

Stomach cancer incidence in Southern Portugal 1998–2006: A spatio-temporal analysis

Ana L. Papoila^{*,1,7}, Andrea Riebler², Antónia Amaral-Turkman^{3,7}, Ricardo São-João^{4,7},
Conceição Ribeiro^{5,7}, Carlos Geraldes^{1,7}, and Ana Miranda⁶

¹ Department of Biostatistics and Informatics, Faculty of Medical Sciences, New University of Lisbon, Campo Mártires da Pátria, 130, 1165-056, Lisboa, Portugal

² Department of Mathematical Sciences, Norwegian University of Science and Technology, Trondheim, Norway

³ Department of Statistics and Operational Research, Faculty of Sciences, University of Lisbon, Edifício C6, 1749-016 Lisboa, Portugal

⁴ School of Management and Technology, Department of Informatics and Quantitative Methods, Polytechnic Institute of Santarém, Complexo Andaluz, Apartado 295, 2001-904 Santarém, Portugal

⁵ Civil Engineering Department, Instituto Superior de Engenharia, University of Algarve, Campus da Penha, 8005-139 Faro, Portugal

⁶ Instituto Português de Oncologia de Lisboa de Francisco Gentil EPE, ROR Sul, Rua Prof. Lima Basto, 1099-023 Lisboa, Portugal

⁷ CEAUL: Center of Statistics and Applications, University of Lisbon, Edifício C6 - Piso 4, Campo Grande, 1749-016 Lisboa, Portugal

Received 15 December 2012; revised 18 December 2013; accepted 10 January 2014

Stomach cancer belongs to the most common malignant tumors in Portugal. Main causal factors are age, dietary habits, smoking, and *Helicobacter pylori* infections. As these factors do not only operate on different time dimensions, such as age, period, or birth cohort, but may also vary along space, it is of utmost interest to model temporal and spatial trends jointly. In this paper, we analyze incidence of stomach cancer in Southern Portugal between 1998 and 2006 for females and males jointly using a spatial multivariate age-period-cohort model. Thus, we avoid age aggregation and allow the exploration of heterogeneous time trends between males and females across age, period, birth cohort, and space. Model estimation is performed within a Bayesian setting assuming (gender specific) smoothing priors. Our results show that the posterior expected rate of stomach cancer is decreasing for all counties in Southern Portugal and that males around 70 have a two times higher risk of getting stomach cancer compared with their female counterparts. We further found that, except for some few counties, the spatial influence is almost constant over time and negligible in the southern counties of Southern Portugal.

Keywords: Bayesian age-period-cohort models; INLA; Portugal; Spatio-temporal analysis; Stomach cancer.



Additional supporting information may be found in the online version of this article at the publisher's web-site

1 Introduction

In Portugal the risk of stomach cancer ranks among the highest in Europe (Lunet and Barros, 2003). According to the number of diagnosed cases, stomach cancer was the fifth-frequent among all cancer

*Corresponding author: e-mail: ana.papoila@fcm.unl.pt, Phone: +351-218-803-052, Fax: +351-218-851-920

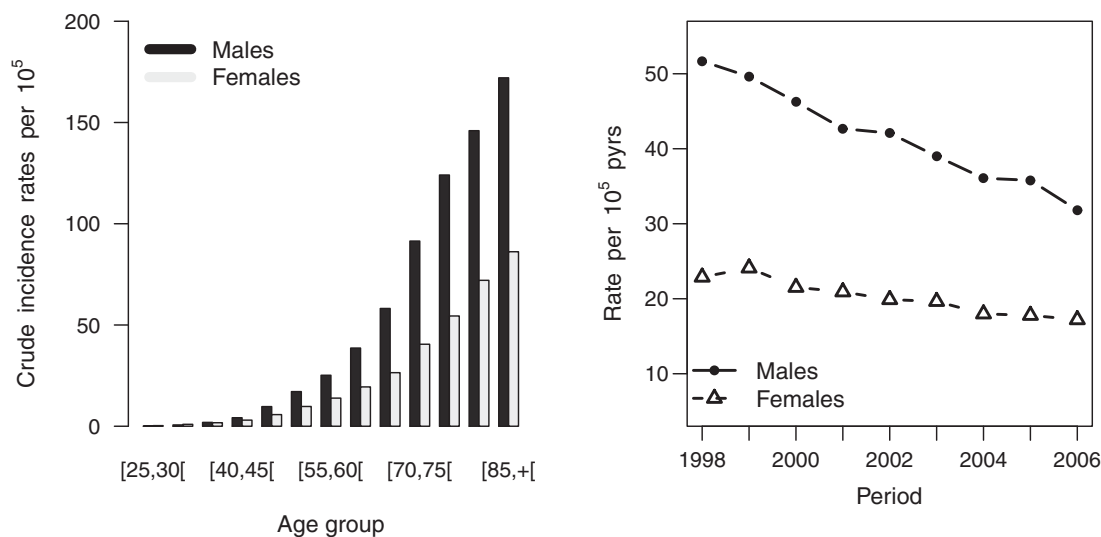


Figure 1 Incidence rates for stomach cancer in ROR-Sul from 1998 to 2006: age-specific crude rates for males and females (left), and age-standardized rates, for all periods, for males and females (right).

types in Portugal in 2006 (Registo Oncológico Nacional, 2006). Regarding the number of deaths, stomach cancer accounts for 10.2% of all malignant cancer deaths, with similar values for both genders (10.4% and 10.0% for males and females, respectively) (Risco de Morrer em Portugal, 2006). Risk factors are, among others, dietary and lifestyle habits, smoking, and *Helicobacter pylori* infections, see for example Malvezzi *et al.* (2010) and Shu *et al.* (2013). These factors depend often on generation and/or age, and in the case of dietary habits also on the geographical region. Furthermore, for almost all cancers, the age of a person is important as it is related to cumulative factors that increase the risk of contracting the disease. Thus, elderly are more likely to get stomach cancer, see also left panel of Fig. 1. When analyzing stomach cancer incidence, particular interest lies therefore in time effects, but also in spatial effects and time-space interactions. In this paper, we analyze stomach cancer incidence data in Southern Portugal from 1998 to 2006. Data are available for males and females, and stratified by age groups, year, and county, which corresponds to the common format of such cancer registry data.

