SPACE-TIME MODELS WITH TIME-DEPENDENT COVARIATES

A. Biggeri[1], C. Lagazio[2], D. Catelan[*] and E. Dreassi[*]
[1]Department of Statistics "G. Parenti". University of Florence.
[2]Department of Statistics. University of Udine.

The aim of ecological studies is to describe the relationship between geographical variation of disease risk and concomitant variation in the level of exposure to a particular factor: for example, an environmental agent or a life-style related characteristic. Both disease rates and covariates could exhibit a strong spatial autocorrelation. Ignoring this space structure might produce overestimates of the association. The model proposed by Besag et al. (1991) decompose the extra-Poisson variability in an independent random area effect and a completely spatially autocorrelated random effect. The introduction of these two terms is a way to control for unmeasured variables and overdispersion that can create the impression of a spurious geographic variation of disease rates.

When both rates of disease and covariates are time-dependent another source of ecological confounding could be the different point in time in which response and predicted variables are related. With regard to mortality for lung cancer, for example, the biology of the process of carcinogenesis suggests that more than 10-15 years (the latency time) should run between exposure and mortality: the model should include time-lags to allow for this latency period.

The main goal of the present work is to investigate the contribution of socioeconomic factors to disease outcome in ecological studies, including in the analysis the relevant etiological periods and putting particular emphasis on model robustness and sensitivity analysis.

We propose hierarchical Bayesian models with time dependent covariates, latent periods, time and space random terms, and misaligned exposure-disease data to study the relation between socioeconomic factors and lung cancer mortality in Tuscany in the period 1971-99. The aim is to jointly estimate the degree of association between lung cancer deaths and risk factors and the weights attributable to the different time-lags. We have made use of two kind of socioeconomic factors: a material deprivation index and a level of education.

From an epidemiological point of view the results of the analysis confirm the external evidence of a great homogeneity of the risk for lung cancer in Tuscany with a general reduction in all the municipalities with historical higher rates and an increase in the rural areas.

The results about the estimation of the parameters measuring the plausibility of each time-lag must, however, be interpreted with care. Indeed, the high complexity

of the models implies that data are weakly informative with respect to the prior assumptions at the level of the hierarchy in which the time-lag terms are specified. We have fitted our models putting different prior emphasis on the time-lag terms to check their robustness. The models are robust with regard to the estimates of the regression parameters but little changes are induced in the estimates of the probabilities attributable to the different latency time. The importance of performing sensitivity analysis is highlighted by the substantial modifications in the results of the study when different assumptions on the spatially structured random term are specified.

Prior assumptions in ecological analysis are crucial even for less complex specifications and cannot typically be checked from the data alone; we emphasize the idea of drawing conclusions on the basis of plausibility and consistency of the results with a preliminary descriptive analysis of the data.

Referencias

Besag, J., York, J. C. y Mollié, A. (1991). Bayesian image restoration, with two applications in spatial statistics (with discussion). *Annals of the Institute of Statistical Mathematics* 43:1–59.