Semiparametric frailty model for clustered interval-censored data

Aysun Çetinyürek Yavuz [1], and Philippe Lambert [1,2]

[1] Faculté des Sciences Sociales, Université de Liège, Liège, Belgium
[2] Institut de Statistique, Université catholique de Louvain, Louvain la Neuve, Belgium

---

Address for correspondence: Faculté des sciences sociales, Université de Liège, Place des orateurs, 3, Quartier Agora (Bat. B31) B-4000 Liège, Belgium.
E-mail: cetinyurek@yahoo.com.

Abstract: The shared frailty model is a popular tool to analyze correlated right-censored time-to-event data. In the shared frailty model, the latent frailty is assumed to be shared by the members of a cluster and is assigned a parametric distribution, typically a gamma distribution due to its conjugacy. In the case of interval-censored time-to-event data, the inclusion of frailties results in complicated intractable likelihoods. Here, we propose a flexible frailty model for analyzing such data by assuming a smooth semiparametric form for the conditional time-to-event distribution and a parametric or a flexible form for the frailty distribution. The results of a simulation study suggest that the estimation of regression parameters is robust to misspecification of the frailty distribution (even when the frailty distribution is multimodal or skewed). Given sufficiently large sample sizes and number of clusters, the flexible approach produces smooth and accurate posterior estimates for the baseline survival function and for the frailty density, and can correctly detect and identify unusual frailty density forms. The methodology is illustrated using dental data from the Signal Tandmobiel® Study.

Key words: Interval censoring; shared frailty; Bayesian; P-splines

1 Introduction

Interval-censored time-to-event data arise frequently in longitudinal studies, where the event of interest is only known to have occurred between the two consecutive visits. A number of parametric, semi-parametric and non-parametric approaches have already been proposed to analyze such data (Peto, 1973; Turnbull, 1976; Komárek et al., 2005; Zhang and Davidian, 2008); including extensions of Cox's proportional hazards (PH) model, the most popular regression model for time-to-event data (Finkelstein, 1986; Cai and Betensky, 2003; Çetinyürek-Yavuz and Lambert, 2011). In the Cox PH model, the observations are assumed to be independent. However, this may not be true in certain situations where the observed

responses are clustered or subject to multiple measurements. A number of approaches generalizing Cox's PH model to handle correlated interval-censored data has been proposed in the literature. Marginal approaches based on generalised estimating equations (GEE) have been presented for the Cox model (Goggins and Finkelstein, 2000; Kim and Xue, 2002; Kor et al., 2013). However, marginal approaches usually do not allow inferences on the relationship or association between the event times. Alternatively, heterogeneity between clusters due to unknown or unobservable risk factors can be accounted for by the introduction of a frailty component in a conditional Cox PH model.

Frailty is a term describing (common) excess risk of failure among individuals in the same cluster. For right-censored time-to-event data, Vaupel et al. (1979) were the first to introduce the concept of frailty assuming a gamma distribution (due to the conjugacy property). The frailty model has been thoroughly studied since then by many authors in an attempt to model the heterogeneity between the clusters or account for the effect of unobserved covariates (Clayton, 1978; Clayton and Cuzick, 1985; Aalen, 1988). Other parametric frailty distributions such as lognormal, uniform, Weibull, positive stable or inverse Gaussian have also been considered (Vaupel and Yashin, 1983; Hougaard, 1984, 1986). However, most softwares limit the choice of the frailty distribution to lognormal and gamma distributions (Therneau, 2012); see for example *coxph* in package *survival* in R (R Development Core Team, 2012). A detailed summary of the frailty model can be found in Hougaard (2000).

Frailty models have also been adapted to handle interval-censored data. Bellamy et al. (2004); Goethals et al. (2009) assumed a Weibull model for the unobserved event times to obtain a closed form for the marginal likelihood using a lognormal or a gamma distribution for the frailty. Zuma and Lurie (2005); Zuma (2007) presented the use of EM algorithm to analyze correlated interval-censored data arising from an HIV study, considering the failure times and the frailty terms as missing in a PH model with a constant baseline hazard and gamma frailty. Zuma and Lurie (2005) also compared the parameter estimates from the EM algorithm with those obtained using a Bayesian approach. Similarly, Lam et al. (2010) proposed to use the EM algorithm with a robust estimator of the covariance matrix to adjust for the possible misspecification of the parametric baseline hazard function in a gamma frailty model. Henschel et al. (2009) proposed a gamma or lognormal frailty model assuming a piecewise constant baseline hazard in a Bayesian framework via a data augmentation approach. Wen and Chen (2013) presented a gamma frailty PH model with bivariate mixed case interval-censored data using a semiparametric maximum likelihood approach.

An important limitation of the above papers is represented by the parametric assumptions made for the baseline hazard and the frailty distribution for computational convenience. However, commonly assumed distributions may be too restrictive in practice to model the within-cluster heterogeneity (Shih and Louis, 1995). It has been shown that the regression parameter estimates are robust to the choice of random effect distribution in specific contexts (Klein et al., 1992), but are sensitive to that choice in other frameworks such as in GLMM (Agresti et al., 2004; Chen et al., 2002). The choice of frailty distribution has received very little attention in the clustered interval-censored data setting (Komárek and Lesaffre, 2008;

Goethals et al., 2009; Henschel et al., 2009) because it is still a challenge using frequentist methods. For these reasons, we aim to present a shared frailty PH model with flexible forms for the time-to-event distribution and possibly also for the frailty distribution following Eilers and Marx (1996); Lambert and Eilers (2009) in a Bayesian framework.

The rest of the paper is organized as follows. In Section 2, we introduce the shared frailty PH model. Our strategies for the specification of a smooth baseline density from interval-censored data and of a flexible frailty distribution are explained. Section 3 is devoted to the Bayesian inference. Section 4 contains the details and results of an extensive simulation study. The methodology is illustrated with a dataset from the Signal Tandmobiel® study in Section 5. A discussion of the results concludes the paper.

2   Shared frailty model

The frailty model is used to explore the effect of risk factors (covariates) on time until the event of interest occurs in the presence of unexplained heterogeneity in the population. Suppose that there are G independent clusters, each containing n_g sub-units. Let t_{gj} denote the (continuous) time until the event of interest occurs, and let x_{gj} denote the $p \times 1$ covariate vector for unit j in cluster g. To model the correlated event times, we assume that t_{gj} follows a shared frailty PH model given by

$$\lambda(t_{gj} \mid \boldsymbol{x_{gj}}, \mathfrak{b}_g) = \lambda_0(t_{gj})\mathfrak{b}_g \exp(\boldsymbol{x'_{gj}\beta}), \tag{2.1}$$

where $\lambda(t_{gj}|\boldsymbol{x_{gj}}, \mathfrak{b}_g)$ is the conditional hazard function for unit j in cluster g for a given covariate vector $\boldsymbol{x_{gj}}$ and regression coefficients $\beta = (\beta_1, ..., \beta_p)^T$, conditional on the frailty $\mathfrak{b}_g$ specific to cluster g. Given the shared frailty PH model, the conditional survival function is

$$S(t_{gj} \mid \boldsymbol{x_{gj}}, \mathfrak{b}_g) = S_0(t_{gj})^{\mathfrak{b}_g \exp(\boldsymbol{x'_{gj}\beta})}, \tag{2.2}$$

where $S_0(t_{gj})$ denotes the baseline survival function for unit j in cluster g. The $\mathfrak{b}_g$s $g = 1, \dots, G$ are the actual values of a sample of density $h_{\mathfrak{b}}$. This model is a shared frailty model because the individuals in cluster g are assumed to share a common frailty $\mathfrak{b}_g$. The survival times are assumed to be conditionally independent given the shared frailty. Individuals with frailty $\mathfrak{b}_g{<}1$ are said to be less frail and thus will tend to survive longer despite having the same covariate pattern. Conversely, individuals with frailty $\mathfrak{b}_g{>}1$ are more frail, for reasons unexplained by the given covariates, and will have an increased risk of failure. Since $\mathfrak{b}_g$ has a multiplicative effect on the hazard function, it can easily be seen that frailty represents the cumulative effect of one or more omitted covariates (Wienke, 2010). The inclusion of the frailty term in the PH model enables us to obtain estimates of standard errors adjusted for the possible effect of within-cluster correlation (Kim and Xue, 2002).

2.1   Notation

We use f_0 to denote the probability density function corresponding to the baseline survival function $S_0(.)$. The baseline survival and hazard functions can be written in terms of $f_0(.)$

as $S_0(t_{gj}) = 1 - \int_0^{t_{gj}} f_0(s)ds$ and $\lambda_0(t_{gj}) = f_0(t_{gj})/S_0(t_{gj})$.

Assume that the event time t_{gj} is not observed exactly, but that the times of the last negative test result, l_{gj}, and of the first positive test result, r_{gj}, are available. Then, t_{gj} is only known to lie within an interval (l_{gj}, r_{gj}) included in the support (a, b) leading to interval-censored data $\{(l_{gj}, r_{gj}) : j = 1, \dots, n_g; g = 1, \dots, G\}$. All the observations are contained in an interval $(a, t_{cens}) \subset (a, b)$, where t_{cens} denotes the end of the study follow-up. Note that $r_{gj} = b$ when no right limit is available (right-censored).

For the sake of computational simplicity, we use the log-frailty $z_g = \log(\mathfrak{b}_g)$ and shall refer to it as *log-frailty* hereafter. Let z_g have a density $h(z)$ with standard deviation α.

2.2   Flexible specification of the baseline distribution

In order to estimate the baseline survival function $S_0(t_{gj})$ in (2.2), we first approximate the underlying density using penalized B-splines (P-splines). We assume some familiarity with P-splines from the reader. If not, information can be found in Eilers and Marx (1996, 2010); Lang and Brezger (2004).

We start by partitioning (a, t_{cens}) into I (more than 100, say) small bins of equal width Δ_i with midpoints u_i $(i = 1, 2, ...I)$ $(I_i = (a_{i-1}, a_i))$. Consider the B-spline basis $\{b_k(\cdot; q)\}_{k=1}^K$ of degree q associated to a rich grid of K pre-specified equidistant K knots and evaluated at the midpoints u_i of the small bins. The probability π_i of observing t_{gj} in small bin I_i is

$$\pi_i = \int_{I_i} f_0(t)dt \approx f_0(u_i)\Delta_i. \tag{2.3}$$

Following Eilers and Marx (1996), we assume that the π_is change smoothly over time. A possible formulation for $\boldsymbol{\pi} = (\pi_1; \dots; \pi_I)$ is the polytomous logistic regression model

$$\pi_i = \zeta \times \frac{e^{\eta_i}}{e^{\eta_1} + e^{\eta_2} + ... + e^{\eta_I}}, \tag{2.4}$$

where $\eta_i = \sum_k \phi_k b_k(u_i)$, $\boldsymbol{\phi}$ is a vector of spline coefficients and ζ denotes $1 - S_0(t_{cens})$. Given that $\pi_i(\phi) = \pi_i(\phi + c)$ for any constant c, an identifiability constraint is imposed on spline coefficients ϕ_k such that $\sum_i^I \exp(\eta_i) = 1$. The density $f_0(u_i)$ is estimated by π_i / Δ_i using Equations (2.3) and (2.4).

If K is too small, we might potentially miss important features in the density. If it is too large, our density estimate will be too wiggly (i.e. capturing all irrelevant features of the data). In practice, we recommend to take a reasonably large value for K (say between 20 and 30). As we take a generous number of B-splines, it is important to avoid overfitting. Eilers and Marx (1996) suggest handling such an issue by penalizing the log-likelihood using a roughness penalty when drawing inference. In a Bayesian setting, the roughness penalty is translated into a prior distribution for the spline parameters. The details appear in Section 3. P-splines were already used in many different contexts; see for example Eilers and Marx (1996); Eilers (2007) in a frequentist framework and Lang and Brezger (2004); Lambert

and Eilers (2005); Lambert (2007, 2013) and Çetinyürek-Yavuz and Lambert (2011) in a Bayesian framework.

2.3 Flexible specification of the frailty distribution

The conventional approach for frailty models is to assume that $h(z)$ follows a known parametric distribution, for instance a normal distribution. Alternatively, one could choose to work with flexible forms and use the data to estimate the density. In this spirit, we investigate the use of P-splines to specify the frailty distribution along the same lines as in Section 2.2. By standardizing the log-frailties using $z_g^* = \frac{z_g - \nu}{\alpha}$, such that expected value of z_g^*, $E(z_g^*)$, is zero, and variance of z_g^*, $V(z_g^*)$, is 1, it is reasonable to assume that all z_g^* take a value in a standard range, say $\Im = (-6, 6)$. This interval is partitioned into $I^* (> 100$) small bins $I_i^* = (a_{i-1}^*, a_i^*)$ of equal width Δ^*, with midpoints $u_i^*; (i = 1, 2, ...I^*)$. The probability of z_g^* being in small bin I_i^* is $\pi_i^* = \int_{I_i^*} h^*(z^*)dz^{(*)}$, where π^* is modeled as a combination of B-splines, as follows:

$$\pi_i^* = \frac{\exp\left(\eta_i^*\right)}{\exp\left(\eta_1^*\right) + \exp\left(\eta_2^*\right) + ... + \exp\left(\eta_I^*\right)}, \tag{2.5}$$

where $\eta_i^* = \sum_\kappa \phi_\kappa^* b_\kappa^*(u_i^*)$ and $\{b_\kappa^*(\cdot; q)\}_{\kappa=1}^{\kappa^*}$ is the B-spline basis. The same identifiability constraint is imposed on the spline coefficients ϕ^* as on ϕ.

The specification of the log-frailty density differs from the previous one for the baseline density $f_0(.)$ (see Section 2.2). Indeed, the log-frailty terms are not subject to censoring; hence, there is no need for a ζ parameter as in (2.4). For identifiability reasons, the mean of log-frailty z_g is fixed to zero. Moreover, z_g^* is also constrained to have a variance of 1.

