

Competing risks joint models using R-INLA

Janet van Niekerk¹, Haakon Bakka¹ and Håvard Rue¹

¹ CEMSE Division, King Abdullah University of Science and Technology, Saudi Arabia

Address for correspondence: Janet van Niekerk, CEMSE Division, King Abdullah University of Science and Technology, Thuwal, 23955, Saudi Arabia.

E-mail: Janet.vanNiekerk@kaust.edu.sa.

Phone: (+966) 12 808 0214.

Fax: .

Abstract: The methodological advancements made in the field of joint models are numerous. None the less, the case of competing risks joint models have largely been neglected, especially from a practitioner's point of view. In the relevant works on competing risks joint models, the assumptions of a Gaussian linear longitudinal series and proportional cause-specific hazard functions, amongst others, have remained unchallenged. In this paper, we provide a framework based on R-INLA to apply competing risks joint models in a unifying way such that non-Gaussian longitudinal data, spatial structures, time dependent splines and various latent association structures, to mention a few, are all embraced in our approach. Our motivation stems from the SANAD trial which exhibits non-linear longitudinal trajectories and competing risks for failure of treatment. We also present a discrete competing risks joint model for longitudinal count data as well as a spatial competing risks joint model, as specific

examples.

Key words: accelerated failure time model; competing risks; integrated nested Laplace approximation; joint model; non-Gaussian; spline

1 Introduction

Time to event data is observed in its simplest form, as the time when a particular event happens or when the monitoring process halts (censoring time). A competing risks model (Hoel, 1972; Moeschberger and David, 1971; Altshuler, 1970) arises in the case where patients are at risk for one of multiple events, i.e. multiple mutually exclusive outcomes. The occurrence of an event precludes all other events, so that each observational unit is at risk for each of the C events, until one event occurs. After this, the risk of the other events are all zero. In classical survival analysis the hazard and survival functions are estimated from the data to infer the risk of one specific event. In competing risks, the risk of each event is dependent on the assumption of surviving all other events, and these classical methods cannot be used for each of the multiple events without the marginalization of all other events. There are mainly two modeling approaches in this respect, the subdistribution (Fine and Gray, 1999) and cause-specific hazards approaches (Prentice et al., 1978). In the subdistribution approach the data enters the model through a predictor term in each of the cumulative incidence functions (CIF), the marginal cumulative probability of experiencing the specific event. On the other hand, the cause-specific hazards approach incorporates the data through a predictor term for each of the cause-specific hazard functions.

Debates surrounding the choice between the two approaches are still ongoing, with the consensus being that the cause-specific hazards approach should be chosen for causal inference of covariate effects. The CIF and the cause-specific hazard functions can be deduced from each other through certain functional relationships, see [Crowder \(2001\)](#) for more details.

In most situations, additional data is available in the form of repeated measurements over time constituting a longitudinal data set. The joint modeling of the competing risks data and the longitudinal data ([Hu et al., 2009](#); [Elashoff et al., 2007](#)) is often necessary since the missing longitudinal observations are not missing at random but rather due to the occurrence of one of the competing events. Ignoring this phenomena, will bias the longitudinal estimates. If the treatment efficacy is being evaluated, then joint modeling enables us to evaluate the treatment effect in both the longitudinal and time to event endpoints, simultaneously. Also, unobserved heterogeneity in the individuals due to some latent effect over time can be successfully accommodated through the joint modeling.

Computational frameworks for competing risks, have been developed in recent years such as the *CFC* package ([Sharabiani and Mahani, 2017](#)) and the *cmprsk* package ([Gray, 2014](#)). The existence of these frameworks is essential for catalyzing the use of competing risks models in practice.

For joint modeling of competing risks and longitudinal data the R package *joiner* ([Philipson et al., 2018](#)) can be used. This package implements the model developed by [Williamson et al. \(2008\)](#). This model is a cause-specific Cox proportional hazards regression model (for two competing events) and a linear mixed effects model

with an assumed latent Gaussian association process. An accelerated failure time (AFT) model, a non-Gaussian distribution for the longitudinal model component or spatial data thus excludes the use of *joiner*. Generic Bayesian software platforms like BayesX, STAN and JAGS can also be used to do inference on competing risks joint models through simulation, like Markov chain Monte Carlo (MCMC). In this paper we will not use simulation-based tools but rather approximate methods which have considerable computational advantage.

Practitioners depend largely on the availability of wieldy computational frameworks to ease the computational burden of implementing such joint models. On the flip side, too many different developments could confuse the potential users and exacerbate the hesitation to applying competing risks models. This is especially true when more complex models are needed and new packages are developed for each additional complexity. This is cumbersome for users since many different packages should be familiarized for common use. The exigency for a more cohesive approach to competing risks joint models is clear. We believe that we address this deficiency in this paper.

To achieve our aim, of presenting a cohesive computational framework for competing risks joint models, we are spurred to characterize competing risks models within the class of latent Gaussian models. This characterization enables us to use the integrated nested Laplace approximation (INLA) method proposed by [Rue et al. \(2009\)](#). Our rationale for this endeavor is the computational efficiency and ability to incorporate various modeling elements available in the accompanying *R-INLA* package ([Rue et al., 2011](#)). The INLA framework has been used in various fields of application since its inception. In survival analysis, [Martino et al. \(2011\)](#) provided the details for basic survival models and [Álvaro-Meca et al. \(2013\)](#) presented an approach for competing

risks models using R-INLA, while [Van Niekerk et al. \(2019\)](#) presented the case of joint survival longitudinal modeling. Competing risks joint models with frailties ([Martino et al., 2011](#)), spatial structures ([Bakka et al., 2018](#)) and non-linear covariate effects ([Lindgren and Rue, 2008](#)) with non-Gaussian longitudinal submodels are all natural possibilities within our framework.

In the SANAD trial ([Marson et al., 2007](#)), the aim was to evaluate the efficacy of various treatments for seizure control in patients with epilepsy. Two treatment failure types, unacceptable adverse effects (UAE) and inadequate seizure control (ISC) were considered as competing events. From exploratory data analysis the non-linear mean trajectory of the titration dosage over time for the two different treatments is clear and also noted by [Williamson et al. \(2008\)](#). A piecewise linear model with two different slopes (with the changepoint at time 500) was used by [Williamson et al. \(2008\)](#) as an attempt to properly model this phenomena. We, however, propose a fully nonparametric spline component to capture this non-linear behavior over time in the form of a random walk order two model. This type of non-linear competing risks model can be formulated as a latent Gaussian model which can then be fitted using the established R-INLA framework for accurate and efficient computation. This endeavor resulted in better results for the SANAD trial as well as the development of a class of latent Gaussian competing risks joint models for various other realistic data situations such as longitudinal count data, spatial data and correlated competing risks to name a few.

