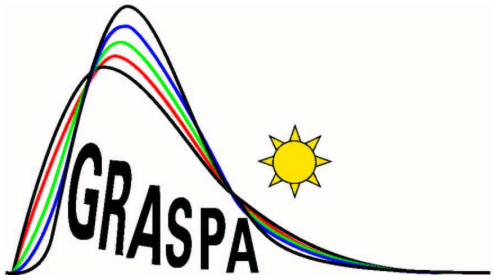
Web Working Papers
by
The Italian Group of Environmental Statistics



Gruppo di Ricerca per le Applicazioni della Statistica
ai Problemi Ambientali

www.graspa.org

A bivariate CAR model for improving the estimation of relative risks

Fedele Greco and Carlo Trivisano

GRASPA Working paper n.28, May 2007

A bivariate CAR model for improving the estimation of relative risks

Fedele Greco, Carlo Trivisano
University of Bologna
Department of Statistics “P. Fortunati”

Abstract: Disease mapping studies have been widely performed at univariate level, that is considering only one disease in the estimated models. Nonetheless, simultaneous modeling of different diseases can be a valuable tool both from the epidemiological and from the statistical point of view. In this paper we propose a model for bivariate disease mapping that generalises the univariate CAR distribution. The proposed model is proven to be an effective alternative to existing bivariate models, mainly because it overcome some restrictive hypotheses underlying models previously proposed in this context. Model performances are checked via a simulation study and via application to some real case studies.

Keywords: Areal data, multivariate Conditional Auto-Regressive (CAR) Models, Disease mapping, Hierarchical Bayesian models.

1. Introduction

Disease mapping has proved to be a useful epidemiological tool for describing geographical variations of diseases. The techniques developed in this field have seen a considerable improvement once disease data have incorporated geographical information about the study units. Because of confidentiality restrictions, disease data are almost always available at small area level and the analyses have to be performed at an aggregate level on count data.

The first step for studying the disease incidence in specific subsets of a study region is to assess what disease incidence should be expected in such subset: then it is possible to compare the expected with the observed incidence. Expected incidence is the one that would be observed if the disease risk was constant over all the study region and the spatial variations in incidence were due uniquely to population density and structure. The expected incidence reflects the “null hypothesis spatial distribution” of the cases. The focus of attention is on identifying features of the spatial distribution of the disease rate that are not captured by the null hypothesis distribution [12]. This identification is performed via the comparison between observed and expected incidence by relative

risk estimation. When the disease under study is rare, data are heavily affected by random variability and the estimates of the relative risk at the small area level are unstable, particularly in areas characterised by low population amounts. The main aim of disease mapping techniques is the depuration of data from random noise, so that the true underlying distribution of the risk can be identified. In the most common models, depuration from random noise is achieved by exploiting the available spatial information: the relative risk estimate in each area is improved by borrowing strength from the neighbouring areas. In this context Bayesian hierarchical models have received considerable attention, because of their ability to easily produce smooth maps depurated from random noise. The smoothing of relative risk spatial distribution is often achieved by including random effects in the model that take account of spatial and non-spatial heterogeneity in the data. As a result, the relative risk estimates' standard error is considerably reduced with respect to the estimates obtained by ignoring spatial information and overdispersion. Maps based on the estimates produced by Bayesian models are more easily interpretable. They can generate hypotheses on the disease aetiology and can help public authority in planning health policies.

Disease mapping studies have been widely performed at univariate level, that is considering relative risk estimation for one disease. Nonetheless, simultaneous modelling of different diseases can be a valuable tool both from the epidemiological and from the statistical point of view. From the epidemiological point of view, joint modelling of diseases can increase the understanding of diseases dynamics and of the relationships between diseases incidence. The merit of joint modelling can be high if the considered diseases share risk factors or if the presence of a disease encourages (or inhibits) the occurrence of a different disease. From the statistical point of view, an evident advantage in joint disease mapping is that, if the disease risks are correlated, standard errors of the estimates obtained via univariate modelling can be sensibly reduced. Moreover, estimates for rare diseases can borrow strength from more diffuse diseases. In fact, correlation between diseases within areas, between areas within diseases and between areas and diseases constitute valuable information contained in the data that can be used to increase statistical efficiency of the estimates.

While models for univariate disease modelling are widely diffused both in the Bayesian and the classical contexts, multivariate models are far less diffuse and have received attention only recently. One of the earlier approaches is based on the multilevel modelling theory: models include fixed and random effects. Parameter estimation is performed using iterative generalised least squares [11, 13]. From the Bayesian point of view, a shared component model has been proposed in [7], useful for modelling variation of two or more diseases that share common risk factors.

The most promising approaches generalise the Conditional Auto-Regressive (CAR) distribution [1] in a multivariate framework [3, 5, 9, 10]. These generalizations face in various ways the problem of

obtaining a valid joint distribution for multivariate risks. In this paper we propose a new extension of the CAR model in the bivariate setting that we dub Bivariate CAR (*BCAR*) model. Our proposal overcome some of the simplifications on which other proposed models are based.

This work is organized as follows: in Section 2 we briefly review some existing approaches to multivariate disease risk modelling. In Section 3 we propose a new approach for bivariate areal data, specifying the joint distribution via bivariate full conditional distributions. In Section 4 a simulation study is performed, while in Section 5 we show results obtained by applying the proposed model to a couple of data sets referring to the Emilia Romagna Region (Italy).

2. From univariate to multivariate CAR models

Consider a study region subdivided in n contiguous areas. Let Y_i and e_i denote respectively the observed and expected counts for a disease in the i-th area. Expected counts for each area can be obtained by applying a standard table of sex and age group-specific rates to the area-specific background population subdivided by age and sex. The likelihood for the observed counts is usually specified as:

$$\begin{aligned} Y_i \mid \theta_i &\sim Poisson\left(e_i\theta_i\right) \\ log\left(\theta_i\right) &= \gamma + \phi_i \end{aligned} \qquad (1)$$

where θ_i is the true relative risk for the disease in area i and γ is an intercept. Conditionally on θ_i, the counts are modelled as independent Poisson variables. Spatial dependence is modelled at the second level of the hierarchy by imposing a probabilistic structure on the distribution of the vector $\boldsymbol{\varphi} = \left(\phi_1,...,\phi_i,...,\phi_n\right)$ that is usually assumed to follow a n-dimensional Normal distribution:

$$\boldsymbol{\varphi} \sim MVN\left(0, \boldsymbol{\Sigma}\right) \qquad (2)$$

The covariance matrix $\boldsymbol{\Sigma}$ is specified in order to model spatial correlation. Model (1)-(2) characterise a Poisson log-Normal mixture. Distribution (2) can be more easily specified in terms of conditional distributions: this is the core idea underlying univariate CAR models [1]. These models can be included in the class of Gaussian Markov Random Field (MRF) models [14]. The dependence among the elements of $\boldsymbol{\varphi}$ is limited to the neighbouring areas, that is the conditional

distribution of the relative risk in a given area is independent from the risk in non-neighbouring areas. Thus, to specify a MRF process, a neighbouring structure has to be selected. Various neighbouring structures have been proposed. In this paper, following the most popular approach, we consider areas as neighbours if they share a boundary, and information about the neighbouring structure is summarized in the $n \times n$ adjacency matrix $\boldsymbol{W}$, with entries $w_{ij} = 1$ if areas *i* and *j* share a common boundary, $w_{ij} = 0$ otherwise. In the univariate setting, the full conditional distributions for the spatial random effects are specified as follows:

$$\phi_i \mid \phi_j, j \sim i, \tau^{-1} \sim N\left(\delta \sum_{j \sim i} c_{ij} \phi_j, \frac{\tau^{-1}}{m_i} \right) \tag{3}$$

where $j \sim i$ denotes that area *i* is a neighbour of region *j* and m_i is the number of neighbours of the *i*-th area. δ is a smoothing parameter controlling the strength of spatial dependence and τ^{-1} is the variance parameter scaled, for each area, by the number of its neighbours so that the conditional variance is inversely proportional to the spatial information. The joint distribution (2) is uniquely determined by the *n* full conditionals (3) and can be expressed as:

$$\boldsymbol{\varphi} \sim MVN\left(0, \tau\left[\boldsymbol{D}\left(\boldsymbol{I} - \delta \boldsymbol{C}\right)\right]^{-1}\right) \tag{4}$$

where $\boldsymbol{D}$ is a diagonal matrix with diagonal elements denoting the number of neighbours of the *i*-th area and $\boldsymbol{C} = \boldsymbol{D}^{-1}\boldsymbol{W}$ is the scaled adjacency matrix with entries $c_{ij} = w_{ij} / m_i$. This specification of $\boldsymbol{C}$ and $\boldsymbol{D}$ matrices guarantees the positive definiteness of the joint distribution covariance matrix in distribution (4), provided $|\delta| < 1$. Various alternative univariate CAR models can be obtained by specifying different structures of $\boldsymbol{D}$ and $\boldsymbol{C}$ matrices.

For multivariate disease mapping, a natural way for approaching statistical modelling, is the generalisation of the CAR distribution to its multivariate version. In what follows, a number of models proposed in the bivariate setting are illustrated.

Let Y_{ik} and e_{ik} denote the observed counts in the *i*-th area, for the *k*-th disease. Like in (1), conditional on the relative risks, the counts are usually modelled as conditionally independent Poisson variables:

$$Y_{ik} \mid \theta_{ik} \sim Poisson\left(e_{ik}\theta_{ik}\right) \qquad i=1,\ldots,n,\ k=1,2 \tag{5}$$

$$\log(\theta_{ik}) = \gamma_k + \phi_{ik}$$

Where γ_k is a disease-specific intercept and ϕ_{ik} is the spatial random effect at area *i* for disease *k*. The bivariate analogous of (2) for the spatial random effects is:

$$\begin{pmatrix} \boldsymbol{\varphi}_{\bullet 1} \\ \boldsymbol{\varphi}_{\bullet 2} \end{pmatrix} \sim MVN\left(\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{12} & \boldsymbol{\Sigma}_2 \end{pmatrix} \right) \tag{6}$$

where $\boldsymbol{\varphi}_{\bullet k}$ is the *n*-dimensional vector of spatial random effects for disease *k*, and the blocks of the covariance matrix describe the correlations characterising each disease in the study region ($\boldsymbol{\Sigma}_k$, *k*=1,2) and the relationships between the risks for the considered diseases ($\boldsymbol{\Sigma}_{12}$).

As in the univariate setting, specification of distribution (6) starts by specifying the full conditional distributions. Following [14], under the MRF assumption, the full conditional distributions analogous to (3) can be specified as:

$$\boldsymbol{\varphi}_{i\bullet} \mid \boldsymbol{\varphi}_{j\bullet}, j \sim i, \boldsymbol{\Gamma}_i \sim N\left(\boldsymbol{A}\sum_{j\sim i} \boldsymbol{\beta}_{ij}\boldsymbol{\varphi}_{j\bullet}, \boldsymbol{\Gamma}_i \right) \qquad i=1,\ldots,n \tag{7}$$

where $\boldsymbol{\varphi}_{i\bullet}^{'} = (\phi_{i1}, \phi_{i2})$ is the 2-dimensional vector of spatial random effects for both diseases in the *i*-th area, while $\boldsymbol{A}$, $\boldsymbol{\beta}_{ij}$ and $\boldsymbol{\Gamma}_i = (m_i \boldsymbol{\Lambda})^{-1}$ are 2×2 matrices. Each spatial random effects bivariate vector, conditionally on the neighbouring areas, follows a bivariate normal distribution. Its mean is obtained as a weighted mean of the neighbouring areas with weights $\boldsymbol{\beta}_{ij}$. Matrix $\boldsymbol{A}$ contains the spatial smoothing parameters and $\boldsymbol{\Gamma}_i$ is the bivariate conditional covariance matrix obtain for each area as the inverse of a common precision matrix $\boldsymbol{\Lambda}$ times the number of neighbours m_i. With $\boldsymbol{F}$ and $\boldsymbol{G}$ being $2n\times 2n$ matrices such that $\boldsymbol{F} = Block(\boldsymbol{A}\boldsymbol{\beta}_{ij})$ and $\boldsymbol{G}$ block-diagonal with diagonal elements $\boldsymbol{\Gamma}_i^{-1}$, the joint distribution becomes:

$$\begin{pmatrix} \boldsymbol{\varphi}_{1\bullet} \\ \boldsymbol{\varphi}_{2\bullet} \\ \cdots \\ \boldsymbol{\varphi}_{n\bullet} \end{pmatrix} \sim N\left(\boldsymbol{0}, \left[\boldsymbol{G}(\boldsymbol{I} - \boldsymbol{F}) \right]^{-1} \right) \tag{8}$$

Different multivariate CAR models can be obtained according to the specification of $\boldsymbol{A}$, $\boldsymbol{\beta}_{ij}$ and $\boldsymbol{\Gamma}_i$ matrices. The most general approach for specifying conditional distributions (7) includes in the $\boldsymbol{A}$ matrix four different smoothing parameters: the diagonal entries describe the spatial correlation characterising the spatial structure of each disease, while the off diagonal element describes the relationship between the considered diseases in the neighbouring areas.

It can be shown [9] that, with matrices $\boldsymbol{C}$ and $\boldsymbol{D}$ previously specified in (4):

$$\begin{pmatrix} \boldsymbol{\varphi}_{\bullet 1} \\ \boldsymbol{\varphi}_{\bullet 2} \end{pmatrix} \sim MVN\left(\begin{pmatrix} \boldsymbol{0} \\ \boldsymbol{0} \end{pmatrix}, \begin{pmatrix} (\boldsymbol{D}-\delta_1\boldsymbol{W})\Lambda_{11} & (\boldsymbol{D}-\delta_3\boldsymbol{W})\Lambda_{12} \\ (\boldsymbol{D}-\delta_3\boldsymbol{W})\Lambda_{12} & (\boldsymbol{D}-\delta_2\boldsymbol{W})\Lambda_{22} \end{pmatrix}^{-1} \right) \tag{9}$$

where smoothing parameters δ_1, δ_2 and δ_3 are functions of model parameters (and in particular of the $\boldsymbol{A}$ and $\boldsymbol{\Lambda}$ matrices) in a way that depend on model specification. One difficulty concerning this specification is that conditions on the smoothing parameters for positive definiteness of the joint distribution precision matrix depend on the unknown matrix $\boldsymbol{\Lambda}$. This causes some difficulties both in determining these conditions and in implementing MCMC algorithms for model estimation. To avoid these difficulties different strategies have been adopted. A naive generalisation of the CAR distribution for the multivariate setting can be obtained through the assumption of equality of the spatial smoothing parameters, obtaining the specification that is dubbed $MCAR(\delta, \boldsymbol{\Lambda})$ in [5]. According to this specification, the condition for positive definiteness of the covariance matrix reduces to $|\delta| < 1$ as in the univariate setting.

