

RESEARCH ARTICLE

A New lifetime model for multivariate survival data with a surviving fraction

Vicente G. Cancho^a, Francisco Louzada^{a*}, Dipak K. Dey^b and Gladys D.C. Barriga^c

^a*ICMC, Universidade de São Paulo, São Carlos, Brazil*

^b*Department of Statistics, University of Connecticut, Storrs, CT, USA*

^a*FEB, State Universidade de São Paulo, Bauru, Brazil*

(v4.2 released May 2008)

In this paper we propose a new lifetime model for multivariate survival data with a surviving fraction. We develop this model assuming that there are m types of unobservable competing risks, where each risk is related to a time of the occurrence of an event of interest. We explore the use of Markov chain Monte Carlo (MCMC) methods to develop a Bayesian analysis for the proposed model. We also perform a simulation study in order to analyze the frequentist coverage probabilities of credible interval derived from posteriors. Our modeling is illustrated through a real data set.

Keywords: Bayesian inference; competing risks; MCMC; Multivariate Survival Models; Cure Rate Models; Cured Fraction.

1. Introduction

Cure rate models play an important role in survival analysis, covering data where a proportion of subjects may remain susceptible to the event of interest. These models, also referred to as survival models with a surviving fraction or long-term survival models, have become very popular due to significant progress and advancements in treatment therapies leading to enhanced cure rates. The proportion of these ‘cured’ units is termed the cured fraction. The terminology is based on clinical studies, where the event of interest can be the death of a patient or a tumor recurrence (which can be attributed to metastasis-component tumor cells left active after an initial treatment), which can not occur due to the event of interest. However, nowadays, it is common to observe applications of cure rate models studies of many other kinds of events, such as criminal recidivism, divorce, child-bearing, unemployment and costumer churn, amongst others.

Two formulations of cure rate models stand out in the literature as being the prevailing approaches. Here we point out a distinguishing feature between them.

*Corresponding author: Francisco Louzada. ICMC-USP, 13566-590, São Carlos, SP, Brazil. Email: louzada@icmc.usp.br

Firstly, the standard mixture cure model [1, 2], where the number of competing causes of the event of interest is a binary random variable on $\{0, 1\}$, and recently, the promotion time cure model [37], where the number of competing causes follows a Poisson distribution. However, these assumptions may not be verified once the number of competing causes is an unobserved random variable (latent variable). In order to overcome this restriction several proposals appeared in the literature considering general distributions such as the geometric, negative binomial, power series, Conway-Maxwell-Poisson distributions, amongst other. The literature on cure rate models is by now vast and expanding rapidly. The books by Maller & Zhou [24] and Ibrahim *et al.* [19], as well as the articles by Chen *et al.* [8], Tsodikov *et al.* [36], Tournoud & Ecochard [35], Cancho *et al.* [4], Perdoná & Louzada-Neto [27], Kim *et al.* [21] and Cobre *et al.* [12] to name a few.

Although extensions of cure rate models were developed, limited attention has been paid to the research on multivariate cure rate models. In the frequentist framework, Chatterjee & Shih [7] proposed a marginal approach using bivariate copula models. Price & Manatunga [28] imposed frailty to account for correlation and conducted the maximum likelihood estimation under a parametric model assumption. Both methods were based on the mixture cure model. In the Bayesian paradigm, Chen *et al.* [9] generalized the work of [37] to multivariate failure time data by introducing a positive stable frailty and Louzada *et al.* [22] proposed bivariate long-term distribution based on the Farlie-Gumbel-Morgenstern copula model.

In this paper, a new multivariate cure rate survival model is developed under a scenario of latent competing causes. In the formulation we consider that there are m types of latent causes of failures, which can produce the correspondent event of interest, where these latent variables are modeled by a multivariate Poisson distribution [20]. In Section 2, we formulate the multivariate cure rate model. We explore the use of Markov chain Monte Carlo methods to develop a Bayesian analysis in Section 3. Simulation study is presented in the Section 4. An application to a real data set is given in Section 5. Finally, Section 6 concludes with some general remarks.

2. The Model

For an individual in the population, let N_k ($k = 1, \dots, m$) be the random variable that denotes the unobservable number of causes of type k of the event of interest for this individual. We assume that $\mathbf{N} = (N_1, N_2, \dots, N_m)$ follows a multivariate Poisson distribution with probability mass function

$$P[N_1 = n_1, \dots, N_m = n_m] = e^{-\{\sum_{i=1}^m \theta_i\}} \prod_{i=1}^m \frac{\theta_i^{n_i}}{n_i!} \sum_{i=0}^s \prod_{j=1}^m \binom{n_j}{i} i! \left(\frac{\theta_0}{\prod_{i=1}^m \theta_i} \right)^i, \quad (1)$$

where $n_j = 0, 1, \dots$, $\theta_j > 0$, $j = 0, 1, \dots, m$ and $s = \min\{n_1, \dots, n_m\}$. The above multivariate distribution allows for positive dependence between the two random variables. Marginally each random variable follows a Poisson distribution with

$E(N_j) = \theta_j + \theta_0$ and, $\text{Cov}(N_i, N_j) = \theta_0, i \neq j = 1, \dots, m$ and hence θ_0 is a measure of dependence between the two random variables. If $\theta_0 = 0$ then the variables are independent and the multivariate Poisson distribution reduces to the product of m independent Poisson distributions. For a comprehensive treatment of the multivariate Poisson distribution the reader can refer to Karlis [20]. It is to be noted that the \mathbf{N} -vector also indicates a multivariate frailty which was also considered by Caroni et al. [6]. Further, this model arrives as boundary case for the compound Poisson model. Several models of these types were studied in Price & Manatunga [28].

The time for the j th competing cause of type k to produce the event of interest is denoted by $Z_{kj}, k = 1, \dots, m, j = 1, 2, \dots$. Given $N_k = n_k$, the Z_{k1}, \dots, Z_{kn_k} are independent and identically distributed random variables with cumulative distribution function $F_k(\cdot) = 1 - S_k(\cdot)$. We also assumed that the latent variables Z_{1j}, \dots, Z_{kj} are independent. The observed times to event are defined by the random variables $Y_k = \min \{Z_{k0}, Z_{k1}, \dots, Z_{kn_k}\}$ with $P(Z_{k0} = \infty) = 1, k=1, 2, \dots, m$. Under this setup we can demonstrate, that the population survival function for $\mathbf{Y} = (Y_1, \dots, Y_m)$ is given by

$$S_{\text{pop}}(\mathbf{y}) = \exp \left\{ - \sum_{i=1}^m \theta_i (1 - S_i(y_i)) - \theta_0 \left(1 - \prod_{i=1}^m S_i(y_i) \right) \right\}. \tag{2}$$

The survival function $S_{\text{pop}}(\mathbf{y})$ in (2) is not a proper survival, that is, $\lim_{y_1, \dots, y_m \rightarrow \infty} S_{\text{pop}}(\mathbf{y}) = \exp \{ - \sum_{i=1}^m \theta_i \} > 0$ (the joint cure fraction). Note that when $\theta_0 = 0$ in (2), the joint survival function reduces to the product of m independent survival functions. From (2) the marginal survival functions are

$$S_{\text{pop}}(y_k) = \exp \{ -(\theta_k + \theta_0) F_k(y_k) \}, k = 1, \dots, m. \tag{3}$$

Equation (3) indicates that the marginal survival function has a cure rate structure with probability of cure $p_{0k} = e^{-\theta_k - \theta_0}$ for $Y_k, k = 1, \dots, m$. It is important to note in (3) that each marginal survival function has the structure of the promotion time cure model [8, 37]. In (3) that each marginal distribution has a proportional hazards structure as long as the covariates, only enter through θ_k and θ_0 . The marginal hazard function is given by, $(\theta_k + \theta_0) f_k(y_k)$ which satisfies the conditions for the proportional hazards model [14]. This is a desirable feature of the proposed model that leads to attractive theoretical properties.