Age-Period-Cohort (APC) models are a common choice to model age-stratified temporal cancer trends, see for example Malvezzi *et al.* (2010) and Rosenberg and William (2011). APC models are used to study the variation of incidence rates according to age at diagnosis (age), date of diagnosis (period), and date of birth (cohort). From a graphical point of view, the effect of these three factors on mortality rates was first considered by Lexis (1875). Later on, Holford (1983) proposed models where the effects of age, period, and cohort are estimated, considering equally spaced age and period intervals. Since then, and taking into account the identifiability problem that arises from the linear dependence between these temporal factors ($cohort = period - age$), several authors have addressed the modeling of incidence and prevalence of diseases by proposing univariate classical approaches (e.g. Clayton and Schifflers, 1987; Holford, 2006).

Bayesian APC models have become popular in the last years (Nakamura, 1986; Berzuini *et al.*, 1993; Chen *et al.*, 2011). As effects adjacent in time are likely to be similar, smoothing priors are typically assumed for age, period, and cohort effects. More recently, multivariate APC models were proposed that analyze age-specific rates not only according to age, period, and cohort, but also to one further stratification variable, such as gender (Hansell *et al.*, 2003; Jacobsen *et al.*, 2004). In fact, age-specific cancer rates are frequently reported in different strata. Due to similar risk factors it seems natural

to analyze data for males and females, say, in a joint analysis treating some sets of time effects, for example the cohort effect, as common for males and females. In a Bayesian framework, Riebler and Held (2010) proposed a multivariate APC model where nonparametric smoothing priors are assigned to all age, period, and cohort effects. Overdispersion is easily incorporated in this hierarchical Bayesian model using additional random effects. The authors showed that the differences of stratum-specific effects are identifiable and interpretable as log relative risks. Riebler et al. (2012) proposed the use of correlated stratum-specific smoothing priors and correlated overdispersion parameters in multivariate APC models. This approach may lead to more precise relative risk estimates, and is particularly attractive to impute and/or project missing data for one particular stratum by borrowing strength from the other strata.

Regarding spatial effects, it is important to understand geographical variations, particularly in small areas. Geographic mapping of diseases (Walter, 2000) is an additional and important tool in the definition of policies, for example for the allocation of resources, and for the identification of clusters with high incidence of disease. In spatial epidemiology, hierarchical Bayesian models are widely applied. They include intrinsic Gaussian Markov random field (GMRF) models (Rue and Held, 2005), such as the conditional autoregressive (CAR) model, to incorporate spatial dependency structures; see, for example, Besag et al. (1991). Through the incorporation of additional temporal effects it is possible to analyze ecological associations between cancer incidence and potential risk factors in aggregated areas over time. Different model specifications for spatial and temporal trends as well as potential space-time interactions have been proposed in this context (Bernardinelli et al., 1995; Knorr-Held, 2000). For the case in which data are also stratified by age Lagazio et al. (2003) proposed a combination of Bayesian age-period-cohort models and disease mapping, which is extended here to a multivariate setting where male and female incidence rates are analyzed jointly within one model. We apply the proposed methodology to analyze the (gender specific) impact of age, time, date of birth, and space on stomach cancer trends in Southern Portugal.

In Section 2, we introduce the dataset and the methodology for implementing Bayesian spatial multivariate APC models. In Section 3, we present the results from applying the described model to our data. A sensitivity analysis to investigate the influence of hyperpriors is presented in Section 4. In Section 5, we discuss the results and draw some conclusions.

2 Data and methods

Data are obtained from the Southern Portugal cancer registry (ROR-Sul), <http://www.ror-sul.org.pt>, which covers an area of about 39,500 km² and 109 counties (from a total of 278 Portuguese counties) plus Madeira island (which was not considered in this study). The registry comprises a population of approximately 4.5 million inhabitants that corresponds to a density of about 114 inhabitants/km². We analyze gender-specific data for 1-year periods from 1998 to 2006, given for five year age groups ([20–25], [25–30], . . . , [80–85]), and each of the 109 counties. Over all counties, ROR-Sul registered a total of 8221 diagnosed cases of stomach cancer in males and females for these age groups from 1998 to 2006. The corresponding resident populations were obtained by Statistics Portugal (INE) (<http://www.ine.pt>). Figure 1 shows age-specific crude rates, and age-standardized rates per 100,000 people of all periods for males and females summed over all counties. Age-standardization was performed using the revised European standard population (Eurostat, 2013). Supporting Information Fig. S1 shows age-standardized region-specific crude rates for males and females per 100,000 people in four selected years. Here, differences between males and females are apparent, but no clear spatial trend is visible.

