FH | JOANNEUM
University of Applied Sciences

The use of electronic Health Records in Clinical Research - The value of CDISC Standards

FH-Prof. Dr. Jozef Aerts
University of Applied Sciences FH Joanneum
Graz, Austria

FH | JOANNEUM
University of Applied Sciences

Who is Jozef Aerts?

- CDISC volunteer since 2001
- Contributed to several CDISC standards
  - ODM, Define-XML, SDM-XML, CTR-XML, …
- Professor in Medical Informatics since 2012 in Graz, Austria

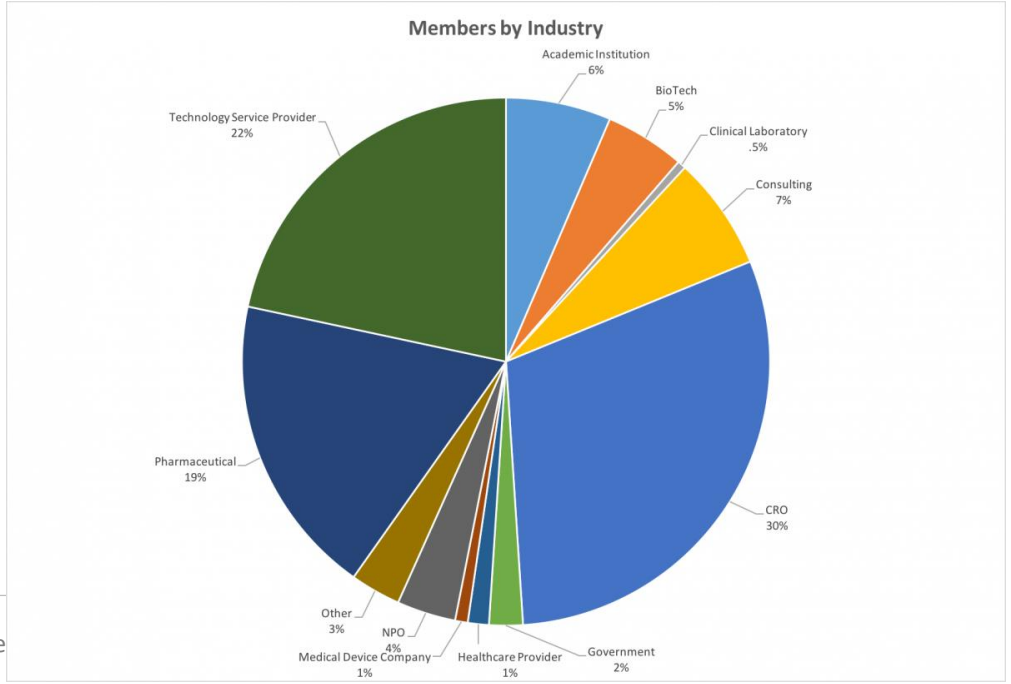
Who is Jozef Aerts?

- Owner and CEO of XML4Pharma

- A consultancy and software company specializing in the implementation of CDISC standards
  - With emphasis on XML technologies

CDISC

- Clinical Data Interchange Standards Consortium
- Consortium structure: >420 members (Japan: >100)
  - Pharma companies
  - CROs
  - Software vendors
  - Academic institutions
  - Consultants
  - Laboratories
  - …
- Volunteer organization
  - Consensus organization



https://www.cdisc.org/membership

Jozef Aerts and CDISC

- Just 1 of the several thousand volunteers
  - But a very active and critical one …
- Started volunteering in 2001
- Member of different CDISC development teams
  - ODM (Operational Data Model)
  - Define-XML (regulatory metadata submission standard)
  - SDM-XML (Study Design in XML)
  - CTR-XML (Clinical Trial Registries standard)
- Integration of healthcare data

CDISC as an SDO

- CDISC is a "standardization organisation" (SDO)
- We create syntactic and semantic standards for use in clinical research
- These are often different than the standards used in healthcare
  - Healthcare is "event" driven
  - Clinical research is "protocol" driven

Use of healthcare data in clinical research

- Healthcare and clinical research data systems are traditionally different and separated

- Same data needs to be entered twice
- Historical data needs to be copy-paste
  - Medical history
  - Prior medications

Use of healthcare data in clinical research

- Nowadays, much more healthcare data is available in electronic form, and we should use it in clinical research

- So we need to automate the use of healthcare data in clinical research

- There are however obstacles:
  - Technical
  - Different semantics
  - Different regulations

Sources of healthcare data for use in clinical research

- Directly from the (own) Hospital Information System
  - Limited to "hospital owned" data
  - Requires custom programming
- From HL7-v2 messages (within or between hospitals)
- From HL7 CDA documents (Electronic Health Records - EHRs)
- From HL7 FHIR documents / messages

Use of healthcare data in clinical research

- Reduces duplication of work
- Data quality can be a problem
  - But this can also be the case in CRF data
- Possibility of retrieval of historical data
  - Prior diseases (MH)
  - Prior medications, therapies and procedures
- Post-marketing data (ADR) - Mihari project
- Do countries already have a life-long EHR?