Without loss of generality, considering the bivariate distribution of (Y_1, Y_2) , then joint survival function is given by

$$S_{\text{pop}}(y_1, y_2) = \exp \{ -\theta_1 (1 - S_1(y_1)) - \theta_2 (1 - S_2(y_2)) - \theta_0 (1 - S_1(y_1) S_2(y_2)) \}. \tag{4}$$

This distribution will be called as a Bivariate cure rate (BCR) model. The parameter θ_0 is a measure of association between (Y_1, Y_2) . As $\theta_0 \rightarrow 0$, this implies less association between (Y_1, Y_2) which can be seen from (4). Following Clayton [11] and Oakes [26], we can compute a local measure of association, denoted by

$\vartheta^*(Y_1, Y_2)$, as a function of θ_0 . This measure of association is defined as

$$\vartheta^*(Y_1, Y_2) = \frac{S_{\text{pop}}(y_1, y_2) \frac{\partial^2}{\partial y_1 \partial y_2} S_{\text{pop}}(y_1, y_2)}{\left(\frac{\partial S_{\text{pop}}(y_1, y_2)}{\partial y_1} \right) \left(\frac{\partial S_{\text{pop}}(y_1, y_2)}{\partial y_2} \right)}. \quad (5)$$

The measure in (5), has the interpretation as the ratio of the hazard rate of the conditional distribution of Y_1 (Y_2), given $Y_2 = y_2$ ($Y_1 = y_1$), to that of Y_1 (Y_2) given $Y_2 > y_2$ ($Y_1 > y_1$). For more discussion of (5), see Clayton [11] and Oakes [26]. For the bivariate cure rate model in (4), $\vartheta^*(y_1, y_2)$ is well defined and is given by

$$\vartheta^*(y_1, y_2) = 1 + \theta_0 \{ [\theta_1 + \theta_0 S_2(y_2)] [\theta_2 + \theta_0 S_1(y_1)] \}^{-1}. \quad (6)$$

We see that $\vartheta^*(y_1, y_2)$ in (6) increases in (y_1, y_2) . That is, the association between (y_1, y_2) is less when (y_1, y_2) are small and the association increases over time.

The following Theorem is a generalization of the proposed model in (2) and provides a natural extension of the univariate survival cure rate models.

Theorem 2.1: Let $\mathbf{N} = (N_1, N_2, \dots, N_m)$ be a random vector with probability generating function, $\varphi_{\mathbf{N}}(w_1, \dots, w_m)$ and random vector $\mathbf{Y} = (Y_1, \dots, Y_m)$ defined in (2). Then the joint survival function of \mathbf{Y} is given by

$$S_{\text{pop}}(\mathbf{y}) = \varphi_{\mathbf{N}}(S_1(y_1), \dots, S_m(y_m)). \quad (7)$$

Proof: Under the conditions of Theorem, we can write the joint survival function of \mathbf{Y} as

$$\begin{aligned} S_{\text{pop}}(\mathbf{y}) &= P[\mathbf{N} = \mathbf{0}] + \sum_{n_1, \dots, n_m=1}^{\infty} P[Z_{11} > y_1, \dots, Z_{1n_1} > y_1, \dots, Z_{m1} > y_m, \dots, Z_{mn_m} > y_m] \\ &\quad \times P[N_1 = n_1, \dots, N_m = n_m] \\ &= P[N_1 = 0, \dots, N_m = 0] \\ &\quad + \sum_{n_1, \dots, n_m=1}^{\infty} P[N_1 = n_1, \dots, N_m = n_m] S_1^{n_1}(y_1) \dots S_m^{n_m}(y_m) \\ &= \sum_{n_1, \dots, n_m=1}^{\infty} P[N_1 = n_1, \dots, N_m = n_m] S_1^{n_1}(y_1) \dots S_m^{n_m}(y_m) \\ &= \varphi_{\mathbf{N}}(S_1(y_1), \dots, S_m(y_m)). \end{aligned}$$

□

The last step comes from the definition of the probability generating function.

Again the survival function $S_{\text{pop}}(\mathbf{y})$ in (7) is not a proper survival, that is, $\lim_{y_1, \dots, y_m \rightarrow \infty} S_{\text{pop}}(\mathbf{y}) = \varphi_{\mathbf{N}}(0, \dots, 0) = P[N_1 = 0, \dots, N_m = 0] > 0$ (joint cure

rate). From (7) the marginal survival function is obtained as

$$S_{pop}(y_k) = \varphi_{N_k}(S_k(y_k)), \quad k = 1, \dots, m. \quad (8)$$

The marginal survival function (8) is the same as the one proposed by Tsodikov et al. [36] and Rodrigues et al. [30]. The marginal cure rate proportion is $S_k(\infty) = \varphi_{N_k}(0) = P[N_k = 0] > 0$. In the case that random variables N_1, \dots, N_m are independent, the survival function $S_{pop}(\mathbf{y})$ in (7) is given by

$$S_{pop}(\mathbf{y}) = \prod_{k=1}^m \varphi_{N_k}(S_k(y_k)).$$

The proposed model in (2) can be obtained from Theorem 2.1. To be more specific, the random vector $\mathbf{N} = (N_1, \dots, N_m)$ with multivariate Poisson given in (1) is defined as $N_k = D_k + D_0$, $k = 1, \dots, m$, where D_i , $i = 0, 1, \dots, m$ are independent univariate Poisson distributed random variables with parameters θ_i . The probability generating function, of \mathbf{N} is defined as

$$\begin{aligned} \varphi_{\mathbf{N}}(w_1, \dots, w_m) &= E \left[w_1^{N_1} w_2^{N_2} \dots w_m^{N_m} \right] = E \left[w_1^{D_1+D_0} \times w_2^{D_2+D_0} \times \dots \times w_m^{D_m+D_0} \right] \\ &= E \left[w_1^{D_1} \times w_2^{D_2} \times \dots \times w_m^{D_m} (w_1 \times \dots \times w_m)^{D_0} \right] \\ &= \varphi_{D_1}(w_1) \times \dots \times \varphi_{D_m}(w_m) \times \varphi_{D_0} \left(\prod_{i=1}^m w_i \right) \\ &= \exp \left\{ - \sum_{i=1}^m \theta_i (1 - w_i) - \theta_0 \left(1 - \prod_{i=1}^m w_i \right) \right\}. \end{aligned} \quad (9)$$

From Theorem 2.1 and (9), we obtain the model given in (2).

