

Cure rate proportional odds models with spatial frailties for interval-censored data

Bao Yiqi^a, Vicente Garibay Cancho^a, Francisco Louzada^a, Adriano Kamimura Suzuki^{1, a}

^aDepartamento de Matemática Aplicada e Estatística, Universidade de São Paulo, Brasil

Abstract

This paper presents proportional odds cure models to allow spatial correlations by including spatial frailty in the interval censored data setting. Parametric cure rate models with independent and dependent spatial frailties are proposed and compared. Our approach enables different underlying activation mechanisms that lead to the event of interest; in addition, the number of competing causes which may be responsible for the occurrence of the event of interest follows a Geometric distribution. Markov chain Monte Carlo method is used in a Bayesian framework for inferential purposes. For model comparison some Bayesian criteria were used. An influence diagnostic analysis was conducted to detect possible influential or extreme observations that may cause distortions on the results of the analysis. Finally, the proposed models are applied for the analysis of a real data set on smoking cessation. The results of the application show that the parametric cure model with frailties under the first activation scheme has better findings.

Keywords: Bayesian inference, influence diagnostics, survival models with a cure fraction, spatial frailty

1. Introduction

Advances in medical and health science have led to several clinical studies that showed non-negligible proportions of patients who respond favorably to a treatment and are non-susceptible to the event of interest. Such a proportion of non-susceptible patients is considered cured or with prolonged disease-free survival which is usually referred to as the cured fraction. Survival models with cure fraction, also known as long-term survival models or cure rate models, have been widely developed to accommodate a cured fraction.

One of the most popular cure rate model is the mixture model introduced by Berkson and Gage (1952). It assumes that a certain proportion of the patients are cured, in the sense that they do not present the event of interest during a long period of time and can be seen as cured or immune from the causes of death under study (Rodrigues *et al.*, 2009). Mixture models can also accommodate the effects of covariances on the cure rate of populations and the survival function of susceptible to (non-cured) individuals via two components of separate regression. An alternative was later proposed by Yakovlev and Tsodikov (1996) and Chen *et al.* (1999) concerning cancer therapy studies. In this context, it is assumed that, for each individual, there is a latent quantity M of cells able to cause cancer, being also assumed that M has a Poisson distribution. The model is known as the promotion time cure model or bounded cumulative hazard model, where an inference can be made concerning

¹ Corresponding author: Departamento de Matemática Aplicada e Estatística, Universidade de São Paulo, Avenida Trabalhador São-carlense, 400 - Centro CEP: 13566-59, São Carlos-SP, Brasil. E-mail: suzuki@icmc.usp.br

the distribution of cells that produce cancer. This subject has attracted the attention of several researchers, because it adheres to the proportional hazard structure and incorporates parameters with clear biological interpretations.

Gu *et al.* (2011) have recently developed another class of cure rate models where the survival function has a proportional odds structure that is an already popular regression model in ordinary survival model (Bennett, 1983; Collett, 1994). The proportional odds models with cure rate retain all the advantages of regular proportional odds models regarding the survival analysis. They can also be derived from the latent factors models conducted by Cooner *et al.* (2007), considering these models share some properties.

In many clinical trial studies, it is very common for patients to be periodically examined for disease occurrence or progression. In such a situation, the exact failure time which patients cannot be observed exists, but can only be allowed to lie in an interval after a sequence of examination times that is known as interval-censored (Peto, 1973). The estimation methods available to the right-censored data, such as the Kaplan-Meier estimator, is not considered adequate to be applied in interval-censored data, because it can lead to biased estimation and invalid inferences. The information concerning interval censorship should be taken into account in modeling (Lindsey and Ryan, 1998; R  cker and Messerer, 1988; Sun and Chen, 2010).

The lifetimes data are sometimes collected from several regions, which can lead to different effects for each observation. To consider these different effects in the cure model for interval-censored data, Xiang *et al.* (2011) adopted the mixture model by including two random effects (also known as frailties) to each cluster to explain the effects on the survival time for susceptible individuals and the effects on the cure fraction. Hu and Xiang (2013) adopted the non-mixture cure model and introduced random effects in the promotion of the cure time model. The development of the geographic information system has allowed several researchers to incorporate geographical information on the subjects under study. They have also developed survival models that account for spatial clustering and variation. Banerjee and Carlin (2004) have specified spatial correlation to fit geographically clustered interval-censored survival data using a parametric cure rate model. They have developed a Bayesian approach to the mixture cure model with spatial random effects on the survival function for subjects at risk and spatial frailties by using a multivariate conditionally autoregressive (MCAR) prior. Pan *et al.* (2014) propose a Bayesian approach under a proportional hazards frailty model to analyze interval-censored survival data with spatial correlation.

This paper analyzes smoking cessation data. In a smoking cessation study, all patients (smokers) were randomized both into a smoking intervention (SI) group, or into a usual care (UC) group that receives no special anti-SI. The SI program treatments were conducted in Rochester city, being located in the center of the map. The details concerning the programs can be found in Murray *et al.* (1998). Here, each patient was observed once a year over the period of five years of monitoring. Our event of interest concerns if they will relapse into smoking (resume smoking) or not. If a smoker starts smoking again after an initial attempt to quit, then only an approximate one-year time interval was observed from the previous observation to the current observation. Thus, the relapse times are interval-censored. In this analysis, we limit our attention to patients who are known to have quit smoking at least once during the period of study and who have an identifiable Minnesota zip code of residence. Thus, the data consists of a total of 223 patients who reside in 51 zip codes in the Southeastern corner of Minnesota; among them, there are 65 patients that have underwent a relapse, which implies the empirical cure rate is approximately 71%. Ma and Xiang (2013) also confirmed the existence of a non-eligible cure fraction in the population.

In this paper, we present a proportional odds cure model (Gu *et al.*, 2011) to allow spatial cor-

relations by including spatial frailty in the interval-censored data setting, to propose parametric cure rate models that include independent and dependent spatial frailties. These new cure rate models enable different underlying activation mechanisms which lead to the event of interest. Also, we develop inference procedures through a Bayesian perspective. The proposed models can be also compared to models introduced by Banerjee and Carlin (2004) and Pan *et al.* (2014) through the deviance information criterion (DIC) proposed by Spiegelhalter *et al.* (2002).

We also conduct influence diagnostics to examine the assumptions of the model and conduct studies on sensitivity to detect possible influential or extreme observations that can cause distortions in the results of the analysis. Here the influence of the diagnostics of case deletion are developed for the posterior joint distribution based on the ψ -divergence (Peng and Dey, 1995; Weiss, 1996).

The remainder of the paper is organized as follows. In Section 2, the introduction of frailty models and some distributions for spatial frailties are presented and followed by the spatial distributional aspects of our modeling. The Bayesian inference for the proposed models is developed in the Section 3. In Section 4, the proposed models are fit into a read data set (smoking cessation study). Finally, Section 5 concludes with some general remarks.

2. Cure rate proportional odds model

Gu *et al.* (2011) have recently proposed a cure rate proportional odds model (CRPO model). Unlike the model proposed by Chen *et al.* (1999), the ratio of hazards for the CRPO model for two covariate values does not remain over time. The model can be characterized by the latent factors model by Cooner *et al.* (2007), which contains a Geometric distribution for the number of latent factors. Supposing that there are I regions and n_i individuals in i^{th} region. Let T_{ij} denotes the random variable for time to the event of the j^{th} individual in the i^{th} region, where $j = 1, \dots, n_i$ and $i = 1, \dots, I$. We suppose that the $(i, j)^{th}$ individual is potentially exposed to M_{ij} latent risks, where M_{ij} denote the initial number of competing causes concerning the occurrence of an event, and assuming M_{ij} has a Geometric distribution with parameter $1/(1 + \theta_{ij})$, the probability mass function is given by

$$P(M_{ij} = m) = \frac{\theta_{ij}^m}{(\theta_{ij} + 1)^{m+1}}, \quad m = 0, 1, 2, \dots, \quad (2.1)$$

where $\theta_{ij} > 0$, $E(M_{ij}) = \theta_{ij}$, and $\text{Var}(M_{ij}) = \theta_{ij}(1 + \theta_{ij})$.

Let Y_{cij} denotes the lifetime (for relapse into smoking) of j^{th} individual in i^{th} region due to the c^{th} ($c = 1, \dots, M_{ij}$) latent cause or risk. Common latent causes of relapse into smoking are giving in to cravings, stress, weight gain and nicotine withdrawal symptoms, which may include dizziness, nausea, constipation, increased appetite, anxiety, depression, and sensation of tightness in the chest (National Cancer Institute at <http://smokefree.gov/withdrawal>). The problem is that the causes or risks are latent in the sense that it is hard to know the exact cause of relapse into smoking.

Given $M_{ij} = m$, $Y_{1ij}, Y_{2ij}, \dots, Y_{mij}$ are assumed to be independent and identically distributed with a common distribution function $F(\cdot) = 1 - S(\cdot)$ that does not depend upon M_{ij} . If we assume that the presence of any latent risk will ultimately lead to the occurrence of the event, the time to the event of interest T_{ij} could be defined as $T_{ij} = \min\{Y_{1ij}, \dots, Y_{mij}\}$ for $M_{ij} \geq 1$. In this case we are assuming the cause, which comes first, not even knowing what it is, cause relapse into smoking. If $M_{ij} = 0$, then the individual is not at risk of occurrence of the event (relapse into smoking) and is considered cured. In this case, we define $T_{ij} = \infty$ with $P(T_{ij} = \infty | M_{ij} = 0) = 1$. Thus, the survival function for

the population is given by

$$S_{\text{pop}}(t_{ij}) = [1 + \theta_{ij}F(t_{ij})]^{-1}. \quad (2.2)$$

Note that this survival function has a proportional odds structure when covariates \mathbf{x}_{ij} are modeled via $\theta_{ij}(\mathbf{x}_{ij})$ and the latent survival $F(t_{ij})$ is free of \mathbf{x}_{ij} , because

$$\frac{1 - S_{\text{pop}}(t_{ij}|\mathbf{x}_{ij})}{S_{\text{pop}}(t_{ij}|\mathbf{x}_{ij})} = \theta_{ij}(\mathbf{x}_{ij})F(t_{ij}).$$

This CRPO model can also be obtained as an extension of the transformation-cure model (Zeng *et al.*, 2006).

The probability density function (pdf) and the hazard function associated to (2.2) are given by

$$f_{\text{pop}}(t_{ij}) = \theta_{ij}f(t_{ij})[1 + \theta_{ij}F(t_{ij})]^{-2} \quad \text{and} \quad h_{\text{pop}}(t_{ij}) = \theta_{ij}f(t_{ij})[1 + \theta_{ij}F(t_{ij})]^{-1},$$

respectively, where $f(t_{ij}) = (\partial/\partial t_{ij})F(t_{ij})$.

