

Knowledge Management System in Preclinical Radiooncology / Radiobiology Research

Wahyu W. Hadiwikarta¹, Nadja Ebert^{1,2}, Mareike Roscher¹, Ina Kurth¹ and Michael Baumann¹

¹ Division of Radiooncology / Radiobiology, Deutsches Krebsforschungszentrum – German Cancer Research Center,

Im Neuenheimer Feld 280, 69120 Heidelberg, Germany

² Department of Radiotherapy and Radiation Oncology, Faculty of Medicine and University Hospital Carl Gustav Carus, Technische Universität,

Fetscherstraße 74, 01307 Dresden, Germany

w.hadiwikarta@dkfz-heidelberg.de

Abstract. Animal research is integral in the settings of clinical research as it provides preclinical evidence to support clinical studies. Unfortunately, these preclinical information and knowledge are not always readily available and presented through a proper knowledge management system. In this poster publication, a concept for such system, that is able to support radiotherapy research for cancer treatment is described. To optimize the value of the data, the concept of Linked Data and the Semantic Web technology is utilized and the interface for scientific users to query through standard SQL query and the more advanced SPARQL query is made available. External applications can connect through WEB-API support of the system.

Type of submission. This poster describes a concept of a software system.

Keywords: cancer, preclinical, animal experiment, knowledge management, semantic web.

1 Introduction

Radiooncology and radiobiology research works on the translational aspects of integrating biological findings of cellular radiotherapy responses into clinical radiooncology studies and to uncover the biological mechanisms behind the observed clinical responses of radiotherapy or combined radiotherapy treatments. Henceforth, knowledge of preclinical research from animal experiments is one of the most valuable assets, particularly when aggregated e.g. compared and combined, with clinical radiotherapy data. From animal welfare aspects and the fastest possible use of research data for clinical applications, it is absolutely essential and necessary to extract as much information as possible from a preclinical translational dataset. Such a knowledge base offers continuity and data security; complete data on a trial is obtained and serve as a planning basis for future trials, avoids redundant experiments and provides a data pool for novel re-analysis. Institutional-level to multicentric pooling of raw preclinical data become an

years. We are aware of the existence of the Animal Study Registry (ASR) [4], launched in January 2019. It is operated by the German Centre for the Protection of Laboratory Animals (Bf3R) and the German Federal Institute for Risk Assessment (BfR). However, there is still an urgent need for a knowledge management system that go beyond a registration system, especially in the field of cancer radiotherapy and radiobiology. This increases the motivation for the research and development of a preclinical knowledge management system as described in this publication.

2.2 Ontology

Ontology is central in the development and utilization of a knowledge base. The *Open World Assumption* (OWA) highlights the importance of reusability of domain specific ontologies. Dekker et al. [5] proposed the *Radiation Oncology Ontology* (ROO) optimized for clinical radiotherapy cancer research. Upon examination, this is currently one of the many ontologies that is proximal into our use-case. Nevertheless, development of extension is required to accommodate preclinical research settings.

2.3 Clinical Trial Platform RadPlanBio

The German Cancer Consortium (DKTK) as a joint initiative of more than 20 institutions and university hospitals, under the umbrella of the German Cancer Research Center (DKFZ) has started the project for the development of a clinical trial data management platform ‘RadPlanBio’ to support multicentric cancer radiotherapy studies in 2012 [6]. In the time of this writing, the platform is running dozens of multicentric clinical studies across DKTK sites and external partners beyond Germany. Currently, it is under development to expand its support for preclinical studies data management. The knowledge management system that is described in this poster publication will be a further expansion to the RadPlanBio system.