In [3] the $MCAR(\delta, \boldsymbol{\Lambda})$ model is extended in order to allow two smoothing parameters obtaining the $MCAR(\delta_1, \delta_2, \boldsymbol{\Lambda})$ model via the Cholesky decomposition of the matrices $\boldsymbol{D}-\delta_k\boldsymbol{W}$ in (9). Under this specification conditions for positive definiteness are $|\delta_1| < 1$ and $|\delta_2| < 1$. A similar result is obtained in [5] via the spectral decomposition of the $\boldsymbol{D}-\delta_k\boldsymbol{W}$ matrices.

In [10] a bivariate CAR model is built by specifying the univariate full conditional distributions $p(\phi_{ik} | \phi_{il}, \boldsymbol{\varphi}_{-i\bullet})$, where $\boldsymbol{\varphi}_{-i\bullet}$ denotes the set of the two-dimensional vectors in all areas but the *i*-th, deriving the following joint distribution:

$$\begin{pmatrix} \boldsymbol{\varphi}_{\bullet 1} \\ \boldsymbol{\varphi}_{\bullet 2} \end{pmatrix} \sim MVN\left(\begin{pmatrix} \boldsymbol{0} \\ \boldsymbol{0} \end{pmatrix}, \begin{pmatrix} \tau_1^{-1}\left(\left(2\boldsymbol{D}^{-1}+1\right)^{-1}-\delta_1\boldsymbol{W}\right) & -(\tau_1\tau_2)^{-1/2}(\delta_0\boldsymbol{I}+\delta_3\boldsymbol{W}) \\ -(\tau_1\tau_2)^{-1/2}(\delta_0\boldsymbol{I}+\delta_3\boldsymbol{W}) & \tau_2^{-1}\left(\left(2\boldsymbol{D}^{-1}+1\right)^{-1}-\delta_2\boldsymbol{W}\right) \end{pmatrix} \right)$$

The conditions on the δ_0, δ_1, δ_2, and δ_3 for positive definiteness of the joint covariance matrix ($|\delta_l| < 1$, $l = 0,...,3$) found in [10] are sufficient but not necessary and extension to *p*-variate framework seems difficult.

The approach proposed in [9] is based on the direct specification of the joint distribution for the multivariate spatial process through the specification of marginal and conditional distributions. In the bivariate setting, the joint distribution for the 2*n* dimensional vector $\boldsymbol{\varphi}$ is:

$$p\begin{pmatrix}\boldsymbol{\varphi}_{\bullet 1}\\ \boldsymbol{\varphi}_{\bullet 2}\end{pmatrix} = p\left(\boldsymbol{\varphi}_{\bullet 1}\middle|\boldsymbol{\varphi}_{\bullet 2}\right) p\left(\boldsymbol{\varphi}_{\bullet 2}\right) \tag{10}$$

where the marginal distribution $p\left(\boldsymbol{\varphi}_{\bullet 2}\right)$ is specified as a univariate CAR distribution: $\boldsymbol{\varphi}_{\bullet 2} \sim N\left(0,\left[\left(\boldsymbol{D}-\delta_2\boldsymbol{W}\right)\tau_2\right]^{-1}\right)$. Conditionally on $\boldsymbol{\varphi}_2$, the distribution of $\boldsymbol{\varphi}_1$ is again CAR $\boldsymbol{\varphi}_{\bullet 1}\middle|\boldsymbol{\varphi}_{\bullet 2} \sim N\left(\left(\eta_0\boldsymbol{I}+\eta_1\boldsymbol{W}\right)\boldsymbol{\varphi}_2,\left[\left(\boldsymbol{D}-\delta_1\boldsymbol{W}\right)\tau_1\right]^{-1}\right)$. We indicate with $GMCAR\left(\delta_1,\delta_2,\eta_0,\eta_1,\tau_1,\tau_2\right)$ the joint distribution of $\boldsymbol{\varphi}$. The relationships between risks for the two diseases is modelled in the mean vector, where parameter η_0 models the relationship between risks for diseases 1 and 2 in the same areas, while parameter η_1 captures the relationship between the risk for disease 1 and risks for disease 2 in the neighbouring areas. In [9] a simulation study shows that the *GMCAR* model outperforms the other models briefly described in this section. The main drawback of the *GMCAR* model is that the conditioning order has effect on the results even if the choice of the conditioning order can be performed by means of some model selection criteria. We believe that, in this context, a "symmetric" modelling approach is more natural.

3. A bivariate CAR (*BCAR*) model

As an extension of the univariate CAR model (2), we start by specifying the bivariate full conditional distributions as:

$$\boldsymbol{\varphi}_{i\bullet}\middle|\boldsymbol{\varphi}_{-i\bullet},\boldsymbol{\Lambda},\boldsymbol{A} \sim N\left(\boldsymbol{\mu}_i,\boldsymbol{\Gamma}_i\right) \qquad i=1,\ldots,n \tag{11}$$

The conditional covariance matrix is $\boldsymbol{\Gamma}_i = \left(m_i\boldsymbol{\Lambda}\right)^{-1}$ where

$$\boldsymbol{\Lambda} = \begin{pmatrix} \tau_1 & \tau_{12} \\ \tau_{12} & \tau_2 \end{pmatrix} = \begin{pmatrix} \sigma_1 & \sigma_{12} \\ \sigma_{12} & \sigma_2 \end{pmatrix}^{-1} \tag{12}$$

The *i*-th conditional covariance matrix is scaled by the number of neighbours, thus the conditional variability is inversely proportional to the number of neighbouring areas. Each conditional mean is modelled as:

$$\mu_{ik} = \frac{1}{2m_i}\left(\alpha_k \sum_{j\sim i}\phi_{jk} + \alpha_{kl}\sum_{j\sim i}\phi_{jl}\right) \qquad j,i=1,\ldots,n \;\; j \neq i\,; \;\; l,k=1,2 \;\; l \neq k$$

i.e. the mean vector is modelled as a weighted mean of the risks for both diseases in the neighbouring areas with α_k and α_{kl} parameters controlling the strength of the association.

Now let $\boldsymbol{A}_{ij} = \begin{pmatrix} \alpha_1 & \alpha_{12} \\ \alpha_{21} & \alpha_2 \end{pmatrix}\frac{w_{ij}}{2m_i} = \boldsymbol{A}\frac{w_{ij}}{2m_i}$. For sake of convenience we set $\boldsymbol{A}_{ii} = -\boldsymbol{I}_{2\times 2}$. Parameters α_1 and α_2 capture the disease specific spatial dependence: in what follows we set them as positive in order to avoid negative spatial correlation. Parameters α_{lk} are bridging parameters controlling the dependence between disease *l* in a given area and disease *k* in the neighbouring areas. Following [14], given the *n* conditional distributions, the joint distribution of $\left(\boldsymbol{\varphi}_{1\bullet}', \boldsymbol{\varphi}_{2\bullet}', ..., \boldsymbol{\varphi}_{n\bullet}'\right)$ is multivariate normal with mean $\boldsymbol{\mu}' = \left(\boldsymbol{\mu}_{1\bullet}', \boldsymbol{\mu}_{2\bullet}', ..., \boldsymbol{\mu}_{n\bullet}'\right)$ and covariance matrix $\boldsymbol{\Sigma} = \left[Block\left(-\boldsymbol{\Gamma}_i^{-1}\boldsymbol{A}_{ij}\right)\right]^{-1}$, provided that conditions for symmetry and positive definiteness are satisfied.