In [Section 2](#) we present an introduction and necessary details of competing risks joint models in general which is then formulated as latent Gaussian models in [Section 3](#). We also provide two specific examples of competing risks joint models which might

arise in practice to give the reader an intuition for the proposed methodology. We apply the proposed method to the SANAD trial in Section 4 and a simulation study for a discrete competing risks joint model (longitudinal count data) is presented in Section 5. Finally, we conclude with a discussion in Section 6 on the proposed method and some exciting possibilities thereof.

2 Joint models for competing risks and longitudinal data

The applications of joint models have increased over the last few years. Essentially, a joint model is a statistical method to model two or more dependent endpoints/responses, simultaneously. An example of joint models in spatial statistics are known as coregionalization models ([Schmidt and Gelfand, 2003](#); [Krainski et al., 2018](#)) where there is dependency between multiple spatial fields. In survival analysis, the use of joint models to infer joint survival longitudinal models is very popular. In these cases, joint modeling is fundamentally correct since the time to event process and the longitudinal data per individual is inherently dependent, since it is governed by the same biological process. Although joint models are popular in survival analysis with one event of interest, the potential for their application in more complicated survival models has yet to be explored. In this paper, we focus on competing risks joint models with various realistic features like multiple events of interest, spatial random effects and non-linear longitudinal trajectories. To this end, we present cause-specific competing risks joint models, with proportional or non-proportional cause-specific hazard functions and non-Gaussian or Gaussian longitudinal series as latent Gaus-

sian models. This venture allows the use of the fast and accurate INLA methodology [Rue et al. \(2009\)](#) for Bayesian inference of the model.

2.1 Cause-specific competing risks joint model

The competing risks joint model is used to do inference for the biological process that underpins the longitudinal and competing risks data. For simplicity we assume C competing risks and one longitudinal marker in this section, even though our approach can be used for various longitudinal markers.

For the longitudinal marker \mathbf{y} , we formulate a linear predictor $\boldsymbol{\eta}^L$ and vector of hyperparameters, $\boldsymbol{\theta}^L$ for N individuals with a total of $N^L = N_1 + N_2 + \dots + N_N$ observations. Note that we do not make any distributional assumptions, hence denote the likelihood of \mathbf{y} as

$$\pi_L(\mathbf{y}|\boldsymbol{\eta}^L, \boldsymbol{\theta}^L) = \prod_{l=1}^{N^L} \pi_L(y_l|\eta_l^L, \boldsymbol{\theta}^L). \quad (2.1)$$

Denote the cause-specific hazard function for cause $j, j = 1, \dots, C$ at time t as

$$h_j(t|\eta^j, \boldsymbol{\theta}^j),$$

where η^j is the linear predictor. The vector of hyperparameters $\boldsymbol{\theta}^j$ is associated with the cause-specific hazard function, for example the baseline hazard parameters for the Cox proportional hazard model or the shape parameter of the Weibull accelerated failure time (AFT) model. The likelihood function for the observed event times $\mathbf{t} = \{t_1, \dots, t_N\}$ with event indicators $\mathbf{d} = \{d_{ij}\}, i = 1, \dots, N, j = 1, \dots, C$ is

$$\pi_S(\mathbf{t}|\boldsymbol{\eta}, \boldsymbol{\theta}) = \prod_{i=1}^N \prod_{j=1}^C h_{ij}(t_i|\eta_i^j, \boldsymbol{\theta}^j)^{d_{ij}} \exp\left(-\int_0^{t_i} h_{ij}(u|\eta_i^j, \boldsymbol{\theta}^j) du\right), \quad (2.2)$$

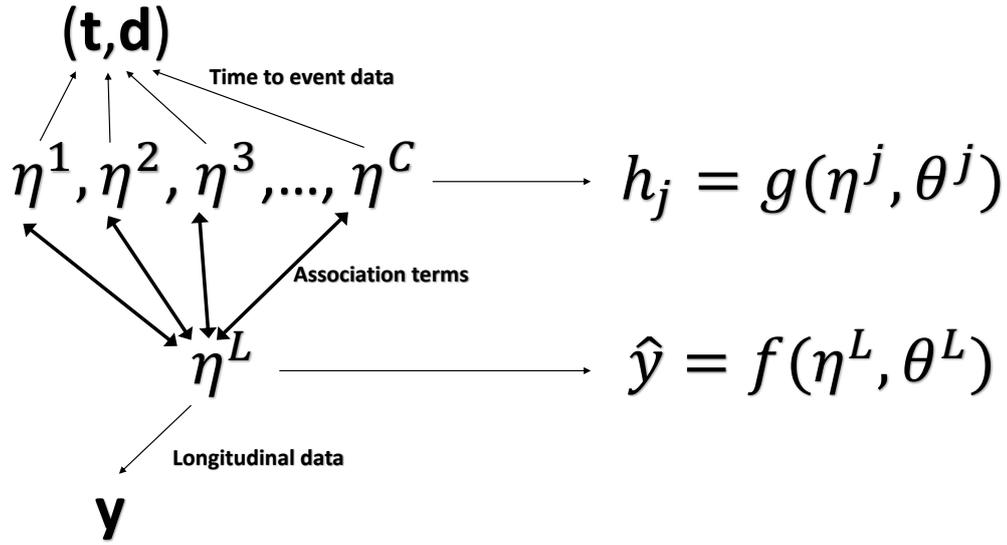


Figure 1: Graphical illustration of the setup of the competing risks joint model

where d_{ij} is an indicator if individual i died from cause j . The model is illustrated in Figure 1. The linear predictors, η^j and η^L , have an additive structure of fixed and random effects as follows:

$$\eta = \beta \mathbf{Z} + \sum_{k=1}^K f_k(u_{ik})$$

where $f_k(u_{ik})$ is the k^{th} random effect of covariate u_{ik} for individual i . This formulation allows us to include spatial fields, spline components and frailties within our approach to joint competing risks models. The form of the linear predictor facilitates the contextualization of this joint competing risks model as a latent Gaussian model.

3 Competing risks joint models as Latent Gaussian models

In this section we briefly present the details of latent Gaussian models (LGMs) and the integrated nested Laplace approximations (INLA) method for approximate Bayesian inference. If we can formulate competing risks joint models as LGMs and consequently use INLA for the inference, then we can fit these types of models very efficiently and with complex modeling components.

