Portal of Medical Data Models
Information Infrastructure for Medical Research and Healthcare

Martin Dugas
dugas@uni-muenster.de
Agenda

• Background

• Portal of Medical Data Models

• Hands-on-training

• Discussion
BACKGROUND
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 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: ~210,000 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

<table>
<thead>
<tr>
<th>Differentialblutbild</th>
<th>Datum:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Haemoglobin</td>
<td>mmol/l</td>
</tr>
<tr>
<td></td>
<td>g/dl</td>
</tr>
<tr>
<td>Leukozyten</td>
<td>GPl/l</td>
</tr>
<tr>
<td></td>
<td>/µl</td>
</tr>
<tr>
<td></td>
<td>10³/µl</td>
</tr>
<tr>
<td></td>
<td>/nl</td>
</tr>
<tr>
<td>Neutrophile</td>
<td>%</td>
</tr>
<tr>
<td>Stabkernige</td>
<td>%</td>
</tr>
<tr>
<td>Segmentkernige</td>
<td>%</td>
</tr>
<tr>
<td>Lymphozyten</td>
<td>%</td>
</tr>
<tr>
<td>Monozyten</td>
<td>%</td>
</tr>
<tr>
<td>Eosinophile</td>
<td>%</td>
</tr>
<tr>
<td>Basophile</td>
<td>%</td>
</tr>
<tr>
<td>Blasten</td>
<td>%</td>
</tr>
<tr>
<td>Monocyte</td>
<td>%</td>
</tr>
<tr>
<td>Myelozyten</td>
<td>%</td>
</tr>
<tr>
<td>Metamyelozyten</td>
<td>%</td>
</tr>
<tr>
<td>Sonstige Zellen</td>
<td>%</td>
</tr>
<tr>
<td>Thrombozyten</td>
<td>GPl/l</td>
</tr>
<tr>
<td></td>
<td>/µl</td>
</tr>
<tr>
<td></td>
<td>10³/µl</td>
</tr>
<tr>
<td></td>
<td>/nl</td>
</tr>
</tbody>
</table>
Semantic Annotation for Data Integration: Concepts

Attribute name NOT sufficient for data integration:
Example “size”:
Body height, Tumor diameter or Foot size ?

=> Semantic code provides clarity,
e.g. Body height C0005890
Semantic annotation for data integration: Permissible values

Data type is NOT sufficient for data integration:
Example “size”, data type FLOAT:
Unit m, cm or mm?

⇒ Unit must be provided,
   e.g. UCUM: m

Code list is NOT sufficient for data integration:
Example “gender”, code list “m=male, f=female”

⇒ Semantic codes provide clarity,
   e.g. male C0024554, female C0015780

Missing semantic annotation in databases. The root cause for data integration and migration problems in information systems.

Dugas M.

Author information

Abstract

Data integration is a well-known grand challenge in information systems. It is highly relevant in medicine because of the multitude of patient data sources. Semantic annotations of data items regarding concept and value domain, based on comprehensive terminologies can facilitate data integration and migration. Therefore it should be implemented in databases from the very beginning.

PMID: 25377893 [PubMed - in process]
How many forms do exist?

- ~ 40 items per form
- ~ 300,000 options per item
- ~ 1,5E171 forms:
Medical information management is a challenge

⇒ Astronomic number of medical forms!
⇒ Independently developed forms are incompatible
⇒ Compatible data only with form sharing
PORTAL OF MEDICAL DATA MODELS (MDM)
http://medical-data-models.org

MDM-Trailer
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**
Welcome to the MDM-Portal!

You are on the website of the world's largest open-access registry for medical forms.

Medical Data Models offers:

| Current forms: | 6319 |
| Form versions: | 13504 |
| Items:         | 477198 |
| Languages:     | over 40 |

Current publications:

- Open Metadata/Open Access to Documentation Forms [Dugas M, et al.]

Contact:

Institute of Medical Informatics
Director: Prof. Dr. Martin Dugas
Albert-Schweitzer-Campus 1, Building A11
48149 Münster, Germany
E-mail: imi@uni-muenster.de
Table of contents:

1. Clinical Trials
   1. Eligibility Determination
   2. Medical History and physical examination
   3. Laboratory
   4. Apparative diagnostics
   5. Treatment
   6. Outcome and Follow-up
   7. Adverse Events
2. Electronic Health Records
3. Registries
4. Quality Assurance
5. Disease-Related Forms
   1. Cardiovascular Diseases
   2. Neurology
   3. Oncology
   4. Diabetes Mellitus
   5. Traumatic Injuries
6. Other

Medical Data Models offers: | Current publications: | Contact: |
--- | --- | --- |
Current forms: 6319 | Portal of medical data models: information infrastructure for medical research and healthcare [Dugas M, et al.] | Institute of Medical Informatics Director: Prof. Dr. Martin Dugas Albert-Schweitzer-Campus 1, Building A11 48149 Münster, Germany E-mail: imi@uni-muenster.de |
Form versions: 13504 | | |
Items: 477198 | Open Metadata/Open Access to Documentation Forms [Dugas M, et al.] | |
WHO-5

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

Created: 2016-05-02 by Martin Dugas
Author / Copyright: Psychiatric Center North Zealand, Hillerød, Denmark
License: CC BY-NC 3.0

Quality of Life
Patient Reported Outcome (PRO)

You need to be logged in to download this form.

