

# Disease mapping and spatio-temporal analysis: importance of expected-case computation criteria

Gonzalo López-Abente<sup>1,2</sup>, Nuria Aragonés<sup>1,2</sup>, Javier García-Pérez<sup>1,2</sup>, Pablo Fernández-Navarro<sup>1,2</sup>

<sup>1</sup>*Environmental and Cancer Epidemiology Unit, National Centre for Epidemiology, Carlos III Institute of Health, Madrid, Spain;* <sup>2</sup>*Consortium for Biomedical Research in Epidemiology and Public Health (CIBERESP), Spain*

**Abstract.** The municipal, spatial pattern of male stomach cancer mortality in Spain, spanning the period 1989-2008, was studied, comparing the results of depicting mortality using different expected-case computation methods in a spatial and spatio-temporal modelling context. Expected cases for each municipality were first calculated by two methods: (i) using reference rates for each 5-year period; and (ii) using average reference rates for the overall period. This was visualised by two types of models: (i) independent maps for each period based on the model proposed by Besag, York and Mollié; and (ii) a series of maps over time based on a model with spatio-temporal interaction terms. An additional model, based on mortality rate ratios as an alternative to the traditional use of standardised mortality ratios, was also fitted. Integrated nested Laplace approximations were used as the Bayesian inference tool. The results show that, in general, the geographical pattern was maintained across the study period, and that the maps differed appreciably according to the method used to obtain the expected number of cases. While the use of average reference rates appears to be the most suitable choice where the aim is to study time trends by area, it may nevertheless mask the spatial pattern in situations where the time trend is very marked and the study period is long. When it comes to studying changes in the spatial pattern of stomach cancer mortality, we feel that it is most useful to plot independent maps by period and use the “local” rates for each period as reference in the computation of expected cases.

**Keywords:** disease mapping, cancer mortality, epidemiology, gastric cancer, spatial epidemiology, Spain.

---

## Introduction

The visual impact of a geographical pattern is a critical component in the interpretation of disease distribution maps and it is strongly influenced by the method used to represent the indicator (Lawson et al., 1999; Elliott et al., 2000). In studies which address the analysis of spatial disease data coming within the category of “small areas”, the indicator most frequently depicted on the maps is a smoothing of the ratio of observed to expected cases (standardised mortality ratios (SMRs) or standard incidence ratios) based on the Besag-York-Mollié spatial autoregressive model (Besag et al., 1991). For this indicator, it is usual for the average rates of the area analysed to be used as reference rates for expected-case computation purposes (Breslow and Day, 1987). Yet, when a variable of period or time is included in the analysis, there are several

possibilities for selecting reference rates: one is to use the average rates for the entire study period, while another, more frequently used approach (Elliott et al., 2000) is to apply the average rates for each period included. The resulting maps are appreciably different, particularly in those cases where there is a marked time trend. The popularity of spatio-temporal modelling (Knorr-Held, 2000; Abellan et al., 2008; Martínez-Beneito et al., 2008; Blangiardo et al., 2013) leads to the adoption of new methods of calculating the indicator to be depicted. The inclusion of interaction terms in smoothing models means that, with a single model, representations of the maps can be obtained for different periods.

The aim of this study was to compare the results of depicting mortality indicators for a given disease using different methods of calculating expected cases in a spatial and spatio-temporal modelling context. This issue is a well known problem in spatial analysis. The disease mortality chosen for analysis was that of stomach cancer in Spain owing to its marked geographical and temporal pattern (Aragonés et al., 2013). Since the geographical distribution of stomach cancer mortality in Spain is very similar in both sexes (Aragonés et al., 2009), we decided to study male mortality in view of the higher number of cases.

---

Corresponding author:  
Gonzalo López-Abente  
Environmental and Cancer Epidemiology Unit  
National Centre for Epidemiology  
Carlos III Institute of Health  
Monforte de Lemos 5, 28029 Madrid, Spain  
Tel. +34 91 822 2634; Fax +34 91 387 7815  
E-mail: glabente@isciii.es

## Materials and methods

### Mortality and population data

As case source, we used individual death entries for the period 1989-2008 corresponding to stomach cancer (International Classification of Diseases, 9<sup>th</sup> Revision (ICD-9) code 151, and 10<sup>th</sup> Revision (ICD-10) code C16) in men in Spain. These data were furnished by the National Statistics Institute (Instituto Nacional de Estadística-INE). The data were broken down by municipality ( $n = 8,073$ ).