The condition for symmetry of the joint covariance matrix is $\boldsymbol{A}_{ij}\boldsymbol{\Gamma}_j = \boldsymbol{\Gamma}_i\boldsymbol{A}_{ji}'$. It can be easily shown that these matrices have the same diagonal entries. Equality of off-diagonal entries can obtained by imposing $\left(\boldsymbol{A}_{ij}\boldsymbol{\Gamma}_j\right)_{12} = \left(\boldsymbol{\Gamma}_i\boldsymbol{A}_{ji}'\right)_{21} = \alpha_0$ where $\left(\boldsymbol{X}\right)_{12}$ denotes the first row and second column entry of a generic matrix ***X***. It turns out that off-diagonal elements of the $\boldsymbol{A}$ matrix can be set to

$$\alpha_{12} = \frac{\alpha_0 - \sigma_{12}\alpha_1}{\sigma_2} \qquad \alpha_{21} = \frac{\alpha_0 - \sigma_{12}\alpha_2}{\sigma_1} \tag{13}$$

The following theorem states the conditions for positive definiteness of the joint covariance matrix $\boldsymbol{\Sigma}$.

Theorem: Given the *n* full conditionals distributions (11), and the bridging parameters defined in (13), the joint distribution of $\boldsymbol{\varphi}$ is a valid multi-normal distribution provided:

i) $0 < \alpha_k < \dfrac{2\sigma_l}{\sigma_l + |\sigma_{12}|}$ $\quad l,k=1,2 \quad l \neq k$; ii) $I_{\alpha_0} < \alpha_0 < S_{\alpha_0}$

where

$$I_{\alpha_0} = max\left\{\alpha_1(\sigma_{12}+\sigma_2)-2\sigma_2, \alpha_2(\sigma_{12}+\sigma_1)-2\sigma_1\right\}$$
$$S_{\alpha_0} = min\left\{\alpha_1(\sigma_{12}-\sigma_2)+2\sigma_2, \alpha_2(\sigma_{12}-\sigma_1)+2\sigma_1\right\}$$

Proof: in the Appendix.

While existing generalisations of the CAR distribution are shown to be valid using a diagonal dominance argument for the joint precision matrix $Block\left(-\boldsymbol{\Gamma}_i^{-1}\boldsymbol{A}_{ij}\right)$, we switch the problem to row-diagonal dominance of the non-symmetric matrix $Block\left(-\boldsymbol{A}_{ij}\right)$. We believe that this approach is useful since the latter matrix has generally a simpler structure with respect to the precision matrix. This conditional specification leads to the following joint distribution:

$$\begin{pmatrix}\boldsymbol{\varphi}_{\bullet 1}\\ \boldsymbol{\varphi}_{\bullet 2}\end{pmatrix}\Big|\boldsymbol{A},\boldsymbol{\Lambda} \sim N\left(\begin{pmatrix}\boldsymbol{0}\\ \boldsymbol{0}\end{pmatrix}, \begin{pmatrix}(\boldsymbol{D}-\delta_1\boldsymbol{W})\tau_1 & (\boldsymbol{D}-\delta_3\boldsymbol{W})\tau_{12}\\ (\boldsymbol{D}-\delta_3\boldsymbol{W})\tau_{12} & (\boldsymbol{D}-\delta_2\boldsymbol{W})\tau_2\end{pmatrix}^{-1}\right) \tag{14}$$

where:

$$\delta_k = \alpha_k - \alpha_0\tau_{12}\left(\rho^2-1\right)+\rho^2\alpha_l \qquad k,l=1,2 \quad k \neq l$$

$$\delta_3 = \alpha_1+\alpha_2-\alpha_0\tau_{12}\left(\frac{\rho^2-1}{\rho^2}\right)$$

Here $\rho^2 = \tau_{12}^2/\tau_1\tau_2$ denotes the conditional correlation coefficient between the risks in a given area. Thus the disease specific smoothing parameters δ_k, k=1,2, generated by the bivariate model are

obtained as the univariate smoothing parameters α_k, k=1,2, modified by the further smoothing parameters in the $\boldsymbol{A}$ matrix. The effect of such modification increases if the between diseases correlation increases. If $\boldsymbol{\Lambda}$ is diagonal, the prior distribution for α_k does not depend on $\boldsymbol{\Lambda}$ and the usual condition for positive definiteness in the univariate setting is recovered. The upper bound is 2 instead of 1 because the scaling factor in the conditional mean is $1/2m_i$ instead of $1/m_i$. Otherwise, the upper bound is lower than 2.

3.1 Bivariate disease mapping using the *BCAR* prior

In what follows we describe how the *BCAR* distribution can be used in the context of bivariate disease modelling. Conditionally on the model parameters, counts are modelled as independent Poisson variables:

$$Y_{ik} \,|\theta_{ik} \sim Poisson\left(e_{il}\theta_{ik}\right) \qquad i=1,\ldots,n \qquad k=1,2$$

where θ_{ik} denotes the relative risk in the i-th area for the disease k which is modelled as follows:

$$\log\left(\theta_{ik}\right) = \gamma_k + \phi_{ik}$$

where γ_k is a disease specific intercept and ϕ_{ik} denotes the spatial random effect. For 2-dimensional column vector vectors $\boldsymbol{\varphi}_{i\bullet}$ the *BCAR* prior distribution (11) is employed. Model hierarchy is completed via prior specification.

The set of hyperparameters for which a prior distribution has to be specified is constituted by the conditional precision matrix $\boldsymbol{\Lambda}$ and by the smoothing parameters in the $\boldsymbol{A}$ matrix. Since the latter depends on the elements of the $\boldsymbol{\Lambda}$ matrix, we first specify a prior distribution for this covariance matrix as

$$\boldsymbol{\Lambda} \sim Wishart\left(\boldsymbol{R}, d\right)$$

Prior distributions for the disease specific smoothing parameters are specified conditionally on $\boldsymbol{\Lambda}$ as uniform distributions with a stochastic extreme.

$$\alpha_k \left| \boldsymbol{\Lambda} \sim U\left(0, \frac{2\sigma_l}{\sigma_l + \left|\sigma_{12}\right|}\right) \right. \quad l,k = 1,2 \tag{15}$$

Conditionally on $\boldsymbol{\Lambda}$, α_1 and α_2, the prior distribution for α_0 is specified as a uniform distribution with stochastic extremes determined in the theorem above.

$$\alpha_0 \left| \boldsymbol{\Lambda}, \alpha_1, \alpha_2 \sim U\left(I_{\alpha_0}, S_{\alpha_0}\right) \right. \tag{16}$$

Since posterior distribution are not obtainable analytically, MCMC methods have to be implemented for model estimation. We implemented our model in the OpenBugs software. Distributions available in OpenBugs for spatial smoothing (`car.normal` and `car.proper`) can not be used to implement the *BCAR* model since we specify conditional distributions for the spatial effects associated with both diseases. Moreover, we point out that the variance of the conditional normal distribution (11) is a function of $\boldsymbol{\Lambda}$ and not $\boldsymbol{\Lambda}$ itself. For this reason, since the Wishart distribution available in OpenBugs can only be used as a prior for a normal distribution covariance matrix, it was necessary to build the Wishart prior by using a "trick" for specifying a new prior distribution within the OpenBugs software(the code is available on request from the authors). We tested this "trick" in a controlled setting (not reported) to assess its right functioning.

4. A simulation study

For evaluating the performances of the *BCAR* distribution, we perform a simulation study and compare results obtained with the *BCAR* model with results obtained by estimating the *GMCAR* model proposed by [9] that has been shown to perform better than other proposed models for multivariate areal data.