2.1 Spatial multivariate APC models

Age-period-cohort models are widely used to analyze temporal time trends, while disease mapping approaches are used to explore the geographical distribution of the risk of a disease. In this paper,

Here, the sum goes over all pairs of adjacent regions r and r' . Writing (2) using a precision matrix $\mathbf{P}_u = \kappa_u \mathbf{R}_u$, the prior density can be written compactly as

$$f(\mathbf{u}|\kappa_u) \propto \exp\left(-\frac{1}{2}\mathbf{u}^\top \mathbf{P}_u \mathbf{u}\right).$$

The structure matrix \mathbf{R}_u has on the diagonal the number of neighbors of region r . The off-diagonal elements are equal to -1 , if $r \sim r'$ and zero otherwise. Since we assume that the temporal trends are different from region to region, but are not structured in space, we assume a type II interaction for the interaction term δ , see Knorr-Held (2000). That means we couple the unstructured spatial effect and the RW2 period effect. The structure matrix \mathbf{R}_δ is obtained as the Kronecker product of the corresponding structure matrices, that is $\mathbf{R}_\delta = \mathbf{R}_v \otimes \mathbf{R}_\varphi$. The rank of \mathbf{R}_δ is thereby $109 \cdot (9 - 2)$ and, hence, the prior distribution is, as the RW2, an intrinsic GMRF.

Finally, we need to place hyperpriors on all six precision parameters. The choice of appropriate prior distributions is a delicate topic. In available software packages, such as WinBugs (Lunn et al., 2000) or INLA (Rue et al., 2009), default prior parameters are given, however, these should not just blindly be overtaken, see for example Sørbye and Rue (2013). Here, we follow the approach of Fong et al. (2010) and assign a gamma distribution with shape equal to 0.5 and rate equal to 0.00149 to the unstructured spatial effect \mathbf{v} . For the remaining parameters we follow Sørbye and Rue (2013), who proposed a novel approach to derive hyperprior distributions specific for intrinsic GMRFs. The idea is to assign prior distributions to scaled precision parameters based on their marginal standard deviations. Thus, for each main effect we calculated the generalized inverse of the corresponding structure matrix. Assuming a fixed precision of one, the generalized inverse has the squared marginal standard deviations on its diagonal. The reference standard deviation $\sigma_{\text{ref}}(\mathbf{x})$ of an arbitrary intrinsic GMRF \mathbf{x} can then be calculated as the geometric mean of these marginal standard deviations $\sigma(x_i)$, $i = 1, \dots, n$, where n is the dimension of \mathbf{x} . Having the marginal and reference standard deviations for each main effect we can compute suitable parameters for the gamma-hyperpriors. Let the shape parameter be equal to one. The rate parameter b follows as

$$b = \frac{U^2 \cdot F^{-1}(\alpha, 1, 1)}{\sigma_{\text{ref}}^2(\mathbf{x})},$$

where $F^{-1}(\cdot)$ denotes the inverse cumulative distribution function for the gamma distribution. The coefficient U can be interpreted as an upper bound for $\sigma(x_i)$, so that $\text{P}(\sigma(x_i) > U) = \alpha$. As in Sørbye and Rue (2013) we set $\alpha = 0.001$ and $U = 0.5$. Table 1 shows the reference standard deviations for all main effects and the corresponding shape parameter of the gamma distribution.

2.3 Specification of linear constraints

For the (gender-specific) age-, period-, and cohort effects we apply the usual sum-to-zero constraints (Holford, 1983). Equally, the structured spatial effects \mathbf{u} are centered around zero. To ensure identifiability of the type II interaction effect δ specific constraints have to be set, as otherwise the interaction effect might be confounded with the period effect φ (Schrödle and Held, 2010). The required linear

Table 1 Reference standard deviations and derived scale parameters for the gamma hyper-prior distributions of all main effects, which have an intrinsic GMRF prior.

| | θ (RW2) | φ (RW2) | ψ (RW2) | u (Besag) | δ (Type II-interact) |
|-----------------------------------|----------------|-----------------|--------------|-------------|-----------------------------|
| $\sigma_{\text{ref}}(\mathbf{x})$ | 1.97 | 1.16 | 23.73 | 0.73 | 1.16 |
| b | 6e-5 | 2e-4 | 4e-7 | 5e-4 | 1e-4 |

constraints are derived following Rue and Held (Chapter 3.2, 2005). We compute the structure matrix $R_\delta = R_\nu \otimes R_\varphi$ and calculate its eigenvalues and eigenvectors. Since the rank of R_δ is 763 there are 218 zero eigenvalues. The eigenvectors belonging to these zero eigenvalues can be used as linear constraints for the estimation of δ leading to a total of 218 linear constraints. To be more precise, let $A\delta = e$ denote the linear constraints. Then, A is row-wise composed by those eigenvectors belonging to the zero eigenvalues, and e is a vector of length 218 where all entries are equal to zero. The number of required linear constraints needed depends on the total number of counties R and periods J , and also the chosen type of interaction. For more details in a similar context we refer to Schrödle and Held (2010).

2.4 Details on the implementation with INLA

For full Bayesian model estimation we use the recently proposed integrated nested Laplace approximations (INLA) for latent Gaussian models (Rue et al., 2009). INLA provides accurate approximations avoiding time-consuming MCMC sampling. For a detailed description of the methodology, please see Rue et al. (2009). INLA can be run under Linux, Windows, and Mac in the software environment R using the R-package `r-inla`, see www.r-inla.org. Here, we use the version updated on July 13, 2013. The complete data and R-code to run the analysis and reproduce all figures in the paper are provided in the Supporting Information.

