Statistical methods for frailty models: studies on old-age mortality and recurrent events

Böhnstedt, M.

Part II

Joint modeling of recurrent events and death

5

Joint modeling of interval counts of recurrent events and death

Abstract

When a recurrent event process is ended by death, this may imply dependent censoring if the two processes are associated. Such dependent censoring would have to be modeled to obtain a valid inference. Moreover, the dependence between the recurrence process and the terminal event may be the primary topic of interest. Joint frailty models for recurrent events and death, which include a separate dependence parameter, have been proposed for exactly observed recurrence times. However, in many situations, only the number of events experienced during consecutive time intervals are available. We propose a method for estimating a joint frailty model based on such interval counts and observed or independently censored terminal events. The baseline rates of the two processes are modeled by piecewise constant functions, and Gaussian quadrature is used to approximate the marginal likelihood. Covariates can be included in a proportional rates setting. The observation intervals for the recurrent event counts can differ between individuals. Furthermore, we adapt a score test for the association between recurrent events and death to the setting in which only individual interval counts are observed. We study the performance of both approaches via simulation studies, and exemplify the methodology in

This chapter has been published as: M. Böhnstedt, H. Putter, A. Dańko, M.J. Dańko, and J. Gampe (2021). Joint modeling of interval counts of recurrent events and death. *Biometrical Journal* 63, 323–340.

a biodemographic study of the dependence between budding rates and mortality in the species *Eleutheria dichotoma*.

5.1 Introduction

Studies of recurrent events, in which an individual can experience the same type of event repeatedly over time, are common in various fields of applications (Cook and Lawless, 2007). Examples range from medical studies of the recurrence of adverse symptoms, such as epileptic seizures, asthma attacks, or tumor relapse; to investigations of repeated insurance claims; to biodemographic studies of fertility (recurrent reproductive events) in particular animal species.

In some cases, the exact occurrence times can be observed, but often only the numbers of events that were experienced in specific time intervals are available. Such interval counts of recurrent events may result if, for example, patients only report the number of adverse events that they experience between two hospital visits, or the number of offspring produced by an animal is collected on a monthly basis only. In the latter example, the observation intervals would be the same for all individuals, while in the former example, we would expect the intervals between the two visits to vary from patient to patient.

The recurrent event process is often terminated by another event – most commonly by death – which usually cannot be assumed to be independent of the recurrent event process. Consequently, the terminal event introduces dependent censoring of the recurrence process, and this has to be taken into account to render a valid inference. Therefore, the two processes, the recurrent event process and the terminal event, have to be modeled jointly.

In many medical applications, the dependence of the two processes will be positive; that is, a higher recurrence rate of the (adverse) symptoms will be accompanied by a higher hazard of death. In other contexts, however, the direction of the association between the recurrence process and the terminal event is not clear at the outset, and will be a matter of interest in itself.

Our motivating example examines fertility and mortality in *Eleutheria dichotoma*, a marine hydrozoan for which the association between fertility and mortality has not previously been studied in detail. Reproduction and survival in *E. dichotoma* were investigated in a laboratory experiment for several months (Dańko et al., 2020). Individual survival times and the number of offspring that were produced by each individual within successive intervals of several days were recorded. The intervals resulted from the laboratory procedures and varied across individuals. These data were used to estimate the patterns of fertility and mortality over age, but the dependence between the two processes is also of biological interest.

It has been suggested that in some species, there is a trade-off between reproduction and survival. This idea is based on the assumption that individuals that produce a higher number of offspring are able to devote fewer resources to maintenance, and will, therefore, tend to die earlier. The claim that there is a cost of reproduction effect is in contrast to the

hypothesis that individuals who are stronger will be able to both produce more offspring and survive longer. Therefore, in addition to modeling the shape of age-specific fertility and mortality, the aim of the analysis is to find out which of the two explanations the data on *E. dichotoma* support. Thus, we will model fertility and mortality jointly to assess how the two processes are related in this species.

Several approaches to jointly modeling recurrent events and death have been proposed. We focus here on the joint frailty model introduced by Liu et al. (2004), because it allows for positive and negative associations between the recurrences and death. The dependence between the two processes is modeled by a shared individual random effect that acts on both the rate of recurrence and the hazard of death, possibly in different directions. In other frailty models, frailty has the same effect on the recurrence rate and the hazard of death. Thus, these models are restricted to a positive association (see Huang and Wang (2004) in the setting with exact recurrence times, or Lancaster and Intrator (1998) in the setting with interval counts). As marginal models for recurrent events in the presence of death leave the dependence between the two processes unspecified, they are not suitable for our purposes (see Cook and Lawless (1997) and Ghosh and Lin (2003) in the setting with exact recurrence times, or Zhao et al. (2013) in the setting with interval counts). Sinha and Maiti (2004) proposed a model similar to that of Liu et al. (2004), which is based on interval counts, but assumes that observation intervals are the same for all individuals, and that the termination time is discrete.

Estimation of the joint frailty model introduced by Liu et al. (2004) has so far only been based on observed recurrence times. For this setting, several methods of estimation have been developed: Liu et al. (2004) used a Monte Carlo EM algorithm, whereas Liu and Huang (2008) and Rondeau et al. (2007) applied Gaussian quadrature to the marginal likelihood. Moreover, a test for the association between recurrent events and a terminal event in the joint frailty model was derived by Balan et al. (2016), which was also based on observed recurrence times. It builds on concepts that are similar to the test proposed by Huang et al. (2004) for the association between two event processes in clustered survival data.

In this chapter, we propose methods for making inferences in the joint frailty model when only individual interval counts of the recurrent events are observed, and these observation intervals can vary between individuals. We will adapt the method of Liu and Huang (2008) for the estimation of the joint frailty model, and we will adjust the score test developed by Balan et al. (2016) to the setting in which only interval counts are available.

The chapter is structured as follows. In Section 5.2, we describe the joint frailty model for recurrent events and death, as well as the setting of individual interval counts. In Section 5.3, we present our approach of using Gaussian quadrature to estimate the joint frailty model based on interval counts, and adapt the score test for association in the joint frailty model. In Section 5.4, we assess the performance of the estimation method and the test in simulation studies. In Section 5.5, we apply the proposed methods to the data on *E. dichotoma*, followed by a discussion in Section 5.6.

5.2 Joint frailty model and interval counts

Before presenting the estimation and test procedure in the next section, we will introduce in the following the joint frailty model, which allows us to model the dependence of the recurrent event process and the terminal event. We then derive the likelihood function for data that only contain interval counts of the recurrent events.

We consider a sample of m independent individuals denoted by $i, i = 1, \ldots, m$. Each individual i is observed from time $t_0 = 0$ until the end of its follow-up X_i. The time X_i is either a censoring time C_i, which is independent of the recurrent event process and the terminal event, such as end of study; or it is the time D_i of the terminal event, whichever comes first: $X_i = \min(C_i, D_i)$. For simplicity, we assume that the terminal event is death, and denote by $\delta_i = \mathbb{1}\{D_i \leq C_i\}$ the death indicator, where $\mathbb{1}\{\cdot\}$ is the indicator function. $Y_i(t) = \mathbb{1}\{t \leq X_i\}$, $t \geq 0$, is the at-risk indicator at time t. We define two additional counting processes $N_i^{D*}(t) = \mathbb{1}\{D_i \leq t\}$ and $N_i^D(t) = \mathbb{1}\{X_i \leq t, \delta_i = 1\}$, where $N_i^{D*}(t)$ refers to the actual (but potentially unobserved) terminal event, whereas $N_i^D(t)$ is the counting process of an observed terminal event, respectively.

For the recurrent event process, we denote with $N_i^{R*}(t)$ the number of events of individual i in the interval $[0, t]$. However, we only observe $N_i^R(t) = N_i^{R*}(\min(t, X_i))$. The increments of the recurrence process over small intervals $[t, t + \mathrm{d}t)$ are $\mathrm{d}N_i^{R*}(t) = N_i^{R*}((t + \mathrm{d}t)^-) - N_i^{R*}(t^-)$. Here, t^- denotes the left-hand limit.

Additional observed characteristics of individual i are collected in the covariate vector z_i, whereas unobserved characteristics are summarized in a frailty value u_i. The u_i are realizations of a positive random variable U, independent across individuals. The observed data on individual i up to time t are collected in $O_i(t) = \{Y_i(s), N_i^R(s), N_i^D(s), 0 \leq s \leq t; z_i\}$.

As in Liu et al. (2004), the recurrence process is characterized by the intensity $Y_i(t)\lambda_i(t)$, for which we assume

$$\begin{aligned} \mathrm{P}(\mathrm{d}N_i^R(t) = 1 \mid \mathcal{F}_{t^-}, D_i \geq t) &= Y_i(t)\lambda_i(t)\mathrm{d}t \quad \text{with} \\ \lambda_i(t)\mathrm{d}t = \mathrm{d}\Lambda_i(t) &= \mathrm{P}(\mathrm{d}N_i^{R*}(t) = 1 \mid z_i, u_i, D_i \geq t), \end{aligned} \tag{5.1}$$

where $\mathcal{F}_t = \sigma\{O_i(s), 0 \leq s \leq t, u_i; i = 1, ..., m\}$.
Analogously, for the terminal event

$$\begin{aligned} \mathrm{P}(\mathrm{d}N_i^D(t) = 1 \mid \mathcal{F}_{t^-}) &= Y_i(t)h_i(t)\mathrm{d}t \quad \text{with} \\ h_i(t)\mathrm{d}t = \mathrm{d}H_i(t) &= \mathrm{P}(\mathrm{d}N_i^{D*}(t) = 1 \mid z_i, u_i, D_i \geq t). \end{aligned}$$

The joint frailty model for recurrent events and death, as proposed by Liu et al. (2004), is then specified as

$$\begin{aligned} \lambda_i(t) &= u_i \, e^{\beta^\top z_i} \lambda_0(t), \\ h_i(t) &= u_i^\gamma \, e^{\alpha^\top z_i} h_0(t), \end{aligned} \tag{5.2}$$

with baseline rates $\lambda_0(t)$ and $h_0(t)$ that are common to all individuals.

The shared frailty u enters both the recurrent event rate and the hazard rate of the terminal event, thereby introducing both dependence between the recurrences within one individual, as well as the association between the recurrences and the terminal event. The parameter γ determines the direction and the strength of the association between the two processes. If $\gamma > 0$, a higher rate of recurrence implies a higher mortality risk; if $\gamma < 0$, a higher rate of recurrence implies a lower mortality risk. If $\gamma = 0$, the rate of recurrence does not affect the mortality risk.

The frailties u_i are often assumed to follow a gamma distribution with mean one and variance θ. In the following, we will more generally assume that the u_i stem from a distribution with density $g_\theta(u)$ with parameter θ.

The covariates z_i enter in (5.2) in a proportional rates or a proportional hazards formulation, but can have different effects $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ on the terminal event and the recurrence process, respectively.

In the simplest setting, the exact times of event occurrence are observed. In many cases, however, only the number of events that occurred in a sequence of consecutive time intervals is available. More precisely, we observe individual interval counts n_{ij} as realizations of $N_{ij} = N_i^R(t_{ij}) - N_i^R(t_{ij-1})$. The N_{ij} give the number of recurrent events experienced by individual i in the interval $I_{ij} = (t_{ij-1}, t_{ij}]$, $j = 1, ..., J_i$. The t_{ij} correspond to the observation times of individual i, for instance, times of hospital visits in medical studies or, as in our example, generated by the lab logistics. Thus, both the positions of the I_{ij} and the total number of intervals J_i can vary across individuals. The follow-up times X_i are still exactly observed so that $t_{iJ_i} = X_i$.

As the frailties u_i are unobservable, the inference is based on the marginal likelihood that is obtained by integrating the conditional likelihood given the frailties u_i over the frailty distribution $g_\theta(u)$. The conditional likelihood of the joint frailty model (5.2) based on exactly observed recurrence times was developed in Liu et al. (2004). For the current setting, the likelihood factor for the recurrence times (formula (7) in Liu et al., 2004) has to be replaced by the contribution of the interval counts of the recurrent events. From (5.1) and (5.2), it follows that $(N_{ij} \mid z_i, u_i, D_i \geq t_{ij})$ has a Poisson distribution with mean $u_i \mu_{ij}$ where $\mu_{ij} = \int_{I_{ij}} e^{\boldsymbol{\beta}^\top z_i} \lambda_0(s) \mathrm{d}s$. Therefore, the likelihood contribution $L_i^{(c)}(u_i)$ of individual i conditional on its frailty value u_i is given by

$$L_i^{(c)}(u_i) = \prod_{j=1}^{J_i} \frac{\exp(-u_i \mu_{ij})(u_i \mu_{ij})^{n_{ij}}}{n_{ij}!} [u_i^\gamma e^{\boldsymbol{\alpha}^\top z_i} h_0(x_i)]^{\delta_i} \exp\left\{ - \int_0^{x_i} u_i^\gamma e^{\boldsymbol{\alpha}^\top z_i} h_0(s) \mathrm{d}s \right\}.$$

This leads to the marginal likelihood contributions

$$L_i = \int_0^\infty L_i^{(c)}(u)\, g_\theta(u)\, \mathrm{d}u\,. \tag{5.3}$$

In general, this integral does not have a closed-form expression.

5.3 Methods

In the first part of this section, we elaborate the estimation procedure of the joint frailty model based on interval counts with different individual observation intervals. Then, in Section 5.3.2, we demonstrate how the score test for dependence between the two processes can be adapted to the case of interval-count data.

5.3.1 Estimation of the joint frailty model based on interval counts

For observed recurrence times, Liu and Huang (2008) suggested using Gaussian quadrature to approximate the marginal likelihood of the joint frailty model. The approximated likelihood can then be maximized directly, since the integral is replaced by a weighted sum of function values. More specifically, Liu and Huang (2008) applied Gauss-Hermite quadrature, which for a function $f(x)$ uses the approximation

$$\int_{-\infty}^{\infty} f(x)\, e^{-x^2} \mathrm{d}x \approx \sum_{q=1}^{Q} w_q\, f(x_q)\,.$$

The quadrature points x_q are the roots of the Q^{th}-order Hermite polynomial, and w_q are the corresponding weights. This approach is applicable to marginal likelihoods that are integrated over normal random effects, for which

$$\int_{-\infty}^{\infty} L^{(c)}(v)\, \phi(v)\, \mathrm{d}v \approx \sum_{q=1}^{Q} \tilde{w}_q\, L^{(c)}(\tilde{x}_q)\, \phi(\tilde{x}_q), \tag{5.4}$$

with the standard normal density $\phi(\cdot)$, and modified quadrature points $\tilde{x}_q = \sqrt{2}x_q$ together with weights $\tilde{w}_q = \sqrt{2}w_q e^{x_q^2}$. For non-normal random effects, Liu and Huang (2008) used the probability integral transformation proposed by Nelson et al. (2006). If the random effect density is $g_\theta(u)$ with corresponding distribution function $G_\theta(u)$, then the integral over the density $g_\theta(u)$ is transformed into an integral over standard normal random effects. This is achieved by noting that $a = \Phi^{-1}(G_\theta(u))$ follows a standard normal distribution if the $G_\theta(u)$, which follow a standard uniform, are transformed by the inverse of the standard normal distribution function $\Phi(\cdot)$.

We apply this quadrature approach to the marginal likelihood of the joint frailty model based on interval counts of recurrent events; see equation (5.3). Substituting $u = G_\theta^{-1}(\Phi(a))$ in the marginal likelihood contributions, we obtain

$$L_i = \int_0^{\infty} L_i^{(c)}(u)\, g_\theta(u)\, \mathrm{d}u = \int_{-\infty}^{\infty} L_i^{(c)}(G_\theta^{-1}(\Phi(a)))\, \phi(a)\, \mathrm{d}a\,.$$

These L_i can directly be approximated using Gauss-Hermite quadrature as

$$L_i \approx \sum_{q=1}^{Q} L_i^{(c)}(G_\theta^{-1}(\Phi(\tilde{x}_q)))\, \phi(\tilde{x}_q)\, \tilde{w}_q,$$

with $\tilde{x}_q$ and $\tilde{w}_q$ as defined in (5.4). The approximate marginal likelihood of the joint frailty model is then given by

$$\prod_{i=1}^{m} \sum_{q=1}^{Q} L_i^{(c)}(G_\theta^{-1}(\Phi(\tilde{x}_q))) \, \phi(\tilde{x}_q) \, \tilde{w}_q. \tag{5.5}$$

To actually maximize the approximate likelihood (5.5), we have to specify the baseline rates $\lambda_0(t)$ and $h_0(t)$, as well as the frailty distribution $g_\theta(u)$. Similar to Liu and Huang (2008), we model the baseline rates as piecewise constant functions

$$\lambda_0(t) = \sum_{k=1}^{K^R} \lambda_{0k} \mathbb{1}\{t \in I_k^R\} \qquad \text{and} \qquad h_0(t) = \sum_{k=1}^{K^D} h_{0k} \mathbb{1}\{t \in I_k^D\}. \tag{5.6}$$

This choice is particularly suitable if no prior knowledge of the shapes of the two rates $\lambda_0(t)$ and $h_0(t)$ is available. The specifications of the intervals $I_k^R = (t_{k-1}^R, t_k^R]$, $k = 1, ..., K^R$, and $I_k^D = (t_{k-1}^D, t_k^D]$, $k = 1, ..., K^D$, can differ between the recurrent event process and the death process in terms of both their lengths $\Delta_k^R = t_k^R - t_{k-1}^R$ and $\Delta_k^D = t_k^D - t_{k-1}^D$ and their total numbers K^R and K^D. (The intervals for the piecewise constant baseline rates should not, however, be confused with the intervals in which the numbers of recurrent events n_{ij} are observed; see Section 5.2.)

The baseline rate $\lambda_0(t)$ of the recurrence process enters the likelihood (5.5) through the conditional means $u_i \mu_{ij}$ of the interval counts N_{ij} given the frailty value u_i. Under the piecewise constant rate model, μ_{ij} is computed as

$$\mu_{ij} = \int_{I_{ij}} e^{\beta^\top z_i} \lambda_0(s) \mathrm{d}s = e^{\beta^\top z_i} \sum_{k=1}^{K^R} \lambda_{0k} \max\{0, \min(t_k^R, t_{ij}) - \max(t_{k-1}^R, t_{ij-1})\}.$$

Piecewise constant baseline rates offer more flexibility than parametric models, such as the Weibull model, while at the same time remaining more tractable than purely nonparametric models. Previous studies have suggested that a moderate number of intervals – i.e., between 8 and 10 intervals – yields satisfactory estimation results (Liu and Huang, 2008; Lawless and Zhan, 1998).

The performance of the piecewise constant model is usually improved when the interval cut-points t_k are based on quantiles of the recurrence and the survival times, respectively. In the current setting in which only interval counts of recurrent events are observed, the exact recurrence times are unknown. If, however, the individual observation intervals I_{ij} are relatively short compared with the total follow-up, we can approximate quantiles by creating a set from the observation times t_{ij}, with each repeated n_{ij} times, $j = 1, ..., J_i$, $i = 1, ..., m$; and then determining the cut-points t_k^R as quantiles of this set of times.

Parameter estimation in the joint frailty model is then done by maximizing the approximate marginal log-likelihood; that is, the logarithm of (5.5). The standard errors for the parameter estimates can be obtained from the inverse of the negative Hessian of the approximate marginal log-likelihood. Further computational details can be found in Section 5.7.1.

5.3.2 Score test for the association between recurrences and death

The joint frailty model for recurrent events and death is rather complex, with estimation procedures that are more involved than those needed for the fitting of two separate models, one for the recurrent event process and one for the survival process. Investigating whether the two processes are associated – and, consequently, whether joint modeling is required – is a useful first step. Moreover, the question of whether there is an association – and, if so, whether it is positive or negative – can be a stand-alone issue, not necessarily followed by fitting a joint model.

Balan et al. (2016) proposed a correlation score test for the association between recurrences and terminal events in the joint frailty model for settings with observed recurrence times. Their test can be performed by fitting separate models for the recurrence times and the survival data, and, thus, without fitting the joint frailty model. In addition, the sign of the test statistic is an indicator of the direction of the association. In the following, we show that this score test can be adapted to the setting with interval counts of recurrent events, provided the survival times are exactly observed.

A test for association in the joint frailty model (5.2) corresponds to a test of $H_0 : \gamma = 0$ against $H_1 : \gamma \neq 0$. All other parameters, including those for the baseline rate models, are treated as nuisance parameters, and denoted by $\boldsymbol{\eta}$. The score test of Balan et al. (2016) is based on the score function for γ under the null hypothesis; that is,

$$U_\gamma(0, \boldsymbol{\eta}) = \frac{\partial}{\partial \gamma} \ell(\gamma, \boldsymbol{\eta}) \mid_{\gamma=0},$$

where ℓ is the marginal log-likelihood of the joint frailty model. The authors showed that the score, evaluated at the maximum likelihood estimate under H_0, that is, $(0, \hat{\boldsymbol{\eta}}_0)$, is proportional to the covariance of the estimated martingale residuals of the terminal event and the 'posterior' estimates of the log-frailties for the recurrent events given the observed data. More formally, defining

$$K_i(u, t) = u^{N_i^R(t-)+\gamma N_i^D(t-)} \exp\left\{-ue^{\boldsymbol{\beta}^\top \mathbf{z}_i} \Lambda_0(t) - u^\gamma e^{\boldsymbol{\alpha}^\top \mathbf{z}_i} H_0(t)\right\}, \tag{5.7}$$

with the cumulative baseline rates $\Lambda_0(t) = \int_0^t \lambda_0(s)\,\mathrm{d}s$ and $H_0(t) = \int_0^t h_0(s)\,\mathrm{d}s$, Balan et al. (2016) derived that

$$\begin{aligned} U_\gamma(0, \hat{\boldsymbol{\eta}}_0) &= \sum_{i=1}^m \left[N_i^D(x_i) - e^{\hat{\boldsymbol{\alpha}}^\top \mathbf{z}_i} \widehat{H}_0(x_i)\right] \frac{\int_0^\infty \ln(u)\ \widehat{K}_i(u, x_i)\ g_\theta(u)\,\mathrm{d}u}{\int_0^\infty \widehat{K}_i(u, x_i)\ g_\theta(u)\,\mathrm{d}u} \\ &= \sum_{i=1}^m \widehat{M_i^D} \cdot \widehat{\ln(u_i)}\,. \end{aligned} \tag{5.8}$$

The $\widehat{M_i^D}$ are estimates of the martingale residuals of the terminal event model $M_i^D = N_i^D(x_i) - \int_0^{x_i} e^{\boldsymbol{\alpha}^\top \mathbf{z}_i}\, h_0(s)\mathrm{d}s$. The $\widehat{\ln(u_i)}$ are the posterior estimates of the log-frailty values

given the observed data from the recurrent event process: $\widehat{\ln(u_i)} = \mathrm{E}[\ln U_i \mid O_i(x_i)]$. For gamma distributed frailties u_i with mean one and variance θ, one obtains

$$\widehat{\ln(u_i)} = \psi\left(\frac{1}{\hat{\theta}} + N_i^R(x_i)\right) - \ln\left(\frac{1}{\hat{\theta}} + e^{\hat{\boldsymbol{\beta}}^\top \boldsymbol{z}_i}\widehat{\Lambda_0(x_i)}\right), \tag{5.9}$$

where $\psi(\cdot)$ is the digamma function.

Because of the zero-mean constraint of the $\widehat{M_i^D}$, the second line of (5.8) is proportional to the correlation $r = \mathrm{cor}(\widehat{\boldsymbol{M^D}}, \widehat{\ln(\boldsymbol{u})})$ between the martingale residuals and the estimated log-frailties. Consequently, Balan et al. (2016) based the correlation score test on the test statistic

$$t = r\sqrt{\frac{m-2}{1-r^2}}, \tag{5.10}$$

which, under the null hypothesis, asymptotically follows a t-distribution with $m-2$ degrees of freedom.

It turns out that in the setting in which only interval counts of the recurrent events are available, but exact survival times (or censoring times) are observed, equation (5.8) still holds. We show this result in Section 5.7.2. Therefore, the test statistic t in (5.10) is still valid. Furthermore, for gamma distributed frailties, the estimates $\widehat{\ln(u_i)}$ can still be determined using formula (5.9). The latter formula involves estimates $\hat{\theta}$ of the frailty variance, the covariate effect $\hat{\boldsymbol{\beta}}$ on the recurrence rate, and the cumulative baseline rate $\widehat{\Lambda_0}$; all determined under H_0. These can be obtained by fitting a mixed Poisson model to the interval counts of the recurrent events (see, for instance, Lawless and Zhan, 1998, who assume a piecewise constant baseline rate function λ_0). We generally recommend to estimate the shared frailty model for the interval counts using a flexible specification such as (5.6) for the baseline rate. A simple parametric model like, for example, the Weibull model, although appealing due to its parsimony, always bears the risk of misspecification and consequently misleading test results. The estimates of the martingale residuals $\widehat{M_i^D}$ can be derived from a Cox proportional hazards model fitted to the survival data $\{X_i, \delta_i, \boldsymbol{z}_i; i = 1, ..., m\}$.

5.4 Simulation study

5.4.1 Performance of the estimation method

To evaluate the performance of the proposed method for estimating the parameters of the joint frailty model based on interval counts of recurrent events and survival times, we conducted a simulation study.

Several different aspects will affect the estimation results, both on the part of the model specification but also on the part of the observable data. The latter include the number and lengths of the intervals for the recurrent event counts and whether they are the same for all individuals in the sample or not. The amount of independent censoring

will also have an impact. Among the former aspects, the size of the frailty variance and the sign of the dependence parameter are expected to influence the estimation, but also the number of intervals used in the piecewise constant specification of the baseline rates (and consequently the total number of parameters to be estimated) will matter.

