Machine Learning algorithm in liver allocation – a promising approach?

P. Ritschl[1], B. Bokelmann[2], L. Wiering[3], F. Butz[1], T. Dziodzio[1], M. Biebl[1], G. Lurje[1], U. Schoeneberg[2], W. Schöning[1], B. Globke[1], I. Sauer[1], M. Schmelzle[1], J. Pratschke[1], R. Öllinger[1]

[1]Charité - Univeristätsmedizin Berlin, Chirurgische Klinik CCM/CVK, Berlin, Germany, Charité - Universitätsmedizin Berlin; [2]Charité - Univeristätsmedizin Berlin, Institut für Biometrie und Klinische Epidemiologie, Berlin, Germany; [3]Charité - Univeristätsmedizin Berlin, Medizinischen Klinik mit Schwerpunkt Hepatologie und Gastroenterologie, Berlin, Germany

Introduction

Most liver allocation systems worldwide are urgency depended, defined by the “Model for End-stage Liver Disease” (MELD) score. These MELD-based liver allocation systems have several disadvantages, as MELD allocation i) is not equally fair for all transplant indications ii) is disadvantages for women iii) may lead to excess mortality in minority candidates iiii) is negatively correlated with post transplantation outcome. The aim of this project is to generate a potent liver transplantation outcome score via machine learning algorithms in 2 different cohorts (Germany and US).

Methods

The datasets used for this analysis were obtain from United Network for Organ Sharing (UNOS) and eurotransplant cooperation (ET). As machine learning method we chose “random forests” (rf) and compared that to logistic regression (lr) and well established outcome scores like the BAR- and a modified SOFT- score. Primary outcome was 3 month survival of the transplant recipients.

Results

A total of 104799 liver transplant recipients were included in this analysis. In both countries the cohorts were split in training sets (80% of all included patients; US: n= 75411; Germany: n= 8429) and test sets (20% of all patients). Applying machine learning achieved a prediction of 3 month survival with an area under the curve (AUC) of 0,68 in Germany and the US (**Figure 1+2**).

This forecasting was comparable with logistic regression (Germany: 0,67; US: 0,69). In the US cohort, where more variable were available, the AUC of SOFT (0,67) and BAR (0,62) were calculated as well, showing no or minimal benefit of the “random forests” method. In general, all predictions showed relatively low accuracy with AUC below 0,7.

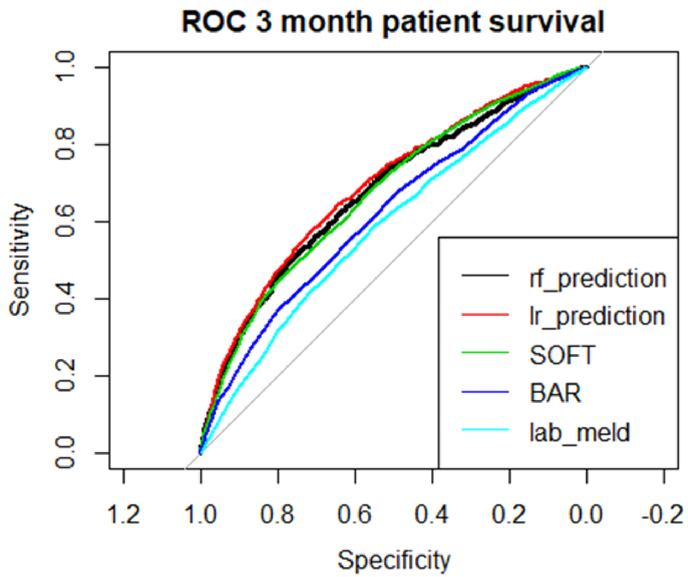

Figure 1. Prediction of 3-month patient survival in the UNOS data set (rf = random forest; lr = logistic regression)

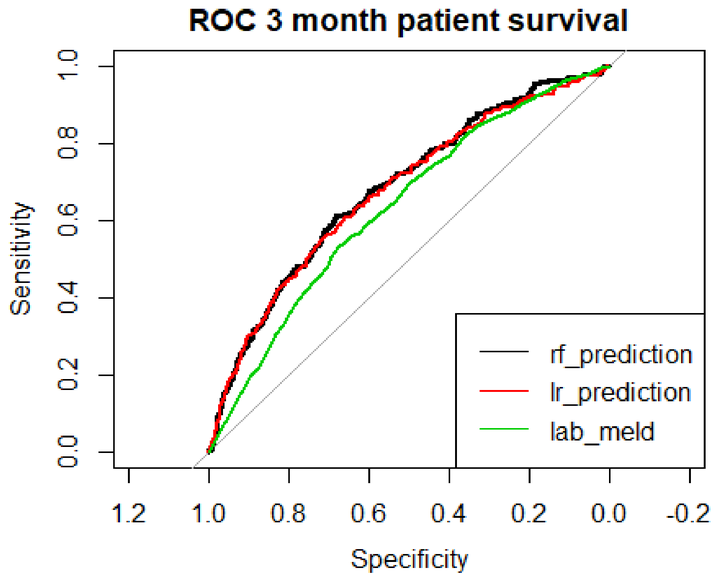

Figure 2. Prediction of 3-month patient survival in the German data set (rf = random forest; lr = logistic regression)

Summary and Conclusion

Machine learning methods as well as conventional outcome predictions show yet insufficient precision if the data sets are from real-life data registries. This may hinder implementation of such result-oriented scores in the process of organ allocation.

paul.ritschl@charite.de

Paul Ritschl
Department of Surgery, Campus Charité Mitte | Campus Virchow-Klinikum
Head: Univ.-Prof. Dr. med. J. Pratschke