Note that, the survival function in (2.2) can also be written as a mixture cure model

$$S_{\text{pop}}(t_{ij}) = (1 + \theta_{ij})^{-1} + \left(1 - (1 + \theta_{ij})^{-1}\right) \left[\frac{\{1 + \theta_{ij}F(t_{ij})\}^{-1} - (1 + \theta_{ij})^{-1}}{1 - (1 + \theta_{ij})^{-1}} \right].$$

Thus, the survival functions of uncured (susceptible) individuals can be expressed by

$$S_{\text{sus}}(t_{ij}) = \frac{[1 + \theta_{ij}F(t_{ij})]^{-1} - (1 + \theta_{ij})^{-1}}{1 - (1 + \theta_{ij})^{-1}}.$$

Now, if we assume another situation where the presence of all latent risks will ultimately lead to the occurrence of the event. In this case we are assuming the cause, which is later, not even knowing what it is, cause relapse into smoking. Thus, the time to the event of interest is defined by the random variable $T_{ij} = \max\{Y_{cij}, c = 1, \dots, M_{ij}\}$ for $M_{ij} \geq 1$ and $T_{ij} = \infty$ if $M_{ij} = 0$ with $P(T_{ij} = \infty | M_{ij} = 0) = 1$. The survival function for the population is given by

$$S_{\text{pop}}(t_{ij}) = 1 + (1 + \theta_{ij})^{-1} - [1 + \theta_{ij}S(t_{ij})]^{-1}. \quad (2.3)$$

The corresponding pdf and the hazard function are given by

$$f_{\text{pop}}(t_{ij}) = \theta_{ij}f(t_{ij})(1 + \theta_{ij}S(t_{ij}))^{-2},$$

and

$$h_{\text{pop}}(t_{ij}) = \frac{\theta f(t_{ij})[1 + \theta_{ij}S(t_{ij})]^{-2}}{1 + (1 + \theta_{ij})^{-1} - (1 + \theta_{ij}S(t_{ij}))^{-1}},$$

respectively. The survival function (2.3) can also be written as a mixture cure model

$$S_{\text{pop}}(t_{ij}) = (1 + \theta_{ij})^{-1} + \left(1 - (1 + \theta_{ij})^{-1}\right) \frac{1 - (1 + \theta_{ij} S(t_{ij}))^{-1}}{1 - (1 + \theta_{ij})^{-1}}.$$

Thus, the survival functions of susceptible individuals is given by

$$S_{\text{sus}}(t_{ij}) = \frac{1 - (1 + \theta_{ij} S(t_{ij}))^{-1}}{1 - (1 + \theta_{ij})^{-1}}.$$

The first situation is also known as first activation (FA) scheme because, in this case, we assume the event of interest occurs when the first possible cause is activated. However, the second situation is known as the last activation (LA) scheme because the event of interest only takes place after all the latent causes have been activated (Cooner *et al.*, 2007). Thus, we denoted the survival functions (2.2) and (2.3) by $S_{\text{pop}}^F(t_{ij})$ and $S_{\text{pop}}^L(t_{ij})$, respectively. There is another kind of situation where the event of interest occurs: when some of the possible causes are activated and, given the number of latent causes M_{ij} , the number of activated causes is a random variable with discrete Uniform distribution $\{1, \dots, M_{ij}\}$. This situation is known as random activation scheme. In this case, the survival function for the population is given by

$$S_{\text{pop}}^R(t_{ij}) = (1 + \theta_{ij})^{-1} + \left(1 - (1 + \theta_{ij})^{-1}\right) S(t_{ij}), \quad (2.4)$$

where the superscript R denotes random activation scheme.

Note that whichever the activation scheme, the density and hazard functions of the cure models are improper functions, since the survival functions are not proper. Its cure fraction is the same for all the activation schemes and can be obtained by $p_{0ij} = \lim_{t_{ij} \rightarrow \infty} S_{\text{pop}}(t_{ij}) = (1 + \theta_{ij})^{-1}$. However, under different activation schemes, the models differ by its survival, density, and hazard functions. Moreover, under the conditions of the models (2.2), (2.3), and (2.4) for any distribution function $F(\cdot)$, we have $S_{\text{pop}}^F(t_{ij}) \leq S_{\text{pop}}^R(t_{ij}) \leq S_{\text{pop}}^L(t_{ij})$ for all $t_{ij} > 0$.

The cure fraction plays a key role in the survival models with cure fraction. So we consider the parametrization of the model in terms of the cure fraction. Since $p_{0ij} = (1 + \theta_{ij})^{-1}$, we have $\theta_{ij} = p_{0ij}^{-1} - 1$. Moreover, we propose that the cured fraction of an individual $(i, j)^{th}$ be associated with covariates \mathbf{x}_{ij} . Thus linking p_{0ij} to covariates \mathbf{x}_{ij} by

$$p_{0ij} = \frac{\exp(\xi_{ij})}{1 + \exp(\xi_{ij})}, \quad j = 1, \dots, n_i, \quad i = 1, \dots, I,$$

where ξ_{ij} is a linear function of covariates, $\xi_{ij} = \mathbf{x}_{ij}' \mathbf{b}$ with \mathbf{b} is a p_1 -dimensional vector which represents the effects of covariates on the cured fraction. Thus, the models (2.2), (2.3), and (2.4) parametrized in terms of p_{0ij} can be written as

$$S_{\text{pop}}^F(t_{ij}) = \left[1 + (p_{0ij}^{-1} - 1) F(t_{ij})\right]^{-1}, \quad (2.5)$$

$$S_{\text{pop}}^L(t_{ij}) = 1 + p_{0ij} - \left[1 + (p_{0ij}^{-1} - 1) S(t_{ij})\right]^{-1} \quad (2.6)$$

and

$$S_{\text{pop}}^R(t_{ij}) = p_{0ij} + (1 - p_{0ij})S(t_{ij}). \quad (2.7)$$

The model in (2.7) is the same considered by Banerjee and Carlin (2004).

Henceforth, we assume the Weibull distribution for the latent variables Y_{cij} , $c = 1, \dots, i = 1, \dots, I$, and $j = 1, \dots, n_i$ in (2.2), (2.3), and (2.4) with $F(y_{cij}|\boldsymbol{\phi}) = \exp(-y_{cij}^\alpha e^{\lambda_{ij}})$ where $\boldsymbol{\phi} = (\alpha, \lambda_{ij})$, $\alpha > 0$ is a shape parameter and $\lambda_{ij} \in \mathbb{R}$ is a scale parameter with $\lambda_{ij} = \mathbf{z}_{ij}'\boldsymbol{\beta}$ is the linear predictor of the covariates, where \mathbf{z}_{ij} represents the covariates of an individual (i, j) and $\boldsymbol{\beta}$ is the p_2 -dimensional vector that represents the effects of covariates on the survival model component. This parametrization of the Weibull distribution is implicitly assumed as a proportional hazards model with the baseline hazard function $h_0(t|\alpha) = \alpha t^{\alpha-1}$. Here, we called the model in (2.2) the Weibull cure rate model (WCRM) under the FA denoted by WCRM-FA and (2.3) by WCRM under the LA denoted by WCRM-LA.

By following Banerjee and Carlin (2004), we introduce the frailties U_i and V_i to evaluate the effect of regions on the lifetime of individuals and on the cured fraction through linear predictor

$$\begin{aligned} \lambda_{ij} &= \mathbf{z}_{ij}'\boldsymbol{\beta} + U_i, \\ \xi_{ij} &= \mathbf{x}_{ij}'\boldsymbol{b} + V_i, \quad \text{for } j = 1, \dots, n_i, i = 1, \dots, I. \end{aligned}$$

Here, the frailties U_i and V_i are spatially correlated across the regions. In this work, we propose two approaches. Firstly, we employ separate independent conditionally auto-regressive (CAR) prior distribution on (\mathbf{U}, \mathbf{V}) .

The CAR model was originally developed by Besag (1974). These priors have become very popular in the Bayesian analysis of real data, especially in disease mapping. Let $\boldsymbol{\psi}_1 = \mathbf{U}$ and $\boldsymbol{\psi}_2 = \mathbf{V}$ denote the random frailties vectors, \mathbf{W} denote the adjacent matrix of the map, so that $W_{ii'} = 1$ if the regions i and i' are adjacent, otherwise it would be 0, while $w_{i+} = \sum_j w_{ij}$ denotes the number of regions adjacent to region i . The zero-centered CAR prior for $\boldsymbol{\psi}_l$, $l = 1, 2$ is an I -dimensional Gaussian distribution with mean $\mathbf{0}$ and precision matrix $\rho_l^{-1}(\mathbf{D}_W - \mathbf{W})$, where $\rho_l > 0$ and \mathbf{D}_W is diagonal with $(\mathbf{D}_W)_{ii} = w_{i+}$ for $i = 1, \dots, I$. For l fixed, the CAR model is denoted by $\text{CAR}(\rho_l)$. Note that the precision matrix $\rho_l^{-1}(\mathbf{D}_W - \mathbf{W})$ is rank deficient and is therefore a non-positive definite matrix that leads the distribution to improper. This singularity, although theoretically awkward, it creates a small problem during the Bayesian implementation, since the identifying sum-to-zero constraint $\sum_{i=1}^I \psi_{il} = 0$ is easily imposed in a Gibbs sampler simply by re-centering the ϕ_i draws around zero after every iteration.

Second, we assume that the spatial priors on (\mathbf{U}, \mathbf{V}) are dependent, presenting MCAR prior distribution. Let $\boldsymbol{\psi} = (\mathbf{U}', \mathbf{V}')'$, we employ a multivariate MCAR distribution with a common smoothness parameter a , that is, $\boldsymbol{\psi}$ has a normal distribution with mean $\mathbf{0}$ and precision matrix $\boldsymbol{\Lambda} \otimes (\mathbf{D}_W - a\mathbf{W})$, where \otimes denotes the Kronecker product, $\boldsymbol{\Lambda}$ is a 2×2 symmetric and positive definite matrix, $a \in (0, 1)$ and \mathbf{W} is standardized so that each of its rows sum to 1. This prior is a proper distribution, the parameter a has a spatial smoothness interpretation. The value of a closer to 1 implies greater weight on the adjacency matrix \mathbf{W} , while a close to 0 implies the adjacency structure has few roles to play in the precision matrix. This prior is denoted by $\text{MCAR}(a, \boldsymbol{\Lambda})$.