We can also obtain an extension of bivariate version of the univariate standard mixture cure rate model. In what follows, consider a bivariate Bernoulli random vector (N_1, N_2) , which takes values from $(0, 0)$, $(0, 1)$, $(1, 0)$ and $(1, 1)$ in the cartesian product space $\{0, 1\}^2 = \{0, 1\} \times \{0, 1\}$ with probability $\theta_{ij} = P[N_1 = i, N_2 = j]$, $i, j = 0, 1$, such that, $\sum_{i=0}^1 \sum_{j=0}^1 \theta_{ij} = 1$. Then the corresponding probability generating function is given by $\varphi_{\mathbf{N}}(w_1, w_2) = \theta_{00} + \theta_{10}w_1 + \theta_{01}w_2 + \theta_{11}w_1w_2$. From Theorem 2.1, we obtain the following results related to the standard mixture cure rate model

$$S_{pop}(y_1, y_2) = \theta_{00} + \theta_{10}S_1(y_1) + \theta_{01}S_2(y_2) + \theta_{11}S_1(y_1)S_2(y_2).$$

Thus the joint cure fraction is $S_{pop}(\infty, \infty) = \theta_{00}$. The marginal survival functions are given by $S_{pop}(y_1) = \theta_{00} + \theta_{01} + (\theta_{10} + \theta_{11})S_1(y_1)$ and $S_{pop}(y_2) = \theta_{00} + \theta_{10} + (\theta_{01} + \theta_{11})S_2(y_2)$.

Now, if we consider a random vector (N_1, N_2) which follows a bivariate geometric

distribution with probability mass function

$$P[N_1 = n_1, N_2 = n_2] = \binom{n_1 + n_2}{n_1} (1 - \theta_1 - \theta_2) \theta_1^{n_1} \theta_2^{n_2},$$

where $n_j = 0, 1, \dots, 0 < \theta_j < 1, j=1,2$ and $0 < 1 - \theta_1 - \theta_2 < 1$. From Theorem 2.1, we obtain a new bivariate cure rate model as

$$S_{\text{pop}}(y_1, y_2) = \frac{1 - \theta_1 - \theta_2}{1 - \theta_1 S_1(y_1) - \theta_2 S_2(y_2)},$$

with the joint cure fraction $S_{\text{pop}}(\infty, \infty) = 1 - \theta_1 - \theta_2$. The marginal survival functions are given by $S_{\text{pop}}(y_1) = \frac{1 - \theta_1 - \theta_2}{1 - \theta_1 - \theta_1 S_1(y_1)}$ and $S_{\text{pop}}(y_2) = \frac{1 - \theta_1 - \theta_2}{1 - \theta_1 - \theta_2 S_2(y_2)}$ respectively. This marginal model is similar to the one proposed by Gu et al. [17].

In the next Section, we develop a Bayesian procedure for inference on bivariate survival cure rate model given in (4). Hereafter we assume a Weibull distribution for the unobserved time Z with $F_k(z|\gamma_k) = 1 - S_k(z|\gamma_k) = 1 - \exp(-z^{\gamma_{k1}} e^{\gamma_{k2}})$ and $f_k(z|\gamma_k) = \gamma_{k1} z^{\gamma_{k1}-1} \exp(\gamma_{k2} - z^{\gamma_{k1}} e^{\gamma_{k2}})$, for $z > 0, \gamma_{k1} > 0, \gamma_{k2} \in \mathbb{R}$ and $\gamma_k = (\gamma_{k1}, \gamma_{k2})^\top, k = 1, 2$. However, in principle, other survival models may be undertaken, taking their specificities into account.

3. Inference

Let us consider the situation when the failure times (Y_1, Y_2) in Section 2 are not completely observed and are subject to right censoring. Let C_{ki} denote the censoring time of k component, $k = 1, 2$. Suppose that $(Y_{1i}; Y_{2i})$ and $(C_{1i}; C_{2i})$ are independent. For each individual i , observed quantities are represented by the random variables $t_{ki} = \min\{Y_{ki}, C_{ki}\}$ and $\delta_{ki} = I(t_{ki} = Y_{ki})$, which denotes a censorship indicator, $k = 1, 2, i = 1, \dots, n$.

Let \mathbf{x}_{ki} denote the vectors of covariates for the i^{th} individual. Extending our model, we propose to relate the parameters θ_{1i} and θ_{2i} of the bivariate Poisson distribution to the covariates by the logarithmic link

$$\theta_{1i} = \exp(\mathbf{x}_{1i}^\top \boldsymbol{\beta}_1) \quad \text{and} \quad \theta_{2i} = \exp(\mathbf{x}_{2i}^\top \boldsymbol{\beta}_2), \quad k = 1, 2, \tag{10}$$

where $\boldsymbol{\beta}_k = (\beta_{k1}, \dots, \beta_{kp_k})^\top$ is the vector of regression coefficients associated with the covariates \mathbf{x}_{ki} . Now with the expression (10) we can express the likelihood of $\boldsymbol{\vartheta} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \gamma_1, \gamma_2, \theta_0)$ under non-informative censoring as,

$$\begin{aligned} L(\boldsymbol{\vartheta}|\mathcal{D}) &= \prod_{i=1}^n S_{\text{pop}}(t_{1i}, t_{2i}) \\ &\times \prod_{k=1}^2 [f_k(t_{ki}|\gamma_k)]^{\delta_{ki}} [\theta_0 + (\theta_{2i} + \theta_0 S_1(t_{1i}|\gamma_1))(\theta_{1i} + \theta_0 S_2(t_{2i}|\gamma_2))]^{\delta_{1i} \delta_{2i}} \\ &\times (\theta_{1i} + \theta_0 S_2(t_{2i}|\gamma_2))^{\delta_{1i}(1-\delta_{2i})} (\theta_{2i} + \theta_0 S_1(t_{1i}|\gamma_1))^{\delta_{2i}(1-\delta_{1i})}, \end{aligned} \tag{11}$$

where $\mathcal{D} = (\mathbf{t}, \boldsymbol{\delta}, \mathbf{x})$, $\mathbf{t} = (\mathbf{t}_1, \mathbf{t}_2)$, $\mathbf{t}_k = (t_{k1}, \dots, t_{kn})^\top$, $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2)$, $\mathbf{x}_k = (\mathbf{x}_{k1}, \dots, \mathbf{x}_{kn})^\top$, $\boldsymbol{\delta} = (\boldsymbol{\delta}_1, \boldsymbol{\delta}_2)$ and $\boldsymbol{\delta}_k = (\delta_{k1}, \dots, \delta_{kn})^\top$, $k = 1, 2$, whereas $S_{\text{pop}}(t_1, t_2)$ is survival function given in (4) and $f_k(t_{ki}|\gamma_k)$ and $S_k(t_{ki}|\gamma_k)$, $k = 1, 2$ are density and survival functions of the Weibull distribution.

