

Using real cross-institutional clinical Data to identify Rare Diseases in practice H Storf¹, J Schaaf¹, D Kadioglu¹, M von Wagner¹, M Boeker², C Haverkamp³, H Binder², C Schade-Brittinger⁴, HU Prokosch⁵, T Wagner⁶

- ¹ Medical Informatics Group (MIG), University Hospital Frankfurt, Frankfurt am Main
- ² Institute of Medical Biometry and Statistics (IMBI), Faculty of Medicine and Medical Center, University of Freiburg, Freiburg i. Br, Germany
- ³ Staff Unit IT Processes, Faculty of Medicine and Medical Center, University of Freiburg, Freiburg i. Br., Germany
- ⁴ Coordinating Center for Clinical Trials, Philipps-University Marburg, Marburg, Germany
- ⁵ Department of Medical Informatics, Biometrics and Epidemiology, Friedrich-Alexander University Erlangen-Nürnberg, Erlangen, Germany
- ⁶ Frankfurt Reference Center for Rare Diseases (FRZSE), Frankfurt University Hospital, Frankfurt, Germany

Introduction

The federal ministry of Education and Research in Germany (BMBF) is funding four consortia for establishing data integration centers (DIC) in University hospitals, one of which is MIRACUM (Medical Informatics For Research and Care in University Medicine), with the goal of fostering linkage and sharing of data between the hospitals. One objective is to support the diagnosis of patients without a not diagnosed rare disease (RD) using data of the Consortium on a large scale (> 12 mio. inpatients). Finding the right diagnosis and therapy for patients with rare conditions is challenging, tedious and associated with many doctor visits. The need to fulfill these patient cases is very high [1]. In the last two decades, different IT approaches have been developed to support clinical decision-making on a given diagnosis (diagnostic reasoning) [2, 3]. Whereas those systems rely on data from the internet or disease in order to cluster similar cases. The MIRACUM partners will develop

Visualization modules are developed for the display of data trajectories of similar patients. Clinicians are especially interested in the visualization of the patient's history in a timeline which has been specifically designed for patients with a chronic disease course. Thus, the diagnostic and therapeutic pathway of a patient can be compared visually to "similar patients" [4]. An alternative might also be to visualize global relationships with so-called sunburst diagrams [5].



and validate a machine learning approach to find similar patients based on the

anonymized clinical data of patients stored within the DIC of all MIRACUM sites.



Figure 1: Overview of all MIRACUM sites

Figure 2: Structure of patient similarity analysis

Conclusion

The main result will be the identification of patients with similar medical characteristics to support the diagnosis of rare disease patients. As part of the project, it will become clear if the described approach can be implemented for all rare diseases. Due to the high number of diseases, the focus will initially be limited to a selection of diseases which will be consented with the experts from the rare

disease centers.

References

 [1] European Commission (2000) Regulation (EC) No 141/2000 of the European Parliament and oft the council of 16December 1999 on orphan medicial products. Official Journal of the European Communities. http://ec.europa.eu/health/ sites/health/files/files/eudralex/vol-1/reg_2000_141_cons-2009-07_en. Accessed: 04. Jan 2018
[2] Dragusin R, Petcu C, Lioma C, Larsen B, Jorgensen HL (2013). FindZebra: a search engine for rare diseases. Int J Med Inform 82(6): 528-538
[3] Koehler S, Schulz M, Krawitz P, Bauer S, Doelken S (2009). Clinical Diagnostics in Human Genetics with Semantic Similarity Searches in Ontologies. Am J Hum Genet 85(4): 457-464
[4] Hirsch JS, Tannenbaum JS, Lipsky Gorman S, Liu C. HARVEST, a longitudinal pa-

Conception

The aim of this project is to design, build and validate a software platform to support the decision-making on the diagnosis of in so far undiagnosed patients. Within the MIRACUM Consortium there are 6 Centers for Rare Diseases at its partner sites. A key element in the implementation is the early and continuous involvement of end users, e.g. collaborating with physicians who focus on undiagnosed diseases and experts in rare diseases who successively can be consulted to confirm a suspected disease. In a next step, the algorithm for identifying similar patients will be developed. The resulting application will send anonymized records to the DIC to identify patients with similar clinical characteristics.

tient record summarizer (2015). J Am Med Inform Assoc 22(2): 263-274

[5] Hripcsak G, Ryan PB, Duke JD, Shah NH, Park RW (2016). Characterizing treatment pathways at scale using the OHDSI network. Proc Natl Acad Sci USA 113 (27): 7329-7336