



Universiteit
Leiden
The Netherlands

Joint modeling of interval counts of recurrent events and death

Bohnstedt, M.; Putter, H.; Danko, A.; Danko, M.J.; Gampe, J.

Citation

Bohnstedt, M., Putter, H., Danko, A., Danko, M. J., & Gampe, J. (2021). Joint modeling of interval counts of recurrent events and death. *Biometrical Journal*, 63(2), 323-340.
doi:10.1002/bimj.201900367

Version: Publisher's Version

License: [Creative Commons CC BY 4.0 license](https://creativecommons.org/licenses/by/4.0/)

Downloaded from: <https://hdl.handle.net/1887/3277348>

Note: To cite this publication please use the final published version (if applicable).

RESEARCH PAPER

Joint modeling of interval counts of recurrent events and death

Marie Böhnstedt^{1,2}  | Hein Putter²  | Aleksandra Dańko³ | Maciej J. Dańko⁴  | Jutta Gampe¹ 

¹ Laboratory of Statistical Demography, Max Planck Institute for Demographic Research, Rostock, Germany

² Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, The Netherlands

³ Department of Biosciences, University of Rostock, Rostock, Germany

⁴ Laboratory of Population Health, Max Planck Institute for Demographic Research, Rostock, Germany

Correspondence

Marie Böhnstedt, Max Planck Institute for Demographic Research, Konrad-Zuse-Str. 1, 18057 Rostock, Germany.

Email: boehnstedt@demogr.mpg.de



This article has earned an open data badge “**Reproducible Research**” for making publicly available the code necessary to reproduce the reported results. The results reported in this article were reproduced partially for data confidentiality reasons and due to their computational complexity.

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

KEYWORDS

frailty, interval counts, joint modeling, recurrent events, score test

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 & 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.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *Biometrical Journal* published by WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

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 (Daňko, Schaible, & Daňko, 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 who 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 & Wang, 2004, in the setting with exact recurrence times, or Lancaster & 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 & Lawless, 1997; Ghosh & Lin, 2003, in the setting with exact recurrence times, or Zhao, Li, & Sun, 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 expectation-maximization (EM) algorithm, whereas Liu and Huang (2008) and Rondeau, Mathoulin-Pelissier, Jacqmin-Gadda, Brouste, and Soubeyran (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, Boonk, Vermeer, and Putter (2016), which was also based on observed recurrence times. It builds on concepts that are similar to the test proposed by Huang, Wolfe, and Hu (2004) for the association between two event processes in clustered survival data.

In this paper, 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 paper is structured as follows. In Section 2, we describe the joint frailty model for recurrent events and death, as well as the setting of individual interval counts. In Section 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 4, we assess the performance of the estimation method and the test in simulation studies. In Section 5, we apply the proposed methods to the data on *E. dichotoma*, followed by a discussion in Section 6.

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 $i, i = 1, \dots, m$. 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 $[t, t + dt)$ are $dN_i^{R*}(t) = N_i^{R*}((t + dt)^-) - N_i^{R*}(t^-)$. Here, t^- denotes the left-hand limit.

Additional observed characteristics of individual i are collected in the covariate vector $\mathbf{z}_i$, 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 $O_i(t) = \{Y_i(s), N_i^R(s), N_i^D(s), 0 \leq s \leq t; \mathbf{z}_i\}$.

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

$$\begin{aligned} P(dN_i^R(t) = 1 \mid \mathcal{F}_{t-}, D \geq t) &= Y_i(t)\lambda_i(t)dt \quad \text{with} \\ \lambda_i(t)dt &= d\Lambda_i(t) = P(dN_i^{R*}(t) = 1 \mid \mathbf{z}_i, u_i, D_i \geq t), \end{aligned} \quad (1)$$

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

Analogously, for the terminal event

$$\begin{aligned} P(dN_i^D(t) = 1 \mid \mathcal{F}_{t-}) &= Y_i(t)h_i(t)dt \quad \text{with} \\ h_i(t)dt &= dH_i(t) = P(dN_i^{D*}(t) = 1 \mid \mathbf{z}_i, u_i, D_i \geq t). \end{aligned}$$

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

$$\begin{aligned} \lambda_i(t) &= u_i e^{\beta' \mathbf{z}_i} \lambda_0(t), \\ h_i(t) &= u_i^\gamma e^{\alpha' \mathbf{z}_i} h_0(t), \end{aligned} \quad (2)$$

