Modeling Breast Cancer screening data using triple interaction with Structured Additive Regression (STAR) models

Elisa Duarte*
Unit of Biostatistics, Department of Statistics and Operations Research, School of Medicine, University of Santiago de Compostela, Santiago de Compostela, Spain - duarte.elisa@gmail.com

Bruno de Sousa
Faculty of Psychology and Education Sciences, University of Coimbra, Coimbra, Portugal - bruno.desousa@fpce.uc.pt
Carmen Cadarso-Suárez
Unit of Biostatistics, Department of Statistics and Operations Research, School of Medicine, University of Santiago de Compostela, Santiago de Compostela, Spain - carmen.cadarso@usc.es
Vítor Rodrigues
Faculty of Medicine, University of Coimbra, Coimbra, Portugal - vrodrigues@fmed.uc.pt
Thomas Kneib
Institute of Statistics and Econometrics, Dept. of Economics, Georg-August-Universität Göttingen, Göttingen, Germany - tkneib@uni-goettingen.de

In the spatial modelling of the risk of breast cancer from a screening program, the proportion of women who attended the program in each municipality is considered a relevant issue to be taken into account in the analysis. The higher is the attendance rate, more the detection rate approximates to the incidence rate of the disease in the population. Also, the neoplasia is detected earlier having the natural consequence of decreasing the mortality rate. So it is of the utmost importance to include as a covariate of the model the interaction of these three rates: Mortality, Attendance and Detection as a way to find out any effect of these variables in the global model. The model is performed using the Structured Additive Regression STAR methodology, because these models are flexible enough to combine in a structured additive form, effects of different types, like linear and non-linear effects, non-linear interaction between two covariates and spatial effect of a location variable. Although, up to now, it has not been developed an automatic procedure to include triple interaction as a covariate in the predictor equation, this study developed a way to overcome this issue, using the Markov Random Fields used for the spatial effect estimation. The results achieved show a model with better performance. The structured spatial effect that is not significant in the model without the triple interaction, is now significant, pointing clearly municipalities with a higher risk of breast cancer, and municipalities with lower purchase power index associated with lower risk of the disease.
Keywords: spatial correlation, markov random fields, attendance rate, detection rate.