3.1 Latent Gaussian models and the integrated nested Laplace approximations method

Latent Gaussian models is a specific subset of hierarchical Bayesian additive models. This class comprises of well-known models such as mixed models, temporal and spatial models amongst many others. An LGM is defined as a model having a specific hierarchical structure, as follows:

1. The observed responses, $\{y_i\}$, are conditionally independent based on the likelihood hyperparameters $\boldsymbol{\theta}_1$ and the linear predictors $\{\eta_i\}$.
2. The linear predictor is formulated as follows:

$$\eta_i = \beta_0 + \boldsymbol{\beta}\mathbf{X} + \sum_{k=1}^K f_k(u_{ik}) + \epsilon_i \quad (3.1)$$

where $\boldsymbol{\beta}$ represent the linear fixed effects of the covariates \mathbf{X} , ϵ is the unstructured random effects and the $\{f_k(\cdot)\}$'s represent unknown (non-linear) functions

of the covariates \mathbf{u} . The unknown functions $\{f_k(\cdot)\}$'s, also known as structured random effects, include spline models, spatial effects, temporal effects, non-separable spatio-temporal effects, frailties, subject or group-specific intercepts and slopes etc.

3. The latent field $\boldsymbol{\mathcal{X}}$ is formed from the structured additive predictor as $\{\beta_0, \boldsymbol{\beta}, \{f_k(\cdot)\}, \boldsymbol{\eta}\}$ with a multivariate Gaussian prior such that the latent field forms a Gaussian Markov random field with sparse precision matrix $\mathbf{Q}(\boldsymbol{\theta}_2)$, i.e. $\boldsymbol{\mathcal{X}} \sim N(\mathbf{0}, \mathbf{Q}^{-1}(\boldsymbol{\theta}_2))$.
4. A prior (possibly non-Gaussian), $\boldsymbol{\pi}(\boldsymbol{\theta})$ can then be formulated for the set of hyperparameters $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$.

From this hierarchical Bayesian formulation the joint posterior distribution is then given by:

$$\boldsymbol{\pi}(\boldsymbol{\mathcal{X}}, \boldsymbol{\theta} | \mathbf{y}) \propto \boldsymbol{\pi}(\boldsymbol{\theta}) \boldsymbol{\pi}(\boldsymbol{\mathcal{X}} | \boldsymbol{\theta}) \prod_i \pi(y_i | \boldsymbol{\mathcal{X}}, \boldsymbol{\theta}) \quad (3.2)$$

Within this framework the joint posterior density (3.2) and subsequently the marginal posterior densities, $\boldsymbol{\pi}(\mathcal{X}_i | \mathbf{y}), i = 1 \dots n$ and $\boldsymbol{\pi}(\boldsymbol{\theta} | \mathbf{y})$ can be efficiently and accurately calculated using the INLA methodology developed by [Rue et al. \(2009\)](#), without using Gibbs sampling.

3.2 Competing risks joint models as LGMs

In this section we show that most competing risks joint models are indeed LGMs and can therefore be defined as one model. This enables the use of the INLA methodology for many different competing risks joint models, for example discrete competing risks joint models, spatial competing risks joint models, non-linear competing risks joint models, correlated competing risks joint models etc.

Consider the case with C competing risks, N individuals with time to event observations \mathbf{t} and event variable \mathbf{d} , and M series' of longitudinal observations each of size $N_l, l = 1, \dots, M$. Let $\mathbf{d}_i, i = 1, \dots, C$ be a vector of indicators if the event type is i .

Now define the stacked response list $\mathbf{y}^{\text{stacked}} = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_M, (\mathbf{t}_1, \mathbf{d}_1), (\mathbf{t}_2, \mathbf{d}_2), \dots, (\mathbf{t}_C, \mathbf{d}_C)\}$.

Also define the stacked linear predictor $\boldsymbol{\eta}^{\text{stacked}} = \{\boldsymbol{\eta}_1^L, \boldsymbol{\eta}_2^L, \dots, \boldsymbol{\eta}_M^L, \boldsymbol{\eta}^1, \boldsymbol{\eta}^2, \dots, \boldsymbol{\eta}^C, \}$ as in (2.1) and (2.2) where each component is of the form (3.1). For each component we define two vectors of hyperparameters, $\boldsymbol{\theta}_1$ for the hyperparameters from the likelihood (like a precision parameter if a Gaussian distribution is assumed, or the shape parameter for a Weibull model etc.) and $\boldsymbol{\theta}_2$ for the hyperparameters from the random effects in the linear predictor (like a range parameter for a spatial effect, or a precision parameter for a spline effect etc.), such that the stacked vector of hyperparameters is $\boldsymbol{\theta}^{\text{stacked}} = \{\boldsymbol{\theta}_{1,1}^L, \boldsymbol{\theta}_{2,1}^L, \boldsymbol{\theta}_{1,2}^L, \boldsymbol{\theta}_{2,2}^L, \dots, \boldsymbol{\theta}_{1,M}^L, \boldsymbol{\theta}_{2,M}^L, \boldsymbol{\theta}_{1,1}, \boldsymbol{\theta}_{2,1}, \boldsymbol{\theta}_{1,2}, \boldsymbol{\theta}_{2,2}, \dots, \boldsymbol{\theta}_{1,C}, \boldsymbol{\theta}_{2,C}\}$. The latent field is then defined as $\boldsymbol{\mathcal{X}} = \{\boldsymbol{\beta}^{\text{stacked}}, \boldsymbol{\eta}^{\text{stacked}}, \{\mathbf{f}(\cdot)\}^{\text{stacked}}\}$, and we assume a multivariate Gaussian prior for $\boldsymbol{\mathcal{X}}$.

From this construction the cause-specific competing risks joint model (in general, with various random effects like spatial fields, spline effects, correlated models etc.) is an LGM and the joint posterior distribution of the latent field and the hyperparameters is

$$\pi(\boldsymbol{\mathcal{X}}, \boldsymbol{\theta}^{\text{stacked}}) \propto \pi(\boldsymbol{\theta}^{\text{stacked}}) \pi(\boldsymbol{\mathcal{X}} | \boldsymbol{\theta}^{\text{stacked}}) \prod_{i=1}^{\text{length}(\mathbf{y}^{\text{stacked}})} \pi(y_i | \boldsymbol{\mathcal{X}}, \boldsymbol{\theta}^{\text{stacked}}).$$

Since this proposition is abstract we provide some specific examples of relevant competing risks joint models in the next section.

3.3 Specific examples

Next we present some specific (non-exhaustive) examples of competing risks joint models that fit in this framework.