The simulation study is based on a spatial grid of n=95 areas constituted by a subset of the Emilia Romagna Region municipalities. The simulation experiment follows closely the design proposed in [9]. Model performances are evaluated by simulating data from a Normal-Normal model instead of from a Poisson log-Normal model in order to speed up computation. We assume that data Y_{ik} arise from a Gaussian model:

$$Y_{ik} \left| \theta_{ik}, \upsilon^2 \sim N\left(\theta_{ik}, \upsilon^2\right) \right. \qquad i=1,\ldots,n, \qquad k=1,2 \tag{17}$$

$$\theta_{ik} = \gamma_k + \phi_{ik} \qquad i=1,\ldots,n, \qquad k=1,2 \tag{18}$$

where γ_k are fixed constants indicating the mean for the vector $\boldsymbol{\theta}_{\bullet k}$ in the study region and ϕ_{ik} is a zero-mean spatial random effect in the *i*-th area.

We perform two different simulation studies. In Study 1 we generate spatial random effects from the *GMCAR* model described in section 2, i.e. $(\boldsymbol{\varphi}_{\bullet 1}, \boldsymbol{\varphi}_{\bullet 2}) \sim GMCAR(\delta_1, \delta_2, \eta_0, \eta_1, \tau_1, \tau_2)$. True parameter values are shown in Table 1.

(*Table* 1 *about here*)

This setting corresponds to true parameter values set in [9].

In Study 2 we generate spatial random effects $(\boldsymbol{\varphi}_{1\bullet}', \boldsymbol{\varphi}_{2\bullet}', ..., \boldsymbol{\varphi}_{n\bullet}')$ from the *BCAR* model (14) with true parameters values reported in Table 2. We stress that the interpretation of the true smoothing parameter values is different in the two studies, due to their range and to the way they enter the joint posterior distribution.

(*Table* 2 *about here*)

Study 2 is conceived in order to maintain the conditional linear correlation between $\boldsymbol{\varphi}_{\bullet 1}$ and $\boldsymbol{\varphi}_{\bullet 2}$ equal to .6 for both settings. In setting 1 the values of (13) are $\alpha_{12} = \alpha_{21} = .56$ while in setting 2 $\alpha_{12} = .2$ and $\alpha_{21} = .6$. Thus in setting 1 the relationship between $\boldsymbol{\varphi}_{\bullet 1}$ and $\boldsymbol{\varphi}_{\bullet 2}$ is symmetric, while in setting 2 the relationship is not symmetric. Moreover, even if the conditional linear correlation between $\boldsymbol{\varphi}_{\bullet 1}$ and $\boldsymbol{\varphi}_{\bullet 2}$ is the same in the two settings, in setting 2 $\boldsymbol{\varphi}_{\bullet 1}$ is more variable than in setting 2: this help us in evaluating the effect of unbalanced dispersions on the results. It is worth noting that this happens in Study 1, where unconditional variances are unbalanced even if $\tau_1 = \tau_2$ because of the different role of these parameters: one is conditional while the other is unconditional.

We simulated *T*=500 data sets for each setting. For each simulated data sets we estimated the *GMCAR* model and the *BCAR* model. At data level both models are specified as in (17) and (18). As prior distributions for parameters υ and γ_k, $IG(1, 0.1)$ and $N(0, 10)$ are respectively assumed. Specification for the compared models differs only for the prior distribution on the spatial random effects. Thus for the *GMCAR* model the prior $(\boldsymbol{\varphi}_{\bullet 1}, \boldsymbol{\varphi}_{\bullet 2}) \sim GMCAR(\delta_1, \delta_2, \eta_0, \eta_1, \tau_1, \tau_2)$ is used.

Normal $N(0,10)$ priors are specified for parameters η_0 and η_1. Vague Gamma priors, specifically $G(1,0.1)$, are assigned to precision parameters τ_1 and τ_2, while uniform $U(0,1)$ priors are assigned to smoothing parameters δ_1 and δ_2. *GMCAR* model is estimated using the OpenBugs code made available at the website www.biostat.umn.edu/~brad/software.html.

As regards prior specification for the *BCAR* model, we specify full conditionals distributions for spatial random effects as $\boldsymbol{\varphi}_{i\bullet}|\boldsymbol{\varphi}_{-i\bullet},\boldsymbol{\Lambda},\boldsymbol{A} \sim N\left(\boldsymbol{\mu}_i,\boldsymbol{\Gamma}_i^{-1}\right)$ as discussed in section 3.1. A Wishart distribution is used for the $\boldsymbol{\Lambda}$ matrix, i.e. $\boldsymbol{\Lambda} \sim Wishart(\boldsymbol{I},2)$, where $\boldsymbol{I}$ is the 2×2 identity matrix and 2 are the smallest possible number of degrees of freedom in order to express vague prior beliefs; prior distributions for the smoothing parameters are specified conditionally on the $\boldsymbol{\Lambda}$ matrix as in (15) and (16).

For each simulated data set and for each model, inference is based on 10.000 samples from the MCMC algorithm, after a burn-in of 15.000 iterations.

Model performances are compared via Deviance Information Criterion (DIC) [15] and Average Mean Squared Error (AMSE). DIC is a model selection criterion according to which the model performance is evaluated as the sum of a measure of fit, the posterior mean of the deviance $\bar{D} = E\left[-2log\left(f\left(y|parameters\right)\right)\right]$, and a measure of complexity, the effective number of parameters p_D, obtained as the difference between the deviance posterior mean and the deviance evaluated at the parameters posterior mean. Thus $DIC = \bar{D} + p_D$: a model is preferred if it shows a lower DIC value.

For each simulation setting, the AMSE is obtained as the mean of squares of differences between true (θ_{ik}^t) and estimated values for each simulated data set ($\hat{\theta}_{ik}^t$), t=1,…,T:

$$AMSE = \frac{1}{2Tn}\sum_{t=1}^{T}\sum_{k=1}^{2}\sum_{i=1}^{n}\left(\hat{\theta}_{ik}^t - \theta_{ik}^t\right)^2$$

We estimated the *GMCAR* model in both conditioning orders and we selected the best performing *GMCAR* conditioning order in terms of DIC. Tables 3 and 4 show model comparison in terms of DIC for the two simulation studies. In these tables, we report the percentiles of the DIC values for the true model (bold) over the simulated data sets. For the antagonist model, the percentiles of DIC differences with respect to the true model are reported. DIC statistics are reported separately for each vector $\boldsymbol{\theta}_{\bullet k}$ and for the 2n dimensional vector $\boldsymbol{\theta}$.

As regards study 1, DIC values for the *BCAR* model appear satisfactory when compared with those of the true model, since the distributions of the estimated DIC differences include 0 between the 2.5^{th} and the 97.5^{th} percentiles.

(*Table* 3 *about here*)

This is confirmed from model comparison by means of AMSE. Table 5.1 shows AMSE values for simulation study 1 obtained by the *GMCAR* model and percentage variation in AMSE obtained by the *BCAR* model. As can be noticed, performances of the *BCAR* model in study 1 are worst in terms of AMSE, mainly because of a considerable increase in AMSE for k=1, the one characterised by higher marginal variability.

(*Table* 4 *about here*)

As regards study 2, in both settings the *BCAR* model shows better performances in terms of DIC (Table 4) and AMSE (Table 5.2). Differences between the *BCAR* and *GMCAR* model reduce in setting 2 where vectors $\boldsymbol{\theta}_{\bullet 1}$ and $\boldsymbol{\theta}_{\bullet 2}$ are characterised by different conditional variances.

(*Tables* 5.1 *and* 5.2 *about here*)

We stress that in Study 1 and in Setting 2 of Study 2, marginal variances of spatial random effects $\boldsymbol{\varphi}_{\bullet 1}$ and $\boldsymbol{\varphi}_{\bullet 2}$ are unbalanced. Under this circumstances, the *GMCAR* model performances are better for the vector with lower marginal variance ($\boldsymbol{\theta}_{\bullet 2}$ in Study 1 and $\boldsymbol{\theta}_{\bullet 1}$ in Setting 2 of Study 2).

