Proceedings

SY03.02 | Data Integration and Modelling

Biobank Semantic Information Management With The Health Intelligence Platform

C. Seebode*, M. Ort†, S. González Álvarez‡, C.R.A. Regenbrecht§

1ORTEC medical GmbH, Berlin, Germany; 2Institute for Pathology Charité Universitätsmedizin Berlin, Berlin, Germany

Introduction/ Background

Traditionally Biobanks were mostly used as repositories of cells and tissues. They are addressing scientific questions mainly for retrospective trials where stored tissues and cells can be used. This paradigm has been changing lately. Biobanks can play a different and more prominent role in the scientific process. In fact, the information contained in biobanks can fuel the scientific reasoning itself more than just accompanying it with management of specimens and data. Information management for biobanks should integrate all kinds of information from various sources including clinical records and documents to support flexible retrieval and analytics based on that information. One of the key concepts to support this is to treat the sample like the patient or donor. Selecting a sample for a research trial may be done based on real-time data from the clinical processes. This closes the loop and opens the iron curtain between clinical processes and clinical research. It is a key requirement for supporting research for precision medicine allowing for dynamic decisions and trial designs. Another important aspect is collaboration and sharing of important information preserving data security and patient safety. Informed consent should be dynamically integrated into the scientific process and support patient/donor literacy for the scientific question.

Aims

More precisely we define requirements for biobank management systems to be:

• Collaboration and global sharing of data;
• Support for Scientific workflows and knowledge management;
• Association of genotype and phenotype data and Integration of patient records;
• Integration with registries and standards integration;
• Support for informed consent.

Methods

We present a biobank architecture based on our product ‘Health Intelligence Platform (HIP)’ that supports these requirements. We describe a way to support scientific reasoning directly with HIP and integrate sample management in a way that it contributes to scientific reasoning and outcome. The NLP semantic information extraction together with information extraction from (semi)structured sources provide a necessary integration to support the link between sample and clinical information. The semantic convergence model of HIP supports semantic queries based on an integrated semantic information base. Selecting samples or running analytics based on this information is possible in real time. The flexible knowledge management supports quick adaption of the knowledge base. Researchers are able to adapt the ontologies by visual interaction with the data sources and semantic tagging. The semantic workbench supports roundtrip engineering of ontologies and conflict resolution. We will present an evaluation framework that compares traditional biobank processes with HIP biobank management. We present the integration capabilities and provide an outlook to allow patient participation in research to address upcoming issues in precision medicine.
Conclusion:
Biobank management with HIP supports biobank management by providing tools and services for horizontal (across stakeholders, workflows and collaborations) and vertical (across institutions and disciplines) semantic integration of data into one common data model. HIP biobank management supports semantic retrieval of sample data and associated scientific and patient centered information. The HIP Core is used already and is used to identify patients for clinical trials based on information from clinical records and documents.