Municipal male population data, broken down by age (18 groups), were obtained from the 1991 and 2001 censuses and the 1996 and 2006 municipal rolls. These years correspond to the midpoints of the four quinquennia that comprised the total study period, i.e. 1989-1993, 1994-1998, 1999-2003 and 2004-2008. The person-years for each 5-year period were obtained by multiplying these populations by 5. The SMR was calculated as the ratio of observed to expected deaths.

### Expected-case computation using reference rates for each period

As the first option for the calculation of expected cases ( $E_{it}$ ), the overall Spanish mortality rates ( $\lambda_{jt}$ ) for the above four 5-year periods (local rates) were multiplied by each municipality's (i) person-years ( $n$ ) by age groups (j) and quinquennium (t).

$$\lambda_{jt} = \sum_{i=1}^N d_{ijt} / \sum_{i=1}^N n_{ijt}$$

and

$$E_{it} = \sum_{j=1}^J (n_{ijt} \lambda_{jt})$$

for  $i = 1, \dots, N$ ;  $t = 1, \dots, T$  where  $d$  were observed deaths,  $N$  the number of municipalities included and  $T$  the number of periods.

### Expected-case computation using average reference rates ( $\lambda_+$ ) for the overall period

The average age-specific reference rates were

$$\lambda_j^* = \sum_{t=1}^T d_{jt} / \sum_{t=1}^T n_{jt}$$

and the expected cases

$$E_{it} = \sum_{j=1}^J n_{ijt} \lambda_j^*$$

### Spatial and spatio-temporal modelling

The smoothing of the SMRs for the two expected-case computation methods was performed in accordance with the following spatial and spatio-temporal modelling:

#### Spatial modelling; the Besag-York-Mollié (BYM) model

For the  $i$ -th area, the number of cancer deaths  $y_i$  is modelled as  $y_i \sim \text{Poisson}(\lambda_i)$ , where the mean  $\lambda_i$  is defined in terms of a rate  $\rho_i$  and the expected number of cancer deaths  $E_i$  as  $\lambda_i = \rho_i E_i$ .

For map plotting purposes, smoothed municipal relative risks (RR) were calculated using the conditional autoregressive BYM model proposed by Besag et al. (1991). This model is based on fitting Poisson spatial models, with observed cases as the dependent variable, expected cases as offset and two types of random effect terms, which take the following into account: (i) municipal contiguity (spatial term  $v_i$ ); and (ii) municipal heterogeneity ( $v_i$ ), where  $\alpha$  is the intercept quantifying the average mortality rate in all the municipalities. In this model the linear predictor is

$$\eta_i = \log(\rho_i) = \alpha + v_i + v_i \quad (\text{Model 1})$$

The spatial term is modelled using an intrinsic conditional autoregressive structure (iCAR) and the heterogeneity term using an exchangeable prior. The smoothed RRs for plotting purposes are  $\zeta_i = \exp(v_i + v_i)$ . A model was fitted for each of the four periods separately and for each of the two methods of expected-case computation.

#### Spatio-temporal models

To obtain a series of maps over time, we used the BYM model with the inclusion of the variable "time". There are several ways to include temporal terms in the model, some with or others without spatio-temporal interactions. The first spatio-temporal model fitted (model 2) was the classic parametric model proposed by Bernardinelli et al. (1995).

$$\eta_{it} = \alpha + v_i + v_i + (\beta + \delta_i) t_i \quad (\text{Model 2})$$

where the term  $\delta_i$  represents the deviation of each area's trend from the overall trend  $\beta$ . This specification assumes a linear effect of time for each area. A sum to zero is imposed on  $\delta$  and  $v$  for "identifiability" purposes.

The second spatio-temporal model used (model 3) overcomes the assumption of linearity of the temporal term, by using a dynamic formulation of the linear predictor. While the terms  $\alpha$ ,  $v_i$  and  $v_t$  have the same “parametrisation” as in the previous models, the term  $\gamma_t$  represents the dynamically modelled structured temporal effect (for details see Blangiardo et al., 2013) taking the neighbourhood structure into account (the mortality time trend in one municipality will resemble that of neighbouring municipalities). The term  $\varphi_t$  is the unstructured temporal term.

$$\eta_{it} = \alpha + v_i + v_t + \gamma_t + \varphi_t \quad (\text{Model 3})$$

The third spatio-temporal model (model 4) is an extension of the previous model, which includes spatially structured ( $v_i$ ) and unstructured components ( $v_t$ ), time-structured ( $\gamma_t$ ) and unstructured components ( $\varphi_t$ ), and a differential trend term ( $\delta_{it}$ ) that identifies the interaction between time and space and would explain differences in mortality time trends in different areas.