We also assume the parameter $\boldsymbol{\psi}$ with an extended MCAR distribution proposed by Gelfand and Vounatsou (2003) and Carlin and Banerjee (2003), which assumes different smoothness parameters for the parameters \mathbf{U} and \mathbf{V} , namely a_1 and a_2 . Let $(\mathbf{D}_W - a_i\mathbf{W})$ denotes the corresponding positive definite matrix and $\mathbf{R}_i'\mathbf{R}_i$ denotes its Cholesky factorization, where \mathbf{R}_i is an upper triangular matrix

with real and positive diagonal entries with $I \times I$ dimension for $i = 1, 2$. The precision matrix can be written as

$$\Lambda \otimes (D_W - \mathbf{a}W) = \begin{bmatrix} \lambda_{11} \mathbf{R}'_1 \mathbf{R}_1 & \lambda_{12} \mathbf{R}'_1 \mathbf{R}_2 \\ \lambda_{21} \mathbf{R}'_2 \mathbf{R}_1 & \lambda_{22} \mathbf{R}'_2 \mathbf{R}_2 \end{bmatrix} = \mathbf{R}' (\Lambda \otimes I_I) \mathbf{R},$$

where λ_{ij} are the elements of Λ . The $\mathbf{R} = \begin{bmatrix} \mathbf{R}_1 & 0 \\ 0 & \mathbf{R}_2 \end{bmatrix}$ has dimension $2I \times 2I$, and the $\Lambda \otimes (D_W - \mathbf{a}W)$ is positive definite since Λ is positive definite. We denoted this prior by $\text{MCAR}(a_1, a_2, \Lambda)$.

3. Bayesian inference

Let $\mathcal{D} = \{(A_{ij}, \mathbf{x}_{ij}, \mathbf{z}_{ij}, \delta_{ij}); j = 1, \dots, n_i, i = 1, \dots, M\}$ denote the observed data, where $A_{ij} = (t_{ijL}, t_{ijR}]$ is the interval during which individual j in cluster i occur the event of interest, \mathbf{x}_{ij} and \mathbf{z}_{ij} are the p_1 -dimensional and p_2 -dimensional vectors of covariates, and δ_{ij} is following interval censoring indicator: $\delta_{ij} = I(t_{ijR} < \infty)$. For a special case where the survival time is right (left) censored, $R_{ij} = +\infty$ ($L_{ij} = 0$), whereas for exact observations, it is $t_{ijL} = t_{ijR}$. By following Finkelstein (1986), the likelihood function regarding the conditional distribution of \mathcal{D} given (\mathbf{U}, \mathbf{V}) for the general interval-censored cure rate model is given by

$$\begin{aligned} L(\boldsymbol{\varphi} | \mathcal{D}, \mathbf{U}, \mathbf{V}) &\propto \prod_{i=1}^I \prod_{j=1}^{n_i} (S_{\text{pop}}(t_{ijL} | \boldsymbol{\varphi}) - S_{\text{pop}}(t_{ijR} | \boldsymbol{\varphi}))^{\delta_{ij}} S_{\text{pop}}(t_{ijL} | \boldsymbol{\varphi})^{1-\delta_{ij}} \\ &\propto \prod_{i=1}^I \prod_{j=1}^{n_i} S_{\text{pop}}(t_{ijL} | \boldsymbol{\varphi}) \left(1 - \frac{S_{\text{pop}}(t_{ijR} | \boldsymbol{\varphi})}{S_{\text{pop}}(t_{ijL} | \boldsymbol{\varphi})} \right)^{\delta_{ij}}, \end{aligned} \quad (3.1)$$

where $\boldsymbol{\varphi} = (\mathbf{b}, \boldsymbol{\beta}, \mathbf{U}, \mathbf{V}, \alpha)$ and α is the shape parameter of the Weibull distribution. For a Bayesian analysis, we assume the follow prior densities for parameters the \mathbf{b}' , $\boldsymbol{\beta}'$, and α

- $b_j \sim N(\mu_b, \sigma_b^2)$, $j = 0, \dots, (p_1 - 1)$, with μ_b and σ_b known;
- $\beta_j \sim N(\mu_\beta, \sigma_\beta^2)$, $j = 1, \dots, p_2$, with μ_β and σ_β known;
- $\alpha \sim N(\mu_\alpha, \sigma_\alpha^2) \mathbf{I}_{(0, \infty)}$, with μ_α and σ_α known;

where $N(\mu, \sigma^2) \mathbf{I}_{(a, b)}$ denotes the truncated normal distribution that is the probability distribution of a normally distributed random variable whose value lies within the interval $-\infty \leq a < b \leq \infty$. In several areas, especially in medicine, it is preferable to use the prior information when it is available. In addition, a truncated normal distribution as prior facilitates the insertion of information in certain regions of the parameter space, since the hyperparameters no longer represent the mean and variance, but still controls the region of higher probability mass.

3.1. Independent assumption

For the independent assumption, we employ the separate independent CAR prior on the random frailties $\mathbf{U} = (U_1, \dots, U_I)'$ and $\mathbf{V} = (V_1, \dots, V_I)'$, that is,

- $U_1, \dots, U_I \sim \text{CAR}(\rho_1)$;
- $V_1, \dots, V_I \sim \text{CAR}(\rho_2)$;

where ρ_1 and ρ_2 are positive unknown hyper-parameters. We assume they have Inverse-Gamma prior with the known shape parameter $a_0 > 0$ and scale parameter $b_0 > 0$. Thus, the joint posterior distribution for $(\boldsymbol{\varphi}, \rho_1, \rho_2)$ is given by

$$\pi(\boldsymbol{\varphi}, \rho_1, \rho_2 | \mathcal{D}) \propto L(\boldsymbol{\varphi} | \mathcal{D}) \pi(\boldsymbol{\varphi}, \rho_1, \rho_2),$$

where $\pi(\boldsymbol{\varphi}, \rho_1, \rho_2) = \pi(\mathbf{b})\pi(\boldsymbol{\beta})\pi(\alpha)\pi(\mathbf{U}|\rho_1)\pi(\mathbf{V}|\rho_2)\pi(\rho_1)\pi(\rho_2)$ is the joint prior distribution for $(\boldsymbol{\varphi}, \rho_1, \rho_2)$ and $L(\boldsymbol{\varphi} | \mathcal{D})$ is the likelihood function given in (3.1). Note that, this joint posterior density is analytically intractable. Thus, we based our inference on Markov chain Monte Carlo (MCMC) simulation methods. We can observed that the full conditional distributions for the parameters $\mathbf{b}, \boldsymbol{\beta}, \alpha, \mathbf{U}$ while \mathbf{V} presents no closed forms; therefore, we will use the Metropolis-Hastings algorithm to generate posterior samples for these parameters. To avoid range restrictions concerning the parameter α , we define $\zeta = \log(\alpha)$ to transform all parameters space into real space (necessary to work with the densities from the Gaussian proposal). Let $\boldsymbol{\vartheta} = (\mathbf{b}, \boldsymbol{\beta}, \boldsymbol{\zeta}, \rho_1, \rho_2)$, using the Jacobian of this transformation, the joint prior density $\pi(\boldsymbol{\vartheta})$ can be expressed by

$$\pi(\boldsymbol{\vartheta}) = \pi(\mathbf{b}, \boldsymbol{\beta}, \boldsymbol{\zeta}^{-1}, \mathbf{U}, \mathbf{V}, \rho_1, \rho_2) \times \exp\left(\sum_{i=1}^Q \zeta_i\right), \quad (3.2)$$

where $\boldsymbol{\zeta}^{-1}$ denotes the inverse function of $\boldsymbol{\zeta}$, that is, $\boldsymbol{\zeta}^{-1} = \{\zeta_i^{-1} = \exp(\zeta_i), i = 1 \dots, Q\}$.

However, the full conditional distributions for parameters ρ_i are given by

$$\begin{aligned} \pi(\rho_i | \boldsymbol{\vartheta}_{-\rho_i}, \mathcal{D}) &\propto \pi(\boldsymbol{\psi}_i | \rho_i) \pi(\rho_i) \\ &\propto (\rho_i)^{-\frac{k}{2}} \exp\left(-\frac{1}{2\rho_i} \boldsymbol{\psi}_i' (\mathbf{D}_W - \mathbf{W}) \boldsymbol{\psi}_i\right) \rho_i^{-a_0-1} \exp(-b_0 \rho_i^{-1}) \\ &\propto \rho_i^{-(a_0 + \frac{k}{2})-1} \exp\left\{-\left(\frac{\boldsymbol{\psi}_i' (\mathbf{D}_W - \mathbf{W}) \boldsymbol{\psi}_i}{2} + b_0\right) \rho_i^{-1}\right\}, \quad i = 1, 2, \end{aligned}$$

where $\boldsymbol{\psi}_1 = \mathbf{U}$, $\boldsymbol{\psi}_2 = \mathbf{V}$ and k is the rank of the matrix $\mathbf{D}_W - \mathbf{W}$. Thus, the full conditional distributions of the parameter ρ_i is an Inverse-Gamma distribution with parameters $a_0 + (k/2)$ e $b_0 + (1/2)(\boldsymbol{\psi}_i' (\mathbf{D}_W - \mathbf{W}) \boldsymbol{\psi}_i)$. In this case, the Gibbs sampler algorithm (see Gamerman and Lopes, 2006) is used to generate a posterior sample.

Thus, the joint posterior density of $\pi(\boldsymbol{\vartheta} | \mathcal{D})$ is proportional to

$$\begin{aligned} L(\boldsymbol{\varphi}) \exp\left\{-\frac{1}{2}\left[\sigma_b^{-2} \sum_{i=0}^{p_1-1} b_i^2 + \sigma_\beta^{-2} \sum_{i=1}^{p_2} \beta_i^2 + \sum_{i=1}^Q \frac{\exp(2\zeta_i)}{\sigma_\alpha^2} + \frac{\mathbf{U}'(\mathbf{D}_W - \mathbf{W})\mathbf{U}}{\rho_1} + \frac{\mathbf{V}'(\mathbf{D}_W - \mathbf{W})\mathbf{V}}{\rho_2}\right]\right. \\ \left. - (a_0 + 1)(\log(\rho_1) + \log(\rho_2)) - \left(\frac{b_0}{\rho_1} + \frac{b_0}{\rho_2}\right) + \sum_{i=1}^Q \zeta_i\right\}. \end{aligned}$$

3.2. Dependent assumption

Now we assume the spatial priors of the parameters (\mathbf{U}, \mathbf{V}) dependent from each other. Let $\boldsymbol{\psi} = (\mathbf{U}', \mathbf{V}')'$, we first employed the parameter $\boldsymbol{\psi}$ as a MCAR distribution with a common smoothness parameter a , i.e.,

$$\boldsymbol{\psi} \sim \text{MCAR}(a, \boldsymbol{\Lambda}).$$

Further, we employ the parameter ψ an extend MCAR distribution that assumes the different smoothness parameters for the parameters U and V , namely a_1 and a_2 , that is,

$$\psi \sim \text{MCAR}(a_1, a_2, \Lambda).$$

Here, we consider the prior distributions for the parameters \mathbf{a} and Λ the same that were used by Carlin and Banerjee (2003), that is,

- $a_i \sim \text{Uniform}(0, 1)$ or $a_i \sim \text{Beta}(18, 2)$, for i ;
- $\Lambda \sim \text{Wishart}(n_0, \Lambda_0)$, with n_0 and Λ_0 known;

where $i = 1$ for $\psi \sim \text{MCAR}(a, \Lambda)$ and $i = 1, 2$ for $\psi \sim \text{MCAR}(a_1, a_2, \Lambda)$.

