Designing and implementing a biobanking IT framework for multiple research scenarios

Hans-Ulrich Prokosch

Chair of Medical Informatics, University of Erlangen-Nuremberg

CIO of Erlangen University Hospital
Biobanking IT projects have been initiated for various topics

- Registrar-based **consent process** for an opt-in biobank in an EHR environment
  - Marsolo K, Corsmo J, Barnes MG, Pollick C, Chalfin J, Nix J, Smith C, Ganta R. Challenges in creating an opt-in biobank with a registrar-based consent process and a commercial EHR. J Am Med Inform Assoc published 9 August 2012, 10.1136/amiajnl-2012-000960 Open Access

- **Identity Management**

- **Ontologies** in federated approaches
Biobanking IT projects have been initiated for various topics

- **Implementation** for one multicenter study

- **Implementation** for an institution

- **Requirements catalogue** for biobanking IT components
Open issues

- How can one institution make efficient use of a set of biobanking IT components for multiple research scenarios?

- What are the core components of a biobanking IT framework which can support a broad variety of biobanking scenarios between
  - institutional patient care based single location biobanking, and
  - multicenter research biobanks?
Objectives of our work

- Design and implement a biobanking IT framework for multiple research scenarios at Erlangen University Hospital.

- Define the requirements for biobanking IT components depending on
  - varying research scenarios,
  - varying patient workflows in routine patient care processes and
  - varying throughput requirements
    (e.g. numbers of samples collected / redistributed for analysis).

- Define and establish reusable biobanking IT components.
Biobanking history at Erlangen University Hospital

1. Germany-wide **multicenter** chronic kidney disease **cohort** with
   - decentralized patient recruitment, patient interviews, sample acquisition (liquid samples) and
   - a centralized biobank
   - high throughput volume (>50,000 samples collected at baseline visit)

2. Comprehensive Cancer Center at **one University Hospital location** with
   - sample acquisition (tissue banking) during patient care
     (**cancer diagnostics and therapy**)  
   - medium throughput volume

3. Germany-wide multicenter biobanking cooperation (for prostate cancer research) with
   - clinical annotations acquired at multiple university hospitals during routine patient care
   - no direct support for sample management
   - requirement: **patient cohort identification tool**
Varying requirements for the different scenarios

1. Germany-wide multicenter chronic kidney disease cohort

- informed patient consent is gathered automatically for every patient recruited into the cohort
- application of a web-based **remote data entry** system for clinical annotations
- direct **pseudonomized patient documentation** in eCRFs
- required **specialized** application support for quality assurance in **sample reception** and distribution in multiple freezers (high throughput)
- required automated **interfaces with aliquoting roboters**
Varying requirements for the different scenarios

2. Comprehensive Cancer Center at one University Hospital location

- sample acquisition during surgery or diagnostic examinations
- informed patient consent only for selected patients by clinical physicians
- sample preparation and storage in pathology
- clinical annotation with patient identification in various EHR modules
Varying requirements for the different scenarios

3. Germany-wide multicenter biobanking cooperation for prostate cancer research

- sample acquisition, sample storage and informed patient consent are issues to be considered separately within every university hospital
- clinical annotation with patient identification in various EHR modules of different commercial vendors
- varying clinical documentation sets need to be aligned to one common minimal dataset
- only anonymized annotations are allowed to be transferred outside the boundaries of one institution
- web-based cohort identification tool shall nevertheless allow to link back to local patients and samples
Biobanking IT Core Components

clinical annotation / patient phenotype data

clinical care setting

EMR
assessment form
patient data with patient identifier

medical research setting

CTMS
eCRF
pseudonymized patient data

sample meta data

BMS
sample data with labid

University Hospital A
Biobanking IT Core Components

Clinical annotation / patient phenotype data
- Clinical care setting
  - EMR
    - Assessment form
    - Patient data with patient identifier
- Medical research setting
  - CTMS
    - eCRF
    - Pseudonymized patient data

Sample meta data
- Medical research setting
  - BMS
    - Sample data with labid

Research data warehouse

Cohort query frontend

University Hospital A
Biobanking IT Core Components

clinical annotation / patient phenotype data

clinical care setting
- assessment form
- patient data with patient identifier

medical research setting
- eCRF
- pseudonymized patient data

sample meta data
- sample data with labid

University Hospital A

EMR

CTMS

BMS

PID generator

research data warehouse

Cohort query frontend

H.U. Prokosch / Chair of Medical Informatics
Biobanking IT Core Components

- Clinical annotation / patient phenotype data
  - Clinical care setting
  - Medical research setting
  - EMR: assessment form, pseudonymized patient data
  - CTMS: eCRF, pseudonymized patient data
  - BMS: sample data with labid

- Clinical care setting
  - EMR: patient data with patient identifier

- Medical research setting
  - CTMS: pseudonymized patient data

- Sample meta data
  - BMS: sample data with labid

- PID generator
  - Pseudonymized patient data

- Research data warehouse
  - Cohort query frontend

University Hospital A
Biobanking IT Core Components

clinical annotation / patient phenotype data

clinical care setting
- assessment form
- patient data with patient identifier

medical research setting
- eCRF
- pseudonymized patient data

sample meta data
- sample data with labid

University Hospital A

- PID generator
- pseudonymized patient data
- research data warehouse
- Cohort query frontend

BMS

CTMS

EMR
Biobanking IT Framework for Multiple Research Scenarios

A. Biobanking Core Components

- Cohort query frontend
- PID generator
- eCRF
- Clinical annotation / patient phenotype data
- Sample meta data

University Hospital A

- EMR
- CTMS
- BMS

- Pseudonymized patient data
- Patient data with patient identifier

University Hospital B

- University Hospital C

- Trans-Institutional i2b2 Integration Database
- Anonymous patient data
- Query definition

- Clinical care setting
- Medical research setting
- CTMS
- EMR

- University Hospital A
- University Hospital B
- University Hospital C
Identification of application independent platform core components

- the biobank management system (BMS),
- the GCP-certified clinical trials management system (CTMS),
- the hospital´s EMR
  - including multidisciplinary clinical assessment forms and
  - informed patient consent documentation
- a user-friendly query frontend being applicable as
  - a local research data warehouse as well as
  - a trans-institutional integration database
- a PID generator serving as the identity management component
Identification of application independent platform core components

- the biobank management system (BMS, STARLIMS®),
- the GCP-certified clinical trials management system (CTMS, secuTrial®),
- the hospital’s EMR (Soarian Clinicals ®)
  - including multidisciplinary clinical assessment forms and
  - informed patient consent documentation
- a user-friendly query frontend (i2b2) being applicable as
  - a local research data warehouse as well as
  - a trans-institutional integration database.
- a PID generator (TMF) serving as the identity management component
Challenges / Take home lessons

- **Workflow Analysis Challenge**

  Transfering such a Biobanking IT Framework to other institutions is possible !!!

- **Ontology Challenge**

- matching of data items between different documentation systems,
- defined independently from each other at different institutions

- **Research Project Independency Challenge**

  - synergy effects are possible with hospital-wide, cross-project approaches

**ColoNet consortium North Germany**

*(Lübeck, Hamburg, Rostock, Greifswald)*

- local “EMR application” for clinical annotations
- local PID Generator
- i2b2 data warehouse the consortium´s integration database
Thank you very much . . .

Thanks to my coworkers:
S. Mate, J. Christoph, A. Beck, F. Köpcke, S. Stephan, M. Ries, T. Bürkle, T. Ganslandt,

the cooperating clinicians:

and our medical informatics partners throughout Germany:
B. Breil, J. Ingenerf, M. Hackmann