3.3.1 Example 1. Poisson competing risks random-effects joint model

Suppose that the longitudinal observations, $y_i(t)$ follow a Poisson distribution with mean $\lambda_i(t)$, i.e., $y_i(t) \sim \text{Poisson}(\lambda_i(t))$ with $\lambda_i(t) = \exp(\eta_i^L(t))$. Also assume three competing causes and a Weibull model for all the cause-specific hazard functions with shape parameters $\alpha_1, \alpha_2, \alpha_3$, i.e. for observed time t_i with event indicators $\mathbf{d}_i = \{d_{i1}, d_{i2}, d_{i3}\}$, $h_{ij}(t) = [\alpha_j(\gamma_{ij})^{\alpha_j} t^{\alpha_j-1}]^{d_{ij}}$ with $\gamma_{ij} = \exp(\eta_i^{(j)})$, $j = 1, 2, 3$.

Now define the associated linear predictors as follows:

$$\begin{aligned}\eta^L(t) &= \boldsymbol{\beta}^L \mathbf{X}^L(t) + \mathbf{v} + \mathbf{w}t \\ \eta^{(1)}(t) &= \boldsymbol{\beta}^{(1)} \mathbf{X}^{(1)}(t) + \gamma^{(1)} \mathbf{v} + \kappa^{(1)} \mathbf{w}t \\ \eta^{(2)}(t) &= \boldsymbol{\beta}^{(2)} \mathbf{X}^{(2)}(t) + \gamma^{(2)} \mathbf{v} + \kappa^{(2)} \mathbf{w}t \\ \eta^{(3)}(t) &= \boldsymbol{\beta}^{(3)} \mathbf{X}^{(3)}(t) + \gamma^{(3)} \mathbf{v} + \kappa^{(3)} \mathbf{w}t,\end{aligned}\tag{3.3}$$

where $\begin{bmatrix} v_i \\ w_i \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \tau_v^{-1} & \rho \\ \rho & \tau_w^{-1} \end{bmatrix}\right)$. Note that we do not enforce the same scaling parameter for the two random effects but accommodate for different scalings of the random intercept and the random slope.

Now define $\boldsymbol{\theta} = \{\alpha_1, \alpha_2, \alpha_3, \tau_v^{-1}, \tau_w^{-1}, \rho\}$ as the vector of hyperparameters. Define the latent field

$$\boldsymbol{\mathcal{X}} = \{\boldsymbol{\beta}^L, \boldsymbol{\beta}^{(1)}, \boldsymbol{\beta}^{(2)}, \boldsymbol{\beta}^{(3)}, \boldsymbol{\eta}^L, \boldsymbol{\eta}^{(1)}, \boldsymbol{\eta}^{(2)}, \boldsymbol{\eta}^{(3)}, \mathbf{v}, \mathbf{w}\},$$

then this competing risks joint model is an LGM following Section 3.

3.3.2 Example 2. Spatial competing risks joint model with smoothed longitudinal predictor as the association term and unknown non-linear time trend

Suppose that the longitudinal observations, y_i follow a Gaussian distribution with mean μ_i and precision parameter τ , i.e., $y_i \sim N(\mu_i, \tau^{-1})$ with $\mu_i = \eta_i^L(s, t)$ with t a time index and s a space index. Also assume three competing causes and an exponential model for all the cause-specific hazard functions i.e. for observed time t_i with event indicators $\mathbf{d}_i = \{d_{i1}, d_{i2}, d_{i3}\}$, $h_{ij}(t) = \lambda_{ij}^{d_{ij}}$ with $\lambda_{ij} = \exp(\eta_i^{(j)})$, $j = 1, 2, 3$.

Now define the associated linear predictors as follows:

$$\begin{aligned}\eta^L(s, t) &= \boldsymbol{\beta}^L \mathbf{X}^L(t) + \mathbf{v} + f(t) + u(s) \\ \eta^{(1)}(s, t) &= \boldsymbol{\beta}^{(1)} \mathbf{X}^{(1)}(t) + \gamma^{(1)} \eta^L(s, t) \\ \eta^{(2)}(s, t) &= \boldsymbol{\beta}^{(2)} \mathbf{X}^{(2)}(t) + \gamma^{(2)} \eta^L(s, t) \\ \eta^{(3)}(s, t) &= \boldsymbol{\beta}^{(3)} \mathbf{X}^{(3)}(t) + \gamma^{(3)} \eta^L(s, t),\end{aligned}\tag{3.4}$$

where $f(t)$ is an unknown non-linear function of time (i.e. a spline component) with hyperparameters $\boldsymbol{\theta}_{\text{time}}$ and $u(s)$ is a spatial field (eg. Besag (Besag, 1974) or BYM (Besag et al., 1991) models for discrete regions or SPDE (eg. Matern) (Gelfand et al., 2010; Lindgren et al., 2011; Krainski et al., 2018) model for continuous space) with hyperparameters $\boldsymbol{\theta}_{\text{space}}$. Now each cause-specific hazard function is associated with the longitudinal process through its own scaling $\gamma^{(j)}$.

Now define $\boldsymbol{\theta} = \{\boldsymbol{\theta}_{\text{time}}, \boldsymbol{\theta}_{\text{space}}, \tau\}$ as the vector of hyperparameters. Define the latent field

$$\boldsymbol{\mathcal{X}} = \{\boldsymbol{\beta}^L, \boldsymbol{\beta}^{(1)}, \boldsymbol{\beta}^{(2)}, \boldsymbol{\beta}^{(3)}, \boldsymbol{\eta}^L, \boldsymbol{\eta}^{(1)}, \boldsymbol{\eta}^{(2)}, \boldsymbol{\eta}^{(3)}, \mathbf{f}(\cdot), \mathbf{u}(\cdot)\},$$

then this spatial competing risks joint model is an LGM following Section 3.

4 Application - SANAD trial

The SANAD trial ([Marson et al., 2007](#)) was a randomized trial study to investigate the effect of seizure control drugs in patient with epilepsy for whom carbamazepine (CBZ) was considered the standard treatment protocol. Two treatment failure types, unacceptable adverse effects (UAE) and inadequate seizure control (ISC) were considered as competing events. The patients were randomly assigned to CBZ, lamotrigine (LTG), gabapentin, oxcarbazepine and topiramate. In the main analysis, LTG appeared to be superior to CBZ but since then various criticisms regarding this finding have risen, mainly due to the difference in titration speed and therefore the dosage. In [Williamson et al. \(2008\)](#) this data was analyzed using a competing risks random-effects joint model with a changepoint at time $t = 500$. In [Figure 2](#) the mean trajectories for the two treatment groups CBZ and LTG are presented. It is clear that the mean trajectories are firstly different between the two treatments and secondly, the changepoint at $t = 500$ seems quite subjective since for CBZ the trajectory does not seem to have a constant linear trend after $t = 500$. We therefore include a spline component in the form of a random walk order two model for the dosage trajectory over time. This spline model is a stochastic spline, i.e. converges to the true trajectory as the number of knots increase. We thus used 50 knots over the range of observed times. Also, we do not enforce equal scaling for the random intercepts and slopes.

