Clinical Genomics: the next step to personalized healthcare

Speed medical research while preserving patient privacy in Academisch Medisch Centrum Amsterdam.

Drs. ir. Juriën Taams
Solution specialist
Healthcare and Life Sciences
IBM Benelux
Agenda

1. Challenges in Medical Science were never this big and complex!
   How biobanks and IT can help to solve the puzzle

2. Clinical Genomics Framework/Biobank Information Management System
   Defining functions and services

3. Building and Implementing at AMC
   Reflections on different project aspects

4. Potential Extensions to the Solution
   Is the implemented solution scalable?
Agenda

1. Challenges in Medical Science were never this big and complex!
   How biobanks and IT can help to solve the puzzle

2. Clinical Genomics Framework/Biobank Information Management System
   Defining functions and services

3. Building and Implementing at AMC
   Reflections on different project aspects

4. Potential Extensions to the Solution
   Is the implemented solution scalable?
Challenges in Medical Science were never this big and complex!

- Our knowledge about the human genetic characteristics has grown tremendously during the last decade.

- The Human Genome Project provides a nearly complete catalogue of all human genes. And yet...this is the beginning

- The big challenge for the medical science will be to understand how genetic make up, environment and the development of diseases relate to each other.
The puzzle contains numerous pieces....

Main objects of study in medical research

- Clinical data/cohorts
- Images
- (analyses of) body tissue

- Genes
- Body processes
- Environment
- Life Style
- Clinical Course
Solving the puzzle will require access to diverse, integrated information

**Challenges**
- Volume and complexity of data
- Integrating massive volumes of disparate data
- Need for sophisticated analytics
- Growing collaboration across ecosystem

1. **Patient Information**
   - Hospital events: admission, surgery, recovery, discharge
   - X-rays, MRI, mamograms, etc
   - Clinical Record

**Access to Diverse Heterogeneous Distributed Data**

- Expression Arrays (various tissues)
- Personal genomics
- Analysis lab notes
Biobanks will play an important role solving the puzzle

- Molecular-level understanding of disease
- Enabling development of targeted treatment solutions
- Improving efficacy, safety and cost-efficiency of patient care

Basic R&D → Clinical Trials → Treatment

Biobank

Phenotype: Clinical Data, Biopsies, Biomedical Image Data
Genotype: High throughput Research Data
Environment: Environmental Data

Molecular-level understanding of disease
Enabling development of targeted treatment solutions
Improving efficacy, safety and cost-efficiency of patient care
Agenda

1. Challenges in Medical Science were never this big and complex!
   How biobanks and IT can help to solve the puzzle

2. Clinical Genomics Framework/Biobank Information Management System
   Defining functions and services

3. Building and Implementing at AMC
   Reflections on different project aspects

4. Potential Extensions to the Solution
   Is the implemented solution scalable?
A Biobank Information Management System

Not only biological samples...

.. but also an information management system
Integration is necessary

Advanced Scientific Data Collection…
... is More Complex...

- Biological Sample Data
- Genotype Data
- Phenotype Data
- Registers

Study Person — Researcher
... Requiring Flexible Data Integration Solutions
Complete Biobank Solution Architecture

BIMS Services
- Public Info
- Study Mgmt
- Data Collection
- Sample Mgmt
- Data Query
- Data Analysis
- Admin Services

BIMS Backend Architecture
- BIMS (federated) Data Repository
- Data Categorization
- Security Services
- Data Integration
- Original Data Sources
Agenda

1. Challenges in Medical Science were never this big and complex!
   How biobanks and IT can help to solve the puzzle

2. Clinical Genomics Framework/Biobank Information Management System
   Defining functions and services

3. Building and Implementing at AMC
   Reflections on different project aspects

4. Potential Extensions to the Solution
   Is the implemented solution scalable?
AMC/IBM partnership in clinical genomics and biobanking

- AMC Strategic IT department approached IBM in 2004:
  - AMC had made a master plan for ICT in the future. An important part was to integrate Clinical and Research environments
  - AMC has chosen to do this with a strong technology partner with proven track record

- IBM filled this in by:
  - Putting forward IBM’s general vision on healthcare
  - Explaining our Clinical Genomics framework
  - Leveraging our Mayo experience and expertise

- AMC embraced IBM’s Clinical Genomics vision and started a joint proof of concept, is now implementing the framework (DIO project Data Integratie voor onderzoek)
AMC - Base Requirements

Based on AMC context, the following requirements needed to be taken into account:

- Clinical data must be accessible by research department
- Clinical environment may not be impacted, in term of performances, by researchers actions
- Data should be kept only at the source database. So replication is not allowed
- Data must be updated from both, clinical and research site
The Clinical Genomics Solution Architecture

- **Query Tool**
  - DDQB – data discovery and query builder

- **De-identification**
  - HDB (Hippocratic database) de identification framework

- **Data Federation**
  - DB2II – DB2 Information Integrator (WebSphere II)

- **Data Sources**
  - IBD
  - DNA bank
  - KlinVir
  - DPS
  - LABZIS

- **IBM proof of concept components**
  - AMC databases
Key functionalities of the framework
Query tool
Example: Dr. Sander’s Query