3.1. Prior and Posterior

Now, some inferential tools are investigated under a Bayesian viewpoint. The normal distribution and gamma distribution with a as shape and b as scale (and mean a/b) are denoted by $N(\mu, \sigma^2)$ and $G(a, b)$. In this context we assume that $\boldsymbol{\beta}_k$, $k = 1, 2$, γ_{k1} , γ_{k2} and θ_0 are a priori independent, that is,

$$\pi(\boldsymbol{\vartheta}) = \prod_{k=1}^2 \pi(\boldsymbol{\beta}_k) \pi(\gamma_{k1}) \pi(\gamma_{k2}) \pi(\theta_0), \quad (12)$$

where $\boldsymbol{\beta}_k \sim N_{p_k}(\mathbf{0}, \Sigma_{k0})$, $\gamma_{k1} \sim G(a_{k0}, b_{k0})$, $\gamma_{k2} \sim N_1(0, \sigma_{\gamma_{k2}}^2)$ $k = 1, 2$ and $\theta_0 \sim G(a_0, b_0)$. Here all the hyper-parameters are specified in order to express non-informative priors.

Combining the likelihood function (11) and the prior distribution in (12), the joint posterior distribution for $\boldsymbol{\vartheta}$ is obtained as $\pi(\boldsymbol{\vartheta}|\mathcal{D}) \propto L(\boldsymbol{\vartheta}|\mathcal{D}) \prod_{k=1}^2 \pi(\boldsymbol{\beta}_k) \pi(\gamma_{k1}) \pi(\gamma_{k2}) \phi(\theta_0)$. This joint posterior density is analytically intractable. Thus the computational problem can be easily handled by using MCMC methods for sampling from the posterior distribution. From (10), (11) and (12), we write the full conditional distributions of $\boldsymbol{\beta}_k$, $\boldsymbol{\gamma}_k = (\gamma_{1k}, \gamma_{2k})$ and θ_0 as

$$\pi(\boldsymbol{\beta}_k | \cdot) \propto \exp\left[-\sum_{i=1}^n \theta_{ki} F_k(t_{ki}|\gamma_k)\right] \prod_{i=1}^n \Delta_i \Lambda_{ki} \pi(\boldsymbol{\beta}_k), \quad k = 1, 2, \quad (13)$$

$$\pi(\boldsymbol{\gamma}_1 | \cdot) \propto \exp\left[-\sum_{i=1}^n (\theta_{1i} F_1(t_{1i}|\gamma_1) - \theta_0 S_1(t_{1i}|\gamma_1) S_2(t_{2i}|\gamma_2))\right] \prod_{i=1}^n f_1(t_{1i}|\gamma_1) \Delta_i \Lambda_{2i} \pi(\boldsymbol{\gamma}_1), \quad (14)$$

$$\pi(\boldsymbol{\gamma}_2 | \cdot) \propto \exp\left[-\sum_{i=1}^n (\theta_{2i} F_2(t_{2i}|\gamma_2) - \theta_0 S_1(t_{1i}|\gamma_1) S_2(t_{2i}|\gamma_2))\right] \prod_{i=1}^n f_2(t_{2i}|\gamma_2) \Delta_i \Lambda_{1i} \pi(\boldsymbol{\gamma}_2), \quad (15)$$

and

$$\pi(\theta_0 | \cdot) \propto \exp\left[-\theta_0 S_1(t_{1i}|\gamma_1) S_2(t_{2i}|\gamma_2)\right] \prod_{i=1}^n \Lambda_{1i} \Delta_i \Lambda_{2i} \pi(\theta_0), \quad (16)$$

where $\Delta_i = [\theta_0 + (\theta_{2i} + \theta_0 S_1(t_{1i}|\gamma_1))(\theta_{1i} + \theta_0 S_2(t_{2i}|\gamma_2))]^{\delta_{1i}\delta_{2i}}$, $\Lambda_{1i} = (\theta_{1i} + \theta_0 S_2(t_{2i}|\gamma_2))^{\delta_{1i}(1-\delta_{2i})}$ and $\Lambda_{2i} = (\theta_{2i} + \theta_0 S_1(t_{1i}|\gamma_1))^{\delta_{2i}(1-\delta_{1i})}$.

Clearly the conditional density functions in (13)-(16) do not belong to any known parametric density family. In order to generate our samples we then implement a Metropolis-Hasting algorithm within Gibbs iterations [10]. For example, to implement the Metropolis-Hastings algorithm for the parameter θ_0 , we consider a target distribution $g_{\theta_0}(\theta_0) = \pi(\theta_0|\cdot)$, and under given model, $\theta_0 > 0$ we consider the transformation $\theta_0 = \exp(\eta)$, where, $-\infty < \eta < \infty$. Then, $g_{\eta}(\eta) = g_{\theta_0}(\eta)e^{\eta}$.

Instead of directly sampling θ_0 , we generate η by choosing a normal proposal $N(\hat{\eta}, \sigma_{\hat{\eta}}^2)$ where $\hat{\eta}$ is the maximizer of the logarithm of $g_{\eta}(\eta)$ and $\sigma_{\hat{\eta}}^2$ is the minus of the inverse of the second derivative of logarithm of $g_{\eta}(\eta)$ evaluated in $\eta = \hat{\eta}$. The algorithm to generate η operates according to: (1) let η be the current value; (2) generate a point η^* from $N(\hat{\eta}, \sigma_{\hat{\eta}}^2)$; (3) a move from η to η^* is made with probability $\min\{1, [g_{\eta}(\eta^*)\phi(\frac{\eta-\hat{\eta}}{\sigma_{\hat{\eta}}})]/[g_{\eta}(\eta)\phi(\frac{\eta^*-\hat{\eta}}{\sigma_{\hat{\eta}}})]\}$, where $\phi(\cdot)$ is the standard normal probability density function.

After we sample η , we obtain $\theta_0 = e^{\eta}$. To implement the Metropolis-Hastings algorithm for the parameter β_k and γ_k , $k = 1, 2$, we proceed in the same way as for the parameter θ_0 , but considering as target distribution the respective conditional density functions are given in (13)-(15).

3.2. Model comparison criteria

There exist a variety of methodologies to compare several competing models for a given data set and to select the one that best fits the data. One of the most used in applied Bayesian research is the deviance information criterion (*DIC*) proposed by Spiegelhalter et al. [34], the expected Akaike information criterion (EAIC)-[3], and the expected Bayesian (or Schwarz) information criterion (EBIC)-[5] can be used. These criteria are based on the posterior mean of the deviance, $d(\boldsymbol{\vartheta})$ which can be approximated from the MCMC output by

$$\bar{d} = \sum_{q=1}^Q d(\boldsymbol{\vartheta}_q)/Q,$$

where the index q indicates the q -th realization of a total of Q realization and $d(\boldsymbol{\vartheta}) = -2 \sum_{i=1}^n \log [g(t_{1i}, t_{2i}|\boldsymbol{\vartheta})]$, where $g(\cdot)$ is the pdf corresponding to our model. For observed data we have that $g(t_{1i}, t_{2i}|\boldsymbol{\vartheta})$ is i -th component of likelihood function.