The random walk models of second order (`rw2`) and also the Besag model (`besag`) to account for spatial structure are directly available in `r-inla`. However, to define the intrinsic GMRF of the type II interaction effect δ , the structure matrix R_δ needs to be explicitly assigned to the field `Cmatrix` of the latent model `generic0`, compare <http://www.math.ntnu.no/inla/r-inla.org/doc/latent/generic0.pdf>. Furthermore, the required linear constraints $A\delta = e$ need to be integrated. This is done by using the field `extraconstr` in the `generic0` model, to which we assign the list `lc<-list(A=A, e=e)`. Note that all linear constraints have to be ordered row-wise in A .

As already noted by Schrödle and Held (2010) the large number of linear constraints slows down the computational efficiency of INLA, when the simplified Laplace (default) or the even more demanding full Laplace approximation is used. For this reason, we used the Gaussian approximation for model estimation. However, note that this may lead to a loss of accuracy in the resulting posterior marginal distributions.

3 Results

Chandanos and Lagergren (2008) pointed out the potential role of oestrogen, which has a protective effect in women, when explaining the male dominance in the incidence of gastric cancer. Since oestrogen level varies across age, we assume in all candidate models gender-specific age effects. In the following, uppercase letters for period (P), cohort (C), Structured-space (S), Unstructured-space (U), or Interaction (I) denote a model where the corresponding effect is assumed to be the same for males and females, while lowercase letters indicate gender-specific effects. To find the best model configuration, we use the deviance information criterion (DIC), which provides a combined measure of model fit and model complexity (Spiegelhalter et al., 2002). Table 2 shows DIC values for different

Table 2 DIC values for different model configurations. The DIC value of the best model is indicated in bold.

| | aPCSUI | apCSUI | aPcSUI | aPCSui | apCSui | aPcSui | aPCsui | apCsui | aPcsui |
|-----|--------|---------------|--------|--------|--------|--------|--------|--------|--------|
| DIC | 12,604 | 12,603 | 12,605 | 12,614 | 12,613 | 12,614 | 12,630 | 12,629 | 12,631 |

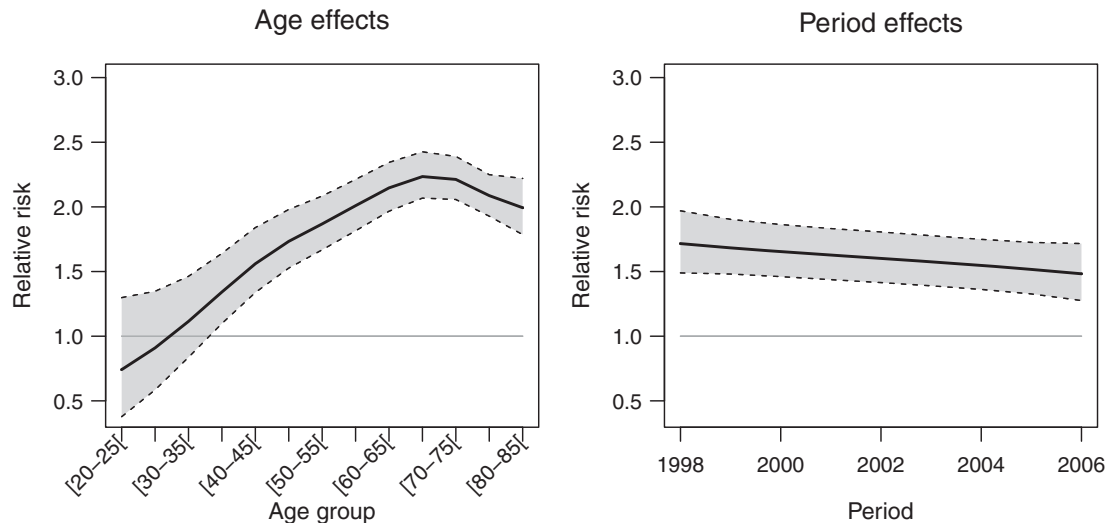


Figure 2 Relative risk (median within 95% point-wise CI) of stomach cancer incidence of males compared with females estimated by a Bayesian apCSUI model.

model configurations. The DIC values for models that assume the same spatial configuration are very close and models assuming a joint spatial structure are considered the best. The model assuming gender-specific age and period effects, and joint cohort and spatial effects (apCSUI) has the lowest DIC value and is hence regarded as the best model. Our results presented in the following are based on this model. Of note, in a preanalysis we also considered correlated gender-specific age, period, and/or cohort effects, see Riebler et al. (2012), as well as a type IV interaction component δ , but DIC values were generally not improved.

Figure 2 shows the relative risks of stomach cancer incidence for males compared with females, estimated by the apCSUI model. Concerning age (on the left), a pronounced increase of the relative risk for males attains a maximum value at around 70 years. Beyond that age, a decrease in the difference between the two genders is observed. Although rates are decreasing for both genders, this fact is more pronounced for males as it is depicted on the right panel of the figure.

To examine temporal changes in the geographical profile we use the adjusted relative risk

$$ARR = \exp(u_r + v_r + \delta_{jr}),$$

as proposed in Knorr-Held (2000). Figure 3 shows the spatial distribution of the posterior median ARR for the years 1998, 2001, 2004, and 2006. Generally, the spatial pattern is a bit more pronounced in the north, and nearly constant over time. However, some counties have interesting time trends, namely Barreiro, a high industrial region, Loures and Odivelas, all of them small urban areas nearby Lisbon. These trends are shown in detail over all years in Fig. 4. Note the different spatial effects of Barreiro and, Loures and Odivelas. While these last two counties have an increasing geographic influence until the year of 2002, decreasing beyond this year, Loures and Odivelas show an opposite spatial influence.