We generated data for $m = 200$ individuals from the joint frailty model (5.2). For the baseline rates $\lambda_0(t)$ and $h_0(t)$, we chose the form of Weibull hazards, $(a/b)\,(t/b)^{a-1}$, with shape parameter a equal to 1.5 or 3 and scale parameter b equal to $1/3$ and 1.35, respectively. A single binary covariate that takes values 0 or 1 with probability 0.5 was included, and had the same effect on the hazard of death and the rate of recurrence ($\alpha = \beta = 1$). Frailties were simulated from a gamma distribution with mean one and different variances $\theta \in \{0.25, 0.5, 0.75\}$. We considered both cases of positive ($\gamma = 1$) and negative association ($\gamma = -1$) between the recurrence process and the terminal event. Additionally, we looked at one setting with a relatively small value for the frailty variance, $\theta = 0.05$, that was inspired by the results of the data set on *Eleutheria dichotoma*, see Section 5.5. In this setting, two values for the dependence parameter $|\gamma| = 1$ or $|\gamma| = 5$ were studied.

Independent censoring was considered in two versions: either an end-of-study censoring at time $t = 2$ for all individuals still alive then or individual censoring times which occurred uniformly over the total follow-up window $[0, 2]$.

Regarding the observation times t_{ij}, which determine the intervals during which the recurrent events are counted, we examined two scenarios. In Scenario I the observation times were the same for all individuals. We set $t_{ij} = 0.2j$, $j = 0, ..., 10$, such that we observed individual interval counts in up to 10 intervals of equal length 0.2. We also studied one scenario with $t_{ij} = 0.1j$, $j = 0, ..., 20$, leading to up to 20 individual interval counts. In Scenario II the observation times t_{ij} varied across individuals. This scenario mimicked a study in which participants have scheduled visit times, but actually could be early or late for their visits. This was implemented as follows: the scheduled times were $t_j^0 = 0.2j$, $j = 0, ..., 10$, but the actual observation times for each individual were obtained by adding a random noise ε_{ij} to the scheduled times so that $t_{ij} = t_j^0 + \varepsilon_{ij}$. The ε_{ij} were drawn from a uniform distribution on $[-0.1, 0.1]$ (except for $\varepsilon_{i0} = 0$ and ε_{i10} as uniform on $[-0.1, 0]$) and hence the t_{ij} varied across individuals.

For the estimation of the joint frailty model, we assumed that the frailties were gamma distributed, and that the baseline rates were specified as piecewise constant functions on 10 intervals of equal length, $t_k^R = t_k^D = 0.2k$. Thus, under Scenario I and $t_{ij} = 0.2j$ the intervals for the rate pieces coincided with the observation intervals for the counts. We also studied the impact of an increase of the number of intervals to 20. Approximation of the likelihood was based on $Q = 30$ quadrature points. All computations were run in `R` (R Core Team, 2019), see Section 5.7.1 for further details. In all settings we ran 200 replications.

The results of the simulation study under Scenario I with frailty variance $\theta = 0.5$ are shown in Figures 5.1 and 5.2, both for positive and negative dependence $|\gamma| = 1$ and for both independent censoring schemes (A: end-of-study censoring at $t = 2$, and B: uniform censoring times in $[0, 2]$).

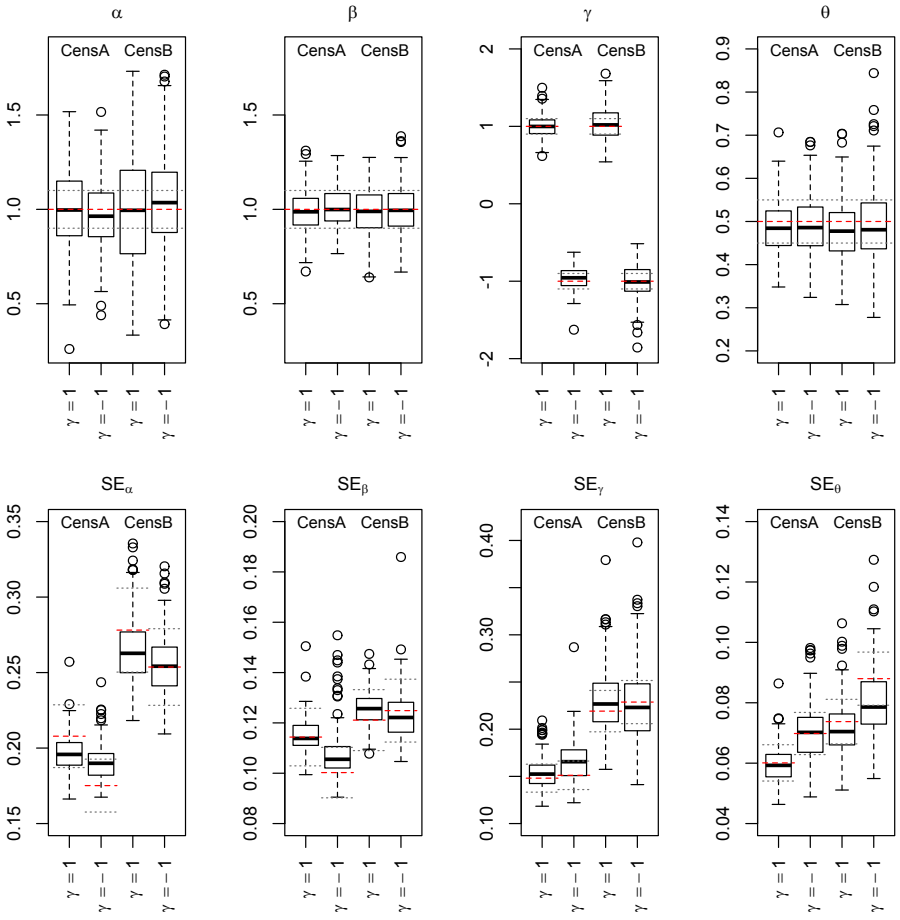

Figure 5.1: Box plots of the parameter estimates (top) and estimated standard errors (bottom) in the joint frailty model for positive ($\gamma = 1$) and negative dependence ($\gamma = -1$) under two schemes of independent censoring (CensA: at end of study $C = 2$, CensB: uniform on $[0, 2]$). Left to right: covariate effect on mortality (α) and on recurrences (β), dependence parameter (γ), and frailty variance (θ), based on 200 samples of size $m = 200$. Red, dashed line marks true parameter value (top) or empirical standard deviation (bottom); gray, dotted lines mark 10% deviations from respective value.

From the box plots of the estimates of the covariate effects α and β, the dependence parameter γ, and the frailty variance θ in Figure 5.1, we can see that the method performs satisfactorily. Censoring scheme A is considerably milder than scheme B (for an illustration see Table 5.4 in Section 5.8.1), and the corresponding loss in information is reflected in a moderately increased variability in the estimates and, as could be expected, also in the estimated standard errors (Figure 5.1, bottom). For the frailty variance θ, variability in estimates and estimated standard errors is higher for negative dependence ($\gamma = -1$) than for positive dependence. This is also true for the dependence parameter γ itself. The estimated standard errors for the covariate effects α and β, the dependence parameter γ, and the frailty variance θ are compared with the empirical standard deviations of the respective estimates across the replications in the bottom panels of Figure 5.1. The general magnitude of the standard errors is well captured.

Figure 5.2 displays the estimates of the cumulative baseline rates for positive dependence for both censoring schemes A and B. The averages of the estimates are very close to the true underlying rates. The stronger loss of observations in scheme B leads to markedly increased variability of the estimates toward the end of the follow-up window, which is an obvious consequence of the decreasing number of individuals under study over time. This low number of observations results in part from individuals experiencing a terminal event. Additionally, even fewer individuals (in this setting, less than 5%) are observed in the interval $[1.8, 2)$, which corresponds to the last piece of the baseline rates, due to independent censoring under scheme B. In contrast, censoring scheme A is benign.

For the other settings of the simulation study we restricted our attention to the independent censoring scheme B. With its uniformly distributed censoring times it produces a relatively challenging loss of observations. So the results we report here are conservative in the sense that they hold despite this appreciable amount of censoring. Section 5.8.1 illustrates the results of the other simulation settings analogous to Figures 5.1 and 5.2. Here we summarize the main results of the various scenarios.

If we change the frailty variance to $\theta = 0.25$ or $\theta = 0.75$, we note that also in these settings the biases of the parameter estimates, if present at all, are small. In general, varying the true underlying frailty variance mainly affects the variability of the estimates. The estimates of the covariate effects and the frailty variance show increased variability for larger frailty variance. In contrast, the estimates of the dependence parameter show less variability for larger frailty variance, because the increased heterogeneity among the individual patterns makes it easier to assess the dependence between the recurrence process and survival.

If the true frailty variance is close to zero ($\theta = 0.05$) the estimates of the dependence parameter are still only modestly biased but the variability of the $\hat{\gamma}$ increases strongly. This had to be expected, since the identification of the dependence hinges on the variation in the individual frailty. If the true value of $|\gamma| = 1$, then the estimated values may turn out with the wrong sign of the dependence parameter, particularly if the true dependence is negative. Thus, weak dependence is difficult to identify in case of low variability in

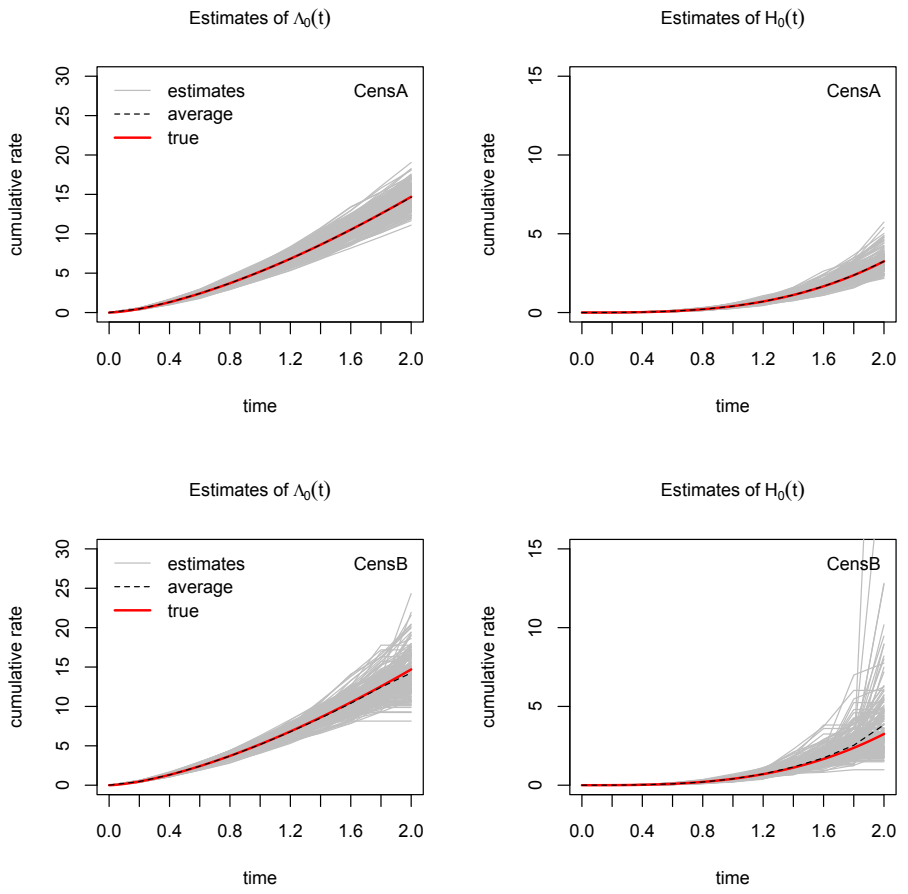

Figure 5.2: Estimates (gray, solid) of the cumulative rate of recurrence (left) and of death (right) based on samples of size $m = 200$ which were generated from a joint frailty model with positive dependence for two schemes of independent censoring. Top: end-of-study censoring at $t = 2$ (CensA), bottom: uniform censoring on $[0, 2]$ (CensB). Red, solid line gives the true cumulative baseline rate; black, dashed line is the mean of the 200 estimates.

frailty. However, if the dependence is strong, $|\gamma| = 5$, then, despite the variability of the estimates, the sign of the dependence is correctly estimated.

If we modify the width of the observation intervals from 0.2 to 0.1 we first note the following: if we keep the specification of the baseline rate of recurrence as piecewise constant on the 10 intervals with $t_k^R = 0.2k$ as before, then the additional information provided by the finer observation intervals would not be used by the estimation method

(see Section 5.7.3 for a proof). Therefore, we also specify the baseline rates as piecewise constant functions on 20 intervals, that is $t_k^R = t_k^D = 0.1k$. Apart from increased variability in the rate estimates near the end of the follow-up window, there is little change to the estimation results with width 0.2.

If we allow the observation intervals to vary across individuals (Scenario II), then we find that the method performs equally well as for fixed observation intervals, with the exception of the estimates for the last baseline rate pieces. In this simulation, we used fixed intervals for the baseline rates though, to unify the presentation of results. In applications we would recommend to choose the cut-points t_k^R and t_k^D based on (approximate) quantiles of the event times, as described earlier, to increase precision of the baseline rate estimates.

If we replace the baseline rates by a parametric Weibull specification (which here is correct), then there is only very modest change in the other parameter estimates of the model as well as their standard errors, so the main advantage is in the less variable estimation of the two baseline rates. This is, however, counterbalanced by the risk of model misspecification so we advise the use of piecewise constant rates unless one has a good understanding of the underlying processes that justifies a parametric choice.

Overall, the proposed method for estimating the joint frailty model based on interval counts yielded reliable results in this simulation study and different scenarios behaved in the way we had anticipated beforehand.

5.4.2 Performance of the score test

In a second set of simulations, we investigated the performance of the score test in the setting with interval counts of recurrent events.

For this purpose, we again generated data for $m = 200$ individuals from the joint frailty model (5.2). The covariate effects, values for the variance of the gamma frailties, and baseline rates were the same as in Section 5.4.1. Here, however, we considered not only different directions but also different strengths of the association between the recurrence process and the terminal event; namely, $\gamma \in \{-1, -0.5, 0, 0.5, 1\}$. Counts were again observed in 10 intervals of equal length 0.2 (Scenario I) or of varying length (Scenario II). We studied two schemes of independent censoring, scheme A with censoring time $C = 2$, and scheme B with uniformly distributed censoring times on $[0, 2]$. We ran 1000 replications for each setting to determine the size or power of the test.

The score test involves fitting separate models to the recurrence data and the survival data. First, we fitted a shared gamma frailty model to the individual interval counts of the recurrent events, assuming that $(N_{ij} \mid u_i)$ follows a Poisson distribution with mean $u_i \tilde{\mu}_{ij}$, where $\tilde{\mu}_{ij} = \int_{I_{ij}} e^{\tilde{\beta}^\top z_i} \tilde{\lambda}_0(s)\, \mathrm{d}s$. The baseline recurrence rate $\tilde{\lambda}_0(\cdot)$ was modeled as piecewise constant, as in (5.6), with pieces defined by the cut-points $t_k^R = 0.2k$, $k = 0, ..., 10$. The estimates $\widehat{\ln(u_i)}$ were then determined according to formula (5.9). Second, we estimated a Cox proportional hazards model from the survival data $\{X_i, \delta_i, z_i; i = 1, ..., m\}$ using function `coxph()` from package `survival` (Therneau and Grambsch,

2000) to obtain the martingale residuals $\widehat{M_i^D}$ of the terminal event. Finally, we calculated the test statistic based on the correlation between the $\widehat{M_i^D}$ and the $\widehat{\ln(u_i)}$.

Table 5.1 reports the size and the power of the score test, performed at a level of 5%, depending on the true underlying dependence parameter γ and frailty variance θ. The proportion of a type I error, that is, of falsely rejecting the hypothesis of no dependence ($\gamma = 0$), was affected most strongly by the censoring scheme. If $C = 2$, which implies a modest proportion of independently censored cases (see Table 5.6 in Section 5.8.2), then the level of the test is met or exceeded only slightly. For strong independent censoring (scheme B), which implies that roughly half of the observations are censored and quite some of them early during the follow-up, then this loss of information increases the proportion of a type I error, particularly for the frailty variance $\theta = 0.75$. This is a noteworthy result, and hence the test should be regarded with caution, if there is strong (and early) censoring in the data.

Table 5.1: Power and size of the score test, performed at the 5% level, in the joint frailty model with varying dependence parameter $\gamma \in \{-1, -0.5, 0, 0.5, 1\}$, frailty variance $\theta \in \{0.25, 0.5, 0.75\}$, and independent censoring $C \sim \mathcal{U}[0, 2]$ or $C = 2$, across 1000 replications each.

		Dependence γ				
censoring	θ	-1	-0.5	0	0.5	1
$C \sim \mathcal{U}[0, 2]$	0.25	0.960	0.586	0.066	0.514	0.932
	0.5	1.000	0.905	0.062	0.872	1.000
	0.75	1.000	0.984	0.091	0.974	1.000
$C = 2$	0.25	0.999	0.788	0.057	0.799	1.000
	0.5	1.000	0.990	0.039	0.992	1.000
	0.75	1.000	1.000	0.040	1.000	1.000

Regarding the power of the score test, as expected, we found that the power increased with the strength of the dependence, that is, $|\gamma|$, and with the magnitude of the frailty variance. In the settings with the larger frailty variances $\theta = 0.5$ and $\theta = 0.75$ and stronger dependence $|\gamma| = 1$, the score test detected the association in all cases. We note that for the samples for which the association was detected by the score test, the direction of the dependence was always identified correctly. An extension of the simulation settings suggested that the score test can detect associations even for small values of the frailty variance as long as the association is sufficiently strong, that is, $|\gamma|$ is sufficiently large (see Table 5.7 in Section 5.8.2).

Furthermore, we assessed the performance of the score test for Scenario II in which the observation intervals vary across individuals. Results for the setting with 10 scheduled visit times and 10 pieces for the baseline rate are displayed in Table 5.10 of Section 5.8.2. The proportion of false rejections of the hypothesis of no dependence are again a bit higher than the nominal level due to the high percentage of censoring. The power of the

test is comparable to the values obtained in Scenario I, although the power in the settings with negative dependence is a bit lower here.

Additional results on the performance of the score test for modifications of Scenario I and II are given in Section 5.8.2.

In conclusion, the results of the simulation study provide further evidence that the score test is a powerful method for assessing the association between recurrent events and the terminal event, also in a setting in which only interval counts of the recurrent events are observed.

5.5 Fertility and mortality in *Eleutheria dichotoma*

To illustrate the proposed methods, we use data from a biodemographic study on the fertility and mortality of *Eleutheria dichotoma* (Dańko et al., 2020), which was briefly introduced in Section 5.1. *E. dichotoma* is a marine organism that passes through several life-cycle stages: i.e., planula larva, polyp, and medusa stages. The colonial polyps (hydroids) asexually produce medusae. The medusae can then reproduce both sexually (by producing larvae) and asexually (by producing medusa buds).

In this study, asexually budded medusae were collected directly from the hydroid colony and reared for three generations – each of which was, in turn, obtained through the asexual budding of medusae. In our analysis, we focus on the age trajectory of the budding rate (asexual reproduction) and the mortality of one of the medusa generations, as well as on the association between the two processes.

Age $t_0 = 0$ of a medusa corresponds to the point in time when it detaches from its hydroid colony or from its ancestor medusa. The medusae were followed individually, and were observed until death or censoring. This occurred when either the study ended, a laboratory accident (e.g., water evaporation) led to the loss of the medusa, or the medusa was absorbed by a large bud of the same individual. The animals were checked for newly released larvae and buds roughly three times per week. The resulting observation times differed across individuals, with interval lengths between 1 and 11 days.

Salinity is an important factor that affects the physiological responses of species like *E. dichotoma*, both at the level of the hydroid colony that produced the medusae under study, and at the level of the medusae themselves. Four combinations of salinity levels were studied here (low(hydroid)–low(medusa), medium–medium, low–medium, and medium–low); for more details, see Dańko et al. (2020).

The data set contains $m = 141$ individuals, with the following group sizes in the four experimental conditions: 36 low–low, 40 medium–medium, 32 low–medium, and 33 medium–low. The individuals produced between 0 and 27 buds over their life course, with a mean of 8.99. Follow-up times varied between 9 and 217 days, with a median of 98 days; and 34 individuals (24%) were censored. Figure 5.3 exemplarily shows the data for 14 individuals.

To assess whether the recurrent budding process and survival are associated in *E. dichotoma*, we first conducted the score test that was presented in Section 5.3.2. For fitting

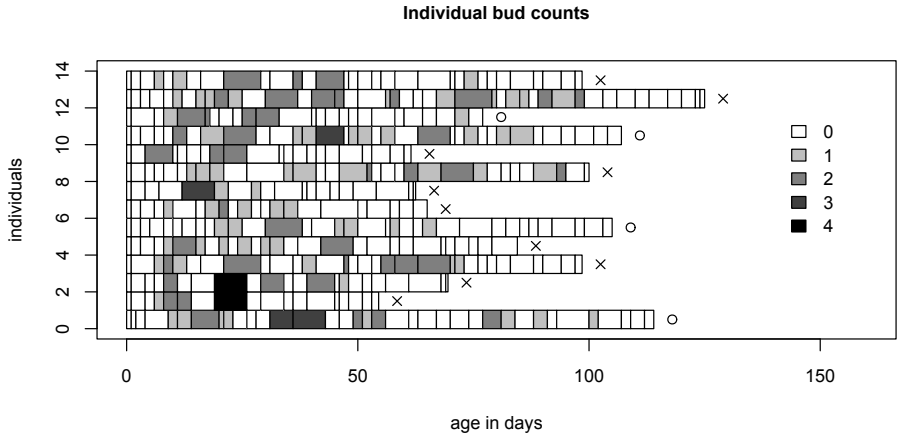

Figure 5.3: Interval counts of budding and survival (dead: cross, censored: circle) for 14 medusae *E. dichotoma*.

the shared frailty model to the individual bud counts, we assumed that the frailty variable was gamma distributed. Moreover, for the piecewise constant baseline budding rate, we defined cut-points t_k^R at 0, 16, 22, 27, 32, 40, 47, 63, 85, 111, and 220 days. These cut-points were obtained as (approximative) deciles of the recurrent event times, as described at the end of Section 5.3.1. We included two binary covariates for the salinity levels at the polyp stage and the medusa stage, respectively.

The martingale residuals of the terminal event were obtained from a Cox proportional hazards model fitted to the survival data, including the two covariates on the salinity levels.

The parameter estimates of the two separately fitted models, the Cox model for the survival data, and the shared frailty model for the budding rate based on interval counts, are shown in Table 5.2.

The correlation between the martingale residuals from the Cox regression and the estimated log-frailties from the budding model is $r = -0.434$, yielding a test statistic of $t = -5.676$ with a p-value of $7.729 \cdot 10^{-8}$. Thus, the result of the score test clearly suggests that reproduction and mortality are negatively associated – i.e., that a higher budding rate is associated with lower mortality – which implies that a joint model should be used to analyze these data on *E. dichotoma*.

Therefore, we estimated a joint frailty model for reproduction and mortality for these data, including the two covariates on salinity. Frailties were again assumed to stem from a gamma distribution. For the baseline budding rate, the same specification was used as in the separate model (see above). For the baseline hazard of death, we used a piecewise

Table 5.2: Parameter estimates (with standard errors) for different models fitted to the *E. dichotoma* data set.

	Joint frailty model	Separate models	
		Shared frailty model	Cox PH model
Mortality			
Polyp (low salinity)	−0.493 (0.299)	–	−0.376 (0.201)
Medusa (low salinity)	2.201 (0.432)	–	1.392 (0.218)
Budding			
Polyp (low salinity)	0.052 (0.069)	0.057 (0.066)	–
Medusa (low salinity)	0.574 (0.071)	0.600 (0.068)	–
Association			
Dependence γ	−4.941 (1.557)	–	–
Frailty variance θ	0.051 (0.018)	0.036 (0.017)	–

constant function with cut-points t_k^D at $0, 60, 67, 72, 85, 98, 103, 114, 132, 153$, and 220 days; again, the cut-points were taken from approximate deciles of the survival times. For the Gaussian quadrature in the marginal likelihood (see equation (5.5)), $Q = 30$ quadrature points were used.

The parameter estimates of the joint model are also displayed in Table 5.2. Interestingly, the dependence parameter γ was estimated as -4.941 along with a frailty variance of 0.051. The negative value of $\hat{\gamma}$ indicated that individuals with higher rates of asexual reproduction tended to have a lower mortality risk than individuals with lower rates of asexual reproduction. This finding that higher rates of reproduction were accompanied by longer survival stands in contrast to the idea of a trade-off between reproduction and survival. Regarding the salinity levels, we found that the salinity experienced by the polyps did not have a noticeable effect on the reproduction or survival of the medusae. In contrast, individuals who were exposed to low salinity at the medusa stage were found to have both higher fertility and mortality rates than those exposed to medium salinity. Finally, Figure 5.4 shows the estimates of the age-specific budding and death rates. The budding rate peaked at about 20 days of age before gradually declining to a non-zero level at later ages. The death rates increased markedly after about 60 days of age.

To conclude, while the results presented here are in agreement with the findings of Dańko et al. (2020), they also provide additional insight into the association between the asexual reproduction and survival of *E. dichotoma*. In addition to the biological implications of the dependence itself, due to the association between the two processes, the recurrent event process is censored by the terminal event (death) non-independently, which warrants the joint modeling of the two processes.

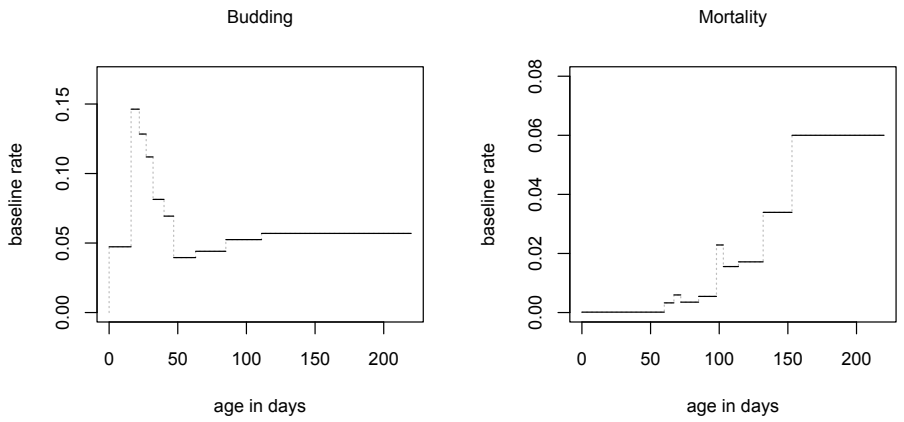

Figure 5.4: Estimated baseline rates of budding (left) and death (right) for the *E. dichotoma* data set.

