A Metadata-based Patient Register for Cooperative Clinical Research

A Case Study in Acute Myeloid Leukemia

Anja S. Fischer
Ulrich Mansmann

Institute for Medical Informatics, Biometry and Epidemiology (IBE)
Ludwig-Maximilians-University Munich, Germany
Clinical research databases from ...

- Invest.-initiated trials (IITs)
- Patient registries
- Biobanks

Meta-analyses
Subgroup analyses
Prognostic factor research
Biomarker research

Simulation of future trials
Determ./Eval. of surrogate endpoints
Health economic research

Secondary use

Combination of several data sources:
- broader evidence for exploratory research
- data sets for validation of findings
- important with rare diseases

Need for Infrastructure for pooling of heterogenous data sources between academic study groups

Clinical research of Acute Myeloid Leukemia in Germany (2005-2011)
- 14 study groups
- 58 IITs
- patient registries and biobanks
Data Pooling across study groups

Patient Registry

**Constraints**
- Time and cost intensive
- Heterogenous environments at sites

**Demands**
- Maximizing the broadness of research

Conflict results in an undesirable limitation to the range of research questions to be answered.

Data Warehouse

Condition: Willingness of participating partners to release complete research databases to the community

Meta-data based patient register

...full data sets from distributed sources transparent ...
...on metadata basis, cooperative projects can be planned ...
... clinical data isn’t shared till the project conditions are set...
Concept of the Meta-data based patient register

**Data about**
- project
- data dictionary
- patients
- data quality

**Research project**
e.g. clinical trial

Clinical research data

---

**Metadata about a data source**

<table>
<thead>
<tr>
<th>Attributes of the research project</th>
<th>e.g. project type, protocol synopsis, population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Status of data management processes</td>
<td>e.g. data capture, data validation, database closure</td>
</tr>
<tr>
<td>Description, structure and content of (e)CRFs</td>
<td>i.e. scheme of study visits and forms</td>
</tr>
<tr>
<td>Description of data items + data validation plan</td>
<td>e.g. item description, data type and precision, location of item in case report form, validation and plausibility checks</td>
</tr>
<tr>
<td>Pseudonyms of patients + „Captured/Missing“ flags</td>
<td>i.e. a true/false flag, indicating on the data item level, if clinical information about a single patient was captured (t) or is missing (f)</td>
</tr>
</tbody>
</table>
Concept of the Meta-data based patient register

Challenges

**Metadata:**
- Exchange format?
- Extraction? Transport? Storage?
- User interfaces
- Support for definition of a project data set

**Data:**
- Extraction of remote clinical data?
- Transformation of heterogenous data to an homogenous project data set
Current status of implementation
- Focus on integrating clinical trial databases -

Metadata format: CDISC Operational Data Model (ODM)
+ ODM extension (for „Captured/Missing“-flag)

ODM Toolbox
- UML 2.0 model and API
- `odm.xml` (JAXB) -> ODM model
- ER Model / PostgreSQL / Hibernate Persistence Mapping
- Tool: DB2ODM

- Meta-data of 5 clinical trials from 3 study groups integrated (~ 7000 patients)
- Sample extraction of clinical data concerning AML diagnosis classifications and provided for statistical analysis
- 2 research projects ongoing (therapy-related AML / health services research)
Summary

It’s a … cross between metadata warehouse and patient register
- No restriction to a common core data set
- No release of research data in advance, but confined to a setting

Benefit for cooperated clinical research

+ Infrastructure to leverage cooperation projects
  - Transparency about data sources
  - Motivation and support for cooperation projects
  - Central repository of data transformation rules

+ Communication
  - Standardization of future CRFs
  - Legacy data → ODM

Further challenges: GUIs, Notation of transformation rules, Record Linkage

Thank you for your attention!