Experiences from the National Demonstrator Study within the German Medical Informatics Initiative

Session S15: Data Banks for Research

Thomas Ganslandt¹, Jannik Schaaf², Josef Schepers³, Holger Storf², Felix Balzer³, Silke Haferkamp⁴, Robert Lodahl⁵, Fabian Prasser⁶, Ulrich Sax⁵, Holger Stenzhorn⁷, Hans-Ulrich Prokosch⁸, Martin Boeker⁹

¹ Heinrich-Lanz-Center for Digital Health, Mannheim University Medicine, Heidelberg University; ² University Hospital Frankfurt; ³ Charité - Universitätsmedizin Berlin; ⁴ University Hospital Aachen; ⁵ University Medical Center Göttingen; ⁶ University Hospital rechts der Isar; ⁷ University Hospital Tübingen; ⁸ Friedrich-Alexander-University Erlangen-Nuremberg; ⁹ Medical Faculty and Medical Center - University of Freiburg
Disclosure

I disclose the following relevant relationships:

• Board member of the German Telematics Platform for Networked Medical Research (TMF) e.V.

• Third-party-funding by German Federal Ministry of Education and Research (Grant 01ZZ1801E) and Baden-Württemberg State Ministry of Science and Culture
Introduction: The German Medical Informatics Initiative (MII)

- Foster re-use of routine clinical data
- Demonstrate utility through clinical use cases
- Strengthen Medical Informatics as a discipline
- 160 M€ funding by BMBF
- Long-term perspective
Introduction:
The 4 MII Consortia

Image source: http://www.medizininformatik-initiative.de/en/node/5
Introduction: Shortcut needed through a four-year Roadmap

Making clinical data available

Organizing access to data

Demonstrating collaborative use
Introduction: Steps towards harvesting low-hanging fruit, early-on

MII Modular Core Dataset
- based on HL7 FHIR
- specified collaboratively by MII Interoperability WG
- modules based on MII use cases and data availability

Readily available data
- German inpatient billing dataset ("§21")
- covers 5 of the 7 basic core dataset modules
- tightly defined, available at all German hospitals
- tooling available from prior work

Limited, reproductive research questions
- based on the §21 billing dataset
- Analysis of Comorbidities
  - using principal and secondary diagnosis codes
  - calculation of published comorbidity scores (Charlson, Elixhauser, Stausberg)
  - correlations with case data
- Geovisualization of rare diseases
  - based on ICD10 diagnosis codes
  - visualization of distance to providing sites
Methods: Focus of MII Demonstrator Study

standardized dataset, established platform

Making clinical data available

Organizing access to data

Harmonized Governance: Ethics & Broad Consent, Data protection concepts, Use & access committees

study protocol based on locally aggregated, anonymized analyses

Demonstrating collaborative use

simple, reproductive research questions

Getting access, Extraction, Integration, Harmonization

Research Data Repositories: Data queries, Result Merge, Analysis, Publication

Methods: Focus of MII Demonstrator Study

standardized dataset, established platform

Making clinical data available

Organizing access to data

Harmonized Governance: Ethics & Broad Consent, Data protection concepts, Use & access committees

study protocol based on locally aggregated, anonymized analyses

Demonstrating collaborative use

simple, reproductive research questions

Getting access, Extraction, Integration, Harmonization

Research Data Repositories: Data queries, Result Merge, Analysis, Publication
Methods: Technical implementation of the Demonstrator

Making clinical data available

Methods:
Technical implementation of the Demonstrator

Demonstrating collaborative use
Results:
Site Participation and Data Volume

- 20 Locations
- 19 Approvals
- 1.8 Mill. patients
- 3.2 Mill. cases

(09/2018 - 03/2019)
Results: Charlson Index vs. Discharge Reason

![Box plot showing the distribution of Charlson Index for different discharge reasons.](image-url)
### Results:

**Charlson comorbidity categories vs. principal diagnosis**

<table>
<thead>
<tr>
<th>Category</th>
<th>Fraction of cases with the comorbidity</th>
</tr>
</thead>
<tbody>
<tr>
<td>16. Certain conditions originating in the perinatal period</td>
<td>-</td>
</tr>
<tr>
<td>15. Pregnancy, childbirth and the puerperium</td>
<td>-</td>
</tr>
<tr>
<td>09. Diseases of the circulatory system</td>
<td>-</td>
</tr>
</tbody>
</table>

- 0.0
- 0.1
- 0.2
- 0.3
Discussion

What was achieved
- rapid demonstration of successful collaboration across all four MII consortia
- open source tools (i2b2, IDRT) were leveraged
- established local access to data & contacts to governance
- results consistent with expectations

Limitations
- even with very simple, aggregated & anonymized approach, it took up to 6 months for some approvals
  - complicated by state-level privacy laws in Germany
- limited scope & potential quality issues of billing data
- insufficient coverage of rare diseases in ICD10

Outlook
- approach will be applied in the upcoming MII cross-consortial use cases (CORD, POLAR)
- next iterations will include further core dataset modules (e.g. laboratory, medication) and leverage FHIR
- approach is transferable
Thank you!

Email me at: thomas.ganslandt@medma.uni-heidelberg.de