3 Inference

3.1 Likelihood

The likelihood contribution of an individual with interval-censored event times (l_{gj}, r_{gj}) is expressed as the difference of the values of survival functions at l_{gj} and r_{gj}. Then, the likelihood of the shared frailty model from interval-censored data is

$$L(\boldsymbol{\phi}, \boldsymbol{\zeta}, \boldsymbol{\beta}, \boldsymbol{z}) = \prod_g \prod_j [S(l_{gj} \mid z_g, x_{gj}) - S(r_{gj} \mid z_g, x_{gj})]. \tag{3.1}$$

The marginal likelihood of the shared frailty model from interval-censored data cannot be evaluated analytically, except in the special case of a gamma frailty distribution Goethals et al. (2009). Therefore, we adopt a Bayesian approach for inference.

3.2   Prior and posterior distributions

The Bayesian approach requires the specification of prior distributions for the model parameters. We start by introducing the prior distributions related to the baseline density specification.

A Bayesian version of the P-spline approach of Eilers and Marx (1996) is presented by Lang and Brezger (2004) for additive models and extensions by replacing difference penalties with their stochastic analogues. In a Bayesian setting, the role of the roughness penalty, τ, is taken over by the prior distribution on the finite (r^{th}) order differences of the spline coefficients as $(\Delta^r \phi_k | \tau) \sim N(0, \tau^{-1})$. As a result, the joint prior for the B-spline coefficients corresponds to a multivariate normal distribution as $p(\boldsymbol{\phi}|\tau) \propto \tau^{K/2} \exp\left\{-\frac{\tau}{2}\boldsymbol{\phi}'\boldsymbol{P}\boldsymbol{\phi}\right\}$, with mean $\boldsymbol{0}$ and variance - covariance matrix $\boldsymbol{P}^{-1}$ where we suggest taking $\boldsymbol{P} = \boldsymbol{D_r'}\boldsymbol{D_r} + \epsilon I$, where D_r is an rth- order difference matrix. The inverse variance τ plays the role of the penalty parameter in the penalized likelihood of the frequentist setting. A noninformative hyperprior with large variance is usually advocated for τ, for example, a gamma distribution $\mathcal{G}(v_1{=}1, v_2{=}0.0001)$ with mean v_1/v_2 and variance v_1/v_2^2 (Lang and Brezger, 2004), leading to almost diffuse priors. Alternative priors are suggested in Jullion and Lambert (2007); Scheipl and Kneib (2009). This idea has been successfully used in many contexts (see e.g. Lambert and Eilers (2009); Lang and Brezger (2004); Lambert (2007)). An improper prior is considered for the regression parameters β. Moreover, a uniform prior on (0,1) is taken for ζ. However, we choose to work with $\xi = \log\left[\zeta/(1 - \zeta)\right]$. The consequent prior density for ξ is thus proportional to $\zeta(1 - \zeta) = \frac{\exp(\xi)}{(1+\exp(\xi))^2}$. In a Bayesian implementation of a frailty model, the log-frailty terms are treated as unknown parameters and assigned a prior $h(z)$. Thanks to the location-scale transformation, one can use equidistant knots on $\mathfrak{I} = (-6, 6)$, where most of the probability mass lies. This allows one to estimate the standard deviation of the log-frailty. Typically, the standard deviation of the log-frailty, α, is assigned a large variance inverse Gamma prior. In the subsequent sections, after presenting the specific prior distributions for each model, the posterior distributions for the frailty are presented in two parts: the parametric frailty and the flexible frailty model formulations.

3.2.1   Parametric frailty model

The log-frailty terms could be assumed, for example, to be independent Gaussian $(z_g|\alpha) \sim N(0, \alpha^2)$. Given the shared frailty PH model and the conditional independence assumption, the joint posterior of the model parameters can be written as

$$p\left(\boldsymbol{\phi}, \tau, \xi, \boldsymbol{\beta}, \boldsymbol{z}, \alpha \mid \mathfrak{D}\right) \propto L(\boldsymbol{\phi}, \boldsymbol{\zeta}, \boldsymbol{\beta}, \boldsymbol{z}) \times p(\boldsymbol{\phi}|\tau) \times p(\tau) \times p(\xi) \times p(\beta)$$
$$\times \left\{\prod_{g=1}^{G} p(z_g|\alpha)\right\} \times p(\alpha), \quad (3.2)$$

given the available data $(\mathfrak{D})$, where $L(\boldsymbol{\phi}, \boldsymbol{\zeta}, \boldsymbol{\beta}, \boldsymbol{z})$ is given by (3.1).

In the parametric frailty model, only the conditional posterior distribution for τ is of a famil-

iar type: $\tau|\mathfrak{D} \sim \mathcal{G}\left(K + \upsilon_1, \left(\phi'_{(m)}P\phi_{(m)}\right)/2 + \upsilon_2\right)$. The conditional posterior distributions for the other model parameters are not of familiar type. Therefore they will be explored using a Markov chain Monte Carlo (MCMC) algorithm (see Section 3.3).

3.2.2   Flexible frailty model

Like for the spline coefficients for the baseline density, the joint prior for the B-spline coefficients corresponds to a multivariate normal distribution $p(\phi^*|\tau^*) \propto (\tau^*)^{\kappa/2} \exp\left\{-\frac{\tau^*}{2}\phi^{*\prime}P^*\phi^*\right\}$, with mean $\mathbf{0}$ and variance-covariance matrix P^{*-1} where $P^* = D^{*\prime}D^* + \epsilon I^*$ is a full-rank matrix for some small quantity ϵ (say 10^{-6}). For τ^*, a noninformative hyperprior (Gamma distribution) with large variance is assumed. The log-frailty terms are assumed to have a nonparametric density, $h^*(z^*)$, which is constructed via (2.5) where $h^*(z^*)$ is a zero mean and unit variance density.

The joint posterior of the model parameters for the shared frailty PH model with a flexible frailty distribution can be written as

$$p\left(\phi, \tau, \xi, \beta, z^*, \phi^*, \tau^*, \alpha \mid \mathfrak{D}\right) \propto L(\phi, \zeta, \beta, z) \times p(\phi|\tau) \times p(\tau) \times p(\xi) \times p(\beta)$$
$$\times \left\{\prod_{g=1}^{G} h^*(z^*|\phi^*)\right\} \times p(\alpha) \times p(\phi^*|\tau^*) \times p(\tau^*) \quad (3.3)$$

where $L(\phi, \zeta, \beta, z)$ is given by (3.1).

Mean zero and unit variance constraint for the frailty distribution is made possible by subtracting a strong penalty of the form $\psi \times [(E(z_g^*) - 0)^2 + (var(z_g^*) - 1)^2]$ from the log-likelihood (where ψ is a large positive quantity, say 1000, see Lambert (2013)).

The conditional posterior distributions for τ and τ^* are of familiar type:

$\tau|\mathfrak{D} \sim \mathcal{G}\left(K + \upsilon_1, \left(\phi'_{(m)}P\phi_{(m)}\right)/2 + \upsilon_2\right)$ and $\tau^*|\mathfrak{D} \sim \mathcal{G}\left(\kappa + \upsilon_1, \left(\phi^*_{(m)}{}'P^*\phi^*_{(m)}\right)/2 + \upsilon_2\right)$. Unfortunately, the conditional posterior distribution of other model parameters are not of familiar type, and hence, will be sampled using MCMC (see Section 3.3).

3.3   Exploring the posterior via MCMC

For both frailty models, parameter ζ is restricted to have a value on [0,1] and α is defined to be positive; therefore the Metropolis steps for updating those parameters are performed after logit and log transformations, respectively.

3.3.1   Parametric frailty model

Let $\boldsymbol{\vartheta} = (\phi, \tau, \xi, \beta, z, \alpha)$ be the vector of parameters of length $H(= K + 1 + 1 + p + G + 1$ and p denote the number of regression parameters. The samples $\left\{\boldsymbol{\vartheta_{(m)}} : m = 1, ..., M\right\}$ will be drawn from the joint posterior $p(\boldsymbol{\vartheta}|\mathfrak{D})$ using a Metropolis-within-Gibbs algorithm.

3.3.2   Flexible frailty model

Let $\boldsymbol{\vartheta} = (\boldsymbol{\phi}, \tau, \xi, \boldsymbol{\beta}, \boldsymbol{z^*}, \boldsymbol{\phi^*}, \tau^*, \alpha)$ be the vector of parameters of length $H(=K + 1 + 1 + p + G + \kappa + 1 + 1)$ and p denote the number of regression parameters. MCMC samples $\{\boldsymbol{\vartheta_{(m)}} : m = 1, ..., M\}$ will be drawn from the joint posterior $p(\boldsymbol{\vartheta}|\mathfrak{D})$ using a Metropolis-within-Gibbs algorithm.

3.3.3   MCMC steps

By running the chains for M iterations (long enough to achieve convergence) and ignoring the first few thousand iterations (say n_b) as an appropriate burn-in period, chains of length $(M - n_b)$ are obtained and thinned (by factor 50 for the spline coefficients due to large auto-correlations). Finally, the resulting $(M - n_b)/50$ iterations can be seen as a random sample from the joint posterior. Point estimates and corresponding credible regions can be calculated based on these random samples.

There are several ways to improve the mixing and the convergence of the chains. Starting the chains at good initial values speeds up convergence. Usually, these can be obtained from restricted frequentist models. Furthermore, the mixing of the chains can be improved by using a Metropolis algorithm on a re-parametrized posterior (Lambert, 2007). For optimal use of the Metropolis algorithm, it is also recommended to tune the acceptance probability to approximately 0.44 in one-dimensional space decreasing to 0.23 in high-dimensional spaces (Gelman et al., 1996; Roberts and Rosenthal, 2001). Good acceptance rates can be achieved via a careful choice of standard deviation δ_h in the generation of proposals in the univariate Metropolis algorithm. For more details on how to improve mixing and convergence, see the Supplement.

4   Simulation study

In order to assess the performance of our modeling strategies, we have performed a series of simulations with various numbers of clusters and cluster sizes under varying interval-censoring schemes. Besides aiming to accurately estimate the regression coefficients and the standard deviation of the log-frailty, we also aim to obtain good estimates of the survival and frailty density functions. In line with these objectives, we considered three scenarios for the frailty density (referred to as "Unimodal", "Bimodal" or "Skewed"; see dashed lines for the assumed frailty distributions in Figure 1). The results are obtained from 300 replications with four different values for the number of clusters G (i.e. 20, 50, 100 and 200), each of which has different sizes n_g (i.e. 4, 6, 10, 20 and 50). It should be noted here that large (resp. small) values of G are combined with small (resp. large) values for n_g. Three different values (0.8, 1.2, and 1.5) were considered for the standard deviation of the frailty. Six different scenarios corresponding to different interval widths (0.5σ, 1.0σ and 1.3σ) and two different amounts of right-censoring (10% and 35%) were used. A balanced binary treatment covariate ($X_1 = 0$ for control subject; $X_1 = 1$ for a subject in

the intervention group) is used with a corresponding regression coefficient chosen as $\log(2)(\cong 0.6931)$. After generating the frailty terms from the specified distribution, given the value of the frailty and the covariate we assumed a PH frailty model with a Weibull baseline hazard, $\lambda_0(t) = 5t^4/70^5$. For each scenario the performances of three estimation strategies named-"Imputation", "Gaussian" and "Semiparametric" are compared. "Imputation" replaces the interval-censored data by the interval midpoints and fits a Cox PH model with a lognormal or gamma frailty. "Gaussian" assumes a flexible baseline survival function for interval-censored data and a lognormal frailty density. "Semiparametric" estimates both the baseline survival and the frailty density using Bayesian P-splines. We considered the compact interval (0,120) as (an approximation to) the support of the target Weibull distribution. The observed range of the considered distribution, $(0, t_{cens})$, changing for different amounts of right-censoring, was divided into small bins of width 1 ($\cong 0.7\sigma$). Cubic B-splines associated to 12 equidistant knots on $(0, t_{cens})$ and a third order penalty were used. A chain of length M=120 000 (including a burn-in period of n_b=54 000 runs) was constructed to explore the posterior distribution of the model parameters.

4.1   Performance in parameter estimation

Although the posterior samples from Bayesian estimation using MCMC contain lots of information, we need to use point estimators to summarize the samples and to enable frequentist comparisons. For this reason we use the mean/median of several quantities of interest. Point estimates for β, ζ, τ and α were calculated using the mean of the generated MCMC sample; $(1-\alpha) \times 100\%$ credible interval can be estimated using the $\alpha/2$ and $(1-\alpha/2)$ sample quantiles of the chain. The proportions of so-defined credible intervals (one for each of the S data sets) containing the true value of the parameter of interest were reported as an estimate of the corresponding coverage. The fitted baseline density $\widehat{f_s}$ for the sth data set corresponds to the MCMC estimate $\frac{1}{M-n_b} \sum_{m=n_b}^{M} \widetilde{f}^{(m)}$ of the posterior mean of the estimated baseline density. The baseline survival is obtained using the estimated baseline density. These quantities can be used to derive a point estimate for the mean, standard deviation and some selected quantiles. Further we report on 5%, 15%, 25%, 35%, 50%, 60%, 75%, 85% and 88% quantiles of the baseline survival function. It should however be noted that for some amounts of right censoring, we cannot get the estimates for baseline survival functions at some quantiles since the density is not observed beyond t_{cens}. The frequentist properties of these Bayesian estimators are measured in terms of relative bias (Rbias), empirical standard errors (ESE) and root mean squared error (RMSE). However only relative biases are reported here. If β (say) is the parameter of interest, then the relative bias is defined as $100\left(\frac{\bar{\bar{\beta}}-\beta}{\beta}\right)$ where $\bar{\bar{\beta}}$ is the mean of the estimates for β over the S data sets.