When vectors $\boldsymbol{\varphi}_{\bullet 1}$ and $\boldsymbol{\varphi}_{\bullet 2}$ are characterised by the same marginal variance and by the same spatial structure (Study 2 Setting 1), *GMCAR* model shows sensibly worse performances with respect to the *BCAR* model both in terms of DIC and AMSE. Moreover, the *BCAR* model performances are comparable for vectors $\boldsymbol{\theta}_{\bullet 1}$ and $\boldsymbol{\theta}_{\bullet 2}$, for which we obtain broadly the same values of DIC and AMSE, in agreement with the true parameters values that specify the same structure for the populations generating $\boldsymbol{\theta}_{\bullet 1}$ and $\boldsymbol{\theta}_{\bullet 2}$. This is not true when *GMCAR* model is estimated: different performances are obtained for vectors $\boldsymbol{\theta}_{\bullet 1}$ and $\boldsymbol{\theta}_{\bullet 2}$ due to the conditional nature of the model specification.

5. Application

In this section two real case studies are examined. Data refer to death counts observed from 1998 to 2001 in the 341 municipalities of the Emilia Romagna Region. We consider bivariate spatial modelling for diseases reported in the following table.

(*Table* 6 *about here*)

The case studies differ in the SMR relative variability and between-diseases correlation. As shown in Table 6, in both case studies, Standardised Mortality Ratios (SMR) show positive correlation. Moreover, while in Case Study 1 SMRs exhibit broadly the same relative variability (broadly similar Coefficients of Variation (CV)), in Case Study 2, SMR's exhibit different Coefficients of Variation and a lower correlation with respect to Case Study 1. For each disease, expected counts e_{ik}, i=1,..,341, k=1,2, are obtained via internal standardisation by applying the overall Emilia Romagna sex-age specific disease rates to the municipalities population. The considered diseases are rare relative to the population in each municipalities, then spatial smoothing of relative risk is needed. As outlined in section 3.1, we assume that:

$$Y_{ik} \mid \theta_{ik} \sim Poisson\left(e_{ik}\theta_{ik}\right) \qquad i=1,\ldots,341 \qquad k=1,2$$

$$\log\left(\theta_{ik}\right) = \gamma_k + \phi_{ik}$$

where θ_{ik} denotes the relative risk in area i for disease k, γ_k is a disease-specific intercept and ϕ_{ik} is the spatial random effect. We use tree different distributions for modelling the zero mean spatial random effects $\boldsymbol{\varphi} = \left(\boldsymbol{\varphi}_{\bullet 1}, \boldsymbol{\varphi}_{\bullet 2}\right)$:

- the *BCAR* distribution;
- the *GMCAR* distribution;
- the *GMCAR* with $\eta_0 = \eta_1 = 0$ which give rise to univariate CAR models.

We chose priors for *BCAR* and *GMCAR* distributions strictly following the specification described in section 4. Inference is based on 10.000 samples from the MCMC algorithm, after 15.000 burn-in iterations for all the estimated models. Convergence has been checked via the graphical examination of the trace plots of sample values versus iteration and of the autocorrelation plot in each chain. A major drawback of the use of the *BCAR* distribution is that computation times are about 6 time higher with respect to computation time when using the *GMCAR* distribution.

As shown in Figure 1, convergence of the elements of the smoothing parameters matrix $\boldsymbol{A}$ is satisfactory, as well as convergence of the parameter α_0 in both case studies.

(*Figure* 1 *about here*)

In Table 7, summary statistics of the parameters posterior distributions are reported. We observe that bridging parameters α_{12} and α_{21} are significantly greater than zero, indicating that the model captures a significant correlation between the spatial processes characterising the considered diseases: this supports the use of joint modelling for the considered diseases. The main difference between the two case studies lies in the fact that, in Case Study 2, one disease (Genitourinary System diseases) shows a weaker spatial structure: this is captured by the small value of parameter α_2 posterior mean. For this disease, a considerable spatial borrowing strength mechanism from the Respiratory System diseases distribution is reflected by the high value of parameter α_{21} posterior mean.

(*Table* 7 *about here*)

In what follows we focus on comparison of results obtained by univariate, *BCAR* and *GMCAR* models. Model performances are compared in terms of DIC. Moreover, in order to asses the plausibility of the estimated models, we make use of a posterior predictive Bayesian p-value [6] based on the following measure of fit proposed in [2] suitable for rare occurrences:

$$D_k = \sum_{i=1}^{341} \left(\sqrt{Y_{ik}} - \sqrt{e_{ik}\theta_{ik}} \right)^2 \qquad k=1,2 \tag{19}$$

Let Y_{ik}^{rep} denote counts in area i for disease k sampled from the posterior predictive distribution; Bayesian p-values are calculated as $p_k = P\left(D_k^{rep} \geq D_k \,|\, \boldsymbol{Y} \right)$ where D_k and $D_k^{rep} = \sum_{i=1}^{341} \left(\sqrt{Y_{ik}^{rep}} - \sqrt{e_{ik}\theta_{ik}} \right)^2$ are computed at each iteration of the MCMC algorithm and p_k is computed as the frequency of iterations where $D_k^{rep} > D_k$. Extreme values of p_k suggest inconsistencies between the model and actual data.

Results concerning DIC and p-values for both case studies are reported in Table 8, separately for each disease. In fact our aim is to underscore different model performances on the couples of diseases considered in the case studies.

Furthermore, in the last row of Table 8, we report the posterior correlations between the relative risks posterior means of diseases 1 and 2 to see the effect of joint modelling on the estimated spatial distribution. In fact, differences between these correlations obtained respectively via univariate and bivariate modelling can help to evaluate the similarity inducted by joint modelling on the relative risks spatial distribution. As regards *GMCAR* model, we estimated the model in both conditioning orders and we selected the best performing *GMCAR* conditioning order in terms of DIC: in Case Study 1 disease 2 is modelled conditionally on disease 1 while in Case Study 2 disease 1 is modelled conditionally on disease 2.

(*Table* 8 *about here*)

In both case studies, bivariate models show comparable performances in terms of DIC while univariate models show slightly higher DIC values. This is in agreement with the fact posterior credibility intervals for the bridging parameters shown in Table 7 do not include zero.

We observe that the total effective number of parameters (p_D) is in both case studies higher for the *BCAR* model with respect to the *GMCAR* model even if the two prior distributions have the same number of parameters. The reason behind this is still unclear to us and will be object of future researches.

For each model, DIC referring to the considered diseases are not comparable, in fact DIC is a measure of fit suitable to compare model performances on the same data. Different behaviours of the considered models on each disease can be checked by comparing Bayesian p-values. As regard *BCAR* model, Bayesian p-values show satisfactory values for both diseases in both case studies. Bayesian p-values obtained by the *GMCAR* model show quite extreme values for the conditioning disease. This reveal a more “symmetric” borrowing strength mechanism when the *BCAR* distribution is used.

In the reported applications, posterior correlations obtained with the *BCAR* model assumes values lying between the posterior correlation obtained with univariate model (lowest) and *GMCAR* model (highest). More specifically, in Case Study 2 correlation obtained with the *GMCAR* model is close to 1: this could be due to a strong attraction of the unconditionally modelled disease on the conditionally modelled disease. We note that, as reported in Table 6, correlation between SMRs is higher in Case Study 1 than in Case Study 2. This ordering is reproduced by *BCAR* and univariate

models while for the *GMCAR* model the correlation between relative risks posterior means is higher in Case Study 2 than in Case Study 1.

