

## Practical experiences on the necessity of external validation

König IR, Ziegler A

*Institut für Medizinische Biometrie und Statistik, Universität zu Lübeck, Deutschland  
inke.koenig@imbs.uni-luebeck.de*

In developing a prognostic model, data dependent methods are usually utilized to optimize the fit in the data at hand. For example, using support vector machines, suitable function parameters for the kernel functions need to be selected. Similarly, the parameters in a logistic regression model are optimally fit to the data. This data dependent optimization increases the risk of overfitting which is even increased with small sample sizes [1].

To estimate the extent of overfitting, a prognostic model often is internally validated. Here, the same data set is artificially split into separate samples in which model development and testing can be performed. Prominent examples include tenfold cross-validation or bootstrapping. By contrast, a stringent model validation requires at least two independent data sets. Depending on the second data set, temporal and external validation are distinguished [2]: temporal validation is based on data points that have been collected at the same centers but at a later time point, external validation uses data points from different centers.

In the presentation, different techniques for validating a prognostic model are described and discussed. Prognostic models that were developed using a number of ensemble methods are compared with regard to differences in temporal and external validation. For this, exemplary data sets to predict functional independence in stroke patients are utilized [3]. Our results demonstrate that classical internal validation techniques are insufficient to estimate prediction quality in temporal validation data. Furthermore, temporal validation only poorly predicts the external generalizability of the model.

### References:

- [1] Harrell FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996; 15: 361-87.
- [2] Altman DG, Royston P. What do we mean by validating a prognostic model? *Stat Med* 2000; 19: 453-73.
- [3] The German Stroke Collaboration. Predicting outcome after acute ischemic stroke: An external validation of prognostic models. *Neurology* 2004; 62: 581-85.