Only part of the results of the simulation study is presented in Tables 1-4 for the sake of brevity. In Table 1, the results for regression coefficient and standard deviation for frailty are summarized when the generated frailty is unimodal (Left: α=0.8, right: α=1.2). The regression coefficients β are properly estimated (Rbias < 4%) with the flexible and the Gaussian frailty models, yielding similar point estimates and 90% credible intervals (except when n_{cl}=20 with 10 observations per cluster). An increase in sample size results in smaller

credible intervals. Moreover, the coverages of 90% and 80% credible intervals are very close to their nominal values. However, the performance of both frailty models deteriorates in the estimation of the standard deviation of the frailty, α, for small cluster sizes (n_i=4 and 6), or when the number of clusters is small (n_{cl}=20). In all settings for both models, the frailty standard deviation is underestimated with low to medium relative bias (between -2% and -11%). The performances of a setting with a larger standard deviation for the log-frailty (α=1.2) are very similar to the setting with α=0.8 except that the estimation of the standard deviation of the log-frailty improves (Rbias < 5%).

Tables 2 presents the results for the "Bimodal" setting with different standard deviations for the log-frailty (left: α=0.8, right: α=1.2). One can notice that the relative biases in the estimation of the regression coefficients are very small (< 3%), except again when n_{cl}=20 and n_i=10 ($\approx$ 7%). The coverages of 90% and 80% credible intervals are very close to their nominal values. Like in the unimodal setting, the models with flexible frailty and the Gaussian frailty produce very similar regression coefficient estimates. When we consider the estimation of frailty standard deviation, the performances of the flexible and of the Gaussian frailty specifications are similar except when n_i is very small: with four observations per cluster the estimation of α produces large relative bias estimates. A similar scenario with a larger mean interval width (1.0σ) was also considered (not presented here): the larger interval width did not affect the results. However, when considering a similar scenario with mean interval width (1.0σ=15) and a larger frailty standard deviation α=1.2 (right of Table 2), the estimation improves. All relative bias estimates are below 5% and the coverages are similar for the flexible and the Gaussian frailty models. This progress in the estimation of α is probably due to the separability of the two modes of the frailty distribution when α is larger. It is also noticeable that the estimation of α is more performant with a smaller bias (less than 5%). The coverages of 90% and 80% intervals tend to be larger than the nominal values. It should be noted that the performances of the two frailty models are similar although the flexible frailty model has more parameters to estimate. The performance of the models in a setting similar to Table 2 but with a larger frailty standard deviation (α=1.5) was also explored (though it is not presented here). The relative bias is less than 5% for β and less than 6% for α. The coverage for α in the semiparametric model is larger than for the Gaussian model. In the bimodal scenario, the value of the frailty standard deviation is an important factor for the separation of two modes of the mixture distribution. The relative bias and coverages do not change when the amount of right-censoring increases. However, a larger percentage of right-censoring results in larger credible intervals.

The results from more challenging settings, ("Skewed" setting), are presented in Table 3. The performances in the estimation of β are quite similar for the flexible and the Gaussian frailty models with relative biases smaller than 4% in all cases. The relative biases are also very small for α (less than 5%). The coverages of the credible intervals tend to be larger than their nominal values in the flexible frailty specification when the number of clusters is small. As sample size increases, the width of credible intervals decreases and coverages become closer to their nominal values. A similar setting with a greater α=1.5 revealed similar results (not presented here). Another simulation setting similar to Table 3 was also run to see the effect of larger amount of right-censoring (35%): the obtained results (not

shown) are consistent with those in Table 3. A number of sensitivity analyses were also carried out to assess the robustness of the results to the prior distributions for penalty parameters τ, τ^* and standard deviation of log-frailty, α. The results are not sensitive to the choice of hyper-parameters in the "non-informative" prior distributions.

The performances of midpoint imputation using a gamma or a lognormal frailty were also compared for all simulated datasets. Except when the widths of the intervals are small (0.5σ), midpoint imputation should be used with caution. Indeed, for a mean interval width 1.0σ, midpoint imputation resulted in relative bias of around 5-10% for the regression coefficient, β, and the frailty standard deviation, α. When the mean interval width increases, the relative biases also increase for both model parameters. Assuming a gamma or a lognormal does not make any difference in the estimation of the regression coefficient, while assuming a lognormal frailty seems preferable when it comes to estimating the standard deviation of the log-frailty. The results are not presented here for the sake of brevity.

To sum up all results so far, we can conclude that the flexible frailty and the Gaussian frailty models perform similarly in the estimation of β. Their performance in the estimation of α depends on different factors (sample size, number of clusters and frailty variance). However, the deterioration in the estimation of α does not impact upon the estimation of β. Therefore, if one is only interested in the estimation of regression coefficients and does not hope to obtain insights from the shape of the frailty distribution, a shared frailty PH model with a flexible baseline hazard and a parametric frailty distribution is probably an adequate choice.

4.2   Curve fitting performance

One of our aims is to investigate the ability of the proposed flexible frailty model to estimate the shape of the frailty density. The quality of curve fit is explored for different numbers of clusters, cluster sizes, and interval widths, different amounts of right-censoring and three shapes for the frailty density (see the dashed curves in each of the columns in Figure 1). For each frailty scenario, we present here only five different settings for the sake of brevity; see Figure 1. From each simulation setting, we simulate 50 datasets and estimate the frailty density. The first row in the figure corresponds to a setting with 20 clusters and 10 observations per cluster (n=200). The third row comes from a setting with n=200 and 50 clusters. The second, fourth and fifth rows correspond to settings with a sample size of 1000 with increasing number of clusters (n_{cl}=20 and n_g=50; n_{cl}=50 and n_g=20; n_{cl}=100 and n_g=10, respectively).

Under the Gaussian frailty scenario (column 1 of Figure 1), the features of the frailty density are properly captured by the flexible frailty model. Under more challenging scenarios (column 2: bimodal; column 3: skewed), the ability of the flexible model (in capturing the features of the frailty density) depends on the number of clusters, cluster sizes and frailty standard deviation. When the number of clusters is small (n_{cl}=20) or when the number of data items per cluster is insufficient (n_g=4), the shape of the frailty density cannot be properly estimated. It improves with the number of clusters, with satisfactory results even
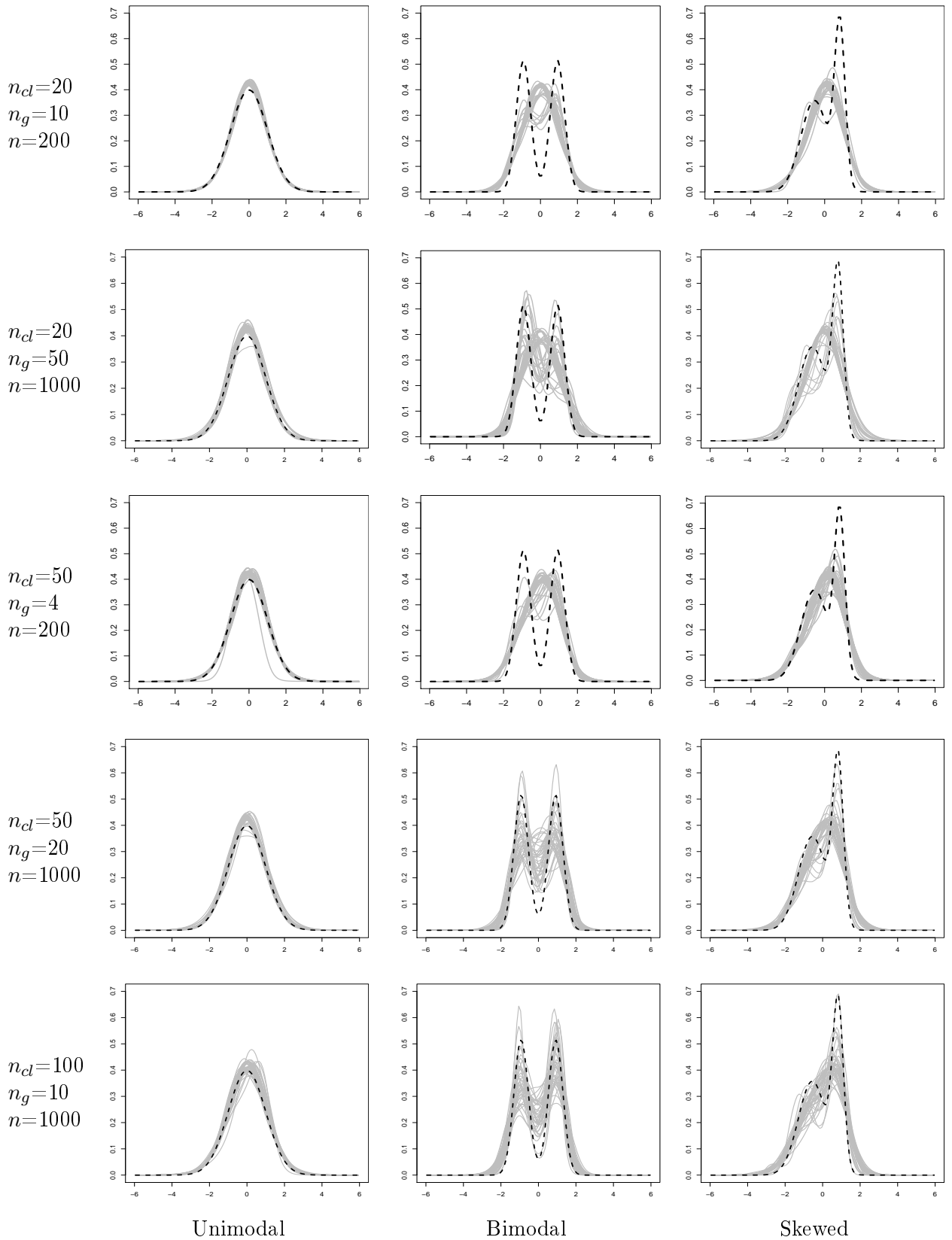
Figure 1: Estimated frailty density aggregated over 50 replications under five different settings for three scenarios

in the challenging "bimodal" frailty scenario (when n_{cl}=50). When the number of clusters and cluster size are small, the estimated frailty density resembles a unimodal normal distribution. This is a consequence of the 3rd order difference penalty used in the estimation of the frailty density. Even with a large number of clusters, a cluster size of less than six might yield a standard normal density. For an accurate representation of the frailty density one needs a large number of clusters of medium to large sizes. The bimodal nature of the frailty density has been well captured with large sample sizes, the largest uncertainty occurring around the modes. It is important to note that the regularization of the estimated frailty density to a Gaussian distribution (for small sample sizes) does not distort the estimation of the regression coefficients: the regression coefficient estimates remained robust to misspecification of the frailty density.

The baseline survival function is properly approximated in many scenarios. We also found that the estimations of the baseline survival densities in the shared frailty PH model are similar to Çetinyürek-Yavuz and Lambert (2011) for comparable settings (ignoring frailty model parameters). However, the performance of the baseline survival functions is affected by the number of clusters and cluster sizes. When the cluster sizes are small, then the estimates of the survival curve at selected quantiles tend to be more biased. These results are presented partially for the sake of brevity (only from the flexible model). In many simulation settings, semiparametric and lognormal frailty specifications resulted in very similar estimates for the survival function.

5 Application: Signal Tandmobiel® Study

We illustrate the aforementioned methodology with data from the Signal Tandmobiel® Study. The data used here is available in the R package *bayesSurv* and contains mainly the information on emergence timing and caries experience reported as interval-censored observations with intervals of approximate length one year. For more details on the design of the study, see Vanobbergen et al. (2000). Adequate knowledge of timing and patterns of tooth emergence is useful for diagnosis and treatment planning in pediatric dentistry and orthodontics. It is anticipated that boys and girls have different tooth emergence distributions. For this reason, the covariate gender (0=boys, 1=girls) was included in the model. Additionally it was of interest to dentists to see whether the distribution of emergence time of a permanent tooth changes when its primary predecessor experiences or does not experience caries. Therefore, a binary score, dmf, was associated to each permanent tooth to indicate whether it was preceded by a decayed primary tooth. The response variable for a particular child reports age (in years) at emergence of each premolar (teeth 14, 15, 24, 25, 34, 35, 44, 45): the data will be analyzed jointly and treated as clustered data. This enables us not only to quantify the impact of different covariates on the emergence time, but also to study the relationship between the emergence times of different teeth.

