Building a Knowledge Base for Variant Annotation using Therapy Recommendations in cBioPortal

Dominik BOEHM [a,b,1], Niklas REIMER [c,d], Jan CHRISTOPH [e], Alexander SCHEITER [f] and Philipp UNBERATH [a,g]

[a] *Medical Center for Information and Communication Technology, Universitätsklinikum Erlangen, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany*

[b] *Bavarian Cancer Research Center (BZKF), Erlangen, Germany*

[c] *Group for Medical Systems Biology, Lübeck Institute of Experimental Dermatology, University of Lübeck, Germany*

[d] *Medical Data Integration Center, University Hospital Schleswig-Holstein, Luebeck, Germany*

[e] *Junior Research Group (Bio-)Medical Data Science, Faculty of Medicine, Martin-Luther-University Halle-Wittenberg, Halle (Saale), Germany*

[f] *Institute of Pathology, University of Regensburg, Regensburg, Germany*

[g] *SRH Fuerth University of Applied Sciences, Fuerth, Germany*

ORCiD ID: Dominik Boehm https://orcid.org/0009-0001-2887-7109, Niklas Reimer https://orcid.org/0000-0002-0491-3929, Jan Christoph https://orcid.org/0000-0003-4369-3591, Alexander Scheiter https://orcid.org/0000-0001-6734-261X , Philipp Unberath https://orcid.org/0000-0002-1269-9360

Abstract. Molecular tumor boards present special challenges when it comes to information collection for case preparation. It is one of the most time-consuming tasks participating pathologists and oncologists face, limiting the number of cases that can be discussed in these specialized tumor boards and in turn can profit from a potential highly personalized therapy. Digital support is a necessity to enable medical professionals to efficiently make use of the vast amount of data available for each patient and their genomic and clinical profile. This includes historically recommended therapies for patients with molecularly similar tumors. To combat this issue, we developed an extension for the MTB-cBioPortal in collaboration with clinicians, enabling users to access previously documented therapy recommendations combined with corresponding follow-up data based on established MII HL7 FHIR profiles and modules. The information is made available through an additional annotation in the MTB-cBioPortal patient view. In doing so we intend not only to improve the efficiency of the case preparation process for molecular tumor boards, but also lay the groundwork for a potential multicentric exchange of anonymized therapy recommendations and follow-up data.

Keywords. Molecular Tumor Board, Therapy Recommendation, Personalized Medicine, Oncology, Genomic Alteration, cBioPortal, Patient Similarity

[1] Corresponding Author: Dominik Böhm, dominik.boehm@uk-erlangen.de

1. Introduction

Molecular tumor boards (MTBs) incorporate medical data from various information systems, including local data storage as well as multiple external databases. Searching for and joining all this information proves to be one of the most resource-intensive steps in the preparation of MTB cases [1]. This is mostly due to the fact that information collection is currently primarily a manual step and requires clinicians to not only search multiple online databases like OncoKB[2], JaxCKB[3] or ClinicalTrials.gov for relevant literature and annotations as well as fitting clinical trials, but also use local data storage to assess whether genetically and clinically similar patients have been treated previously. In practice, especially the local data is often disregarded for MTB annotation, due to the lack of structured storage and informatic solutions to make previous cases searchable. While the tools in use, as well as the documentation storage and recollection approaches differ from site-to-site, recent analyses have solidified our assumption that the challenges clinicians face in the preparation of MTB cases are very similar across German hospitals. While several previous projects have aimed to solve parts of these issues and created targeted solutions, many problems remain unsolved and resulting in a lock-in of valuable knowledge of therapeutic options at individual physicians and single sites.

2. Methods

2.1. MTB-cBioPortal

The MTB-cBioPortal (previously MIRACUM-cBioPortal) itself is an extension of the popular genomic data analysis and visualization platform cBioPortal[4]. It was developed in the context of the MIRACUM consortiums use case 3 project which aimed to provide digital support for personalized medicine through molecular tumor boards [5,6]. The Extension includes several features that our work leverages. First and foremost is the ability to document therapy recommendations in a structured manner. The information can be documented and submitted through a user interface (see Figure 1) and is then translated into FHIR Resourced by the FhirSpark component, leveraging the Molecular Genomics Report FHIR profiles by the Medical Informatics Initiative Germany (MII) [7]. This presents the base of our extension, as the availability of structured documentation for therapy recommendations and follow-up data is a prerequisite for the recollection of recommendations based on targeted alterations.

Figure 1. MTB-cBioPortal Documentation Interface for Therapy Recommendations.