To avoid range restrictions in parameters a_i , considering the transformations $\rho_i = \log(a_i/(1-a_i)) \in \mathbb{R}$, then, the joint prior density for the parameters $\boldsymbol{\theta} = (\mathbf{b}, \boldsymbol{\beta}, \kappa, \boldsymbol{\zeta}, \psi, \Lambda, \boldsymbol{\rho})$ can be written as

$$\pi(\boldsymbol{\theta}) = \pi(\mathbf{b})\pi(\boldsymbol{\beta})\pi(\boldsymbol{\zeta})\pi(\psi|\Lambda, \boldsymbol{\rho})\pi(\Lambda)\pi(\boldsymbol{\rho}) \times \exp\left(\sum_{i=1}^Q \zeta_i\right) \prod_{i=1}^2 \frac{\exp(-\rho_i)}{(1 + \exp(-\rho_i))^2}.$$

The joint posterior density is given by

$$\begin{aligned} \pi(\boldsymbol{\theta}|\mathbf{D}) \propto L(\boldsymbol{\varphi}^{-1}) \exp\left\{-\frac{1}{2}\left[\sigma_b^{-2} \sum_{i=0}^{p_1} b_i^2 + \sigma_\beta^{-2} \sum_{i=1}^{p_2} \beta_i^2 + \sum_{i=1}^Q \frac{\exp(2\zeta_i)}{\sigma_\alpha^2}\right]\right. \\ \left. + \boldsymbol{\psi}' [\Lambda \otimes (\mathbf{D}_W - \mathbf{a}W)] \boldsymbol{\psi} + \log |\Lambda \otimes \mathbf{a}W| + \frac{n_0 - 4}{2} \log |\Lambda| - \frac{1}{2} \text{tr}(\Lambda_0^{-1} \Lambda) + \sum_{i=1}^Q \zeta_i\right\} \pi(\boldsymbol{\rho}), \end{aligned}$$

where $\boldsymbol{\varphi}^{-1} = (\mathbf{b}, \boldsymbol{\beta}, \boldsymbol{\zeta}^{-1}, U, V)$ and $\pi(\rho_i) = 1$ if $a_i \sim \text{Uniform}(0, 1)$ and $\pi(\rho_i) = \{1/B(18, 2)\} \{\exp(17\rho_i)/(1 + \exp(\rho_i))^{18}\}$ if $a_i \sim \text{Beta}(18, 2)$, where $B(18, 2) = 17!/18! = 1/18$.

This joint posterior density is analytically intractable. Thus, we based our inference on MCMC simulation methods. We can observed that the full conditional distributions for the parameters $\mathbf{b}, \boldsymbol{\beta}, \boldsymbol{\zeta}, \psi$ and $\boldsymbol{\rho}$ do not present closed form; therefore, we will use the Metropolis-Hastings algorithm to generate posterior samples for these parameters. However, the Gibbs sampler algorithm is used to generate a posterior sample for the parameter Λ , because its full conditional distribution has a closed form. The full conditional distribution $\pi(\Lambda|\boldsymbol{\theta}_{(-\Lambda)}, \mathbf{D}_{\text{obs}})$ is proportional to

$$\begin{aligned} \pi(\psi|\Lambda, \mathbf{a})\pi(\Lambda) \propto |\Lambda \otimes \mathbf{D}_W - \mathbf{a}W|^{\frac{1}{2}} \exp\left(-\frac{1}{2}\boldsymbol{\psi}'(\mathbf{D}_W - \mathbf{a}W)\boldsymbol{\psi}\right) |\Lambda|^{\frac{n_0-4}{2}} \exp\left(-\frac{1}{2}\text{tr}(\Lambda_0^{-1} \Lambda)\right) \\ \propto |\Lambda|^{\frac{I+n_0-4}{2}} \exp\left\{-\frac{1}{2}\text{tr}((\Lambda_0^{-1} + \mathbf{B}) \Lambda)\right\}, \end{aligned} \quad (3.3)$$

where

$$\mathbf{B} = \begin{bmatrix} \text{tr}(\mathbf{R}_1 U (\mathbf{R}_1 U)') & \text{tr}(\mathbf{R}_1 U (\mathbf{R}_2 V)') \\ \text{tr}(\mathbf{R}_2 V (\mathbf{R}_1 U)') & \text{tr}(\mathbf{R}_2 V (\mathbf{R}_2 V)') \end{bmatrix}.$$

Thus, the full conditional distribution for Λ can be taken by the Wishart distribution with scale matrix $(\Lambda_0^{-1} + \mathbf{B})^{-1}$ and degrees of freedom $I + n_0$.

3.3. Model comparison criteria

There are several Bayesian criteria to compare competing models for a given data set and to select the one that best fits the data. One of the most used in applied works is based on the posterior mean of deviance and called the DIC. For a model, the statistic DIC is defined as

$$\text{DIC} = \bar{d} + \text{pd},$$

where $\bar{d} = E[D(\boldsymbol{\varphi})]$, $\text{pd} = E[D(\boldsymbol{\varphi})] - D[E(\boldsymbol{\varphi})]$ and $D(\boldsymbol{\varphi})$ is the deviance function of the model defined by $-2 \log L(\boldsymbol{\varphi})$. L is the likelihood function of the model. Spiegelhalter *et al.* (2002) provide evidences that pd is a suitable measure of model complexity even in hierarchical settings, and thus, DIC is considered as a sensible generalization of the expected Akaike information criterion to hierarchical settings. The model, with the smallest value of DIC, is commonly taken as the preferred model describe the data set given.

3.4. Bayesian case influence diagnostics

Performing a sensitivity analysis is advisable since regression models are sensitive to underlying model assumptions. One of the most used ways of evaluating the influence of an observation in the fitted model is a case-deletion (Cook and Weisberg, 1982), where the effects are studied by completely removing cases from the analysis. This reasoning forms the basis of the Bayesian global influence methodology and makes it possible to determine which subjects might influence the analysis. The Bayesian case-deletion influence diagnostic measures for the joint posterior distribution based on the ψ -divergence (Peng and Dey, 1995; Weiss, 1996) is now introduced as follows.

Let $D_\psi(P, P_{(-i)})$ denote the ψ -divergence between P and $P_{(-i)}$, where P denotes the posterior distribution of $\boldsymbol{\theta}$ for full data, and $P_{(-i)}$ denotes the posterior distribution of $\boldsymbol{\theta}$ without the i th case. Therefore,

$$D_\psi(P, P_{(-i)}) = \int \psi \left(\frac{\pi(\boldsymbol{\theta}|\mathcal{D}^{(-i)})}{\pi(\boldsymbol{\theta}|\mathcal{D})} \right) \pi(\boldsymbol{\theta}|\mathcal{D}) d\boldsymbol{\theta}, \quad (3.4)$$

where ψ is a convex function with $\psi(1) = 0$. Several choices concerning the ψ are given by Dey and Birmiwal (1994). For example, $\psi(z) = -\log(z)$ defines the Kullback-Leibler (K-L) divergence, $\psi(z) = (z - 1) \log(z)$ gives J -distance (or the symmetric version of K-L divergence), $\psi(z) = 0.5|z - 1|$ defines the variational distance (or L_1 norm) and $\psi(z) = (z - 1)^2$ defines the χ^2 -square divergence.

Let $\boldsymbol{\theta}^{(1)}, \dots, \boldsymbol{\theta}^{(Q)}$ be a size Q sample of $\pi(\boldsymbol{\theta}|\mathcal{D})$, $D_\psi(P, P_{(-i)})$ can be calculated numerically by

$$\widehat{D}_\psi(P, P_{(-i)}) = \frac{1}{Q} \sum_{q=1}^Q \psi \left(\frac{\widehat{\text{CPO}}_i}{L(y_i|\boldsymbol{\theta}^{(q)})} \right), \quad (3.5)$$

where $\widehat{\text{CPO}}_i = [(1/Q) \sum_{q=1}^Q 1/\{L(y_i|\boldsymbol{\theta}^{(q)})\}]^{-1}$ is the numerical approximation of the conditional predictive ordinate statistic of i^{th} observation (Ibrahim *et al.*, 2001).

Note that $D_\psi(P, P_{(-i)})$ can be interpreted as the ψ -divergence of the effect of deleting the i -th case from the full data on the joint posterior distribution of $\boldsymbol{\theta}$. As pointed by Peng and Dey (1995), Weiss (1996), and Cancho *et al.* (2010), it may be difficult for a practitioner to judge the cutoff point of the divergence measure to determine if a small subset of observations is influential. In this context, we will use the proposal given by Peng and Dey (1995) and Weiss (1996) by considering a biased coin,

which has success probability p . Then the ψ -divergence between the biased and an unbiased coin is

$$D_\psi(f_0, f_1) = \int \psi\left(\frac{f_0(x)}{f_1(x)}\right) f_1(x) dx, \quad (3.6)$$

where $f_0(x) = p^x(1-p)^{1-x}$ and $f_1(x) = 0.5$, $x = 0, 1$. Now if $D_\psi(f_0, f_1) = d_\psi(p)$, then it would be easy to check if d_ψ satisfies the following equation

$$d_\psi(p) = \frac{\psi(2p) + \psi(2(1-p))}{2}. \quad (3.7)$$

It is not difficult to see for the divergence measures considered that d_ψ increases as p moves away from 0.5. In addition, $d_\psi(p)$ is symmetric about $p = 0.5$ and d_ψ achieves its minimum at $p = 0.5$. At this point, $d_\psi(0.5) = 0$, and $f_0 = f_1$. Therefore, if we consider $p > 0.90$ (or $p \leq 0.10$) as a strong bias in a coin, then $d_{K-L}(0.90) = 0.51$, $d_J(0.90) = 0.88$, $d_{L_1}(0.90) = 0.4$, and $d_{\chi^2}(0.90) = 0.64$. This equation implies that i^{th} case is influential when $d_{L_1} > 0.4$ or $d_{\chi^2} > 0.64$. Thus, if we use the K-L divergence, we can consider an influential observation when $d_{K-L} > 0.51$. Similarly, using the J-distance, an observation which implies $d_J > 0.88$ can be considered influential.