The DIC criterion can be estimated using the MCMC output by $\widehat{DIC} = \bar{d} + \hat{\rho}_d = 2\bar{d} - \hat{d}$, with ρ_D is the effective number of parameters, which is defined as $E\{d(\boldsymbol{\vartheta})\} - d\{E(\boldsymbol{\vartheta})\}$, where $d\{E(\boldsymbol{\vartheta})\}$ is the deviance evaluated at the posterior mean and estimated as

$$\widehat{D} = d \left(\frac{1}{Q} \sum_{q=1}^Q \beta_1^{(q)}, \frac{1}{Q} \sum_{q=1}^Q \beta_2^{(q)}, \frac{1}{Q} \sum_{q=1}^Q \gamma_1^{(q)}, \frac{1}{Q} \sum_{q=1}^Q \gamma_2^{(q)}, \frac{1}{Q} \sum_{q=1}^Q \theta_0^{(q)} \right).$$

Similarly, the EAIC and EBIC criteria can be estimated by means of $\widehat{EAIC} = \bar{d} + 2\#(\boldsymbol{\vartheta})$ and $\widehat{EBIC} = \bar{d} + \#(\boldsymbol{\vartheta}) \log(n)$, where $\#(\boldsymbol{\vartheta})$ is the number of model parameters. The EBIC criterion used in this work is the same as that in Sahu & Dey [32]. Comparing alternative models, the preferred model is the one with the smallest criterion values.

Another criterion which is one of the most used in applied works is derived from the conditional predictive ordinate (*CPO*) statistic. For a detailed discussion on the *CPO* statistic and its applications to model selection, see [16] and Geisser & Eddy [15]. Let \mathcal{D} denote the full data and $\mathcal{D}^{(-i)}$ denote the data with the i -th observation deleted. We denote the posterior density of $\boldsymbol{\vartheta}$ given $\mathcal{D}^{(-i)}$ by $\pi(\boldsymbol{\vartheta}|\mathcal{D}^{(-i)})$, $i = 1, \dots, n$. For the i -th observation, CPO_i can be written as

$$CPO_i = \int_{\boldsymbol{\vartheta} \in \Theta} g(t_{1i}, t_{2i}|\boldsymbol{\vartheta}) \pi(\boldsymbol{\vartheta}|\mathcal{D}^{(-i)}) d\boldsymbol{\vartheta} = \left\{ \int_{\boldsymbol{\vartheta}} \frac{\pi(\boldsymbol{\vartheta}|\mathcal{D})}{g(t_{1i}, t_{2i}|\boldsymbol{\vartheta})} d\boldsymbol{\vartheta} \right\}^{-1}. \quad (17)$$

For the proposed model a closed form of the CPO_i is not available. However, a Monte Carlo estimate of CPO_i can be obtained by using a single MCMC sample from the posterior distribution $\pi(\boldsymbol{\vartheta}|\mathcal{D})$. Let $\boldsymbol{\vartheta}^{(1)}, \dots, \boldsymbol{\vartheta}^{(Q)}$ be a sample of size Q of $\pi(\boldsymbol{\vartheta}|\mathcal{D})$ after the burn-in. A Monte Carlo approximation of CPO_i [19] is given by

$$\widehat{CPO}_i = \left\{ \frac{1}{Q} \sum_{q=1}^Q \frac{1}{g(t_{1i}, t_{2i}|\boldsymbol{\vartheta}^{(q)})} \right\}^{-1}.$$

For model comparison we use the log pseudo marginal likelihood (LPML) defined by $LPML = \sum_{i=1}^n \log(\widehat{CPO}_i)$. The larger is the value of $LPML$, the better is the fit of the model.

4. Simulation study

To evaluate the performance of the Bayesian estimation approach for the bivariate survival cure rate model, we carry on a simulation study. In this study we consider the bivariate cure rate model with the Weibull distribution for the event time (Z_{kj} , $k = 1, 2$, $j = 1, 2, \dots$), with parameter, $\gamma_{k1} = 1.4$ and $\gamma_{k2} = 2.0$. For each individual i , $i = 1, \dots, n$, the number of competing risks of the event of interest for this individual (N_1, N_2) is generated from the bivariate distribution with parameter $\theta_0 = 0.5$, $\theta_{1i} = \exp(\beta_{10} + \beta_{11}x_i)$ and $\theta_{2i} = \exp(\beta_{20} + \beta_{21}x_i)$, where $\beta_{10} = -1.6$, $\beta_{11} = 1.3$, $\beta_{20} = -1.5$, $\beta_{21} = 1.1$ and the covariates x_i are generated from a Bernoulli distribution with parameter 0.5. The censoring times C_{ki} are sampled from the uniform distribution on the interval $(0, \tau_k)$, where τ_k is set in order to control the proportion of censored observations. In this study the proportion of censored observations was on an average approximately equal to 50%.

We consider a great proportion of censored observations since it is common we find such amount of censoring in practice. For instance, a great censoring proportion

Table 1.: Simulation results of the posterior summaries over 5,000 replications.

n		Parameter								
		β_{10}	β_{11}	γ_{11}	γ_{12}	β_{20}	β_{21}	γ_{21}	γ_{12}	θ_0
200	AE	-1.489	1.201	2.018	1.409	-1.482	1.060	2.030	1.410	0.489
	SD	0.348	0.381	0.155	0.161	0.368	0.345	0.154	0.153	0.068
	RMSE	0.368	0.401	0.156	0.163	0.373	0.361	0.157	0.1549	0.072
	PC	0.942	0.943	0.949	0.942	0.953	0.950	0.948	0.949	0.945
400	AE	-1.592	1.283	2.008	1.391	-1.501	1.010	2.021	1.403	0.502
	SD	0.268	0.252	0.133	0.130	0.368	0.256	0.143	0.139	0.034
	RMSE	0.269	0.254	0.137	0.131	0.373	0.258	0.145	0.141	0.034
	PC	0.951	0.948	0.952	0.948	0.948	0.951	0.949	0.951	0.952

is observed in the data set obtained from the diabetic retinopathy study [18] of time of blindness in each eye of diabetic patients with diabetic retinopathy, which will be considered further in Section 5.

The sample sizes are taken as $n = 200$ and 400 . For each simulated data set, the posterior summaries and 95% HPD intervals of the model parameters were obtained. A amount of 80,000 MCMC posterior samples are generated for each parameter, from which 20,000 iterations, regarded as the burn-in time, are eliminated for obtaining a sample of size 60,000. The autocorrelation of these sampled values are reduced by taking a spacing of size 10, thus resulting in 6,000 samples. For each sample, we calculate the estimates of each parameter and then by using these estimates we obtain the average (AE), standard deviation (SD), the root mean square error (RMSE) and the coverage probability (PC). The results are all presented in Table 1, in mean values over 5,000 Monte Carlo repetitions. We observe that the RMSE decreases as sample size increases. Also, we observe that the difference between the AE's and the parameter values are quite small which suggests that the estimates possess very little bias. Finally, we note that the simulated coverage probabilities are quite close to the nominal level.