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 $\mathbf{z}_i$ enter in (2) in a proportional rates or a proportional hazards formulation, but can have different effects α and β 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 $N_{ij} = N_i^R(t_{ij}) - N_i^R(t_{ij-1})$. N_{ij} shows the number of recurrent events experienced by

individual i in the interval $I_{ij} = (t_{ij-1}, t_{ij}]$, $j = 1, \dots, J_i$. 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 (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 (1) and (2), it follows that $(N_{ij} \mid \mathbf{z}_i, u_i, D_i \geq t_{ij})$ has a Poisson distribution with mean $u_i \mu_{ij}$, where $\mu_{ij} = \int_{I_{ij}} e^{\beta' \mathbf{z}_i} \lambda_0(s) ds$. Therefore, the likelihood contribution $L_i^{(c)}(u_i)$ of individual i conditional on its frailty value u_i is given by

$$L_i^{(c)}(u_i) = \prod_{j=1}^{J_i} \frac{1}{n_{ij}!} \exp(-u_i \mu_{ij}) (u_i \mu_{ij})^{n_{ij}} [u_i^\gamma e^{\alpha' \mathbf{z}_i} h_0(x_i)]^{\delta_i} \exp \left\{ - \int_0^{x_i} u_i^\gamma e^{\alpha' \mathbf{z}_i} h_0(s) ds \right\}.$$

This leads to the marginal likelihood contributions

$$L_i = \int_0^\infty L_i^{(c)}(u) g_\theta(u) du. \quad (3)$$

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

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 3.2, we demonstrate how the score test for dependence between the two processes can be adapted to the case of interval-count data.

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 because 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

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

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

$$\int_{-\infty}^{\infty} L^{(c)}(v) \phi(v) dv \approx \sum_{q=1}^Q \tilde{w}_q L^{(c)}(\tilde{x}_q) \phi(\tilde{x}_q), \quad (4)$$

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 nonnormal 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 (3). Substituting $u = G_{\theta}^{-1}(\Phi(a))$ in the marginal likelihood contributions, we obtain

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

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 (4). The approximate marginal likelihood of the joint frailty model is then given by

$$\prod_{i=1}^m \sum_{q=1}^Q L_i^{(c)}(G_{\theta}^{-1}(\Phi(\tilde{x}_q))) \phi(\tilde{x}_q) \tilde{w}_q. \quad (5)$$

To actually maximize the approximate likelihood (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

$$\lambda_0(t) = \sum_{k=1}^{K^R} \lambda_{0k} \mathbb{1}\{t \in I_k^R\} \quad \text{and} \quad h_0(t) = \sum_{k=1}^{K^D} h_{0k} \mathbb{1}\{t \in I_k^D\}. \quad (6)$$

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 $I_k^R = (t_{k-1}^R, t_k^R]$, $k = 1, \dots, K^R$, and $I_k^D = (t_{k-1}^D, t_k^D]$, $k = 1, \dots, K^D$, 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 2.)

The baseline rate $\lambda_0(t)$ of the recurrence process enters the likelihood (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

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

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, that is, between 8 and 10 intervals, yields satisfactory estimation results (Lawless & Zhan, 1998; Liu & Huang, 2008).

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, $j = 1, \dots, J_i$, $i = 1, \dots, m$; 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). 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 A.1.

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 (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 η . The score test of Balan et al. (2016) is based on the score function for γ under the null hypothesis; that is,

$$U_\gamma(0, \eta) = \frac{\partial}{\partial \gamma} \ell(\gamma, \eta) |_{\gamma=0},$$

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, $(0, \hat{\eta}_0)$, 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

$$K_i(u, t) = u^{N_i^R(t-) + \gamma N_i^D(t-)} \exp \left\{ -u e^{\beta' z_i} \Lambda_0(t) - u^\gamma e^{\alpha' z_i} H_0(t) \right\}, \quad (7)$$

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

$$\begin{aligned} U_\gamma(0, \hat{\eta}_0) &= \sum_{i=1}^m \left[N_i^D(x_i) - e^{\hat{\alpha}' z_i} \hat{H}_0(x_i) \right] \frac{\int_0^\infty \log(u) \hat{K}_i(u, x_i) g_\theta(u) du}{\int_0^\infty \hat{K}_i(u, x_i) g_\theta(u) du} \\ &= \sum_{i=1}^m \widehat{M}_i^D \cdot \widehat{\log(u_i)}. \end{aligned} \quad (8)$$

The $\widehat{M}_i^D$ are estimates of the martingale residuals of the terminal event model $M_i^D = N_i^D(x_i) - \int_0^{x_i} e^{\alpha' z_i} h_0(s) ds$. The $\widehat{\log(u_i)}$ are the posterior estimates of the log-frailty values given the observed data from the recurrent event process: $\widehat{\log(u_i)} = E[\log U_i | O_i(x_i)]$. For gamma distributed frailties u_i with mean one and variance θ , one obtains

$$\widehat{\log(u_i)} = \psi \left(\frac{1}{\hat{\theta}} + N_i^R(x_i) \right) - \log \left(\frac{1}{\hat{\theta}} + e^{\hat{\beta}' z_i} \Lambda_0(x_i) \right), \quad (9)$$

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

Because of the zero-mean constraint of the $\widehat{M}_i^D$, the second line of (8) is proportional to the correlation $r = \text{cor}(\widehat{M}^D, \widehat{\log(u)})$ between the martingale residuals and the estimated log-frailties. Consequently, Balan et al. (2016) based the correlation score test on the test statistic

$$t = r \sqrt{\frac{m-2}{1-r^2}}, \quad (10)$$

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 (8) still holds. We show this result in Section A.2. Therefore, the test statistic t in (10) is still valid. Furthermore, for gamma distributed frailties, the estimates $\widehat{\log(u_i)}$ can still be determined using formula (9). The latter formula involves estimates $\hat{\theta}$ of the frailty variance, the covariate effect $\hat{\beta}$ 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 & 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 (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 $\{X_i, \delta_i, \mathbf{z}_i; i = 1, \dots, m\}$.

