vital Signs</td>
</tr>
<tr>
<td>Physical Exam</td>
<td>Physical Exam</td>
</tr>
<tr>
<td>Adverse Events</td>
<td>Allergies</td>
</tr>
<tr>
<td>Lab Test Results</td>
<td>Coded Results</td>
</tr>
<tr>
<td>ECG Test Results</td>
<td>Coded Results</td>
</tr>
</tbody>
</table>
REUSE: “Retrieve and Integrate Forms for Data capture”

Patient care (HL7)

Hospital n° 1

DxCare

Hospital n° 2

Orbis

Hospital n° N

Orbis

Send metadata (eCRF) (ODM transaction)

Clinical Research (CDISC)

CRO n° 1

Clinical trial n° 1

XClinical

CRO n° 2

Clinical trial n° 2

OpenClinica

Send metadata+data (CDISC ODM transaction)
Clinical research
- 20 clinical trials conducted between 2001 and 2009 by the HEGP Clinical Research Unit (CRU)
- 3739 data elements including 2669 unique data elements

Patient care
- 861 forms created between 2000 and 2009
- 6587 data elements
Normalization

- Problem: Synonymy, homonymy, typos...
- Normalization
  - Elimination of accents, change to lower case
  - Elimination of “empty words” (e.g. and, or...)
  - Simplification (singular/plural, male/female)
  - Classification of words in alphanumerical order

Mesure ambulatoire de la pression artérielle => ambulatoir;arteriel,mesur;pression
Similarity between terms & expressions

- Levenshtein distance

  \[ \text{Hyper}\star\text{tension}\star \]
  \[ \| \| \]
  \[ \text{Hyper-tensions} \]

- 3 classes
  - Exact match \( \text{Lev}(v1, v2) = 1 \)
  - Close match \( 0.9 \geq \text{Lev}(v1, v2) > 1 \)
  - Possibly close match \( 0.85 \geq \text{Lev}(v1, v2) > 0.9 \)
Mapping patient care & clinical research data element - Results

- Preliminary mapping
- Among 3739 clinical research data elements
  - 648 exact matches (17%)
  - 56 close matches
  - 81 possibly close matches (50 % false positive)
- Need of pivot terminology
  - Reference terminology: e.g SNOMED CT
Summary

- **REUSE project**
  - Single source: “Retrieve and Integrate Forms for Data capture”
  - Increase secondary use of clinical data
- **Semantic interoperability framework**
  - Pivot terminology (e.g. SNOMED CT) & model (CDA r2)
  - CDA r2 to ODM mediator
Discussion - Conclusions

- **REUSE project**
  - Integrated solution for both clinicians and researchers between EHR and CDMS
  - Differs from IHE RFD integration profile since EHR is the single source of data

- **Benefits**
  - For healthcare providers: avoiding double data entry
  - For patients: in our approach, considering the EHR as the single source of data may improve patient safety as ALL data collected in a biomedical research setting is kept in the EHR, which is currently not the case for most research studies

- **Next steps**
  - Further work on semantic interoperability
  - Tests with other Clinical Research software
  - International tests: collaboration with other European hospitals within the EHR4CR project (Innovative Medicine Initiative of the European Commission and the EU Pharma Industry)
Strength through collaboration.

www.cdisc.org