

# CDISC ODM tools with open access: MDM portal and ODMconverter

Martin Dugas

dugas@uni-muenster.de

# Agenda

- About IMI
- Portal of Medical Data Models
- ODMconverter
- Discussion

About IMI

**INSTITUTE OF MEDICAL INFORMATICS  
UNIVERSITY OF MÜNSTER, GERMANY**

# About University Hospital of Münster (UKM)

~ 60.000 inpatients per year

~ 500.000 outpatients per year

1450 beds

30 clinics



# Institute of Medical Informatics:

## **Informatics for Personalised Medicine**

- Health Informatics
- Biomedical Informatics

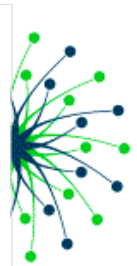
# Single Source



**Routine  
Care**



**Medical  
Research**



# Electronic Health Records for Clinical Research

In 4 domains...

## Emerging and Future EHR-Research Connectivity and Complexity



**Current Divided**

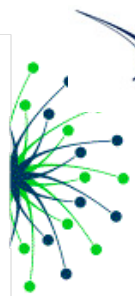
- EHR systems holding source for data used in clinical trials may not meet essential regulatory requirements
- Clinical care data required for clinical trials is entered into EHR and into clinical trial systems
- Research holds source data not part of normal clinical care
- Redundant data entry, transcription errors and source issues cost healthcare and industry time and money.

- Integrated Future**
1. Optimizing clinical protocol designs
  2. Enhanced patient recruitment
  3. Optimized clinical trial data collection
  4. Facilitated Serious Adverse Event reporting

**Interoperable Ideal**

- EHR and research systems work seamlessly together so that groups benefit from data access and mining capabilities across healthcare and research data sets
- Research systems and healthcare systems sit on the same spine
- Systems conform to the same data exchange standards