3 Concepts

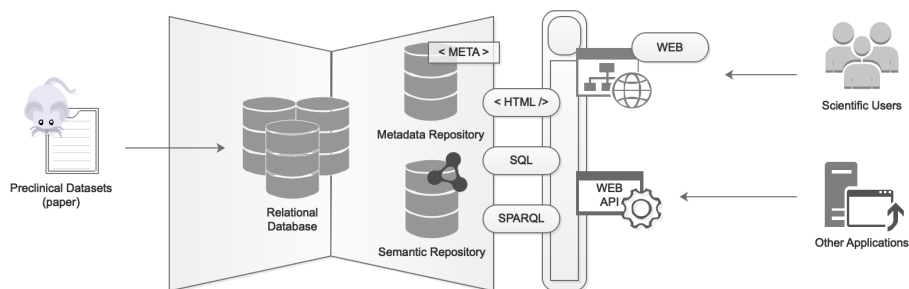


Figure 2. Concept illustration of the architecture of the knowledge management system that is currently under development

The knowledge management system (see Figure 2 for the concept illustration of the architecture) that currently under development is comprised of three central repositories. The first repository is a relational database system that stores the preclinical datasets in digital format. Currently, efforts of digitization from paper are done through manual labor. The scheme in this relational database accommodates the requirement that data need to be stored in a structure reflecting the uniqueness of each project study protocol, the category of tumor model and the model of treatment. Each of the entered data, models an animal subject tagged by a unique subject identifier. The identifier will allow identification not just of the unique subject, but also of the treatment cohort and the type of treatment or drug combination used for this particular unique mouse subject.

The second repository is a metadata repository. To ensure consistency of the study parameters and their meaning across projects, we employ a metadata repository that operates as a reference of registered annotations for the preclinical parameters used in a study. Knowing that different study projects can have different parameters, including novel parameters, the metadata repository is constantly updated for new information.

The third repository is a semantic repository comprising RDF stores exposing the preclinical datasets to a set of ontologies that links multiple data sources e.g. projects, that will allow for potential new knowledge discovery. The semantic repository supports the implementation formats of the RDF/RDFS/OWL standards.

To interact with the repositories, the system supports a user web interface as an application accessible through the web browser which will allow a scientific user to do a query on the database by using SQL query and in the same way to the knowledge base by using SPARQL query. To increase usability on the interface, a translator from natural language i.e. English, to SPARQL will be utilized. The system will also have WEB-API support for external applications to connect and to access the data and the knowledge base.

4 Conclusions

Preclinical research data has proven to be indispensable in translational cancer research. Unfortunately, guidelines for proper practice of preclinical cancer data and knowledge management in radiotherapy are scarce. Meanwhile, animal research itself has been under pressure over the last years. Every time new regulations come out, the procedure and the required ethical administration to do animal experiment become more stringent. Hence, there is an urgency to have a preclinical knowledge management system that may support future preclinical research in cancer radiotherapy. The existence of such system may potentially reduce the scale of required preclinical study results and hence the number of animals needed as redundant experiments no longer occur and even more prediction for radiotherapy responses can be acquired through machine-supported analysis on the knowledge base.

References

1. Aria, M., Cuccurullo, C.: bibliometrix: An R-tool for comprehensive science mapping analysis, *Journal of Informetrics*, 11(4), 959-975 (2017).
2. Web of Science Homepage, <http://www.webofknowledge.com>, last accessed 2020/05/14.
3. Web of Science Support Page, https://support.clarivate.com/ScientificandAcademicResearch/s/article/KeyWords-Plus-generation-creation-and-changes?language=en_US, last accessed 2020/08/14.
4. Bert, B., Heintz, C., Chmielewska, J., Schwarz, F., Grune, B., Hensel, A., Greiner, M., Schönfelder, G.: Refining animal research: The Animal Study Registry, *PLoS Biology*, 17(10), e3000463.
5. Traverso, A., van Soest, J., Wee, L., Dekker, A.: The radiation oncology ontology (ROO): Publishing linked data in radiation oncology using semantic web and ontology techniques, *Medical physics*, 45(10), e854–e862 (2018).
6. Skripcak, T., et. al.: Creating a data exchange strategy for radiotherapy research: towards federated databases and anonymised public datasets, *Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology*, 113(3), 303–309 (2014).