5. Application

In this Section we fit the proposed model on a data set obtained from the diabetic retinopathy study [18] of time of blindness in each eye of the 197 diabetic patients with diabetic Retinopathy. One eye of each patients was randomly selected for treatment (the effectiveness of laser photo coagulation in delaying the onsets of blindness) and other eye was observed without treatment. A binary age covariate (0 for juvenile and 1 for adult) is available. The first component of the bivariate survival time is the time of the blindness on the treated eye (Y_1) and the second component is the similar time for the untreated eye (Y_2). One notable feature of the data set is that 159 patients experienced some form of censoring. For all these patients, censoring time is the same. For 80 patients, the censoring times were greater than or equal to the failure times of the untreated eyes. The reverse happened for only 16 patients. The remaining 38 patients experienced failure before

the censoring time in both eyes of which simultaneous failures were observed for 6 patients.

Huster et al. [18] analyzed the data using a frequentist approach with the Clayton-Oakes model and Weibull marginal distributions. Manatunga & Oakes [25] considered random effect models and Sahu & Dey [32] considered exponential and Weibull bivariate distributions with a Bayesian approach. and Romeo et al. [31] presented the Bayesian counterpart to the frequentist work of Shih & Louis [33], based on Archimedean copulas with parametric estimation for the marginals. All proposals do not consider the possibility that patients who may not remain susceptible to the event of interest. Recently Louzada et al. [22] analyzed the data considering bivariate long-term distribution based on the Farlie-Gumbel-Morgenstern copula model. The objective here is to see the effect of age in the surviving of patients, and to see whether there is any association between treated and untreated time to event and to investigate whether there is any difference in the proportion of patients of non-blindness (cure fraction) between eyes treated and untreated.

Then, the bivariate cure rate (BCR) model proposed in (4) is fitted to the data. For model fitting the following independent priors are adopted in the Bayesian computations $\beta_{kj} \sim N(0, 10^4)$, $k = 1, 2$ and $j = 0, 1$, $\gamma_{k1} \sim G(1, 0.01)$ and $\gamma_{k2} \sim N(0, 10^4)$. Thus, our choice is to assume weakly but informative prior. Since our prior is still informative, the posterior is always proper. 80,000 MCMC posterior samples are generated for each parameter, from which the 20,000 iterations are eliminated for obtaining 60,000 samples. The autocorrelation of these sampled values are reduced by taking a spacing of size 10 thus resulting in 6,000 samples. Figure 1 shows the behavior of the chains of the estimated parameters.

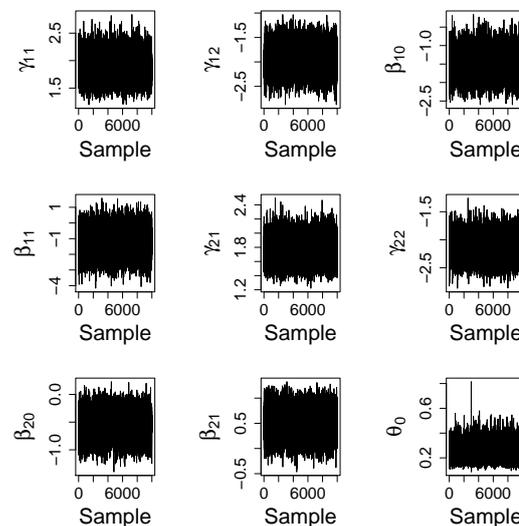


Figure 1.: Sequence plots of the chains

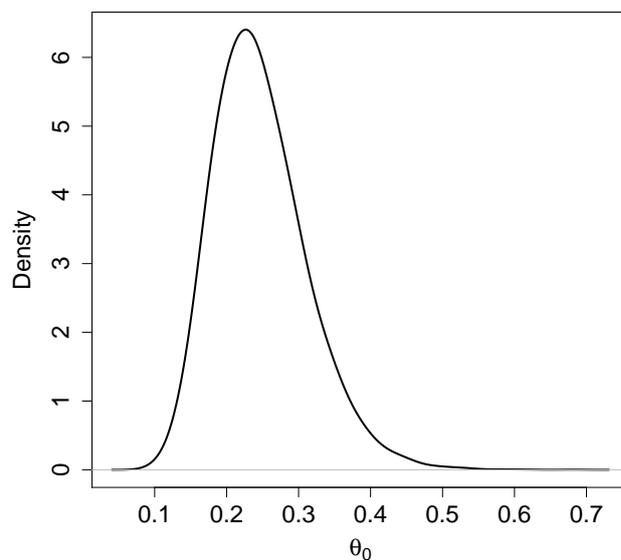
Table 2 gives posterior estimates of $\beta = (\beta_1, \beta_2)$, (γ_1, γ_2) , and θ_0 , where $\beta_k = (\beta_{k0}, \beta_{k1})$, and $\gamma_k = (\gamma_{1k}, \gamma_{2k})$, $k=1, 2$. We see from Table 2 that all of the highest posterior density (HPD) intervals for the regression coefficients of the covariates contain 0, this implies that the age do not have a significant effect on the surviving

Table 2.: Posterior summaries of the parameters for the bivariate cure rate model for the diabetic retinopathy data set.

Time to blindness	Parameter	Mean	Standard Deviation	95%HPD	
				Lower	Upper
Treated eye	γ_{11}	1.869	0.2217	1.473	2.335
	γ_{12}	-1.934	0.2633	-2.440	-1.420
	β_{10}	-1.449	0.347	-2.129	-0.776
	β_{11}	-1.216	0.8207	-2.854	0.380
Untreated eye	γ_{21}	1.746	0.165	1.443	2.101
	γ_{22}	-2.112	0.2113	-2.529	-1.699
	β_{20}	-0.563	0.221	-0.997	-0.129
	β_{21}	0.488	0.251	-0.012	0.976
	θ_0	0.247	0.064	0.145	0.393

of patients. Also, from Table 2, we see that the posterior mean of θ_0 is 0.247, with a 95% HPD interval of (0.145 , 0.393) which indicates significant association between Y_1 and Y_2 . A plot of the marginal posterior distribution of θ_0 is given in Figure 1. We see in this figure that the posterior distribution of θ_0 appears to be slightly skewed.

To evaluate the robustness of the model with regard to the choice of the hyperparameters of the prior distributions, a small sensitivity study is carried out with larger standard deviations for the prior distributions. The posterior summaries of the parameters do not display much difference and do not alter the results presented in Table 2.

Figure 2.: The marginal posterior density of θ_0 .

We compare the fitting of the BCR model with bivariate long-term distribution based on the Farlie-Gumbel-Morgenstern (FGM) copula model [22] with Weibull mixture marginal distributions by considering DIC , $EAIC$, $EBIC$ and $LPLM$ criteria. These information criteria furnish the values given in the Table 3. According to all the criteria, we have indication in favor of the BCR model.