MIHARI
Medical Information for Risk Assessment Initiative

Use of EHRs in clinical research - Regulations

- ICH (International Commitee of Harmonisation)
  - E.g. audit trails (USA: 21 CRF Part 11)

> 4.9.3 Any change or correction to a CRF should be dated, initialed, and explained (if necessary) and should not obscure the original entry (i.e. an audit trail should be maintained); this applies to both written and electronic changes or corrections (see 5.18.4 (n)). Sponsors should provide guidance to investigators and/or the investigators' designated representatives on making such corrections. Sponsors should have written procedures to assure that changes or corrections in CRFs made by sponsor's designated representatives are documented, are necessary, and are endorsed by the investigator. The investigator should retain records of the changes and corrections.

https://www.pmda.go.jp/files/000156725.pdf

- FDA: "Electronic Source Data in Clinical Investigations" https://www.fda.gov/downloads/drugs/guidances/ucm328691.pdf

The value of CDISC standards

- For new medications or therapies, subject data must be submitted to the PMDA in a standardized way
- The CDISC SDTM (Submission Data Tabulation Model) provides the structure and contents of these data
- The data are not (and cannot) be captured in this way
  - SDTM contains a good number of derived data
  - Data needs to be transformed, categorized
  - Not every data point must be submitted

SDTM Dataset Example: Medical History (MH)

	STUDYID	DOMAIN	USUBJID	MHSEQ	MHSPID	MHTERM	MHLLT	MHDECOD	MHHLT	MH
1	CDISCPILOT01	MH	01-701-1015	9		ALZHEIMER'S DISEASE				
2	CDISCPILOT01	MH	01-701-1015	1	E01	VERBATIM_0135	PALPITATIONS	PALPITATIONS	HLT_0493	HLG
3	CDISCPILOT01	MH	01-701-1015	8		VERBATIM_0140	SUBTOTAL HYS...	HYSTERECTOMY	HLT_0675	HLG
4	CDISCPILOT01	MH	01-701-1015	2	E03	VERBATIM_0301	HEADACHE	HEADACHE	HLT_0064	HLG
5	CDISCPILOT01	MH	01-701-1015	10		VERBATIM_0539	GALLBLADDER ...	CHOLELITHIASIS	HLT_0084	HLG
6	CDISCPILOT01	MH	01-701-1015	3	E05	VERBATIM_0825	TINNITUS	TINNITUS	HLT_0570	HLG
7	CDISCPILOT01	MH	01-701-1015	4	E02	VERBATIM_0841	HEARTBURN	DYSPEPSIA	HLT_0244	HLG
8	CDISCPILOT01	MH	01-701-1015	6		VERBATIM_1004	THYROIDECTO...	THYROIDECTO...	HLT_0243	HLG
9	CDISCPILOT01	MH	01-701-1015	11		VERBATIM_1230	SORE THROAT	PHARYNGOLAR...	HLT_0130	HLG
10	CDISCPILOT01	MH	01-701-1015	7		VERBATIM_1716	TONSILLECTOMY	TONSILLECTOMY	HLT_0202	HLG
11	CDISCPILOT01	MH	01-701-1015	5	E04	VERBATIM_1779	NUMBNESS IN ...	HYPOAESTHESIA	HLT_0014	HLG
12	CDISCPILOT01	MH	01-701-1023	16		ALZHEIMER'S DISEASE				
13	CDISCPILOT01	MH	01-701-1023	21		VERBATIM_0073	CELLULITIS OF ...	CELLULITIS	HLT_0613	HLG
14	CDISCPILOT01	MH	01-701-1023	1	E02	VERBATIM_0249	GASTROESOPH...	GASTROOESOP...	HLT_0407	HLG
15	CDISCPILOT01	MH	01-701-1023	19		VERBATIM_0291	LOW BACK PAIN	BACK PAIN	HLT_0115	HLG
16	CDISCPILOT01	MH	01-701-1023	15		VERBATIM_0342	HYPERTENSION	HYPERTENSION	HLT_0228	HLG
17	CDISCPILOT01	MH	01-701-1023	17		VERBATIM_0388	PENILE PROST...	PENILE PROST...	HLT_0448	HLG
18	CDISCPILOT01	MH	01-701-1023	2	E06	VERBATIM_0461	MUSCULAR PAIN	MYALGIA	HLT_0700	HLG
19	CDISCPILOT01	MH	01-701-1023	12		VERBATIM_0596	NUMBNESS OF ...	HYPOAESTHESIA	HLT_0014	HLG
20	CDISCPILOT01	MH	01-701-1023	3	E03	VERBATIM_0699	HIATAL HERNIA	HIATUS HERNIA	HLT_0159	HLG
21	CDISCPILOT01	MH	01-701-1023	8		VERBATIM_0788	APPENDICITIS	APPENDICITIS	HLT_0703	HLG
22	CDISCPILOT01	MH	01-701-1023	4	E04	VERBATIM_0800	INDIGESTION	DYSPEPSIA	HLT_0244	HLG
23	CDISCPILOT01	MH	01-701-1023	5	E05	VERBATIM_0929	HEADACHE	HEADACHE	HLT_0064	HLG

A few important questions

- How can clinical research data be captured so that it can easily and smoothly be used for the submission?

- How can healthcare data be used in such a way that it can easily and smoothly be used for the submission?

Answer:
CDISC standards from the start

A simple example - Adverse event

- Adverse event severity (AESEV)

Your protocol / CRF:
1 - very mild, unrelated
2 - mild
3 - mild to severe, no action required
4 - severe, apply concommitant medication
5 - hospitalization required
6 - fatal

PMDA expected categories - (CDISC AESEV)
- Mild
- Moderate
- Severe

PMDA expected categories - (CDISC AEOUT)
- Recovered/Resolved
- …
- Fatal

CDISC Controlled Terminology:
https://www.cdisc.org/standards/semantics/terminology

PMDA expectation (SDTM-AE)
AEHOSP = Y or AEHOSP= N

Solution: use CDISC Standards from the start

- CDASH: Clinical Data Acquisition Standards Harmonization
- Defines 20 standardized CRFs
  - Standardized content
  - Standardized (CDISC) controlled terminology
- Makes it very easy to transform captured data to submission data
- Can easily be implemented in any study design tool (XML representation available)

https://www.cdisc.org/standards/foundational/cdash

CDASH example: AE form with annotations

Were any adverse events experienced? NOT SUBMITTED AEYN	O Yes O No
Adverse Event Category: *Defaulted* AECAT	**Sponsor Defined**
Adverse Event Subcategory: *Defaulted* AESCAT	**Sponsor Defined**
What is the adverse event identifier? AESPID	_____
What is the adverse event term? AETERM	_____
Start Date AESTDTC AESTDAT	_ _ / _ _ _ / _ _ _ _
Is the adverse event ongoing? AEENRF / AEENRTPT AEENTPT AEONGO	O Yes O No
End Date AEENDTC AEENDAT	_ _ / _ _ _ / _ _ _ _
What is the severity of the adverse event? AESEV	O MILD O MODERATE O SEVERE
Was the adverse event serious? AESER	O Yes O No