5.6 Discussion

We have presented a method for estimating the joint frailty model for recurrent events and death in situations in which only individual interval counts of the recurrence process are observed. When modeling the baseline rates as piecewise constant, the marginal likelihood can be approximated using Gaussian quadrature, and can then be maximized directly. In addition, we have shown that the score test for the association between recurrences and death (Balan et al., 2016) is also applicable in the setting with interval counts. The test is based on the correlation between the martingale residuals of the terminal event and the estimates of the log-frailties, which are obtained by separately fitting a Cox proportional hazards model to the survival data and a shared frailty model to the interval counts of recurrent events. Our simulation studies demonstrated that both the estimation method and the score test perform well.

We also found that when applying the proposed methods to data on fertility and mortality in *E. dichotoma*, the rate of asexual reproduction and the mortality risk are negatively associated. While this finding is interesting from a biological point of view, it also demonstrates the necessity of allowing positive and negative dependence in a model such as the relatively complex joint frailty model (5.2). Moreover, the *E. dichotoma* example illustrates the advantages of using the piecewise constant rate model. On the one hand, the shape of the budding rate, which is characterized by a sharp peak at younger ages and a gradual leveling off to a non-zero level at older ages, is hard to capture using a simple parametric model. On the other hand, the data structure of the interval counts,

in which the observation times vary across individuals, makes it difficult to construct a purely nonparametric estimate of the baseline rate.

Implementing the estimation method in the joint frailty model involves making several choices, such as decisions regarding the distribution of the frailties, the number of quadrature points, and the number of pieces included in the baseline rate models. For the frailty distribution, we assumed that frailties follow a gamma distribution in both the simulation study and the application. The use of a gamma distribution is a popular choice for the distribution of frailties, and, with respect to the score test, it has the benefit of yielding closed-form expressions for the estimated log-frailties. However, the quadrature approach is equally able to accommodate the log-normal distribution or any frailty distribution that has a closed-form inverse distribution function. Yet, the performance of the quadrature approach relies on the quality of the approximation of the marginal likelihood, which, in turn, depends on the number of quadrature points. Liu and Huang (2008) suggested using $Q = 30$ quadrature points for gamma frailty models, and in our experience, this number yields reliable results. In practice, we recommend fitting the model for several increasing values of Q until the estimates stabilize. For the piecewise constant rate models, using a moderate number of up to 10 intervals for the rates seems to produce good results. Using a larger number of intervals for the baseline rates generates a larger number of parameters to be estimated, which, in turn, increases computational costs, and might affect the numeric stability of the method.

One of the limitations of the approach presented here is that, in the joint frailty model, the association between recurrences and death is modeled via a dependence parameter acting on a shared frailty. If, however, the individuals are not sufficiently heterogeneous – that is, if the frailty variance is not sufficiently large – the dependence parameter is not meaningful, and the association cannot be assessed.

Another restriction is imposed by the piecewise constant rate models. While these models can capture a variety of different shapes of the baseline rates, using more flexible rates – and, in particular, smooth rates – might be desirable in some applications. Further work is needed on how to incorporate smooth rate models with automatic smoothing parameter selection.

Finally, the observation times in our application are fixed by the experimental set-up, even though in real-world applications, particularly in medicine, the observation times may depend on the recurrence process. For instance, patients may visit the doctor more often when they are in worse condition. This is taken into account in a model developed by Zhao et al. (2013), which considers interval counts of recurrent events in the presence of death and a dependent observation process. However, the authors adopted a marginal approach that left the association between the recurrences and death unspecified.

5.7 Appendix

5.7.1 Computational details

We implemented our estimation approach in `R` (R Core Team, 2019), using function `gauss.quad()` from package `statmod` (Smyth, 1998) to determine the quadrature points and weights, and function `nlm()` for numerical optimization of the approximate marginal log-likelihood.

The Hessian of the log-likelihood can be obtained directly from the output of the function `nlm()`. However, in some cases, computation of the Hessian using function `hessian()` from package `numDeriv` (Gilbert and Varadhan, 2019) yields more stable results.

To ensure that the Hessian of the approximate log-likelihood with piecewise constant baseline rates is invertible, it can be necessary to fit the joint frailty model with small, fixed ridge penalties on the logarithm of the baseline rates. The standard errors are then calculated based on the Hessian of the penalized log-likelihood.

5.7.2 Derivation of the score test

In this section, we show that the score $U_\gamma(\gamma, \boldsymbol{\eta})$ of the joint frailty model (5.2) has the same form independently of whether exact recurrence times or only interval counts of the recurrent events are available, as long as the recurrence process is observed up to exactly known follow-up times.

Let us start by rewriting the individual contributions (5.3) to the marginal likelihood for interval counts

$$
\begin{aligned}
L_i = &\int_0^\infty \prod_{j=1}^{J_i} \left[\frac{1}{n_{ij}!} \exp\left\{ -\int_{I_{ij}} u e^{\boldsymbol{\beta}^\top \mathbf{z}_i} \lambda_0(s)\, \mathrm{d}s \right\} \left(\int_{I_{ij}} u e^{\boldsymbol{\beta}^\top \mathbf{z}_i} \lambda_0(s)\, \mathrm{d}s \right)^{n_{ij}} \right] \\
&\cdot \left[u^\gamma e^{\boldsymbol{\alpha}^\top \mathbf{z}_i} h_0(x_i) \right]^{\delta_i} \exp\left\{ -\int_0^{x_i} u^\gamma e^{\boldsymbol{\alpha}^\top \mathbf{z}_i} h_0(s) \mathrm{d}s \right\} g_\theta(u)\, \mathrm{d}u \\
= &\prod_{j=1}^{J_i} \left[\frac{1}{n_{ij}!} \left(\int_{I_{ij}} e^{\boldsymbol{\beta}^\top \mathbf{z}_i} \lambda_0(s)\, \mathrm{d}s \right)^{n_{ij}} \right] \left[e^{\boldsymbol{\alpha}^\top \mathbf{z}_i} h_0(x_i) \right]^{\delta_i} \cdot \int_0^\infty \left[\prod_{j=1}^{J_i} u^{n_{ij}} \right] \\
&\cdot u^{\gamma \delta_i} \exp\left\{ -u \sum_{j=1}^{J_i} \int_{I_{ij}} e^{\boldsymbol{\beta}^\top \mathbf{z}_i} \lambda_0(s)\, \mathrm{d}s \right\} \exp\left\{ -\int_0^{x_i} u^\gamma e^{\boldsymbol{\alpha}^\top \mathbf{z}_i} h_0(s) \mathrm{d}s \right\} g_\theta(u)\, \mathrm{d}u \\
= &\prod_{j=1}^{J_i} \left[\frac{1}{n_{ij}!} \mu_{ij}^{n_{ij}} \right] \left[e^{\boldsymbol{\alpha}^\top \mathbf{z}_i} h_0(x_i) \right]^{\delta_i} \\
&\cdot \int_0^\infty u^{n_i + \gamma \delta_i} \exp\left\{ -u e^{\boldsymbol{\beta}^\top \mathbf{z}_i} \Lambda_0(x_i) \right\} \exp\left\{ -u^\gamma e^{\boldsymbol{\alpha}^\top \mathbf{z}_i} H_0(x_i) \right\} g_\theta(u)\, \mathrm{d}u
\end{aligned}
$$

$$= \prod_{j=1}^{J_i} \left[\frac{1}{n_{ij}!} \mu_{ij}^{n_{ij}} \right] \left[e^{\boldsymbol{\alpha}^\top \mathbf{z}_i} h_0(x_i) \right]^{\delta_i} \cdot \int_0^\infty K_i(u, x_i)\, g_\theta(u)\, \mathrm{d}u,$$

with $n_i = \sum_{j=1}^{J_i} n_{ij}$ and $K_i(u,t)$ as defined in (5.7).

The marginal log-likelihood $\ell(\gamma, \boldsymbol{\eta}) = \sum_{i=1}^m \ln L_i$ takes the form

$$\sum_{i=1}^m \sum_{j=1}^{J_i} \left[-\ln(n_{ij}!) + n_{ij} \ln(\mu_{ij}) \right] + \delta_i [\boldsymbol{\alpha}^\top \mathbf{z}_i + \ln(h_0(x_i))] + \ln \int_0^\infty K_i(u, x_i)\, g_\theta(u)\, \mathrm{d}u,$$

and the score with respect to γ reads

$$U_\gamma(\gamma, \boldsymbol{\eta}) = \sum_{i=1}^m \frac{\int_0^\infty \left[\frac{\partial}{\partial \gamma} K_i(u, x_i) \right] g_\theta(u)\, \mathrm{d}u}{\int_0^\infty K_i(u, x_i)\, g_\theta(u)\, \mathrm{d}u}.$$

Since

$$\begin{aligned}
\frac{\partial}{\partial \gamma} K_i(u,t) &= u^{N_i^R(t-) + \gamma N_i^D(t-)} \cdot N_i^D(t-) \cdot \ln(u) \cdot \exp\left\{ -u e^{\boldsymbol{\beta}^\top \mathbf{z}_i} \Lambda_0(t) - u^\gamma e^{\boldsymbol{\alpha}^\top \mathbf{z}_i} H_0(t) \right\} \\
&\quad - u^{N_i^R(t-) + \gamma N_i^D(t-)} \cdot \exp\left\{ -u e^{\boldsymbol{\beta}^\top \mathbf{z}_i} \Lambda_0(t) - u^\gamma e^{\boldsymbol{\alpha}^\top \mathbf{z}_i} H_0(t) \right\} \\
&\quad \cdot u^\gamma \cdot e^{\boldsymbol{\alpha}^\top \mathbf{z}_i} H_0(t) \cdot \ln(u) \\
&= K_i(u,t) \left[N_i^D(t-) \cdot \ln(u) - u^\gamma \cdot e^{\boldsymbol{\alpha}^\top \mathbf{z}_i} H_0(t) \cdot \ln(u) \right],
\end{aligned}$$

we have

$$U_\gamma(\gamma, \boldsymbol{\eta}) = \sum_{i=1}^m \frac{\int_0^\infty \left[N_i^D(x_i) - u^\gamma \cdot e^{\boldsymbol{\alpha}^\top \mathbf{z}_i} H_0(x_i) \right] \ln(u)\, K_i(u, x_i)\, g_\theta(u)\, \mathrm{d}u}{\int_0^\infty K_i(u, x_i)\, g_\theta(u)\, \mathrm{d}u}.$$

If we now evaluate the score at $(\gamma, \boldsymbol{\eta}) = (0, \hat{\boldsymbol{\eta}}_0)$, we obtain

$$U_\gamma(0, \hat{\boldsymbol{\eta}}_0) = \sum_{i=1}^m \frac{\int_0^\infty \left[N_i^D(x_i) - e^{\hat{\boldsymbol{\alpha}}^\top \mathbf{z}_i} \widehat{H}_0(x_i) \right] \ln(u)\, \widehat{K}_i(u, x_i)\, g_\theta(u)\, \mathrm{d}u}{\int_0^\infty \widehat{K}_i(u, x_i)\, g_\theta(u)\, \mathrm{d}u},$$

which is the first line of (5.8).

5.7.3 Likelihood with fixed observation times

In this section, we study the likelihood of the joint frailty model (5.2) based on individual interval counts of recurrent events in one particular setting. Specifically, the observation times are assumed to be the same for all individuals and the baseline rate of recurrence $\lambda_0(t)$ is modeled as a piecewise constant function. We will show that if the

observation intervals are finer than the intervals for the rate pieces, the score depends on the individual's interval counts only through the sums of these counts over each baseline rate piece.

For that purpose, suppose the observation times t_{ij} are given by $t_{ij} = t_j$ for $j = 0, ..., J_i - 1$, and the last observation time $t_{iJ_i} = X_i$ is equal to the follow-up time X_i. The interval counts n_{ij}, observed over the intervals $I_{ij} = (t_{ij-1}, t_{ij}]$, enter the likelihood contribution $L_i^{(c)}(u_i)$ of individual i given its frailty value u_i (see (5.3)) only through the factor

$$\prod_{j=1}^{J_i} \frac{1}{n_{ij}!} \exp(-u_i\mu_{ij})(u_i\mu_{ij})^{n_{ij}} = \left(\prod_{j=1}^{J_i} \frac{1}{n_{ij}!}\right) \cdot u_i^{n_i} \cdot \exp\left(-u_i \sum_{j=1}^{J_i} \mu_{ij}\right) \prod_{j=1}^{J_i} \mu_{ij}^{n_{ij}} \quad (5.11)$$

with $n_i = \sum_{j=1}^{J_i} n_{ij}$ and $\mu_{ij} = \int_{I_{ij}} e^{\beta^\top z_i} \lambda_0(s)\, \mathrm{d}s$.

For a baseline rate of recurrence specified as a piecewise constant function on intervals $I_k^R = (t_{k-1}^R, t_k^R]$ of length $\Delta_k^R = (t_k^R - t_{k-1}^R)$, $k = 1, ..., K^R$, as in (5.6), we have

$$\mu_{ij} = e^{\beta^\top z_i} \sum_{k=1}^{K^R} \lambda_{0k} \int_0^\infty \mathbb{1}\{s \in I_{ij} \cap I_k^R\} \mathrm{d}s = e^{\beta^\top z_i} \sum_{k=1}^{K^R} \lambda_{0k} |I_{ij} \cap I_k^R|, \quad (5.12)$$

where $|I|$ denotes the length of the interval I. Now let K_i be the index k of the interval I_k^R for which $X_i \in I_{K_i}^R$. We can write the sum in the third term in (5.11) as

$$\begin{aligned}\sum_{j=1}^{J_i} \mu_{ij} &= e^{\beta^\top z_i} \sum_{j=1}^{J_i} \sum_{k=1}^{K^R} \lambda_{0k} |I_{ij} \cap I_k^R| = e^{\beta^\top z_i} \sum_{k=1}^{K^R} \lambda_{0k} \sum_{j=1}^{J_i} |I_{ij} \cap I_k^R| \\ &= e^{\beta^\top z_i} \sum_{k=1}^{K^R} \lambda_{0k} \Big|\bigcup_{j=1}^{J_i} I_{ij} \cap I_k^R\Big| = e^{\beta^\top z_i} \left\{ \sum_{k=1}^{K_i - 1} \lambda_{0k} \Delta_k^R + \lambda_{0K_i} (X_i - t_{K_i - 1}^R) \right\}, \quad (5.13)\end{aligned}$$

which depends only on the individual's follow-up time X_i, but not on the other observation times t_{ij}, $j = 0, ..., J_i - 1$.

Using (5.12), the last term in (5.11) equals

$$\prod_{j=1}^{J_i} \mu_{ij}^{n_{ij}} = (e^{\beta^\top z_i})^{n_i} \prod_{j=1}^{J_i} \left(\sum_{k=1}^{K^R} \lambda_{0k} |I_{ij} \cap I_k^R| \right)^{n_{ij}}.$$

Suppose now we choose the intervals I_k^R for the baseline rate of recurrence to be rougher than the observation intervals, but such that the t_k^R are a subset of the t_j. Then, each I_{ij}, $j = 1, ..., J_i$, is a subset of exactly one I_k^R, thus,

$$\prod_{j=1}^{J_i} \left(\sum_{k=1}^{K^R} \lambda_{0k} |I_{ij} \cap I_k^R| \right)^{n_{ij}} = \prod_{j=1}^{J_i} \prod_{k=1}^{K^R} \left(\lambda_{0k} |I_{ij} \cap I_k^R| \right)^{n_{ij} \cdot \mathbb{1}\{I_{ij} \subset I_k^R\}}$$

$$= \prod_{k=1}^{K^R} \prod_{j=1}^{J_i} \lambda_{0k}^{n_{ij} \cdot \mathbb{1}\{I_{ij} \subset I_k^R\}} |I_{ij} \cap I_k^R|^{n_{ij} \cdot \mathbb{1}\{I_{ij} \subset I_k^R\}}$$

$$= \left(\prod_{k=1}^{K^R} \lambda_{0k}^{\sum_{j=1}^{J_i} n_{ij} \cdot \mathbb{1}\{I_{ij} \subset I_k^R\}} \right) \left(\prod_{k=1}^{K^R} \prod_{j=1}^{J_i} (|I_{ij} \cap I_k^R|)^{n_{ij} \cdot \mathbb{1}\{I_{ij} \subset I_k^R\}} \right). \tag{5.14}$$

Note that $\tilde{n}_{ik} = \sum_{j=1}^{J_i} n_{ij} \cdot \mathbb{1}\{I_{ij} \subset I_k^R\}$ gives the number of recurrent events which individual i experiences over the time interval $[0, X_i] \cap I_k^R$. Hence, the first factor of expression (5.14) depends on the counts n_{ij} which were observed over the smaller intervals I_{ij} only via the counts $\tilde{n}_{ik}$ corresponding to the larger intervals I_k^R.

Inserting (5.13) and (5.14) into (5.11) yields

$$\left(\prod_{j=1}^{J_i} \frac{1}{n_{ij}!} \right) \cdot u_i^{n_i} \cdot \exp\left\{ -u_i e^{\boldsymbol{\beta}^\top \mathbf{z}_i} \left(\sum_{k=1}^{K_i - 1} \lambda_{0k} \Delta_k^R + \lambda_{0K_i} (X_i - t_{K_i-1}^R) \right) \right\}$$

$$\cdot (e^{\boldsymbol{\beta}^\top \mathbf{z}_i})^{n_i} \prod_{k=1}^{K^R} \lambda_{0k}^{\tilde{n}_{ik}} \prod_{k=1}^{K^R} \prod_{j=1}^{J_i} (|I_{ij} \cap I_k^R|)^{n_{ij} \cdot \mathbb{1}\{I_{ij} \subset I_k^R\}}$$

$$\propto (u_i \, e^{\boldsymbol{\beta}^\top \mathbf{z}_i})^{n_i} \, \exp\left\{ -u_i e^{\boldsymbol{\beta}^\top \mathbf{z}_i} \left(\sum_{k=1}^{K_i - 1} \lambda_{0k} \Delta_k^R + \lambda_{0K_i} (X_i - t_{K_i-1}^R) \right) \right\} \left(\prod_{k=1}^{K^R} \lambda_{0k}^{\tilde{n}_{ik}} \right).$$

Consequently, if the observation intervals are finer than the rate intervals the likelihood is proportional to an expression that depends only on recurrent event interval counts corresponding to the larger rate intervals.

5.8 Supplementary material: Extended simulation study

In this section, we present additional figures and tables illustrating the results of our simulation study. Subsection 5.8.1 focuses on the performance of the method for estimating the joint frailty model based on individual interval counts of recurrent events. Subsection 5.8.2 deals with the performance of the score test for the association between the recurrent events and death.

5.8.1 Performance of the estimation method

The basic set-up for the simulation study was described in Section 5.4.1. Table 5.3 gives an overview of the scenarios covered in this supplementary material. All samples include $m = 200$ individuals. Data were generated from a joint frailty model with a binary covariate having effects $\alpha = \beta = 1$ on the hazard of death and the rate of recurrence. Independent censoring occurs uniformly over the total follow-up window $[0, 2]$.

In Section 5.4.1, we focused on scenarios in which the values of the frailty variance θ and the dependence parameter γ are fixed at $\theta = 0.5$ and $\gamma = \pm 1$. Here, we also study the effect of varying the frailty variance by including scenarios with $\theta \in \{0.25, 0.75\}$. Moreover, we present settings with $\theta = 0.05$ which are motivated by the estimates obtained in the application (see Section 5.5).

Regarding the observation times t_{ij}, that define the intervals for which the recurrent event counts are observed, these times can be fixed and the same for all individuals (Scenario I) or varying across individuals (Scenario II), as explained in Section 5.4.1. In Scenarios I.G and II.D, the widths of these observation intervals are changed as well.

The baseline rates are generally specified as piecewise constant (pwc) functions on $K^R = K^D = K^*$ intervals of equal length, where we also assess the impact of increasing K^* from 10 to 20 (Scenarios I.F and I.G). In contrast, in Scenario I.E, a simple parametric model is used for the baseline rates by assuming (in this case correctly) a Weibull model.

Concerning the independent censoring, in Section 5.4.1, we have compared uniform censoring over the total follow-up window with the alternative of end-of-study censoring at $t = 2$. Table 5.4 reports on the resulting proportions of censoring under the different mechanisms across 200 replications.

Table 5.3: Simulation settings.

Scenario	$\lvert\gamma\rvert$	θ	obs. times t_{ij}	rates	K^*	Figure
Varying the frailty variance in Scenario I						
I.A	1	0.25	$0.2j$	pwc	10	5.5
I.B	1	0.75	$0.2j$	pwc	10	5.6
I.C	1	0.05	$0.2j$	pwc	10	5.7
I.D	5	0.05	$0.2j$	pwc	10	5.8
Varying the specification of the baseline rates in Scenario I						
I.E	1	0.5	$0.2j$	Weibull	-	5.9
I.F	1	0.5	$0.2j$	pwc	20	5.10
Varying the width of the observation intervals in Scenario I						
I.G	1	0.5	$0.1j$	pwc	20	5.11
Varying the frailty variance in Scenario II						
II.A	1	0.25	$0.2j + \varepsilon_{ij}$	pwc	10	5.12
II.B	1	0.5	$0.2j + \varepsilon_{ij}$	pwc	10	5.13
II.C	1	0.75	$0.2j + \varepsilon_{ij}$	pwc	10	5.14
Varying the width of the observation intervals in Scenario II						
II.D	1	0.5	$0.1j + \varepsilon_{ij}$	pwc	10	5.15

The ε_{ij} are drawn from a uniform distribution (see Section 5.4.1).

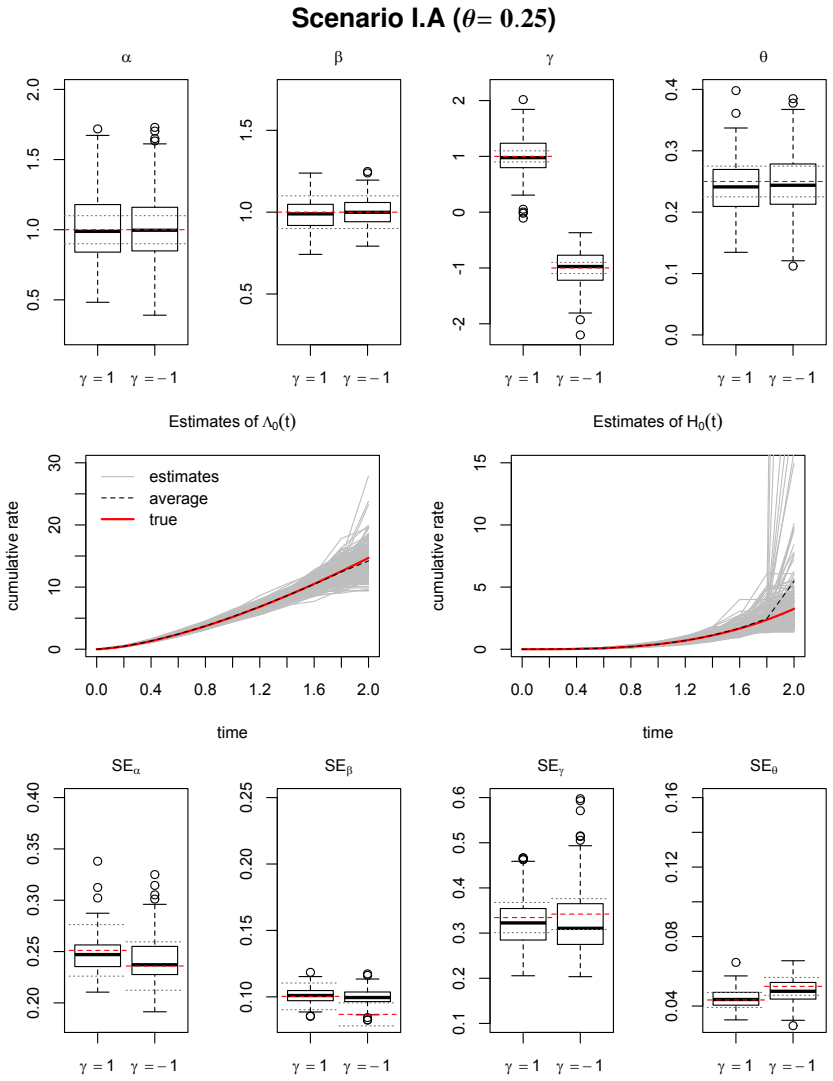

Figure 5.5: Box plots of the parameter estimates (top row) and estimated standard errors (bottom row) in the joint frailty model for positive ($\gamma = 1$) and negative dependence ($\gamma = -1$). Left to right: covariate effect on mortality (α) and on recurrences (β), dependence parameter (γ), and frailty variance (θ), based on 200 samples of size $m = 200$. Red, dashed line marks true value (top row) or empirical standard deviation (bottom row); gray, dotted lines mark 10% deviations from respective value. Middle row: estimates (gray, solid) of the cumulative rate of recurrence (left) and of death (right) in the joint frailty model with $\gamma = 1$.