A random sample of 150 boys and 150 girls is used for inference, resulting in a dataset of n=2400 with $n_{cl} = 300$ clusters. For a better fit, we shifted the time origin to five years
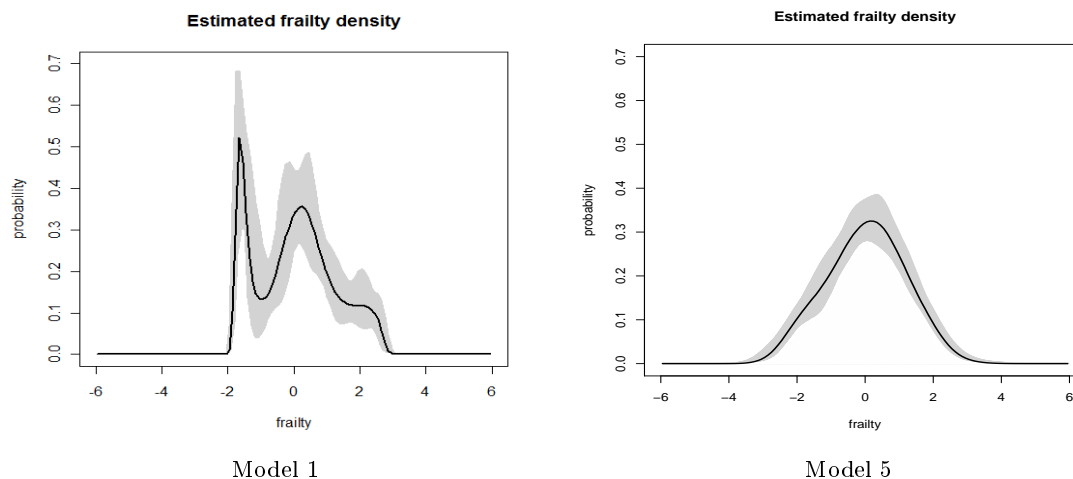
Figure 2: Signal Tandmobiel® data: Estimated frailty density

of age which is clinically the minimal emergence time for permanent teeth. We replaced t_{gj} by t_{gj}-5 in the model specification. We used 15 knots from -6 to 6 and a third order difference penalty. All computations are performed using an R code that calls a Fortran function for the MCMC. The joint posterior for the flexible frailty model was sampled using MCMC, yielding a chain length of 154 000 thinned by 50:1 ratio. Sampling of the 350 parameters required approximately 7 hours. Five frailty models were fitted to the data: The first one includes only *dmf*, the second *dmf* and *gender*, and the third *dmf*, *gender* and their interaction. The fourth model extends the third model by including the *horizontal symmetry* and its interaction with *gender*. It has been shown by Leroy et al. (2003) that there is horizontal symmetry with respect to teeth emergence: the same emergence distributions can be assumed for horizontally symmetric positions. Thus three dummy variables were created: *Man4*, *Max5* and *Man5*, for mandibular first premolars (teeth 34,44), maxillary second premolars (teeth 15,25) and mandibular second premolars (teeth 35,45), respectively. The fourth model includes the main effects and their interactions of *Man4*, *Max5* and *Man5* with *gender*, respectively. For the first four models, the estimated frailty density was bimodal or trimodal (see Fig. 2). This suggested that important covariates were missing. Finally, in the fifth model the interactions of *dmf* with *Man4*, *Max5* and *Man5* and the three-way interactions of *gender*, *dmf* and *horizontal symmetry* were also included. The estimated frailty density for the fifth model turned out to be unimodal. For the sake of brevity, we only present the estimates from Model 1 and 5. The estimated survival curves based on model 5 are shown in Figure 3. As the main interest in this analysis was the effect of dmf on emergence, we provide the plots of the estimated survival curves for dmf=0 and dmf> 0 for boys and girls and the four pairs of horizontally symmetric teeth based on Model 5 (see Figure 3). The figures show that the difference between children with dmf=1 (> 0) and dmf=0 is larger for boys than for girls and that the emergence process starts later for boys. Moreover, subjects who have had caries on the primary predecessor significantly tend to have the permanent successor earlier for maxillary teeth. It can also be noticed that

there is almost no difference for mandibular teeth between teeth with and without caries on the primary predecessor. Usually, the emergence process is late for second premolars when compared to first premolars. Our conclusions and figures are in line with Komárek and Lesaffre (2007) although we use a completely different modeling approach to analyze the data. The estimates from a naive shared frailty PH model using midpoint imputation and assuming a lognormal frailty is also presented in Table 5. The 90% confidence intervals are relatively similar to the 90% credible intervals for most of the regression coefficients in Model 5, while the estimates of the regression coefficients are smaller than with the proposed models. It is also noticeable that the width of the intervals are smaller in the naive approach, based on mid-point imputation.

The convergence of the chains was checked using several diagnostic measures (Cowles and Carlin, 1996). The traces of all model parameters were examined critically. Moreover, autocorrelation coefficients was also computed and plotted. Geweke and Heidelberger-Welch diagnostics also support convergence(Geweke, 1992; Heidelberger and Welch, 1983). We also ran five chains with overdispersed starting points to check the convergence(Gelman and Rubin, 1992). The Gelman-Rubin statistics were found to be at most 1.01 for all β (15 coefficients in total), 1.00 for ζ, and 1.01 for α. These checks suggest convergence of the MCMC algorithm. More details on the data are available in the Supplement.

6　Discussion

In this paper, we have presented rather general Bayesian approaches to address an important class of models, the shared frailty PH model for clustered interval-censored data. Inference ignoring the correlation between observations can be misleading. The main consequence of ignoring the frailty is a reduction in the standard errors, which might lead to wrongly significant findings. The contribution of this paper revolves around building flexible Bayesian models for clustered interval-censored data.

On the other hand, it has been indicated in the literature in different contexts that the misspecification of the random effect distribution can influence the estimation of quantities of primary interest, like the fixed effects. To circumvent such misspecification, we have suggested modeling the distribution of the frailty in a flexible way using P-splines. The biggest advantage of using a flexible specification for the density of frailty arise when its shape is of specific interest. If it is considered as a nuisance, assuming a simpler lognormal frailty would be an adequate solution to draw conclusions related to other model parameters, such as regression coefficients and variance of the frailty. Indeed, it was shown in the simulation study that the regression parameter estimates in a shared frailty PH model are robust to the misspecification of the frailty density. However the use of a flexible form for the frailty does not cause any loss in the precision of the regression parameter estimates when compared to the simpler parametric frailty model. Both models provide a possibility of visualizing the baseline density and survival functions.
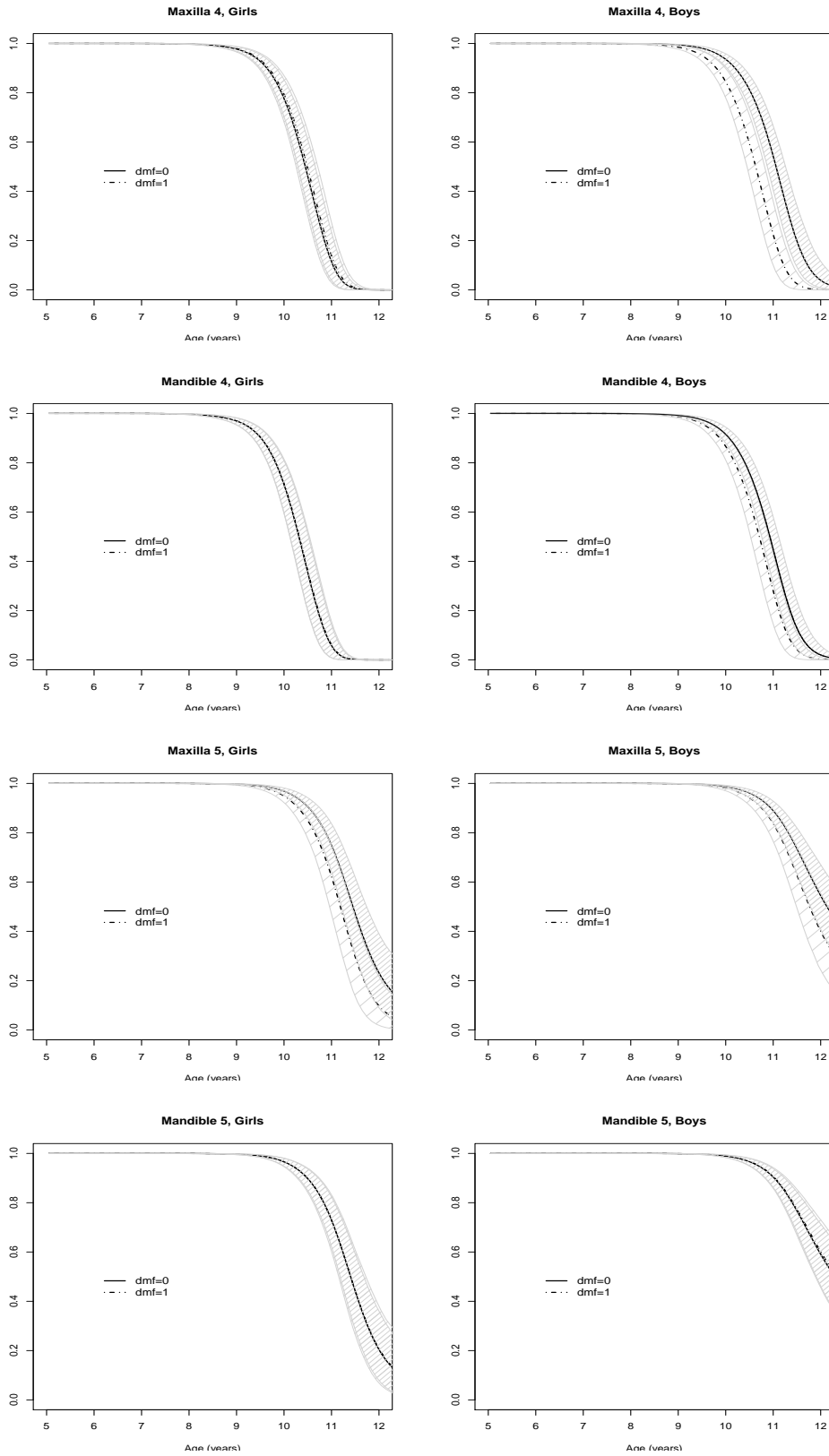
Figure 3: Signal Tandmobiel® Study data: Estimated survival curves based on Model 5 - dmf=0 versus dmf=1

Acknowledgement

References

Aalen, O.O. (1988). Heterogeneity in Survival Analysis. *Statistics in Medicine*, **7**, 1121–1137.

Agresti, A., Caffo, B., and Ohman-Strickland, P. (2004). Examples in which misspecifications of a random effects distribution reduces efficiency, and possible remedies. *Computational Statistics and Data Analysis*, **47**, 639–653.

Bellamy, H.W., Li, Y., Ryan, L.M., Lipsitz, S., Canner, M.J., and Wright, R. (2004). Analysis of clustered and interval-censored data from a community-based Study in asthma. *Statistics in Medicine*, **23**, 3607–3621.

Cai, T. and Betensky, R.A. (2003). Hazard regression for interval-censored data with penalized spline. *Biometrics*, **59**, 570–579.

Çetinyürek-Yavuz, A. and Lambert, P. (2011). Smooth estimation of survival functions and hazard ratios from interval-censored data using Bayesian penalized B-splines. *Statistics in Medicine*, **30**, 75–90.

Chen, J., Zhang, D., and Davidian, M. (2002). A Monte Carlo EM algorithm for generalized linear mixed models with flexible random effects distribution. *Biostatistics*, **3(3)**, 347–360.

Clayton, D.G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*, **65**, 141–151.

Clayton, D. and Cuzick, J. (1985). The semi-parametric Pareto model for regression analysis of survival times. *Proceedings of the Centenary Session of the International Statistical Institute*, **47**, 467–485.

Cowles, M.K. and Carlin, B.P. (1996). Markov Chain Monte Carlo Convergence Diagnostics: A Comparative Review. *Journal of the American Statistical Association*, **91**(434), 883–904.

Eilers, P.H.C. and Marx, B.D. (1996). Flexible smoothing with B-splines and penalties (with comments and rejoinder). *Statistical Science*, **11**, 89–121.

Eilers, P.H.C. (2007). Ill-posed problems with counts, the composite link model and penalized likelihood *Statistical Modelling*, **7(3)**, 239–254.

Eilers, P.H.C. and Marx, B.D. (2010). Splines, knots, and penalties. *WIREs Comp Stat*, **2**, 637–653.

Finkelstein, D.M. (1986). A proportional hazards model for interval-censored failure time data. *Biometrics*, **42**, 845–854.

Gelman, A., Roberts, G.O., Gilks, W.R. (1996). Efficient Metropolis jumping rules. *Bayesian Statistics*, **5**, 599–607.

Gelman, M.Y. and Rubin, V.G. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, **7**(4): 457–472.

Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to calculating posterior moments. *In Bayesian Statistics 4* (ed JM Bernado, JO Berger, AP Dawid and AFM Smith). Clarendon Press, Oxford, UK. : 169–193.

Goethals, K., Ampe, B., Berkvens, D., Laevens, H., Janssen, P., and Duchateau, L. (2009). Modeling interval-censored, clustered cow udder quarter infection times through the shared gamma frailty model. *Journal of Agricultural Biological and Environmental Statistics*, **26**, 769–781.

Goggins, W.B., and Finkelstein, D.M. (2000). A proportional hazards model for multivariate interval-censored failure time data. *Biometrics*, **56**, 940–943.

Heidelberger P, Welch PD. (1983). Simulation run length control in the presence of an initial transient. *Operations Research* **31**: 1109–1144.

Henschel, V., Engel, J., Hölzel,D., and Mansmann,U. (2009). A semiparametric Bayesian proportional hazards model for interval censored data with frailty effects. *BMC Medical Research Methodology*, **9**, 9.

Hougaard, P. (1984). Life table methods for heterogeneous populations: distributions describing the heterogeneity. *Biometrika*, **71**, 75–84.

Hougaard, P. (1986). Survival models for heterogeneous populations derived from stable distributions. *Biometrika*, **73**, 387–396.

Hougaard, P. (2000). *Analysis of multivariate survival data*. Springer, New York.

Jullion, A. and Lambert, P. (2007). Robust specification of roughness penalty prior distribution in spatially adaptive Bayesian P-splines model. *Computational Statistics and Data Analysis*, **51**, 2542–2558.

Kim, M. and Xue, X. (2002). The analysis of multivariate interval-censored survival data. *Statistics in Medicine*, **21**, 3715–3726.

Klein, J.P., Moeschberger, M.L., Li, Y.H., and Wang, S.T. (1992). Estimating random effects in the Framingham heart study. *Survival Analysis: State of the Art. Kluwer Academic, Boston,MA*, **211**, 99–120.

Komárek, A., Lesaffre, E., and Hilton, J.F. (2005). Accelerated failure time model for arbitrarily censored data with smoothed error distribution. *Journal of Computational and Graphical Statistics*, **14(3)**, 726–745.