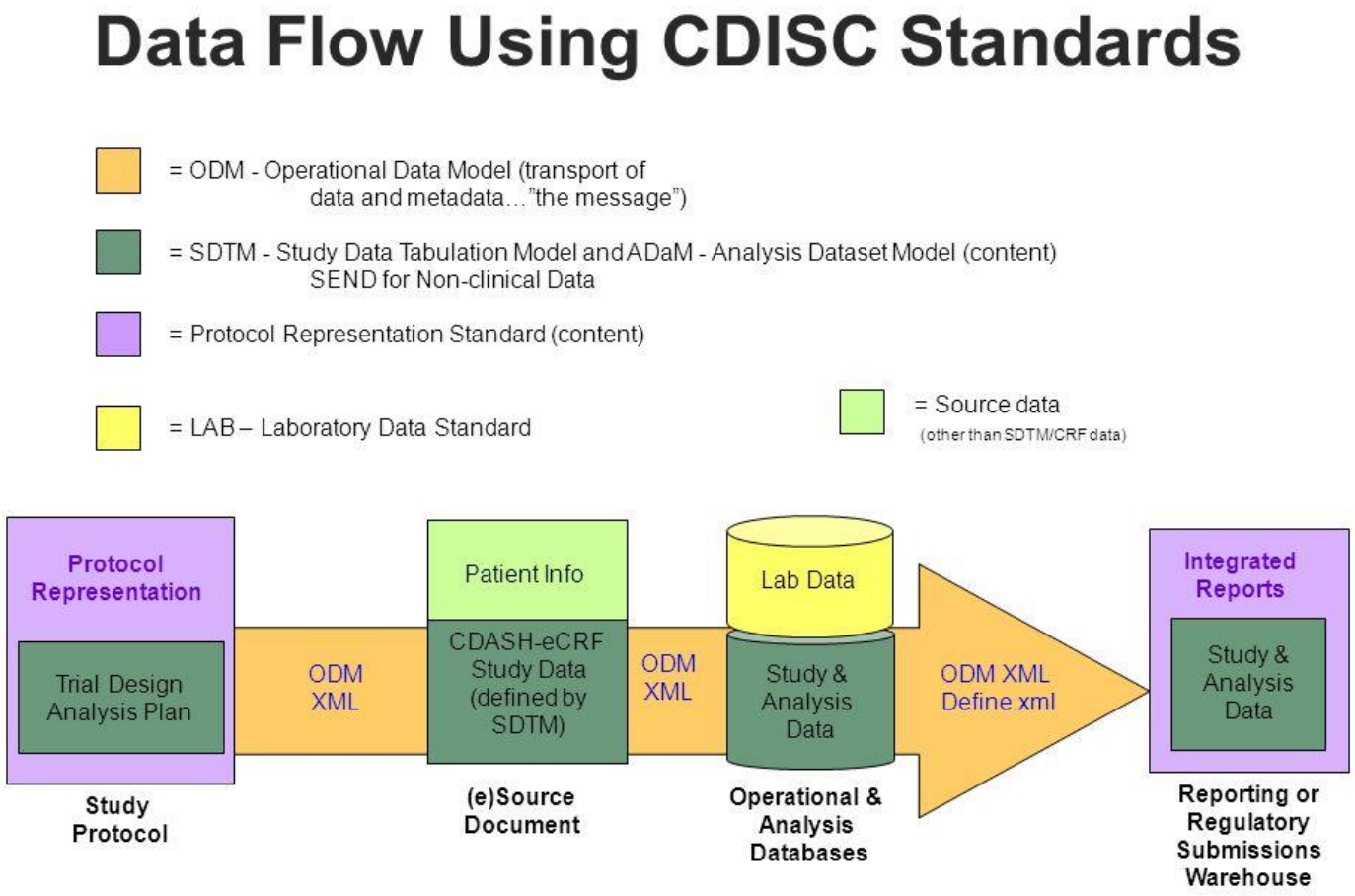
How do we exchange the data between systems?