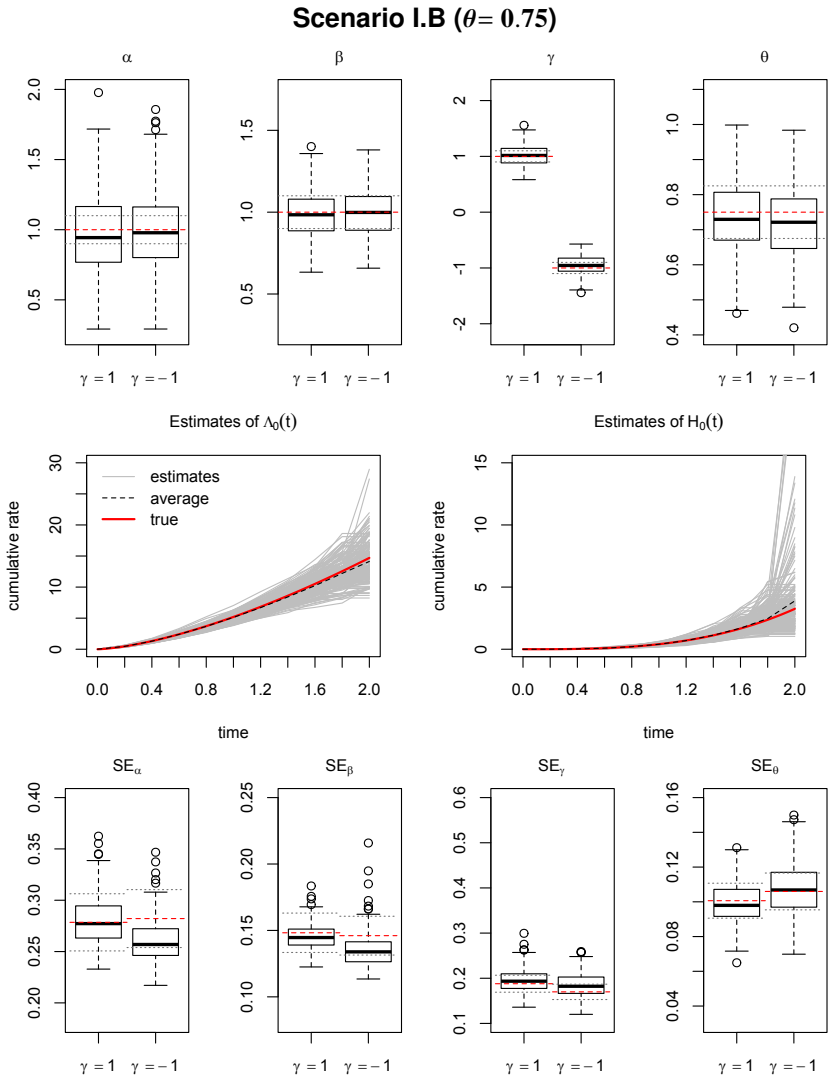

Figure 5.6: Box plots of the parameter estimates (top row) and estimated standard errors (bottom row) in the joint frailty model for positive ($\gamma = 1$) and negative dependence ($\gamma = -1$). Left to right: covariate effect on mortality (α) and on recurrences (β), dependence parameter (γ), and frailty variance (θ), based on 200 samples of size $m = 200$. Red, dashed line marks true value (top row) or empirical standard deviation (bottom row); gray, dotted lines mark 10% deviations from respective value. Middle row: estimates (gray, solid) of the cumulative rate of recurrence (left) and of death (right) in the joint frailty model with $\gamma = 1$.

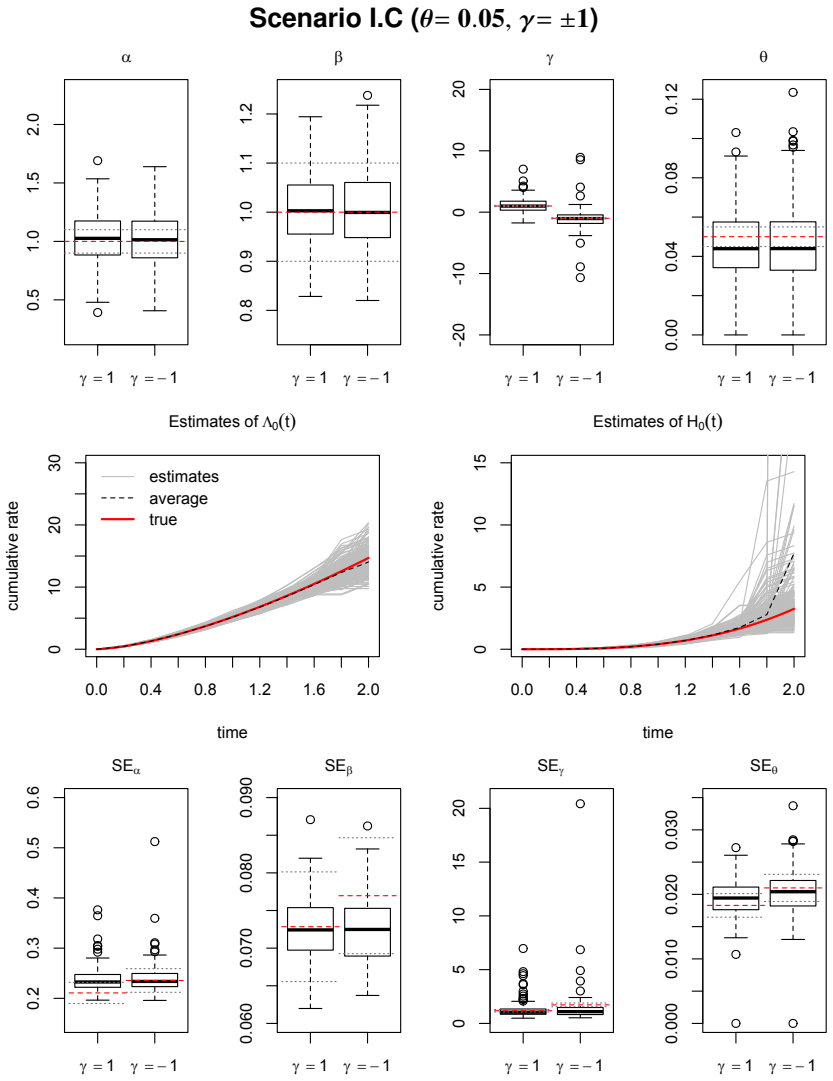

Figure 5.7: Box plots of the parameter estimates (top row) and estimated standard errors (bottom row) in the joint frailty model for positive ($\gamma = 1$) and negative dependence ($\gamma = -1$). Left to right: covariate effect on mortality (α) and on recurrences (β), dependence parameter (γ), and frailty variance (θ), based on 200 samples of size $m = 200$. Red, dashed line marks true value (top row) or empirical standard deviation (bottom row); gray, dotted lines mark 10% deviations from respective value. Middle row: estimates (gray, solid) of the cumulative rate of recurrence (left) and of death (right) in the joint frailty model with $\gamma = 1$.

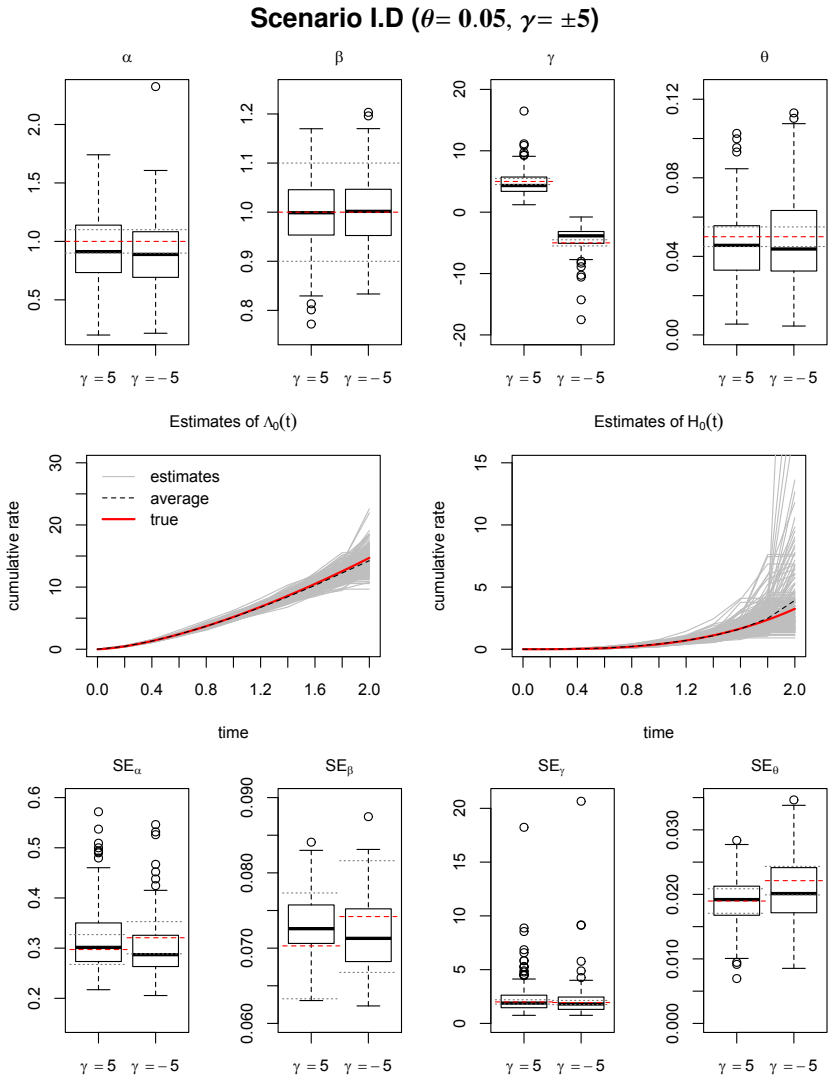

Figure 5.8: Box plots of the parameter estimates (top row) and estimated standard errors (bottom row) in the joint frailty model for positive ($\gamma = 5$) and negative dependence ($\gamma = -5$). Left to right: covariate effect on mortality (α) and on recurrences (β), dependence parameter (γ), and frailty variance (θ), based on 200 samples of size $m = 200$. Red, dashed line marks true value (top row) or empirical standard deviation (bottom row); gray, dotted lines mark 10% deviations from respective value. Middle row: estimates (gray, solid) of the cumulative rate of recurrence (left) and of death (right) in the joint frailty model with $\gamma = 5$.

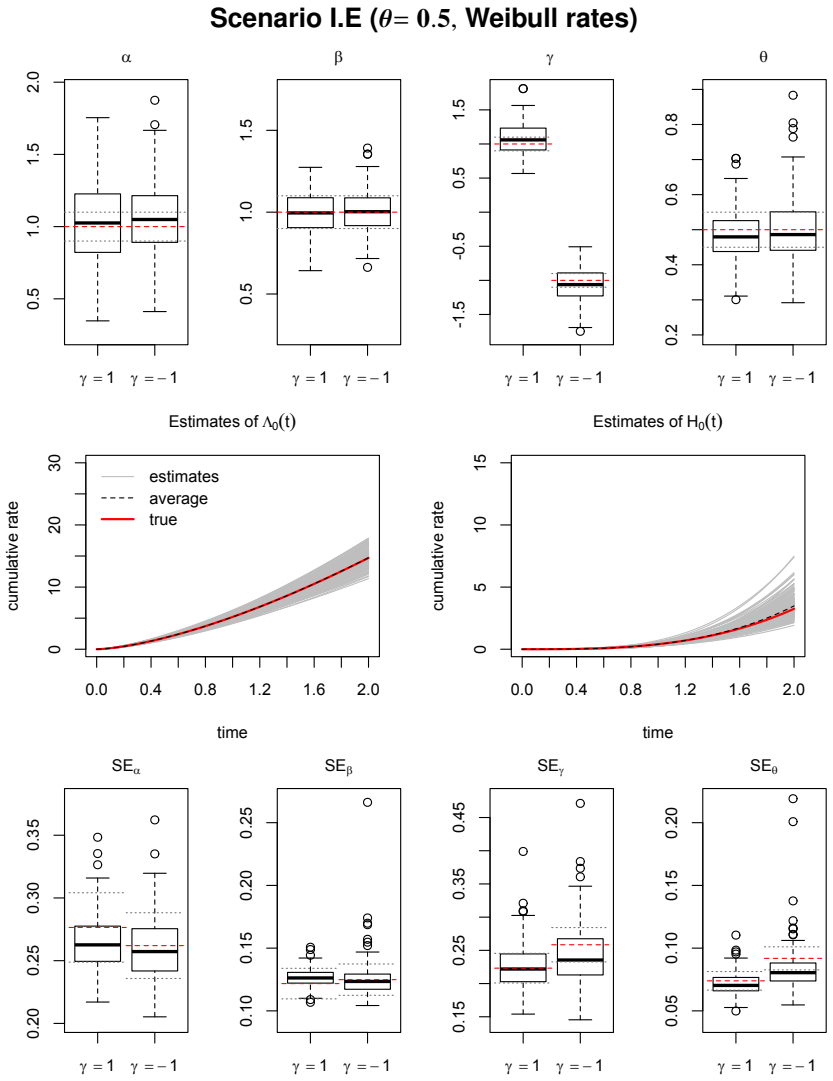

Figure 5.9: Box plots of the parameter estimates (top row) and estimated standard errors (bottom row) in the joint frailty model for positive ($\gamma = 1$) and negative dependence ($\gamma = -1$). Left to right: covariate effect on mortality (α) and on recurrences (β), dependence parameter (γ), and frailty variance (θ), based on 200 samples of size $m = 200$. Red, dashed line marks true value (top row) or empirical standard deviation (bottom row); gray, dotted lines mark 10% deviations from respective value. Middle row: estimates (gray, solid) of the cumulative rate of recurrence (left) and of death (right) in the joint frailty model with $\gamma = 1$.

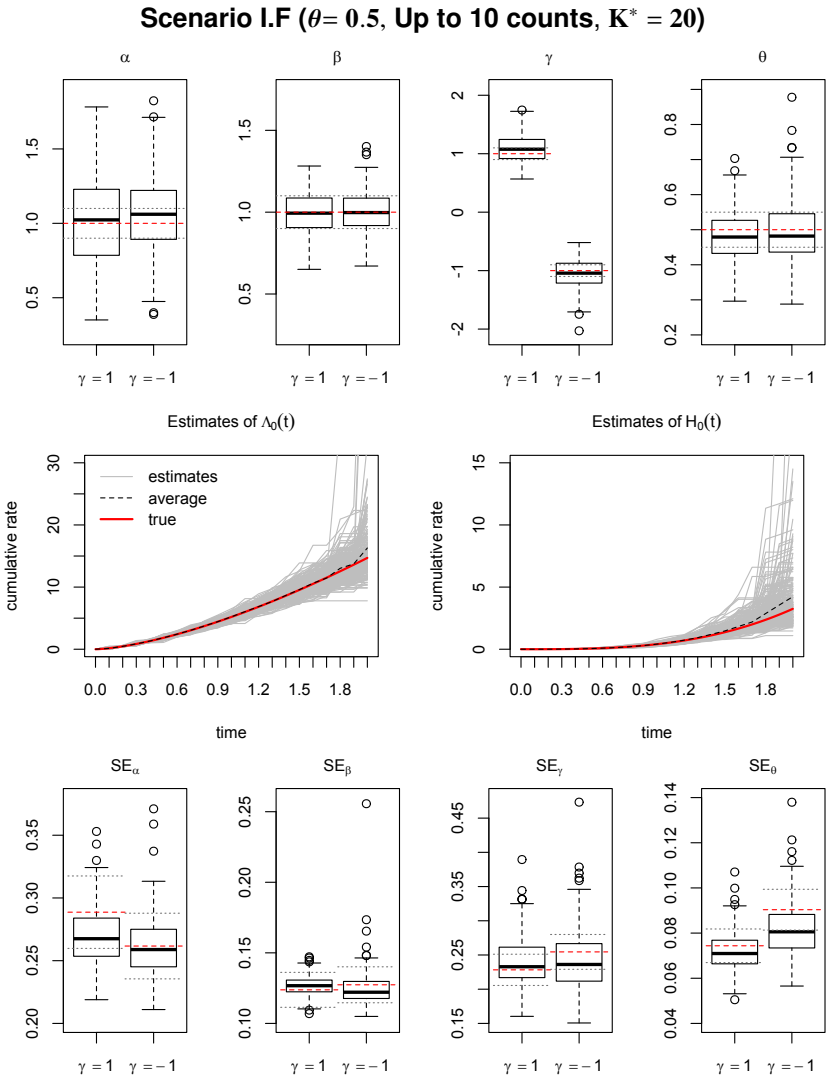

Figure 5.10: Box plots of the parameter estimates (top row) and estimated standard errors (bottom row) in the joint frailty model for positive ($\gamma = 1$) and negative dependence ($\gamma = -1$). Left to right: covariate effect on mortality (α) and on recurrences (β), dependence parameter (γ), and frailty variance (θ), based on 200 samples of size $m = 200$. Red, dashed line marks true value (top row) or empirical standard deviation (bottom row); gray, dotted lines mark 10% deviations from respective value. Middle row: estimates (gray, solid) of the cumulative rate of recurrence (left) and of death (right) in the joint frailty model with $\gamma = 1$.

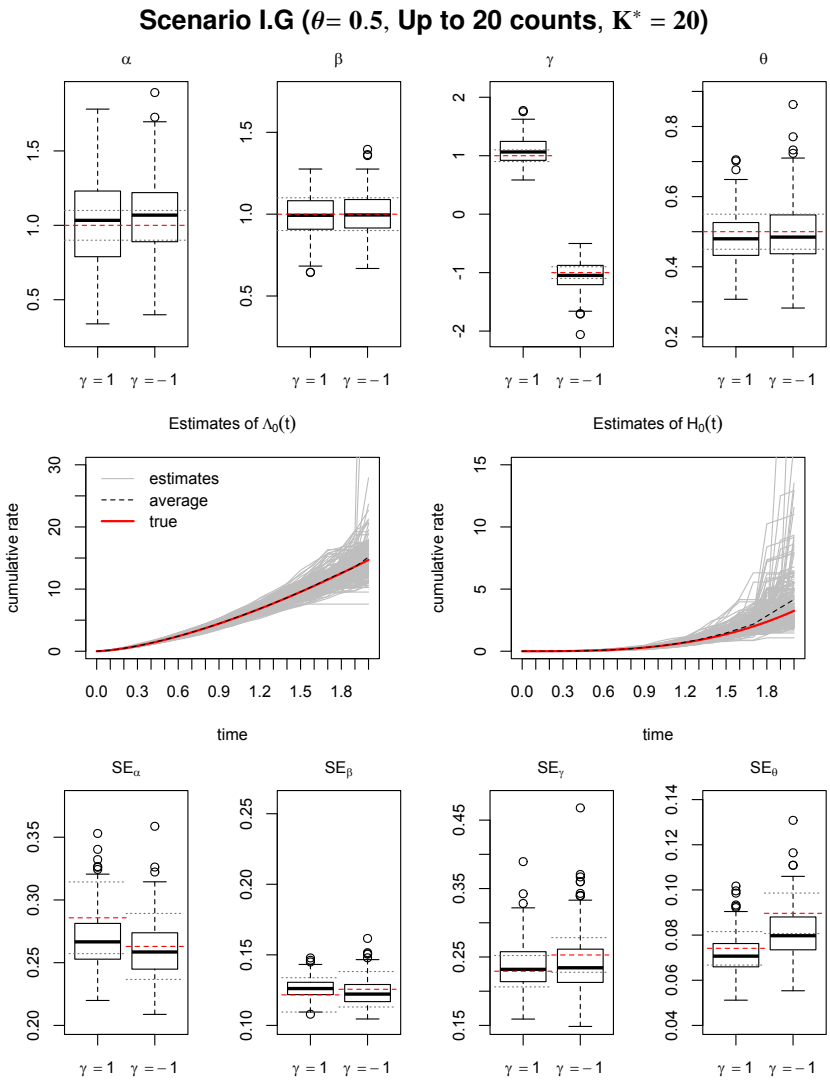

Figure 5.11: Box plots of the parameter estimates (top row) and estimated standard errors (bottom row) in the joint frailty model for positive ($\gamma = 1$) and negative dependence ($\gamma = -1$). Left to right: covariate effect on mortality (α) and on recurrences (β), dependence parameter (γ), and frailty variance (θ), based on 200 samples of size $m = 200$. Red, dashed line marks true value (top row) or empirical standard deviation (bottom row); gray, dotted lines mark 10% deviations from respective value. Middle row: estimates (gray, solid) of the cumulative rate of recurrence (left) and of death (right) in the joint frailty model with $\gamma = 1$.

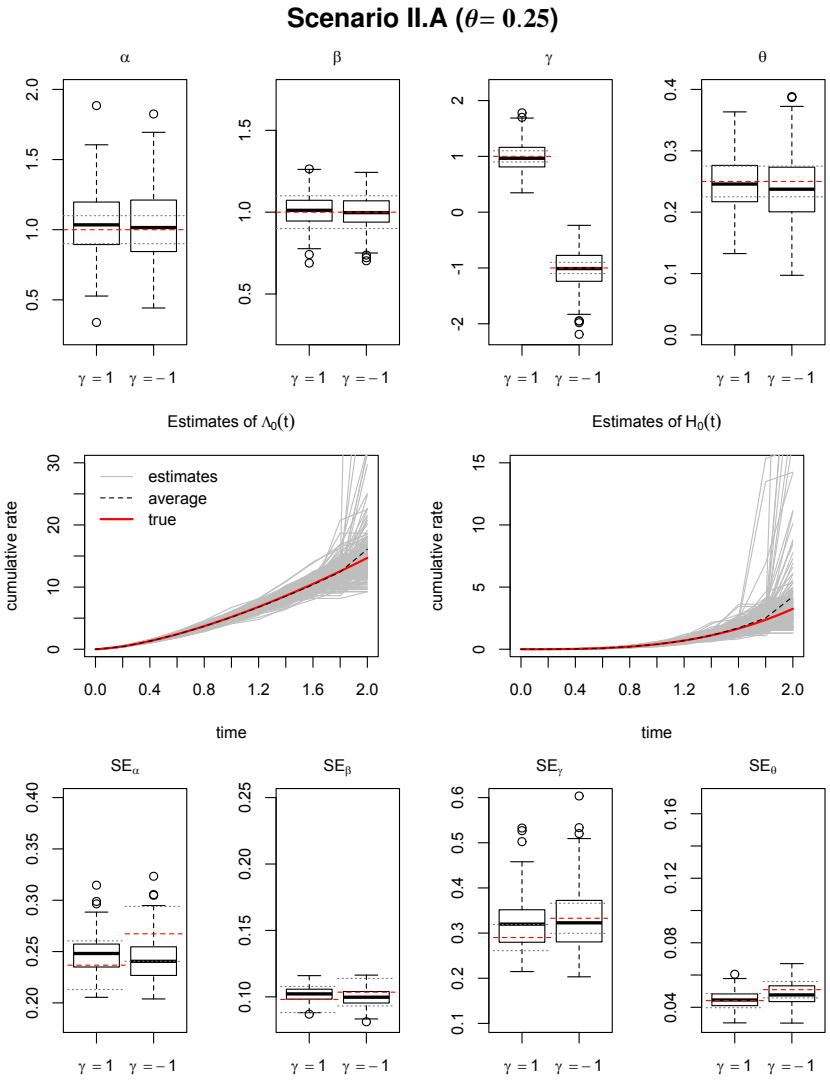

Figure 5.12: Box plots of the parameter estimates (top row) and estimated standard errors (bottom row) in the joint frailty model for positive ($\gamma = 1$) and negative dependence ($\gamma = -1$). Left to right: covariate effect on mortality (α) and on recurrences (β), dependence parameter (γ), and frailty variance (θ), based on 200 samples of size $m = 200$. Red, dashed line marks true value (top row) or empirical standard deviation (bottom row); gray, dotted lines mark 10% deviations from respective value. Middle row: estimates (gray, solid) of the cumulative rate of recurrence (left) and of death (right) in the joint frailty model with $\gamma = 1$.

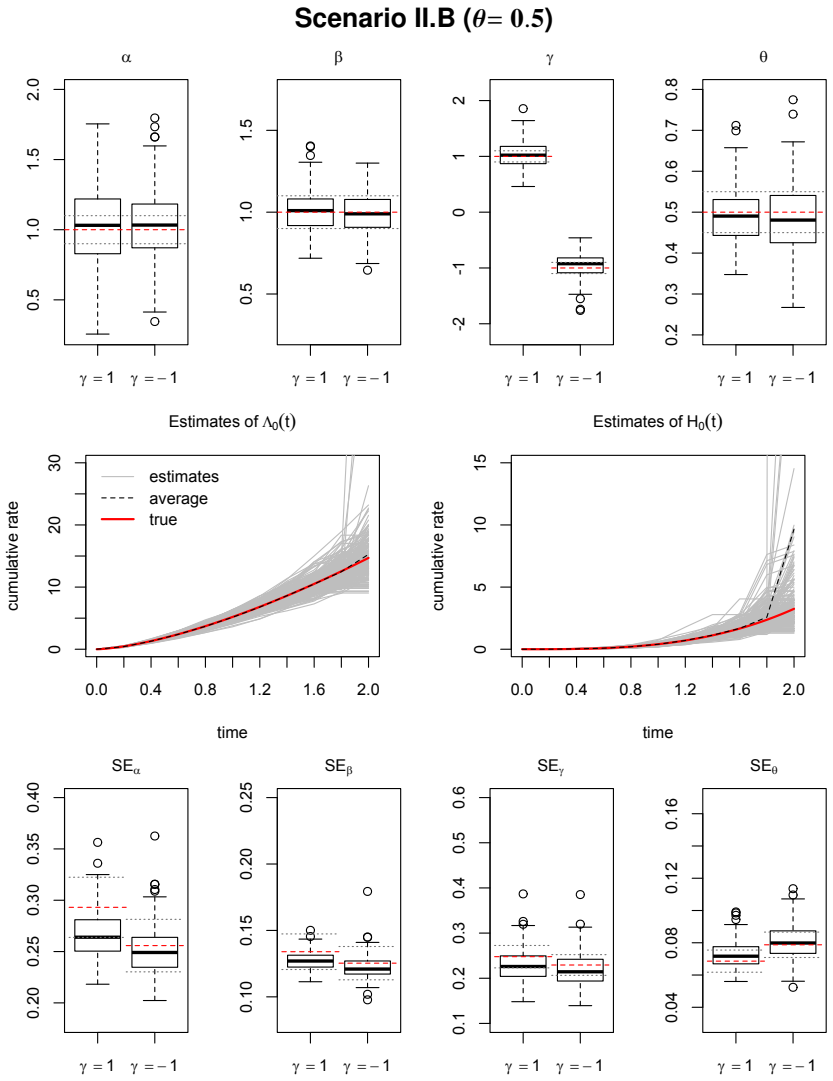

Figure 5.13: Box plots of the parameter estimates (top row) and estimated standard errors (bottom row) in the joint frailty model for positive ($\gamma = 1$) and negative dependence ($\gamma = -1$). Left to right: covariate effect on mortality (α) and on recurrences (β), dependence parameter (γ), and frailty variance (θ), based on 200 samples of size $m = 200$. Red, dashed line marks true value (top row) or empirical standard deviation (bottom row); gray, dotted lines mark 10% deviations from respective value. Middle row: estimates (gray, solid) of the cumulative rate of recurrence (left) and of death (right) in the joint frailty model with $\gamma = 1$.

