Autoregressive spatio-temporal disease mapping

M.A. Martínez-Beneito\textsuperscript{1}, A. López-Quílez\textsuperscript{2} and P. Botella-Rocamora\textsuperscript{2}

\textsuperscript{1} D.G. de Salud Pública, Conselleria de Sanitat-Generalitat Valenciana, Micer Mascó 31, 46010 Valencia, Spain
\textsuperscript{2} Dept. d’Estadística i I.O., Universidad de Valencia, Doctor Moliner 50, 46100 Burjassot, Spain

Keywords: Bayesian Hierarchical Models; Lattice data; Epidemiology; Lung cancer.

1 Introduction

During the last decade statistical techniques for disease mapping have become very popular in epidemiology. These collection of methods enable to smooth ecological health indicators from the geographical structure of the units under study. As a consequence, more reliable risk estimates in less populated areas are obtained due to the sharing of information between neighboring regions, which are intended to share common risk factors.

Among all those proposals to perform risk smoothing appeared in the literature, the one stated by Besag, York and Molliè (BYM, 1991) has had an special impact and it can be found lots of applications of such model. This approach decomposes the risk in every region as a sum of two effects, one spatially dependent and one independent for all the geographical units.

But although the BYM model is a benchmark in spatial problems it ignores the temporal evolution of risks on the region under study. So risk estimates provided with such proposal are supposed to be static in time and this assumption can not be always too realistic, mainly in those problems with a wide time window.

Bernardinelli et al. (1995) propose to model spatio-temporal data as an ecological regression model with spatially structured errors (CAR-normal) both in the intercept and the covariate, the time period. In that case risk evolution in time for every region depends on the evolution in its neighbors, but it only admits a linear trend. Following this approach, Assunção et al. (2001) introduce a quadratic term in time trend but they also rely on a parametric description of such evolution, which they considered appropriate for short time periods. In Sun et al. (2000) it is also proposed a linear
modeling of time trend but with a CAR proper variable instead of the spatial effect of the BYM model.

From a very different approach, in Waller et al. (1997) and Knorr-Held and Besag (1998) risks are modelled for every period as independent (in time) spatial and heterogeneous effects. In these works the main effect for every period is modelled as a gaussian random walk in time, and precisions of random effects are supposed to have a common distribution in the first case and they are supposed equal in the second one. This way, although random effects are independent in these two works, the main effects and precisions share information among different periods. Following this approach Nobre et al. (2005) structure the distribution of the former precisions as gaussian random walks in time, so closer years will have a similar performance. As it can be seen, there is neither a spatio-temporal model with the agreement of BYM in the spatial field nor a wide consensus on how to describe in a proper way the temporal and spatial evolution at a single time. In this work it is introduced a spatio-temporal approach to disease mapping combining autoregressive time series ideas, to link risk in time at every site, and the BYM proposal to link risk in space at every moment. The model proposed is applied to the study of the spatio-temporal distribution of lung cancer mortality in women during the period 1987-2004 in Comunidad Valenciana (Spain) with a yearly temporal aggregation.

2 Autoregressive linking of spatial patterns

The main aim of the approach that we are going to introduce is to define a spatio-temporal structure in which the relative risks are both spatial and temporal dependent at the same time. So, we agree with the philosophy of Bernardinelli et al. (1995), Sun et al. (2000) and Assunção et al. (2001) of defining a similar temporal evolution in those places that are geographically close. But our intention is to define a time structure smoother than a linear trend to capture non-linear shaped evolutions in every geographical unit under study.

It is necessary to define an structure on the relative risk for every area and time interval under study that takes profit of spatial and temporal relations to get reliable estimates. With that aim we define the log-relative risk at the first observed season as sum of an intercept, and two random effects as in the BYM model. We include both random effects with the hope of capturing the spatial pattern of risk and to ensure enough flexibility to permit very different risk estimates in close places. For the next time intervals we incorporate an autoregressive term of order 1 for every region. This way the expected value for every region will depend not only in their neighbors, but also in their estimates on previous seasons. Alternatively, if we condition on the values of the spatial random effect, an autoregressive process of order 1 is obtained.
In this case the prior mean of the relative risk does not depend only on the actual values of the log-relative risk in its neighbors but it also depends of their values in the past. So it is induced a temporal dependence between risk estimations on every site.

Moreover, the differences in risk between seasons have spatial structure. So, that difference will be geographically smooth and as risk in every period is based on risk estimation on the former one it will provide a smooth evolution of the risk on time.

Lastly, the log-relative risk evolution on time at every moment has been defined as a linear function of such value in the previous season instead of using a linear trend for the whole period under study. So, it is used non-parametric modelling to describe temporal evolutions in a similar way to that used for the geographical term. We hope that this modelling could reflect non-linear evolutions on risk in a proper way.

3 Distribution of lung cancer in women

Comunidad Valenciana is one of the 17 autonomous regions in Spain, placed beside of its mediterranean coast. It has been described for this region (Martinez-Beneito et al., 2005) that lung cancer mortality in women has been the tumoral cause of mortality with a higher increase (53.3%) during the decade 1991-2000. A spatio-temporal study was planned to determine if such increase was been experienced in the whole region in a similar way or if it is not, to determine the regions with higher raising.

A comparison of several models are made in terms of the deviance information criterion (DIC) (Spiegelhalter et al., 2002) to assess the performance of our proposal. The models that we have evaluated are the following:

- **BYM model** (Besag et al, 1991) ignoring temporal evolution of risk.
- **Ecological Regression on Time** in a similar way to that proposed in Bernardinelli et al. (1995) but with heterogenous and spatial random effects in both the intercept and the linear time trend.
- **Ecological Regression on Time** including a quadratic term in time trend for every municipality as proposed in Assunção et al. (2001).
- **Independent Risks** estimates for every period with gaussian random walk prior on the intercept term and precision of random effects (spatial and heterogenous), as proposed in Nobre et al. (2005).
- **AutoRegressive Spatio-temporal** estimation of risks for every region with spatial structure on temporal evolution.

The autoregressive spatio-temporal proposal has proved very valuable in this situation. The risk is not evolving in the same way in the whole region
but it is increasing much more in the south-east region than in the rest of Comunidad Valenciana. On the contrary risk rising is not as evident in other regions as, for example the northern or the west central side where risk estimates are maintaining along the whole sequence.

Acknowledgments: This work was supported by the Conselleria de Sanitat (Comunidad Valenciana) and by the project MTM2004-03290 funded by Ministerio de Educació y Ciencia (Spain), European Regional Development Fund, and Conselleria de Empresa, Universitat i Ciència (Comunidad Valenciana).

References


