Statistica Sinica **34** (2024), 67-86
doi:https://doi.org/10.5705/ss.202021.0411

GENERALIZED ODDS RATE FRAILTY MODELS FOR CURRENT STATUS DATA WITH INFORMATIVE CENSORING

Yang Xu[1], Shishun Zhao[1], Tao Hu*[2] and Jianguo Sun[3]

[1]*Jilin University,* [2]*Capital Normal University and* [3]*University of Missouri*

Abstract: Current-status data occur in many areas, and the analysis of such data attracted much attention. In this study, we consider a regression analysis of current-status data in the presence of informative censoring, for which most existing methods either apply only to limited situations or are computationally unstable. Here, we propose a new sieve maximum likelihood estimation procedure under the class of semiparametric generalized odds rate frailty models. The proposed method uses the latent variable to describe the informative censoring or relationship between the failure time of interest and the censoring time. We develop a novel expectation-maximization algorithm for determining the proposed estimators, and establish their asymptotic consistency and normality. The results of a simulation study show that the proposed method performs well in practical situations. In addition, we demonstrate the proposed method by applying it to a set of real data arising from a tumorigenicity experiment.

Key words and phrases: EM algorithm, generalized odds rate frailty models, informative censoring, sieve approach.

1. Introduction

Current-status data occur in many areas, including clinical studies, epidemiology studies, and sociological surveys. By current-status data, we usually mean that in a failure-time study, each study subject is observed only once and, thus, the failure time of interest is known only to be either smaller or larger than the observation time, rather than being observed exactly. In other words, each observation is either left or right censored (Sun (2006); Jewell and Emerson (2013)). In this study, we examine a regression analysis of such data in the presence of dependent or informative censoring, meaning that the failure time of interest and the observation time may be correlated.

An area that often produces current-status data with informative censoring is that of tumorigenicity experiments about the time to tumor onset. In such experiments, it is usually the case that only the death or sacrifice time of an animal is observed, and one knows only the presence or absence of a tumor at the

*Corresponding author

observation time. Thus, only current-status data are available on the tumor onset time. It is well known that if a tumor is between lethal and nonlethal, as is often the case, then the tumor onset time and the death time are correlated. We discuss an example of this in Section 6. Although many methods have been proposed for the analysis of tumorigenicity experiments, most of the existing methods are parametric approaches (Dinse and Lagakos (1983); Lagakos and Louis (1988)).

The analysis of current-status data has recently attracted a great deal of attention, particularly regression analysis of such data with independent or non-informative censoring. For example, Huang (1996), Rossini and Tsiatis (1996), and Lin, Oakes and Ying (1998) discuss the problem under the proportional hazards (PH) model, proportional odds (PO) model, and additive hazards model, respectively. Chen, Tong and Sun (2009) also consider the problem, but for multivariate situations. Note that, as pointed out by Shiboski (1998) and others, ignoring informative censoring could yield biased or misleading results or conclusions.

Two types of semiparametric methods have been proposed for regression analyses of current-status data with informative censoring. The first is the copula model-based methods, which use a copula function to describe the relationship between the failure time of interest and the observation time. The second is the frailty-based approaches. Examples of the former include the methods of Ma, Hu and Sun (2015) and Zhao et al. (2015), who examine situations in which the failure time of interest marginally follows the PH model or the additive hazards model, respectively. Du, Hu and Sun (2019), Xu et al. (2019), and Xu, Zhao and Sun (2020) developed similar estimation procedures for cases in which the failure time of interest marginally follows a class of generalized probit models, a class of linear transformation models, and the accelerated failure time model, respectively. Note that, as pointed out by Ma, Hu and Sun (2015) and others, a drawback of the copula model-based approach is that it is usually difficult or impossible to estimate the copula function and association parameter without imposing some strong assumptions.

Frailty-based methods employ latent variables to describe the association between the failure time of interest and the observation time. Among others, Zhang, Sun and Sun (2005) and Li et al. (2017) proposed such methods under the assumption that the failure time of interest follows the additive hazards frailty model and the PH frailty model, respectively. It is well known that the two models may not provide an appropriate fit to the data, because they are individual models. As a result, Chen et al. (2012) investigated using the transformation frailty model with a piecewise constant baseline hazard function, and developed an expectation-maximization (EM) algorithm. However, their EM algorithm can be inefficient and needs a large computational effort.

The class of generalized odds rate (GOR) models has recently attracted attention, and includes the PH and PO models as special cases. For example,

Scharfstein, Tsiatis and Gilbert (1998) and Banerjee et al. (2007) discuss fitting the GOR model to right-censored failure time data. In addition, Zhou, Zhang and Lu (2017) proposed an EM algorithm for the maximum likelihood estimation of the GOR model based on interval-censored failure time data, meaning that the failure time of interest is observed only to belong to some intervals (Sun (2006)). However, all existing methods apply only to the case of independent censoring. Therefore, we consider a class of GOR frailty models for a regression analysis of current-status data with informative censoring, and propose a sieve maximum likelihood estimation procedure (Shen and Wong (1994); Shen (1997)). In particular, we use gamma-Poisson latent variables to develop a novel and efficient EM algorithm.