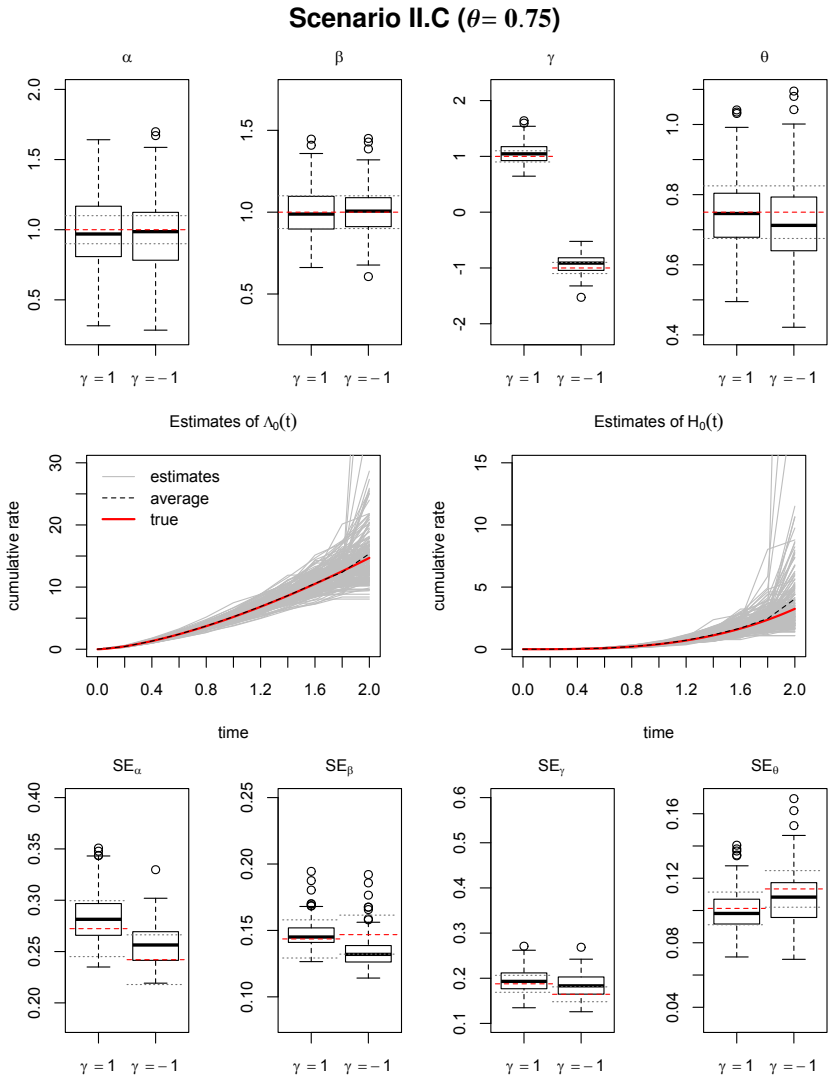

Figure 5.14: Box plots of the parameter estimates (top row) and estimated standard errors (bottom row) in the joint frailty model for positive ($\gamma = 1$) and negative dependence ($\gamma = -1$). Left to right: covariate effect on mortality (α) and on recurrences (β), dependence parameter (γ), and frailty variance (θ), based on 200 samples of size $m = 200$. Red, dashed line marks true value (top row) or empirical standard deviation (bottom row); gray, dotted lines mark 10% deviations from respective value. Middle row: estimates (gray, solid) of the cumulative rate of recurrence (left) and of death (right) in the joint frailty model with $\gamma = 1$.

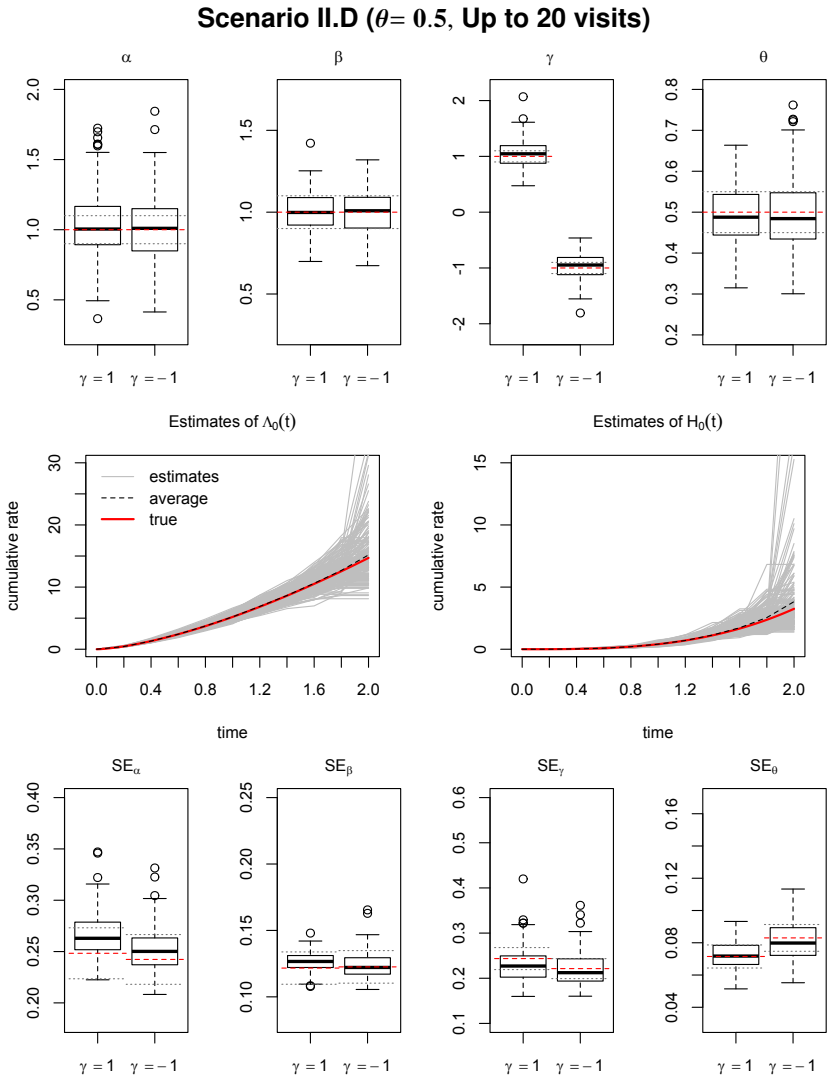

Figure 5.15: Box plots of the parameter estimates (top row) and estimated standard errors (bottom row) in the joint frailty model for positive ($\gamma = 1$) and negative dependence ($\gamma = -1$). Left to right: covariate effect on mortality (α) and on recurrences (β), dependence parameter (γ), and frailty variance (θ), based on 200 samples of size $m = 200$. Red, dashed line marks true value (top row) or empirical standard deviation (bottom row); gray, dotted lines mark 10% deviations from respective value. Middle row: estimates (gray, solid) of the cumulative rate of recurrence (left) and of death (right) in the joint frailty model with $\gamma = 1$.

Table 5.4: Proportion of censoring under different censoring mechanisms for the joint frailty model with frailty variance $\theta = 0.5$ in 200 replications of Scenario I.

		censoring proportion		
γ	censoring	min	mean	max
1	$C \sim \mathcal{U}[0, 2]$	0.445	0.564	0.650
	$C = 2$	0.040	0.091	0.140
-1	$C \sim \mathcal{U}[0, 2]$	0.365	0.484	0.580
	$C = 2$	0.000	0.032	0.065

5.8.2 Performance of the score test

The general design of the simulation study regarding the score test was laid out in Section 5.4.2. In this supplementary material, we present the results for additional scenarios as summarized in Table 5.5. The samples of $m = 200$ individuals each are drawn from a joint frailty model with a binary covariate with effects $\alpha = \beta = 1$ on the hazard of death and the rate of recurrence. Independent censoring is taken to occur uniformly over the total follow-up window $[0, 2]$. We assess the power or size, respectively, of the score test for the different combinations of the dependence parameter γ varying in $\Gamma = \{-1, -0.5, 0, 0.5, 1\}$ and the frailty variance θ varying in $\Theta = \{0.25, 0.5, 0.75\}$.

In Scenario I.A, we extend this to consider stronger dependence, $\gamma = \pm 5$, and a smaller frailty variance, $\theta = 0.05$, inspired by the estimates we obtained for the real data (see Section 5.5).

To study the effect of the width of the observation intervals for the recurrent event counts, the observation times t_{ij} are fixed and the same for all individuals in Scenario I, but varying across individuals in Scenario II. Moreover, the test performance based on finer observation intervals is examined for each of the scenarios (Scenarios I.C and II.B).

The baseline rate of recurrence is specified as a piecewise constant function on $K^R = 10$ or $K^R = 20$ intervals of equal length.

In Section 5.4.1, two versions of independent censoring were considered: uniform censoring over the total follow-up window or end-of-study censoring at $t = 2$. The resulting proportions of censored individuals under the two mechanisms across 1000 replications are given in Table 5.6.

Table 5.5: Simulation settings for the score test with dependence parameter varying in $\Gamma = \{-1, -0.5, 0, 0.5, 1\}$ and frailty variance varying in $\Theta = \{0.25, 0.5, 0.75\}$.

Scenario	γ	θ	obs. times t_{ij}	K^R	Table
Varying the frailty variance in Scenario I					
I.A	$\Gamma \cup \{\pm 5\}$	$\Theta \cup \{0.05\}$	$0.2j$	10	5.7
Varying the specification of the baseline rate in Scenario I					
I.B	Γ	Θ	$0.2j$	20	5.8
Varying the width of the observation intervals in Scenario I					
I.C	Γ	Θ	$0.1j$	20	5.9
Varying the observation times across individuals (Scenario II)					
II.A	Γ	Θ	$0.2j + \varepsilon_{ij}$	10	5.10
Varying the width of the observation intervals in Scenario II					
II.B	Γ	Θ	$0.1j + \varepsilon_{ij}$	10	5.11

The ε_{ij} are drawn from a uniform distribution (see Section 5.4.1).

Table 5.6: Proportion of censoring under different censoring mechanisms for the joint frailty model with dependence parameter $\gamma = 0$ in 1000 replications of Scenario I.

		censoring proportion		
θ	censoring	min	mean	max
0.25	$C \sim \mathcal{U}[0, 2]$	0.365	0.516	0.625
	$C = 2$	0.000	0.019	0.070
0.5	$C \sim \mathcal{U}[0, 2]$	0.400	0.516	0.645
	$C = 2$	0.000	0.020	0.055
0.75	$C \sim \mathcal{U}[0, 2]$	0.405	0.515	0.645
	$C = 2$	0.000	0.019	0.060

Table 5.7: Power and size of the score test, performed at the 5% level, in the joint frailty model with varying dependence parameter $\gamma \in \{-5, -1, -0.5, 0, 0.5, 1, 5\}$ and frailty variance $\theta \in \{0.05, 0.25, 0.5, 0.75\}$, across 1000 replications each.

	Dependence γ						
θ	-5	-1	-0.5	0	0.5	1	5
0.05	0.938	0.290	0.142	0.081	0.116	0.263	0.931
0.25	1.000	0.960	0.586	0.066	0.514	0.932	1.000
0.5	1.000	1.000	0.905	0.062	0.872	1.000	1.000
0.75	1.000	1.000	0.984	0.091	0.974	1.000	1.000

Table 5.8: Power and size of the score test, performed at the 5% level, in the joint frailty model with varying dependence parameter $\gamma \in \{-1, -0.5, 0, 0.5, 1\}$, frailty variance $\theta \in \{0.25, 0.5, 0.75\}$, and $K^R = 20$, across 1000 replications each.

	Dependence γ				
θ	-1	-0.5	0	0.5	1
0.25	0.948	0.525	0.058	0.530	0.929
0.5	0.998	0.878	0.070	0.870	1.000
0.75	1.000	0.974	0.087	0.975	1.000

Table 5.9: Power and size of the score test, performed at the 5% level, in the joint frailty model with varying dependence parameter $\gamma \in \{-1, -0.5, 0, 0.5, 1\}$, frailty variance $\theta \in \{0.25, 0.5, 0.75\}$, up to 20 counts, and $K^R = 20$, across 1000 replications each.

	Dependence γ				
θ	-1	-0.5	0	0.5	1
0.25	0.955	0.556	0.062	0.533	0.940
0.5	0.999	0.890	0.063	0.873	1.000
0.75	1.000	0.979	0.090	0.977	1.000

Table 5.10: Power and size of the score test, performed at the 5% level, in the joint frailty model with varying dependence parameter $\gamma \in \{-1, -0.5, 0, 0.5, 1\}$, frailty variance $\theta \in \{0.25, 0.5, 0.75\}$, and observation times varying across individuals (Scenario II), across 1000 replications each.

	Dependence γ				
θ	-1	-0.5	0	0.5	1
0.25	0.943	0.533	0.071	0.543	0.962
0.5	1.000	0.879	0.075	0.875	0.998
0.75	1.000	0.978	0.065	0.971	1.000

Table 5.11: Power and size of the score test, performed at the 5% level, in the joint frailty model with varying dependence parameter $\gamma \in \{-1, -0.5, 0, 0.5, 1\}$, frailty variance $\theta \in \{0.25, 0.5, 0.75\}$, up to 20 visits, and $K^R = 10$, across 1000 replications each.

	Dependence γ				
θ	-1	-0.5	0	0.5	1
0.25	0.941	0.540	0.079	0.511	0.947
0.5	1.000	0.890	0.074	0.863	1.000
0.75	1.000	0.989	0.069	0.981	1.000

Supporting information

Source code to reproduce the simulation studies and perform the analysis of Section 5.5 is available as Supporting Information at https://onlinelibrary.wiley.com/doi/10.1002/bimj.201900367.

References

Balan, T. A., S. E. Boonk, M. H. Vermeer, and H. Putter (2016). Score test for association between recurrent events and a terminal event. *Statistics in Medicine 35*, 3037–3048.

Cook, R. J. and J. F. Lawless (1997). Marginal analysis of recurrent events and a terminating event. *Statistics in Medicine 16*, 911–924.

Cook, R. J. and J. F. Lawless (2007). *The Statistical Analysis of Recurrent Events*. Statistics for Biology and Health. New York: Springer.

Dańko, A., R. Schaible, and M. J. Dańko (2020). Salinity effects on survival and reproduction of hydrozoan *Eleutheria dichotoma*. *Estuaries and Coasts 43*, 360–374.

Ghosh, D. and D. Y. Lin (2003). Semiparametric analysis of recurrent events data in the presence of dependent censoring. *Biometrics 59*, 877–885.

Gilbert, P. and R. Varadhan (2019). *numDeriv: Accurate Numerical Derivatives*. R package version 2016.8-1.1. https://CRAN.R-project.org/package=numDeriv.

Huang, C.-Y. and M.-C. Wang (2004). Joint modeling and estimation for recurrent event processes and failure time data. *Journal of the American Statistical Association 99*, 1153–1165.

Huang, X., R. A. Wolfe, and C. Hu (2004). A test for informative censoring in clustered survival data. *Statistics in Medicine 23*, 2089–2107.

Lancaster, T. and O. Intrator (1998). Panel data with survival: Hospitalization of HIV-positive patients. *Journal of the American Statistical Association 93*, 46–53.

Lawless, J. F. and M. Zhan (1998). Analysis of interval-grouped recurrent-event data using piecewise constant rate functions. *The Canadian Journal of Statistics / La Revue Canadienne de Statistique 26*, 549–565.

Liu, L. and X. Huang (2008). The use of Gaussian quadrature for estimation in frailty proportional hazards models. *Statistics in Medicine 27*, 2665–2683.

Liu, L., R. A. Wolfe, and X. Huang (2004). Shared frailty models for recurrent events and a terminal event. *Biometrics 60*, 747–756.

Nelson, K. P., S. R. Lipsitz, G. M. Fitzmaurice, J. Ibrahim, M. Parzen, and R. Strawderman (2006). Use of the probability integral transformation to fit nonlinear mixed-effects models with nonnormal random effects. *Journal of Computational and Graphical Statistics 15*, 39–57.

R Core Team (2019). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rondeau, V., S. Mathoulin-Pelissier, H. Jacqmin-Gadda, V. Brouste, and P. Soubeyran (2007). Joint frailty models for recurring events and death using maximum penalized likelihood estimation: application on cancer events. *Biostatistics 8*(4), 708–721.

Sinha, D. and T. Maiti (2004). A Bayesian approach for the analysis of panel-count data with dependent termination. *Biometrics 60*, 34–40.

Smyth, G. K. (1998). Numerical integration. In P. Armitage and T. Colton (Eds.), *Encyclopedia of Biostatistics*, pp. 3088–3095. London: Wiley.

Therneau, T. M. and P. M. Grambsch (2000). *Modeling Survival Data: Extending the Cox Model*. New York: Springer.

Zhao, H., Y. Li, and J. Sun (2013). Analyzing panel count data with a dependent observation process and a terminal event. *The Canadian Journal of Statistics / La Revue Canadienne de Statistique 41*, 174–191.

6

Incorporating delayed entry into the joint frailty model for recurrent events and a terminal event

Abstract

In studies of recurrent events, joint modeling approaches are often needed to allow for potential dependent censoring by a terminal event such as death. Joint frailty models for recurrent events and death with an additional dependence parameter have been studied for cases in which individuals are observed from the start of the event processes. However, the samples are often selected at a later time, which results in delayed entry. Thus, only individuals who have not yet experienced the terminal event will be included in the study. We propose a method for estimating the joint frailty model from such left-truncated data. The frailty distribution among the selected survivors differs from the frailty distribution in the underlying population if the recurrence process and the terminal event are associated. The correctly adjusted marginal likelihood can be expressed as a ratio of two integrals over the frailty distribution, which may be approximated using Gaussian quadrature. The baseline rates are specified as piecewise constant functions, and the covariates are assumed to have multiplicative effects on the event rates. We assess the performance of the

This chapter has been submitted for publication as: M. Böhnstedt, J. Gampe, M.A.A. Caljouw, and H. Putter. Incorporating delayed entry into the joint frailty model for recurrent events and a terminal event.

estimation procedure in a simulation study, and apply the method to estimate age-specific rates of recurrent urinary tract infections and mortality in an older population.

6.1 Introduction

Repeated occurrences of the same type of event in one individual arise in various applications. Examples of such recurrent event data include incidents of myocardial infarction, recurrent infections, fractures, or tumor relapses.

If the individual is additionally at risk of experiencing a terminal event such as death, which will stop the recurrent event process, this might induce dependent censoring of the recurrence process. Therefore, approaches for jointly modeling the two processes of the recurrent events and the terminal event have been developed. Moreover, studies might explicitly address the question of whether there is an association between the processes by asking, for instance, whether individuals who experience more recurrences also have a higher risk of experiencing the terminal event. Specific joint models can provide additional insights into the direction and the strength of the association between the event processes.

The choice of the time scale t, along which the recurrence process and the terminal event process are assumed to evolve, depends on the specific application. In clinical studies, the time since randomization is often used as the time scale; whereas in demographic studies of fertility, for instance, the most relevant time scale is age.

For a given time scale, the event processes can, in some cases, be observed from the time origin t_0. Consider, for example, a medical study in which the time since the disease onset or diagnosis is used as the time scale. If each individual is observed from the disease onset or diagnosis onwards, then all individuals enter the study at time $t_0 = 0$.

However, studies are often initiated at a later point in time when the two processes have already started. If in the above setting patients do not participate in the clinical study until some period of time after their diagnosis, individual times of study entry will differ from 0, and might also vary between patients. Similarly, in studies of certain diseases in old-age populations, such as in register-based studies of cardiovascular disease and mortality, age can be considered the natural time scale (see, for instance, Modig et al., 2013). The individuals included in such analyses have already reached a certain advanced age by the beginning of the study period.

In these cases, the data are left-truncated in the sense that the individuals can enter the study only if they have not yet experienced the terminal event. If the recurrence process and the terminal event process are associated, this sample of survivors is not a random sample of the underlying population. Rather, the sample is comprised of individuals who tend to have a lower risk of experiencing the terminal event, and – if there is a positive association between the two processes – to also have a lower risk of experiencing recurrent events. Thus, to obtain valid inferences, correctly adjusting for the truncation is crucial.

The example we use in this study to illustrate our method focuses on recurrent urinary tract infections (UTIs) in older residents of long-term care facilities (LTCF). The original

study by Caljouw et al. (2014) investigated whether cranberry capsules are effective in preventing UTIs. Given that around 34% of the elderly study population died during the follow-up period, recurrent UTIs and mortality had to be modeled jointly. The original study used time since randomization as the time scale. By contrast, we have chosen to use age as the main time scale of the event processes, as mortality and, presumably, UTI recurrences naturally depend on age. Because the participants were between 64 and 102 years old when they entered the study, the observations are then left-truncated. In our analysis, we seek to assess the effects of the cranberry treatment, to estimate the age-specific rates of UTIs and death, and to determine whether there is an association between UTIs and death risks.

While different joint models for recurrent events and death have been proposed, we focus here on the joint frailty model introduced by Liu et al. (2004). This model has been applied repeatedly in medical studies (e.g., to study recurrent cancer events as in Rondeau et al., 2007, or recurrent heart failure hospitalizations as in Rogers et al., 2016), and extended in several directions (e.g., to the setting of nested case-control studies in Jazić et al., 2019). The joint frailty model allows us to examine the shape of the rates of recurrence and death, as well as the potential dependence between the two processes. The dependence is introduced by a shared individual random effect entering both the rate of recurrence and the hazard of death. An additional parameter determines whether the processes are positively or negatively associated, and how strong this association is. In other models, the frailty affects both event rates in the same way (Huang and Wang, 2004), or the dependence between the processes is left completely unspecified (Cook and Lawless, 1997; Ghosh and Lin, 2003). A common approach to estimating the joint frailty model that we consider here is using Gaussian quadrature to approximate the marginal likelihood (Liu and Huang, 2008; Rondeau et al., 2007). We will see that in this framework, the setting with left-truncated data can be handled in a straightforward manner by adapting the likelihood.

Incorporating left truncation, which is also called delayed entry, has received varying levels of attention in studies based on different frailty models and joint models. Several studies have discussed handling left truncation in shared frailty models for clustered survival data (e.g., Jensen et al., 2004; van den Berg and Drepper, 2016). In a recurrent event setting, Balan et al. (2016) considered event dependent selection; i.e., individuals were included in the study only if they had experienced at least one recurrent event in a given time period. Recurrent event studies with selection dependent on survival were briefly discussed in Cook and Lawless (2007), but not specifically in the context of frailty models. It has been argued that for the joint frailty model for recurrent events and death, as specified by Liu et al. (2004), delayed entry can be easily incorporated (Rondeau et al., 2007). However, to our knowledge, no detailed account of this approach has previously been provided. At the time of writing, the `R` package `frailtypack`, which can be used for fitting a variety of frailty models, does not provide functionality for estimating the joint frailty model from left-truncated data (according to the manual of version 3.3.2, date 2020-10-07, Rondeau et al., 2020). Extensions to left-truncated data have been considered

in some other joint models for recurrent events and death. Emura et al. (2017) introduced a joint frailty-copula model for two event times that can be adapted to accommodate left truncation and recurrent event data. Outside of the class of shared frailty models, Cai et al. (2017) proposed a model for longitudinal measurements, recurrent events, and a terminal event with inferences based on estimation equations that can be generalized to allow for left-truncated data. Liu et al. (2012) presented an estimating equation procedure for a partial marginal model of recurrent events in the presence of a terminal event with left truncation. In the context of joint models for longitudinal data and death, estimation procedures based on left-truncated data have been derived for different models with shared random effects (see, for example, van den Hout and Muniz-Terrera, 2016; Crowther et al., 2016; Piulachs et al., 2021).

In this chapter, we propose a method for estimating the joint frailty model for recurrent events and a terminal event, as introduced by Liu et al. (2004), when data are left-truncated. Our approach adapts the method of Liu and Huang (2008), who use Gaussian quadrature to approximate the marginal likelihood of the joint frailty model.

The chapter is structured as follows. Section 6.2 first presents the joint frailty model and the corresponding likelihood in the usual setting without truncation, and then shows the adjustments for left-truncated data. The method of estimation is detailed in Section 6.3, and its performance is assessed via simulation studies in Section 6.4. We illustrate the approach using the data set on recurrent UTIs in Section 6.5, and conclude with a discussion in Section 6.6.

6.2 Joint frailty model and left truncation

The joint frailty model for recurrent events and a terminal event has been applied most frequently in situations in which the time of the terminal event is subject to independent right-censoring only. In the following, we will first present the model and the corresponding likelihood for such right-censored data. Then, we will lay out how certain assumptions and, in particular, the likelihood function are adjusted to the case of left-truncated data. Throughout, we will often refer to the terminal event as death for the sake of simplicity.

6.2.1 Joint frailty model

We consider independent individuals i, $i = 1, ..., m$, who can experience recurrent events between time $t_0 = 0$ and the time D_i of the terminal event. Let C_i denote a censoring time, which is assumed to be independent of the recurrence and terminal event processes. An individual can then be observed only up to his or her follow-up time $X_i = \min(C_i, D_i)$, and $\delta_i = \mathbb{1}\{D_i \leq C_i\}$ will indicate whether the terminal event occurred before censoring, with the indicator function $\mathbb{1}\{\cdot\}$. The at-risk indicator at time $t \geq 0$ is given by $Y_i(t) = \mathbb{1}\{t \leq X_i\}$, if individuals enter the study at $t_0 = 0$.