- CDISC ODM (Operational Data Model) standard

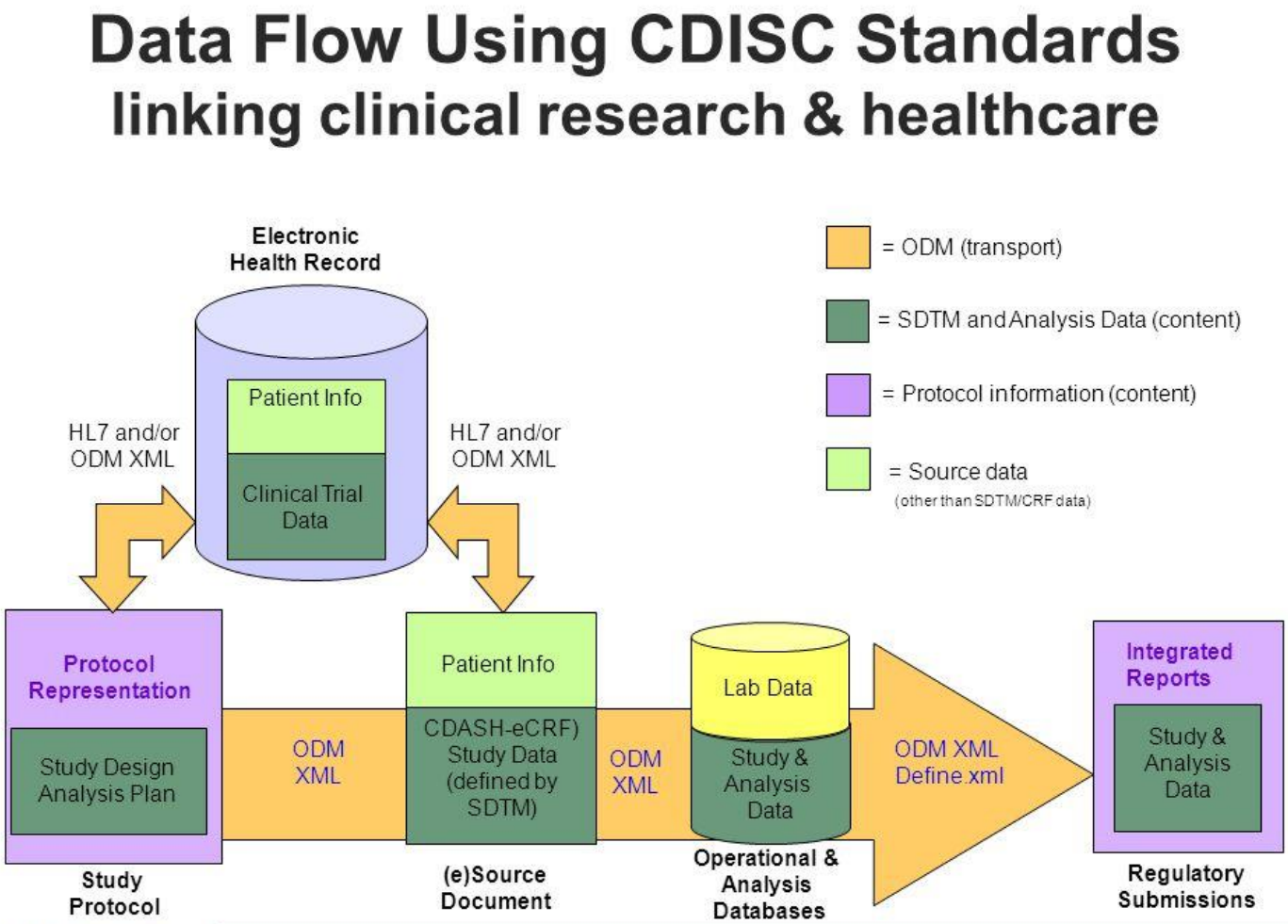
- XML-based standard

- Metadata (study design)
- AdminData
- ReferenceData (non-subject data)
- ClinicalData
  - Possibility for full audit trails

Clinical Research Data Flow using CDISC-ODM



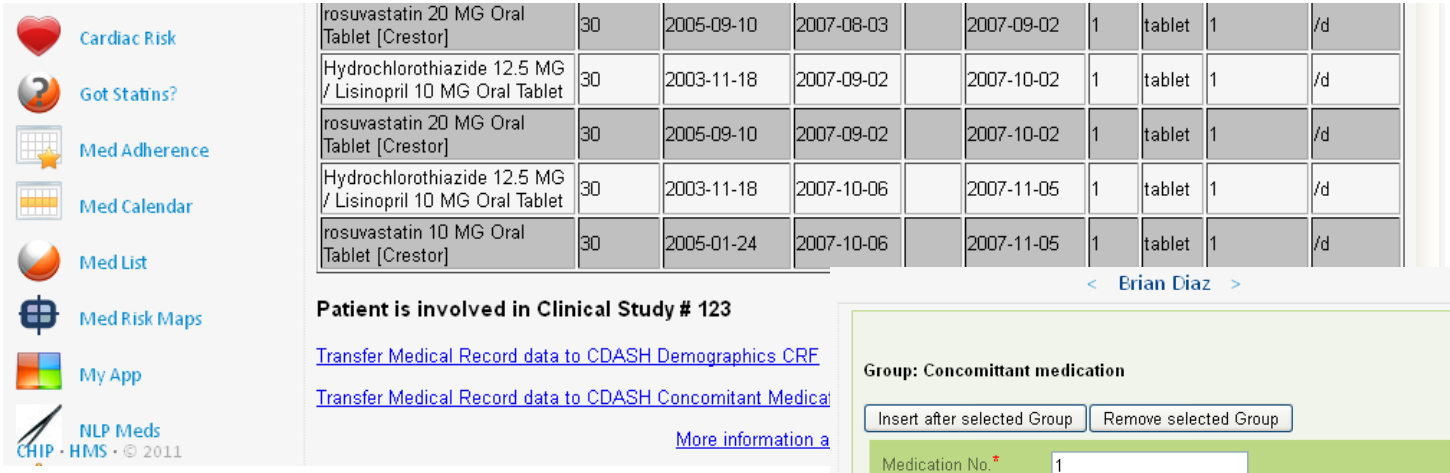
Healthcare Data Flow in Clinical Research - CDISC ODM



FH JOANNEUM
University of Applied Sciences

Healthcare data into Clinical Research data
Data from Hospital Information Systems

- Example: Smart Challenge (2010)