Komárek, A. and Lesaffre, E. (2007). Bayesian accelerated failure time model for correlated censored data with a normal mixture as an error distribution. *Statistica Sinica*, **17**, 549–569.

Komárek, A. and Lesaffre, E. (2008). Bayesian accelerated failure time model with multivariate doubly interval-censored data and flexible distributional assumptions. *Journal of American Statistical Association*, **103(482)**, 523–533.

Kor, C.T., Cheng, K.F. and Chen, Y.H. (2013). A method for analyzing clustered interval-censored data based on Cox's model. *Statistics in Medicine*, **32**, 822–832.

Lam, K.F., Xu, Y., Cheung, T.L. (2010). A multiple imputation approach for clustered interval-censored survival data. *Statistics in Medicine*, **29**, 680–693.

Lambert,P. and Eilers, P.H.C. (2005). Bayesian proportional hazards model with time varying regression coefficients: a penalized Poisson regression approach. *Statistics in Medicine* 2005, **24**: 3977–3989.

Lambert, P. (2007). Archimedean copula estimation using Bayesian splines smoothing techniques. *Computational Statistics and Data Analysis*, **51**, 6307–6320.

Lambert, P. and Eilers, P.H.C. (2009). Bayesian density estimation from grouped continuous data. *Computational Statistics and Data Analysis*, **53**, 1388–1399.

Lambert, P. (2013). Nonparametric additive location-scale models for interval censored data. *Statistics and Computing*, **23(1)**, 75–90.

Lang, S. and Brezger, A. (2004). Bayesian P-splines. *Journal of Computational and Graphical Statistics*, **13**, 183–212.

Leroy, R., Bogaerts, K., Lesaffre, E. and Declerck, D. (2003). The emergence of permanent teeth in Flemish children. *Community Dentistry and Oral Epidemiology*, **31**, 30–39.

Lesaffre, E., Komárek, A., and Declerck, D. (2005). An overview of methods for interval-censored data with an emphasis on applications in dentistry. *Statistical Methods in Medical Research*, **14**, 539–552.

Peto, R. (1973). Experimental survival curves for interval-censored data. *Journal of the Royal Statistical Society (Series C)*, **22**, 86–91.

R Development Core Team (2012). *R: A Language and Environment for Statistical Computing*. Vienna, Austria. URL http://www.R-project.org/.

Roberts, G.O. and Rosenthal, J.S. (2001). Optimal scaling for various Metropolis- Hastings algorithms. *Statistical Science*, **16**, 351–367.

Scheipl, F. and Kneib, T. (2009). Locally adaptive Bayesian P-splines with a normal-exponential-gamma prior. *Computational Statistics and Data Analysis*, **53**, 3533–3552.

Shih, J.H. and Louis, T.A. (1995). Assessing gamma frailty models for clustered failure time data. *Lifetime Data Analysis*, **1**, 205–220.

Therneau, T. (2012). survival: Survival Analysis, Including Penalised Likelihood. *R package version 2.36-14*, URL http://CRAN.R-project.org/package=survival.

Turnbull, B.W. (1976). The empirical distribution function with arbitrarily grouped, censored, and truncated data. *Journal of the Royal Statistical Society, Series B (Methodological)*, **38**, 290–295.

Vanobbergen, J., Martens, L., Lesaffre, E., and Declerck, D. (2000) The Signal-Tandmobiel project-a longitudinal intervention health promotion study in Flanders (Belgium): baseline and first year results. *European Journal of Paediatric Dentistry*, **2**, 87–96.

Vaupel, J.W., Manton, K.G. and Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, **16**, 439–454.

Vaupel, J.W. and Yashin, A.I. (1983). The deviant dynamics of death in heterogeneous populations. *RR-83-1*, International Institute for Applied System Analysis.

Wen, C.C. and Chen, Y.H. (2013). A frailty model approach for regression analysis of bivariate interval-censored survival data. *Statistica Sinica*, **16**, 439–454.

Wienke, A. (2010). *Frailty models in survival analysis*. Chapman & Hall, CRC Press, New York.

Zhang, M. and Davidian, M. (2008). Smooth semiparametric regression for arbitrarily censored time-to-event data. *Biometrics*, **64**(2), 567–576.

Zuma, K. and Lurie, M. (2005). Application and comparison of methods for analysing correlated interval-censored data from sexual partnerships. *Journal of Data Science*, **3**, 241–256.

Zuma, K. (2007). A Bayesian analysis of correlated interval-censored data. *Communications in Statistics - Theory and Methods*, **36**, 725–730.

Table 1: The mean, relative bias (Rbias in %), 90% credible intervals and corresponding empirical coverages (EC) for β and α with 10% of right-censoring and a mean width of 0.5σ for a varied number of clusters and cluster sizes under a "Unimodal" frailty density with a standard deviation of α=0.8 (left) and α=1.2 (right) in S=300 replications

Left panel (α=0.8)

β=0.693

n_{cl}	n_g	N	Semiparametric Mean	90% CI	Rbias (%)	EC 90	EC 80	Gaussian Mean	90% CI	Rbias (%)	EC 90	EC 80
20	10	200	0.75	(0.47-1.06)	7.7	88	74	0.75	(0.49-1.08)	7.8	88	75
	20	400	0.70	(0.51-0.90)	0.8	89	79	0.70	(0.51-0.90)	0.8	90	79
	50	1000	0.70	(0.58-0.83)	0.7	90	80	0.70	(0.58-0.83)	0.7	89	80
50	4	200	0.72	(0.43-1.02)	3.7	88	80	0.72	(0.43-1.02)	3.8	88	82
	6	300	0.70	(0.48-0.94)	0.8	89	81	0.70	(0.48-0.95)	1.1	90	79
	10	500	0.69	(0.53-0.87)	-0.3	91	82	0.69	(0.54-0.87)	-0.1	90	82
	20	1000	0.70	(0.59-0.81)	0.7	92	82	0.70	(0.60-0.81)	0.7	92	82
100	4	400	0.71	(0.53-0.90)	2.1	92	78	0.71	(0.53-0.90)	2.6	91	77
	6	600	0.70	(0.55-0.85)	0.4	91	81	0.70	(0.55-0.86)	0.6	92	80
	10	1000	0.69	(0.58-0.80)	-0.4	92	84	0.69	(0.58-0.80)	-0.1	93	84
200	4	800	0.70	(0.56-0.84)	1.1	90	77	0.70	(0.56-0.85)	1.6	89	79
	6	1200	0.68	(0.59-0.77)	-2.1	89	77	0.69	(0.61-0.78)	-0.3	91	80
	10	2000	0.68	(0.60-0.78)	-1.4	89	76	0.68	(0.59-0.78)	-1.2	89	79

α=0.8

n_{cl}	n_g	N	Semiparametric Mean	90% CI	Rbias (%)	EC 90	EC 80	Gaussian Mean	90% CI	Rbias (%)	EC 90	EC 80
20	10	200	0.73	(0.47-1.06)	-8.1	84	77	0.73	(0.49-1.08)	-7.6	83	75
	20	400	0.76	(0.52-1.03)	-4.5	89	82	0.76	(0.53-1.02)	-4.7	87	78
	50	1000	0.76	(0.52-1.00)	-5.1	88	78	0.76	(0.53-0.99)	-5.5	86	76
50	4	200	0.72	(0.20-1.02)	-9.8	88	78	0.72	(0.21-1.02)	-9.7	88	75
	6	300	0.74	(0.55-0.93)	-7.4	87	79	0.74	(0.54-0.93)	-7.3	85	78
	10	500	0.76	(0.57-0.95)	-4.8	88	78	0.77	(0.60-0.95)	-3.4	87	75
	20	1000	0.78	(0.65-0.94)	-2.5	94	82	0.78	(0.64-0.94)	-2.8	90	77
100	4	400	0.76	(0.61-0.93)	-5.0	91	76	0.77	(0.62-0.93)	-3.8	91	76
	6	600	0.77	(0.65-0.91)	-3.3	92	82	0.78	(0.65-0.91)	-2.5	91	82
	10	1000	0.77	(0.65-0.90)	-3.3	89	79	0.79	(0.67-0.90)	-1.6	90	80
200	4	800	0.77	(0.63-0.87)	-4.2	88	80	0.78	(0.64-0.89)	-2.3	91	81
	6	1200	0.77	(0.68-0.87)	-3.6	89	81	0.78	(0.69-0.88)	-1.9	92	82
	10	2000	0.78	(0.70-0.86)	-2.5	91	83	0.79	(0.70-0.87)	-1.6	92	81

Right panel (α=1.2)

β=0.693

n_{cl}	n_g	N	Semiparametric Mean	90% CI	Rbias (%)	EC 90	EC 80	Gaussian Mean	90% CI	Rbias (%)	EC 90	EC 80
20	10	200	0.71	(0.43-0.99)	2.8	90	77	0.71	(0.44-0.99)	3.0	90	78
	20	400	0.70	(0.50-0.89)	1.7	89	78	0.71	(0.50-0.89)	1.8	88	79
	50	1000	0.69	(0.58-0.81)	-0.1	91	83	0.69	(0.58-0.81)	0.1	91	83
50	4	200	0.73	(0.43-1.03)	5.7	88	77	0.74	(0.44-1.04)	6.1	87	76
	6	300	0.71	(0.49-0.93)	1.9	91	84	0.71	(0.49-0.93)	2.1	93	83
	10	500	0.70	(0.54-0.88)	0.8	91	82	0.70	(0.55-0.87)	0.9	92	82
	20	1000	0.70	(0.58-0.83)	1.4	87	81	0.70	(0.58-0.83)	1.5	88	81
100	4	400	0.72	(0.54-0.93)	3.0	90	81	0.72	(0.54-0.94)	3.5	89	81
	6	600	0.71	(0.56-0.84)	2.4	93	85	0.71	(0.56-0.84)	2.8	93	85
	10	1000	0.69	(0.58-0.82)	-0.1	91	82	0.69	(0.58-0.82)	0.1	91	81
200	4	800	0.69	(0.53-0.84)	-0.5	89	82	0.69	(0.54-0.84)	-0.1	89	79
	6	1200	0.70	(0.59-0.85)	1.2	90	80	0.70	(0.59-0.85)	1.4	88	80
	10	2000	0.70	(0.60-0.78)	0.3	88	80	0.70	(0.60-0.78)	0.4	89	80

α=1.2

n_{cl}	n_g	N	Semiparametric Mean	90% CI	Rbias (%)	EC 90	EC 80	Gaussian Mean	90% CI	Rbias (%)	EC 90	EC 80
20	10	200	1.16	(0.77-1.60)	-3.3	88	78	1.15	(0.76-1.56)	-3.9	85	74
	20	400	1.16	(0.79-1.55)	-3.2	90	77	1.15	(0.79-1.52)	-4.0	85	73
	50	1000	1.17	(0.81-1.50)	-2.8	89	80	1.15	(0.80-1.48)	-4.1	88	75
50	4	200	1.16	(0.79-1.50)	-3.7	86	76	1.15	(0.80-1.49)	-3.9	84	73
	6	300	1.15	(0.89-1.40)	-4.0	89	79	1.15	(0.90-1.39)	-4.5	87	77
	10	500	1.18	(0.97-1.41)	-1.5	94	88	1.17	(0.95-1.39)	-2.5	91	82
	20	1000	1.20	(1.01-1.44)	0.1	93	84	1.19	(1.00-1.43)	-1.0	89	78
100	4	400	1.18	(0.93-1.41)	-1.9	89	79	1.19	(0.95-1.41)	-1.2	87	76
	6	600	1.18	(1.00-1.38)	-1.3	92	82	1.18	(1.01-1.39)	-1.3	90	81
	10	1000	1.20	(1.04-1.38)	0.1	90	82	1.19	(1.04-1.37)	-0.4	90	78
200	4	800	1.17	(1.01-1.35)	-2.5	89	76	1.19	(1.03-1.35)	-1.0	86	76
	6	1200	1.19	(1.06-1.34)	-0.9	91	84	1.20	(1.08-1.34)	-0.1	91	77
	10	2000	1.19	(1.08-1.31)	-0.7	91	85	1.19	(1.08-1.33)	-0.5	90	81

Table 2: The mean, relative bias (Rbias in %), 90% credible intervals and corresponding empirical coverages (EC) for β and α with a mean width of 0.5σ (left) and 1.0σ (right) for a varied number of clusters and cluster sizes under a "Bimodal" frailty density with standard deviation of α=0.8 (left) and α=1.2 (right) in S=300 replications