- Patients diagnosed with Acute Myeloid Leukemia
- Diagnosed while less than 60 years of age
- Diagnosed while living in Illinois
- Normal White Blood Cell Count
- Male
- Living
- Mention of Prednisone treatment in Clinical Notes
- For which good quality microarray expression data exists
A Physician driven abstract query tool
De-indentification
Functionality of Hippocratic Database

- Active Enforcement
  - Policy creation
  - Enforce the policy during retrieval
- De-identification
- Re-identification
- Compliance Auditing
  - Logging of queries and query results

The HDB application is a joint development program of AMC and IBM Almaden Research Center, San Jose, California
Important Benefits for AMC

- Make **optimal use of the large amount information** in AMC by integration of diverse clinical and biomedical data
- **Significant time savings** when gathering data (Minutes in stead of weeks or months)
- **No database technology knowledge necessary** by users
- **No errors** due to manual data manipulation
- **Privacy policies** can be enforced across the organization
- **Simplification of IT-control.** Reduction of complexity of the IT-environment with clear connections to databases and applications
- **Easy interface** with analysis tools (SAS, SPSS, Spotfire, others)
- The system will **show associations** within the clinical and research data that **helps** the researcher **formulate or explore new hypotheses** of the genetic and biological basis of disease. This will enable more accurate diagnoses and tailored treatment, **resulting in improved patient outcomes**
Phases of implementation in the AMC of CGF

**Phase 1**
- Full functionality, scalable framework
- Robust authentication
- Authorization
- Robust de-identification
- Performance tuning
- Embedding in IT infrastructure

**Phase 2**
- New research line: CVA
- More supporting databases (surgical reports, medication)
- Embedding in organization

**Proof of Concept**
- Six databases integrated
- Significant functionality
- Simple de-identification

---

**Functionality and Complexity**

**Time**

- 2004
- 2005
- 2006
Building Phase 1 en 2 - Success Factors

- Cooperation between IT department, research staff and IBM
- Expectation management
- Iterative development process
- Speak the language of the users
- Modestly sized core team (2 - 4 people)
- Employ generalist IT personnel (no super-specialists)
- Bring in expert knowledge as needed
- Encourage short lines of communication
Building Phase 1 en 2 - Lessons learned

- The involvement of many departments and professions makes for a complex project organization, make sure to have short lines to DMU’s in the client organization
- Plan as far in advance as possible, as detailed as possible.
- Building the DDQB abstraction model (DAM) took more time than expected
- Creation of database wrappers took less time than expected
- Technology is not the bottle neck, culture and human factors are key to realize sharing data and collaboration
Agenda

1. Challenges in Medical Science were never this big and complex!
   How biobanks and IT can help to solve the puzzle

2. Clinical Genomics Framework/Biobank Information Management System
   Defining functions and services

3. Building and Implementing at AMC
   Reflections on different project aspects and lessons learned

4. Potential Extensions to the Solution
   Is the implemented solution scalable?
Possible future applications within AMC

The solution could give possibilities of further usage of the information for example:

- Data mining
- Management reporting
- Consistency checking, e.g. patient records vs. DBC-registration
- Quality control
Possible future applications outside AMC

Potential Extensions to the solution:
... three-dimensional scalability guidelines
Het Parelsnoer project is logical extension

- Last year, the Dutch Academic IBD Group ICC and IBM developed a business case for a national biobank on Inflammatory Bowel Diseases.

- The greatest strength of this biobank is the ability to extend the current local patient data collection to a national or even international data collection. This would imply a significant increase in the reliability and probability of statistical evidence for future medical research.

- This business case showed, that indeed such a biobank could be a self-funding operation, subject to a significant size of patient groups, suppliers and clients of the intended biobank.

- This plan was taken a step further by the NFU. Could this idea of a self-funding operation to reach national coverage, and which turned out to be true for the ICC, be replicated for other disease areas?

- This resulted in the initiative called ‘het Parelsnoer’ to be born. The initiative aims to realize an integrated environment in and between the eight UMC’s in order to be able to collaborate and exchange clinical and research information for eight disease areas and research lines.
Clinical Genomics & Biobanking: Other Project Examples

**The Mayo Clinic:** Transform the effectiveness and economics of health care, by focusing on new techniques to harness patient data to improve diagnoses; deep computing power to model diseases to find cures; and new devices to access information to transform how patients and physicians interact, leading to more personalized care.

**Karolinska Institutet:** Link multiple “biobanks” at KI to accelerate the understanding of the underlying mechanisms of disease, conduct more focused clinical research and ultimately transform healthcare delivery through information-based medicine.

**iCapture/University of British Columbia:** Determine the relationships between genomic, phenotypic, and environmental data in relation to one of the developed worlds leading killers - heart, lung, and blood vessel disease.

**Moffitt Cancer Center:** Translate new molecular technologies into accelerated, personalized cures for leading causes of cancer death including lung cancer by leveraging an integrated clinical genomics information infrastructure.

**The Cleveland Clinic:** Accelerate the rapid translation of scientific discoveries to better, more targeted patient care for Abdominal Aortic Aneurysms by depicting associations within the clinical and research data that helps the researcher formulate or explore new hypotheses of the genetic and biological basis of disease.
THANK YOU!

Scientific discovery
New drugs and treatments
Transformation in healthcare

Medicine

Information Technology
IBM Healthcare and Life Sciences innovation that matters

http://www.ibm.com/industries/healthcare