4 | SIMULATION STUDY

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 (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 *E. dichotoma*, see Section 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 $t_{ij} = 0.2j$, $j = 0, \dots, 10$, such that we observed individual interval counts in up to 10 intervals of equal length 0.2. We also studied one scenario with $t_{ij} = 0.1j$, $j = 0, \dots, 20$, 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 $t_j^0 = 0.2j$, $j = 0, \dots, 10$, 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 A.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 1 and 2, both

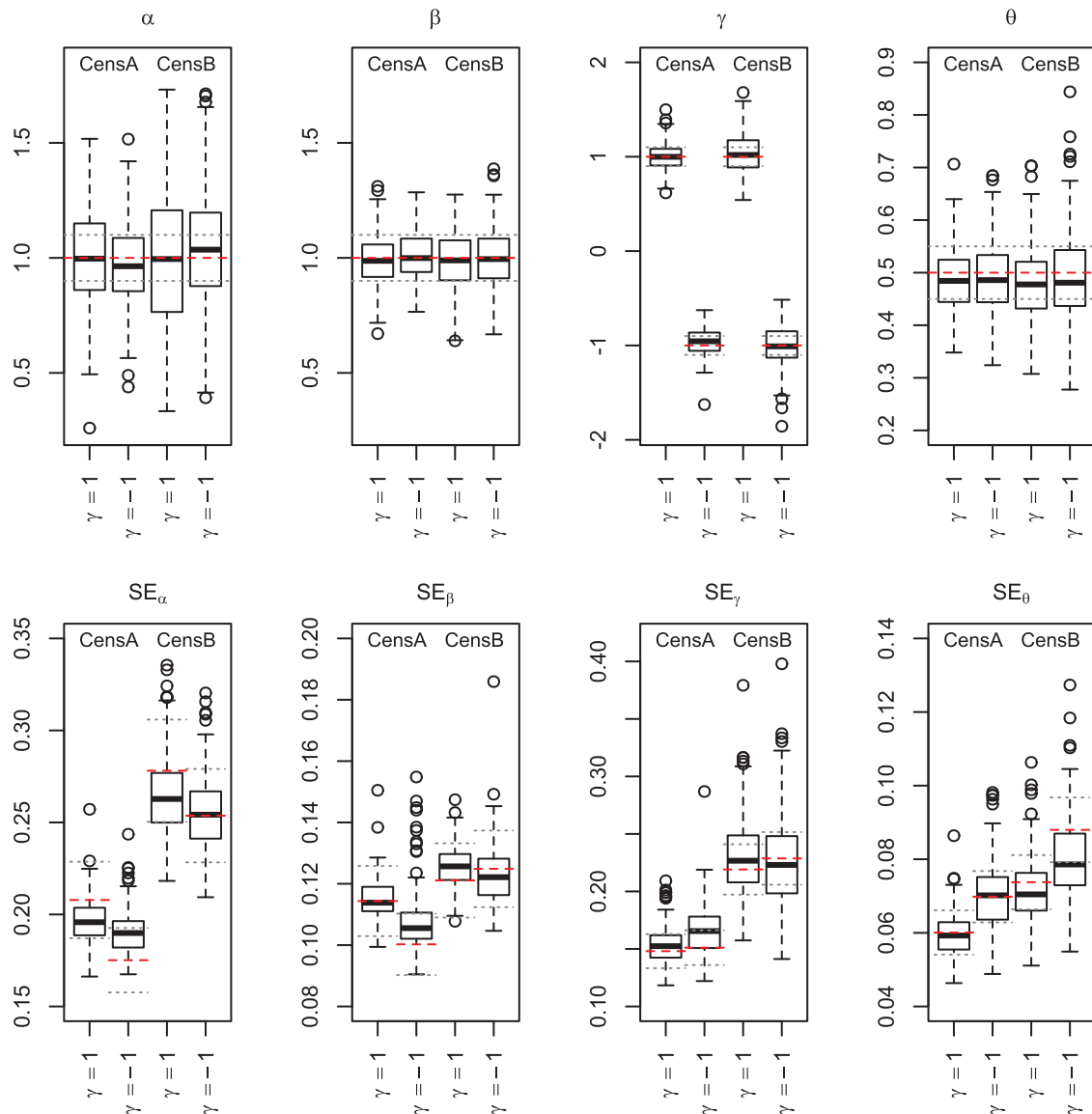


FIGURE 1 Box plots of the parameter estimates (top) and estimated standard errors (bottom) in the joint frailty model for positive ($\gamma = 1$) and negative dependence ($\gamma = -1$) under two schemes of independent censoring (CensA: at end of study $C = 2$, CensB: uniform on $[0, 2]$). Left to right: 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

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 $[0, 2]$).

