Age-space-time models for mapping cancer mortality rates

Goicoa, T.[1,*], Ugarte, M.D.[1], Etxeberria, J.[1], and Militino, A.F.[1]

[1] *Department of Statistics and Operations Research. Public University of Navarre. Campus de Arrosadia, 31006, Pamplona, Spain; tomas.goicoa@unavarra.es, lola@unavarra.es, jaione.etxeberria@unavarra.es, militino@unavarra.es*
[*] *Goicoa, T*

Abstract. *Accurate and precise knowledge about the distribution and evolution of a disease in space and time is crucial to develop health policies and to help researchers to look into risk factors related to the disease. During the last years, the availability of modern computers has made it possible the development of statistical models and estimation techniques to analyze spatio-temporal data. Such spatio-temporal models have been used in disease mapping to study how a disease evolves in space throughout the years. It is common in practice to study age-standardized mortality or incidence risks or rates such that a single measure is provided for the whole region and all age groups. However, if the evolution of the disease is not the same among the age groups, age-specific rates within each region should be provided. In this work, several age-space-time models are considered and fitted to study the evolution of age-specific rates along time in different small areas. Spanish prostate cancer mortality data during the period 1986-2010 will be used for illustration.*

Keywords. *Age-specific space-time modelling; Cancer mortality; CAR models; INLA*

1 Introduction

Disease mapping comprises a collection of statistical tools to describe and analyze the spatial or the spatio-temporal distribution of a disease. A deep knowledge and understanding of such spatio-temporal trends is very important for epidemiologists and health policy makers to allocate the more and more scarce resources, to detect and reduce risk factors and to implement prevention-intervention programmes.

Conditional autoregressive (CAR) models have been widely used to smooth risks in space and space-time (see for instance Ugarte et al., 2006 and Knorr-Held, 2000, for space and space-time models respectively), typically aggregating counts over all age groups in the population, based on the assumption that the region and time effects are the same over all age groups. However, if the age groups are not equally

affected by the disease, this may not be informative, and consequently the geographical and temporal evolution of mortality or incidence is different among the age groups. To deal with the effect of age (or other variables), some research has been conducted to model relative risks or rates by age groups. For example, Dean et al. (2001) consider a CAR model to smooth risks with a region-age interaction term and develop a score test to assess its significance, while Zhang et al. (2006) consider rates and propose a model to describe interaction effects among demographic variables such as age, region effects and temporal effects on colorectal cancer incidences.

In this paper mortality rates by region, time and age groups are smoothed to have a complete picture of the spatial variation and temporal evolution of mortality. The main goal of this work is to detect different spatio-temporal patterns according to the age groups the population is divided into. Additionally, the use of rates makes it possible to directly compare the mortality figures among regions and age groups. Firstly, additive models with CAR structures for the region, time and age effects will be taken into account, but these models may not be appropriate in practice as space-time, space-age and time-age interactions can occur (see for instance Knorr-Held and Besag, 1998). Consequently, interaction models of the type described by Knorr-Held (2000) are considered. An approximate method for Bayesian inference based on nested Laplace approximations, INLA (Rue et al, 2009), will be used for model fitting and inference. A spatial prior for the region effect proposed by Leroux et al. (1999) will be used instead of the widely used spatial prior given by Besag et al. (1991). The reason is that this latter prior has been shown to yield negative correlations for regions far apart (see MacNab, 2011; Botella-Rocamora et al., 2013). The Leroux spatial prior can be implemented in INLA (see Ugarte et al., 2014). The results will be illustrated with Spanish prostate cancer mortality data during the period 1986-2010.

Acknowledgments. This work has been supported by the Spanish Ministry of Science and Innovation (project MTM2011-22664 which is co-funded by FEDER grants). We would like to thank the staff of the National Center of Epidemiology in Spain for providing the data.

References

[1] Besag, J., York, J., and Mollié, A. (1991). Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics* **43**, 1-20.

[2] Botella-Rocamora, P., López-Quílez, A., and Martínez- Beneito, M. A. (2013). Spatial Moving Average Risk Smoothing. *Statistics in Medicine* **32**, 2595–2612.

[3] Dean, C. B., Ugarte, M. D., and Militino, A. F. (2001). Detecting interaction between random region and fixed age effects in disease mapping. *Biometrics* **57**, 197–202.

[4] Knorr-Held, L. (2000). Bayesian modelling of inseparable space-time variation in disease risk. *Statistics in Medicine* **19**, 2555–2567.

[5] Knorr-Held, L. and Besag, J. (1998). Modelling risk from a disease in time and space. *Statistics in Medicine* **17**, 2045-2060.

[6] Leroux, B. G., Lei, X., and Breslow, N. (1999). Estimation of disease rates in small areas: A new mixed model for spatial dependence. In: Halloran M and Berry D (eds) *Statistical models in epidemiology, the environment and clinical trials*, pp. 135–178. New York, Springer-Verlag.

[7] MacNab, Y. C. (2011). On Gaussian Markov random fields and Bayesian disease mapping. *Statistical Methods in Medical Research* **20**, 49–68.

[8] Rue, H., Martino, S., and Chopin, N. (2009). Approximate Bayesian Inference for Latent Gaussian Models Using Integrated Nested Laplace Approximations (with discussion). *Journal of the Royal Statistical Society, Series B* **71**, 319–392

[9] Ugarte, M. D., Adín, A., Goicoa, T., and Militino, A. F. (2014). On fitting spatio-temporal disease mapping models using approximate Bayesian inference. *Statistical Methods in Medical Research*, doi:10.1177/0962280214527528.

[10] Ugarte, M. D., Ibáñez, B., and Militino, A. F. (2006). Modelling risks in disease mapping. *Statistical Methods in Medical Research* **15**, 21–35.

[11] Zhang, S., Sun, D., He, C. Z., and Schootman, M. (2006). A Bayesian semi-parametric model for colorectal cancer incidences. *Statistics in Medicine* **25**, 285–309.