$$\eta_{it} = \alpha + v_i + v_t + \gamma_t + \varphi_t + \delta_{it} \quad (\text{Model 4})$$

In this model, the two unstructured effects,  $v_i$  and  $\varphi_t$ , are assumed to interact. The model was fitted according to the Knorr-Held (2000) criteria (Blangiardo et al., 2013) and corresponds to the type I interaction. The deviance information criterion (DIC) was used to assess the fit of these models according to Spiegelhalter et al. (2002).

#### *Model with mortality rates*

Lastly, we fitted an alternative model based on mortality rates. Apart from the random-effects terms determined by the BYM model, the model includes person-years as offset (py), as fixed-effects terms the covariate “age” as a factor in nine groups (0-9, ..., 80+) and the variable “period” (t) as a continuous variable categorised into four groups. In this model the mean  $\lambda_i$  is defined in terms of a rate  $\rho_i$  and the person-years  $py_i$  as  $\lambda_i = \rho_i py_i$ .

$$\eta_{it} = \log(\rho_i) = \alpha + v_i + v_t + \beta_1 t + \sum_j w_j \text{age} \quad (\text{Model 5})$$

This model does not suffer from the comparability problems of RRs obtained from previous models based on the SMRs, since it includes person-years as

denominators and age groups. The estimator shown is a smoothing of the rates ratios. The use of RRs obtained from a model based on the mortality rates, which includes the variable age in its formulation, is equivalent to the calculation of rate ratios from age-adjusted rates by direct method of standardization. The smoothing of SMRs based on the ratio of observed and expected cases, calculated using reference rates (indirect method), is traditionally used in disease mapping.

#### *Statistical software*

Integrated nested Laplace approximations (INLAs) were used as a tool for Bayesian inference. The INLA has been developed as a very efficient alternative to the Markov chain Monte Carlo (MCMC) methods normally used for Bayesian computations. For this purpose, we used R-INLA (Rue et al., 2009), a package available in the R environment (R Development Core Team, 2005), with the option of simplified Laplace estimation of the parameters. We assumed the default specification of R-INLA for the distribution of the hyper-parameters for all the models investigated here. A total of 8,073 municipalities were included, and the spatial data on municipal contiguities were obtained by processing the official INE maps.

#### **Results**

Between 1989 and 2008, there were 75,750 deaths due to stomach cancer in men in Spain. The age-adjusted mortality rate (European standard population) decreased from 21.26 per 100,000 in the quinquennium 1989-1993 (20,176 deaths) to 12.85 in 2004-2008 (17,611 deaths).

#### *Spatial modelling*

Fig. 1 depicts the locations of the different provinces in Spain. Fig. 2 shows maps generated on the basis of the spatial model, using the two methods of expected-case computation. These are maps generated with independent models for each period (model 1). Fig. 2a plots the maps of municipal distribution of RRs of stomach cancer mortality for the respective 5-year periods, with each using its own reference rates. This map clearly shows that the geographical pattern changed very little over the 20 years studied (1989-2008). In general, geographical differences lessened as result of a reduction in the number of municipalities in the highest RR category, with areas of highest mortal-



Fig. 1. Spain’s provinces and Autonomous regions (Comunidades Autónomas).

ity, such as the north of Lerida and Girona, apparently tending to merge with the remainder. Areas which maintained above-average mortality were broad stretches of Castile-Leon and municipalities along the Galician coast (Pontevedra and Corunna) having higher rates than those of Burgos and Palencia. The maps that used average reference rates (model 1) are shown in Fig. 2b. This method of depiction highlights the mortality time trend, indicating the important reduction in stomach cancer mortality registered by men in Spain.