α=0.8 (mean width 0.5σ)

n_{cl}	n_y	N	Semiparametric β=0.693					Semiparametric α=0.8					Gaussian β=0.693					Gaussian α=0.8				
			Mean	90% CI	Rbias (%)	EC 90	EC 80	Mean	90% CI	Rbias (%)	EC 90	EC 80	Mean	90% CI	Rbias (%)	EC 90	EC 80	Mean	90% CI	Rbias (%)	EC 90	EC 80
20	10	200	0.74	(0.47-1.02)	6.6	90	80	0.74	(0.33-1.01)	-8.1	89	82	0.74	(0.47-1.02)	6.7	88	80	0.75	(0.53-0.98)	-5.8	93	86
	20	400	0.70	(0.54-0.85)	0.3	94	85	0.77	(0.60-0.93)	-3.8	97	92	0.70	(0.53-0.85)	0.5	94	85	0.77	(0.60-0.92)	-3.9	97	91
	50	1000	0.69	(0.57-0.81)	0.1	91	78	0.77	(0.63-0.90)	-3.7	99	96	0.69	(0.57-0.80)	0.1	90	78	0.77	(0.65-0.90)	-4.2	99	96
50	4	200	0.71	(0.42-0.99)	1.9	88	79	0.59	(0.26-0.95)	-25.8	60	54	0.71	(0.44-0.99)	2.6	89	79	0.73	(0.40-0.96)	-8.6	91	77
	6	300	0.70	(0.46-0.91)	0.8	91	81	0.73	(0.29-0.93)	-8.2	86	80	0.71	(0.48-0.90)	1.1	91	81	0.77	(0.61-0.93)	-3.4	95	86
	10	500	0.71	(0.54-0.87)	2.0	91	81	0.79	(0.66-0.92)	-1.8	97	91	0.70	(0.54-0.86)	1.9	90	81	0.79	(0.66-0.92)	-1.2	97	91
	20	1000	0.70	(0.57-0.82)	0.4	89	81	0.79	(0.69-0.90)	-1.4	95	89	0.69	(0.57-0.82)	0.2	88	81	0.80	(0.69-0.90)	-0.6	98	91
100	4	400	0.69	(0.52-0.91)	0.1	89	83	0.63	(0.29-0.92)	-21.5	64	60	0.70	(0.53-0.92)	1.2	91	83	0.78	(0.63-0.93)	-2.1	92	84
	6	600	0.70	(0.54-0.86)	1.0	88	77	0.78	(0.64-0.90)	-2.8	91	85	0.70	(0.54-0.86)	0.8	88	77	0.79	(0.66-0.91)	-0.7	93	87
	10	1000	0.70	(0.57-0.83)	1.5	86	78	0.79	(0.69-0.89)	-1.2	92	82	0.70	(0.57-0.83)	0.9	88	78	0.80	(0.70-0.90)	0.1	95	86
200	4	800	0.69	(0.58-0.81)	0.1	89	80	0.63	(0.38-0.88)	-21.0	54	48	0.69	(0.58-0.80)	-0.5	91	80	0.81	(0.71-0.89)	0.6	93	82
	6	1200	0.69	(0.57-0.80)	-0.5	90	80	0.79	(0.72-0.86)	-1.1	95	87	0.68	(0.56-0.79)	-1.3	89	79	0.80	(0.72-0.87)	0.4	96	92
	10	2000	0.69	(0.61-0.78)	0.2	89	79	0.79	(0.73-0.86)	-1.3	90	81	0.69	(0.58-0.78)	-0.3	91	80	0.80	(0.75-0.87)	0.5	97	92

α=1.2 (mean width 1.0σ)

n_{cl}	n_y	N	Semiparametric β=0.693					Semiparametric α=1.2					Gaussian β=0.693					Gaussian α=1.2				
			Mean	90% CI	Rbias (%)	EC 90	EC 80	Mean	90% CI	Rbias (%)	EC 90	EC 80	Mean	90% CI	Rbias (%)	EC 90	EC 80	Mean	90% CI	Rbias (%)	EC 90	EC 80
20	10	200	0.72	(0.44-1.00)	4.3	93	82	1.18	(0.85-1.53)	-1.7	95	90	0.72	(0.45-1.01)	4.4	93	82	1.17	(0.88-1.50)	-2.4	95	91
	20	400	0.70	(0.50-0.90)	0.9	93	82	1.16	(0.91-1.40)	-3.5	97	92	0.70	(0.50-0.90)	0.8	92	82	1.15	(0.91-1.39)	-4.1	97	90
	50	1000	0.70	(0.58-0.84)	1.3	91	78	1.16	(0.95-1.36)	-3.4	97	94	0.70	(0.58-0.84)	1.3	91	78	1.16	(0.96-1.36)	-3.7	99	94
50	4	200	0.69	(0.41-1.02)	-0.3	90	81	1.21	(0.93-1.54)	1.1	94	86	0.69	(0.41-1.02)	-0.3	91	83	1.21	(0.93-1.53)	1.1	94	83
	6	300	0.71	(0.48-0.95)	2.0	81	82	1.20	(0.97-1.44)	0.3	95	87	0.71	(0.47-0.95)	1.5	82	81	1.20	(0.98-1.42)	0.3	95	89
	10	500	0.70	(0.51-0.88)	1.7	91	91	1.19	(1.03-1.38)	-0.8	97	91	0.70	(0.50-0.88)	1.3	90	91	1.19	(1.04-1.40)	0.2	98	93
	20	1000	0.70	(0.58-0.84)	0.8	90	91	1.16	(1.02-1.32)	-3.1	97	85	0.70	(0.51-0.91)	0.3	88	91	1.16	(1.04-1.35)	-0.9	98	92
100	4	400	0.70	(0.50-0.91)	1.3	93	83	1.22	(1.01-1.40)	1.3	93	85	0.70	(0.50-0.91)	0.7	91	83	1.23	(1.05-1.43)	2.8	93	86
	6	600	0.69	(0.50-0.87)	-0.4	87	77	1.19	(1.02-1.37)	-0.4	90	80	0.69	(0.51-0.85)	-1.6	90	77	1.23	(1.06-1.39)	2.3	93	85
	10	1000	0.70	(0.56-0.84)	0.6	87	78	1.18	(1.03-1.33)	-2.0	81	74	0.70	(0.56-0.82)	-0.2	90	78	1.21	(1.10-1.35)	1.1	97	92
200	4	800	0.69	(0.57-0.81)	1.5	77	77	1.22	(1.04-1.42)	2.0	90	80	0.69	(0.57-0.81)	-0.2	92	80	1.26	(1.13-1.39)	4.7	83	72
	6	1200	0.69	(0.59-0.80)	0.1	88	79	1.17	(1.04-1.27)	-2.4	93	79	0.69	(0.59-0.80)	-1.2	90	81	1.23	(1.15-1.33)	3.2	92	81
	10	2000	0.70	(0.62-0.78)	0.3	90	82	1.17	(1.07-1.26)	-2.4	87	80	0.70	(0.62-0.76)	-0.4	92	85	1.23	(1.16-1.30)	2.4	93	85

Table 3: The mean, relative bias (Rbias in %), 90% credible intervals and corresponding empirical coverages (EC) for β and α with 10% of right-censoring and a mean width of 1.0σ for a varied number of clusters and cluster sizes under a "Skewed" frailty density with standard deviation of α=1.2 in S=300 replications

| | | | Semiparametric | | | | | Gaussian | | | | |
| | | | | β=0.693 | | | | | β=0.693 | | | |
n_{cl}	n_g	N	Mean	90% CI	Rbias (%)	EC 90	EC 80	Mean	90% CI	Rbias (%)	EC 90	EC 80
	10	200	0.71	(0.46-1.01)	2.4	92	81	0.71	(0.43-1.01)	2.4	92	81
20	20	400	0.70	(0.51-0.87)	1.0	94	85	0.70	(0.51-0.87)	0.9	94	86
	50	1000	0.70	(0.57-0.82)	0.5	91	80	0.70	(0.57-0.82)	0.4	90	81
	4	200	0.70	(0.41-1.01)	0.9	91	79	0.70	(0.41-1.01)	0.8	91	79
50	6	300	0.70	(0.48-0.94)	1.3	91	84	0.70	(0.48-0.94)	1.4	91	84
	10	500	0.70	(0.48-0.88)	0.5	88	76	0.70	(0.49-0.89)	0.5	89	76
	20	1000	0.69	(0.55-0.84)	0.1	88	80	0.69	(0.56-0.83)	0.1	87	80
	4	400	0.70	(0.46-0.91)	0.3	89	77	0.69	(0.48-0.92)	0.1	89	76
100	6	600	0.69	(0.51-0.88)	-0.6	87	80	0.69	(0.51-0.87)	-0.9	90	80
	10	1000	0.69	(0.56-0.81)	-0.5	91	82	0.69	(0.55-0.81)	-0.7	91	81
	4	800	0.70	(0.57-0.83)	0.4	90	83	0.69	(0.56-0.81)	-0.4	90	81
200	6	1200	0.68	(0.56-0.80)	-1.8	88	76	0.68	(0.55-0.80)	-2.4	89	78
	10	2000	0.69	(0.61-0.76)	-0.5	86	77	0.69	(0.61-0.76)	-0.7	90	82
				α=1.2					α=1.2			
n_{cl}	n_g	N	Mean	90% CI	Rbias (%)	EC 90	EC 80	Mean	90% CI	Rbias (%)	EC 90	EC 80
	10	200	1.20	(0.83-1.63)	-0.2	93	81	1.18	(0.83-1.59)	-2.0	89	79
20	20	400	1.16	(0.86-1.46)	-3.0	96	90	1.14	(0.88-1.45)	-4.6	94	89
	50	1000	1.15	(0.87-1.44)	-3.8	95	88	1.14	(0.88-1.42)	-4.9	94	84
	4	200	1.17	(0.83-1.51)	-2.9	91	79	1.15	(0.84-1.50)	-4.1	87	75
50	6	300	1.19	(0.94-1.47)	-0.7	94	83	1.17	(0.93-1.46)	-2.1	93	82
	10	500	1.18	(0.96-1.41)	-1.3	94	85	1.17	(0.96-1.39)	-2.2	92	83
	20	1000	1.18	(1.00-1.36)	-1.9	96	88	1.18	(1.00-1.36)	-1.9	96	86
	4	400	1.18	(0.97-1.39)	-1.5	91	84	1.18	(0.96-1.39)	-1.7	91	83
100	6	600	1.19	(1.01-1.38)	-1.2	92	83	1.18	(1.02-1.39)	-1.2	92	82
	10	1000	1.17	(1.00-1.32)	-2.4	90	83	1.18	(1.03-1.33)	-1.5	94	82
	4	800	1.17	(1.03-1.32)	-2.7	80	68	1.20	(1.09-1.35)	0.3	91	84
200	6	1200	1.16	(1.03-1.31)	-3.0	84	73	1.19	(1.07-1.32)	-0.9	94	82
	10	2000	1.17	(1.07-1.26)	-2.8	88	74	1.19	(1.10-1.28)	-0.9	92	85

Table 4: Relative bias (Rbias in %), empirical standard errors (ESE) and root mean squared error (RMSE) for baseline survival at selected quantiles (5%, 15%, 25%, 35%, 50%, 60%, 75%, 85% and 88%) of T for 10% right-censoring, a mean interval width of $1.0\sigma \approx 15$ and frailty standard deviation $\alpha = 1.2$ for a sample of size n=200 and n=1000 in S=300 replications (using the semiparametric frailty model)

| Frailty | | n_{cl}=20,n_g=10 | | | n_{cl}=20,n_g=50 | | | n_{cl}=50,n_g=4 | | | n_{cl}=50,n_g=20 | | |
		Rbias (%)	ESE	RMSE	Rbias (%)	ESE	RMSE	Rbias (%)	ESE	RMSE	Rbias (%)	ESE	RMSE
	S(39)	0.1	0.014	0.521	-0.1	0.01	0.52	-0.1	0.017	0.521	-0.3	0.008	0.519
	S(49)	0.5	0.034	0.446	0.1	0.025	0.442	0.4	0.037	0.445	-0.6	0.019	0.438
	S(55)	1.1	0.051	0.382	0.3	0.038	0.375	1.0	0.052	0.379	-1.0	0.029	0.367
	S(59)	1.8	0.062	0.329	0.5	0.047	0.324	1.9	0.061	0.333	-1.3	0.035	0.318
Uni.	S(65)	4.6	0.074	0.299	1.0	0.058	0.294	5.7	0.071	0.297	-1.9	0.043	0.295
	S(69)	8.7	0.077	0.312	1.3	0.061	0.315	10.6	0.073	0.303	-2.8	0.044	0.314
	S(75)	22.7	0.073	0.365	5.4	0.057	0.389	28.3	0.070	0.361	-1.2	0.04	0.398
	S(80)	48.1	0.064	0.421	26.4	0.051	0.442	61.0	0.064	0.407	16.3	0.035	0.453
	S(81)	55.4	0.062	0.431	35.0	0.050	0.447	70.7	0.063	0.413	23.8	0.034	0.455
	S(39)	0.1	0.019	0.523	-0.4	0.017	0.519	-0.1	0.016	0.521	-0.4	0.012	0.519
	S(49)	0.2	0.050	0.444	-1.9	0.044	0.434	-0.9	0.038	0.435	-0.8	0.032	0.435
	S(55)	0.3	0.075	0.379	-1.2	0.066	0.375	-1.8	0.053	0.367	-0.8	0.047	0.372
	S(59)	0.8	0.091	0.343	-1.4	0.082	0.332	-2.2	0.063	0.324	-1.1	0.058	0.325
Bimodal	S(65)	4.4	0.108	0.305	-3.8	0.099	0.296	-0.5	0.073	0.292	-3.7	0.069	0.289
	S(69)	11.5	0.112	0.315	-5.4	0.100	0.339	4.9	0.075	0.311	-5.7	0.070	0.330
	S(75)	38.7	0.109	0.363	10.1	0.096	0.395	28.9	0.071	0.365	8.2	0.067	0.389
	S(80)	93.2	0.102	0.384	65.5	0.095	0.402	79.6	0.066	0.385	60.9	0.065	0.403
	S(81)	110.4	0.100	0.389	84.8	0.094	0.415	95.7	0.065	0.393	79.4	0.065	0.411
	S(39)	0.3	0.018	0.523	-0.5	0.015	0.517	-0.2	0.017	0.520	-0.5	0.012	0.517
	S(49)	0.6	0.047	0.447	-1.4	0.040	0.433	-1.1	0.040	0.436	-1.5	0.029	0.433
	S(55)	1.0	0.072	0.385	-2.1	0.061	0.367	-2.4	0.057	0.366	-2.3	0.043	0.362
	S(59)	1.6	0.090	0.336	-2.9	0.075	0.324	-3.4	0.068	0.314	-3.3	0.053	0.317
Skew.	S(65)	5.3	0.109	0.311	-6.2	0.089	0.303	-3.1	0.078	0.300	-7.2	0.063	0.297
	S(69)	12.4	0.114	0.317	-8.5	0.089	0.334	1.2	0.079	0.319	-10.1	0.063	0.329
	S(75)	38.6	0.109	0.356	4.4	0.085	0.399	22.6	0.073	0.366	1.5	0.059	0.398
	S(80)	89.9	0.098	0.389	53.6	0.083	0.417	69.4	0.067	0.400	48.9	0.057	0.453
	S(81)	105.7	0.095	0.390	70.9	0.083	0.418	84.4	0.066	0.403	65.6	0.057	0.455