2.2. Determination of similarity measures

Determining the relevant genomic and clinical similarity of patients to enable a useful selection of previous therapy recommendations, is in itself a difficult task. Even though there is a high number of potential parameters available, discussions with several MTB-participants led us to the decision to use matching genomic alterations, including gene and protein change to represent genomic similarity. This simple approach allows us to easily identify most of the relevant therapy recommendations without the potentially error prone calculation of a proprietary score. While other clinical and genomic data points are generally of relevance when evaluating previously applied therapeutic approaches, in this context they are secondary and should be provided as supplementary information instead of being a selection criterion. As differing tumor entities might offer insights into the applicability of certain therapies this information is highlighted in the visualization, however it does not influence what therapy recommendations are shown.

2.3. Integration into the MTB-cBioPortal

The mutation tables located in cBioPortal's patient view already offer several annotations based on specific genomic alterations – e.g. OncoKB and CIViC –, presenting a fitting entry point for additional information. Besides the required front-end components, the MTB-cBioPortal middleware FhirSpark needed to be altered with additional interfaces for alteration-based therapy recommendation and follow-up requests to allow the MTB-cBioPortal frontend components access to the required information from the underlying FHIR server.

3. Results

The additional annotation is accessible through the mutation tables in cBioPortal's patient view (Figure 2). Availability of previous therapy recommendations is shown through a leveled color-coded helix Symbol (In decreasing order: green – matching alteration and tumor entity, yellow – only matching alteration, red – only matching diagnosis, grey – no matches). This coding information together with a short explanation is also available in the corresponding column legend. On mouseover the tooltip offers an *Overview* of the available information, including the number of data points with matching alterations and diagnoses for locally available previous MTB documentation.

Figure 2. Overview of Available Local Therapy Recommendations.

The secondary tab *Local* (Figure3) presents detailed information on the available therapy recommendations. Besides columns with the identifying *Alteration* and *Cancer Type*, the

clinical data column contains clinical information that has been included as reasoning for the recommended therapy. The *Therapy Response* column contains a summarization of the available follow-up data in the form of Response evaluation criteria in solid tumors (RECIST)[2] for 3/6/12 months. A search bar is available that allows users to search for specific alterations, cancer types or therapy responses through free-text matching. One of the key take-aways from the discussions with clinicians was that while a quick overview on a specific recommendation in the form of a table was helpful, more information is necessary to actively use the previously recommended therapies for future patients. To solve this issue locally documented patients are linked to be accessible via cBioPortal, allowing clinicians to revisit cases with all their clinical and genomic data if necessary.

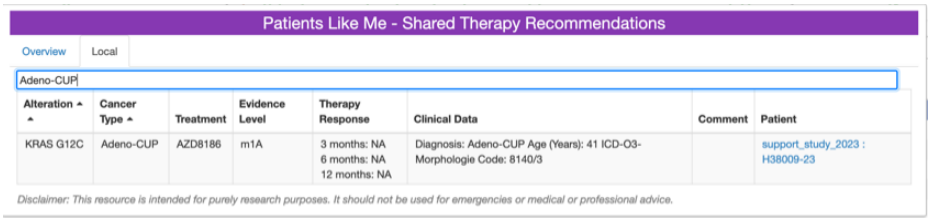

Figure 3. Table Visualization of Locally Available Therapy Recommendations.

The annotation has been incorporated into the 2024q1 release of the MTB-cBioPortal project and is available on Github[3].

4. Discussion and Conclusions

This work features an extension of the MTB-cBioPortal with an additional annotation that features previously recommended therapies based on a patient's molecular tumor profile. One must note that while effective, exclusively using matching alterations as a similarity measure is a significant simplification and does not encompass all potentially relevant cases. The determination of a more complete metric is an objective for future work. Current projects like PM4Onco[4] include work aiming to identify appropriate patient similarity measures for the molecular tumor board context [8].

The implementation also lays the groundwork for further developments, aiming at incorporating data from multiple MTBs to improve data availability and increase coverage for rare diseases and molecular profiles. This is primarily focused on a national exchange, as our extension expects the usage of the established MII-FHIR profiles. When sharing therapy recommendations, data quality is one of the core concerns. Due to missing automated data integration, caused by heterogeneous IT-infrastructures and the resulting need for proprietary ETL-processes to support MTB applications, clinical and genomic information tend to be copied manually. This leads to transitory errors, missing information and a general decline in data quality. Lack of access to the underlying data sets limits the trust in the available shared data and in turn leads to significantly reduced usability for participating clinicians. These circumstances pose the question whether such information can and should be used to inform treatment decisions. However, we

[2] https://recist.eortc.org/.
[3] https://github.com/buschlab/Mtb-cbioportal.
[4] https://pm4onco.de/.

expect that with good documentation and standardized procedures including rigorous data curation these issues can be mitigated in the future. Additionally current national legislature like §64e and the resulting initiatives like GenomDE require similar processes to enable a research infrastructure for genomic data. This provides a significant amount of groundwork to overcome the identified difficulties as well as entry points for future work.