*Spatio-temporal modelling*

Table 1 shows the DIC corresponding to the three spatio-temporal models that used the two methods of expected-case computation. The non-parametric model which displayed the best fit in both situations was model 4, and the comments in this section refer to the results yielded by this particular model. Fig. 3 plots the map of the RRs of stomach cancer mortality generated on the basis of the model with spatio-temporal interaction (model 4), using the two methods of expected-

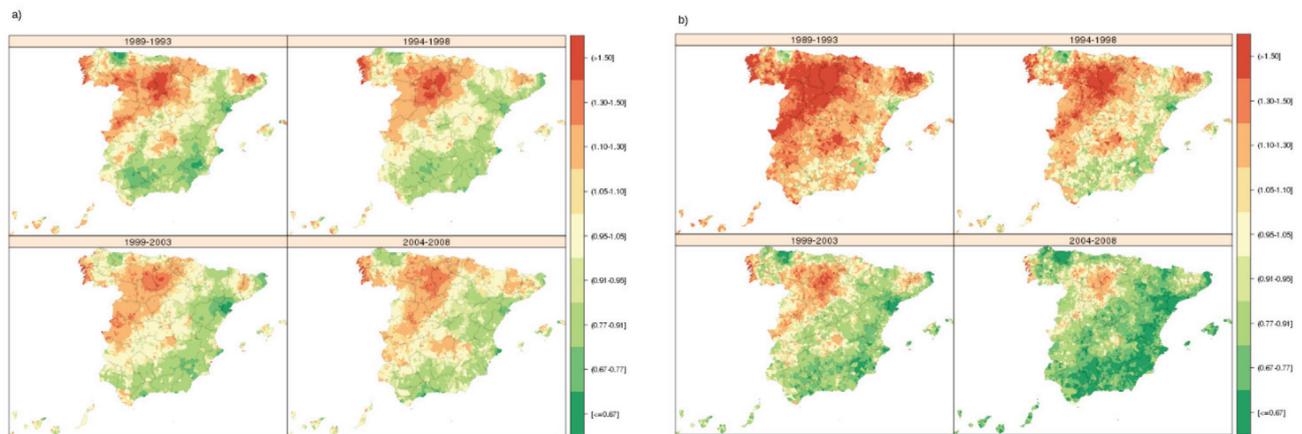


Fig. 2. Municipal distribution of RR of stomach cancer mortality in men for each 5-year period between 1989-2008 yielded by model 1. Independent maps for each quinquennium based on: (a) expected-case calculation with reference rates for each period; and (b) expected-case calculation with average reference rates for the overall period.

Table 1. Deviance information criterion (DIC) for the three spatio-temporal models considered as defined by equations 2, 3 and 4 .

Model	Reference rates for each period <sup>a</sup>			Average reference rates <sup>b</sup>		
	D	pD	DIC	D	pD	DIC
2 (parametric)	68,719.29	1,207.70	69,927.00	68,783.01	1,036.51	69,819.52
3	68,739.92	1,180.89	69,920.81	68,738.60	1,186.57	69,925.18
4	68,654.57	1,254.02	69,908.59	68,638.28	1,276.50	69,914.77

D is the posterior mean of the deviance measuring model fit, and pD the effective number of parameters representing model complexity. (a) Expected-case calculation using reference rates for each period; and (b) expected-case calculation using average reference rates for the overall period.

case computation. While these maps are generally similar in appearance to those obtained from the independent models shown in Fig. 2, some differences can nevertheless be seen. They are fundamentally centring on a spatial smoothing of the somewhat smaller RR and the fact that the maps generated with reference rates for the four periods are practically identical. The maps generated using average reference rates depict the temporal variation in mortality more clearly, though visually they give the impression that the spatial pattern is disappearing, something that would not seem to be true if the spatial patterns shown in Fig. 2a are borne in mind.

The overall spatial pattern obtained from model 4 with spatio-temporal interaction ( $\zeta_i = \exp(v_i + v_j)$ ) is shown in Fig. 4a for the 8,073 municipalities in Spain, while Fig. 4b depicts the measure of uncertainty of the posterior probability of RR >1 ( $\zeta_i : p(\zeta_i > 1 | y)$ ). A spatial relative risk greater than 1 can be seen in some parts of the country with posterior probabilities above 0.8, indicating a relatively small level of associated uncertainty. The time trend obtained from model 2 is included in Fig. 4c and shows a fairly uniform decrease

in the stomach-cancer mortality risk across the period 1989-2008. The global maps of smoothed risks and posterior probabilities of RR being >1, obtained with model 4 using average reference rates or reference rates by period, were practically identical, so only the former is shown. With respect to the general trend in stomach cancer mortality ( $\exp(\gamma_t + \phi_t)$ , model 4) for the four study periods, it can be said that if the expected cases are calculated using the reference rates for the four quinquennia (local rates), the general trend would be flat even if the frequency of stomach cancer mortality is decreasing.