We further found a strong decreasing trend for all counties. Figure 5 shows the age-standardized mean posterior expected rate for all counties, by gender. Within each year, a similar pattern is obvious that decreases from year to year. The rates for males are higher than for females, but show a stronger decreasing trend.

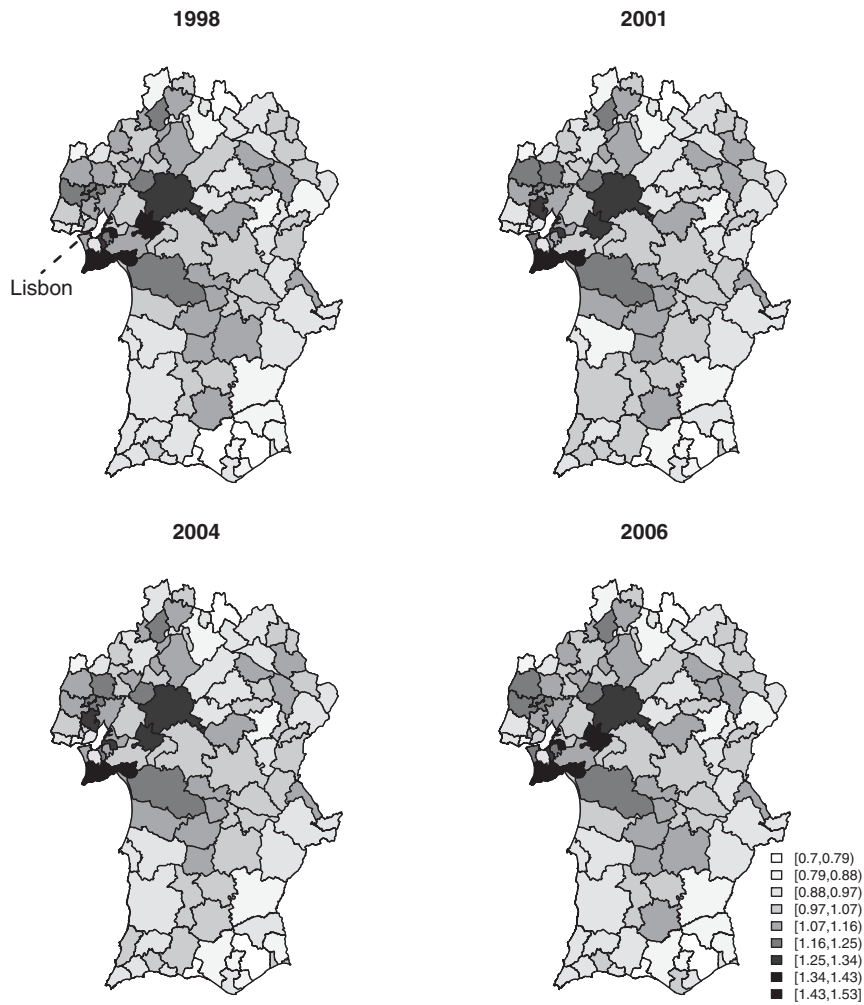


Figure 3 Adjusted relative risks (median) for stomach cancer incidence in Southern Portugal.

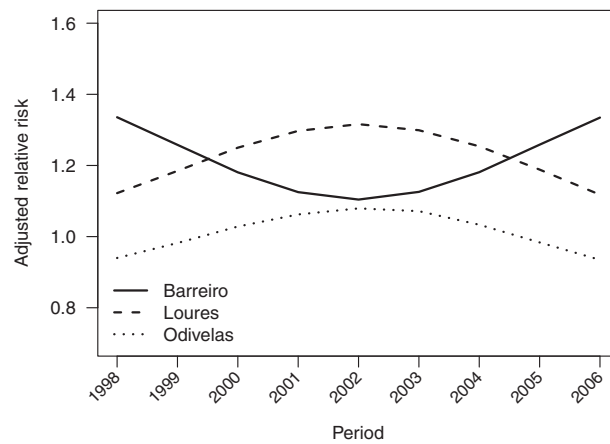


Figure 4 Adjusted relative risks (median) by period of Barreiro, Loures, and Odivelas counties.

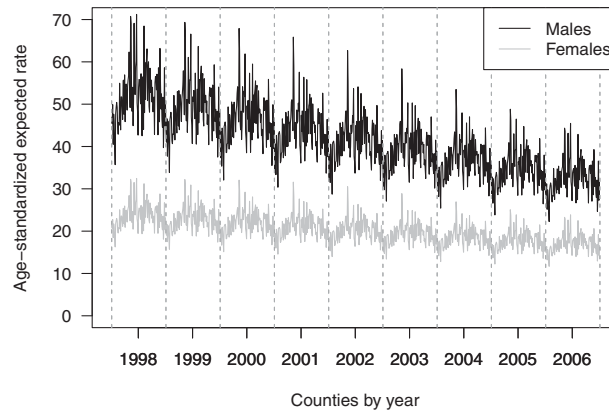


Figure 5 Age-standardized mean posterior expected rate for stomach cancer for each year and county, by gender.

4 Sensitivity analysis

In a Bayesian data analysis, obtained results might depend on the underlying prior specification. It is, therefore, important to check whether results are sensitive to the choice of the hyperpriors.

In the following, we tested five different prior settings. The first three rely on the approach of Sørbye and Rue (2013) described in Section 2.2. Here, the sensitivity analysis can be performed for all parameters simultaneously by varying the parameter U . Since U represents an upper limit for the marginal standard deviation, changing this limit is equal to changing the allowed influence of a random effect in the model. Instead of $U = 0.5$, we set $U = 1$ and $U = 5$ and recalculate all hyperpriors of the five intrinsic GMRFs components as described in Section 2.2. As an alternative, we assume for all five components the same hyperprior distributions. We use, thereby, either the default choice of WinBugs, namely a gamma prior $G(\epsilon, \epsilon)$ where ϵ is small (here 0.001), or the default of INLA, a gamma prior $G(1, 1e-5)$. The detailed prior specifications for all effects and the five scenarios considered are shown in Table 3.