4. Simulation study

This section presents simulation studies for a cure rate model under FA with the dependent assumption of examining their performances. The interval-censored survival times $(t_{ijL}, t_{ijR}, \delta_{ij})$ with the cure fraction under the FA are generated similarly to what was introduced by Yau and Ng (2001) with some modifications.

We generate the latent Geometric variable M_{ij} , which denotes the initial number of competing causes regarding the event, with parameter $p_{0ij} = [1 + \exp(-(b_0 + b_1 x_{ij}) + v_i)]^{-1}$ for the j^{th} individual in the i^{th} region, $j = 1, \dots, n_i$, $i = 1, \dots, I$, where the covariate x_{ij} is generated from the Bernoulli distribution with parameter 0.5. Interval-censored data $(t_{ijL}, t_{ijR}, \delta_{ij})$ are then generated as:

- (i) If $M_{ij} = 0$, then let $t_{ij} = t_{ijL}$ from the Exponential distribution with hazard rate 10 and let the censoring indicator $\delta_{ij} = 0$.
- (ii) If $M_{ij} > 0$, then we generate M_{ij} latent Weibull variables with parameter $\alpha = 0.30$ and $\lambda_{ij} = (\beta x_{ij} + u_i)$ with $\beta = -0.15$. Let t_{ij} takes the lowest generated variable in case of generating the variables of the model under FA. The censoring times c_{ij} were sampled from Uniform $U(0, \tau)$, where $\tau > 0$ was set in order to control the proportion of censored observations to be approximately 40% on average and, the indicator variables will be of $\delta_{ij} = 1$ if $t_{ij} \leq c_{ij}$ and $\delta_{ij} = 0$, otherwise.
- (iii) For $\delta_{ij} = 0$, let $0 < t_{ijL} < t_{ijR} = \infty$.
- (iv) For $\delta_{ij} = 1$, we create len_{ij} from the distribution $U(0.2, 0.7)$ and l_{ij} from $U(0, 0.01)$. Then, from $(0, l_{ij}]$, $(l_{ij}, l_{ij} + \text{len}_{ij}]$, \dots , $(l_{ij} + k\text{len}_{ij}, \infty]$, $k = 1, 2, \dots$, the interval $(t_{ijL}, t_{ijR}]$ which satisfies $t_{ijL} < t_{ij} \leq t_{ijR}$ is chosen.

Table 1: Simulation results for the Weibull cure model under the first activation with depended spatial frailties

Parameter	True value	Estimate mean	SD of the estimate	Bias	MSE	SDs mean
Prior 1: $\psi \sim \text{MCAR}(a, \Lambda)$, $a \sim \text{Beta}(18, 2)$, $\Lambda_0 \sim \text{Wishart}(2, \text{Diag}(0.9, 1))$						
b_0	-1.50	-1.4812	0.0708	0.0188	0.0054	0.2685
b_1	-0.50	-0.5209	0.1280	-0.0209	0.0168	0.2485
β	-0.15	-0.1360	0.0466	0.0140	0.0024	0.1915
α	0.30	0.1930	0.0521	-0.1070	0.0142	0.0677
Λ_{11}	4.00	4.0062	0.1597	0.0062	0.0255	2.4728
Λ_{22}	4.00	4.0120	0.1918	0.0120	0.0369	2.6151
Λ_{12}	0.00	-0.4541	0.1343	-0.4541	0.2242	1.9196
a	0.90	0.9001	0.0016	0.0001	0.0000	0.0653
Prior 2: $\psi \sim \text{MCAR}(a_1, a_2, \Lambda)$, $a_1, a_2 \sim \text{Beta}(18, 2)$, $\Lambda_0 \sim \text{Wishart}(2, \text{Diag}(0.9, 1))$						
b_0	-1.50	-1.4902	0.0701	0.0098	0.0050	0.2583
b_1	-0.50	-0.5376	0.1330	-0.0376	0.0191	0.2227
β	-0.15	-0.1295	0.0493	0.0205	0.0028	0.1870
α	0.30	0.1863	0.0443	-0.1137	0.0149	0.0536
Λ_{11}	4.00	4.1638	0.1676	0.1638	0.0549	2.5070
Λ_{22}	4.00	4.2657	0.1819	0.2657	0.1036	2.6919
Λ_{12}	0.00	-0.5809	0.1472	-0.5809	0.3591	1.9647
a_1	0.90	0.8999	0.0015	-0.0002	0.0000	0.0655
a_2	0.90	0.9002	0.0015	0.0002	0.0000	0.0654

SD = standard deviation; Bias = average bias; MSE = mean square error; SDs mean = average SD; MCAR = multivariate conditionally autoregressive.

In this study, we consider $I = 5$ regions (Zip) with the corresponding adjacent matrix

$$\begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}.$$

Random effects U_i and U_i are generated from the Normal distribution with mean $\mathbf{0}$ and precision matrix $\Lambda \otimes (D_W - aW)$, where W is the standardized adjacent matrix so that each of its rows sum one, $D_W = \text{Diag}(1, 1, 2, 1, 1)$ is a diagonal matrix, and then we fixed $a = 0.9$ and $\Lambda = \text{Diag}(4, 4)$, i.e., we fixed $\Lambda_{11} = 4$, $\Lambda_{22} = 4$, and $\Lambda_{12} = \Lambda_{21} = 0$. By considering 100 individuals in the simulation studies. The corresponding zip codes for each individual were distributed using sample with replace, thus the number of individuals in each region n_i , $i = 1, \dots, 5$ are varied; therefore, these five regions could present different numbers of individuals with $\sum_{i=1}^5 n_i = 100$. Thus, we have sample size $n = 100$ and we fixed the parameters $b_0 = -1.50$, $b_1 = -0.50$, $\beta = -0.15$, $\alpha = 0.30$. In simulations, we used to consider around 40% of the censored data for each generated sample with 500 repeated samples simulated for each model. The priors for the parameters b_0 , b_1 , β , and α used in the studies are $b_0 \sim N(0, 3^2)$, $b_1 \sim N(0, 3^2)$, $\beta \sim N(0, 3^2)$, and $\alpha \sim N(0, 10^2)I_{(0, \infty)}$.

For each generated sample, we simulate one chain of size 10,000 for each parameter, disregarding the first 1,000 iterations to eliminate the effect of the initial values and avoid correlation problems and thinning to every third iteration, thus obtaining an effective sample of size 3,000 upon which the posterior is based on. To evaluate the performance of the parameter estimates, the average bias (Bias), standard deviation (SD) of the estimate, average SD (SDs mean), and the mean square error (MSE) are calculated for a cure rate model under the FA (Table 1).

We can note that the bias and the MSE of the parameter Λ_{12} are larger than other parameters. The estimator of the Λ_{12} presents a negative biases; however, the biases and MSEs are always near zero.

Table 2: Simulation results of the perturbed cases for the Weibull cure rate models under the first activation

Setup	Perturbed case	Prior 1					Prior 2				
		Parameters	Mean	SD	Bias	MSE	Parameters	Mean	SD	Bias	MSE
A	None	b_0	-1.482	0.278	0.018	0.000	b_0	-1.341	0.277	0.159	0.025
		b_1	-0.859	0.259	-0.359	0.129	b_1	-0.554	0.248	-0.054	0.003
		β	-0.036	0.193	0.114	0.013	β	-0.105	0.197	0.045	0.002
		α	0.228	0.068	-0.072	0.005	α	0.112	0.072	-0.188	0.035
		Λ_{11}	4.172	2.533	0.172	0.030	Λ_{11}	4.052	2.486	0.052	0.003
		Λ_{22}	4.152	2.693	0.152	0.023	Λ_{22}	4.012	2.636	0.012	0.000
		Λ_{12}	-0.502	1.945	-0.502	0.252	Λ_{12}	-0.540	1.905	-0.540	0.291
		a	0.902	0.064	0.002	0.000	a_1	0.899	0.068	-0.001	0.000
							a_2	0.899	0.067	-0.001	0.000
		b_0	-1.526	0.263	-0.026	0.001	b_0	-1.572	0.275	-0.072	0.005
		b_1	-0.439	0.253	0.061	0.004	b_1	-0.622	0.260	-0.122	0.015
		β	-0.155	0.189	-0.005	0.000	β	-0.099	0.190	0.051	0.003
		α	0.231	0.061	-0.069	0.005	α	0.332	0.073	0.032	0.001
		Λ_{11}	4.057	2.433	0.057	0.003	Λ_{11}	4.056	2.494	0.056	0.003
B	{18}	Λ_{22}	3.947	2.630	-0.053	0.003	Λ_{22}	3.913	2.561	-0.087	0.008
		Λ_{12}	-0.497	1.904	-0.497	0.247	Λ_{12}	-0.406	1.991	-0.406	0.164
		a	0.900	0.065	0.000	0.000	a_1	0.899	0.065	-0.001	0.000
							a_2	0.900	0.065	0.000	0.000
		b_0	-1.510	0.275	-0.010	0.000	b_0	-1.524	0.263	-0.024	0.001
		b_1	-0.586	0.257	-0.086	0.007	b_1	-0.679	0.264	-0.179	0.032
		β	-0.139	0.194	0.011	0.000	β	-0.083	0.193	0.067	0.005
		α	0.255	0.082	-0.045	0.002	α	0.235	0.078	-0.065	0.004
		Λ_{11}	3.832	2.406	-0.168	0.028	Λ_{11}	4.018	2.386	0.018	0.000
		Λ_{22}	3.535	2.423	-0.465	0.216	Λ_{22}	3.919	2.649	-0.081	0.007
		Λ_{12}	-0.115	1.835	-0.115	0.013	Λ_{12}	-0.302	1.969	-0.302	0.091
		a	0.900	0.063	0.000	0.000	a_1	0.899	0.064	-0.001	0.000
							a_2	0.900	0.068	0.000	0.000
		b_0	-1.460	0.259	0.040	0.002	b_0	-1.599	0.266	-0.099	0.010
C	{80}	b_1	-0.348	0.248	0.152	0.023	b_1	-0.316	0.246	0.184	0.034
		β	-0.203	0.189	-0.053	0.003	β	-0.208	0.191	-0.058	0.003
		α	0.187	0.059	-0.113	0.013	α	0.253	0.061	-0.047	0.002
		Λ_{11}	4.205	2.569	0.205	0.042	Λ_{11}	4.259	2.541	0.259	0.067
		Λ_{22}	4.086	2.634	0.086	0.007	Λ_{22}	4.002	2.549	0.002	0.000
		Λ_{12}	-0.515	1.971	-0.515	0.265	Λ_{12}	-0.588	1.946	-0.588	0.346
		a	0.900	0.064	0.000	0.000	a_2	0.901	0.065	0.001	0.000
							a_2	0.898	0.068	-0.002	0.000
		b_0	-1.460	0.259	0.040	0.002	b_0	-1.599	0.266	-0.099	0.010
		b_1	-0.348	0.248	0.152	0.023	b_1	-0.316	0.246	0.184	0.034
		β	-0.203	0.189	-0.053	0.003	β	-0.208	0.191	-0.058	0.003
		α	0.187	0.059	-0.113	0.013	α	0.253	0.061	-0.047	0.002
		Λ_{11}	4.205	2.569	0.205	0.042	Λ_{11}	4.259	2.541	0.259	0.067
		Λ_{22}	4.086	2.634	0.086	0.007	Λ_{22}	4.002	2.549	0.002	0.000
		Λ_{12}	-0.515	1.971	-0.515	0.265	Λ_{12}	-0.588	1.946	-0.588	0.346
		a	0.900	0.064	0.000	0.000	a_2	0.901	0.065	0.001	0.000
							a_2	0.898	0.068	-0.002	0.000