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 1, we can see that the method performs satisfactorily. Censoring scheme A is considerably milder than scheme B (for an illustration, see Table S2), 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 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 1. The general magnitude of the standard errors is well captured.

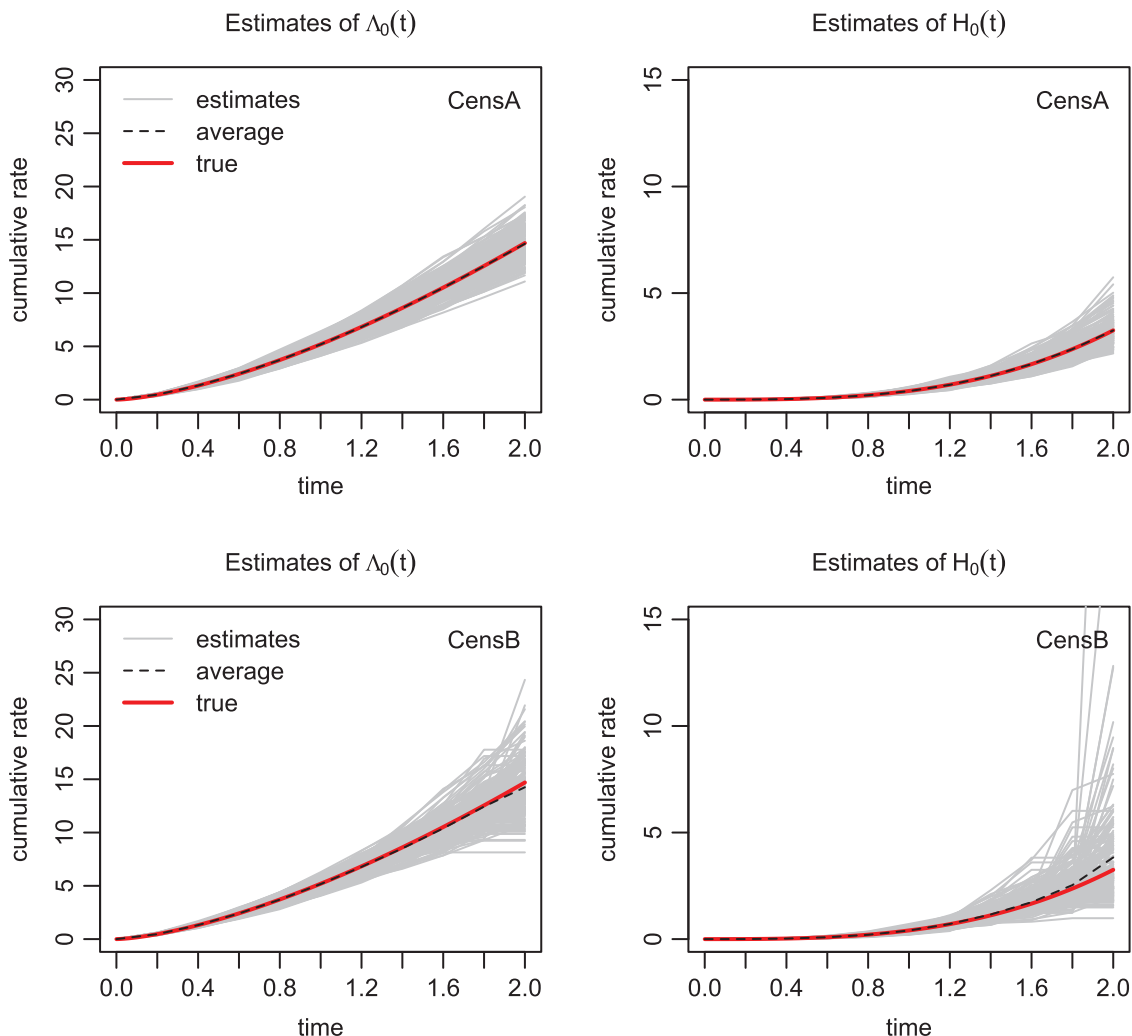


FIGURE 2 Estimates (gray, solid) of the cumulative rate of recurrence (left) and of death (right) based on samples of size $m = 200$, which were generated from a joint frailty model with positive dependence for two schemes of independent censoring. Top: end-of-study censoring at $t = 2$ (CensA), bottom: 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

Figure 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. The Supporting Information illustrates the results of the other simulation settings analogous to Figures 1 and 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 because 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 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 A.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.