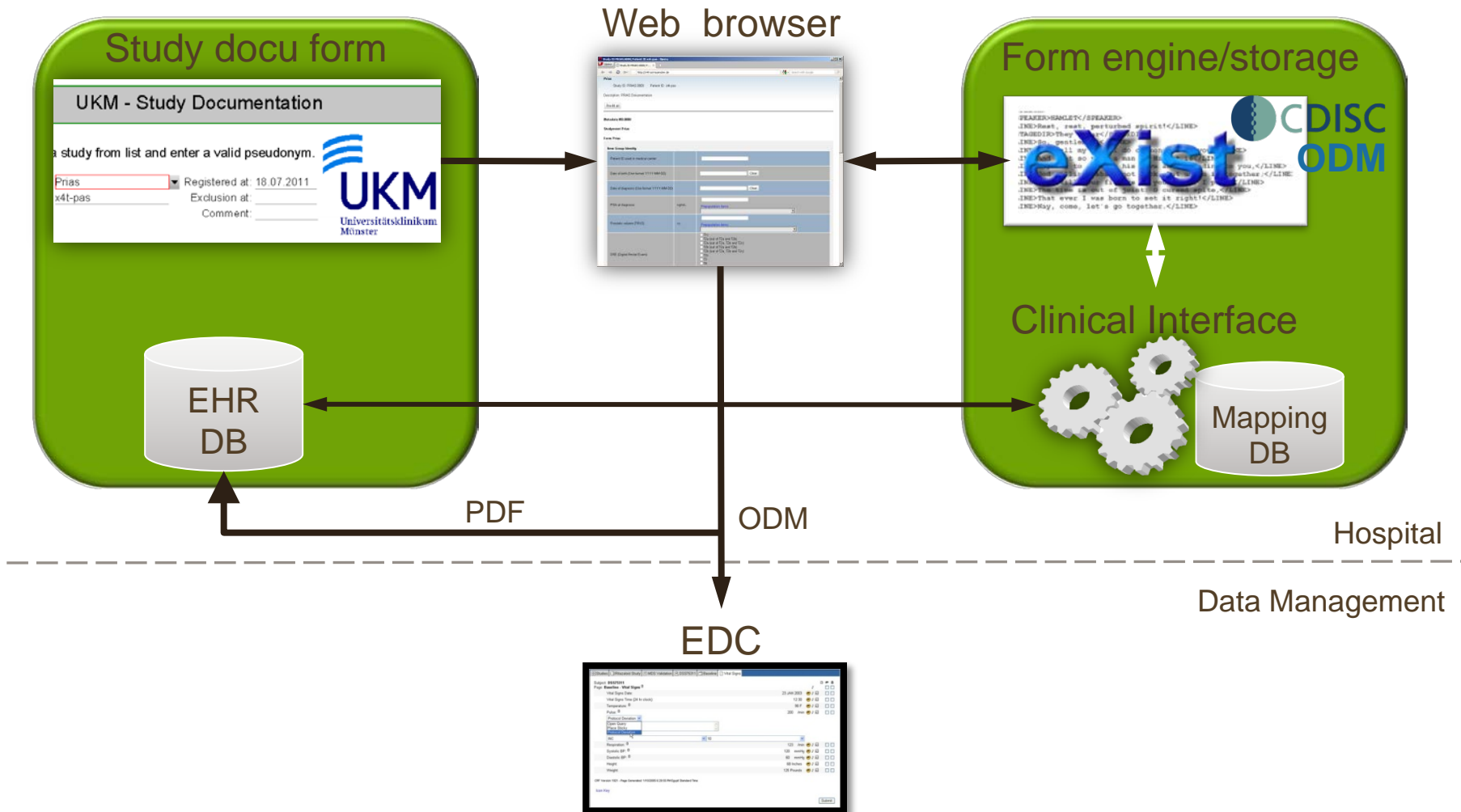
# Consortium





# x4T

## Clinical registries with EHR integration



# MoPat

## PROs and QoL within EHR



iPad 12:52 PM 100%

<http://192.168.2.114/response/81/2>

vorherige Frage DLQI Frage: 3 von 11 nächste Frage

Tägliche Aktivitäten

Wie sehr hat Ihre Hauterkrankung Sie in den vergangenen 7 Tagen bei Einkäufen oder bei Haus- oder Gartenarbeit behindert?

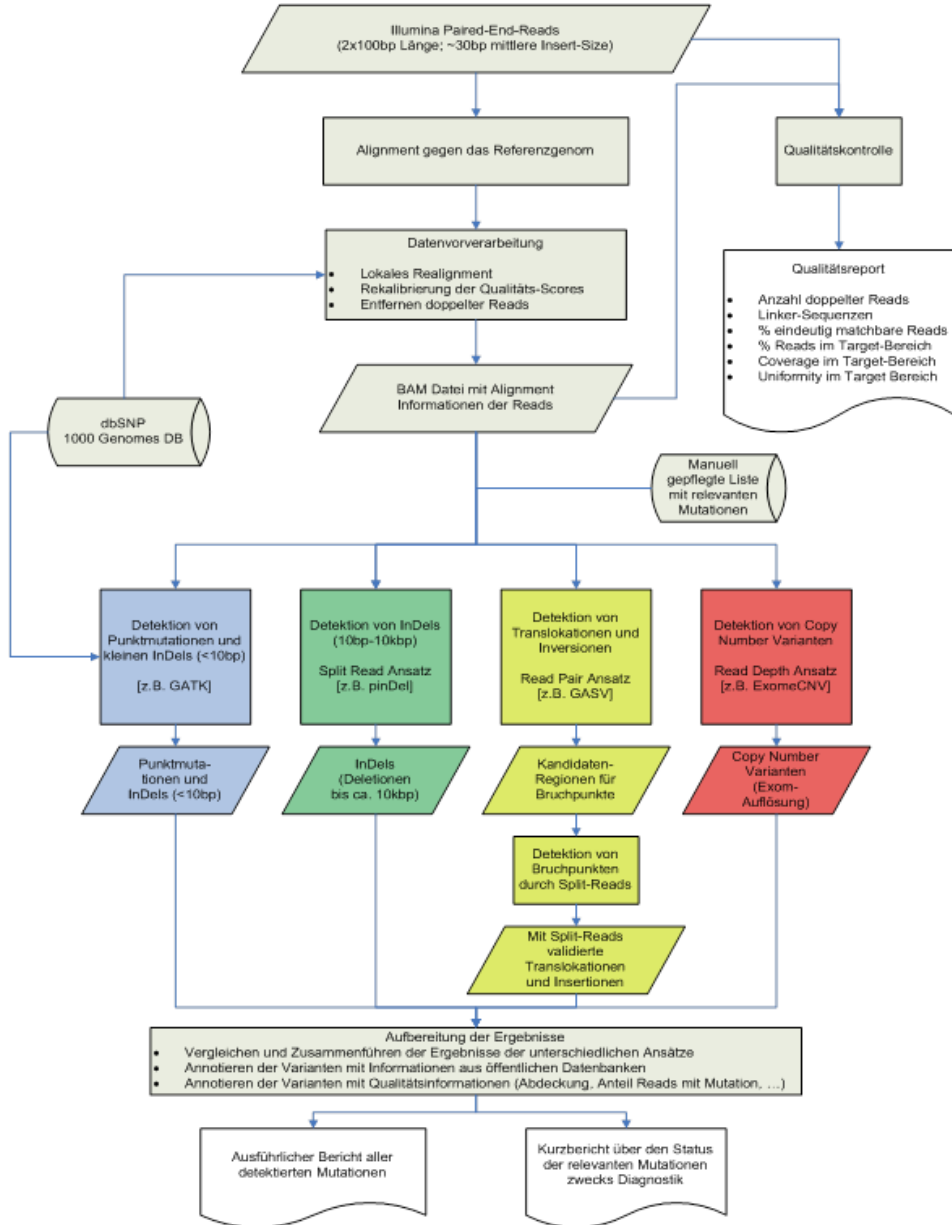
- sehr
- ziemlich
- ein bisschen
- überhaupt nicht
- Frage betrifft mich nicht