SD = standard deviation; Bias = average bias; MSE = mean square error.

Moreover, the simulation results for the cure model considering the prior 1 are close to those obtained considering the prior 2.

Influence of outlying observations

A goal of this study is to show the need for robust models to deal with the presence of outliers in data. We consider two cases for perturbation with the parameter's values and the setup the same as in the simulation studies; therefore, four data sets of size 100 were generated from the cure model under the FA with dependent spatial frailties.

We selected the cases 18 and 80 for perturbation. To create influential observations concerning the data set, we choose one or two of these selected cases and perturbed the response variable as: $\widetilde{t}_{kL} = t_{kL} + 10SD_L$ and $\widetilde{t}_{kR} = t_{kR} + 10SD_L$, for $k = 18$ and 80 , where SD_L is the SD of the t_{ijL} . Note that this perturbation method will not change the time interval of the perturbation candidate. Here,

Table 3: ψ -divergence measures of the perturbed cases and the values of DIC for simulated data sets

Activation	Prior	Setup	Perturbed case	No. of case	d_{KL}	d_J	d_{L1}	d_{χ^2}	DIC
First	1	A	None	18	0.006	0.011	0.042	0.012	142.754
				80	0.030	0.060	0.097	0.067	
		B	{18}	18	0.080	0.171	0.159	0.238	164.476
		C	{80}	80	0.246	0.602	0.277	2.128	153.082
		D	{18,80}	18	0.069	0.143	0.147	0.181	185.709
				80	0.282	0.677	0.304	1.742	
	2	A	None	18	0.007	0.014	0.046	0.014	140.186
				80	0.033	0.067	0.102	0.075	
		B	{18}	18	0.036	0.075	0.106	0.084	164.446
		C	{80}	80	0.294	0.940	0.288	14.544	149.760
		D	{18,80}	18	0.062	0.131	0.138	0.176	184.934
				80	0.120	0.269	0.190	0.498	

DIC = deviance information criterion.

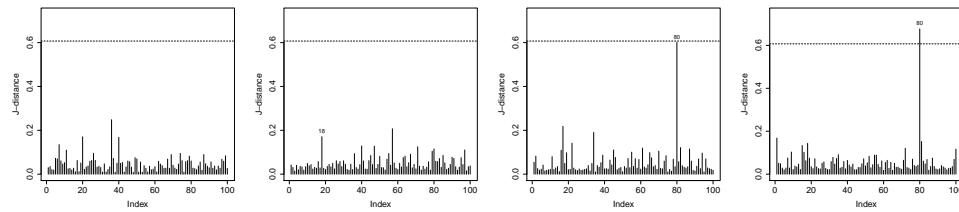


Figure 1: Index plots of J -distance measure from the fitted of the Weibull cure rate model-first activation. From left to right, the data sets of the figures correspond to the setup A, B, C, and D.

we consider four setups in the study. Setup A: original dataset, without outliers, Setup B: data with outlier 18, Setup C: data with outlier 80, and Setup D: data with outliers 18 and 80. The MCMC computations were similar to those in the last section; in addition, Gewekes convergence diagnostic proposed by Geweke (1992) was used to monitor the convergence of the Gibbs samples.

Table 2 reports the posterior mean, the SD, the bias and the MSE of the parameters of WCRM-FA. We can note that the estimative of parameter Λ_{11} creasing in the perturbation cases when prior 1 is used. However, considering prior 2 for the parameters, the estimative of all parameters of cases B, C, and D are very closed the case A, which means that the parameters are not sensitive to perturbations.

For each simulated data set, the four divergence measures (d_{KL} , d_J , d_{L1} , d_{χ^2}) of the perturbed cases and the DIC values for the Weibull and the proportional hazards (PH) cure rate models were calculated (Table 3).

We note that all measures provide larger ψ -divergence measures when compared to the non-perturbed setup (setup A) and that the difference between the measures of the perturbed cases and non-perturbed cases is more clear for the PH cure rate models than the WCRMs. Furthermore, we observe that the values of the measures obtained from the cure models, whether the parameters have the prior 1 or 2, are similar.

To better show present the results, we plot the J -distance measure from the fitted models considering the prior 1 for the parameters. The Figure 1 presents the divergence measures before the perturbation (setup A), the model indicates the absence of outline observations, and after perturbation observations (setups B, C, and D). Note that outline observation 18 cannot be easy detected for the cure rate model under the FA.

Table 4: Real dataset: Bayesian criteria for the fitted models

Cure rate model under first activation		Criteria	
Model	Priors	DIC	pd
1	$U \sim \text{CAR}(\rho_1)$, $V \sim \text{CAR}(\rho_2)$, $\rho_1, \rho_2 \sim \text{InvGamma}(0.01, 0.01)$	416.3917	11.5105
2	$\psi \sim \text{MCAR}(a, \Lambda)$, $a \sim \text{Uniform}(0, 1)$, $\Lambda \sim \text{Wishart}(2, \text{Diag}(0.1, 0.1))$	416.9691	11.6971
3	$\psi \sim \text{MCAR}(a_1, a_2, \Lambda)$, $a_1, a_2 \sim \text{Uniform}(0, 1)$, $\Lambda \sim \text{Wishart}(2, \text{Diag}(0.1, 0.1))$	417.4942	12.1639
4	$\psi \sim \text{MCAR}(a, \Lambda)$, $a \sim \text{Beta}(18, 2)$, $\Lambda \sim \text{Wishart}(2, \text{Diag}(0.1, 0.1))$	417.1560	12.6639
5	$\psi \sim \text{MCAR}(a_1, a_2, \Lambda)$, $a_1, a_2 \sim \text{Beta}(18, 2)$, $\Lambda \sim \text{Wishart}(2, \text{Diag}(0.1, 0.1))$	416.6700	11.7184
Cure rate model under last activation		Criteria	
Model	Priors	DIC	pd
6	$U \sim \text{CAR}(\rho_1)$, $V \sim \text{CAR}(\rho_2)$, $\rho_1, \rho_2 \sim \text{InvGamma}(0.01, 0.01)$	419.2723	11.8914
7	$\psi \sim \text{MCAR}(a, \Lambda)$, $a \sim \text{Uniform}(0, 1)$, $\Lambda \sim \text{Wishart}(2, \text{Diag}(0.1, 0.1))$	419.1585	12.2811
8	$\psi \sim \text{MCAR}(a_1, a_2, \Lambda)$, $a_1, a_2 \sim \text{Uniform}(0, 1)$, $\Lambda \sim \text{Wishart}(2, \text{Diag}(0.1, 0.1))$	419.4509	12.6852
9	$\psi \sim \text{MCAR}(a, \Lambda)$, $a \sim \text{Beta}(18, 2)$, $\Lambda \sim \text{Wishart}(2, \text{Diag}(0.1, 0.1))$	417.8917	13.6477
10	$\psi \sim \text{MCAR}(a_1, a_2, \Lambda)$, $a_1, a_2 \sim \text{Beta}(18, 2)$, $\Lambda \sim \text{Wishart}(2, \text{Diag}(0.1, 0.1))$	418.6432	13.9228

DIC = deviance information criterion.

5. Application

Recall the interval-censored smoking cessation data briefly presented in the introduction section. In the smoking cessation study, all patients (smokers) were randomized into either a SI group, or a UC group that received no special anti-SI. The treatments of the SI program were realized in Rochester city localizing in the center of the maps. The program detail can be found in Murray *et al.* (1998). Here, each patient was observed annually over five years of monitoring. Our event of interest concerns if they will relapse (resume smoking). If a smoker starts smoking again after an initial attempt to quit, then only an approximate one-year time interval was observed from the previous observation to the current observation. Thus, the relapse times are interval-censored. In this analysis, we limit our attention to those patients who are known to have quit smoking at least once during the study period and who have an identifiable Minnesota zip code of residence. Thus, the data consists of a total of 223 patients who reside in 51 zip codes in the Southeastern corner of Minnesota, among them there are 65 patients that have undergone relapse, which implies the empirical cure rate is approximately 71%. Moreover, Ma and Xiang (2013) have also confirmed the existence of non-eligible cure fraction in the population.

We fitted cure rate model under first and LA, considering the different spatial frailties in the models to the data set. The prior distributions for the parameters b , β , and α are:

- $b_j \sim N(0, 100)$, $j = 0, \dots, 4$;
- $\beta_j \sim N(0, 100)$, $j = 1, \dots, 4$;
- $\alpha \sim N(0, 100)I_{(0, \infty)}$.

Because of the high computational cost, we implement the MCMC algorithms in the C programming language; subsequently, the results were analyzed in the R language (R Development Core Team, 2010) through the “coda” package (Plummer *et al.*, 2005). All of our MCMC algorithms ran a total of 60,000 iterations discarding the first 20,000 realizations as burn-in and thinning to every fifth iteration. Posterior results are then based on 8,000 realizations of the Markov chain. Our Metropolis acceptance rate for these parameters ranged from 25% to 50%. The convergence was checked using the Geweke diagnostic which did not indicate a lack of convergence. The models are compared using DIC criterion.