Lastly, while the user feedback collected during development was very positive, a detailed evaluation is necessary to determine and quantify the impact the developed extension has on the preparation of MTB cases and to identify necessary changes to further improve its usability.

Acknowledgment

This work was funded by the German Federal Ministry for Education and Research through the PM4Onco project (01ZZ2322R, 01ZZ2322N, 01ZZ2322S), and by the Bavarian Cancer Research Center (BZKF). This work was performed in (partial) fulfilment of the requirements for obtaining the degree “Dr. rer. biol. hum.” at the Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU).

References

[1] M.-P. Schapranow, F. Borchert, N. Bougatf, H. Hund, and R. Eils, Software-Tool Support for Collaborative, Virtual, Multi-Site Molecular Tumor Boards, *Sn Comput. Sci.* **4** (2023) 358. doi:10.1007/s42979-023-01771-8.

[2] D. Chakravarty, J. Gao, S. Phillips, R. Kundra, H. Zhang, J. Wang, J.E. Rudolph, R. Yaeger, T. Soumerai, M.H. Nissan, M.T. Chang, S. Chandarlapaty, T.A. Traina, P.K. Paik, A.L. Ho, F.M. Hantash, A. Grupe, S.S. Baxi, M.K. Callahan, A. Snyder, P. Chi, D.C. Danila, M. Gounder, J.J. Harding, M.D. Hellmann, G. Iyer, Y.Y. Janjigian, T. Kaley, D.A. Levine, M. Lowery, A. Omuro, M.A. Postow, D. Rathkopf, A.N. Shoushtari, N. Shukla, M.H. Voss, E. Paraiso, A. Zehir, M.F. Berger, B.S. Taylor, L.B. Saltz, G.J. Riely, M. Ladanyi, D.M. Hyman, J. Baselga, P. Sabbatini, D.B. Solit, and N. Schultz, OncoKB: A Precision Oncology Knowledge Base, *JCO Precis. Oncol.* (2017) 1–16. doi:10.1200/PO.17.00011.

[3] S.E. Patterson, R. Liu, C.M. Statz, D. Durkin, A. Lakshminarayana, and S.M. Mockus, The clinical trial landscape in oncology and connectivity of somatic mutational profiles to targeted therapies, *Hum. Genomics*. **10** (2016) 4. doi:10.1186/s40246-016-0061-7.

[4] J. Gao, B.A. Aksoy, U. Dogrusoz, G. Dresdner, B. Gross, S.O. Sumer, Y. Sun, A. Jacobsen, R. Sinha, E. Larsson, E. Cerami, C. Sander, and N. Schultz, Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal, *Sci. Signal.* **6** (2013) pl1. doi:10.1126/scisignal.2004088.

[5] P. Buechner, M. Hinderer, P. Unberath, P. Metzger, M. Boeker, T. Acker, F. Haller, E. Mack, D. Nowak, C. Paret, D. Schanze, N. von Bubnoff, S. Wagner, H. Busch, M. Boerries, and J. Christoph, Requirements Analysis and Specification for a Molecular Tumor Board Platform Based on cBioPortal, *Diagnostics*. **10** (2020) 93. doi:10.3390/diagnostics10020093.

[6] N. Reimer, P. Unberath, H. Busch, M. Börries, P. Metzger, A. Ustjanzew, C. Renner, H.-U. Prokosch, and J. Christoph, Challenges and Experiences Extending the cBioPortal for Cancer Genomics to a Molecular Tumor Board Platform, *Stud. Health Technol. Inform.* **287** (2021) 139–143. doi:10.3233/SHTI210833.

[7] N. Reimer, P. Unberath, H. Busch, and J. Ingenerf, FhirSpark - Implementing a Mediation Layer to Bring FHIR to the cBioPortal for Cancer Genomics, *Stud. Health Technol. Inform.* **281** (2021) 303–307. doi:10.3233/SHTI210169.

[8] I. Manuilova, J. Bossenz, A.B. Weise, D. Boehm, C. Strantz, P. Unberath, N. Reimer, P. Metzger, T. Pauli, S.D. Werle, S. Schulze, S. Hiemer, A. Ustjanzew, H.A. Kestler, H. Busch, B. Brors, and J. Christoph, Identifications of Similarity Metrics for Patients With Cancer: Protocol for a Scoping Review, *JMIR Res. Protoc.* **13** (2024) e58705. doi:10.2196/58705.