# NGS in Myelodysplastic Syndrome

- TRIAGE-MDS: TRanslational Implementation of Genetic Evidence in the management of MDS (ERA-NET on Translational Cancer Research, TRANSCAN)
- NGS for ~1000 Patients (EU MDS registry Nimwegen)

# Clinical Translation: NGS for routine diagnostics

Clinical software tools with EHR integration



CDISC ODM tools with open access

# **PORTAL OF MEDICAL DATA MODELS (MDM)**

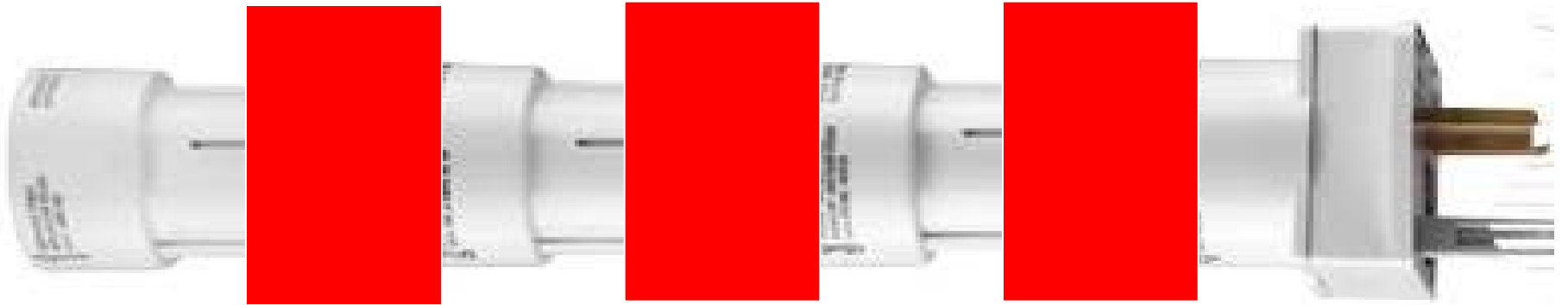
# The Current Information System Landscape in Healthcare

