

**NONPARAMETRIC ESTIMATION OF HAZARD RATIOS OF CONTINUOUS PREDICTORS IN MULTISTATE MODELS**

Luís Meira-Machado<sup>1</sup>, Carmen Cadarso-Suárez<sup>2</sup>, Francisco Gude<sup>3</sup>

<sup>1</sup> *Department of Mathematics for Science and Technology, University of Minho, Portugal  
Spain.*

<sup>2</sup> *Department of Statistics and Operational Research, University of Santiago de Compostela,*

<sup>3</sup> *Clinical Epidemiology Unit, Hospital Clínico Universitario, University of Santiago de  
Compostela, Santiago de Compostela, Spain.*

An important aim in survival analysis is to assess the possible effect of a set of prognostic factors on the course of a disease. Usually, the effect of the continuous covariates on the log-hazards is often assumed to have a linear functional form. Nevertheless, if the true effect is highly nonlinear, this erroneous assumption of linearity can lead to nonproportional hazards and/or misleading statistical conclusions (bias and decreased power of tests for statistical significance). The relative lack of flexibility of parametric survival models has, in recent years, led to the development of nonparametric regression methods based on additive Cox regression models. This regression has the advantage of not assuming a parametric relationship between the continuous covariates and the response, and eliminates the need for the researcher to impose functional assumptions. To introduce flexibility into the Cox model, several smoothing methods may be applied, but B-splines, smoothing splines or P-splines are being the most considered in this context. The goal of this paper was to propose a flexible approach using P-splines, to allow for nonlinear relationships between continuous predictors and possible outcomes in the multi-state framework. For such models to be directly interpreted, employment of an effect measure, such as the hazard ratio (HR), which summarizes the effects, proves extremely useful. To better understand the effects that each continuous covariate has on the outcome at each transition, results are expressed in terms of smooth HR curves, taking a specific covariate value as reference. Confidence bands for these curves are also derived. The proposed methodology was applied to a database on breast cancer, using a progressive three-state model. The use of the breast cancer data illustrates the advantages of using these flexible methods for analyzing quantitative predictors, providing an effect measure that greatly simplifies the information and yielding important biological insights of the disease. Software (in R) implementing the methods proposed in this paper were developed by the authors.