The number of recurrent events experienced by individual i up to time t is recorded by the actual recurrent event process $N_i^{R*}(t)$. Similarly, we define the actual counting process

of the terminal event as $N_i^{D*}(t) = \mathbb{1}\{D_i \leq t\}$. However, due to right-censoring, we can only observe the processes $N_i^D(t) = \mathbb{1}\{X_i \leq t, \delta_i = 1\}$ and $N_i^R(t) = \int_0^t Y_i(s)\mathrm{d}N_i^{R*}(s)$ $= N_i^{R*}(\min(t, X_i))$. Here, $\mathrm{d}N_i^{R*}(t) = N_i^{R*}((t + \mathrm{d}t)^-) - N_i^{R*}(t^-)$ is the increment of the recurrence process, equal to the number of events in the small interval $[t, t + \mathrm{d}t)$, with t^- as the left-hand limit.

The observed data of individual i up to time t are given by $\mathcal{O}_i(t) = \{Y_i(s), N_i^R(s), N_i^D(s), 0 \leq s \leq t; \boldsymbol{z}_i\}$, including the observed time-fixed covariate vector $\boldsymbol{z}_i$. Individual risks will depend on the covariates as well as on the unobservable frailty value u_i, where the frailties u_i are independent realizations of a positive random variable U.

In the joint frailty model introduced by Liu et al. (2004), the observed recurrence process is assumed to have the intensity $Y_i(t)\lambda_i(t|u_i)$ with

$$\begin{aligned} \mathrm{P}(\mathrm{d}N_i^R(t) = 1 \mid \mathcal{F}_{t^-}, D_i \geq t) &= Y_i(t)\lambda_i(t|u_i)\mathrm{d}t \\ \lambda_i(t|u_i)\mathrm{d}t = \mathrm{d}\Lambda_i(t|u_i) &= \mathrm{P}(\mathrm{d}N_i^{R*}(t) = 1 \mid \boldsymbol{z}_i, u_i, D_i \geq t). \end{aligned} \tag{6.1}$$

Here, $\mathcal{F}_t = \sigma\{\mathcal{O}_i(s), 0 \leq s \leq t, u_i; i = 1, ..., m\}$ denotes the σ-algebra generated by the frailty and the observed data. The terminal event process is characterized by the intensity $Y_i(t)h_i(t|u_i)$ with

$$\begin{aligned} \mathrm{P}(\mathrm{d}N_i^D(t) = 1 \mid \mathcal{F}_{t^-}) &= Y_i(t)h_i(t|u_i)\mathrm{d}t \\ h_i(t|u_i)\mathrm{d}t = \mathrm{d}H_i(t|u_i) &= P(\mathrm{d}N_i^{D*}(t) = 1 \mid \boldsymbol{z}_i, u_i, D_i \geq t). \end{aligned} \tag{6.2}$$

The first lines in (6.1) and (6.2) follow from the assumption that the censoring mechanism is conditionally independent of the two event processes given the process history.

Following Liu et al. (2004), we specify the intensities as

$$\begin{aligned} \lambda_i(t|u_i) &= u_i\, e^{\boldsymbol{\beta}^\top \boldsymbol{z}_i}\, \lambda_0(t), \\ h_i(t|u_i) &= u_i^\gamma\, e^{\boldsymbol{\alpha}^\top \boldsymbol{z}_i}\, h_0(t). \end{aligned} \tag{6.3}$$

The baseline rates of recurrence and death, $\lambda_0(t)$ and $h_0(t)$, are affected by the covariates $\boldsymbol{z}_i$ through a multiplicative model with effects $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$, respectively. The inclusion of the frailty u in the recurrence rate accommodates heterogeneity across individuals and dependence between the recurrences within one individual. The association between the recurrent events and death results from the fact that the shared frailty u also enters the hazard of death. Due to the additional parameter γ, the model can capture associations of variable magnitudes and in different directions. For positive $\gamma > 0$, individuals with a higher rate of recurrence will also be subject to a higher hazard of death. For $\gamma < 0$, a higher rate of recurrence implies a lower hazard of death. If $\gamma = 0$, the intensities in (6.3) do not share any parameters, and the censoring of the recurrence process by death is non-informative.

A common choice for the distribution of the frailty U is a gamma distribution with a mean of one and a variance of θ. We will more generally consider that the frailties u_i follow a distribution with a density of $g_\theta(u)$ and a corresponding distribution function $G_\theta(u)$

with parameter θ. This assumption refers to the initial distribution of frailties in the population at time $t_0 = 0$. However, if $\gamma \neq 0$, the distribution of frailties in the population will change over time due to selection effects, which will cause the population at time t to be composed of survivors with lower mortality risks.

We now formulate the likelihood of the joint frailty model (6.3) when individuals are observed from time $t_0 = 0$. Let t_{ij}, $j = 1, ..., J_i$, be the observed recurrence times of individual i. Based on the arguments stated in Liu et al. (2004), the conditional likelihood contribution of individual i given his or her frailty value u_i can be written as

$$\begin{aligned} L_i^{(c)}(u_i) &= \left[\prod_{j=1}^{J_i} \lambda_i(t_{ij}|u_i)\right] \exp\left\{-\int_0^\infty Y_i(s)\lambda_i(s|u_i)\,\mathrm{d}s\right\} \\ &\qquad \cdot h_i(x_i|u_i)^{\delta_i} \exp\left\{-\int_0^\infty Y_i(s)h_i(s|u_i)\,\mathrm{d}s\right\} \\ &= \left[\prod_{j=1}^{J_i} u_i e^{\boldsymbol{\beta}^\top \mathbf{z}_i}\lambda_0(t_{ij})\right] \exp\left\{-\int_0^{x_i} u_i e^{\boldsymbol{\beta}^\top \mathbf{z}_i}\lambda_0(s)\,\mathrm{d}s\right\} \\ &\qquad \cdot \left[u_i^\gamma e^{\boldsymbol{\alpha}^\top \mathbf{z}_i} h_0(x_i)\right]^{\delta_i} \exp\left\{-\int_0^{x_i} u_i^\gamma e^{\boldsymbol{\alpha}^\top \mathbf{z}_i} h_0(s)\,\mathrm{d}s\right\}. \end{aligned} \tag{6.4}$$

The marginal likelihood L_i of the observed data of individual i is obtained by integrating the above expression over the frailty distribution,

$$L_i = \int_0^\infty L_i^{(c)}(u)\,\mathrm{d}G_\theta(u) = \int_0^\infty L_i^{(c)}(u)\,g_\theta(u)\,\mathrm{d}u. \tag{6.5}$$

6.2.2 Adjusting for left truncation

We will now extend the above framework to allow for left truncation, that is, individuals entering the study at times that may be later than $t_0 = 0$. Before deriving the likelihood for the left-truncated data, we introduce some additional notations and assumptions.

A sample of m_V independent individuals i, $i = 1, ..., m_V$, is left-truncated if the individuals i enter the study only at times $V_i \geq t_0$, with strict inequality for at least some individuals. Then, the observation of individual i is conditional on his or her survival up to the entry time, $D_i > V_i$, and events can only be observed in the interval $[V_i, X_i]$. Hence, the at-risk indicator $Y_i(t)$ of Section 6.2.1 is replaced by ${}_VY_i(t) = \mathbb{1}\{V_i \leq t \leq X_i\}$.

As a consequence, the observed recurrent event process ${}_VN_i^R(t) = \int_0^t {}_VY_i(s)\mathrm{d}N_i^{R*}(s)$ $= [N_i^{R*}(\min(t, X_i)) - N_i^{R*}(V_i)]\mathbb{1}\{t > V_i\}$ in this setting records only the recurrences after study entry at V_i. Analogously, the left-truncated counting process of the terminal event is given by ${}_VN_i^D(t) = \mathbb{1}\{V_i \leq X_i \leq t, \delta_i = 1\}$. The observed data for individual i are then ${}_V\mathcal{O}_i(t) = \{{}_VY_i(s), {}_VN_i^R(s), {}_VN_i^D(s), V_i \leq s \leq t; \mathbf{z}_i; V_i\}$, and the σ-algebra is modified as ${}_V\mathcal{F}_t = \sigma\{{}_V\mathcal{O}_i(s), 0 \leq s \leq t, u_i; i = 1, ..., m_V\}$.

In addition to the assumption of conditionally independent censoring already made in Section 6.2.1, we further assume that the truncation times V_i are conditionally independent of the recurrence and terminal event processes given the process history. Hence, the intensity of the observed recurrence process is given by ${}_VY_i(t)\lambda_i(t|u_i)$ and (6.1) is adapted as

$$\mathrm{P}(\mathrm{d}\,{}_VN_i^R(t) = 1 \mid {}_V\mathcal{F}_{t-}, D_i \geq t) = {}_VY_i(t)\lambda_i(t|u_i). \tag{6.1'}$$

The intensity of the observed terminal event process is, correspondingly, ${}_VY_i(t)h_i(t|u_i)$, such that (6.2) is modified as

$$\mathrm{P}(\mathrm{d}\,{}_VN_i^D(t) = 1 \mid {}_V\mathcal{F}_{t-}) = {}_VY_i(t)h_i(t|u_i). \tag{6.2'}$$

Based on this, we can develop the likelihood of the joint frailty model (6.3) for left-truncated data. The conditional likelihood contribution of individual i given u_i is constructed in analogy to (6.4), with $Y_i(s)$ replaced by ${}_VY_i(s)$, and appropriately restricting to the observed recurrence times $t_{ij} \geq v_i$; that is,

$$\begin{aligned}
{}_VL_i^{(c)}(u_i) &= \left[\prod_{t_{ij}\geq v_i} \lambda_i(t_{ij}|u_i)\right] \exp\left\{-\int_0^\infty {}_VY_i(s)\lambda_i(s|u_i)\,\mathrm{d}s\right\} \\
&\qquad \cdot h_i(x_i|u_i)^{\delta_i} \exp\left\{-\int_0^\infty {}_VY_i(s)h_i(s|u_i)\,\mathrm{d}s\right\} \\
&= \left[\prod_{t_{ij}\geq v_i} u_i e^{\boldsymbol{\beta}^\top \mathbf{z}_i}\lambda_0(t_{ij})\right] \exp\left\{-\int_{v_i}^{x_i} u_i e^{\boldsymbol{\beta}^\top \mathbf{z}_i}\lambda_0(s)\,\mathrm{d}s\right\} \\
&\qquad \cdot \left[u_i^\gamma e^{\boldsymbol{\alpha}^\top \mathbf{z}_i} h_0(x_i)\right]^{\delta_i} \exp\left\{-\int_{v_i}^{x_i} u_i^\gamma e^{\boldsymbol{\alpha}^\top \mathbf{z}_i} h_0(s)\,\mathrm{d}s\right\}.
\end{aligned} \tag{6.6}$$

The marginal likelihood contribution is again obtained by integrating out the frailty. However, as the frailty distribution in the sample of survivors differs from the frailty distribution at time t_0, we need to integrate over the conditional frailty distribution given survival to the time of entry into the study. This point has previously been discussed in the context of clustered survival data by van den Berg and Drepper (2016) and Eriksson et al. (2015) and for general state duration models by Lawless and Fong (1999). More formally, the marginal likelihood contribution of individual i is thus

$${}_VL_i = \int_0^\infty {}_VL_i^{(c)}(u)\,\mathrm{d}G_\theta(u \mid D_i > v_i, V_i = v_i, \mathbf{z}_i) = \int_0^\infty {}_VL_i^{(c)}(u)\,\mathrm{d}G_\theta(u \mid D_i > v_i, \mathbf{z}_i), \tag{6.7}$$

under the assumption that the truncation time V_i is independent of u.

In particular, if $\gamma > 0$ such that the recurrence process and the mortality process are positively associated, individuals who survived up to time v will tend to have lower frailty values than individuals who died before time v, for given $\mathbf{z}_i$. Hence, the frailty

distribution among survivors beyond time v, $G_\theta(u \mid D > v)$, will tend to have more probability mass at lower values u than the frailty distribution $G_\theta(u)$ in the underlying population at time t_0. Consequently, neglecting the effect of the survivor selection on the frailty distribution in the sample and constructing a marginal likelihood as

$$ {}_vL_i^{\text{naive}} = \int_0^\infty {}_vL_i^{(c)}(u)\,\mathrm{d}G_\theta(u) = \int_0^\infty {}_vL_i^{(c)}(u)\,g_\theta(u)\,\mathrm{d}u, \tag{6.8} $$

would lead to an invalid inference if $\gamma \neq 0$. We will illustrate the resulting biases in the parameter estimates in a simulation study in Section 6.4.

For computing the correct marginal likelihood (6.7), we first apply Bayes' theorem to find that

$$ g_\theta(u|D_i > v_i, \boldsymbol{z}_i) = \frac{\mathrm{P}(D_i > v_i \mid u, \boldsymbol{z}_i)\,g_\theta(u)}{\mathrm{P}(D_i > v_i \mid \boldsymbol{z}_i)} = \frac{\exp\left\{-\int_0^{v_i} h_i(s|u)\,\mathrm{d}s\right\} g_\theta(u)}{\int_0^\infty \mathrm{P}(D_i > v_i \mid u)\,g_\theta(u)\,\mathrm{d}u}, \tag{6.9} $$

where we suppress the dependence on the covariates $\boldsymbol{z}_i$ in the last expression for notational convenience. Combining equations (6.6), (6.7), and (6.9), we can express the marginal likelihood contribution ${}_vL_i$ of individual i with left-truncated data as

$$ \frac{\int_0^\infty \left[\prod_{t_{ij}\geq v_i} \lambda_i(t_{ij}|u)\right] \exp\left\{-\int_{v_i}^{x_i} \lambda_i(s|u)\,\mathrm{d}s\right\} h_i(x_i|u)^{\delta_i} \exp\left\{-\int_0^{x_i} h_i(s|u)\,\mathrm{d}s\right\} g_\theta(u)\,\mathrm{d}u}{\int_0^\infty \mathrm{P}(D_i > v_i \mid u)\,g_\theta(u)\,\mathrm{d}u}. \tag{6.10} $$

Interestingly, the formula for ${}_vL_i$ in (6.10) could have been equivalently derived as the marginal (with respect to the frailty) probability of the recurrence and follow-up data on individual i, conditional on individual i being included in the study, $D_i > v_i$. To see this, let us denote by E_i the event that individual i has follow-up time x_i with indicator δ_i and the observed recurrence times t_{ij} over $[v_i, x_i]$, and consider

$$ \mathrm{P}(E_i \mid D_i > v_i) = \frac{\mathrm{P}(E_i \cap \{D_i > v_i\})}{\mathrm{P}(D_i > v_i)} = \frac{\mathrm{P}(E_i)}{\mathrm{P}(D_i > v_i)} = \frac{\int_0^\infty \mathrm{P}(E_i \mid u)\,g_\theta(u)\,\mathrm{d}u}{\int_0^\infty \mathrm{P}(D_i > v_i \mid u)\,g_\theta(u)\,\mathrm{d}u}. $$

As the integrals over the frailty distribution in (6.10) will not, in general, have closed-form expressions, we will use numerical integration in the following.

6.3 Estimation of the joint frailty model

Liu and Huang (2008) proposed using Gaussian quadrature to approximate the marginal likelihood of frailty proportional hazards models, including the joint frailty model (6.3). In combination with a piecewise constant specification of the baseline rates, this approach allows for the direct maximization of the approximated likelihood.

The aim of Gaussian quadrature is to replace the integral of a function with a weighted sum of function values. The Gauss-Hermite quadrature rule gives an approximation for a specific integral of a function $f(x)$,

$$\int_{-\infty}^{\infty} f(x)e^{-x^2}\mathrm{d}x \approx \sum_{q=1}^{Q} w_q f(x_q).$$

The quadrature points x_q need to be determined as the roots of the Q^{th}-order Hermite polynomial, while w_q specify corresponding weights. As the marginal likelihood of a model with normal random effects is easily rewritten in the above form, it follows that such a likelihood can be approximated as

$$\int_{-\infty}^{\infty} L^{(c)}(b)\,\phi(b)\,\mathrm{d}b \approx \sum_{q=1}^{Q} \tilde{w}_q\, L^{(c)}(\tilde{x}_q)\,\phi(\tilde{x}_q), \tag{6.11}$$

where $\phi(\cdot)$ denotes the standard normal density, while $\tilde{x}_q = \sqrt{2}x_q$ and $\tilde{w}_q = \sqrt{2}w_q e^{x_q^2}$ are modified quadrature points and weights, respectively. Marginal likelihoods integrated over non-normal random effects can be expressed as an integral of the form in (6.11) by applying the probability integral transformation (see Nelson et al., 2006; Liu and Huang, 2008). We show in Section 6.7.1 how this quadrature approach can be used to derive an approximation of the marginal likelihood of the joint frailty model with left truncation.

The evaluation of the approximate marginal likelihood depends on the specific form of the baseline rates $\lambda_0(t)$ and $h_0(t)$. As suggested by Liu and Huang (2008), we adopt a piecewise constant model for these functions,

$$\lambda_0(t) = \sum_{k=1}^{K^R} \lambda_{0k}\,\mathbb{1}\{t \in I_k^R\} \quad \text{and} \quad h_0(t) = \sum_{k=1}^{K^D} h_{0k}\,\mathbb{1}\{t \in I_k^D\}.$$

with intervals $I_k^R = (t_{k-1}^R, t_k^R]$, $k = 1, ..., K^R$, and $I_k^D = (t_{k-1}^D, t_k^D]$, $k = 1, ..., K^D$. Specifications with a moderate number of up to 10 intervals and the cut-points t_k, $k \geq 1$, which have been determined based on the quantiles of the observed event times, are generally expected to lead to good results in practice (see Cook and Lawless, 2007; Liu and Huang, 2008). In the setting with left truncation, appropriate choices have to be made for the starting points of the first intervals, t_0^R and t_0^D. Depending on the study design, they might be set equal to the lowest study entry time, $\min_i v_i$, or a lower value $t^* \geq t_0$.

The direct maximization of the marginal likelihood would also be possible if the baseline rates were assumed to follow a simple parametric model, such as the Weibull model. Nonetheless, we recommend the use of the more flexible piecewise constant rate models, unless prior knowledge allows for an informed choice of a specific parametric model.

Finally, the parameter estimates in the joint frailty model with left-truncated data are obtained by maximizing the approximate marginal log-likelihood. The calculation of the

standard errors is based on the inverse of the negative Hessian matrix of the approximate marginal log-likelihood. We give additional details on the implementation in Section 6.7.2.

6.4 Simulation study

To evaluate the performance of the proposed method for estimating the parameters of the joint frailty model in case of left-truncated data, we conducted a simulation study. We will also demonstrate which biases can arise if the likelihood is not correctly adjusted to the survivor selection, in particular, to the selection effects on the frailty distribution.

Estimator performance will depend on various aspects of the observation scheme. One aspect is the distribution of the study entry times V_i, in which both the range of the distribution and its shape matter. Furthermore, the censoring mechanism – that is, the length of the individual follow-up periods and the number of additional drop-outs – will influence the performance of the method. To study these issues, we will first present a base scenario, and will then assess how different observational settings affect the results.

6.4.1 Settings

In the base scenario, we generated data from a joint frailty model (6.3). The time scale t is the age of the individual. The hazard of death and the rate of recurrence are each affected by a single binary covariate, which is drawn from a Bernoulli distribution with parameter 0.5. The regression coefficients are $\alpha = 0.5$ (death) and $\beta = 0.5$ (recurrence), respectively. The frailty values are realizations of a gamma distribution with a mean of one and a variance of $\theta = 0.5$. The values of the dependence parameter γ were chosen to cover a positive ($\gamma = 0.5$) and a negative ($\gamma = -0.5$) association between the recurrence process and death, as well as the case in which the recurrence rate does not affect the mortality risk ($\gamma = 0$).

The baseline rates $h_0(t)$ and $\lambda_0(t)$ were designed to mimic a study in an older population among whom the death rates as well as the recurrent event rates increase exponentially with age. Hence, we chose for both baseline rates the Gompertz-Makeham form, $ae^{bt} + c$, where $t = 0$ corresponds to age 75. By setting $a = 0.984$, $b = 0.045$, and $c = 0$ for the recurrence process ($\lambda_0(t)$), and $a = 0.108$, $b = 0.07$, and $c = 0.12$ for the survival process ($h_0(t)$), the baseline rates were comparable to the estimated rates for the high risk group in the data example in Section 6.5.

To arrive at the left-truncated samples, the following steps were combined. For each individual, a survival time D (i.e., age > 75) and an entry time into study V were simulated. Only those individuals who survived beyond his or her entry time – that is, for whom $D > V$ – were included in the final sample (i.e., were 'observed'). Therefore, the distribution of entry times V that are observed in the final sample depends on both the mortality model in (6.3) and the initial distribution of the truncation times before selection.

In the base scenario, our aim was to have entry times in the final sample that were distributed across the total age range – here, ages 75 to 95 – with higher numbers of study entries at the younger ages than at the older ages. This scheme will be referred to as truncation pattern A in the following.

To obtain a final observed sample with such characteristics, the entry times V were drawn from a truncated normal distribution defined on the age range 75 to 95. More specifically, the truncated normal distribution was specified to have a mode equal to the maximum age of 95 with parameter values chosen so that the distribution of the observed study entry times in the truncated sample had the desired shape (the left panel of Figure 6.1 illustrates this procedure). The initial number of generated survival times was chosen such that the final truncated samples had an average size of about $m_V = 500$ individuals.

An independent censoring mechanism was imposed in the following way. For most individuals, the censoring times were the end of a planned individual follow-up period of $t_C = 4$ years. However, some of the follow-up times were longer than four years, and some premature random drop-outs occurred. Again, this was done in response to the situation that we observed in the data application of Section 6.5. Accordingly, we generated random durations from a mixture distribution with an 85% point mass at t_C, a 10% uniform distribution on $[0, t_C]$, and a 5% uniform distribution on $[t_C, t_C + 0.5]$, with the latter two covering the drop-outs before t_C and the longer follow-up periods, respectively. These random durations were added to the individual V_i, and the individual censoring time C_i was the minimum of this sum and age 95.

The right panel of Figure 6.1 illustrates how the mechanisms of truncation and censoring jointly determine the number of individuals at risk at any time t across the age range $[75, 95]$. Truncation pattern A causes the number of individuals at risk to increase steeply at the early ages, and then to decrease only gradually across the age range. However, due to the relatively short individual follow-up times, the number of individuals at risk across ages is considerably smaller than the total sample size of about 500.

In the setting with a positive association between the recurrence and mortality process ($\gamma = 0.5$), additional simulation scenarios were set up by varying the censoring and truncation patterns.

First, we considered the effect of changing the planned individual follow-up times to $t_C = 1$ year or $t_C = 8$ years, respectively. Longer individual follow-up times increase the number of individuals who are under observation at a certain time t, and are therefore expected to improve the estimator performance.

Second, we explored a scenario with a more unimodal distribution of the study entry times in which relatively few individuals entered the study at the youngest and the oldest ages (see Figure 6.1). This is truncation pattern B. To obtain a final sample with these characteristics, we simulated the initial truncation times again from a truncated normal distribution on the age range 75 to 95. However, in this scenario, the distribution had a mode equal to 90, that is, within the above age range.

Finally, we examined a setting with a wider age range of $[64, 105]$. If $t = 0$ was now expected to correspond to age 64, but the Gompertz-Makeham rates were expected

to agree with the rates of the base scenario over $[75, 95]$, the parameters needed to be adapted. This was achieved by maintaining the values of b and c, but setting $a = 0.6$ or $a = 0.05$ for the recurrence and death processes, respectively. The initial distributions of the study entry times were adapted to produce truncation patterns A or B on the wider age range $[64, 105]$. In all of the additional scenarios, the truncated samples again had a target size of $m_V = 500$ individuals. The parameter values for the distributions of the study entry times and the initial sample sizes for the different scenarios can be found in Section 6.7.3.

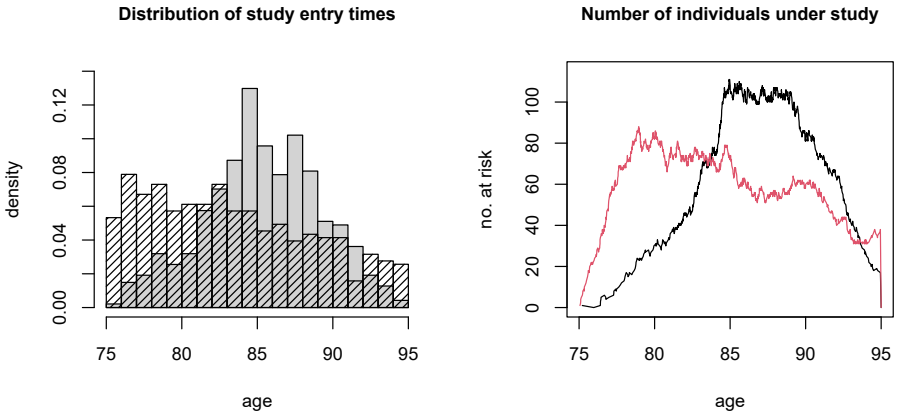

Figure 6.1: Distribution of the ages at study entry (left) and the number of individuals at risk across the age range $[75, 95]$ (right) for one simulated sample from the base scenario with truncation pattern A (shaded bars, red line) or truncation pattern B (gray bars, black line), both with planned individual follow-ups of four years.

6.4.2 Estimation and results

The estimation of the joint frailty model was carried out under the assumptions of gamma distributed frailties with a mean of one and piecewise constant models for the two baseline rates $\lambda_0(t)$ and $h_0(t)$. For both rates, 10 intervals were used that were denoted by I_k^R (recurrence process) and I_k^D (mortality), $k = 1, ..., 10$. The intervals were determined by the deciles of the observed recurrence and survival times, respectively. We set $t_0^R = t_0^D = 75$ (or $t_0^R = t_0^D = 64$) equal to the starting point of the respective age range and $t_{10}^R = t_{10}^D$ equal to the maximum follow-up time in the sample. The marginal likelihood was approximated using non-adaptive Gauss-Hermite quadrature with $Q = 30$ quadrature points. We ran 200 replications in each setting. All computations were performed in R (R Core Team, 2020). Further details on the implementation are provided in Section 6.7.2.