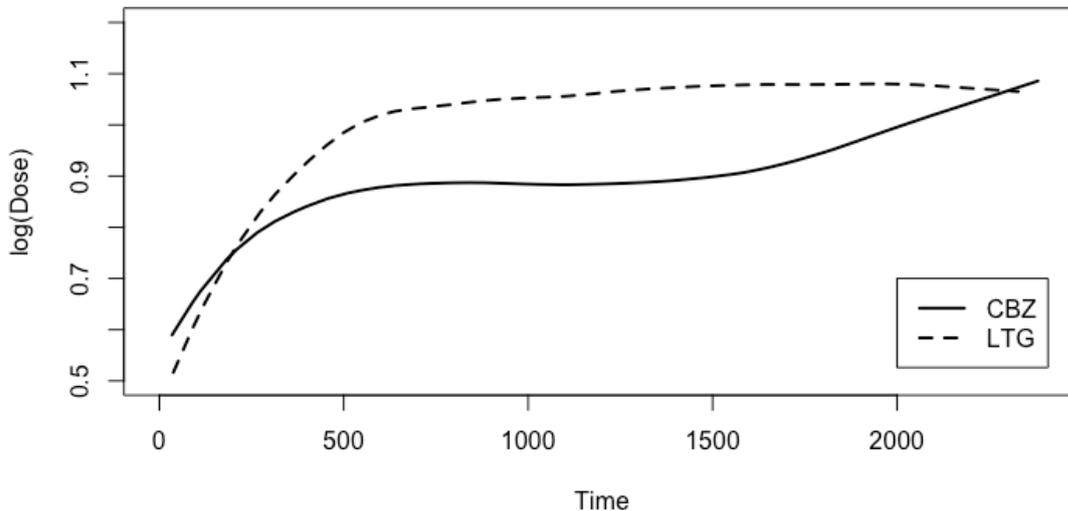


Figure 2: Mean trajectories for CBZ and LTG treatment groups

4.1 Bayesian competing risks joint model

We propose the following competing risks joint model with non-linear longitudinal trajectories for the SANAD trial data at time t ,

$$\begin{aligned}\eta^L(t) &= \beta_0 + \beta_1 \text{CBZ} + \text{CBZ} f_{\text{CBZ}}(t) + \text{LTG} f_{\text{LTG}}(t) & (4.1) \\ \eta^{(\text{UAE})}(t) &= \beta_0 + \beta_1^{(1)} \text{CBZ} + \gamma_1 \eta^L(t) \\ \eta^{(\text{ISC})}(t) &= \beta_0 + \beta_1^{(2)} \text{CBZ} + \gamma_2 \eta^L(t),\end{aligned}$$

where f_{CBZ} is a random walk order two model (Lindgren and Rue, 2008) with hyperparameter $\tau_{f_{\text{CBZ}}}$ on the interaction of the time and CBZ, the CBZ treatment indicator, and similarly for LTG. We assume a Gaussian longitudinal distribution with noise precision τ_L , and Weibull models for the cause-specific hazard functions

with hyperparameters $\alpha_{\text{UAE}}, \alpha_{\text{ISC}}$. The priors for the latent field are all Gaussian i.e.,

$$\begin{aligned}\beta_0 &\sim N(0, \tau_{\beta_0}^{-1}), & \beta_1 &\sim N(0, \tau_{\beta_1}^{-1}), & \beta_1^{(1)} &\sim N(0, \tau_{\beta_1^{(1)}}^{-1}), \\ \beta_1^{(2)} &\sim N(0, \tau_{\beta_1^{(2)}}^{-1}), & \gamma_1 &\sim N(0, \tau_{\gamma_1}^{-1}), & \gamma_2 &\sim N(0, \tau_{\gamma_2}^{-1}),\end{aligned}$$

with penalizing complexity priors (Simpson et al., 2017) for most of the hyperparameters, such that

$$\begin{aligned}\sigma_L &= \tau_L^{-0.5} \sim \text{Exp}(\lambda_L), & \sigma_{\beta_0} &= \tau_{\beta_0}^{-0.5} \sim \text{Exp}(\lambda_{\beta_0}), & \sigma_{\beta_1} &= \tau_{\beta_1}^{-0.5} \sim \text{Exp}(\lambda_{\beta_1}), \\ \sigma_{\beta_1^{(1)}} &= \tau_{\beta_1^{(1)}}^{-0.5} \sim \text{Exp}(\lambda_{\beta_1^{(1)}}), & \sigma_{\beta_1^{(2)}} &= \tau_{\beta_1^{(2)}}^{-0.5} \sim \text{Exp}(\lambda_{\beta_1^{(2)}}), & \sigma_{\gamma_1} &= \tau_{\gamma_1}^{-0.5} \sim \text{Exp}(\lambda_{\gamma_1}), \\ \sigma_{\gamma_2} &= \tau_{\gamma_2}^{-0.5} \sim \text{Exp}(\lambda_{\gamma_2}), & \sigma_{f_{\text{CBZ}}} &= \tau_{f_{\text{CBZ}}}^{-0.5} \sim \text{Exp}(\lambda_{\text{CBZ}}), & \sigma_{f_{\text{LTG}}} &= \tau_{f_{\text{LTG}}}^{-0.5} \sim \text{Exp}(\lambda_{\text{CBZ}}), \\ 10 \ln(\alpha_{\text{UAE}}) &\sim N(0, \tau_{\text{UAE}}^{-1}), & 10 \ln(\alpha_{\text{ISC}}) &\sim N(0, \tau_{\text{ISC}}^{-1}).\end{aligned}$$

The user-defined parameter $\lambda_{[\cdot]}$ is defined through the probability $P(\tau_{[\cdot]}^{-0.5} > 1)$, as an example $P(\tau_{[\cdot]}^{-0.5} > 1) = 0.01$ produces $\lambda_{[\cdot]} = \ln(0.01)$. Note that the exponential prior on the standard deviation $\sigma_{[\cdot]} = \tau_{[\cdot]}^{-0.5}$ translates to a Gumbel type 2 prior on $\tau_{[\cdot]}$. The motivation for employing penalizing complexity priors for the precision hyperparameters are founded in the fact that the usual priors for the variance components, i.e. independent inverse-gamma priors as in Huang et al. (2018), overfits in the sense that it cannot contract to the simpler model in which the respective model component has trivial variance. This is especially important in the case of joint models since the effect of overfitting is exacerbated by the influence of the shared random effect on all the linear predictors.