Figure 6 shows the relative risk estimates (median and 95% credible intervals (CI)) of age and period effects for males compared with females, and the joint structured spatial effect \mathbf{u} under all five hyperprior configurations. For \mathbf{u} we selected, for the ease of visibility, 28 counties and only show the median estimate. In particular, the relative age effects do not seem to be very sensitive to a hyperprior change. For certain period effects or counties, the Sørbye and Rue priors based on $U = 5$ and the $G(\epsilon, \epsilon)$ priors produce slightly different estimates as the other three settings. One reason might be, that the marginal standard deviations are allowed to be very high setting $U = 5$, which might lead to

Table 3 Gamma-prior specifications for intrinsic GMRFs used in the sensitivity analysis.

| | Sørbye & Rue $U = 0.5$ | Sørbye & Rue $U = 1$ | Sørbye & Rue $U = 5$ | WinBugs default | INLA default |
|------------------|---------------------------|-------------------------|-------------------------|-----------------|--------------|
| κ_θ | $G(1, 6e-5)$ | $G(1, 3e-4)$ | $G(1, 6e-3)$ | $G(1e-3, 1e-3)$ | $G(1, 1e-5)$ |
| κ_φ | $G(1, 2e-4)$ | $G(1, 7e-4)$ | $G(1, 2e-2)$ | $G(1e-3, 1e-3)$ | $G(1, 1e-5)$ |
| κ_ψ | $G(1, 4e-7)$ | $G(1, 2e-6)$ | $G(1, 4e-5)$ | $G(1e-3, 1e-3)$ | $G(1, 1e-5)$ |
| κ_μ | $G(1, 5e-4)$ | $G(1, 2e-3)$ | $G(1, 5e-2)$ | $G(1e-3, 1e-3)$ | $G(1, 1e-5)$ |
| κ_δ | $G(1, 1e-4)$ | $G(1, 7e-4)$ | $G(1, 2e-2)$ | $G(1e-3, 1e-3)$ | $G(1, 1e-5)$ |

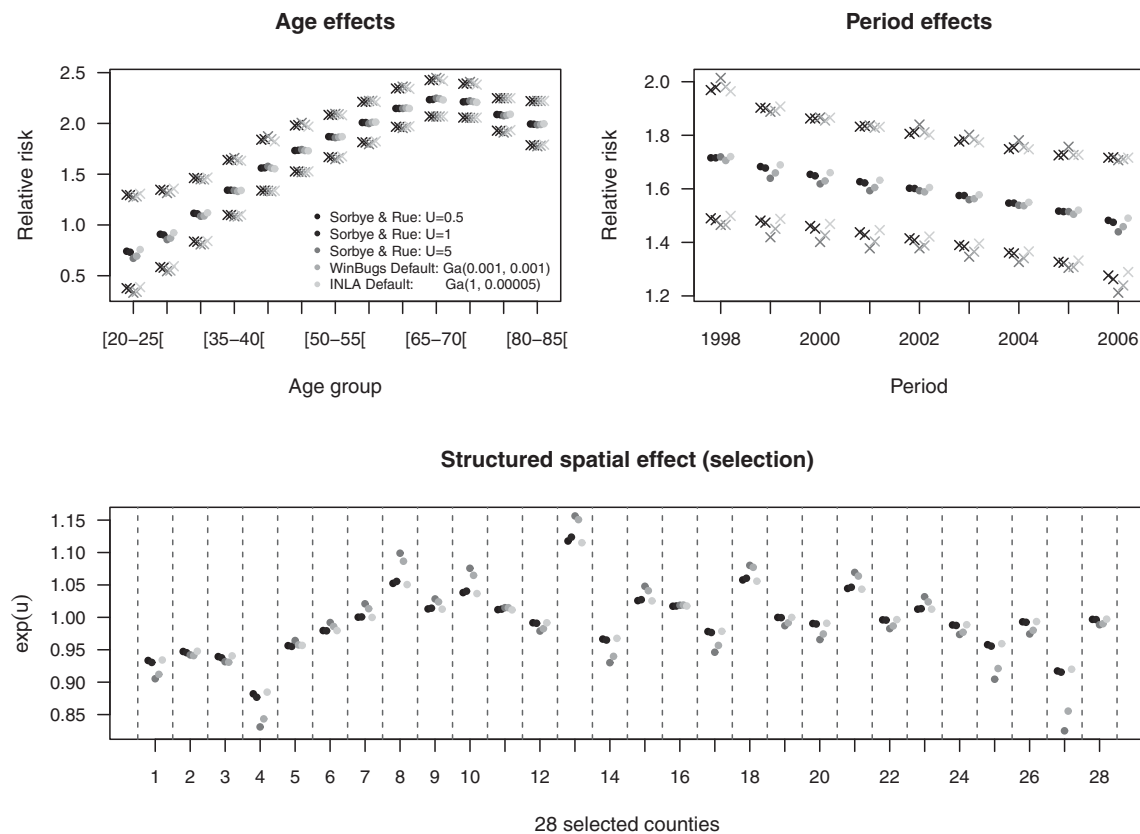


Figure 6 Relative risk estimates of age and period effects for males compared with females (median: filled dot, 95% CI: cross), and the joint structured spatial effect (median: filled dot) for selected counties under all hyperprior configurations.

structures that are not supported by the data. The Gamma(ϵ , ϵ) prior has been criticized for causing unwanted results (Lunn *et al.*, 2009), and this might be confirmed here.