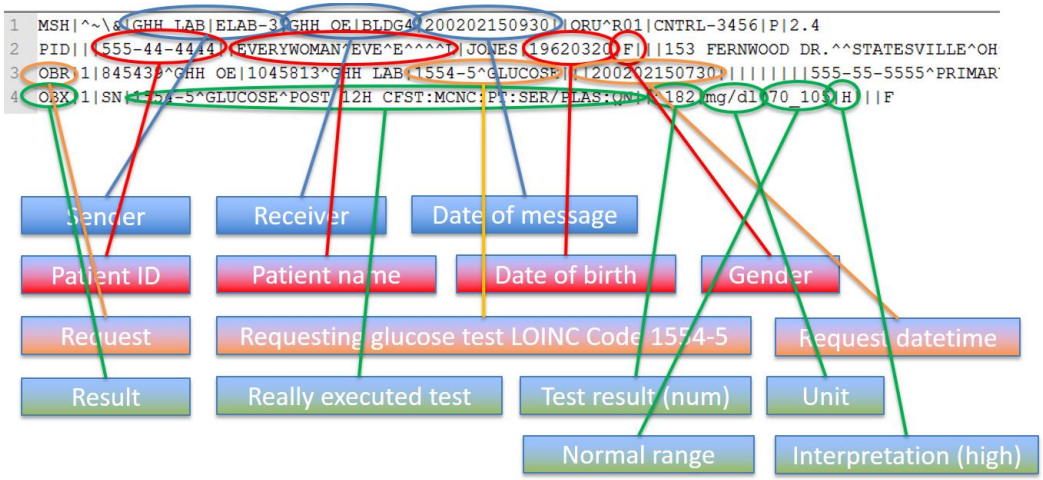


http://www.xml4pharma.com/Smart_Challenge/

Requires custom programming versus HIS-API

Healthcare data to research data from HL7-v2 messages

- Is very well feasible
- Not often done, as:
  - Messages are usually not kept in the HIS (not persistent) - they flow into the core HIS (extract)
    Exception: SS-MIX HL7-v2.5 (Japan)
  - Most of the data from messages never appears in the discharge documents or transfer-of-care document



Clinical Research Data from HL7-CDA documents

- HL7-CDA (XML) is in more and more countries used for:
  - Discharge documents
  - Transfer-of-care documents
  - Laboratory and radiology reports
  - E-Medication

- However:
  - CDAs usually contain an extract of the collected data (incomplete information)
  - Use of the information for clinical research heavily depends on which (semantic) coding systems are used

Clinical Research Data from HL7-CDA documents

Some history

- 2006: CDISC starts "Healthcare Link" (Landen Bain)

  "*CDISC, which is spearheading the demonstration project, views data capture as a 'single-source encounter between an investigator/physician and subject/patient,' says Landen Bain, CDISC liaison to healthcare. The 'key idea' is to have a data-capture form appear on the computer screen in an EHR session. 'If you have an EHR, it's contorted to go to a different computer and a different application to enter data for a clinical trial when the data originates inside the EHR and that's where you do [most of] your work.' "*

Clinical Research Data from HL7-CDA documents
Technical framework

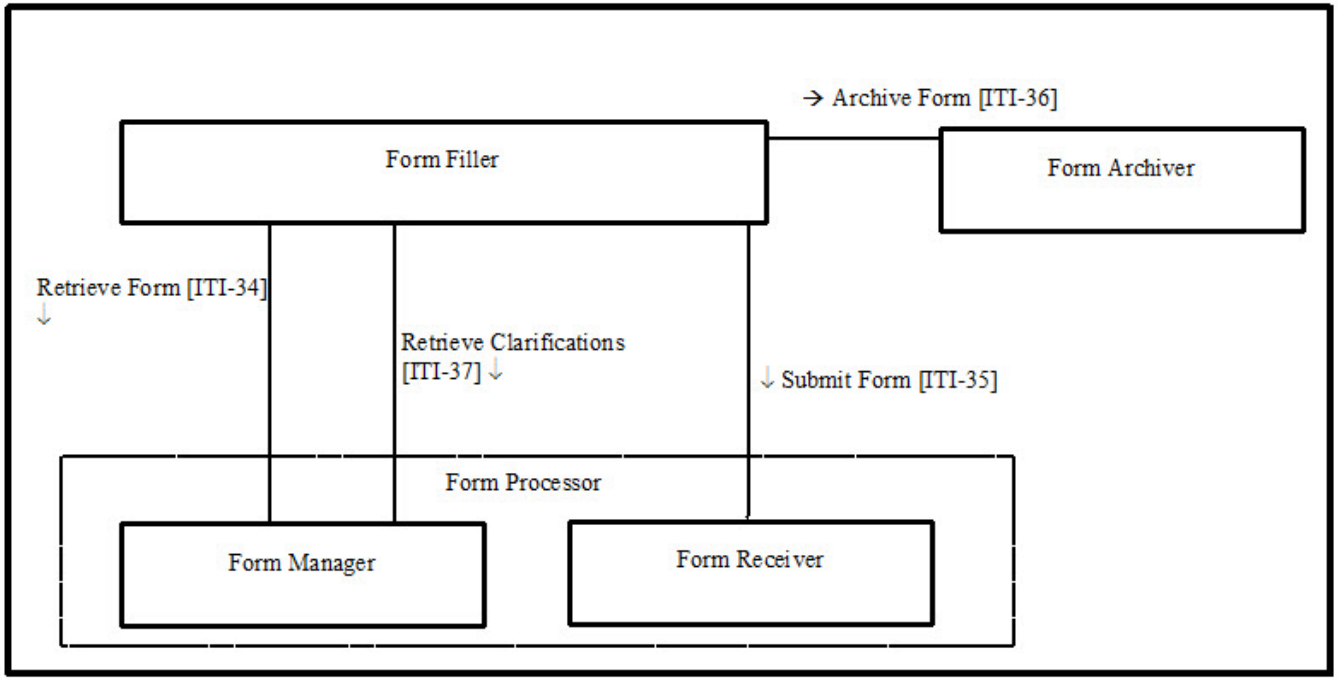
- Another healtcare consortium, IHE (Integrating the Healtcare Enterprise) defines (vendor-neutral) standardized process flows for use in healthcare

- Provides the technical framework for such exchanges

- These are named "IHE Profiles"
  "Retrieve Form for Data Capture" (RFD)

IHE Profiles for use of healthcare data in clinical research

- RFD: Retrieve Form for Data Capture
- CRD: Clinical Research Document
- Both use existing standards and technologies (XML, SOAP, …)