#### Modelling of mortality rates

Lastly, global municipal maps (Figs. 5a and 5b) were generated for the entire period using the additional model based on mortality rates (model 5). Since the pattern of smoothed rates ratios and posterior probabilities of RR being >1 did not differ substantially from the global map obtained with the SMRs of the spatial or spatio-temporal model, the validity of the map based on SMRs can be assumed.

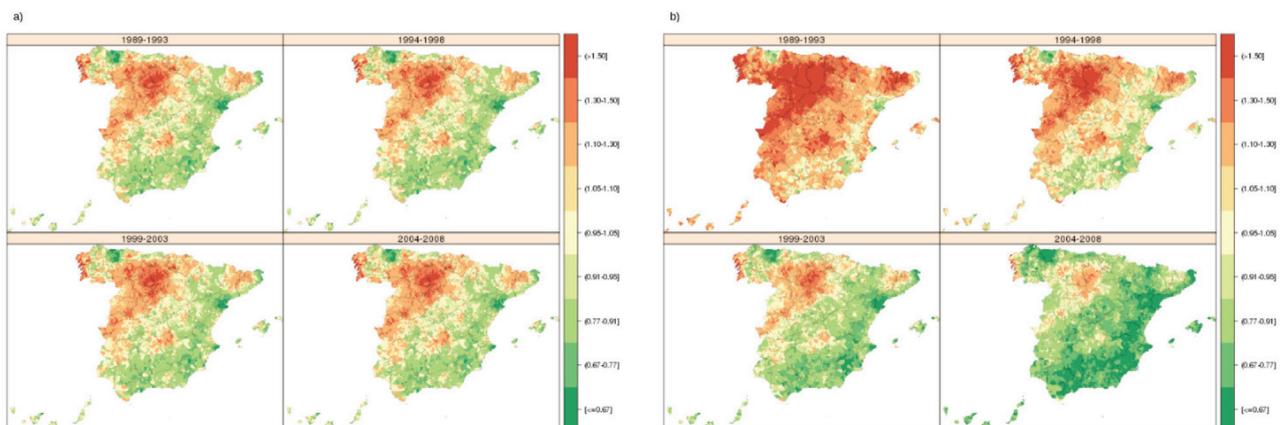


Fig. 3. Municipal distribution of RR of stomach cancer mortality in men for each 5-year period between 1989 and 2008 yielded by model 4. Maps obtained from the model with spatio-temporal interaction based on: (a) expected-case calculation with reference rates for each period; and (b) expected-case calculation with average reference rates for the overall period.

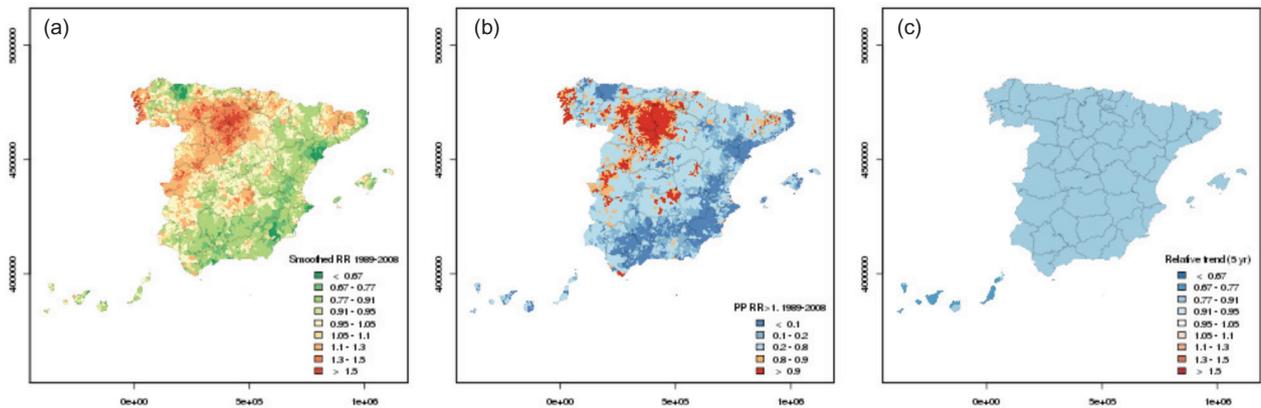


Fig. 4. Municipal distribution of stomach cancer mortality in men over the period 1989-2008 estimated based on average reference rates and model 4. (a) Map of spatial pattern of relative risk of disease  $\zeta_i = \exp(v_i + v_j)$ ; (b) map of uncertainty for the spatial effect  $\zeta_i : p(\zeta_i > 1 | y)$ ; and (c) map of relative time trend.

