Abstract

This thesis has been written in the framework of a project of the research unit DFG 534 “Statistical modeling and data analysis in clinical epidemiology“ at the Institute of Medical Biometry and Medical Informatics of the University Medical Center Freiburg. It focusses on the investigation of approaches for selecting and modelling time-varying effects in extended Cox models.

In the analysis of studies with long-term follow-up and several covariates, the Cox model as the standard approach may be inappropriate due to violation of the proportional hazards assumption caused by time-varying effects. Ignoring the presence of these time-varying effects results in incorrect models and possibly false conclusions thereof. However, beyond testing to detect the time-dependency of effects appropriate modelling of their shape is at least as important, as wrong shapes can likewise lead to false conclusions.

Several different approaches have been proposed which extend the Cox model to allow for time-varying effects. However, analytical results are hardly available and larger simulation studies on properties of the approaches and advice on which techniques to use are rare. To give guidance on the usability and performance of different techniques for modelling time-varying effects, we compare five recent approaches based on promising techniques such as non-parametric cumulative regression functions, splines and fractional polynomials, including frequentist as well as Bayesian inference, in a real life example and a simulated data set.

The main focus of this work lies on the investigation of the Fractional Polynomial Time (FPT) approach which on top of the Multivariable Fractional Polynomial (MFP) approach tests and models time-varying effects based on fractional polynomials in an extended Cox model. We investigate the properties of the FPT approach using bootstrap resampling and a large simulation study. In contrast to the other investigated approaches, FPT is theoretically and technically feasible for such investigations and allows automated selection and modelling of time-varying effects. Besides evaluating the performance of the test procedure on time-varying effects in terms of type I and II error, we put special emphasis on the appropriate modelling of time-varying effects. This requires the definition of more general criteria for the appropriateness of time-varying effects which take the shape of selected effects into account. Additionally we quantify the distance between true and estimated effects as the area between the two curves. The fit of selected models is evaluated by means of the prediction error.
The results show that the FPT approach holds the nominal significance level and yields satisfactory power for stronger time-varying effects. Based on these results we consider FPT as a suitable approach to improve multivariable models for survival data with long-term follow-up.