6. Concluding remarks

In this paper we propose a bivariate CAR model for areal data. The model is built by specifying full conditionals bivariate distributions. Model development strictly follows the theory of multivariate MRF as stated in [14]. In this context, difficulties arise because conditions for symmetry and positive definiteness of the joint distribution covariance matrix depend both on the conditional covariance matrix $\boldsymbol{\Lambda}$ and on the matrix of smoothing parameters $\boldsymbol{A}$, whose values depend on the $\boldsymbol{\Lambda}$ matrix as well. In [5] it is claimed that this make model fitting practically intractable. We approach these problems by specifying prior distributions for the elements of the $\boldsymbol{A}$ matrix conditionally on the matrix $\boldsymbol{\Lambda}$. Our approach is non-standard in that we obtain sufficient conditions for positive definiteness by working on the non-symmetric matrix $Block\left(-\boldsymbol{A}_{ij}\right)$ by using a theorem stated in [4]. On one hand this makes the problem tractable and on the other hand this allows obtaining a joint covariance matrix that is positive definite even if it is not necessary diagonally dominant. By means of a simulation experiment and a couple of data examples we show the effectiveness of the *BCAR* model when compared with the *GMCAR* model that has been shown to perform better than other proposed models for multivariate areal data [9].

The work can be extended in two main directions. First of all the conditions we found are shown to be sufficient but it is not clear how far they are from necessary conditions. This may be important in applied contexts because conditions could not allow to properly capture all the spatial information characterising the data generating process. Moreover the extension of the proposed model to the case $p > 2$ is non trivial and well be object of future research.

REFERENCES

1. Besag J. Spatial interaction and the statistical analysis of lattice systems (with discussion). *Journal of the Royal Statistical Society, Series B* 1974; **36**:192-236.
2. Brooks SP, Catchpole EA, Morgan BJT. Bayesian Animal Survival Estimation. *Statistical Science* 2000; **15**: 357-376.

3. Carlin BP, Banerjee S. Hierarchical multivariate CAR models for spatio-temporally correlated survival data (with discussion). In *Bayesian Statistics 7*, Bernardo JM, Bayarri MJ, Berger JO, Dawid AP, Heckerman D, Smith AFM, West M (eds) 2003; 45-63. Oxford: Oxford University Press.
4. Drazin MP, Haynswort EV. Criteria for the reality of matrix eigenvalues. *Matematische Zeitschrift* 1962; **78**:449-452.
5. Gelfand AE, Vounatsou P. Proper multivariate conditional autoregressive models for spatial data analysis. *Biostatistics* 2003; **4**:11-25.
6. Gelman A, Meng X, Stern H. Posterior predictive assessment of model fitness via realized discrepancies. *Statistica Sinica* 1996; **6**:733-807.
7. Held L, Natario I, Fenton SE, Rue H, Becker N. Towards joint disease mapping. *Statistical Methods in Medical Research* 2005; **14**:61-82.
8. Horn RA, Johnson CR. *Matrix Analysis* 1990; Cambridge, UK, Cambridge University Press.
9. Jin X, Carlin BP, Banerjee S. Generalized Hierarchical Multivariate CAR Models for Areal Data. *Biometrics* 2005; **61(4)**:950-961.
10. Kim H, Sun D, Tsutakawa RK. A bivariate Bayes methods for improving the estimates of mortality rates with a twofold conditional autoregressive model. *Journal of the American Statistical Association* 2001; **96**:1506-1521.
11. Langford IH, Leyland AH, Rasbash J, Goldstein H. Multilevel modelling of the geographical distributions of diseases. *Applied Statistics* 1999; **48**:253-268.
12. Lawson AB. *Statistical methods in spatial epidemiology.* 2001; Whiley. Chichester.
13. Leyland AH, Langford IH, Rasbash J, Goldstein H. Multivariate spatial models for event data. *Statistics in Medicine* 2000; **19**:2469-2478.
14. Mardia KV. Multi-dimensional multivariate Gaussian Markov random fields with application to image processing. *Journal of Multivariate Analysis* 1998; **24**:265-284.
15. Spiegelhalter DJ, Best N, Carlin BP, Van der Linde A. Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society, Series B* 2002; **64**:583-639.

Appendix

Proof: Following [14], a sufficient condition for matrix $\boldsymbol{\Sigma}$ to be positive definite is that the non-symmetric matrix $Block\left(-\boldsymbol{A}_{ij}\right)$ is positive definite, where:

$$Block\left(-\boldsymbol{A}_{ij}\right)=\begin{bmatrix} \boldsymbol{I}_{2\times 2} & -\boldsymbol{A}_{12} & \cdots & -\boldsymbol{A}_{1j} & \cdots & -\boldsymbol{A}_{1n} \\ -\boldsymbol{A}_{21} & \boldsymbol{I}_{2\times 2} & \cdots & -\boldsymbol{A}_{2j} & \cdots & -\boldsymbol{A}_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ -\boldsymbol{A}_{n1} & -\boldsymbol{A}_{n2} & \cdots & -\boldsymbol{A}_{nj} & \cdots & \boldsymbol{I}_{2\times 2} \end{bmatrix}$$

We derive conditions for row diagonal dominance of matrix $Block\left(-\boldsymbol{A}_{ij}\right)$. Note that, for a generic odd row, diagonal dominance requires that

$$1>\sum_{\substack{j\neq i \\ j=1}}^{n}\frac{\alpha_1 w_{ij}}{2m_i}+\sum_{\substack{j\neq i \\ j=1}}^{n}\frac{|\alpha_{12}|\,w_{ij}}{2m_i}=\frac{\alpha_1}{2}+\frac{|\alpha_{12}|}{2} \tag{A.1}$$

Diagonal dominance for a generic even row requires that

$$1>\sum_{\substack{j\neq i \\ j=1}}^{n}\frac{\alpha_2 w_{ij}}{2m_i}+\sum_{\substack{j\neq i \\ j=1}}^{n}\frac{|\alpha_{21}|\,w_{ij}}{2m_i}=\frac{\alpha_2}{2}+\frac{|\alpha_{21}|}{2} \tag{A.2}$$

In both equations (A.1) and (A.2), absolute values for parameters α_1 and α_2 are dropped since those parameters are restricted to be positive.

Firstly conditions for row diagonal dominance are derived for parameters α_1 and α_2, by posing $\alpha_0=0$ (in order to obtain conditions that are independent by α_0).

$$2>\alpha_l+|\alpha_{lk}|=\alpha_l+\frac{|\alpha_l\sigma_{12}|}{\sigma_k} \qquad\qquad l,k=1,2 \quad l\neq k\,;$$

which, solved with respect to α_l gives the following conditions on α_1 and α_2 which do not depend by α_0

$$0 < \alpha_l < \frac{2\sigma_k}{\sigma_k + |\sigma_{12}|}$$

Condition on α_0 is derived as a function of α_1 and α_2 in order to preserve row diagonal dominance, i.e. by imposing that $2 > \alpha_l + |\alpha_{lk}|$ l,k=1,2 $l \neq k$ .We obtain that parameter α_0 must satisfy at the same time:

$$\alpha_0 \in \left[\alpha_1 (\sigma_{12} + \sigma_2) - 2\sigma_2 ; \alpha_1 (\sigma_{12} - \sigma_2) + 2\sigma_2 \right]$$
$$\alpha_0 \in \left[\alpha_2 (\sigma_{12} + \sigma_1) - 2\sigma_1 ; \alpha_2 (\sigma_{12} - \sigma_1) + 2\sigma_1 \right]$$

It turns out that α_0 must satisfy condition $I_{\alpha_0} < \alpha_0 < S_{\alpha_0}$.