Discussion

The study of the time trend in spatial mortality patterns can be approached using different methodologies. This study compared the results of various possibilities of spatial and spatio-temporal modelling of stomach cancer mortality among men in Spain. Independent maps were plotted for each study period resulting in a series of maps over time on the basis of a model with spatio-temporal interaction terms. The methods were applied using two ways of obtaining expected cases. Furthermore, an additional model was fitted based on age-adjusted mortality rates ratios as an alternative to the traditional use of SMRs in these types of models. These decisive factors were used to study the spatial distribution of stomach cancer mortality in view of the fact that this disease displays a marked spatial and temporal pattern in Spain.

Our results show that the choice of reference rates

for the calculation of expected cases is a critical decision when it comes to disease mapping and constructing mortality atlases. It is advisable to use reference rates corresponding to each study period when checking spatial patterns, since the use of average reference rates for a very broad time window may mask the geographical pattern with increasing distance from the reference rates. The reason is that it accentuates the temporal component, as occurs in the model applied by us in the case of gastric cancer, due to the disease’s visible decline throughout Spain over the study period. The use of reference rates for each of the four 5-year periods intrinsically includes the mortality time trend. The general time trend is thus “deducted from” the geographical representation in this method of calculating expected cases. Even so, it appears to be the most advisable option for visually detecting possible changes in the spatial pattern.

The method of visualising a map series using average

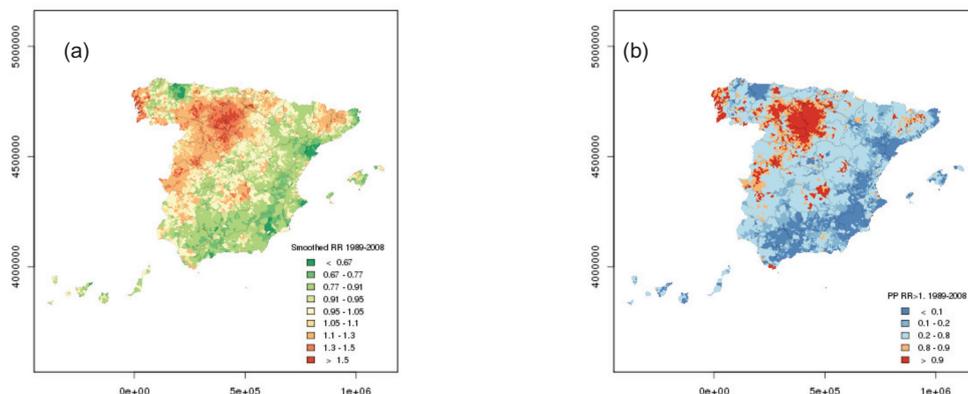


Fig. 5. Municipal distribution of RR ( $\zeta_i = \exp(v_i + v_j)$ ) of stomach cancer mortality in men over the period 1989-2008 and map of uncertainty for the spatial effect  $\zeta_i : p(\zeta_i > 1 | y)$ . Maps based on the smoothing of age- (a) and period-adjusted (b) mortality rates ratios (model 5).

rates highlights the time trend in cases where this is very pronounced. In the case of stomach cancer, the areas of Castile-Leon and Atlantic coast of Galicia, shown to have excess risk on the map corresponding to the last 5-year period, could be interpreted as: (i) areas with a lag in the rate decrease, since the average reference rates would correspond to those for 1998-1999; and/or (ii) areas with differential risk factors vis-à-vis the country as a whole. The plotting of independent maps for each period (model 1) affords advantages as well as disadvantages. One advantage is conceptual simplicity because the same model is applied for different periods; another is the possibility of detecting self-limited problems in time (appearing and disappearing clusters, which are masked in the spatio-temporal smoothing). As regards disadvantages, the reduction in the number of cases as a consequence of separating the data by period should be mentioned as well as the fact of having to fit different models to obtain each map. This criterion can be applied in the study of the most frequent cancers but it is of questionable interest in the least frequent tumours.