4.2 Results

The estimated mean dosage trajectories for the two treatments, CBZ and LTG, on a log scale is presented in Figure 3. The effect of the joint model on the longitudinal

trajectories is clear when compared to Figure 2, in which the dropout data is not used. The dropout process thus affects the longitudinal trajectories in a meaningful way.

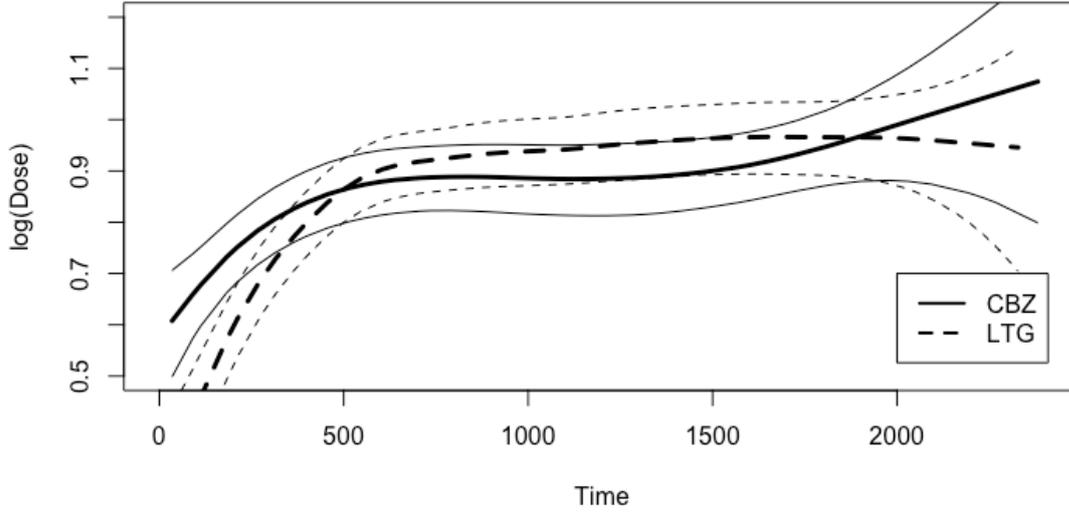


Figure 3: Mean estimated dosage trajectories for the CBZ and LTG treatment groups in bold (with 95% credible intervals)

The estimated competing risks joint model is:

$$\begin{aligned}\hat{\eta}^L(t) &= 2.438(2.319; 2.556) + 0.305(0.145; 0.465)\text{CBZ} + \text{CBZ}f_{\text{CBZ}}(t) + \text{LTG}f_{\text{LTG}}(t) \\ \hat{\eta}^{(\text{UAE})}(t) &= 2.438(2.319; 2.556) - 0.726(-1.104; -0.371)\text{CBZ} - 1.1(-1.192; -1.02)\eta^L(t) \\ \hat{\eta}^{(\text{ISC})}(t) &= 2.438(2.319; 2.556) - 0.11(-0.406; 0.175)\text{CBZ} - 0.998(-1.087; -0.928)\eta^L(t),\end{aligned}$$

with $\tau_{f_{\text{CBZ}}} = 14.879$, $\tau_{f_{\text{LTG}}} = 7.055$, $\alpha_{\text{UAE}} = 0.84(0.05; 0.915)$ and $\alpha_{\text{ISC}} = 1.21(1.19; 1.33)$.

The negative association between the cause-specific hazard functions of UAE and ISC, respectively, and the smoothed dosage level should be interpreted with the shape

parameters, α_{UAE} and α_{ISC} of the weibull models. For UAE, the estimated shape parameter is significantly smaller than 1 which implies that as the dose increase over time as captured by $\hat{\eta}^L(t)$ the linear predictor $\hat{\eta}^{(\text{UAE})}(t)$ will decrease due to the negative association and even more so for treatment CBZ. The decreased $\hat{\eta}^{(\text{UAE})}(t)$ together with $\alpha_{\text{UAE}} = 0.84$ leads to the conclusion that the hazard of UAE is lower for LTG than for CBZ, and increasing over time since higher doses correspond to larger times. The converse is true for the hazard of ISC since α_{ISC} is significantly larger than 1. Due to the insignificant effect of treatment on the hazard of ISC, the two treatments under consideration provide similar seizure control.

Our findings are thus consistent with those of [Williamson et al. \(2008\)](#) with the added advantage of the flexible estimated mean longitudinal trajectories for the two treatment dosages.

5 Simulation study - Discrete competing risks random-effects joint model

In this section we simulate longitudinal count data as well as time to event data with three competing risks (Causes 1, 2 and 3). We also included a non-linear time component in the longitudinal trajectory. The data is simulated from a competing risks joint random-effects model with a random intercept (similar to (3.4)), such that for N individuals, each with $N_i, i = 1, \dots, N$ longitudinal counts and time to event t_i

with event indicator $d_i \in \{0, 1, 2, 3\}$ the model is

$$\begin{aligned} y_i(t) &\sim \text{Poisson}(\exp(t^{1.2} + u_i)) \\ h_1(t) &= \exp(0.3u_i + 0.01\text{Age}) \\ h_2(t) &= \exp(-0.1u_i + 0.015\text{Age}) \\ h_3(t) &= \exp(0.2u_i + 0.0003\text{Age}), \end{aligned}$$

with Age a discrete explanatory variable within the range (15; 75) and $u_i \sim N(0, 1)$ is a random intercept which associates the longitudinal series with the cause-specific hazard functions, $h_j(t)$, $j = 1, 2, 3$. Note that we also include a non-linear trend in the longitudinal trajectory, $t^{1.2}$. In the simulation setup, we induce a negative association between the hazard of event 2 and the longitudinal series, whereas the hazard of events 1 and 3 are positively correlated with the longitudinal series. From the above, the hazard functions are constant over time and an exponential distribution is thus implied.

5.1 Results

In Figure 4 the longitudinal series and the mean estimated trajectory from a random walk order two model with precision τ_r is presented. The results are presented in Table 1 and it is clear that this challenging model is recovered well.