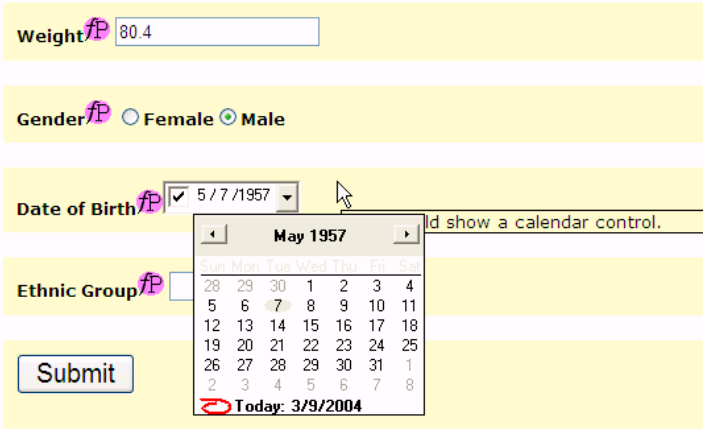
Implementations of RFD and CRD

- Concentrating on HL7-CDA on the healthcare data side
  - Especially CCD (Continuity of Care Document - USA)
- Works very well with CDASH forms
- Flow to EDC works best with CDISC ODM

Own work in the field (2005-2010)

- Demonstrate feasability of the concept

- Using following technologies & standards:
  - CDISC CDASH
  - W3C XForms
  - CDISC ODM



- Currently being used by several vendors (OpenClinica, ClinPal, …)

Problems with RFD / CRD

- Semantic mapping is the problem
- Works well for US CCD standard
- Must be adapted when other codes / coding systems are used
- CDISC controlled terminology is often not granular enough

Problems with RFD / CRD

- CCD ("Continuity of Care Document") is a "transfer of care" document
- Such documents usually only contain a limited amount of (the most relevant) information
- Much information that is in the Hospital Information System never appears in the CCD

Europe: The EHR4CR Project

- EU-funded programm 2011-2015
- 16 Million € (>2 Billion ¥)
- 35 academic and private partners (10 pharmaceutical companies)
- Goal: to build a platform to use de-identified data from hospital EHR systems

- Result: real life implementation: InSite platform
  - Used by several sponsors and (university) hospitals

EHR4CR: Participants

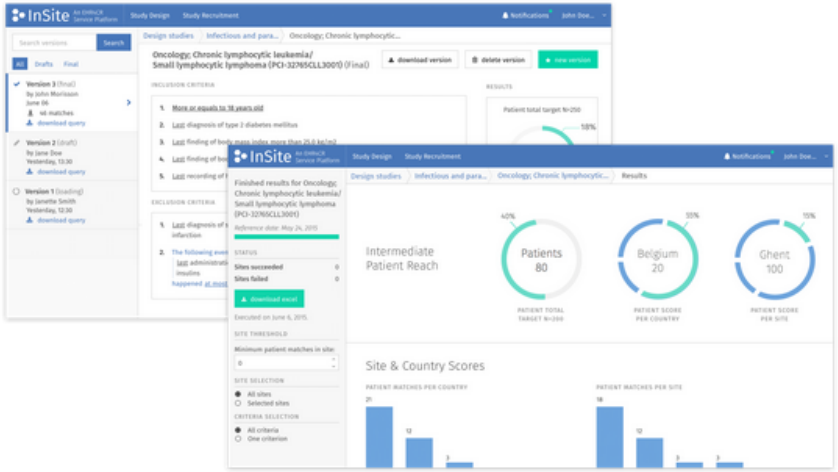
http://www.ehr4cr.eu/9april2014/presentations/EHR4CR%20-%20April%209%20-%20Sundgren.pdf

EHR4CR: main topics

- Protocol feasibility platform
  - Will it be hard to find subjects for this study?

- Patient recruitment
  - Site tools to:
    - Identify potential candidates (based on IE criteria)
    - Contact treating physicians
    - Evaluate and recruit patients

EHR4CR: The InSite platform



Deploying All Over Europe in 2016-2017

InSite is a private initiative determined to define and realise a **sustainable eco-system for Real World Data driven clinical trial design and execution**. We build upon the outcome of several international research projects aimed at accelerating research through re-use of care data (EHR, primary care, Personal Health Records, ...) and the collaboration with pharmaceutical industry through the EHR4CR - Electronic Health Records for Clinical Research project.

Supported by a number of pharmaceutical companies, CROs and hospitals we are shaping a trustworthy environment for clinical data re-use from which all stakeholders — industry, care providers, patients — benefit.

https://www.insiteplatform.com/

- Although (partially) paid by tax money, not open

Use of EHRs for clinical research in Europe

- Healthcare in Europe is hopelesly divided
- Each country has ist own healthcare system, laws, …
  - Sometimes even huge differences between provinces
- The only common denominator is the EU data protection regulation

Use of EHRs for clinical research in Austria

- Austria has a national EHR exchange system (ELGA)

- ELGA system manages interoperable health documents (CDA) between healthcare providers / patients
  - Discharge letters
  - Laboratory / radiology reports
  - E-Medication