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 (2). The covariate effects, values for the variance of the gamma frailties, and baseline rates were the same as in Section 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 $(N_{ij} | u_i)$ follows a Poisson distribution with mean $u_i \tilde{\mu}_{ij}$, where $\tilde{\mu}_{ij} = \int_{I_{ij}} e^{\beta' z_i} \tilde{\lambda}_0(s) ds$. The baseline recurrence rate $\tilde{\lambda}_0(\cdot)$ was modeled as piecewise constant, as in (6), with pieces defined by the cut-points $t_k^R = 0.2k$, $k = 0, \dots, 10$. The estimates $\widehat{\log(u_i)}$ were then determined according to formula (9). Second, we estimated a Cox proportional hazards model from the survival data $\{X_i, \delta_i, z_i; i = 1, \dots, m\}$ using function `coxph()` from package `survival` (Therneau & 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 $\widehat{\log(u_i)}$.

Table 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, 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 S4), 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

TABLE 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 1,000 replications each

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

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.

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 S5).

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 S8. 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 the Supporting Information.

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 | FERTILITY AND MORTALITY IN *ELEUTHERIA DICHOTOMA*

To illustrate the proposed methods, we use data from a biodemographic study on the fertility and mortality of *E. dichotoma* (Daňko et al., 2020), which was briefly introduced in Section 1. *Eleutheria dichotoma* is a marine organism that passes through several life-cycle stages: that is, 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 Daňko et al. (2020).

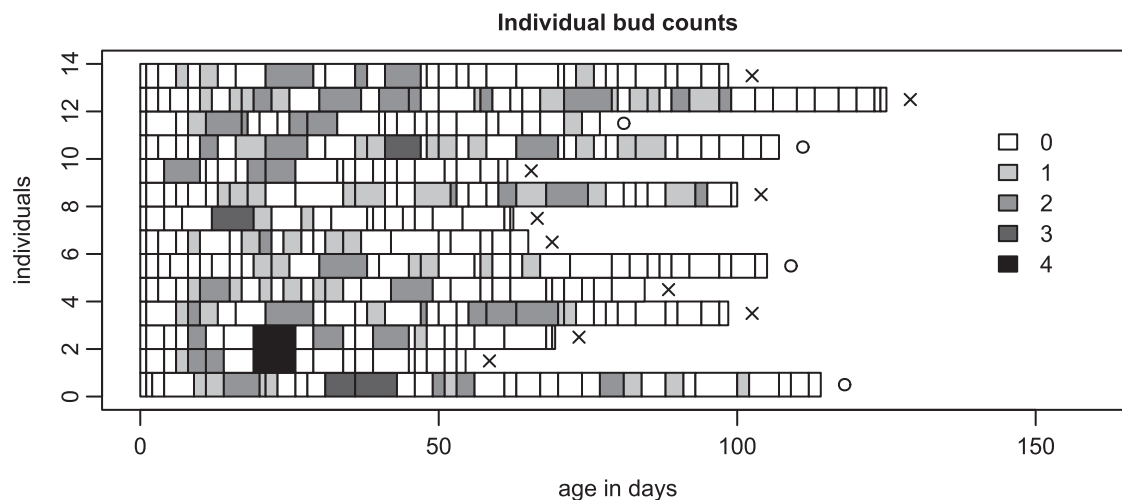


FIGURE 3 Interval counts of budding and survival (dead: cross, censored: circle) for 14 medusae *E. dichotoma*

TABLE 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)	–

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 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 3.2. For fitting 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 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 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, that is, 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 constant

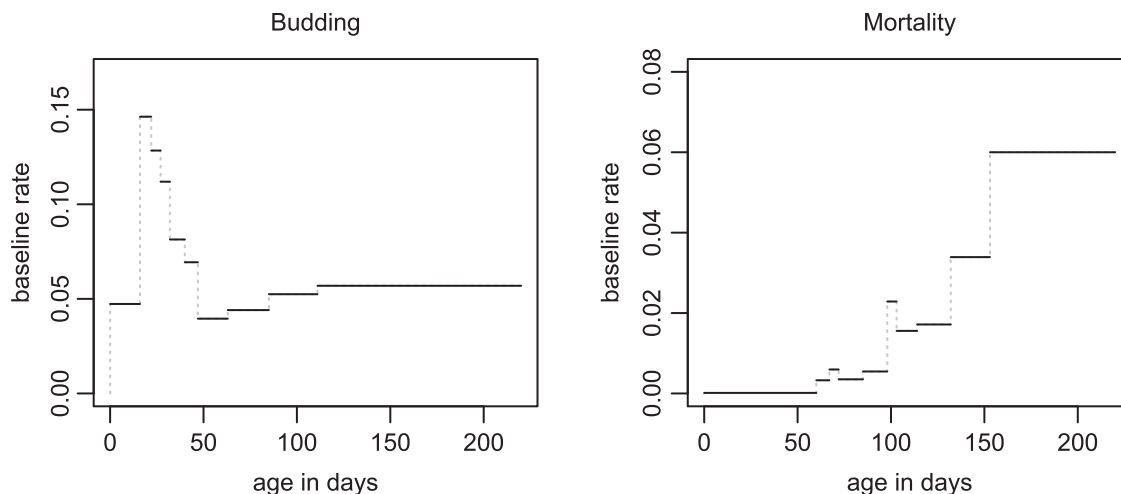


FIGURE 4 Estimated baseline rates of budding (left) and death (right) for the *E. dichotoma* data set

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)), $Q = 30$ quadrature points were used.