- Electronic Health Records (EHRs)
  - **25% - 40%** of physicians daily workload
  - Proprietary data models for each hospital
  - **> 1000 documentation forms** per hospital with uncontrolled redundancy
- Research documentation (EDC)
  - Large and increasing amount of CRFs per study:  
55 pages (1999-2002) => **180 pages** (2003-2006)
  - Large overlap with routine documentation, especially to exclude adverse events

# Medical Data Models: A LARGE-Scale Problem

- Disease-specific: ICD-10: ~13.300 diagnoses
- Complex vocabulary: ~300.000 terms (SNOMED)
- Language specific: EHR in the local language
- EHR-specific: >800 EHR-products in Europe
- Trial-specific: ~ 150.000 trials  
with >1000 data items per trial

Medical Data Models  
are complex and heterogeneous ...



**and >99% are secret !!!**



>99% of Medical data models are secret

**Differentialblutbild** Datum:    .

<b>Haemoglobin</b>		<input type="radio"/> mmol/l <input type="radio"/> g/dl		Eosinophile		%
<b>Leukozyten</b>		<input type="radio"/> GPt/l <input type="radio"/> / $\mu$ l <input type="radio"/> $10^3$ / $\mu$ l <input type="radio"/> /nl		Basophile		%
Neutrophile			%	Blasten		%
oder	Stabkernige		%	Immunkelozyten		%
	Segmentkernige		%	Myelozyten		%
Lymphozyten			%	Metamyelozyten		%
Monozyten			%	Sonstige Zellen		%
				<b>Thrombozyten</b>		<input type="radio"/> GPt/l <input type="radio"/> / $\mu$ l <input type="radio"/> $10^3$ / $\mu$ l <input type="radio"/> /nl

Top Secret ????

# Open Medical Data Models: Goals

- **Improved design** of data models (Best Practice)
- **Accelerated development** of documentation forms
- **Standardized** forms
- Better **EHR-EDC-Integration** (semantic annotation)
- More **efficient clinical trials**

<http://Medical-Data-Models.org>

Largest public portal for medical forms worldwide

- Contents: Focus on oncology forms
  - ~5.400 Forms / 180.000 data elements
  - Pediatric oncology (GPOH): Ewing, EWOG MDS 2006, EWOG SAA 2010, CWS 2007-HR, SIOP CNS GCT II
  - Leukemia: CML TIGER, AML-AZA
  
- Semantic annotations:  
UMLS, SNOMED, MedDRA, LOINC codes
  
- CDISC ODM-Editor

# Prototype: Portal for Medical Data Models

- Download in various technical formats  
(CDISC ODM, pdf, excel, sql, spss, CDA)
- Online-Discussion of forms:  
Comments regarding items, itemgroups and forms
- Version control
- Forms in multiple languages (up to 26)

# MDM: Semantic Medical Forms

Logged in as Martin Dugas | [Log out](#)



## Medical Data Models



Portal



My forms



Search

A-Z

Keywords



Administration

Information:

Signed in successfully.

### Portal

Best rated forms:	Medical Data Models offers:	Most frequent keywords:
Follow Up ★★★★★ EHR4CR data inventory ★★★★★ NCI Standard Adverse Event/Serious A... ★★★★★ EORTC QLQ-C30 ★★★★★ register Finish Cancer Registry cancer ★★★★★ CDA discharge letter VHitG 1.50 CDA_d... ★★★★★ CDASH ECG – Scenario 2 (local process... ★★★★★ CDASH Vital Signs ★★★★★ Pancreatic Cancer NCT00058149 Qualit... ★★★★★ WHO (Five) Well-Being Index ★★★★★	<b>Forms:</b> 5487 <b>Current versions:</b> 4081 <b>Itemgroups:</b> 27183 <b>Items:</b> 184168 <b>Tags:</b> 2800 <b>Ratings:</b> 84 <b>Comments:</b> 36	Cancer Breast Cancer Leukemia Documentation Diabetes mellitus AML Sarcoma, Ewing Multiple Sclerosis Hospital Information Systems Diabetes Mellitus, Type 2