Table 3.: Bayesian criteria for the fitted models.

Model	Criterion			
	LPLM	DIC	EAIC	EBIC
FGM	-689.32	1589.67	1597.40	1626.95
BCR	-469.12	932.87	940.92	970.47

Romeo *et al.* [31] modeled this data set using Archimedean copulas with Weibull marginal distribution. The estimated LPLM and BIC measures for Clayton copula were respectively, -643.402 and 1308.02 . Also, Sahu & Dey [32] considered a bivariate Weibull model, the estimated BIC measure was 1703.03 . Based on Bayesian criteria, there is a clear positive evidence in favor of BCR model, indicating that the model can be seen as a competitor to the bivariate survival with cure fraction models.

We turn our attention to the role of the covariates on the cured fraction p_0 . Table 4 shows the posterior summaries of the cured fraction stratified by group (juvenile and adult). A plot of the marginal posterior distribution of θ_0 is given in Figure 2. The plot and the table both reveals that the mean cure fraction for old patients with treated eye is higher than that of juvenile patients. In the case of juvenile patients with untreated eye the mean cure fraction is higher than that of the old patients. For both age groups, in the case of patients with treated eye the mean cure fraction is greater than that of the patients with untreated eye.

Table 4.: Posterior summaries of the cured fraction stratified by juvenile and adult patients.

Eye	Patient	Mean	Standard Deviation	95%HPD	
				Lower	Upper
Treated	Juvenile	0.611	0.053	0.498	0.704
	Adult	0.710	0.063	0.565	0.805
Untreated	Juvenile	0.430	0.059	0.320	0.551
	Adult	0.309	0.060	0.195	0.428

Figure 4 exhibits the Kaplan–Meier estimates of the survival function together with the Bayesian estimates of the marginal survival function based on the multivariate cure rate model. The covariates are not used in constructing plots. We see that the two curves in both plots on the left and right panel are nearly similar and appear to plateau after approximately 6 years of follow-up. Figure 5 shows a

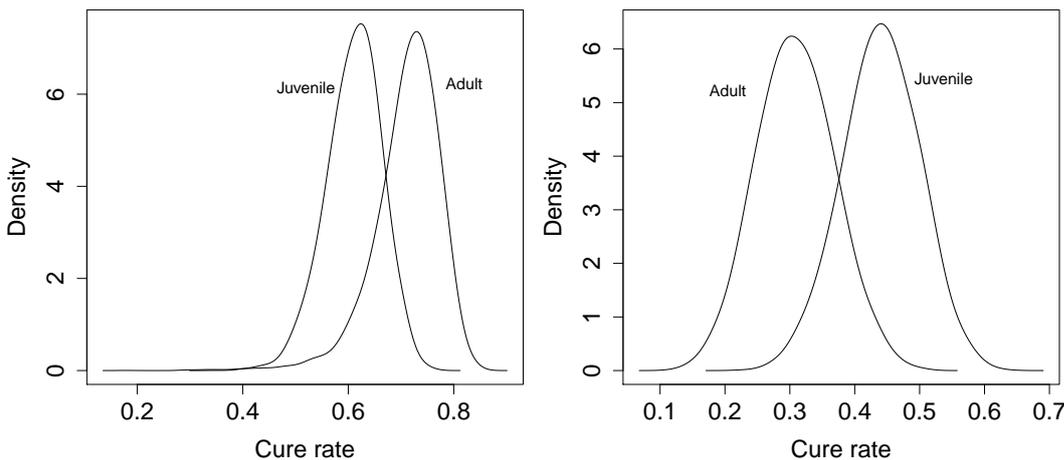


Figure 3.: Posterior density estimates for the cure rates corresponding to treated eye (left panel) and untreated eye (right panel) patients.

three-dimensional plot of the posterior mean survival surface. We observe in this plot how the survival curve plateaus for each failure time variable.

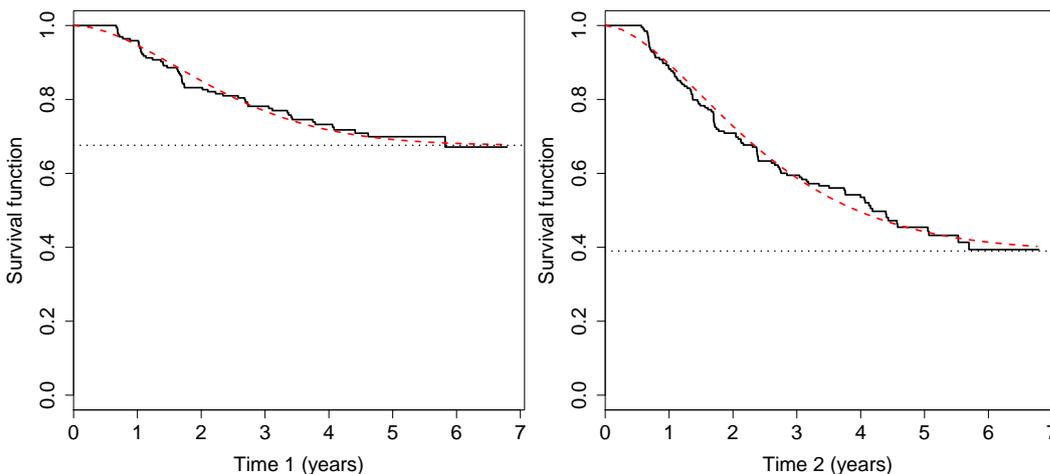


Figure 4.: Kaplan–Meier curve together with Bayesian estimates of the survival for patients with treated eye (left panel) and untreated eye (right panel).

6. Conclusions

In this paper we proposed a new multivariate survival model with cure rate and have examined some of its properties. This model is useful for jointly modeling any type of time to event data with a surviving fraction. We also develop a multivariate extension of the univariate survival cure rate model proposed by Tsodikov et al.

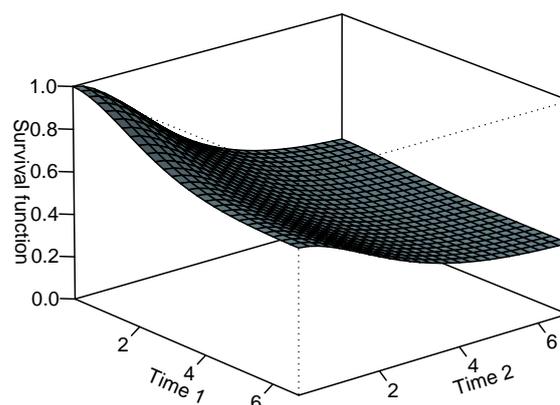


Figure 5.: The bivariate posterior survival surface.

[36] and Rodrigues *et al.* [30]. In the application to the diabetic retinopathy data set we discovered that the Bivariate cure rate model provides the best fit.

The model considered in this article can be fitted using standard available software [29], which makes the approach quite powerful and accessible to practitioners in the field.