The parameter estimates of the joint model are also displayed in Table 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 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 nonzero level at later ages. The death rates increased markedly after about 60 days of age.

To conclude, although the results presented here are in agreement with the findings of Daňko 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) nonindependently, which warrants the joint modeling of the two processes.

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. This finding is interesting from a biological point of view, and it also demonstrates the necessity of allowing positive and negative dependence in a model such as the relatively complex joint frailty model (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 nonzero 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. Although 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 setup, 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.

CONFLICT OF INTEREST

The authors have declared that there is no conflict of interest.

OPEN RESEARCH BADGES

 This article has earned an Open Data badge for making publicly available the digitally-shareable data necessary to reproduce the reported results. The data is available in the [Supporting Information](#) section.

This article has earned an open data badge “**Reproducible Research**” for making publicly available the code necessary to reproduce the reported results. The results reported in this article were reproduced partially for data confidentiality reasons and due to their computational complexity.

ORCID

Marie Böhnstedt  <https://orcid.org/0000-0001-5465-4534>

Hein Putter  <https://orcid.org/0000-0001-5395-1422>

Maciej J. Dańko  <https://orcid.org/0000-0002-7924-9022>

Jutta Gampe  <https://orcid.org/0000-0001-8189-2853>

REFERENCES

- Balan, T.-A., Boonk, S. E., Vermeer, M. H., & Putter, H. (2016). Score test for association between recurrent events and a terminal event. *Statistics in Medicine*, 35, 3037–3048.
- Cook, R. J., & Lawless, J. F. (1997). Marginal analysis of recurrent events and a terminating event. *Statistics in Medicine*, 16, 911–924.
- Cook, R. J., & Lawless, J. F. (2007). *The statistical analysis of recurrent events*. New York: Springer.

- Daňko, A., Schaible, R., & Daňko, M. J. (2020). Salinity effects on survival and reproduction of hydrozoan *Eleutheria dichotoma*. *Estuaries and Coasts*, 43, 360–374.
- Ghosh, D., & Lin, D. Y. (2003). Semiparametric analysis of recurrent events data in the presence of dependent censoring. *Biometrics*, 59, 877–885.
- Gilbert, P., & Varadhan, R. (2019). *numDeriv: Accurate numerical derivatives*. R package version 2016.8-1.1. Retrieved from <http://optimizer.r-forge.r-project.org/>
- Huang, C.-Y., & Wang, M.-C. (2004). Joint modeling and estimation for recurrent event processes and failure time data. *Journal of the American Statistical Association*, 99, 1153–1165.
- Huang, X., Wolfe, R. A., & Hu, C. (2004). A test for informative censoring in clustered survival data. *Statistics in Medicine*, 23, 2089–2107.
- Lancaster, T., & Intrator, O. (1998). Panel data with survival: Hospitalization of HIV-positive patients. *Journal of the American Statistical Association*, 93, 46–53.
- Lawless, J. F., & Zhan, M. (1998). Analysis of interval-grouped recurrent-event data using piecewise constant rate functions. *Canadian Journal of Statistics/La Revue Canadienne de Statistique*, 26, 549–565.
- Liu, L., & Huang, X. (2008). The use of Gaussian quadrature for estimation in frailty proportional hazards models. *Statistics in Medicine*, 27, 2665–2683.
- Liu, L., Wolfe, R. A., & Huang, X. (2004). Shared frailty models for recurrent events and a terminal event. *Biometrics*, 60, 747–756.
- Nelson, K. P., Lipsitz, S. R., Fitzmaurice, G. M., Ibrahim, J., Parzen, M., & Strawderman, R. (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., Mathoulin-Pelissier, S., Jacqmin-Gadda, H., Brouste, V., & Soubeyran, P. (2007). Joint frailty models for recurring events and death using maximum penalized likelihood estimation: Application on cancer events. *Biostatistics*, 8, 708–721.
- Sinha, D., & Maiti, T. (2004). A Bayesian approach for the analysis of panel-count data with dependent termination. *Biometrics*, 60, 34–40.
- Smyth, G. K. (2005). Numerical integration. *Encyclopedia of Biostatistics*, 3088–3095.
- Therneau, T. M., & Grambsch, P. M. (2000). *Modeling survival data: Extending the Cox model*. New York: Springer.
- Zhao, H., Li, Y., & Sun, J. (2013). Analyzing panel count data with a dependent observation process and a terminal event. *Canadian Journal of Statistics/La Revue Canadienne de Statistique*, 41, 174–191.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Böhnstedt M, Putter H, Daňko A, Daňko MJ, Gampe J. Joint modeling of interval counts of recurrent events and death. *Biometrical Journal*. 2021;63:323–340. <https://doi.org/10.1002/bimj.201900367>

APPENDIX A

A.1 | Computational details

We implemented our estimation approach in R (R Core Team, 2019) using function `gauss.quad()` from package `statmod` (Smyth, 2005) 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 & 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.