### Portal of Martin Dugas

Latest forms:	My contributions:	My observations:
Multiple Sclerosis - Current treatment a... today Multiple Sclerosis Registry - Consultatio... today Multiple Sclerosis Registry - Inclusion (... today Anal Cancer NCT00324415 Toxicity - A... yesterday Multiple Sclerosis - Past treatment (RE... yesterday Kidney Cancer NCT00326898 Treatmen... yesterday Ovarian Cancer NCT00993655 Treatme... yesterday Melanoma (Skin) NCT01274338 Treat... yesterday Melanoma (Skin) NCT00110019 Treat... yesterday Breast Cancer NCT01101438 Pre-Study... yesterday	<b>ODM Files:</b> 713 <b>Ratings:</b> 66 <b>Comments:</b> 5	





# Medical Data Models



Portal



My forms



Search

A-Z

Keywords



Administration

Form data

View in detail

## WHO (Five) Well-Being Index



WHO (Five) Well-Being Index  
(1998 version)  
<http://www.who-5.org>

**Form family:** WHO (Five) Well-Being Index

**Version:** 6

**Copyright:** Psychiatric Research Unit, WHO Collaborating Center for Mental Health, Frederiksborg General Hospital, DK-3400 Hillerød

**Created at:** 2014-10-24

**Created by:** Martin Dugas

**Your rating:** ★★☆☆☆

**Average rating:** ★★★★★

**Keywords:** Quality of Life, Patient Reported Outcome (PRO)



Download

Comments (0)

Languages (26)

Versions (6)

Compare versions

### WHO (Five) Well-Being Index

#### Over the last two weeks

##### I have felt cheerful and in good spirits

- All of the time(5)
- Most of the time(4)
- More than half of the time(3)
- Less than half of the time(2)
- Some of the time(1)
- At no time(0)

##### I have felt calm and relaxed

- All of the time(5)
- Most of the time(4)
- More than half of the time(3)
- Less than half of the time(2)
- Some of the time(1)
- At no time(0)

##### I have felt active and vigorous

- All of the time(5)
- Most of the time(4)
- More than half of the time(3)
- Less than half of the time(2)
- Some of the time(1)
- At no time(0)

##### I woke up feeling fresh and rested

- All of the time(5)
- Most of the time(4)
- More than half of the time(3)
- Less than half of the time(2)
- Some of the time(1)
- At no time(0)

# User comments

Logged in as Martin Dugas | [Log out](#)



Medical Data Models



Portal



My forms



Search

A-Z

Keywords



Administration

Form data

View in detail

## Follow Up



Follow-Up documentation for automated Kaplan-Meier-Plots

### Please notice

This is not the latest version of this type of form.

Form family: Follow Up

Version: 13

Copyright: none

Created at: 2014-06-10

Created by: Martin Dugas

Your rating: ★★★★★

Average rating: ★★★★★

Keywords: Cancer, Follow-Up Studies, Kaplan-Meier Estimate



Download

## Comments (1)

Dr. Vojtech Huser 2014-06-10  
nice to see Czech version.

Languages (13)

Versions (14)

Compare versions

## Follow-Up Oncology

### Identity

Patient Last Name

Patient First Name

Patient Day Of Birth

### Diagnosis

Diagnosis text

Date of Diagnosis

ICD10 Code

### Therapy data

Therapy Text

Therapy End Date

# ODM-Editor

WHO (Five) Well-Being Index (WHO-5)		en	Form manager	
Over the last two weeks	Over the last two weeks			
<b>Lately cheerful</b>	I have felt cheerful and in good spirits	<b>C3261637</b>	integer	5=All of the time 4=Most of the time 3=More than half of the time 2=Less than half of the time 1=Some of the time 0=At no time
<b>Lately relaxed</b>	I have felt calm and relaxed	<b>C3261651</b>	integer	5=All of the time 4=Most of the time 3=More than half of the time 2=Less than half of the time 1=Some of the time 0=At no time
<b>Lately active and vigorous</b>	I have felt active and vigorous		integer	5=All of the time 4=Most of the time 3=More than half of the time 2=Less than half of the time 1=Some of the time 0=At no time
<b>Lately slept well</b>	I woke up feeling fresh and rested		integer	5=All of the time 4=Most of the time 3=More than half of the time 2=Less than half of the time 1=Some of the time 0=At no time
<b>Lately daily life interesting</b>	My daily life has been filled with things that interest me		integer	5=All of the time 4=Most of the time 3=More than half of the time 2=Less than half of the time 1=Some of the time 0=At no time
<b>New item</b>				
<b>New itemgroup</b>				
		<b>Download ODM</b>		<b>Upload to Portal</b>