Table 5: Real dataset: Bayesian criteria for the fitted models

Cure rate model under first activation			Criteria	
Model	Priors		DIC	pd
11	$U \sim \text{CAR}(\rho_1), V \sim \text{CAR}(\rho_2), \rho_1, \rho_2 \sim \text{InvGamma}(0.01, 0.01)$		414.3751	8.2277
12	$\psi \sim \text{MCAR}(a, \Lambda), a \sim \text{Uniform}(0, 1), \Lambda \sim \text{Wishart}(2, \text{Diag}(0.1, 0.1))$		414.7620	10.9257
13	$\psi \sim \text{MCAR}(a_1, a_2, \Lambda), a_1, a_2 \sim \text{Uniform}(0, 1), \Lambda \sim \text{Wishart}(2, \text{Diag}(0.1, 0.1))$	1	414.7484	10.8353
14	$\psi \sim \text{MCAR}(a, \Lambda), a \sim \text{Beta}(18, 2), \Lambda \sim \text{Wishart}(2, \text{Diag}(0.1, 0.1))$		414.3483	10.9108
15	$\psi \sim \text{MCAR}(a_1, a_2, \Lambda), a_1, a_2 \sim \text{Beta}(18, 2), \Lambda \sim \text{Wishart}(2, \text{Diag}(0.1, 0.1))$		414.4925	10.9244
Cure rate model under last activation			Criteria	
Model	Priors		DIC	pd
16	$U \sim \text{CAR}(\rho_1), V \sim \text{CAR}(\rho_2), \rho_1, \rho_2 \sim \text{InvGamma}(0.01, 0.01)$		418.1336	9.4144
17	$\psi \sim \text{MCAR}(a, \Lambda), a \sim \text{Uniform}(0, 1), \Lambda \sim \text{Wishart}(2, \text{Diag}(0.1, 0.1))$		417.2717	11.7087
18	$\psi \sim \text{MCAR}(a_1, a_2, \Lambda), a_1, a_2 \sim \text{Uniform}(0, 1), \Lambda \sim \text{Wishart}(2, \text{Diag}(0.1, 0.1))$		416.8620	11.4699
19	$\psi \sim \text{MCAR}(a, \Lambda), a \sim \text{Beta}(18, 2), \Lambda \sim \text{Wishart}(2, \text{Diag}(0.1, 0.1))$		416.8979	11.6343
20	$\psi \sim \text{MCAR}(a_1, a_2, \Lambda), a_1, a_2 \sim \text{Beta}(18, 2), \Lambda \sim \text{Wishart}(2, \text{Diag}(0.1, 0.1))$		416.8782	11.5676

DIC = deviance information criterion.

Table 6: Posterior summaries of the parameter of the Model 15 for the smoking cessation data

Parameter	Survival model					Cure rate				
		Mean	SD	2.5%	97.5%		Mean	SD	2.5%	97.5%
Intercept						b_0	1.3736	0.5789	0.2600	2.5442
Sex (male = 0)	β_1	-0.1562	0.4551	-1.0626	0.6992	b_1	-0.5096	0.3718	-1.2711	0.2052
SI/UC (UC = 0)	β_2	0.8427	0.5191	-0.1659	1.8703	b_2	0.8601	0.4310	0.0671	1.7192
Cigarettes per day	β_3	-0.1148	0.0378	-0.1809	-0.0322	b_3	-0.0728	0.0290	-0.1345	-0.0201
Duration as smoker	β_4	-0.0246	0.0343	-0.1003	0.0345	b_4	0.0197	0.0230	-0.0264	0.0651
α		2.4097	0.3073	1.8113	3.0065					
$a_1(a_u)$		0.8968	0.0676	0.7261	0.9874					
$a_2(a_v)$							0.8994	0.0670	0.7370	0.9881
Λ_{11}		2.6769	0.6377	1.5543	4.0673					
Λ_{22}							2.5743	0.6413	1.4924	3.9718
Λ_{12}		-0.0104	0.4625	-0.9321	0.8919					
Σ_{11}		0.4113	0.1075	0.2531	0.6754					
Σ_{22}							0.4298	0.1236	0.2561	0.7298
$\Sigma_{12}/(\Sigma_{11}\Sigma_{22})^{1/2}$		0.0035	0.1828	-0.3655	0.3559					

Where Λ_{ij} is the element of precision matrix Λ in position (i, j) , and Σ_{ij} is the element of matrix $\Sigma = \Lambda^{-1}$ in position (i, j) , this Σ_{11} is the spatial variance component of U and Σ_{22} is the spatial variance component of V , $\Sigma_{12}/(\Sigma_{11}\Sigma_{22})^{1/2}$ denote their correlation. SD = standard deviation; SI = smoking intervention group; UC = usual care group.

Table 4 provides the DIC scores for a variety of effects of the Weibull cure model under FA and LA. The DIC scores of the Model 1 and 5 show the best models despite DIC values that are close to each other. We also can note that the cure rate models under the FA are more adequate than models under the LA.

For the comparison with the models proposed by Carlin and Banerjee (2003), we consider the same prior distributions for the parameters b and β as considered by these authors. Table 5 reports the DIC scores for a variety of effects for the WCRM. We observe that the DIC scores in Table 5 are smaller than the values in Table 4. However, the DIC scores for the cure rate model under FA are very close each other and indicates that these models are equivalent. Moreover, similarly to the previous case, the DIC values of the cure rate model under FA are smaller than the cure rate model under LA. Comparing the obtained DIC scores in Table 4 and 5 with the DIC scores presented in the study conducted by Carlin and Banerjee (2003), where they proposed the mixture cure model with the spatial frailty, we can conclude that all our models are more adequate since the DIC scores are

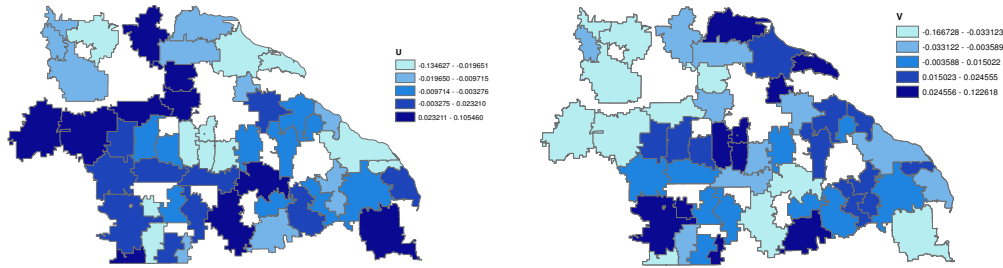


Figure 2: Maps of posterior means for frailties U (left panel) and V (right panel) in Model 15.

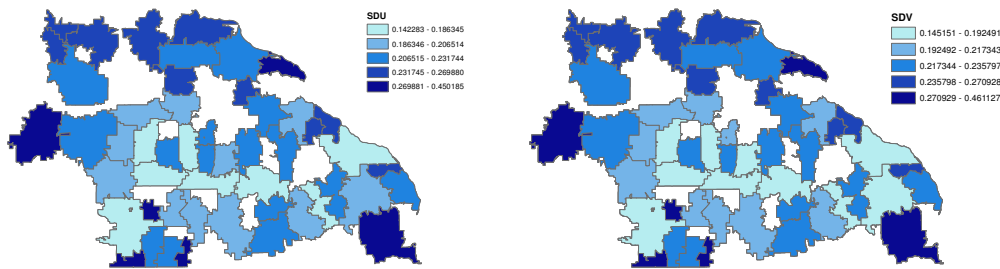


Figure 3: Maps of posterior standard deviation for frailties U (left panel) and V (right panel) in Model 15.

smaller. Here, we select the Model 15 as our working model.

Table 6 presents the posterior summary of the parameters of the Model 15. We note that only the parameters b_0 , b_2 , b_3 , and β_3 are significant. In the cure rate, the negative value of b_3 means that the individuals with higher levels of cigarette consumption have a lower probability to quit smoking, while the positive value of b_2 implies the individuals with special intervention have higher probability to quit smoking than those with UC.

The survival model shows that the special intervention and the number of cigarettes smoked per day have negative effects on the hazard rate of the relapse time; therefore, individuals with special intervention do not present lower hazard rates for the relapse time when compared to those who attend UC. However, the individuals with a higher level of cigarette consumption do not present high hazard rates.

The estimated $SD \Sigma_{11}^{1/2}$ of random spatial effects in the survival model is 0.4113, and the estimated $SD \Sigma_{22}^{1/2}$ of the random spatial effects in the cure rate is 0.4268, which indicates that there is a considerable heterogeneity among clusters. It is also observed that there are no correlations between the spatial effects U and V .

The Figure 2 maps the posterior means of the frailties U and V in the Model 15. For the frailties U of which the high value presents the high relapse rate, we note that the city of Owatonna and some southern cities have higher values with individuals in these regions indicating higher relapse rates than others. However, the city of Rochester (in the central region) suggests slightly better than average cessation behavior, which can also be observed by the frailties V . Note that the high value of V presents the high cure probability.

The Figure 3 maps the posterior SD of the frailties U and V corresponding to the posterior means

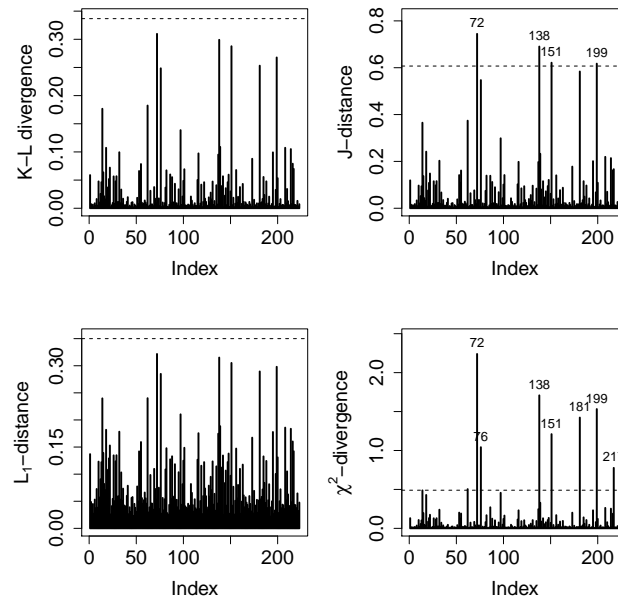
Figure 4: Estimates of ψ -divergence measures for Model 15. K-L = Kullback-Leibler.