Gershgorin's disc theorem [8] allows determining a region where eigenvalues of a complex square $n \times n$ matrix lie. More precisely, for non symmetric matrices, the eigenvalues lie in the intersection of two regions. The first region is determined by the union of the *n* discs centered at each (positive) diagonal elements and with radius given by the sum of the absolute values of the off-diagonal elements of the row. The second region is determined by the union of the *n* discs centered at each (positive) diagonal elements and with radius given by the sum of the absolute values of the off-diagonal elements of the column. Thus, because of row diagonal dominance of matrix $Block\left(-\boldsymbol{A}_{ij}\right)$, Gershgorin's disc theorem implies that eigenvalues of $Block\left(-\boldsymbol{A}_{ij}\right)$ lie in the half right plane, that is they have positive real part even if they could be complex since $Block\left(-\boldsymbol{A}_{ij}\right)$ in non-symmetric. There is to show that eigenvalues are real. Note that

$$\boldsymbol{\Sigma}^{-1} = \left[Block\ diag\left(\boldsymbol{\Gamma}_1, \boldsymbol{\Gamma}_2, ..., \boldsymbol{\Gamma}_n\right) \right] \left[Block\left(-\boldsymbol{A}_{ij}\right) \right] = \boldsymbol{G} \left[Block\left(-\boldsymbol{A}_{ij}\right) \right]$$

This means that an hermitian positive definite matrix $\boldsymbol{G}$ exists such that product $\boldsymbol{G}\left[Block\left(-\boldsymbol{A}_{ij}\right) \right]$ is hermitian. Moreover, $rank\left(\boldsymbol{G}\right) = 2n$. As stated by Theorem 1 in [4], this is a necessary and sufficient condition for matrix $Block\left(-\boldsymbol{A}_{ij}\right)$ to have 2*n* real eigenvalues. Thus, all its eigenvalues are real, so $Block\left(-\boldsymbol{A}_{ij}\right)$ is positive definite. This implies positive definiteness of the covariance matrix $\boldsymbol{\Sigma}$.

Table 1: *Study 1, $\boldsymbol{\varphi}$ simulated from the GMCAR model*

	γ_1	γ_2	τ_1	τ_2	η_0	η_1	υ^2	δ_1	δ_2
Setting 1	-2	-5	10	10	0.9	0.5	0.01	0.2	0.9

Table 2: *Study 2, $\boldsymbol{\varphi}$ simulated from the BCAR model*

	γ_1	γ_2	τ_1	τ_2	τ_{12}	υ^2	α_1	α_2	α_0
Setting 1	-2	-5	10	10	-6	0.01	1.2	1.2	0.2
Setting 2	-2	-5	30	10	-10	0.01	1.2	1.2	0.09

Table 3: *Percentiles of DIC for the GMCAR* (*true*) *model, reported in bold. Percentiles of DIC differences between GMCAR model and the BCAR model* (*Study 1*).

		k=1			*k*=2			Total		
Data	Model	2.5%	50%	97.5%	2.5%	50%	97.5%	2.5%	50%	97.5%
S1	***GMCAR***	**-62.18**	**-45.83**	**-29.05**	**-94.63**	**-77.39**	**-61.34**	**-159.44**	**-122.55**	**-93.67**
S1	*BCAR*	-8.98	7.60	26.30	-6.26	8.96	24.45	-15.16	17.29	50.46

Table 4: *Percentiles of DIC for the BCAR* (*true*) *model, reported in bold. Percentiles of DIC differences between BCAR model and the GMCAR model* (*Study 2*)

		k=1			*k*=2			Total		
Data	Model	2.5%	50%	97.5%	2.5%	50%	97.5%	2.5%	50%	97.5%
S1	***BCAR***	**-88.37**	**-73.66**	**-53.67**	**-89.52**	**-74.32**	**-56.78**	**-149.43**	**-119.26**	**-89.76**
S1	*GMCAR*	9.22	21.12	39.04	9.82	22.01	41.58	19.18	42.89	80.47
S2	***BCAR***	**-122.30**	**-111.30**	**-94.15**	**-96.02**	**-86.25**	**-72.65**	**-220.19**	**-198.22**	**-170.87**
S2	*GMCAR*	2.03	8.03	20.97	7.58	15.42	29.75	10.59	23.73	50.76

Table 5.1: *Percentage changes in Average mean squared error of: BCAR model relative to GMCAR* (*true*) *model*

Data	Model	*k*=1	*k*=2	Total
S1	***GMCAR***	**0.00946**	**0.00674**	**0.00810**
S1	*BCAR*	17.31%	1.47%	10.72%

Table 5.2: *Percentage changes in Average mean squared error of: GMCAR model relative to BMCAR* (*true*) *model*

Data	Model	*k*=1	*k*=2	Total
S1	***BCAR***	**0.00746**	**0.00760**	**0.00753**
S1	*GMCAR*	19.10%	36.84%	28.05%
S2	***BCAR***	**0.00490**	**0.00762**	**0.00626**
S2	*GMCAR*	4.17%	12.95%	9.52%

Table 6: *Analysed case studies: summary statistics*

	Case Study 1		Case Study 2	
Disease	Cirrhosis (1)	Liver cancer (2)	Respiratory System (1)	Genitourinary System (2)
SMR's Coefficient of variation	0.63	0.58	0.30	0.58
Correlation between SMRs	0.39		0.20	

Table 7: *Posterior summaries of the BCAR model parameters*

	Case Study 1		Case Study 2	
	Posterior Mean	95% Posterior C.I.	Posterior Mean	95% Posterior C.I.
α_1	1.054	0.244 – 1.546	1.374	0.753 – 1.748
α_{12}	0.801	0.339 – 1.559	0.458	0.053 – 1.091
α_2	1.231	0.518 – 1.618	0.639	0.024 – 1.431
α_{21}	0.656	0.279 – 1.334	0.965	0.249 – 1.723
α_0	0.372	0.197 – 0.579	0.137	0.049 – 0.225
γ_1	-0.050	-0.149 – 0.055	-0.029	-0.088 – 0.021
γ_2	0.016	-0.089 – 0.122	-0.007	-0.071 – 0.051
σ_1	0.340	0.234 – 0.463	0.113	0.076 – 0.159
σ_{12}	0.129	0.054 – 0.227	0.048	0.013 – 0.089
σ_2	0.309	0.191 – 0.452	0.172	0.099 – 0.281

Table 8: *Some measures of model performances*

		Case Study 1			Case Study 2		
		BCAR	*GMCAR*	Univariate	*BCAR*	*GMCAR*	Univariate
Disease 1	Deviance	1597	1612	1607	2063	2069	2071
	p_D	120.0	103.1	114.5	134.6	127.9	130.6
	DIC	1717.0	1715.1	1721.5	2197.6	2196.9	2201.6
	p-value	0.19	0.09	0.15	0.38	0.41	0.44
Disease 2	Deviance	1543	1537	1546	1474.82	1489	1494
	p_D	99.7	107.3	105.3	58.3	41.8	53.0
	DIC	1642.7	1644.3	1651.3	1533.1	1530.8	1547.0
	p-value	0.59	0.66	0.58	0.80	0.94	0.95
	COR^{++}	0.91	0.92	0.78	0.86	0.95	0.46

$^{++}$Correlation between relative risks posterior means

Figure 1: *Chains trace plot of the **A** matrix and α_0 parameters for Case Studiy* 1 (*left column*) *and Case Study 2* (*right column*)