- Use of ELGA documents for clinical research is NOT allowed

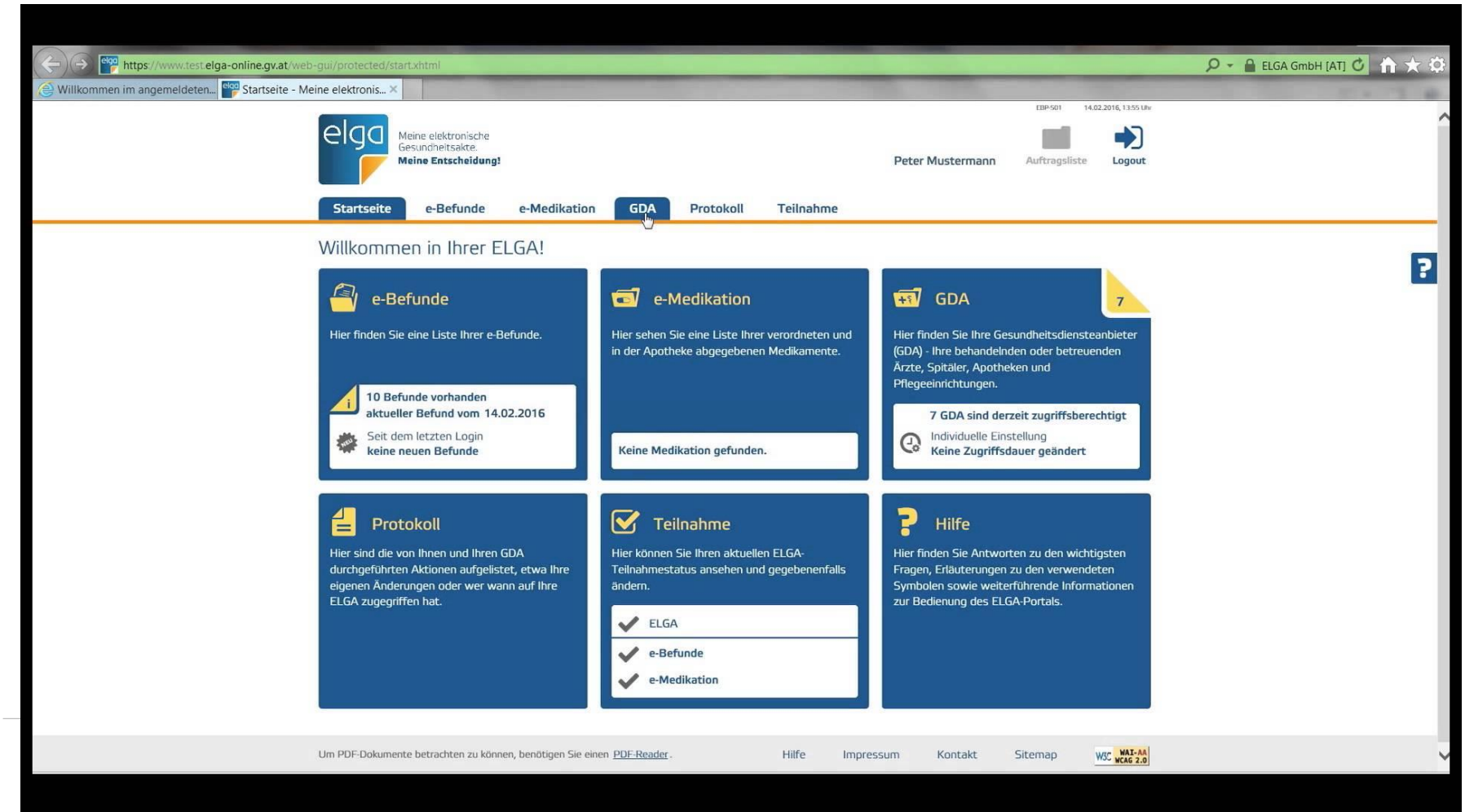
The ELGA-EHR System in Austria

- Technically, based on HL7-CDA documents
- Exchange over IHE-XDS profile
  (Cross-Enterprise Document Sharing)
- Used for "continuity of care"
- E-Medication currently rolled out