Future developments of our work may involve other parametric or semi-parametric extensions of the multivariate lifetime distribution under the proposed setup, as well as, the relaxation of assumption that $Y_k = \min \{Z_{k0}, Z_{k1}, \dots, Z_{kN_k}\}$. Indeed, we may consider the modeling of its counterpart, $Y_k = \max \{Z_{k0}, Z_{k1}, \dots, Z_{kN_k}\}$, corresponding to a complementary risk scenarios as discussed by [23]. Moreover, following Cooner *et al.* [13], we envisage a generalization of our framework by assuming Y_k as random, such as we may scan all possible Y_k values from the first to last order statistics.

References

- [1] Berkson, J. & Gage, R. P. (1952). Survival curve for cancer patients following treatment. *Journal of the American Statistical Association*, **47**, 501–515.
- [2] Boag, J. W. (1949). Maximum likelihood estimates of the proportion of patients cured by cancer therapy. *Journal of the Royal Statistical Society B*, **11**, 15–53.
- [3] Brooks, S. P. (2002). Discussion on the paper by Spiegelhalter, Best, Carlin, and van der Linde (2002). *Journal of the Royal Statistical Society B*, **64**, 616–618.
- [4] Cancho, V., Rodrigues, J. & de Castro, M. (2011). A flexible model for survival data with a cure rate: a Bayesian approach. *Journal of Applied Statistics*, **38**(1), 57–70.
- [5] Carlin, B. P. & Louis, T. A. (2001). *Bayes and Empirical Bayes Methods for Data Analysis*. Chapman & Hall/CRC, Boca Raton, second edition.
- [6] Caroni, C., Crowder, M. & Kimber, A. (2010). Proportional hazards models with discrete frailty. *Lifetime data analysis*, **16**(3), 374–384.
- [7] Chatterjee, N. & Shih, J. (2001). A bivariate cure-mixture approach for modeling familial association in diseases. *Biometrics*, **57**(3), 779–786.

- [8] Chen, M.-H., Ibrahim, J. G. & Sinha, D. (1999). A new Bayesian model for survival data with a surviving fraction. *Journal of the American Statistical Association*, **94**, 909–919.
- [9] Chen, M.-H., Ibrahim, J. G. & Sinha, D. (2002). Bayesian inference for multivariate survival data with a cure fraction. *Journal of Multivariate Analysis*, **80**(1), 101–126.
- [10] Chib, S. & Greenberg, E. (1995). Hierarchical analysis of survival models with extensions to correlated serial errors and time-varying parameter models. *Journal of Econometrics*, **68**(2), 339–360.
- [11] Clayton, D. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*, **65**(1), 141–151.
- [12] Cobre, J., Perdoná, G. S. C., Peria, F. M. & Louzada, F. (2013). A mechanistic breast cancer survival modelling through the axillary lymph node chain. *Statistics in Medicine*, **32**, 1536–1546.
- [13] Cooner, F., Banerjee, S., Carlin, B. P. & Sinha, D. (2007). Flexible cure rate modeling under latent activation schemes. *Journal of the American Statistical Association*, **102**, 560–572.
- [14] Cox, D. & Oakes, D. (1984). *Analysis of Survival Data*. Chapman & Hall, London.
- [15] Geisser, S. & Eddy, W. (1979). A predictive approach to model selection. *Journal of the American Statistical Association*, pages 153–160.
- [16] Gelfand, A., Dey, D. & Chang, H. (1992). Model determination using predictive distributions with implementation via sampling based methods (with discussion). *Bayesian statistics 4*. eds: J. Bernardo et al. Bernardo, J.M., Berger, J.O., Dawid, A.P., Smith, A.F.M. (editors). Oxford University Press, **1**(4), 7–167.
- [17] Gu, Y., Sinha, D. & Banerjee, S. (2011). Analysis of cure rate survival data under proportional odds model. *Lifetime data analysis*, **17**(1), 123–134.
- [18] Huster, W., Brookmeyer, R. & Self, S. (1989). Modelling paired survival data with covariates. *Biometrics*, pages 145–156.
- [19] Ibrahim, J. G., Chen, M.-H. & Sinha, D. (2001). *Bayesian Survival Analysis*. Springer, New York.
- [20] Karlis, D. (2003). An EM algorithm for multivariate Poisson distribution and related models. *Journal of Applied Statistics*, **30**(1), 63–77.
- [21] Kim, S., Chen, M. & Dey, D. (2011). A new threshold regression model for survival data with a cure fraction. *Lifetime data analysis*, **17**(1), 101–122.
- [22] Louzada, F., Suzuki, A. & Cancho, V. (2013). The FGM long-term bivariate survival copula model: Modeling, Bayesian estimation, and case influence diagnostics. *Communications in Statistics—Theory and Methods*, **42**(4), 673–691.
- [23] Louzada-Neto, F. (1999). Poly-hazard regression models for lifetime data. *Biometrics*, **55**, 1121–1125.
- [24] Maller, R. A. & Zhou, X. (1996). *Survival Analysis with Long-Term Survivors*. Wiley, New York.
- [25] Manatunga, A. & Oakes, D. (1999). Parametric analysis for matched pair survival data. *Lifetime Data Analysis*, **5**(4), 371–387.
- [26] Oakes, D. (1982). A model for association in bivariate survival data. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 414–422.
- [27] Perdoná, G. S. C. & Louzada-Neto, F. (2011). A general hazard model for lifetime data in the presence of cure rate. *Journal of Applied Statistics*, **38**, 1395–1405.
- [28] Price, D. & Manatunga, A. (2001). Modelling survival data with a cured fraction using frailty models. *Statistics in medicine*, **20**(9-10), 1515–1527.
- [29] R Core Team (2012). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
- [30] Rodrigues, J., Cancho, V. G., de Castro, M. & Louzada-Neto, F. (2009). On the unification of the long-term survival models. *Statistics & Probability Letters*, **79**, 753–759.
- [31] Romeo, J., Tanaka, N. & Pedrosa-de Lima, A. (2006). Bivariate survival modeling: a Bayesian approach based on copulas. *Lifetime Data Analysis*, **12**(2), 205–222.
- [32] Sahu, S. & Dey, D. (2000). A comparison of frailty and other models for bivariate survival data. *Lifetime data analysis*, **6**(3), 207–228.
- [33] Shih, J. & Louis, T. (1995). Inferences on the association parameter in copula models for bivariate survival data. *Biometrics*, pages 1384–1399.
- [34] Spiegelhalter, D. J., Best, N. G., Carlin, B. P. & van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society B*, **64**, 583–639.
- [35] Tournoud, M. & Ecochard, R. (2007). Application of the promotion time cure model with time-changing exposure to the study of HIV/AIDS and other infectious diseases. *Statistics in Medicine*, **26**, 1008–1021.
- [36] Tsodikov, A. D., Ibrahim, J. G. & Yakovlev, A. Y. (2003). Estimating cure rates from survival data: An alternative to two-component mixture models. *Journal of the American Statistical Association*, **98**, 1063–1078.
- [37] Yakovlev, A. Y. & Tsodikov, A. D. (1996). *Stochastic Models of Tumor Latency and Their Biostatistical Applications*. World Scientific, Singapore.