Table 5: The Signal Tandmobiel® Study data: Parameter estimates and 90% confidence (or credible) intervals for the effect of covariates and frailty standard deviation

		Naive Cox PH with Frailty		Gaussian Frailty		Semiparametric Frailty	
		β	90% CI	β	90% Cred. Int.	β	90% Cred. Int.
Model 1	dmf	0.30	(0.17;0.43)	0.32	(0.17 ; 0.46)	0.34	(0.18 ; 0.48)
	sd(frailty)	1.92	NA	1.52	(1.37 ; 1.68)	1.48	(1.35 ; 1.65)
Model 5	Girl	0.84	(0.59;1.09)	1.47	(0.90 ; 2.08)	1.32	(0.69 ; 1.89)
	dmf	0.69	(0.39;0.99)	0.92	(0.57 ; 1.28)	0.92	(0.55 ; 1.28)
	Girl - dmf	-0.81	(-1.25;-0.38)	-1.02	(-1.56 ;-0.57)	-1.01	(-1.52 ;-0.48)
	Mandibular 4	0.15	(-0.10;0.40)	0.29	(0.01 ; 0.56)	0.28	(-0.01 ; 0.57)
	Maxillary 5	-1.66	(-1.95;-1.38)	-1.80	(-2.12 ;-1.47)	-1.68	(-2.01 ;-1.36)
	Mandibular 5	-1.76	(-2.05;1-.47)	-1.94	(-2.27 ;-1.63)	-1.83	(-2.16 ;-1.49)
	Girl - mandibular 4	0.07	(-0.29;0.42)	0.05	(-0.37 ; 0.47)	0.03	(-0.39 ; 0.45)
	Girl - maxillary 5	-0.22	(-0.61;0.18)	-0.42	(-0.85 ; 0.03)	-0.41	(-0.84 ; 0.04)
	Girl - mandibular 5	-0.01	(-0.40;0.38)	-0.18	(-0.61 ; 0.27)	-0.17	(-0.61 ; 0.28)
	dmf - mandibular 4	-0.28	(-0.67;0.10)	-0.45	(-0.89 ; 0.01)	-0.44	(-0.91 ; 0.03)
	dmf - maxillary 5	-0.37	(-0.78;0.04)	-0.52	(-1.00 ;-0.07)	-0.49	(-0.96 ; 0.00)
	dmf - mandibular 5	-0.77	(-1.22;-0.32)	-1.00	(-1.50 ;-0.49)	-0.96	(-1.48 ;-0.44)
	Girl - dmf - mandibular 4	0.38	(-0.18;0.93)	0.52	(-0.10 ; 1.18)	0.51	(-0.17 ; 1.20)
	Girl - dmf - maxillary 5	1.01	(0.41;1.61)	1.15	(0.42 ; 1.80)	1.09	(0.41 ; 1.75)
	Girl - dmf - mandibular 5	0.88	(0.24;1.52)	1.14	(0.41 ; 1.84)	1.04	(0.30 ; 1.80)
	sd(frailty)	3.58	NA	2.80	(2.55 ; 3.07)	2.24	(2.01 ; 2.53)

Supplement for Semiparametric frailty model for clustered interval-censored data

Aysun Çetinyürek Yavuz [1], **and Philippe Lambert** [1,2]

[1] Faculté des Sciences Sociales, Université de Liège, Liège, Belgium
[2] Institut de Statistique, Université catholique de Louvain, Louvain la Neuve, Belgium

Address for correspondence: Faculté des sciences sociales, Université de Liège, Place des orateurs, 3, Quartier Agora (Bat. B31) B-4000 Liège, Belgium.
E-mail: `cetinyurek@yahoo.com`.

Abstract: This document contains supplementary materials for the article of the same title submitted for publication to the Statistical Modelling: An international journal.

1 Introduction

This document contains supplementary materials for the article of the same title submitted for publication to Statistical Modelling: An international journal. Notation not defined in this document is defined as in the main paper. Section 2 presents some details about the MCMC, e.g. the frequentist density estimation procedure from interval-censored data and some strategies for improving mixing and convergence of the chains, that are not addressed in the main paper. Section 3 provides further information on an extensive simulation study while Section 3.1 presents some sensitivity analysis for the considered prior distributions. Section 4 provides some details a data set from Signal Tandmobiel® study.

2 MCMC

MCMC is an important tool for estimating statistical models. However, especially with complex problems, MCMC can require massive computing resources and converge too slowly. In the following sections, we present various useful approaches in order to improve mixing and convergence of the chains.

2.1 Initialization

Starting the chains at good initial values fasten convergence. Usually, these could be obtained from restricted frequentist models as described in the following. The initial state of the chain $\boldsymbol{\vartheta_{(0)}} = \left(\boldsymbol{\phi}_{(0)}, \tau_{(0)}, \xi_{(0)}, \beta_{(0)}, \boldsymbol{z}_{(0)}, \alpha_{(0)} \right)^T$ is chosen as follows: First, a value for $\boldsymbol{\phi}_{(0)}$ can be obtained using the frequentist density estimation procedure (described in detail in Section 2.2) (Çetinyürek-Yavuz and Lambert, 2011). Then, we define $\tau_{(0)}$ as the value of τ yielding the smallest BIC for different values taken in a grid. $\zeta_{(0)}$ is taken as the proportion of pseudo-counts corresponding to small bins located below t_{cens}. In accordance with the estimation of the spline coefficients, we start by ignoring possible covariate effects: $\boldsymbol{\beta}_{(0)} = 0$. We can obtain $\boldsymbol{z}_{(0)}$ from a gamma frailty model applied on data resulting from mid-point imputation, where the estimated variance is used for $\alpha_{(0)}$.

For the nonparametric specification of the frailty, the initial values of the chain $\boldsymbol{\vartheta}$ are chosen along the same lines as described for the parametric frailty model, except for $(\boldsymbol{\phi}^*_{(0)}, \tau^*_{(0)})$. For obtaining an initial estimate of the frailty density, we also applied the same frequentist procedure as described in Section 2.2. However, as frailty is not observed directly, initial values for the latent frailty terms need to be obtained. It is made in two steps: we first fit a parametric (gamma or log-normal) frailty model using midpoint imputation for the interval-censored data. Then, using the estimated frailty terms as if they were actually observed, a value for $\boldsymbol{\phi}^*_{(0)}$ can be obtained from the frequentist procedure after obtaining an approximation for the density of $\boldsymbol{z^*}_{(0)}$. Similarly, $\tau^*_{(0)}$ is defined as the value of τ^* evaluated on a grid yielding the smallest BIC.

2.2 Frequentist estimation of baseline density

Initial values of the spline coefficients and penalty parameters can be obtained using the naive frequentist models. In this spirit, we shall explain a frequentist density estimation procedure from time-to-event data when the covariates and the possible heterogeneity are ignored ($\boldsymbol{\beta} = 0$, $\boldsymbol{b_g} = 1$ (or $\boldsymbol{z} = 0$)). We start by partitioning the support of t into small bins (more than 100 small bins of equal width) for obtaining an accurate estimate of the density for time-to-event data. Then, following an approach similar to Eilers and Marx (1996) we calculate the number of observations in each small bin, namely the pseudo-counts. These pseudo-counts, which are calculated from the C matrix of the composite link model defined in Eilers and Marx (1996), are later used to build the density estimate. The relationship between the intervals (l_{gj}, r_{gj}) and the small bins is provided by a $G \mathrm{x} I \mathrm{x} n_g$ array $C = [c_{gji}]$ such that $c_{gji} = 1$ if the ith small bin $I_i \subset (l_{gj}, r_{gj})$ and 0 otherwise. In this spirit, for cluster g, each element of a row in the C matrix is divided by the sum of the elements in that row ($C_{gj.} = \sum_i C_{gji}$). The so-obtained numbers, $W_{gji} = C_{gji}/C_{gj.}$, provide the contribution of the concerned observation in cluster g (e.g. a patient for a multicenter clinical trial) for each small bin partitioning (a, t_{cens}). Then, the contributions of each observation for the ith small bin, I_i, are summed over all observations ($W_{..i} = \sum_g \sum_j W_{gji}$) and rounded to the nearest integer value y_i in order to get the pseudo-count for that small bin. Note that π_i denotes the probability to have an event time in I_i, then the likelihood for these pseudo-

counts is proportional to $\prod_i^I \pi_i^{y_i}$. Alternatively, the well known link between the Poisson and the multinomial distributions suggests to assume that the pseudo-counts, y_i, have a Poisson distribution with mean $\mu_i = \pi_i y_+$ conditional on the total number of observations $y_+ = \sum_i^I y_i$. Using a rich B-spline basis as regressors in a log-linear model for the mean, one obtains the likelihood

$$\mathcal{L}(\boldsymbol{\phi}|y) = \sum_{i=1}^N y_i \log(\mu_i) - \sum_{i=1}^N \mu_i,$$

where $\log(\mu_i) = \eta_i = \sum_{k=1}^K \phi_k b_k(u_i)$. Then, by subtracting the 2nd order penalty (say) and a small ridge penalty from $\mathcal{L}(\boldsymbol{\phi}|y)$, one obtains the penalized log likelihood function

$$\mathcal{L}_p(\boldsymbol{\phi}|y, \tau) = \mathcal{L}(\boldsymbol{\phi}|y) - \frac{\tau}{2}\boldsymbol{\phi}'\boldsymbol{P}\boldsymbol{\phi},$$

where $\boldsymbol{\phi}'\boldsymbol{D}'\boldsymbol{D}\boldsymbol{\phi} = \sum_k (\phi_k - 2\phi_{k-1} + \phi_{k-2})^2$ and $\boldsymbol{P} = \boldsymbol{D}'\boldsymbol{D} + \epsilon I$. The function $\mathcal{L}_p$ can be optimized by solving the score equations $B^T(y - \mu) = \tau\boldsymbol{P}\boldsymbol{\phi}$, using iteratively reweighted least squares (IRWLS): iteratively solve (for ϕ)

$$(B^T \tilde{W} B + \tau P)\phi = B^T \tilde{W}(y - \tilde{\mu}) + B^T \tilde{W} B \tilde{\phi},$$

where $\widetilde{W}$ is a diagonal matrix with elements $\mu_i(\widetilde{\phi})$ and $\tilde{\phi}$ and $\tilde{\mu}$ are current approximations to the solution. The variance-covariance matrix for the estimated spline coefficients $\boldsymbol{\phi}$ is given by (at convergence),

$$\Sigma_0 = (B^T W B + \tau\boldsymbol{P})^{-1}. \tag{2.1}$$

More detailed information can be found in Eilers and Marx (1996). Information criteria such as AIC or BIC could be used for choosing the initial optimal (plausible) value of the penalty parameter τ. In our experience, BIC is preferable to AIC as AIC tends to undersmooth the target curve, which was also mentioned by other authors (Strasak et al., 2009).

2.3   Automatic tuning of the algorithm

Good acceptance rates can be achieved via a careful choice of the standard deviation δ_h in the generation of proposals in a Metropolis algorithm. For an optimal use of Metropolis algorithm, it is recommended to tune the acceptance probability to approximately 0.44 in one dimensional space decreasing to 0.23 in high dimensional spaces (Gelman et al., 1996; Roberts and Rosenthal, 2001). Let δ denote the tuning parameter of interest. The value of δ at iteration $m + 1$ can be adjusted using the value at iteration m using (with $\bar{\eta} = 0.44$)

$$\sqrt{\delta_{m+1}} = h\left(\sqrt{\delta_m} + \gamma_m\left(\alpha(\boldsymbol{\theta}_{(h)}, \boldsymbol{\theta}_{(h-1)}) - \bar{\eta}\right)\right)$$

with

$$h(x) = \begin{cases} \epsilon & \text{if } x < \epsilon \\ x & \text{if } x \in (\epsilon, A) \\ A & \text{if } x > A \end{cases}$$

where ϵ is a very small number (say 0.0001) and A a large one (say 10000). If the targeted acceptance level is not achieved, these constants should be changed. The series $\{\gamma_m\}$ is a non-increasing sequence of positive real numbers such that $|\gamma_m - \gamma_{m-1}| \leq m^{-1}$. Possible choices for γ_m are $\frac{10}{m}$ or $\frac{1}{m}$. Practically, the MCMC algorithm is run for a few hundred iterations with the δ_m's automatically updated to achieve the targeted acceptance rate (Atchadé and Rosenthal, 2005). Then, the last value of δ_m in the so-generated chain can be used in a non-adaptive version of the modified Metropolis algorithm to produce the long chain(s) that will be used for inference.

2.4   Reparametrizing the posterior

The mixing of the chain could be improved by using a Metropolis algorithm on a re-parametrized posterior (Lambert, 2007). In this sense, one can use an approximation to the 2nd order dependence structure of the conditional posterior. The variance covariance matrix, $\mathbf{\Sigma_0}$, of the penalized maximum likelihood estimator of the spline parameters $\boldsymbol{\phi}$ could be calculated for a fixed and reasonably chosen value of the roughness penalty parameter τ. Then, the posterior can be re-parametrized using φ (Equation 2.1) with $\phi = \phi_0 + L\varphi$ where L denotes the lower triangular matrix obtained from the Cholesky decomposition of $\mathbf{\Sigma_0}$. Then, the univariate Metropolis algorithm described before can be employed on the re-parametrized posterior. This also fastens convergence.

