Fachhochschule Dortmund
University of Applied Sciences and Arts

WisPerMed
Wissens- und datenbasierte Personalisierung von Medizin am Point of Care

UNIVERSITÄT DUISBURG ESSEN
Offen im Denken

RP6 - Analysis of genome data

Research question

Personalized representation of the reliability of genome variants

Diagnosis and therapeutic decisions in oncology are based on gene variants or tumor mutational burden [1,2].

Technology: Currently gene panels (few genes, deep sequencing); soon genome sequencing (WGS)

WGS advantage: More options for therapy decisions

WGS disadvantage: low and heterogeneous sequencing depth; observed variants of varying reliability

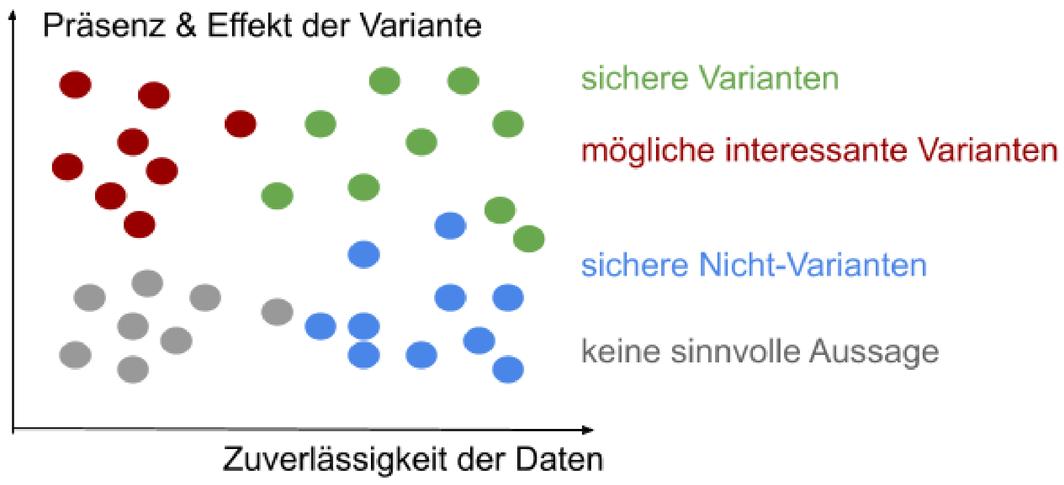
Question: How can we distinguish between

- „safe non-variants“ und
- „non-safe variants“ ?

Objective: Doctors at the point of care can prioritise relevant reliable variants and probably exclude ineffective therapy options.

Solution approach

One-dimensional representation (p-value or a posterior probability) does not do justice to the problem,
equally non-high-dimensional representation from many quality indicators (information overload).
→ Development of a 2D to 3D representation



Program: Development of intuitively understandable, visualisable quantitative measures for both dimensions; development of communication and presentation at the PoC; support on questions of acceptance by FP 11

State of the art

Numerous methods and tools exist for variant identification from genome data [3,4,5,6].

However: No method reports "safe non-variants". No method distinguishes between "safe non-variants" and "non-safe variants”:

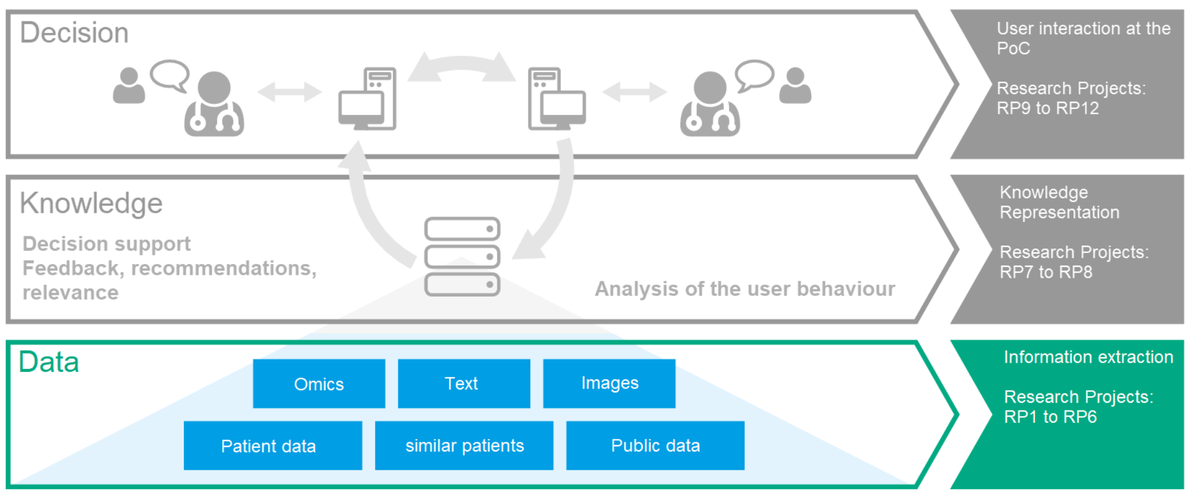
- GATK (Broad Institute): many different quality indicators, several threshold values such as for coverage, beach bias lead to information overload [6].
- Bayesian method (e.g. FreeBayes [4], Varlociraptor [5]):

  a-posteriori probability for the presence of the variant; low for little data and indistinguishable from a non-variant.

We have extensive experience in the presentation and interpretation of variants [1,2,5].

Integration

- Primary domain: Information extraction from omics data
- Communication of information via the Smart Hospital Information Platform (SHIP) at the point of care.
- Prospective testing of the use of information in the Molecular Tumor Board in melanoma patients* with existing exome or WGS data
- Integration of the representation in OpenEHR model for the patient file



Prof. Dr. rer. nat.
Sven Rahmann

Universitätsklinikum Essen
Universität Duisburg-Essen

Prof. Dr. med.
Dirk Schadendorf

Universitätsklinikum Essen

Literatur

1. Schramm A., Köster J., Assenov Y., Althoff K., Peifer M., Mahlow E., Odersky A., Beisser D., Ernst C., Henssen A. G., Stephan H., Schroder C., Heukamp L., Engesser A., Kahlert Y., Theissen J., Hero B., Roels F., Altmuller J., Nurnberg P., Astrahantseff K., Gloeckner C., De Preter K., Plass C., Lee S., Lode H. N., Henrich K. O., Gartlgruber M., Speleman F., Schmezer P., Westermann F., **Rahmann S.*,** Fischer M.*, Eggert A.*, und Schulte J.H.* (*: gemeinsame Letztautorenschaft).
   “Mutational dynamics between primary and relapse neuroblastomas”. *Nature Genetics* Aug 2015, 47(8):872-877.
2. Van Allen EM, Miao D, Schilling B, Shukla SA, Blank C, Zimmer L, Sucker A, Hillen U, Foppen MHG, Goldinger SM, Utikal J, Hassel JC, Weide B, Kaehler KC, Loquai C, Mohr P, Gutzmer R, Dummer R, Gabriel S, Wu CJ, **Schadendorf D**, Garraway LA. “Genomic correlates of response to CTLA-4 blockade in metastatic melanoma”. *Science* 2015 Oct 9;350(6257):207-211.
3. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. “The variant call format and VCFtools.” *Bioinformatics* 2011;27:2156–8.
4. Garrison E, Marth G. “Haplotype-based variant detection from short-read sequencing”. *arXiv*:1207.3907 (2012)
5. Köster J**, Diskstra L J, Marschall T**, Schönhuth A. “Enhancing sensitivity and controlling false discovery rate in somatic indel discovery”. Preprint. *bioRxiv* 741256 (doi: https://doi.org/10.1101/741256; **: ehemalige Doktoranden von S. Rahmann)
6. DePristo M, Banks E, Poplin R, Garimella K, Maguire J, Hartl C, Philippakis A, del Angel G, Rivas MA, Hanna M, McKenna A, Fennell T, Kernytsky A, Sivachenko A, Cibulskis K, Gabriel S, Altshuler D, Daly M. “A framework for variation discovery and genotyping using next-generation DNA sequencing data”. *Nature Genetics* 2011, 43:491-498