**MDM-PORTAL: LIVE DEMO**  
**[HTTP://MEDICAL-DATA-MODELS.ORG](http://MEDICAL-DATA-MODELS.ORG)**

CDISC ODM tools with open access

# **ODMCONVERTER**

# Different stakeholders in clinical research use different IT tools

- Medical personnel:  
Office tools (e.g. Microsoft Word, Excel)
- Data managers:  
EDC systems (e.g. medidata Rave)
- Statisticians:  
Statistical software (e.g. SAS, SPSS)

# Different stakeholders in clinical research use different IT tools

- CRFs are complex (~180 pages)

=> high cost and errors  
due to manual metadata conversion

e.g. Word-CRF => eCRF => SAS-dataset

# ODMconverter

- Automated metadata transformation between several technical formats
- Supported formats: CDISC ODM, Microsoft Excel, R, CDA
- Available at <http://cran.r-project.org/>
- [Dugas PLoS One 2014, PMID 24587378]

# ODMconverter: Excel => ODM

**Table 1.** Simplified example of data model in office format (spreadsheet).

	A	B	C	D	E	F
1	StudyOID	S.0000				
2	Sponsor	Testsponsor				
3	Condition	Testcondition				
4	StudyName	ODM Test Study				
5	StudyDescription	Test of ODM tools				
6	Form	ODM-Test				
7	FirstName	Test				
8	LastName	Testname				
9	Organization	Test organization				
10						
11	Type	Name	en	UMLS CUI	SNOMED CT 2010_0731	LOINC
12	itemgroup	Info	General Information	C0332118	106227002	
13	boolean	Willingness	Willingness to participate in clinical trials	C1516879		
14	integer	Age	Age		102518004	
15	date	DOB	Date of Birth		152322001	
16	integer	Gender	Gender		139865004	
17	codelistitem	1	male	C0024554	248153007	
18	codelistitem	2	female	C0015780	248152002	
19	string	DiagnosisTx	Diagnosis text		439401001	
20	string	DiagnosisCd	Diagnosis code			
21	float	Crea	Creatinine			38483-4
22	time	labTime	Time of lab value			

The header (line 1–9) contains general information about the study. Line 13–22 provide data items of different data types (column A). Column C presents item labels (en = english). Columns D,E,F contain semantic codes for each data item.

doi:10.1371/journal.pone.0090492.t001

**> office2ODM("ODM-Test.xlsx"))**

# ODMconverter: ODM => R

```
1 <?xml version="1.0" encoding="UTF-8" ?>
2 <ODM xmlns="http://www.cdisc.org/ns/odm/v1.3" Description="ODM-Test S.0000 Testcondition"
3 ODMVersion="1.3.1" CreationDateTime="2013-05-27T21:38:03+01:00" FileOID="ODM-Test S.0000.xml" FileType="Snapshot">
4 <Study OID="S.0000">
5 <GlobalVariables>
10 <BasicDefinitions>
15 <MetaDataVersion OID="MD.1" Name="Metadataversion">
16 <Protocol><StudyEventRef StudyEventOID="SE.1" OrderNumber="1" Mandatory="Yes" /></Protocol>
17 <StudyEventDef OID="SE.1" Name="Testcondition" Repeating="No" Type="Unscheduled">
20 <FormDef OID="F.1" Name="ODM-Test" Repeating="No">
23 <ItemGroupDef OID="IG.1" Name="Info" Repeating="No">
39 <ItemDef OID="I.1001" Name="Willingness" DataType="boolean">
40 <Question>
41 <TranslatedText xml:lang="en">Willingness to participate in clinical trials</TranslatedText>
42 </Question>
43 <Alias Context="UMLS CUI" Name="C1516879" />
44 </ItemDef>
45 <ItemDef OID="I.1002" Name="Age" DataType="integer">
52 <ItemDef OID="I.1003" Name="DOB" DataType="date">
59 <ItemDef OID="I.1004" Name="Gender" DataType="integer">
67 <ItemDef OID="I.1005" Name="DiagnosisTx" DataType="string">
74 <ItemDef OID="I.1006" Name="DiagnosisCd" DataType="string">
80 <ItemDef OID="I.1007" Name="Crea" DataType="float">
87 <ItemDef OID="I.1008" Name="labTime" DataType="time">
93 <CodeList OID="CL.1" Name="Gender" DataType="string">
111 </MetaDataVersion>
112 </Study>
113
114 <AdminData>
128 </ODM>
```