Table 7: Possible influential observations are detected by four divergence measures

Observations	Sex	Duration	Intervention	No. of cigarettes	Relapse	Time interval	Zip
72	0	20	1	25	1	(3.159, 3.929)	55987
138	1	25	1	20	1	(2.998, 3.992)	55021
151	0	39	1	10	1	(0.923, 3.962)	55057
199	0	22	0	20	1	(3.885, 5.013)	55904

mapped in Figure 2. We note the posterior SD of the frailties U and V have approximated values. Both maps show that the central region cities have lower values and that some periphery cities have higher values.

Figure 4 presents the estimates of ψ -divergence measures, which were obtained by the posterior sample of the parameters of the model that are used to detect possible influential observations in the posterior distribution of the parameters of Model 15. It shows some individuals can be influential observations that can be detected by divergence measures. Here, we will analyze individuals 72, 138, 151, and 199 were detected by the J -distance and the χ^2 -divergence. Table 7 presents information on them, so that we can note that these four individuals had special interventions and did not consume high amounts of cigarettes per day, while the individuals 72, 138, and 151 had relapse, but the individual 199 did not. To reveal the impact of these possible influential observations on the parameter estimates and inferences, we removed such observations, refitting the models. We also calculated the relative variations (RV) for the posterior mean of the parameters. The RV is defined by $RV = (\hat{\vartheta}_{d,-I} - \hat{\vartheta}_d) / \hat{\vartheta}_d$, for all d , where I denotes a set of influential observations, d is the index of the parameters, $\hat{\vartheta}_{d,-I}$ denotes the posterior mean of $\vartheta_{d,-I}$, after the set of observations I was removed. In this case, we have $I = \{72, 138, 151, 199\}$.

The posterior summaries of the parameters for the readjust Model 15 and RV for the posterior mean of the parameters are presented in Table 8. We can note that only the values of RV for the

Table 8: Posterior summaries of the parameter of the Model 15 and relative variations adjusted for the smoking cessation data without detected individuals 72, 138, 151, and 199

Parameter	Survival model					Cure rate				
		Mean	SD	2.5%	97.5%		Mean	SD	2.5%	97.5%
Intercept						b_0 (-0.0028)	1.3697	0.5981	0.2041	2.5265
Sex (male = 0)	β_1	-0.2101 (-0.3455)	0.4326	-1.0848	0.6344	b_1	-0.5657 (0.1101)	0.3425	-1.2511	0.0902
SI/UC (UC = 0)	β_2	1.0138 (0.2031)	0.5071	0.0486	2.0473	b_2 (0.0491)	0.9023	0.4002	0.1406	1.7522
Cigarettes per day	β_3	-0.1048 (-0.0873)	0.0377	-0.1800	-0.0304	b_3 (-0.1602)	-0.0611	0.0257	-0.1221	-0.0183
Duration as smoker	β_4	-0.0308 (0.2518)	0.0361	-0.1073	0.0375	b_4	0.0184 (-0.0656)	0.0230	-0.0280	0.0629
α		2.6474 (0.0987)	0.3515	1.9792	3.3582					
a_1 (a_u)		0.9006 (0.0043)	0.0665	0.7388	0.9880					
a_2 (a_v)							0.9002 (0.0009)	0.0650	0.7374	0.9864
Λ_{11}		2.6749 (-0.0008)	0.6497	1.5769	4.0686					
Λ_{22}							2.5717 (-0.0010)	0.6364	1.5178	3.9480
Λ_{12}		0.0113 (-2.0812)	0.4709	-0.9257	0.9130					
Σ_{11}		0.4122 (0.0023)	0.1086	0.2515	0.6701					
Σ_{22}							0.4307 (0.0021)	0.1208	0.2575	0.7207
Σ_{12}		-0.0054 (-2.5601)	0.1839	-0.3637	0.3568					

SD = standard deviation; SI = smoking intervention group; UC = usual care group.

posterior means of the parameters Λ_{12} and Σ_{12} are more than one, but they still have posterior means near zero, and others parameters have posterior means near the obtained values for the completed data set. In this case, there are no inferential changes after removing the observations.

6. Conclusions

This paper described an approach to extend the proportional odds cure models to allow to spatial correlations by including spatial frailties in the interval-censored data setting. We used MCMC methods through Bayesian inference and the DIC for the model comparison. The results of the application show that the parametric cure model with frailties under the FA scheme have better fittings. A comparison of the proposed models with the model introduced by Carlin and Banerjee (2003) indicated that our model is more adequate. Moreover, the both models proposed are not sensitive to influence observations, which can be observed by the influence diagnostic analysis in the application. The interpretation of the covariates was easy due to the parametrization of the models considered in terms of the cure rate. The MCAR prior can also be used even if frailties do not present or do present low correlations. We assume M_{ij} as Geometric distributed; however, we believe that the methodology presented here may be theoretically extended by considering other distributions, such as Negative Binomial and Power Series.

Acknowledgements

The research is partially supported by the Brazilian organizations CAPES (via Science without Borders-PDSE), CNPq and FAPESP.

References

- Banerjee S and Carlin BP (2004). Parametric spatial cure rate models for interval-censored time-to-relapse data, *Biometrics*, **60**, 268–275.
- Bennett S (1983). Analysis of survival data by the proportional odds model, *Statistics in Medicine*, **2**, 273–277.
- Berkson J and Gage RP (1952). Survival curve for cancer patients following treatment, *Journal of the American Statistical Association*, **47**, 501–515.
- Besag J (1974). Spatial interaction and the statistical analysis of lattice systems, *Journal of the Royal Statistical Society, Series B (Methodological)*, **36**, 192–236.
- Cancho VG, Ortega EMM, and Paula GA (2010). On estimation and influence diagnostics for log-Birnbaum-Saunders Student-*t* regression models: full Bayesian analysis, *Journal of Statistical Planning and Inference*, **140**, 2486–2496.
- Carlin BP and Banerjee S (2003). Hierarchical multivariate CAR models for spatio-temporally correlated survival data, *Bayesian Statistics*, **7**, 45–63.
- Chen MH, Ibrahim JG, and Sinha D (1999). A new Bayesian model for survival data with a surviving fraction, *Journal of the American Statistical Association*, **94**, 909–919.
- Cook RD and Weisberg S (1982). *Residuals and Influence in Regression*, Chapman & Hall/CRC, FL.
- Collett D (1994). Modelling survival data. In D Collett (Ed), *Modelling Survival Data in Medical Research* (pp. 53–60), Springer, New York.
- Cooner F, Banerjee S, Carlin BP, and Sinha D (2007). Flexible cure rate modeling under latent activation schemes, *Journal of the American Statistical Association*, **102**, 560–572.
- Dey DK and Birmiwal LR (1994). Robust Bayesian analysis using divergence measures, *Statistics & Probability Letters*, **20**, 287–294.
- Finkelstein DM (1986). A proportional Hazards model for interval-censored failure time data, *Biometrics*, **42**, 845–854.
- Gamerman D and Lopes HF (2006). *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference* (2nd ed.), Chapman & Hall/CRC, Boca Raton.
- Gelfand AE and Vounatsou P (2003). Proper multivariate conditional autoregressive models for spatial data analysis, *Biostatistics*, **4**, 11–15.
- Geweke J (1992). Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In JM Bernardo, J Berger, AP Dawid, and JFM Smith (Eds), *Bayesian Statistics 4* (pp. 169–193), Clarendon Press, Oxford.
- Gu Y, Sinha D, and Banerjee S (2011). Analysis of cure rate survival data under proportional odds model, *Lifetime Data Analysis*, **17**, 123–134.
- Hu T and Xiang L (2013). Efficient estimation for semiparametric cure models with interval-censored data, *Journal of Multivariate Analysis*, **121**, 139–151.
- Ibrahim JG, Chen MH, and Sinha D (2001). *Bayesian Survival Analysis*, Springer, New York.
- Lindsey JC and Ryan LM (1998). Methods for interval-censored data, *Statistics in Medicine*, **17**, 219–238.
- Ma X and Xiang L (2013). Testing for the presence of a cure fraction in clustered interval-censored survival data, *Australian & New Zealand Journal of Statistics*, **55**, 173–190.

- Murray RP, Anthonisen NR, Connett JE, *et al.* (1998). Effects of multiple attempts to quit smoking and relapses to smoking on pulmonary function, *Journal of Clinical Epidemiology*, **51**, 1317–1326.
- Pan C, Cai B, Wang L, and Lin X (2014). Bayesian semiparametric model for spatially correlated interval-censored survival data, *Computational Statistics & Data Analysis*, **74**, 198–208.
- Peng F and Dey DK (1995). Bayesian analysis of outlier problems using divergence measures, *Canadian Journal of Statistics*, **23**, 199–213.
- Peto R (1973). Experimental survival curves for interval-censored data. *Journal of the Royal Statistical Society Series C (Applied Statistics)*, **22**, 86–91.
- Plummer M, Best N, Cowles K, and Vines K (2005). Output analysis and diagnostics for MCMC: R package version 0.10-3, Retrieved November 2, 2017, from: <http://cran.rproject.org>
- R Development Core Team (2010). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna.
- Rodrigues J, Cancho VG, de Castro M, and Louzada-Neto F (2009). On the unification of long-term survival models, *Statistics & Probability Letters*, **79**, 753–759.
- Rücker G and Messerer D (1988). Remission duration: an example of interval-censored observations, *Statistics in Medicine*, **7**, 1139–1145.
- Spiegelhalter DJ, Best NG, Carlin BP, and van der Linde A (2002). Bayesian measures of model complexity and fit, *Journal of the Royal Statistical Society B (Statistical Methodology)*, **64**, 583–639.
- Sun X and Chen C (2010). Comparison of Finkelstein's method with the conventional approach for interval-censored data analysis, *Statistics in Biopharmaceutical Research*, **2**, 97–108.
- Weiss R (1996). An approach to Bayesian sensitivity analysis, *Journal of the Royal Statistical Society Series B (Methodological)*, **58**, 739–750.
- Xiang L, Ma X, and Yau KK (2011). Mixture cure model with random effects for clustered interval-censored survival data, *Statistics in Medicine*, **30**, 995–1006.
- Yakovlev AY and Tsodikov AD (1996). *Stochastic Models of Tumor Latency and Their Biostatistical Applications*, World Scientific, Singapore.
- Yau KKW and Ng ASK (2001). Long-term survivor mixture model with random effects: application to a multi-centre clinical trial of carcinoma, *Statistics in Medicine*, **20**, 1591–1607.
- Zeng D, Yin G, and Ibrahim JG (2006). Semiparametric transformation models for survival data with a cure fraction, *Journal of the American Statistical Association*, **101**, 670–684.