5 Conclusions and discussion

Gender differences and the known fact that age, birth cohort, date of diagnosis, and geographic region, influence the incidence of stomach cancer, have motivated the choice of a Bayesian hierarchical multivariate age-period-cohort (APC) model that also takes into account the geographical component. We found strong gender-specific differences in the analysis of age-specific stomach cancer incidence of 109 counties of Southern Portugal from 1998 to 2006. Men have a higher risk for stomach cancer, and in particular elderly around an age of 70 have a two-times higher risk than females of the same age group. Oestrogen may explain some of these differences (Chandan and Lagergren, 2008), as it acts protectively on women. Several factors, suggested to cause stomach cancer such as dietary habits or smoking, are often attributed to vary between birth cohorts (Malvezzi *et al.*, 2010). However, in the present APC study we have found no indication to assume gender-specific cohort effects. Relative period effects showed that the risk for males is about 1.5 times higher than for females but slowly decreasing. A decreasing behavior over time is present for all counties, and coincides with Malvezzi

et al. (2010), who found a recent decreasing trend in gastric cancer mortality for most European countries. They explained this, partly, by advancements in management (diagnosis and treatment) of gastric cancer and earlier detection through gastroscopy. Furthermore, a better control of *Helicobacter pylori* infections may explain the decreasing incidence in Portugal. Although rates have been decreasing for both genders, the effect is more pronounced for males. To analyze geographical changes over time we inspected the adjusted relative risk (ARR), which was assumed to be the same for males and females. The ARRs are slightly higher in the north, however, in general the spatial pattern is not very clustered, which was already indicated by the data. Except for single small urban counties nearby Lisbon, the ARRs are almost constant over time. The time-varying pattern in the ARRs shown by three counties was guessed to be related to administrative policies. However, this guess was not confirmed by authorities of the Southern Portugal cancer registry (ROR-Sul).

Our model assuming gender-specific age and period effects, joint cohort and spatial effects for males and females and, an interaction term, was selected based on the deviance information criterion (DIC). Hyperprior distributions were thereby carefully chosen according to Fong et al. (2010) and Sørbye and Rue (2013). A sensitivity analysis was carried out to analyze the influence of the chosen hyperprior distributions on the posterior estimates. Average risk estimates of age and period effects for males compared with females, and joint structured spatial effects were compared for five different hyperprior settings, and overall the estimates were relatively stable. Due to the relevance of a sensitivity analysis in the context of Bayesian inference, new approaches are being proposed (Roos and Held, 2011). We inspected the calibration of our model (and other potential model configurations) using a histogram of the probability integral transform (PIT) values (Gneiting et al., 2007). We found that the PIT values deviated from a uniform distribution with more values close to one (results not shown). This might be related to the limited available time span of the data combined with the different aggregation of age groups and periods, which results in a rare occurrence of single cohort groups over time, see Riebler and Held (2010). Omitting cohort effects, however, was not supported by DIC and did not improve the PIT values.

Several recent research studies took place in Northern Portugal to identify stomach cancer risk factors (e.g. Peleteiro et al., 2011, 2012) and, the present population-based study is just another step toward the contribution to a better understanding of stomach cancer incidence trends in Southern Portugal. Data for whole Portugal are, unfortunately, not available to us at the moment. Extending the analysis to the whole country, however, we would expect a stronger spatial pattern, because in the north of Portugal some dietary habits, such as the preference for salty/salted/cured foods, favor the onset of stomach cancer (Shu et al., 2013). The presented modeling framework is also relevant for the spatio-temporal analysis of other diseases where it is relevant to keep the age-specific data structure.

Acknowledgment The authors wish to thank the Editor and two anonymous reviewers for their helpful comments and suggestions. Research was partially sponsored by national funds through the Fundação Nacional para a Ciência e Tecnologia, Portugal - FCT under the projects PEst-OE/MAT/UI0006/2014 and PTDC/MAT/118335/2010.

Conflict of interest

The authors have declared no conflict of interest.

References

- Bernardinelli, L., Clayton, D., Pascutto, C., Montomoli, C., and Ghislandi, M. (1995). Bayesian analysis of space-time variation in disease risk. *Statistics in Medicine* **14**, 2433–2443.
- Berzuni, C., Clayton, D., and Bernardinelli, L. (1993). Bayesian inference on the Lexis diagram. *Bulletin of the International Statistical Institute* **50**, 149–164.

