MODELLING MULTIVARIATE CANCER RATES WITH A LATENT STRUCTURE MIXTURE MODEL

T.C. Bailey[1]; P.J.Hewson[†2]

[1] *University of Exeter, Devon, UK,* [2] *University of Plymouth, Devon, UK*

[†] E-mail: *paul.hewson@plymouth.ac.uk*

Joint modelling of area disease rates has experienced some recent interest in spatial epidemiology. Motivation includes identification of similarities or differences in the risk distribution of diseases, and 'borrowing of strength' to improve risk estimates for any one of the diseases studied. Amongst models proposed, latent structure models are popular, i.e. multivariate GLMMs with correlated random effects formulated in terms of independent (latent) factors. Typically these assume exchangeability for random effects over areas, but this may be an over-simplification. Accordingly, we superimpose a discrete mixture on the latent structure, so allowing our model to accommodate situations where a fully exchangeable latent structure may be restrictive, yet where knowledge of the structural hierarchy is unknown. We illustrate results of applying this to simultaneous modelling of multiple cancer rates (1999 to 2001) in all primary care trusts (PCTs) in the UK. Using a Bayesian (MCMC) approach, an over-large number of components was initially fitted and then sequentially reduced based on an approximation to the Kullback-Leibler distance between current and most plausible collapsed model. A three latent variable, two component mixture model finally emerged and revealed interesting urban/rural epidemiological distinctions in these data when compared with results from models proposing an exchangeable latent structure for all areas.