A.2 | Derivation of the score test

In this section, we show that the score $U_\gamma(\gamma, \eta)$ of the joint frailty model (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 (3) to the marginal likelihood for interval counts

$$L_i = \int_0^\infty \prod_{j=1}^{J_i} \left[\frac{1}{n_{ij}!} \exp \left\{ - \int_{I_{ij}} u e^{\beta' z_i} \lambda_0(s) ds \right\} \left(\int_{I_{ij}} u e^{\beta' z_i} \lambda_0(s) ds \right)^{n_{ij}} \right]$$

$$\begin{aligned}
& \cdot [u^\gamma e^{\alpha' z_i} h_0(x_i)]^{\delta_i} \exp \left\{ - \int_0^{x_i} u^\gamma e^{\alpha' z_i} h_0(s) ds \right\} g_\theta(u) du \\
&= \prod_{j=1}^{J_i} \left[\frac{1}{n_{ij}!} \left(\int_{I_{ij}} e^{\beta' z_i} \lambda_0(s) ds \right)^{n_{ij}} \right] [e^{\alpha' z_i} h_0(x_i)]^{\delta_i} \\
& \quad \cdot \int_0^\infty \left[\prod_{j=1}^{J_i} u^{n_{ij}} \right] \cdot u^{\gamma \delta_i} \exp \left\{ -u \sum_{j=1}^{J_i} \int_{I_{ij}} e^{\beta' z_i} \lambda_0(s) ds \right\} \exp \left\{ - \int_0^{x_i} u^\gamma e^{\alpha' z_i} h_0(s) ds \right\} g_\theta(u) du \\
&= \prod_{j=1}^{J_i} \left[\frac{1}{n_{ij}!} \mu_{ij}^{n_{ij}} \right] [e^{\alpha' z_i} h_0(x_i)]^{\delta_i} \\
& \quad \cdot \int_0^\infty u^{n_i + \gamma \delta_i} \exp \left\{ -u e^{\beta' z_i} \Lambda_0(x_i) \right\} \exp \left\{ -u^\gamma e^{\alpha' z_i} H_0(x_i) \right\} g_\theta(u) du \\
&= \prod_{j=1}^{J_i} \left[\frac{1}{n_{ij}!} \mu_{ij}^{n_{ij}} \right] [e^{\alpha' z_i} h_0(x_i)]^{\delta_i} \cdot \int_0^\infty K_i(u, x_i) g_\theta(u) du,
\end{aligned}$$

with $n_i = \sum_{j=1}^{J_i} n_{ij}$ and $K_i(u, t)$ as defined in (7).

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

$$\sum_{i=1}^m \sum_{j=1}^{J_i} [-\log(n_{ij}!) + n_{ij} \log(\mu_{ij})] + \delta_i [\alpha' z_i + \log(h_0(x_i))] + \log \int_0^\infty K_i(u, x_i) g_\theta(u) du,$$

and the score with respect to γ reads

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

As

$$\begin{aligned}
\frac{\partial}{\partial \gamma} K_i(u, t) &= u^{N_i^R(t-) + \gamma N_i^D(t-)} \cdot N_i^D(t-) \cdot \log(u) \cdot \exp \left\{ -u e^{\beta' z_i} \Lambda_0(t) - u^\gamma e^{\alpha' z_i} H_0(t) \right\} \\
&\quad - u^{N_i^R(t-) + \gamma N_i^D(t-)} \cdot \exp \left\{ -u e^{\beta' z_i} \Lambda_0(t) - u^\gamma e^{\alpha' z_i} H_0(t) \right\} \\
&\quad \cdot u^\gamma \cdot e^{\alpha' z_i} H_0(t) \cdot \log(u) \\
&= K_i(u, t) \left[N_i^D(t-) \cdot \log(u) - u^\gamma \cdot e^{\alpha' z_i} H_0(t) \cdot \log(u) \right],
\end{aligned}$$

we have

$$U_\gamma(\gamma, \eta) = \sum_{i=1}^m \frac{\int_0^\infty \left[N_i^D(x_i) - u^\gamma \cdot e^{\alpha' z_i} H_0(x_i) \right] \log(u) K_i(u, x_i) g_\theta(u) du}{\int_0^\infty K_i(u, x_i) g_\theta(u) du}.$$

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

$$U_\gamma(0, \hat{\eta}_0) = \sum_{i=1}^m \frac{\int_0^\infty \left[N_i^D(x_i) - e^{\hat{\alpha}' z_i} \hat{H}_0(x_i) \right] \log(u) \hat{K}_i(u, x_i) g_\theta(u) du}{\int_0^\infty \hat{K}_i(u, x_i) g_\theta(u) du},$$

which is the first line of (8).