Figures 6.2 and 6.3 illustrate the results of the base scenario with different underlying associations, $\gamma \in \{-0.5, 0, 0.5\}$. The top panels of Figure 6.2 show that the covariate effects α and β, the dependence parameter γ, and the frailty variance θ are estimated without significant bias. The estimated standard errors of these parameters in the bottom panels of Figure 6.2 are largely in line with the empirical standard deviations of the respective parameter estimates across the replications. Nevertheless, we notice that the estimator performance varies for different true values of the dependence parameter.

This pattern can be explained to some extent by different survivor selection effects. The truncated sample consists of survivors, who tend to have lower mortality risks. If the recurrence process and the mortality process are positively associated, this implies that the frailty values and the recurrence rates are lower in the sample of survivors. In the current setting, this lower frailty variance in the sample is favorable for the estimation of θ; whereas the low recurrence rate, which is associated with higher probabilities of having no observed recurrent event, increases the variability in the corresponding estimated covariate effect $\hat{\beta}$. The opposite effects are observed if the event processes are negatively associated. If the recurrence rate has no effect on survival ($\gamma = 0$), the method still yields reliable results, and the parameters exclusively affecting survival are estimated with higher levels of precision.

The estimates of the baseline rates, displayed for the base scenario with positive dependence $\gamma = 0.5$ in Figure 6.3, also perform satisfactorily.

It is instructive to look at how the results change if the effects of survivor selection on the frailty distribution in the sample are not taken into account correctly. As Figures 6.4 and 6.5 show, if the inference is based on the naively constructed marginal likelihood (6.8), biases can be seen in all parameter estimates in case the recurrence process and the mortality process are associated. Moreover, as the estimated standard errors for the covariate effects are substantially smaller than those obtained using the correct likelihood, they do not adequately reflect the uncertainty in the parameter estimates. The baseline rates of recurrence and death are increasingly underestimated for advancing age in the base scenario with positive dependence ($\gamma = 0.5$), as depicted in Figure 6.5. This is because in a setting with a positive association, the distribution of frailty among the survivors tends to be concentrated at lower values. Accordingly, for negative associations, the recurrence rate will be overestimated at the older ages, while the hazard of death will again be underestimated at the older ages. Hence, failing to construct the marginal likelihood based on the correct distribution of the frailty, see (6.7), introduces marked biases in the estimates and the standard errors. Only if the event rates are not associated ($\gamma = 0$) is the distribution of frailty among survivors equal to the initial frailty distribution G_θ, such that the naive marginal likelihood coincides with the correct marginal likelihood (6.7) and yields valid inferences.

Lastly, we want to examine the results for the additional simulation scenarios with modified censoring and truncation patterns. The figures illustrating these results can be found in Section 6.7.3. In the scenario with a planned individual follow-up of only $t_C = 1$ year, we find increased variability in all parameter estimates (see Figures 6.8 and 6.9).

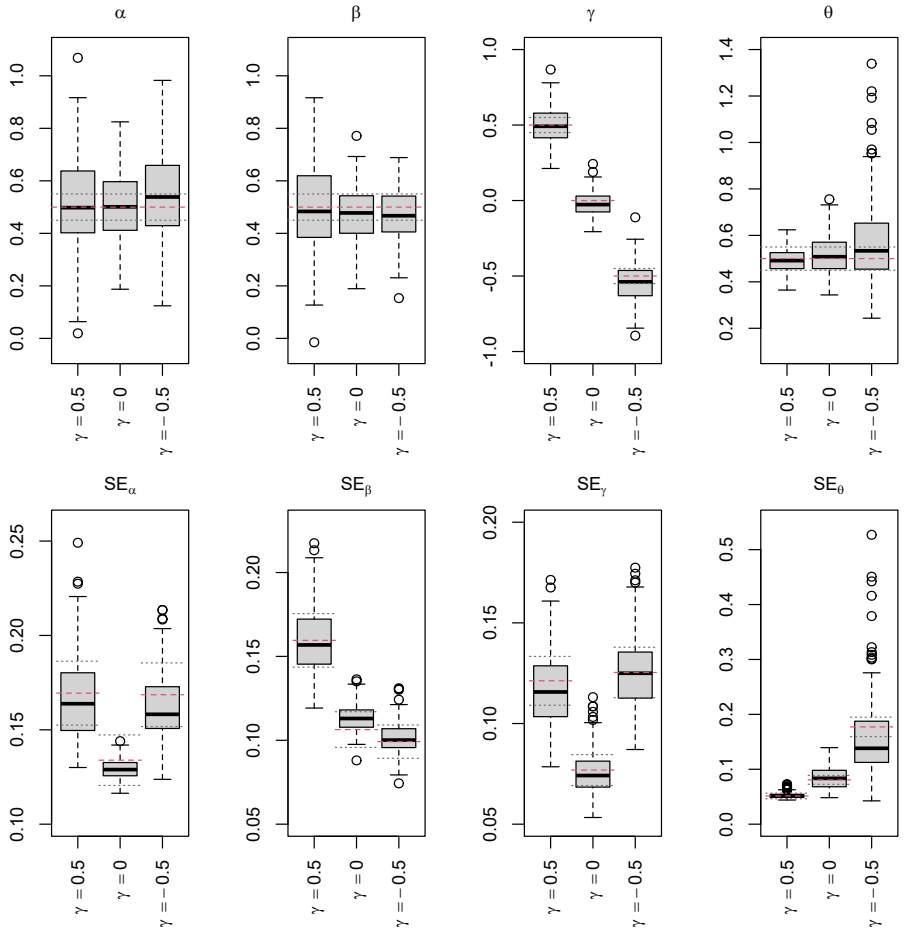

Figure 6.2: Box plots of the parameter estimates (top) and estimated standard errors (bottom) in the joint frailty model for positive ($\gamma = 0.5$), no ($\gamma = 0$), or negative ($\gamma = -0.5$) dependence under the base scenario. Left to right: covariate effect on mortality (α) and on recurrences (β), dependence parameter (γ), and frailty variance (θ) based on 200 truncated samples with a target size of 500. The red dashed line marks the true parameter value (top) or empirical standard deviation (bottom); the gray dotted lines mark 10% deviations from the respective value.

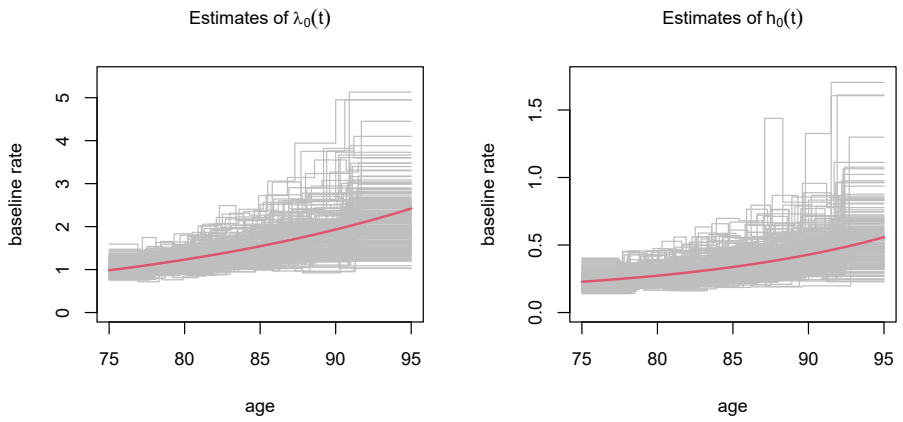

Figure 6.3: Estimates (gray) of the baseline rate of recurrence (left) and of death (right) based on 200 truncated samples with a target size of 500 generated from a joint frailty model with positive dependence ($\gamma = 0.5$) under the base scenario. The red bold line gives the true baseline rate.

This is expected, because with shorter individual follow-up times, fewer individuals are observed at a given age t than in the base scenario. Further extending the planned individual follow-up times of the base scenario from $t_C = 4$ to $t_C = 8$ years does not lead to considerable improvements, apart from some reduced variability in the estimates of the frailty variance and the baseline rates.

A change in the distribution of the study entry times can markedly influence the estimation results. In the modified base scenario with truncation pattern B, the estimated covariate effects $\hat{\alpha}$ and $\hat{\beta}$ are more variable than under truncation pattern A, occasionally with negative estimates (see Figure 6.10). In addition, the first piece of each of the baseline rates shows an upward bias (cf. top panels of Figure 6.11) because few individuals entered the study at the younger ages. Although the intervals for the rate pieces were constructed to contain roughly equal numbers of observed events, the first intervals cover a relative large age range with few individuals under study at a given age t due to the delayed entry.

The last scenario combines a wider age range $[64, 105]$ and truncation times spread across the whole age range according to pattern A or B, with individual follow-ups planned for $t_C = 4$ years. This setting is more demanding because the amount of information available at a given age t is considerably smaller than it is in the scenarios with age range $[75, 95]$. Therefore, the variability in the estimates tends to increase, and the estimates of the dependence parameter and the frailty variance exhibit a small downward bias (see Figure 6.10).

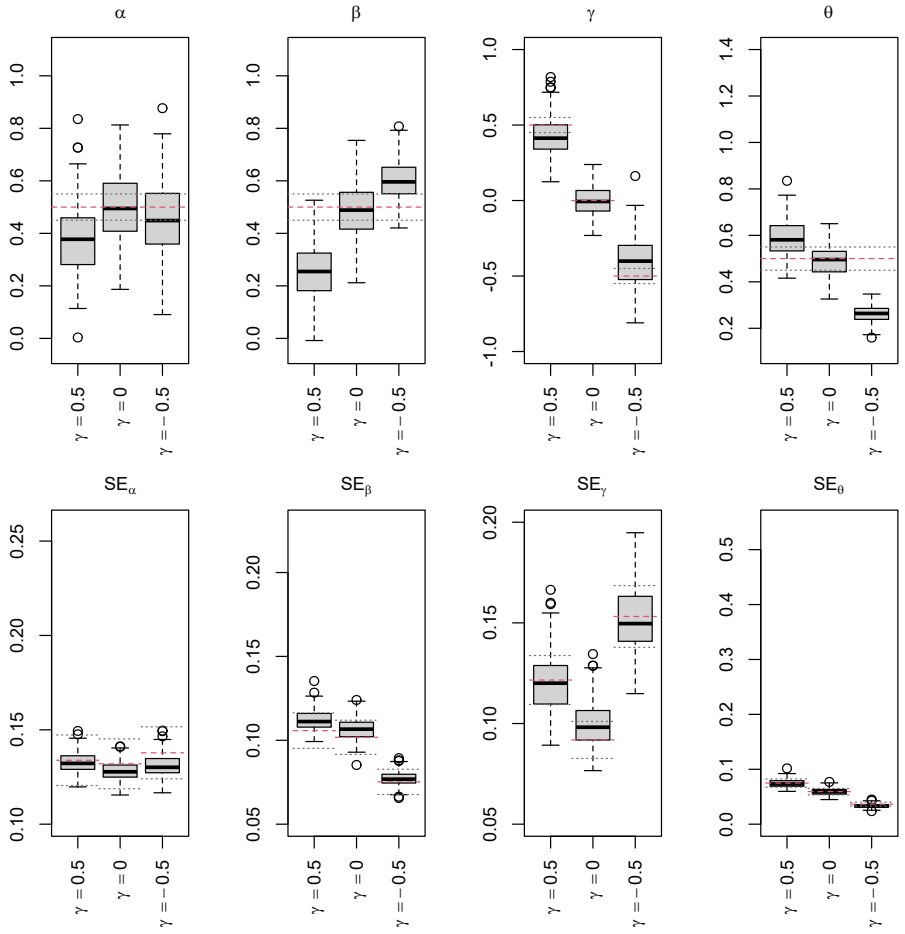

Figure 6.4: Box plots of the parameter estimates (top) and estimated standard errors (bottom) based on the naive likelihood of the joint frailty model for positive ($\gamma = 0.5$), no ($\gamma = 0$), or negative ($\gamma = -0.5$) dependence under the base scenario. Left to right: covariate effect on mortality (α) and on recurrences (β), dependence parameter (γ), and frailty variance (θ) based on 200 truncated samples with a target size of 500. The red dashed line marks the true parameter value (top) or empirical standard deviation (bottom); the gray dotted lines mark 10% deviations from the respective value.

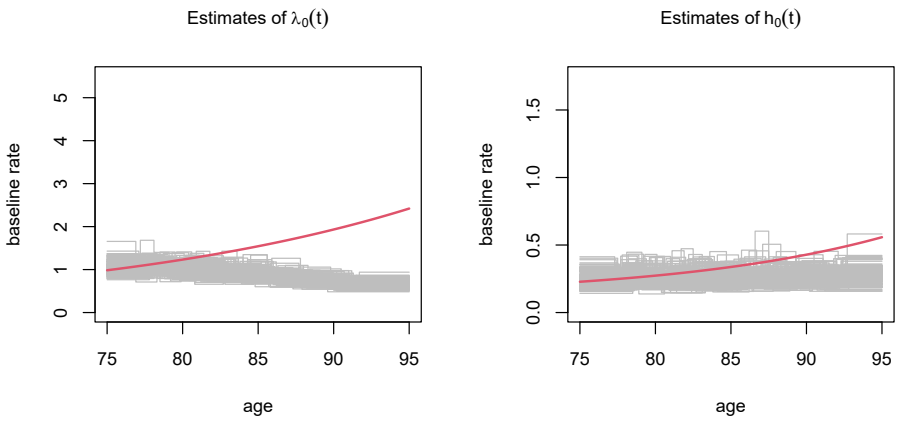

Figure 6.5: Estimates (gray) of the baseline rate of recurrence (left) and of death (right) based on the naive likelihood for 200 truncated samples with a target size of 500 generated from a joint frailty model with positive dependence ($\gamma = 0.5$) under the base scenario. The red bold line gives the true baseline rate.

Overall, the simulation studies suggest that the proposed method for the estimation of the joint frailty model based on left-truncated data performs satisfactorily. The parameter estimates are largely unbiased if the study design ensures that a reasonable number of individuals are under observation across time t. Including a relatively large number of individuals early on and a preferably stable number of study entries across the remaining time range benefits the estimation. In addition, the individual follow-up times should be sufficiently long given the total time window and the sample size. As expected, the patterns of censoring and truncation that cause more information to be lost negatively affect the estimator performance.

6.5 Recurrent infections and mortality in an older population

We use the proposed method to analyze recurrent urinary tract infections and mortality in an institutionalized elderly population. The data come from a double-blind, randomized, placebo-controlled trial in long-term care facilities that aimed to assess the effect of cranberry capsules on the occurrence of UTIs in vulnerable older persons (Caljouw et al., 2014). At baseline, the participants were stratified into two groups of high or low UTI risk depending on whether they had diabetes mellitus, a urinary catheter, or at least one treated UTI in the preceding year. Within these two strata, the participants were randomly assigned to the treatment or the control group. The participants took cranberry or placebo capsules twice a day over a period of one year. Occurrences of UTIs were recorded by

the treating physicians according to a clinical definition based on international practice guidelines for LTCF residents, and a strict definition based on scientific criteria. We focus here on the occurrence of the more broadly defined clinical UTIs.

The final study population consisted of 928 individuals, most of whom were women (703; 75.8%). Of these individuals, 516 were considered to be at high baseline UTI risk, while 412 were considered to be at low baseline UTI risk. Individuals entered the study between ages 64 and 102, as shown in the left panel of Figure 6.6, and were followed on average for about a year (mean: 332 days, median: 372 days). A total of 317 participants (34.2%) died during the study period. The number of observed UTIs per individual ranged from zero to 10, with 62.2% of the individuals having no UTIs, 20.8% having one UTI, and 17.0% experiencing two or more UTIs during the follow-up period.

Unlike in the original study, we modeled recurrent UTIs and mortality to evolve with age, where $t_0 = 0$ corresponded to age 64. Because of the specific distribution of the ages at study entry in conjunction with the short individual follow-up times, relatively few individuals were under observation at a given age, in particular at the youngest and oldest ages, as the right panel of Figure 6.6 shows.

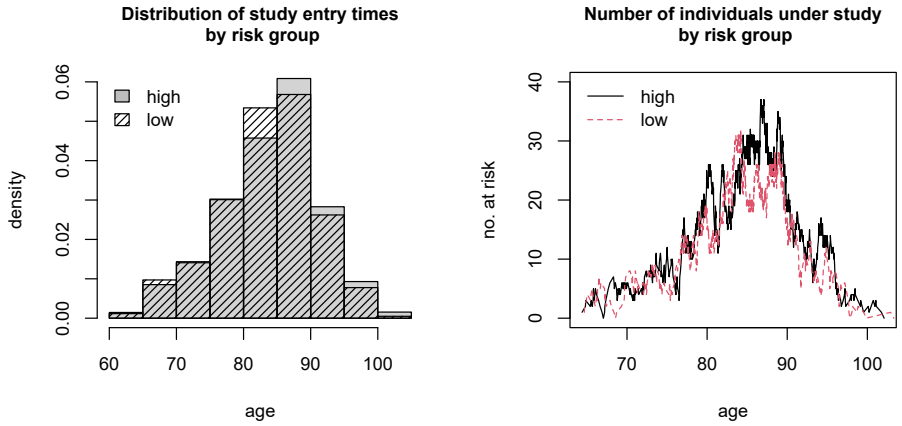

Figure 6.6: Distribution of ages at study entry (left) and number of individuals under observation across the age range (right) in the cranberry data set, separately for the groups with high baseline UTI risk (gray bars, black solid line) and low baseline UTI risk (shaded bars, red dashed line).

We estimated the joint frailty model for UTIs and overall mortality with age as the time scale separately for the groups with high and low baseline UTI risk. Two binary covariates for treatment and gender were included, and frailties were assumed to follow a gamma distribution with a mean of one. The baseline rate of UTI recurrence and the hazard of death were specified as piecewise constant functions with 10 intervals over the age range 64 to 103 in the high risk group and 64 to 104 in the low risk group.

Separately for the two risk groups, the cut-points for the intervals were determined based on the deciles of the observed recurrence or death times, respectively. The likelihood was approximated using non-adaptive Gaussian quadrature with 30 nodes.

The parameter estimates for both risk groups are reported in Table 6.1. In the group with a high baseline UTI risk, the infection rates varied between participants with an estimated frailty variance of $\hat{\theta} = 0.380$ (SE: 0.086). In particular, individuals with a higher rate of recurrent infections tended to also experience higher mortality risks, as indicated by the positive estimate of the dependence parameter, $\hat{\gamma} = 0.181$ (SE: 0.084). The participants in the low risk group seemed to be more heterogeneous ($\hat{\theta} = 1.122$, SE: 0.316), but the analysis did not detect an association between the occurrence of UTIs and survival ($\hat{\gamma} = 0.058$, SE: 0.044). The results suggest that the cranberry capsules did not have a noticeable effect on the occurrence of UTIs irrespective of the baseline UTI risk. When we look at gender differences, we see that males and females experienced similar infection rates, while males had higher mortality levels than females in both groups.

Table 6.1: Parameter estimates (with standard errors) for the joint frailty model fitted to the cranberry data set, separately by risk group.

	High baseline UTI risk	Low baseline UTI risk
Recurrent UTIs		
Treatment (cranberry)	0.000 (0.161)	0.189 (0.217)
Gender (male)	0.061 (0.218)	−0.384 (0.381)
Mortality		
Treatment (cranberry)	0.107 (0.152)	−0.001 (0.197)
Gender (male)	0.396 (0.178)	0.787 (0.210)
Association		
Dependence γ	0.181 (0.084)	0.058 (0.044)
Frailty variance θ	0.380 (0.086)	1.122 (0.316)

The estimated baseline rates displayed in Figure 6.7 demonstrate nicely the age dependence of the recurrence rate and the hazard of death. For the individuals with a high baseline UTI risk, both the rate of recurrent infection and the mortality risk showed a general tendency to increase with age, although the small number of observations leads to considerable uncertainty at the highest ages. In addition, the individuals with a high baseline UTI risk tended to experience higher rates of recurrent infection and death than the individuals with a low baseline UTI risk.

The original study, which used a different time scale, reported a positive treatment effect of the intake of cranberry capsules only in the group with a high baseline UTI risk and only for the outcome of UTI incidence (first infection during follow-up). When all recurrent UTIs were analyzed in a gamma-frailty model, no treatment effect was detected, which is in line with the findings presented here.

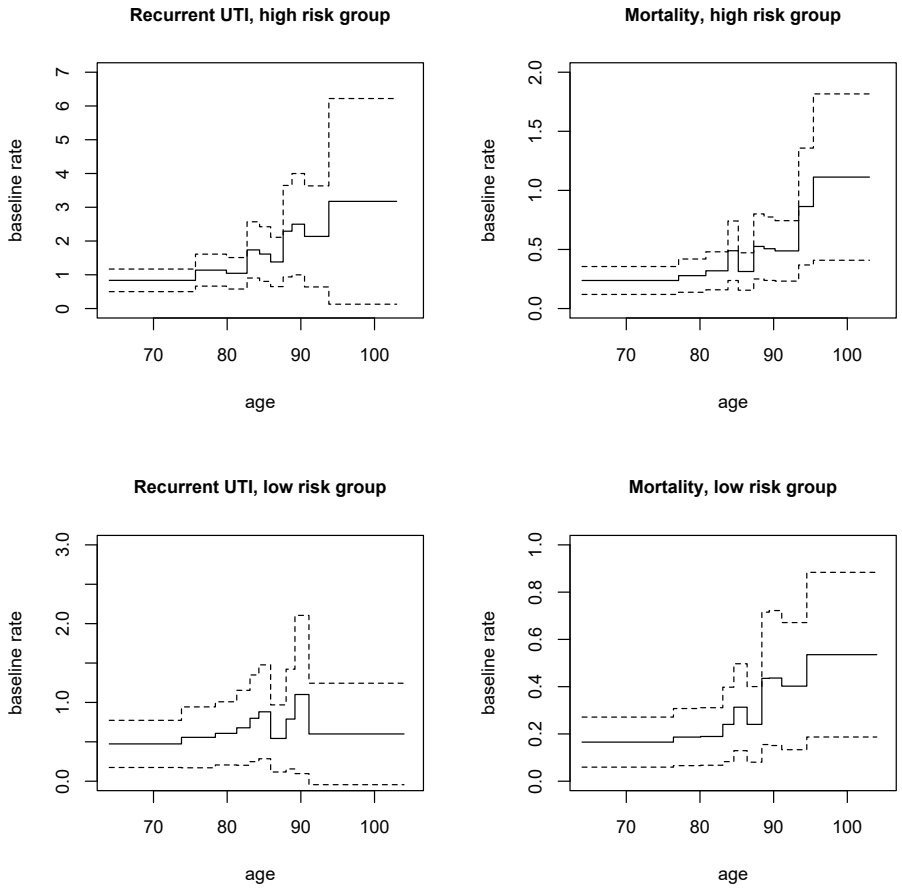

Figure 6.7: Estimated baseline rates (solid) of recurrence (left) and mortality (right) with ±2 SE-confidence bounds (dashed) for the cranberry data, separately for the groups with high baseline UTI risk (top) and low baseline UTI risk (bottom).

6.6 Discussion

We have proposed a method for estimating the joint frailty model for recurrent events and a terminal event based on left-truncated data. The marginal likelihood of the model can be expressed as a ratio of two integrals over the frailty distribution, each of which is approximated using Gauss-Hermite quadrature. The direct maximization of the approximate marginal likelihood is possible if the baseline rates are specified as piecewise constant functions.

The simulation studies presented here have shown that the estimation procedure performs satisfactorily in general, and have demonstrated how different observation schemes affect the estimator performance. While any pattern of truncation or censoring results in incomplete information, study designs should still aim to provide enough information to meet the needs of a model as complex as the joint frailty model. Having a sufficient number of individuals under observation across most of the time range, and especially at the start of the process, seems to be crucial for the method to yield reliable results.

Allowing for left truncation in frailty models requires us to consider carefully how the frailty distribution in the sample of survivors may differ from the frailty distribution in the underlying population due to selection effects. We illustrated through simulations the biases that can arise in the parameter estimates of the joint frailty model if this difference in the frailty distributions is ignored.

Extending the framework of the joint frailty model to incorporate delayed entry allowed us to study age-specific rates of recurrent urinary tract infections and death in an older population. Similarly, the proposed approach enables researchers to use the joint frailty model in a wider variety of contexts in which subjects are included in a study only if they have not yet experienced the terminal event. Apart from clinical studies with delayed entry, these contexts may include register-based studies of event processes evolving with age as the main time scale, with individuals entering at different ages.

For a complete specification of the model and the approximate likelihood function, we need to choose a frailty distribution as well as the number of quadrature points and the intervals for the baseline rates. Although the simulation study and the application covered only the common choice of a gamma distribution for the frailties, the quadrature approach can be employed with other frailty distributions that have a closed-form inverse distribution function, or a log-normal distribution. The number of quadrature points then determines the accuracy of the integral approximations in the marginal likelihood, as well as the computation time. In line with previous recommendations for gamma frailty models (see Liu and Huang, 2008), we used $Q = 30$ quadrature points, which produced good results in a reasonable period of time in our settings. Regarding the baseline rate functions, the number of intervals corresponds to the number of parameters for the rates, and should thus be selected to allow for sufficient flexibility of the shape of the rates, while retaining numerical stability and the computational feasibility of the method. Adequate results can often be obtained with moderate numbers of up to 10 intervals.

Nevertheless, in some applications, it may seem appealing to aim for smooth rate estimates, such as through the use of penalized splines. However, automatic smoothing parameter selection in joint frailty models is an issue that needs further investigation.

Moreover, the current approach is limited to applications in which there is heterogeneity in recurrence rates. Due to the specific dependence structure in the joint frailty model considered here, the association between the recurrence process and the terminal event process cannot be assessed if the frailty variance is close to zero.

Finally, in some applications, it might be of interest to extend the proposed method to use information on recurrences before entry into the study. These additional event

times can be included in the marginal likelihood, and are expected to lead to increased precision in the estimation of the model for the recurrence process. However, when using such an approach, researchers should reflect critically on the quality of the retrospectively collected data, as recollections by study participants may be less reliable than data drawn from other sources, such as registries.

