**Kolloquium „Statistische Methoden in der empirischen Forschung“**

Wann: 08. November 2016, 17:00 – 18:30 Uhr

Wo: Robert Koch-Institut | Nordufer 20 | 13353 Berlin (Wedding),
S41, S42, U9 Westhafen | U9, Bus 142 Amrumer Str

**Jörg Rahnenführer (Universität Dortmund)**

## Automatic model selection with MBO: A bridge between high-dimensional survival analysis end efficient model optimization

Model selection is a common task in survival analysis. For high-dimensional genetic data as covariables in survival models, the selection of an optimal model is challenging with respect to (i) the large number of potential models, (ii) long computation times of single models, and (iii) the required tuning of hyperparameters for many models. Exhaustively trying out all models and all parameter combinations is computationally unfeasible, especially with continuous hyperparameters.

We present a modern promising approach to this problem setting called model based optimization (MBO). MBO moves efficiently through the model space and simultaneously tunes hyperparameters. A so-called surrogate regression model is fitted, with models as predictors and prediction performance as target variable. MBO iteratively suggests new models with an exploration-exploitation trade-off that prefers models with either good performance estimates or high performance uncertainty.

We apply MBO to cancer data sets with high-dimensional gene expression covariates. Overall MBO outperforms all single algorithms with autotuning of parameters as well as a classical benchmarking approach that can be regarded standard in today's practice.