A.3 | Likelihood with fixed observation times

In this section, we study the likelihood of the joint frailty model (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 this purpose, suppose the observation times t_{ij} are given by $t_{ij} = t_j$ for $j = 0, \dots, J_i - 1$, 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 (3)) only through the factor

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

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

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 $\Delta_k^R = (t_k^R - t_{k-1}^R)$, $k = 1, \dots, K^R$, as in (6), we have

$$\mu_{ij} = e^{\beta' z_i} \sum_{k=1}^{K^R} \lambda_{0k} \int_0^\infty \mathbb{1}\{s \in I_{ij} \cap I_k^R\} ds = e^{\beta' z_i} \sum_{k=1}^{K^R} \lambda_{0k} |I_{ij} \cap I_k^R|, \quad (\text{A.2})$$

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 (A.1) as

$$\begin{aligned} \sum_{j=1}^{J_i} \mu_{ij} &= e^{\beta' z_i} \sum_{j=1}^{J_i} \sum_{k=1}^{K^R} \lambda_{0k} |I_{ij} \cap I_k^R| = e^{\beta' z_i} \sum_{k=1}^{K^R} \lambda_{0k} \sum_{j=1}^{J_i} |I_{ij} \cap I_k^R| = e^{\beta' z_i} \sum_{k=1}^{K^R} \lambda_{0k} \left| \bigcup_{j=1}^{J_i} I_{ij} \cap I_k^R \right| \\ &= e^{\beta' z_i} \left\{ \sum_{k=1}^{K_i-1} \lambda_{0k} \Delta_k^R + \lambda_{0K_i} (X_i - t_{K_i-1}^R) \right\}, \end{aligned} \quad (\text{A.3})$$

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

Using (A.2), the last term in (A.1) equals

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

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} , $j = 1, \dots, J_i$, is a subset of exactly one I_k^R , thus,

$$\begin{aligned} \prod_{j=1}^{J_i} \left(\sum_{k=1}^{K^R} \lambda_{0k} |I_{ij} \cap I_k^R| \right)^{n_{ij}} &= \prod_{j=1}^{J_i} \prod_{k=1}^{K^R} (\lambda_{0k} |I_{ij} \cap I_k^R|)^{n_{ij} \cdot \mathbb{1}\{I_{ij} \subset I_k^R\}} \\ &= \prod_{k=1}^{K^R} \prod_{j=1}^{J_i} \lambda_{0k}^{n_{ij} \cdot \mathbb{1}\{I_{ij} \subset I_k^R\}} |I_{ij} \cap I_k^R|^{n_{ij} \cdot \mathbb{1}\{I_{ij} \subset I_k^R\}} \\ &= \left(\prod_{k=1}^{K^R} \lambda_{0k}^{\sum_{j=1}^{J_i} n_{ij} \cdot \mathbb{1}\{I_{ij} \subset I_k^R\}} \right) \left(\prod_{k=1}^{K^R} \prod_{j=1}^{J_i} (|I_{ij} \cap I_k^R|)^{n_{ij} \cdot \mathbb{1}\{I_{ij} \subset I_k^R\}} \right). \end{aligned} \quad (\text{A.4})$$

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 (A.4) 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 (A.3) and (A.4) into (A.1) yields

$$\begin{aligned} & \left(\prod_{j=1}^{J_i} \frac{1}{n_{ij}!} \right) \cdot u_i^{n_i} \cdot \exp \left\{ -u_i e^{\beta' z_i} \left(\sum_{k=1}^{K_i-1} \lambda_{0k} \Delta_k^R + \lambda_{0K_i} (X_i - t_{K_i-1}^R) \right) \right\} \\ & \cdot (e^{\beta' z_i})^{n_i} \prod_{k=1}^{K^R} \lambda_{0k}^{\tilde{n}_{ik}} \prod_{k=1}^{K^R} \prod_{j=1}^{J_i} (|I_{ij} \cap I_k^R|)^{n_{ij} \cdot \mathbb{1}\{I_{ij} \subset I_k^R\}} \\ & \propto (u_i e^{\beta' z_i})^{n_i} \exp \left\{ -u_i e^{\beta' z_i} \left(\sum_{k=1}^{K_i-1} \lambda_{0k} \Delta_k^R + \lambda_{0K_i} (X_i - t_{K_i-1}^R) \right) \right\} \left(\prod_{k=1}^{K^R} \lambda_{0k}^{\tilde{n}_{ik}} \right). \end{aligned}$$

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.