6.7 Supplementary material

6.7.1 Approximation of the marginal likelihood using Gaussian quadrature

In this section, we will elaborate on the use of Gauss-Hermite quadrature for approximating the marginal likelihood of the joint frailty model. We will first recap the quadrature approach as proposed by Liu and Huang (2008) in the setting with right-censoring only, and then show how to adapt the method to the setting with left truncation.

An approximation to marginal likelihoods integrated over normal random effects based on Gauss-Hermite quadrature was already presented in Section 6.3. But the marginal likelihoods of the joint frailty model given in Section 6.2 involve integrals over non-normal random effects. Thus, we use the probability integral transformation (Nelson et al., 2006; Liu and Huang, 2008) to rewrite the integrals over the random effect u with distribution function $G_\theta(u)$ as integrals over standard normal random effects. This relies on the fact that the $G_\theta(u)$ have a standard uniform distribution, such that their transformations $a = \Phi^{-1}[G_\theta(u)]$ follow a standard normal distribution, if $\Phi(\cdot)$ denotes the standard normal distribution function.

For the marginal likelihood contribution (6.5) of individual i in the joint frailty model without truncation, the substitution $u = G_\theta^{-1}[\Phi(a)]$ yields

$$L_i = \int_0^\infty L_i^{(c)}(u)\, g_\theta(u)\, \mathrm{d}u = \int_{-\infty}^\infty L_i^{(c)}(G_\theta^{-1}[\Phi(a)])\, \phi(a)\, \mathrm{d}a.$$

We can then employ Gauss-Hermite quadrature as in (6.11) to arrive at the approximate marginal likelihood contribution of the form

$$L_i \approx \sum_{q=1}^{Q} L_i^{(c)}(G_\theta^{-1}[\Phi(\tilde{x}_q)])\, \phi(\tilde{x}_q)\, \tilde{w}_q.$$

In the setting with left-truncated data, the marginal likelihood of the joint frailty model has a slightly more complex structure. In (6.10), the marginal likelihood contribution ${}_VL_i$ of individual i is expressed as a ratio of two integrals over the density $g_\theta(u)$. Hence, we will approximate the likelihood by applying the above approach separately to the two integrals,

$${}_VL_i = \frac{\int_0^\infty {}_VL_i^{(c)}(u)\, \exp\{-H_i(v_i|u)\}\, g_\theta(u)\, \mathrm{d}u}{\int_0^\infty \exp\{-H_i(v_i|u)\}\, g_\theta(u)\, \mathrm{d}u}$$

$$\approx \frac{\sum_{q=1}^{Q} {}_{V}L_i^{(c)}(G_\theta^{-1}[\Phi(\tilde{x}_q)])\, \exp\{-H_i(v_i|G_\theta^{-1}[\Phi(\tilde{x}_q)])\}\, \phi(\tilde{x}_q)\, \tilde{w}_q}{\sum_{q=1}^{Q} \exp\{-H_i(v_i|G_\theta^{-1}[\Phi(\tilde{x}_q)])\}\, \phi(\tilde{x}_q)\, \tilde{w}_q},$$

with ${}_{V}L_i^{(c)}(u)$ given in (6.6). The approximate marginal likelihood of the joint frailty model with left truncation is then given by

$$\prod_{i=1}^{m_V} \frac{\sum_{q=1}^{Q} {}_{V}L_i^{(c)}(G_\theta^{-1}[\Phi(\tilde{x}_q)])\, \exp\{-H_i(v_i|G_\theta^{-1}[\Phi(\tilde{x}_q)])\}\, \phi(\tilde{x}_q)\, \tilde{w}_q}{\sum_{q=1}^{Q} \exp\{-H_i(v_i|G_\theta^{-1}[\Phi(\tilde{x}_q)])\}\, \phi(\tilde{x}_q)\, \tilde{w}_q}.$$

In the naive likelihood (6.8), there is only one integral over the frailty distribution, and hence only one approximation is required,

$${}_{V}L_i^{\text{naive}} = \int_0^\infty {}_{V}L_i^{(c)}(u)\, g_\theta(u)\, \mathrm{d}u \approx \sum_{q=1}^{Q} {}_{V}L_i^{(c)}(G_\theta^{-1}[\Phi(\tilde{x}_q)])\, \phi(\tilde{x}_q)\, \tilde{w}_q.$$

6.7.2 Computational details

We used `R` (R Core Team, 2020) to implement the estimation procedure. The quadrature points and weights were calculated using function `gauss.quad()` from package `statmod` (Smyth, 1998). For numerical optimization of the approximate marginal log-likelihood, we applied function `nlm()`, which performs minimization based on a Newton-type algorithm, to the negative log-likelihood. The Hessian of the marginal log-likelihood was approximated numerically using function `hessian()` from package `numDeriv` (Gilbert and Varadhan, 2019).

As the frailty variance and the parameters of the piecewise constant baseline rates are restricted to be non-negative, the log-likelihood was maximized with respect to the log-transform of these parameters, which guaranteed non-negative estimates. The delta-method was then applied to derive the respective standard errors.

However, when specifying the baseline rates as piecewise constant functions, the numerically computed Hessian of the log-likelihood may in some cases not be invertible. To overcome this issue, one can add to the log-likelihood small, fixed ridge penalties (e.g., with penalty parameter 10^{-6}) on the logarithm of the rate parameters and derive the Hessian matrix from this penalized log-likelihood.

6.7.3 Supplement to the simulation studies

Generation of different truncation patterns

A brief description of how the truncation patterns A and B were generated was already given in Section 6.4. For both patterns, truncation times were simulated from a truncated normal distribution defined on the corresponding age range $[75, 95]$ or $[64, 105]$. However, the parameters μ and σ^2 of the underlying normal distribution were chosen such that

the density of the resulting truncated normal distribution was either increasing over the whole age range with mode equal to the maximum age or unimodal with a mode within the age range. These two distinct shapes yielded the desired truncation patterns A (TrA) or B (TrB) in the final samples. Table 6.2 reports, for the different settings, the parameter values for the distribution of the truncation times as well as the initial sample size M needed to obtain truncated samples with an average size of $m_V = 500$.

When implementing the simulation study in `R`, we used function `rtruncnorm()` from package `truncnorm` (Mersmann et al., 2018) for drawing random numbers from a truncated normal distribution.

Table 6.2: Initial sample size M and parameter values for the distribution of the truncation times.

γ	age range	pattern	M	μ	σ^2
0.5	[75, 95]	TrA	$1.07 \cdot 10^4$	109	124
0.5	[75, 95]	TrB	$1.20 \cdot 10^4$	90	18
0.5	[64, 105]	TrA	$7.65 \cdot 10^4$	120	225
0.5	[64, 105]	TrB	$4.75 \cdot 10^4$	93	52
0	[75, 95]	TrA	$3.50 \cdot 10^4$	115	110
–0.5	[75, 95]	TrA	$3.44 \cdot 10^4$	115	109

Additional figures for the simulation results

The following figures illustrate the impact of different censoring and truncation patterns on the performance of the estimation procedure. All settings are modifications of the base scenario with positive dependence $\gamma = 0.5$, that was presented in Section 6.4. In particular, samples with a target size of $m_V = 500$ were generated from a joint frailty model with covariate effects $\alpha = \beta = 0.5$, frailty variance $\theta = 0.5$, and Gompertz-Makeham baseline rates.

- In Figures 6.8 and 6.9, the effect of different censoring mechanisms is studied by comparing the estimation results for different lengths of the planned individual follow-up $t_C \in \{1, 4, 8\}$ years.
- Changes in the distribution of truncation times are examined in Figures 6.10 and 6.11. The displayed settings assume different distributions of the study entry times in the final sample, both in terms of the shape (TrA: truncation pattern A, or TrB: truncation pattern B) and the support of the distribution (75+: ages 75 to 95, or 64+: ages 64 to 105).

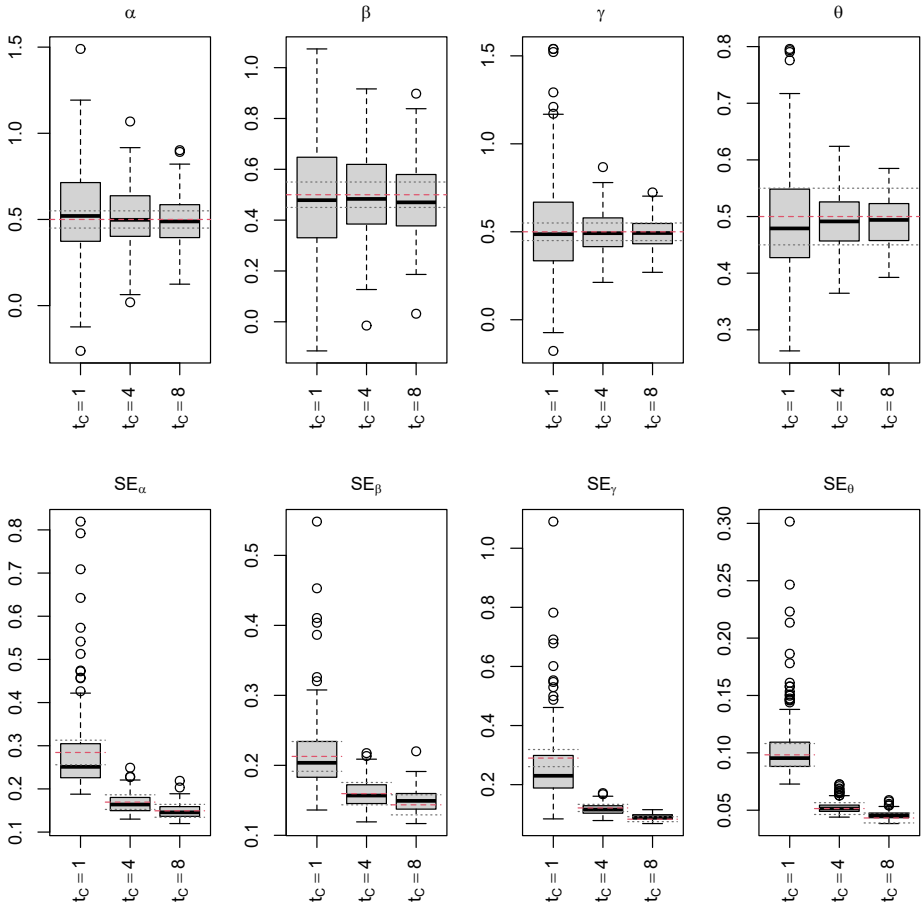

Figure 6.8: Box plots of the parameter estimates (top) and estimated standard errors (bottom) in the joint frailty model with positive dependence ($\gamma = 0.5$) as in the base scenario with truncation pattern A for ages 75+, but varying the planned individual follow-up $t_C \in \{1, 4, 8\}$. Left to right: covariate effect on mortality (α) and on recurrences (β), dependence parameter (γ), and frailty variance (θ) based on 200 truncated samples with a target size of 500. The red dashed line marks the true parameter value (top) or empirical standard deviation (bottom); the gray dotted lines mark 10% deviations from the respective value.

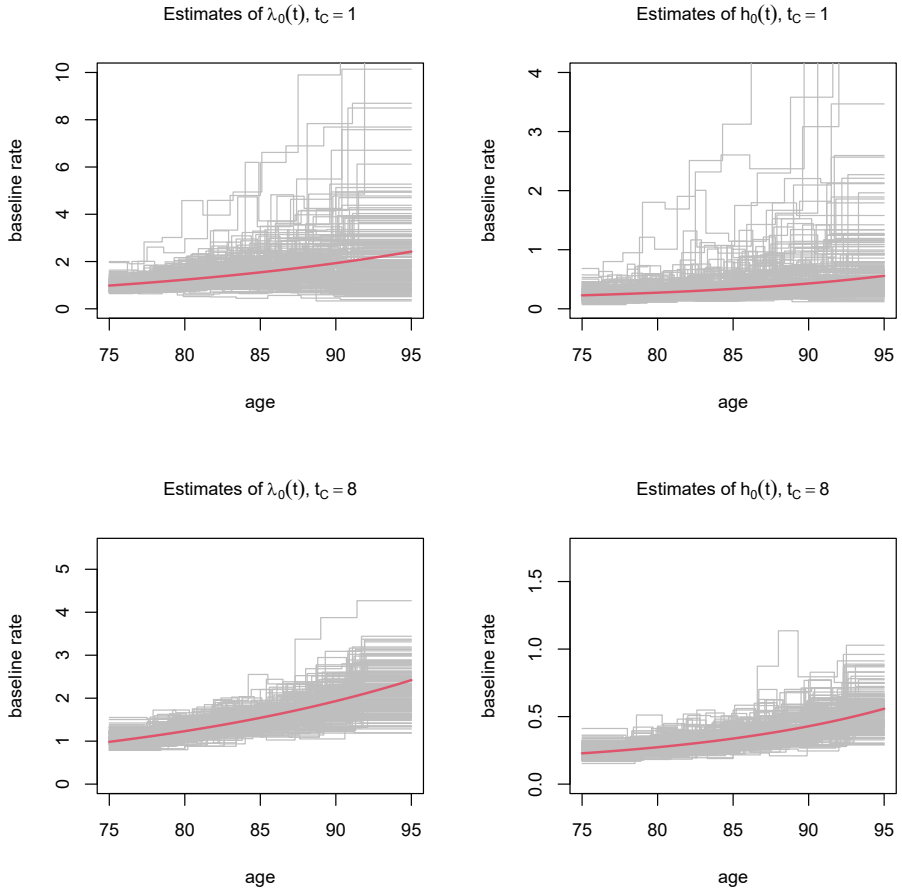

Figure 6.9: Estimates (gray) of the baseline rate of recurrence (left) and of death (right) based on 200 truncated samples with a target size of 500 generated from a joint frailty model with positive dependence ($\gamma = 0.5$). As in the base scenario truncation follows pattern A for ages 75+, but planned individual follow-up is $t_C = 1$ (top) or $t_C = 8$ (bottom) year(s). The red bold line gives the true baseline rate. (Note the different scales of the vertical axes.)

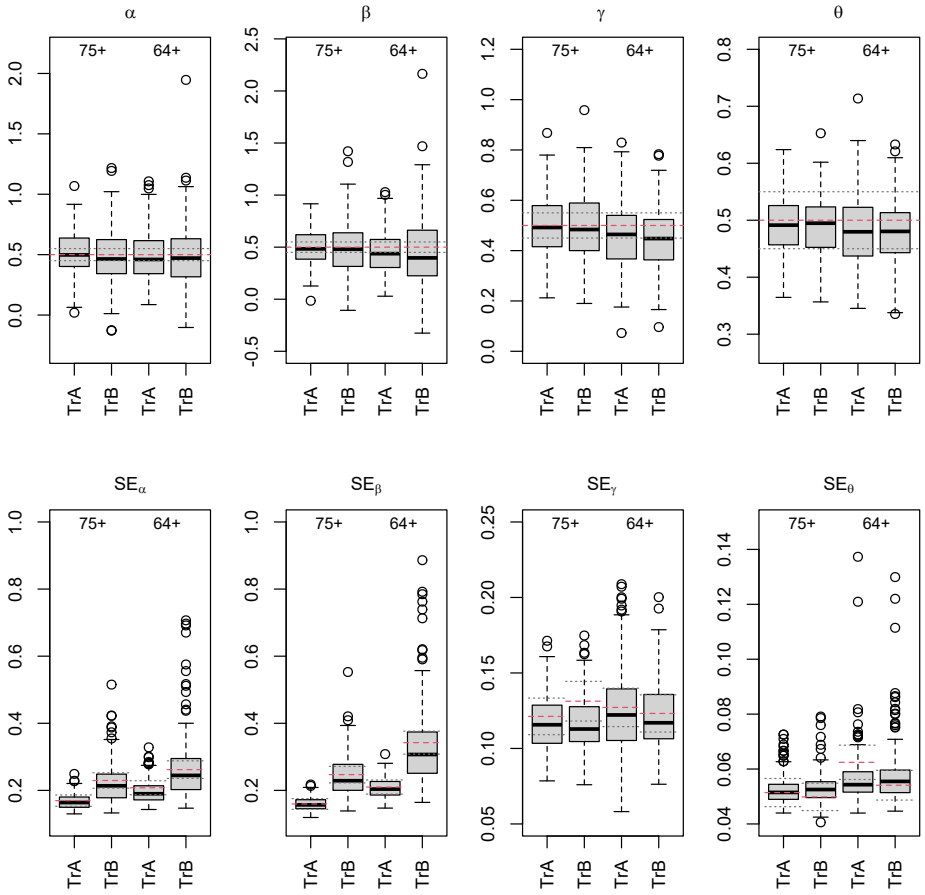

Figure 6.10: Box plots of the parameter estimates (top) and estimated standard errors (bottom) in the joint frailty model with positive dependence ($\gamma = 0.5$) as in the base scenario with $t_C = 4$, but truncation times distributed with different shapes (TrA: truncation pattern A, TrB: truncation pattern B) and across different age ranges (75+: $[75, 95]$, 64+: $[64, 105]$). Left to right: covariate effect on mortality (α) and on recurrences (β), dependence parameter (γ), and frailty variance (θ) based on 200 truncated samples with a target size of 500. The red dashed line marks the true parameter value (top) or empirical standard deviation (bottom); the gray dotted lines mark 10% deviations from the respective value.

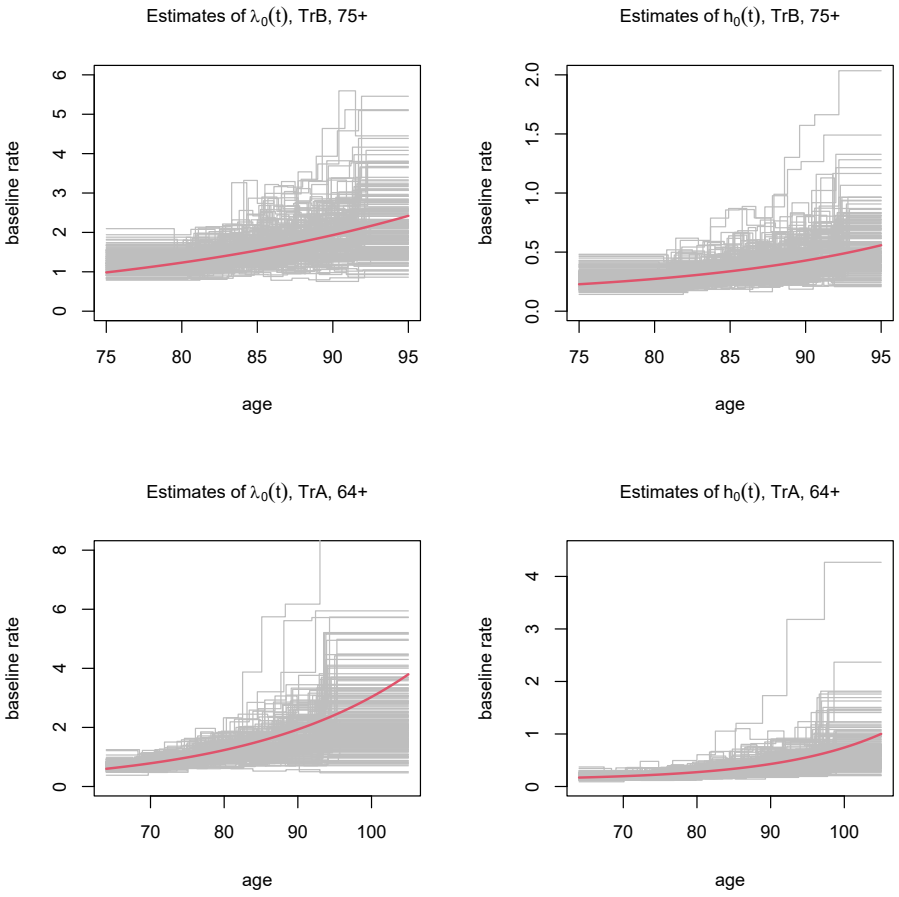

Figure 6.11: Estimates (gray) of the baseline rate of recurrence (left) and of death (right) based on 200 truncated samples with a target size of 500 generated from a joint frailty model with positive dependence ($\gamma = 0.5$) and planned individual follow-up of $t_C = 4$ years as in the base scenario. Truncation times are distributed according to pattern B across ages $[75, 95]$ (top; TrB, 75+) or according to pattern A across ages $[64, 105]$ (bottom; TrA, 64+). The red bold line gives the true baseline rate. (Note the different scales of the horizontal and vertical axes.)

References

Balan, T. A., M. A. Jonker, P. C. Johannesma, and H. Putter (2016). Ascertainment correction in frailty models for recurrent events data. *Statistics in Medicine 35*, 4183–4201.

Cai, Q., M.-C. Wang, and K. C. G. Chan (2017). Joint modeling of longitudinal, recurrent events and failure time data for survivor's population. *Biometrics 73*, 1150–1160.

Caljouw, M. A. A., W. B. van den Hout, H. Putter, W. P. Achterberg, H. J. M. Cools, and J. Gussekloo (2014). Effectiveness of cranberry capsules to prevent urinary tract infections in vulnerable older persons: A double-blind randomized placebo-controlled trial in long-term care facilities. *Journal of the American Geriatrics Society 62*, 103–110.

Cook, R. J. and J. F. Lawless (1997). Marginal analysis of recurrent events and a terminating event. *Statistics in Medicine 16*, 911–924.

Cook, R. J. and J. F. Lawless (2007). *The Statistical Analysis of Recurrent Events*. Statistics for Biology and Health. New York: Springer.

Crowther, M. J., T. M.-L. Andersson, P. C. Lambert, K. R. Abrams, and K. Humphreys (2016). Joint modelling of longitudinal and survival data: incorporating delayed entry and an assessment of model misspecification. *Statistics in Medicine 35*, 1193–1209.

Emura, T., M. Nakatochi, K. Murotani, and V. Rondeau (2017). A joint frailty-copula model between tumour progression and death for meta-analysis. *Statistical Methods in Medical Research 26*, 2649–2666.

Eriksson, F., T. Martinussen, and T. H. Scheike (2015). Clustered survival data with left-truncation. *Scandinavian Journal of Statistics 42*, 1149–1166.

Ghosh, D. and D. Y. Lin (2003). Semiparametric analysis of recurrent events data in the presence of dependent censoring. *Biometrics 59*, 877–885.

Gilbert, P. and R. Varadhan (2019). *numDeriv: Accurate Numerical Derivatives*. R package version 2016.8-1.1. https://CRAN.R-project.org/package=numDeriv.

Huang, C.-Y. and M.-C. Wang (2004). Joint modeling and estimation for recurrent event processes and failure time data. *Journal of the American Statistical Association 99*, 1153–1165.

Jazić, I., S. Haneuse, B. French, G. MacGrogan, and V. Rondeau (2019). Design and analysis of nested case-control studies for recurrent events subject to a terminal event. *Statistics in Medicine 38*(22), 4348–4362.

Jensen, H., R. Brookmeyer, P. Aaby, and P. K. Andersen (2004). *Shared Frailty Model for Left-Truncated Multivariate Survival Data*. Biostatistisk Afdeling: Museum Tusculanum.

Lawless, J. F. and D. Y. T. Fong (1999). State duration models in clinical and observational studies. *Statistics in Medicine 18*, 2365–2376.

Liu, D., D. E. Schaubel, and J. D. Kalbfleisch (2012). Computationally efficient marginal models for clustered recurrent event data. *Biometrics 68*, 637–647.

Liu, L. and X. Huang (2008). The use of Gaussian quadrature for estimation in frailty proportional hazards models. *Statistics in Medicine 27*, 2665–2683.

Liu, L., R. A. Wolfe, and X. Huang (2004). Shared frailty models for recurrent events and a terminal event. *Biometrics 60*, 747–756.

Mersmann, O., H. Trautmann, D. Steuer, and B. Bornkamp (2018). *truncnorm: Truncated Normal Distribution*. R package version 1.0-8. https://CRAN.R-project.org/package=truncnorm.

Modig, K., T. Andersson, S. Drefahl, and A. Ahlbom (2013). Age-specific trends in morbidity, mortality and case-fatality from cardiovascular disease, myocardial infarction and stroke in advanced age: Evaluation in the Swedish population. *PLOS ONE 8*. doi: 10.1371/journal.pone.0064928.

Nelson, K. P., S. R. Lipsitz, G. M. Fitzmaurice, J. Ibrahim, M. Parzen, and R. Strawderman (2006). Use of the probability integral transformation to fit nonlinear mixed-effects models with nonnormal random effects. *Journal of Computational and Graphical Statistics 15*(1), 39–57.

Piulachs, X., E.-R. Andrinopoulou, M. Guillén, and D. Rizopoulos (2021). A Bayesian joint model for zero-inflated integers and left-truncated event times with a time-varying association: Applications to senior health care. *Statistics in Medicine 40*(1), 147–166.

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rogers, J. K., A. Yaroshinsky, S. J. Pocock, D. Stokar, and J. Pogoda (2016). Analysis of recurrent events with an associated informative dropout time: Application of the joint frailty model. *Statistics in Medicine 35*, 2195–2205.

Rondeau, V., J. R. Gonzalez, Y. Mazroui, A. Mauguen, A. Diakite, A. Laurent, M. Lopez, A. Król, C. L. Sofeu, J. Dumerc, and D. Rustand (2020). *frailtypack: General Frailty Models: Shared, Joint and Nested Frailty Models with Prediction; Evaluation of Failure-Time Surrogate Endpoints*. R package version 3.3.2. https://CRAN.R-project.org/package=frailtypack.

Rondeau, V., S. Mathoulin-Pelissier, H. Jacqmin-Gadda, V. Brouste, and P. Soubeyran (2007). Joint frailty models for recurring events and death using maximum penalized likelihood estimation: application on cancer events. *Biostatistics 8*(4), 708–721.

Smyth, G. K. (1998). Numerical integration. In P. Armitage and T. Colton (Eds.), *Encyclopedia of Biostatistics*, pp. 3088–3095. London: Wiley.

van den Berg, G. J. and B. Drepper (2016). Inference for shared-frailty survival models with left-truncated data. *Econometric Reviews 35*, 1075–1098.

van den Hout, A. and G. Muniz-Terrera (2016). Joint models for discrete longitudinal outcomes in aging research. *Journal of the Royal Statistical Society, Series C 65*, 167–186.