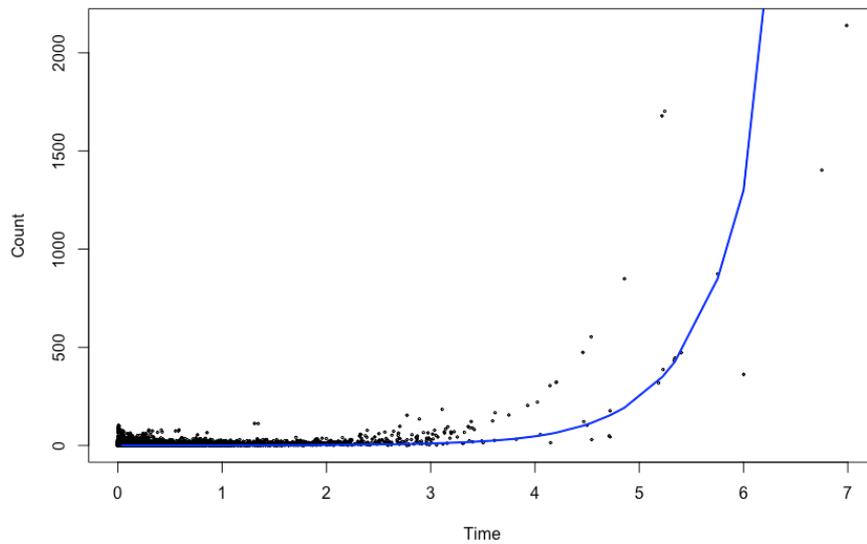


Figure 4: Simulated longitudinal data and estimated mean trajectory

From the estimates of the covariate effects in the cause-specific hazard functions we can deduce various other quantities of interest like the cause-specific survival-like function, CIF etc. In Figure 5 the survival-like functions of the three risks are illustrated. Certain functions in the *R-INLA* framework can be used to estimate the marginal distributions of these quantities.

Parameter	True value	Estimate	95% credible interval
σ_u	1	1.017	(0.925; 1.114)
β_1	0.01	0.012	(0.005; 0.018)
β_2	0.015	0.052	(0.009; 0.019)
β_3	0.0003	0.0001	(-0.003; 0.005)
γ_1	0.3	0.285	(0.192; 0.369)
γ_2	-0.1	-0.108	(-0.206; -0.011)
γ_3	0.2	0.208	(0.062; 0.352)
τ_r	—	28.513	(11.709; 116.878)

Table 1: Estimates from the discrete competing risks joint model using R-INLA

6 Discussion

Here we presented a general method to apply a wide class of competing risks joint models using the R-INLA framework. The main stimulus for this research was the necessity for a proper non-linear competing risks joint model as motivated from the SANAD trial. Although we addressed this issue, we also developed a class of latent Gaussian competing risks joint models which enables the practitioner to use our method for various realistic situations which demands different types of competing risks joint models. We presented the details of two specific realistic situations, one being a competing risks model with three competing events and one longitudinal series of count data, the other a spatial competing risks joint model. The application of the non-linear competing risks joint model to the SANAD trial proved fruitful as this model properly captures the non-linear trajectory.

The possibility of a non-linear competing risks joint model based on a random walk

order two model emerged through the R-INLA framework. The INLA methodology (Rue et al., 2009) has been proved to be a very efficient and popular computational platform for Bayesian models. The use of the R-INLA framework necessitated the formulation of the proposed model as a latent Gaussian model, which in turn facilitated the formulation of various other competing risks joint models that could arise in practical situations, such as multiple longitudinal series, discrete longitudinal distributions, spatial models and frailty models to name but a few. Quantities of interest like the cause-specific survival function, the all-cause survival function, the median survival time etc. can be computed in R-INLA.

The possibilities of realistic competing risks joint models that can be inferred through the proposed methodology is vast and exciting. In particular, spatial/spatio-temporal random effects can be included in the cause-specific hazard functions and/or the longitudinal submodel to develop, for example, a spatial competing risks joint model. The works of Hickey et al. (2018) on the SANAD trial, Huang et al. (2010) on outlying values in the longitudinal submodel, Rajeswaran et al. (2018) on multivariate longitudinal data and the model presented by Elashoff et al. (2008) are some examples in the literature that can be applied with our approach through R-INLA. A thorough presentation of all the possibilities is not possible in this paper but we believe that the general method presented herein catalyzes future research, especially different instances of competing risks joint models motivated from a modeling perspective.

7 Appendix

The code for this article is available at <http://www.statmod.org/smij/archive.html>.

References

- Altshuler, B. (1970). Theory for the measurement of competing risks in animal experiments. *Mathematical Biosciences*, **6**, 1–11.
- Álvaro-Meca, A., Akerkar, R., Alvarez-Bartolome, M., Gil-Prieto, R., Rue, H., and de Miguel, Á. G. (2013). Factors involved in health-related transitions after curative resection for pancreatic cancer. 10-years experience: a multi state model. *Cancer epidemiology*, **37**(1), 91–96.
- Bakka, H., Rue, H., Fuglstad, G.-A., Riebler, A., Bolin, D., Illian, J., Krainski, E., Simpson, D., and Lindgren, F. (2018). Spatial modeling with R-INLA: A review. *Wiley Interdisciplinary Reviews: Computational Statistics*, **10**(6), e1443. doi: 10.1002/wics.1443. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/wics.1443>.
- Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society: Series B (Methodological)*, **36**(2), 192–225.
- Besag, J., York, J., and Mollié, A. (1991). Bayesian image restoration, with two applications in spatial statistics. *Annals of the institute of statistical mathematics*, **43**(1), 1–20.

- Crowder, M. J. (2001). *Classical competing risks*. Chapman and Hall/CRC.
- Elashoff, R. M., Li, G., and Li, N. (2007). An approach to joint analysis of longitudinal measurements and competing risks failure time data. *Statistics in Medicine*, **26**(14), 2813–2835.
- Elashoff, R. M., Li, G., and Li, N. (2008). A joint model for longitudinal measurements and survival data in the presence of multiple failure types. *Biometrics*, **64**(3), 762–771.
- Fine, J. P. and Gray, R. J. (1999). A proportional hazards model for the subdistribution of a competing risk. *Journal of the American statistical association*, **94**(446), 496–509.
- Gelfand, A. E., Diggle, P., Guttorp, P., and Fuentes, M. (2010). *Handbook of spatial statistics*. CRC press.
- Gray, B. (2014). *cmprsk: Subdistribution Analysis of Competing Risks*. URL <https://CRAN.R-project.org/package=cmprsk>.
- Hickey, G. L., Philipson, P., Jorgensen, A., and Kolamunnage-Dona, R. (2018). A comparison of joint models for longitudinal and competing risks data, with application to an epilepsy drug randomized controlled trial. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, **181**(4), 1105–1123.
- Hoel, D. G. (1972). A representation of mortality data by competing risks. *Biometrics*, pages 475–488.
- Hu, W., Li, G., and Li, N. (2009). A bayesian approach to joint analysis of longitudinal measurements and competing risks failure time data. *Statistics in medicine*, **28**(11), 1601–1619.