**Comments (0)**
You need to be logged in to create comments.

**Languages (29)**

**Versions (15)**

**WHO (Five) Well-Being Index**

**Over the last two weeks (0)**

**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 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)

**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)

**My daily life has been filled with things that interest me**
- 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)
WHO (Five) Well-Being Index (1998 version)
http://www.who-5.org

Created: 2016-05-02 by Martin Dugas
Author / Copyright: Psychiatric Center North Zealand, Hillerød, Denmark
License: CC BY-NC 3.0

WHO (Five) Well-Being Index

Over the last two weeks
Name: Over the last two weeks
Description: Over the last two weeks

<table>
<thead>
<tr>
<th>Item</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>I have felt cheerful and in good spirits</td>
<td></td>
</tr>
</tbody>
</table>

Datatype: Integer
Aliases: 
UMLS CUI [1] C3261637
Measurement Units
Range checks
Codelist items
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)
http://Medical-Data-Models.org
Largest public portal for medical forms worldwide

- Contents: Focus on oncology forms
  - ~6,300 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
http://Medical-Data-Models.org

- Keyword search (MeSH, multilingual)
- Download formats:
  - PDF
  - Excel, CSV
  - ODM
  - REDCap
  - CDA
  - FHIR
  - SQL
  - SPSS, R

- Upload of new forms by users
Contents: Time course

[Dugas et al., Database 2016, PMID 26868052]
Frequent keywords
Frequency distribution of UMLS codes: „long tail“
http://Medical-Data-Models.org

• Online-Discussion of forms: Comments at item-, item group- or form level

• Version management

• Multilingual forms: overall > 40 languages
<table>
<thead>
<tr>
<th>WHO (Five) Well-Being Index (WHO-5)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Over the last two weeks</strong></td>
<td><strong>Over the last two weeks</strong></td>
</tr>
<tr>
<td><strong>Lately cheerful</strong></td>
<td><strong>C3261637</strong></td>
</tr>
<tr>
<td>I have felt cheerful and in good spirits</td>
<td>integer</td>
</tr>
<tr>
<td><strong>Lately relaxed</strong></td>
<td><strong>C3261651</strong></td>
</tr>
<tr>
<td>I have felt calm and relaxed</td>
<td>integer</td>
</tr>
<tr>
<td><strong>Lately active and vigorous</strong></td>
<td></td>
</tr>
<tr>
<td>I have felt active and vigorous</td>
<td>integer</td>
</tr>
<tr>
<td><strong>Lately slept well</strong></td>
<td></td>
</tr>
<tr>
<td>I woke up feeling fresh and rested</td>
<td>integer</td>
</tr>
<tr>
<td><strong>Lately daily life interesting</strong></td>
<td></td>
</tr>
<tr>
<td>My daily life has been filled with things that interest me</td>
<td>integer</td>
</tr>
</tbody>
</table>

New item

New itemgroup

Download ODM

Upload to Portal
MDM - Portal of Medical Data Models

https://medical-data-models.org

Contact:
Prof. Dr. Martin Dugas

Details:
Average number of national users per year: 300
Average number of European users per year: 150
Average number of international users per year: 50
Main status: Operational from 2011
RI Type: Single-sited

Description:
The Portal of Medical Data Models (MDM) is a multi-lingual repository of medical data models for the scientific community. It contains data models from medical research (clinical studies) and routine patient care (electronic health records). MDM is an officially registered research infrastructure funded by the German Research Foundation (DFG). It was established in 2011 as an open access repository. To date there are >4,000 data models with >300,000 data elements available for download. Tumor medicine is a key topic in MDM and there are close links with the European Leukemia Network (ELN). MDM is the largest system of its kind world-wide. It is developed and maintained in close collaboration with the University Library of Münster.

Keywords:
Medical data model, clinical studies, electronic health records (EHR), CDISC ODM, HL7 CDA

Research Service(s):
Metadata for Information Systems in biomedical research
The MDM portal provides metadata for information systems in clinical research and routine healthcare.

Equipment:
Internet Portal for Medical Data Models
MDM content is available to the scientific community via https://medical-data-models.org

Last updated on: 2015-09-24 at 14:43
MDM-PORTAL: LIVE DEMO
HTTP://MEDICAL-DATA-MODELS.ORG
HANDS-ON-TRAINING
DISCUSSION
• Test of MDM portal: download, upload, form rating & discussion, what contents should be added? what technical formats should be supported?

• Contribute ODM forms

• Test of EDC systems regarding ODM compatibility
Portal of Medical Data Models
Information Infrastructure for
Medical Research and Healthcare

Martin Dugas
dugas@uni-muenster.de