3   Simulation Study

Our data generation and simulation strategy contain the following steps:

1. Firstly, we generate the log-frailty terms z_g from one of the specified frailty distributions.

2. Then the values of the covariate, x_{gj} are generated.

3. Afterwards, given the values of frailty terms and the covariate, the observations t_{gj} $(g = 1, ..., G; j = 1, ..., n_g)$ are generated using the selected proportional hazards frailty model.

4. Each observation t_{gj} is converted into an interval of width w_{gj}, where w_{gj} is generated from a Gamma distribution with a mean equal to the targeted mean width (0.5σ, 1.0σ and 1.3σ) and a variance equal to one fifth of the mean. The interval corresponding to t_{gj} was finally defined as $(L_{gj}, R_{gj}) = (t_{gj} - u_{gj} \cdot w_{gj}, t_{gj} + u_{gj} \cdot w_{gj})$ where u_{gj} is randomly generated from a uniform distribution on (0,1).

5. For each simulated data set, initial values for the spline parameters were obtained using the strategy described in Section 2.2.

6. We sample the posterior for the parameters of interest using MCMC (see Section 3.3).

7. Steps 1-6 were repeated for all data sets ($S = 300$ times) to obtain the Monte Carlo estimates for the quantities of interest.

3.1   Sensitivity analysis

Following the advice from the referees, some of the simulation studies have been run again with different prior specifications, namely Gamma (2,0.01) for the penalty parameters, and inverse-gamma(1,1) and inverse-gamma(2,1) for the standard deviation of the frailty. The detailed results are presented in the tables below. The results in Tables 1 and 2 can be compared to the results in the main paper. Table 1 and Table 2 presents two different sensitivity analysis with different prior distribution specifications, namely Gamma (2,0.01) prior for the penalty parameters, and inverse-gamma(1,1) and inverse-gamma(2,1) priors for the standard deviation of the frailty.

- The unimodal setting with α=0.8 can be compared to left half of Table 1 (main paper) for the given sample sizes.

- The bimodal setting with α=0.8 can be compared to left half of Table 2 (main paper) for the given sample sizes.

- The bimodal setting with α=1.2 can be compared to right half of Table 2 (main paper) for the given sample sizes.

- The skewed setting with α=1.2 can be compared to Table 3 (main paper) for the given sample sizes.

It reveals that changes in prior specification have a limited impact on bias and coverage of credible intervals for the regression parameters, with occasionally a slight improvement (over our standard prior) in the estimation of the standard deviation of the log-frailty. It should also be stressed here that sensitivity analyses were performed for the smallest sample sizes where the chosen prior has the biggest potential impact.

4   Application: Signal Tandmobiel® Study

The Signal Tandmobiel® data set results from a longitudinal prospective dental study performed in Flanders (northern Belgium) between 1996 and 2001, using 4468 randomly selected children attending the first year of primary school at the beginning of the study. Then annual dental examinations were performed on the selected cohort by one of 16 trained dentists.

Table 1: Sensitivity analysis I: The mean, relative bias (Rbias in %), 90% credible intervals and corresponding empirical coverages (EC) for β and α in S=300 replications using Gamma(2,0.001) prior for penalty parameters τ and τ^*, Inverse-gamma(1,1) prior for standard deviation of frailty α

| | | | | | Semiparametric | | | | | Gaussian | | | | |
| | | | | | β=0.693 | | | | | β=0.693 | | | | |
Distribution	α	n_{cl}	n_g	N	Mean	90% CI	Rbias(%)	EC90	EC80	Mean	90% CI	Rbias(%)	EC90	EC80
Unimodal	0.8	20	10	200	0.75	(0.51-0.98)	8.4	90	80	0.75	(0.51-0.99)	8.6	91	80
		50	4	200	0.75	(0.46-1.03)	8.5	89	77	0.75	(0.46-1.04)	8.7	89	77
		50	6	300	0.74	(0.52-0.95)	6.7	91	78	0.74	(0.52-0.95)	6.8	89	77
		100	4	400	0.71	(0.51-0.91)	2.4	89	78	0.71	(0.51-0.91)	2.7	90	78
Bimodal	0.8	20	10	200	0.76	(0.51-1.06)	9.9	86	80	0.76	(0.50-1.05)	9.1	86	80
		50	4	200	0.75	(0.49-1.04)	8.3	89	79	0.76	(0.48-1.05)	8.0	88	79
		50	10	500	0.71	(0.55-0.88)	3.1	93	84	0.71	(0.55-0.87)	2.9	92	84
Bimodal	1.2	20	10	200	0.71	(0.42-1.02)	2.9	90	81	0.71	(0.40-1.01)	2.1	90	81
		50	4	200	0.72	(0.37-1.00)	4.5	88	75	0.73	(0.37-1.02)	4.7	89	75
		50	6	300	0.72	(0.48-0.96)	3.6	91	82	0.71	(0.46-0.95)	2.4	90	82
		100	4	400	0.71	(0.51-0.93)	2.4	90	80	0.71	(0.49-0.93)	2.9	90	80
Skewed	1.2	20	10	200	0.73	(0.45-1.02)	5.2	92	81	0.73	(0.45-1.01)	5.4	91	81
		50	4	200	0.72	(0.39-1.01)	3.5	90	77	0.72	(0.39-1.02)	3.5	90	79
		50	6	300	0.72	(0.48-0.96)	3.4	91	77	0.72	(0.47-0.96)	3.4	90	77
		100	4	400	0.70	(0.48-0.91)	0.4	88	81	0.69	(0.46-0.92)	0.1	89	80
					α					α				
Distribution	α	n_{cl}	n_g	N	Mean	90% CI	Rbias(%)	EC90	EC80	Mean	90% CI	Rbias(%)	EC90	EC80
Unimodal	0.8	20	10	200	0.87	(0.67-1.12)	8.3	94	87	0.85	(0.65-1.11)	6.2	93	86
		50	4	200	0.84	(0.68-1.06)	4.5	94	86	0.83	(0.68-1.07)	4.0	93	85
		50	6	300	0.83	(0.65-1.01)	3.5	95	85	0.82	(0.65-1.00)	2.7	94	84
		100	4	400	0.80	(0.65-0.95)	0.1	94	84	0.80	(0.65-0.95)	0.5	92	82
Bimodal	0.8	20	10	200	0.76	(0.51-1.06)	-5.5	86	80	0.88	(0.70-1.06)	9.8	96	86
		50	4	200	0.69	(0.55-0.88)	-7.7	84	76	0.84	(0.66-1.04)	5.4	95	86
		50	10	500	0.77	(0.65-0.87)	-3.3	95	93	0.82	(0.70-0.94)	2.9	99	91
Bimodal	1.2	20	10	200	1.28	(0.89-1.47)	7.8	91	83	1.27	(0.91-1.57)	6.2	93	83
		50	4	200	1.27	(0.98-1.57)	5.8	92	83	1.25	(0.98-1.55)	4.1	92	82
		50	6	300	1.12	(0.95-1.30)	-6.3	94	85	1.23	(1.03-1.45)	2.9	96	90
		100	4	400	1.23	(1.04-1.43)	2.4	92	86	1.23	(1.04-1.43)	2.5	93	84
Skewed	1.2	20	10	200	1.27	(0.96-1.65)	5.5	94	88	1.23	(0.93-1.61)	2.9	94	87
		50	4	200	1.25	(0.96-1.60)	4.0	92	84	1.22	(0.94-1.56)	1.9	92	83
		50	6	300	1.23	(0.95-1.52)	2.4	91	81	1.20	(0.93-1.49)	0.3	88	81
		100	4	400	1.21	(1.00-1.45)	1.0	92	83	1.20	(0.99-1.42)	-0.3	92	81

Table 2: Sensitivity analysis II: The mean, relative bias (Rbias in %), 90% credible intervals and corresponding empirical coverages (EC) for β and α in S=300 replications using Gamma (2,0.001) prior for penalty parameters τ and τ^*, Inverse-gamma(2,1) prior for standard deviation of frailty α

| | | | | | Semiparametric | | | | | Gaussian | | | | |
| | | | | | β=0.693 | | | | | β=0.693 | | | | |
Distribution	α	n_{cl}	n_g	N	Mean	90% CI	Rbias(%)	EC90	EC80	Mean	90% CI	Rbias(%)	EC90	EC80
Unimodal	0.8	20	10	200	0.75	(0.49-1.03)	8.6	87	77	0.75	(0.49-1.02)	8.8	86	77
		50	4	200	0.74	(0.46-1.02)	7.0	89	80	0.74	(0.46-1.02)	7.1	88	80
		50	6	300	0.72	(0.52-0.95)	4.1	91	82	0.72	(0.52-0.96)	4.3	91	83
		100	4	400	0.71	(0.53-0.91)	1.9	89	81	0.71	(0.53-0.91)	2.1	90	80
Bimodal	0.8	20	10	200	0.74	(0.46-1.01)	7.3	89	73	0.74	(0.45-1.01)	6.7	87	73
		50	4	200	0.74	(0.45-1.02)	6.2	88	76	0.74	(0.45-1.04)	6.6	88	76
		50	10	500	0.72	(0.55-0.89)	3.4	89	80	0.71	(0.55-0.90)	3.2	89	80
Bimodal	1.2	50	4	200	0.72	(0.40-1.05)	3.3	89	79	0.72	(0.39-1.05)	3.4	90	80
		50	6	300	0.72	(0.47-0.99)	3.4	89	78	0.71	(0.46-0.99)	2.2	88	78
		100	4	400	0.69	(0.49-0.90)	-0.1	91	84	0.69	(0.48-0.90)	-0.2	91	83
Skewed	1.2	20	10	200	0.72	(0.45-1.00)	4.1	92	82	0.72	(0.45-1.00)	4.2	91	82
		50	4	200	0.70	(0.41-1.01)	1.7	91	83	0.70	(0.42-1.02)	1.7	91	84
		50	6	300	0.70	(0.47-0.95)	1.1	91	82	0.70	(0.47-0.95)	1.1	91	81
		100	4	400	0.68	(0.46-0.90)	-1.5	89	77	0.68	(0.46-0.89)	-1.6	88	78
					α					α				
Distribution	α	n_{cl}	n_g	N	Mean	90% CI	Rbias(%)	EC90	EC80	Mean	90% CI	Rbias(%)	EC90	EC80
Unimodal	0.8	20	10	200	0.82	(0.60-1.07)	2.3	96	86	0.81	(0.59-1.05)	1.6	92	85
		50	4	200	0.76	(0.59-1.01)	-4.4	93	78	0.76	(0.59-1.01)	-4.5	90	77
		50	6	300	0.78	(0.62-0.96)	-2.0	89	79	0.79	(0.63-0.97)	-1.7	95	83
		100	4	400	0.78	(0.65-0.94)	-2.6	96	89	0.78	(0.66-0.93)	-2.0	95	87
Bimodal	0.8	20	10	200	0.71	(0.57-0.86)	-11.6	91	82	0.81	(0.64-1.01)	1.4	99	92
		50	4	200	0.63	(0.34-0.83)	-21.6	72	59	0.81	(0.64-0.99)	1.1	97	90
		50	10	500	0.76	(0.66-0.87)	-5.2	96	90	0.81	(0.70-0.92)	0.8	98	94
Bimodal	1.2	20	10	200	1.20	(0.94-1.47)	0.4	95	89	1.19	(0.95-1.46)	-0.5	93	88
		50	4	200	1.09	(0.90-1.26)	-9.3	88	76	1.19	(0.98-1.40)	-0.6	94	89
		50	6	300	1.21	(1.03-1.41)	0.6	94	83	1.21	(1.03-1.40)	0.8	94	88
Skewed	1.2	20	10	200	1.16	(0.87-1.54)	-3.1	94	86	1.14	(0.86-1.51)	-4.6	93	83
		50	4	200	1.14	(0.88-1.43)	-5.0	94	80	1.13	(0.88-1.40)	-6.1	88	79
		50	6	300	1.17	(0.91-1.40)	-2.6	93	85	1.15	(0.90-1.36)	-4.1	90	81
		100	4	400	1.18	(0.97-1.39)	-2.7	90	78	1.17	(0.97-1.39)	-2.7	90	78

References

Atchadé, Y.F. and Rosenthal, J.S. (2005). On adaptive Markov chain Monte Carlo algorithms. *Bernoulli*, **11 (5)**, 815–828.

Çetinyürek-Yavuz, A. and Lambert, P. (2011). Smooth estimation of survival functions and hazard ratios from interval-censored data using Bayesian penalized B-splines. *Statistics in Medicine*, **30**, 75–90.

Eilers, P.H.C. and Marx, B.D. (1996). Flexible smoothing with B-splines and penalties (with comments and rejoinder). *Statistical Science*, **11**, 89–121.

Gelman, A., Roberts, G.O., Gilks, W.R. (1996). Efficient Metropolis jumping rules. *Bayesian Statistics*, **5**, 599–607.

Lambert, P. (2007). Archimedean copula estimation using Bayesian splines smoothing techniques. *Computational Statistics and Data Analysis*, **51**, 6307–6320.

Roberts, G.O. and Rosenthal, J.S. (2001). Optimal scaling for various Metropolis- Hastings algorithms. *Statistical Science*, **16**, 351–367.

Strasak, A.M., Lang, S., Kneib, T., Brant, L.J., Klenk, J., Hilbe, W., Oberaigner, W., Ruttmann, E., Kaltenbacha, L., Concin, H., Diem, G., Pfeiffer, K.P., Ulmer, H., The VHM PP Study Group. (2009) Use of penalized splines in extended Cox-type additive hazard regression to flexibly estimate the effect of time-varying serum uric acid on risk of cancer incidence: A prospective, population-based study in 78,850 men. *Annals of Epidemiology*, **19 (1)**, 15–24.