- Huang, X., Li, G., and Elashoff, R. M. (2010). A joint model of longitudinal and competing risks survival data with heterogeneous random effects and outlying longitudinal measurements. *Statistics and its interface*, **3**(2), 185.
- Huang, Y., Lu, X., Chen, J., Liang, J., and Zangmeister, M. (2018). Joint model-based clustering of nonlinear longitudinal trajectories and associated time-to-event data analysis, linked by latent class membership: with application to AIDS clinical studies. *Lifetime data analysis*, **24**(4), 699–718.
- Krainski, E. T., Gómez-Rubio, V., Bakka, H., Lenzi, A., Castro-Camilo, D., Simpson, D., Lindgren, F., and Rue, H. (2018). *Advanced spatial modeling with stochastic partial differential equations using R and INLA*. Chapman and Hall/CRC.
- Lindgren, F. and Rue, H. (2008). On the second-order model for irregular locations. *Scandinavian Journal of Statistics*, **35**(4), 691–700.
- Lindgren, F., Rue, H., and Lindström, J. (2011). An explicit link between gaussian fields and gaussian markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **73**(4), 423–498.
- Marson, A., Al-Kharusi, A., Alwaidh, M., Appleton, R., Baker, G., Chadwick, D., Cramp, C., Cockerell, O., Cooper, P., Doughty, J., and Eaton, B. (2007). Carbamazepine versus gabapentin, lamotrigine, oxcarbazepine or topiramate for partial epilepsy: results from arm a of the sanad trial. *The Lancet*, **369**, 1000–1015.
- Martino, S., Akerkar, R., and Rue, H. (2011). Approximate Bayesian inference for survival models. *Scandinavian Journal of Statistics*, **38**(3), 514–528.

- Moeschberger, M. and David, H. (1971). Life tests under competing causes of failure and the theory of competing risks. *Biometrics*, pages 909–933.
- Philipson, P., Sousa, I., Diggle, P. J., Williamson, P., Kolamunnage-Dona, R., Henderson, R., Hickey, G. L., and Sudell, M. (2018). *joineR: Joint Modelling of Repeated Measurements and Time-to-Event Data*. URL <https://CRAN.R-project.org/package=joineR>.
- Prentice, R. L., Kalbfleisch, J. D., Peterson Jr, A. V., Flournoy, N., Farewell, V. T., and Breslow, N. E. (1978). The analysis of failure times in the presence of competing risks. *Biometrics*, pages 541–554.
- Rajeswaran, J., Blackstone, E. H., and Barnard, J. (2018). Joint modeling of multivariate longitudinal data and competing risks using multiphase sub-models. *Statistics in Biosciences*, **10**(3), 651–685.
- Rue, H., Martino, S., and Chopin, N. (2011). *INLA: Approximate Bayesian Inference using Integrated Nested Laplace Approximations*. URL www.r-inla.org.
- Rue, H., Martino, S., and Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested laplace approximations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **71**(2), 319–392.
- Schmidt, A. M. and Gelfand, A. E. (2003). A bayesian coregionalization approach for multivariate pollutant data. *Journal of Geophysical Research: Atmospheres*, **108** (D24).
- Sharabiani, M. T. and Mahani, A. S. (2017). *CFC: Cause-Specific Framework for Competing-Risk Analysis*. URL <https://CRAN.R-project.org/package=CFC>.

Simpson, D., Rue, H., Riebler, A., Martins, T. G., Sørbye, S. H., et al. (2017). Penalising model component complexity: A principled, practical approach to constructing priors. *Statistical Science*, **32**(1), 1–28.

Van Niekerk, J., Bakka, H., and Rue, H. (2019). Joint models as latent gaussian models-not reinventing the wheel. *arXiv preprint arXiv:1901.09365*.

Williamson, P., Kolamunnage-Dona, R., Philipson, P., and Marson, A. (2008). Joint modelling of longitudinal and competing risks data. *Statistics in medicine*, **27**(30), 6426–6438.

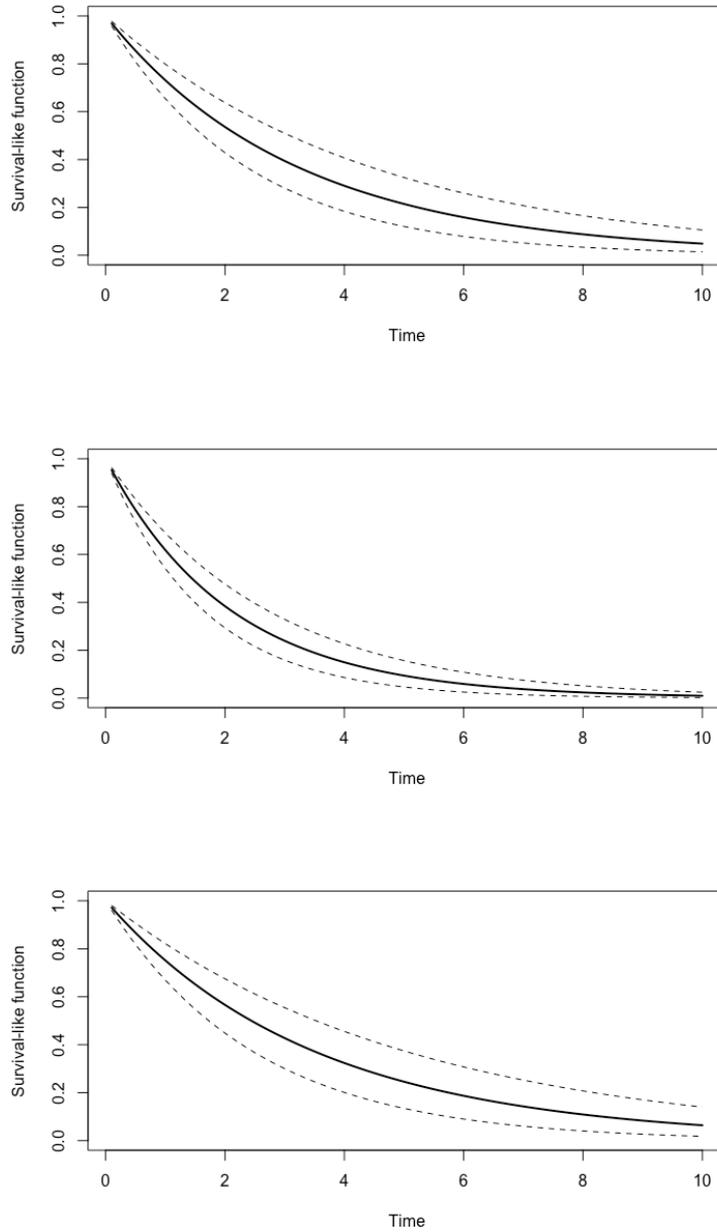


Figure 5: Cause-specific survival-like functions of the three risks with 95% credible interval for risks 1,2 and 3 (from top to bottom)