FH JOANNEUM
University of Applied Sciences

The ELGA-EHR System in Austria - Patient Portal

- Login through "e-card" or mobile phone TAN

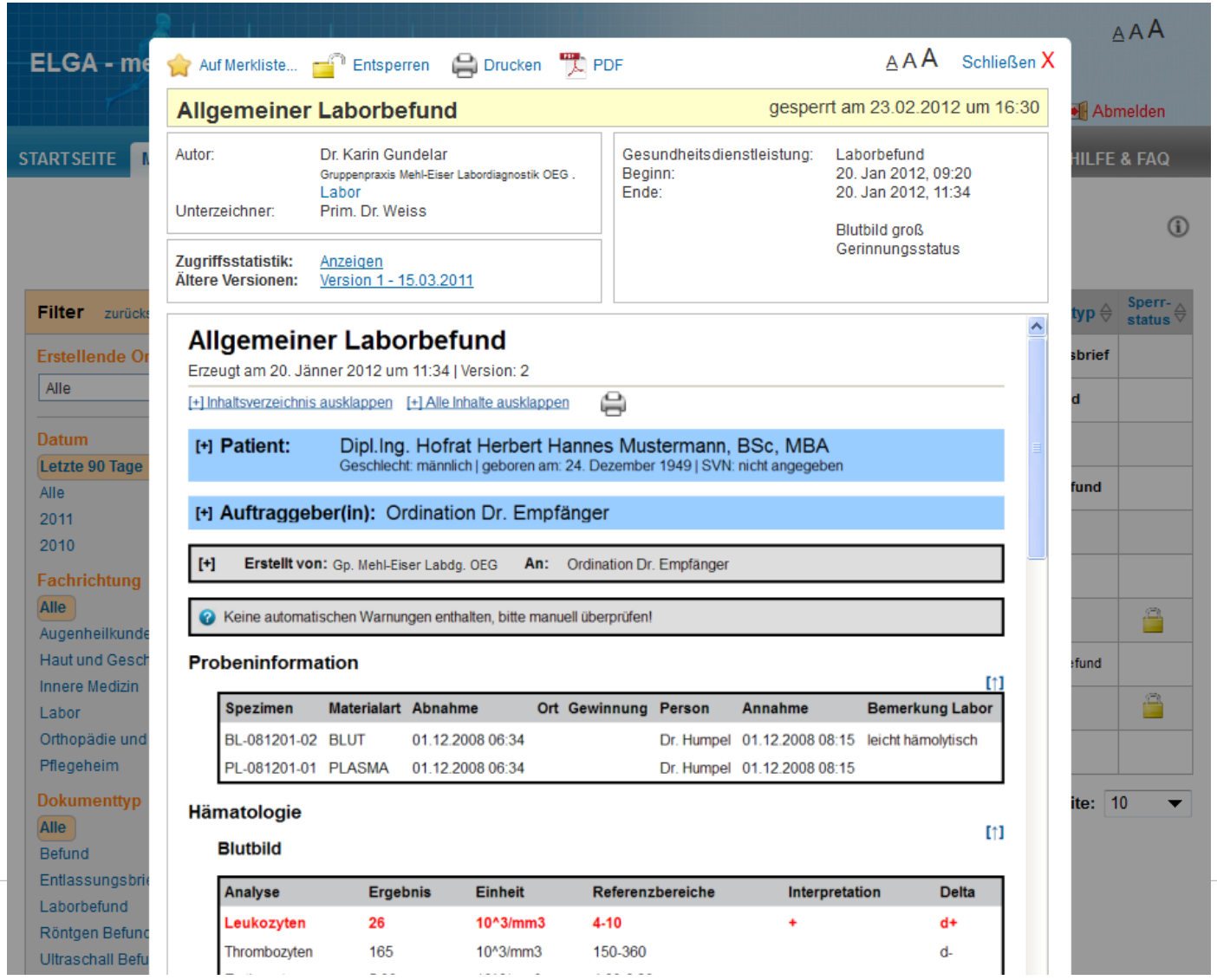


The ELGA-EHR System in Austria - Patient Rights

- Patient can "opt-out" completely
- Patient can do partial opt-outs
- Patient can make individual documents invisible for specific healthcare providers

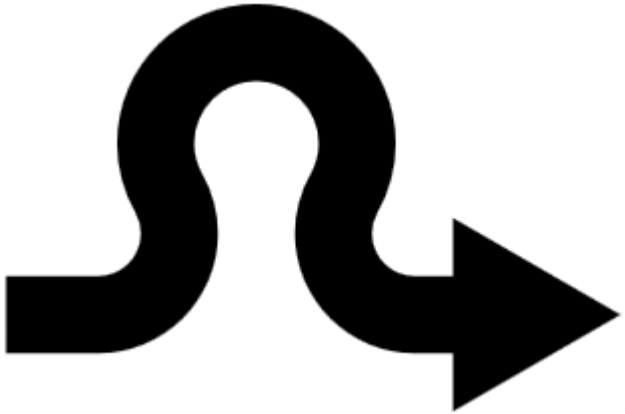
- Patient can view all generated documents
  - But currently now download the CDA-XML document
- Patient can see which healtcare providers have inspected each document and when (logging)

The ELGA-EHR System in Austria - Document View



Use of EHRs for clinical research in Austria

- Essentially not allowed
- Workaround:
  - Investigator must have a "treatment relationship" with patient
  - After informed consent, investigator loads ELGA documents and extracts the information into his own HIS
  - Investigator uses information from the own HIS in the CRF



FH | JOANNEUM
University of Applied Sciences

Use of EHR data for clinical research in other European countries

- Very limited based on EU privacy protection rules
- In some countries (e.g. Germany) EHR data may not be used even when the patient agrees
- Initiatives from different organizations to change (local) regulations
  - Use of anonymized data for research

Opportunities for use of EHRs in clinical research

- Searching for eligible subjects
  - Usually still blocked by legislation
  - Social media are of very good help
- Use of life-long EHRs
  - Detection of long-term effects of medications and treatments (post-marketing surveillance - MIHARI project)
  - Austria: medication data are destroyed after 1 year
  - Usually, hospitals must keep medical records 10-30 years
- Data quality
  - EHR information on "the past" is usually of higher quality (than the patient's memory)

Future developments

- For exchange (and even storage) of healthcare data, the HL7-FHIR standard is coming up
  - XML, JSON, RDF
  - RESTful web services
  - Well-defined basic building bricks: resources
  - Messages and Documents
    - Mappings between HL7-v2 and FHIR exist (see e.g. https://fhirblog.com/2014/10/05/mapping-hl7-version-2-to-fhir-messages/)

- "Networks" of data

FHIR and CDISC

- CDISC is currently working on ODMv2
- Will use many of the ideas of FHIR
- ODMv2 will have a RESTful Web Service API
- Unlike what is currently done, data can lie anywhere (cloud …), one only needs to know the address
  - Distributed data
  - Moving away from "files"

Current gaps

- CDISC really wants to have one transport format from study design to submission: ODM-XML

- Regulatory authorities (including PMDA) still require SAS-XPT (30 years old) for submissions
  - 8-, 40-, 200-character limitations
  - Only for 2D (tabular data)
    - No possibility e.g. to add an EHR data point to the submission data point
  - No support for the Japanese language
    - Some workarounds exist
  - Not well usable with modern technologies such as RESTful web services

Closing the gap

- CDISC developed an ODM-based standard for exchange of submission data: Dataset-XML

- FDA/PMDA is still reluctant to use Dataset-XML

- PMDA could make a great step forward by accepting submissions in Dataset-XML
  - Japanese language
  - Easy integration of EHR into submissions
  - Allowing for modern IT methodologies like Web Services & Artificial Intelligence

FH JOANNEUM
University of Applied Sciences

What we even did not talk about …

- Use of genomic data
- Use of telemonitoring data
- Use of lifestyle data

Obstacles to the use of EHRs in clinical research

- Legislation
- Technical issues are the minor problems

- Major problem are the semantics
  - Different coding systems
  - Over-simple CDISC Controlled Terminology
    - Required to be used in PMDA & FDA submissions
  - Many codes for same CDISC controlled term
    - Example: "ALB" (Albumin measurement) in CDISC
    - LOINC: >50 albumin test codes

Conclusions

- EHR information can be used (and is used) in clinical research
- Technical issues are the minor ones
- Distinction between local EHRs and interoperable EHRs
- Legal aspects can be obstacles
- Long-life EHR ideal, but not often there yet

Where to find more information?

- CDISC: www.cdisc.org
- CDISC CDASH: https://www.cdisc.org/standards/foundational/cdash
- CDISC ODM: https://www.cdisc.org/standards/transport/odm
- IHE profiles for clinical research
  - RFD: https://wiki.ihe.net/index.php/Retrieve_Form_for_Data_Capture
  - CRD: http://wiki.ihe.net/index.php/Clinical_Research_Document