The idea of obtaining independent maps in time was applied by Waller et al. (1997), who used a nested model where the BYM model is applied to each time point separately. The model does not have spatial main effects and therefore allows for the spatial patterns at each time point to be completely different. There is less emphasis on modelling temporal development in disease risk, as time is treated as essentially exchangeable (Knorr-Held, 2000). The use of spatio-temporal models also has its advantages and disadvantages. Advantages include the use of a single model to obtain a series of maps over time, the provision of an estimator of the time trend for each area enabling the detection of areas which diverge from the general trend, and the possibility of obtaining a picture of the global spatial pattern for the entire study period. The disadvantages entail the model's complexity, which is due to the inclusion of temporal strata leading to a considerable computational burden posed by massive data analysis, such as that described in this study. The fitting of these spatio-temporal models for a large number of areas was inconceivable until a few years ago. However, the use of INLAs makes it possible for such analysis to be tackled within reasonable computational time limits (Schrödle and Held, 2010; Blangiardo et al., 2013).

Although the use of the indicator (SMRs *versus* rates-ratios adjusted by the direct method) has received considerable attention, the choice of reference rates has been less discussed. Historically speaking, the

use of the SMR has a long tradition and is the indicator most frequently used for producing disease atlases and undertaking small-area studies. It is also an indicator that is highly criticised due to the intrinsic difficulty of comparison among SMRs, since the weighting of each SMR depends on the age structure of the study population (Julious et al., 2001; Rothman et al., 2008; Ocaña-Riola, 2010). Nevertheless it continues to be used in practice. One advantage of the SMR is that it affords an indicator of the deviation of each area's mortality from that of the general population, which is easily interpretable and considered to have greater numerical stability (Breslow and Day, 1987). Furthermore, the number of parameters to be estimated is considerably reduced in comparison with a model that would also include age groups as a covariate. Theoretically, this could make for improved efficiency in terms of estimating parameters. In the case of stomach cancer in Spain, the provincial pattern displayed by rates age-adjusted by the direct method (Lopez-Abente et al., 1996) is similar to that obtained by the smoothing procedure based on SMRs.

Since the middle of the 20<sup>th</sup> century, gastric cancer incidence and mortality has decreased in tandem with socio-economic development. In Spain, the 8,805 deaths registered in both sexes accounted for 16.8% of all cancer-related deaths in 1975, a percentage that has fallen to 5.3% by 2011 (with 5,597 deaths). While the causes of this decline are still not wholly understood, this trend and the pronounced geographical differences shown in the mortality due to this tumour suggest that environmental factors are strongly implicated in its aetiology. The discovery of *Helicobacter pylori* in 1983 (Marshall and Warren, 1984) and of its role in gastric carcinogenesis (IARC, 1994) meant a sharp change in the aetiological understanding of this tumor. Currently, the decrease in the frequency of this cancer is attributed to the improvement in living standards coinciding with the availability and wider use of antibiotics. Both factors have reduced the incidence and prevalence of *H. pylori* infection. Other risk factors that modulate the development of stomach cancer are genetic factors, smoking habits and diet (González and Agudo, 2012). Worth mentioning in this connection is a recent paper describing the spatio-temporal trend in gastric cancer mortality at the provincial level in Spain (Aragónés et al., 2013) using mortality for the entire study period as reference. This study highlights the surprising fact that, despite the differences in risk which exist in the different provinces, risk of dying from cancer is decreasing at a similar rate in all areas regardless of

their level of risk, and in both sexes alike.

The analysis reported here extends the results of this provincial analysis, and, by also presenting the results using a temporal reference for each 5-year period, it clearly shows that the spatial pattern of the distribution of gastric cancer mortality in Spain has been maintained with little variation across the four periods studied, even though they cover a time span totalling 20 years. The risk factors that give rise to this pattern and the reasons for its persistence are unknown. The maps shown allow for a clearer identification of the municipalities along Galicia's Atlantic coast that display a higher mortality rate and the decrease in the number of areas in Castile-Leon that display excess mortality. While the decline in mortality is fairly uniform throughout Spain, the risk factors involved in the different regions having areas of higher mortality can be notably different, e.g. Atlantic Galicia *versus* interior of Burgos with very different dietary patterns. Lastly, by way of a validation instrument, it might be of interest to model mortality, avoiding the problems inherent in SMRs by using rates as the indicator for this purpose. This would, however, make it necessary for age to be included in the models, thus increasing the computational burden considerably.

## Conclusions

Visual perception of spatial patterns is significantly influenced by reference rates chosen for the purpose of calculating expected cases. While the use of average reference rates for an entire period appears to be the most suitable choice where the aim is to study time trends by area, it may nevertheless mask the spatial pattern in situations where the time trend is very marked and the study period is long. When it comes to studying changes in the spatial pattern of stomach cancer mortality, we feel that it is more useful to plot independent maps by period and use the "local" rates for each period as reference in the computation of expected cases. The above analysis clearly shows that stomach cancer mortality in Spain maintains its singular spatial pattern in spite of the continuous decline in rates witnessed in recent decades.