- Besag, J., York, J., and Mollie, A. (1991). Bayesian image restoration with two application in spatial statistics. *Annals of the Institute in Mathematical Statistics* **43**, 1–59.
- Chandanos, E. and Lagergren, J. (2008). Oestrogen and the enigmatic male predominance of gastric cancer. *European Journal of Cancer* **44**, 2397–2403.
- Chen, W. Q., Zheng, R. S., and Zeng, H. M. (2011). Bayesian age-period-cohort prediction of lung cancer incidence in China. *Thoracic Cancer* **2**, 149–155.
- Clayton, D. and Schifflers, E. (1987). Models for temporal variation in cancer rates. II: age-period-cohort models. *Statistics in Medicine* **6**, 469–481.
- Eurostat. *Revision of the European Standard Population—Report of Eurostat's Task Force*: http://epp.eurostat.ec.europa.eu/portal/page/portal/product_details/publication?p_product_code=KS-RA-13-028.
- Fong, Y., Rue, H., and Wakefield, J. (2010). Bayesian inference for generalized linear mixed models. *Biostatistics* **11**, 397–412.
- Gneiting, T. and Raftery, A.E. (2007). Strictly proper scoring rules, prediction and estimation. *Journal of American Statistical Association: Series B* **102**, 359–378.
- Hansell, A., Knorr-Held, L., Best, N., Schmid, V., and Aylin, P. (2003). COPD mortality trends 1950–1999 in England and Wales—Did the 1956 Clean Air Act make a detectable difference? *Epidemiology* **14**, S55.
- Heuer, C. (1997). Modeling of time trends and interactions in vital rates using restricted regression splines. *Biometrics* **53**, 161–177.
- Holford, T. R. (1983). The estimation of age, period and cohort effects for vital rates. *Biometrics* **39**, 311–324.
- Holford, T. R. (2006). Approaches to fitting age-period-cohort models with unequal intervals. *Statistics in Medicine* **25**, 977–993.
- Jacobsen, R., Von Euler, M., Osler, M., Lynge, E., and Keiding, N. (2004). Women's death in Scandinavia—what makes Denmark different? *European Journal of Epidemiology* **19**, 117–121.
- Knorr-Held, L. (2000). Bayesian modelling of inseparable space-time variation in disease risk. *Statistics in Medicine* **19**, 2555–2567.
- Lagazio, C., Biggeri, A., and Dreassi, E. (2003). Age-period-cohort models and disease mapping. *Environmetrics* **14**, 475–490.
- Lexis, W. (1875). *Einleitung in die Theorie der Bevölkerungsstatistik*. Karl J. Trübner, Straßburg.
- Lunet, N. and Barros, H. (2003). Helicobacter Pylori infection and gastric cancer: facing the enigmas. *International Journal of Cancer*, **106**, 953–960.
- Lunn, D. J., Thomas, A., Best, N., and Spiegelhalter, D. (2000). WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing* **10**, 325–337.
- Lunn, D., Spiegelhalter, D., Thomas, A., and Best, N. (2009). The BUGS project: evolution, critique and future directions. *Statistics in Medicine* **28**, 3049–3067.
- Malvezzi, M., Bonifazi, M., Bertuccio, P., Levi, F., La Vecchia, C., Decarli, A., and Negri, E. (2010). An age-period-cohort analysis of gastric cancer mortality from 1950 to 2007 in Europe. *Annals of Epidemiology* **20**, 898–905.
- Nakamura, T. (1986). Bayesian cohort models for general cohort table analyses. *Annals of the Institute of Statistical Mathematics* **38**, 353–370.
- Peleteiro, B., Lopes, C., Figueiredo, C., and Lunet, N. (2011). Salt intake and gastric cancer risk according to *Helicobacter pylori* infection, smoking, tumour site and histological type. *British Journal of Cancer* **104**, 198–207.
- Peleteiro, B., Lunet, N., Xiaogang, W., Afonso, L. P., Mendes, N., Barros, R., Carneiro, F., Almeida, R., and Barros, H. (2012). Association between environmental factors and CDX2 expression in gastric cancer patients. *European Journal of Cancer Prevention* **21**, 423–431.
- Registo Oncológico Nacional (2006). Instituto Português de Oncologia de Lisboa de Francisco Gentil - EPE.
- Riebler, A. and Held, L. (2010). The analysis of heterogeneous time trends in multivariate age-period-cohort models. *Biostatistics* **11**, 57–69.
- Riebler, A., Held, L., and Rue, H. (2012). Estimation and extrapolation of time trends in registry data—borrowing strength from related populations. *The Annals of Applied Statistics* **6**, 304–333.
- Risco de Morrer em Portugal 2006 (2009). Direcção-Geral de Saúde, Direcção de Serviços de Epidemiologia e Estatísticas de Saúde, Divisão de Epidemiologia, Lisboa, PT.
- Rosenberg, P. S., and William, F. A. (2011). Age-period-cohort models in cancer surveillance research: ready for prime time? *Cancer Epidemiology Biomarkers & Prevention* **20**, 1263–1268.
- Roos, M. and Held, L. (2011). Sensitivity analysis in Bayesian generalized linear mixed models for binary data. *Bayesian Analysis* **6**, 259–278.

- Rue, H. and Held, L. (2005). Gaussian Markov Random Fields: Theory and Applications, volume 104 of Monographs on Statistics and Applied Probability. Chapman & Hall, London, UK.
- Rue, H., Martino, S., and Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society B* **71**, 319–392.
- Schrödle, B. and Held, L. (2010). Spatio-temporal disease mapping using INLA. *Environmetrics* **22**, 725–734.
- Shu, L., Wang, X. Q., Wang, S. F., Wang, S., Mu, M., Zhao, Y., Sheng, J., and Tao, F. B. (2013). Dietary patterns and stomach cancer: a meta-analysis. *Nutrition and Cancer* **65**, 1105–1115.
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and van der Linde, A. (2002). Bayesian measures of complexity and fit (with discussion). *Journal of the Royal Statistical Society: Series B* **64**, 583–639.
- Sørbye, S. H. and Rue, H. (2013). Scaling intrinsic Gaussian Markov random field priors in spatial modelling. *Spatial Statistics*, doi: 10.1016/j.spasta.2013.06.004.
- Walter, S. D. (2000). Disease mapping: a historical perspective. In: Elliott, P., Wakefield, J. C., Best, N. G., and Briggs, D. J. (Eds.), *Spatial Epidemiology: Methods and Applications*. Oxford University Press, NY.