> ODM2R("ODM-Test\_S.0000.xml")

# ODMconverter

```
R Console
'data.frame':  0 obs. of  8 variables:
 $ I.1001: logi
 $ I.1002: int
 $ I.1003: chr
 $ I.1004: Factor w/ 2 levels "1","2":
 .. attr(*, "labels")= chr  "male" "female"
 .. attr(*, "labels2")= chr  "maennlich" "weiblich"
 .. attr(*, "Alias_Contexts")= chr  "UMLS CUI" "SNOMED CT 2010_0731"
 .. attr(*, "UMLS_CUI")= chr  "C0024554" "C0015780"
 .. attr(*, "SNOMED CT 2010_0731")= chr  "248153007" "248152002"
 $ I.1005: chr
 $ I.1006: chr
 $ I.1007: num
 $ I.1008: chr
- attr(*, "StudyOID")= chr "S.0000"
- attr(*, "Sponsor")= chr "Testsponsor"
- attr(*, "Condition")= chr "Testcondition"
- attr(*, "StudyName")= chr "ODM Test Study"
- attr(*, "StudyDescription")= chr "Test of ODM tools"
- attr(*, "Form")= chr "ODM-Test"
- attr(*, "FirstName")= chr "Test"
- attr(*, "LastName")= chr "Testname"
- attr(*, "Organization")= chr "Test organization"
- attr(*, "varlabels")= chr  "Willingness" "Age" "DOB" "Gender" ...
- attr(*, "varlabels_en")= chr  "Willingness to participate in clinical trials" "Age" "Date of Birth" "Gender" ...
- attr(*, "varlabels_de")= chr  "Willingness to participate in clinical trials" "Alter" "Geburtsdatum" "Geschlecht" ...
- attr(*, "DataTypes")= chr  "boolean" "integer" "date" "integer" ...
- attr(*, "itemgroups")= chr  "IG.1" "IG.1" "IG.1" "IG.1" ...
- attr(*, "itemgroups_unique")= chr "IG.1"
- attr(*, "itemgroupnames_unique")= chr "Info"
- attr(*, "itemgroupnames_unique_en")= chr "General Information"
- attr(*, "itemgroupnames_unique_de")= chr "Allgemeine Angaben"
- attr(*, "IGAlias")= chr  "IG.1" "UMLS CUI" "C0332118" "IG.1" ...
- attr(*, "Alias_Items")= chr [1:6, 1:3] "I.1001" "I.1002" "I.1003" "I.1004" ...
> |
```



# ODMconverter: Further tools

- office2ODM
- ODM2office
- ODM2R
- R2ODM
- ODM2CDA
- CDA2ODM
  
- Huge set of R tools, e.g. R => SPSS, SAS

# DISCUSSION

# Discussion:

## French CDISC User Group:

- Test of MDM portal:  
download, upload, form rating & discussion,  
what contents should be added?
- Contribute ODM forms from France
- CRF translation into French
- Test of ODMconverter
- Test of EDC systems regarding ODM compatibility
- Inform CDISC community about these tools
- Develop further ODM functions for R

# Merci pour votre attention!

Martin Dugas