## Acknowledgements

The study was partially supported by a research grant from the Spanish Health Research Fund (FIS PI11/00871). Mortality data were supplied by the Spanish National Statistics Institute in accordance with a specific confidentiality protocol.

## References

- Abellan JJ, Richardson S, Best N, 2008. Use of space-time models to investigate the stability of patterns of disease. *Environ Health Perspect* 116, 1111-1119.
- Aragónés N, Goicoa T, Pollán M, Militino AF, Pérez-Gómez B, López-Abente G, Ugarte MD, 2013. Spatio-temporal trends in gastric cancer mortality in Spain: 1975-2008. *Cancer Epidemiol* 37, 360-369.
- Aragónés N, Pérez-Gómez B, Pollán M, Ramis R, Vidal E, Lope V, García-Pérez J, Boldo E, López-Abente G, 2009. The striking geographical pattern of gastric cancer mortality in Spain: environmental hypotheses revisited. *BMC Cancer* 9, 316.
- Bernardinelli L, Clayton D, Montomoli C, 1995. Bayesian estimates of disease maps: how important are priors? *Stat Med* 14, 2411-2431.
- Besag J, York J, Mollié A, 1991. Bayesian image restoration, with two applications in spatial statistics (with discussion). *Ann Inst Stat Math* 43, 1-59.
- Blangiardo M, Cameletti M, Baio G, Rue H, 2013. Spatial and spatio-temporal models with R-INLA. *Spat Spatiotemporal Epidemiol* 4, 33-49.
- Breslow NE, Day N, 1987. *Statistical methods in cancer research: volume 2 - the design and analysis of cohort studies*. IARC Scientific Publications No 82. IARC, 406 pp.
- Elliott P, Wakefield JC, Best N, Briggs S, 2000. *Spatial epidemiology - methods and applications*. Oxford: Oxford University Press, 494 pp.
- González CA, Agudo A, 2012. Carcinogenesis, prevention and early detection of gastric cancer: where we are and where we should go. *Int J Cancer* 130, 745-753.
- IARC, 1994. *Schistosomes, liver flukes and Helicobacter pylori*. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans No 61. IARC, 270 pp.
- Julious SA, Nicholl J, George S, 2001. Why do we continue to use standardized mortality ratios for small area comparisons? *J Public Health Med* 23, 40-46.
- Knorr-Held L, 2000. Bayesian modelling of inseparable space-time variation in disease risk. *Stat Med* 19, 2555-2567.
- Lawson A, Biggeri A, Bohning D, Lesaffre E, Biggeri A, Viel JF, Best N, 1999. *Disease mapping and risk assessment for public health*. New York: John Wiley and Sons, 482 pp.
- López-Abente G, Pollán M, Escolar A, Errezola M, Abaira V, 1996. *Atlas de mortalidad por cáncer y otras causas en España, 1978-1992*. Fundación Científica de la Asociación Española Contra el Cáncer, 338 pp.
- Marshall BJ, Warren JR, 1984. Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration. *Lancet* 1(8390), 1311-1315.
- Martínez-Beneito MA, López-Quilez A, Botella-Rocamora P, 2008. An autoregressive approach to spatio-temporal disease

- mapping. *Stat Med* 27, 2874-2889.
- Ocaña-Riola R, 2010. Common errors in disease mapping. *Geospat Health* 4, 139-154.
- R Development Core Team, 2005. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.
- Rothman KJ, Greenland S, Lash TL, 2008. *Modern epidemiology* (third editions). Lippincott Williams & Wilkins, 758 pp.
- Rue H, Martino S, Chopin N, 2009. Approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations (with discussion). *J R Stat Soc Ser B* 71, 319-392.
- Schrödle B, Held L, 2010. A primer on disease mapping and ecological regression using INLA. *Comp Stat* 26, 241-258.
- Spiegelhalter DJ, Best NG, Carlin BP, Van der Linde A, 2002. Bayesian measures of model complexity and fit (with discussion). *J R Stat Soc Ser B* 64, 583-640.
- Waller LA, Carlin BP, Xia H, Gelfand AE. 1997. Hierarchical spatio-temporal mapping of disease rates. *J Am Stat Assoc* 92, 607-617.