# An Integrated Information Platform for a Biomedical Research Network: Concept and First Experiences

**Petra Knaup[a], Christian D. Kohl[a,b], Justo Lorenzo Bermejo[a], Hartmut Dickhaus[a], Meinhard Kieser[a]**

[a] *Institute of Medical Biometry and Informatics, University of Heidelberg, Heidelberg, Germany*
[b] *Coordination Center for Clinical Trials, University of Heidelberg, Heidelberg, Germany*

### Abstract and objective

*In biomedical research a variety of data like clinical, genetic and proteomic data are processed to gain new insights in diseases and therapies. In a transregional research network (TRN) different projects work together with data in different resources and in different formats. Providing an IT infrastructure that integrates these data enables cross-project analyses and provides an overview of available data and resources (blood, tissue etc.). For a German TRN on liver cancer we develop an integrated information platform for research data of 22 projects. Data will be stored together with meta-data to enable a generic approach. Major challenges are harmonization efforts on procedures and data structures as well as security issues.*

***Keywords:***

Biomedical research, Information platform, Data standards and ontologies

## Methods

The information platform will be built up in an iterative process. We start with a case study in two selected project groups and analyze their data structures, identify overlapping data and ambiguous data structures. We will harmonize the terminology and develop a common data set. We will compare the results with current bioinformatics standards and validate their applicability for the information platform. After these first steps further TRN projects can apply for integrating their data. Existing specifications and terminologies will be continuously enhanced and transformed into generic specifications. Standard operating procedures will be developed, e.g. for transformation of local data structures and meta-data. An authorization concept will be developed as well as a tool for the management of data requests.

## Results

In the information platform meta-data have to be generated and stored together with the project data. The information platform contains exclusively pseudonymous data. Data describing the same subjects will be joined. We will use a flexible approach: Data are stored together with meta-data, which describe their structure. Thus, the architecture can be easily enhanced for new data structures and research questions.

All projects will be requested to submit meta-data that describe their local information resources. Therefore, it becomes transparent to all TRN projects which data and tissue samples for joint analyses might be available in which research project. TRN research projects that want to share their data in the common information platform can apply for the integration process. Two committees have to be constituted:

1.) Terminology and Data Harmonization Committee: Generic concepts for terminology, data structures and meta-data have to be agreed on and approved by all TRN projects.

2.) Access Authorization Board: Projects can apply for analyzing the data. They specify the research question and the necessary data. When the board permits access, the ‘owner’ of the data will be informed and maintains a right to veto.

For analysis the requested data are extracted from the information platform und transmitted to the requesting project.

Several tools will be necessary: a management system to support harmonization, integration interfaces, provision of central SOPs, a system that processes and integrates meta-data and research data (pre-processing, data validation, storage).

## Discussion and Conclusion

Medical informatics can provide resources to support biomedical research. Nevertheless, there is a lack of interoperability among data resources from independent institutions. This heterogeneity may hamper the development of new interventive strategies and their translation from bench to bedside. Providing an IT infrastructure that integrates the data of the different projects can enable cross-project analyses and provides an overview of available resources. Major challenges when establishing an integrated information platform for research networks are harmonization efforts on procedures and data structures as well as security issues.

### Acknowledgements

The TRN: ‘SFB/TRR 77 – Liver Cancer. From molecular pathogenesis to targeted therapies’ is supported by the German Research Foundation (DFG).