The remainder of the paper is organized as follows. We first introduce some notation and the GOR frailty model, and then describe the resulting likelihood function in Section 2. The proposed estimation approach is presented in Section 3, along with the development of the novel EM algorithm for the implementation of the proposed estimators. The method uses I-spline functions to approximate the unknown functions involved, and in Section 4, we establish the asymptotic properties of the proposed estimators. In Section 5, we present results from an extensive simulation study conducted to assess the finite-sample performance of the proposed approach that show that it works well in practice. We apply the proposed method to data from a tumorigenicity experiment in Section 6, and conclude the paper in Section 7.

2. Assumptions, Models, and the Likelihood Function

Consider a failure time study consisting of n independent subjects. For subject i, let T_i denote the failure time of interest and X_i be the corresponding p-dimensional vector of covariates. Furthermore, for subject i, let C_i and C_i^c denote two times related to the observation on T_i, where C_i may be related to T_i and C_i^c is independent of T_i. Suppose that each subject is observed only at time $\tilde{C}_i = \min(C_i, C_i^c)$. That is, T_i is either left or right censored, and we have current-status data T_i' only. In the tumorigenicity experiment example, C_i represents the time of death and C_i^c is the sacrifice time. Define $\Delta_i = I(\tilde{C}_i = C_i)$, the administrative censoring indicator, and $\delta_i = I(T_i \leq \tilde{C}_i)$, the observed censoring indicator. Then, the observed data have the form $\boldsymbol{O} = \{\boldsymbol{O_i} = (\tilde{C}_i, \Delta_i, \delta_i, X_i), i = 1, \ldots, n\}$.

To describe the covariate effects and the relationship between T_i and C_i, suppose that there exists a latent variable b_i with mean one and variance η. Furthermore, suppose that given X_i and the random effect b_i, T_i follows the GOR frailty model with the cumulative hazard function given by

$$\Lambda_T\left(t|X_i, b_i\right) = G_r\left\{\Lambda_1(t)\exp\left(X_i'\boldsymbol{\beta}_1\right)b_i\right\}, \tag{2.1}$$

where $\Lambda_1(t)$ denotes an unknown baseline cumulative hazard function, $\boldsymbol{\beta}_1$ is a p-dimensional vector of regression parameters, and $G_r(\cdot)$ is a prespecified increasing transformation function indexed by a nonnegative argument r. It is easy to see that this model includes many commonly used models as special cases. For example, by setting $G_0(x) = x$ and $G_r(x) = \log(1 + rx)/r$, with $r = 1$, model (2.1) gives the PH frailty model and the PO frailty model, respectively. Under model (2.1) with $G_r(x) = \log(1 + rx)/r$, the survival function of T_i given X_i and b_i can be written as

$$S_T(t) = \begin{cases} \exp\left\{-\Lambda_1(t)e^{X_i'\boldsymbol{\beta}_1}b_i\right\} & r = 0 \\ \left\{1 + r\Lambda_1(t)e^{X_i'\boldsymbol{\beta}_1}b_i\right\}^{-1/r} & r > 0 \end{cases}.$$

In practice, covariates may also affect C_i. To describe this, assume that given X_i and b_i, the cumulative hazard function of C_i has the form

$$\Lambda_C(t|X_i, b_i) = \Lambda_2(t)\exp(X_i'\boldsymbol{\beta}_2)\, b_i, \tag{2.2}$$

where $\Lambda_2(t)$ and $\boldsymbol{\beta}_2$ are defined as $\Lambda_1(t)$ and $\boldsymbol{\beta}_1$, respectively. That is, C_i follows the PH frailty model. Let $S_C(t) = \exp\{-\Lambda_2(t)\exp(X_i'\boldsymbol{\beta}_2)\, b_i\}$ and $f_C(t) = d\Lambda_2(t)\exp(X_i'\boldsymbol{\beta}_2)\, b_i S_C(t)$ be the survival and density functions, respectively, of C_i given X_i and b_i, and assume that T_i and C_i are independent given b_i. Then, the likelihood function of $\boldsymbol{\theta} = (\boldsymbol{\beta}_1', \boldsymbol{\beta}_2', \eta, \Lambda_1, \Lambda_2)'$ can be written as

$$\begin{aligned} L_n(\boldsymbol{\theta}|\boldsymbol{O}) = \prod_{i=1}^{n} & \left[\int_0^\infty \left\{1 - S_T(\tilde{C}_i)\right\} f_C(\tilde{C}_i) f(b_i;\eta)\, \mathrm{d}b_i\right]^{\delta_i\Delta_i} \\ & \times \left\{\int_0^\infty S_T(\tilde{C}_i) f_C(\tilde{C}_i) f(b_i;\eta)\, \mathrm{d}b_i\right\}^{(1-\delta_i)\Delta_i} \\ & \times \left[\int_0^\infty \left\{1 - S_T(\tilde{C}_i)\right\} S_C(\tilde{C}_i) f(b_i;\eta)\, \mathrm{d}b_i\right]^{\delta_i(1-\Delta_i)} \\ & \times \left\{\int_0^\infty S_T(\tilde{C}_i) S_C(\tilde{C}_i) f(b_i;\eta)\, \mathrm{d}b_i\right\}^{(1-\delta_i)(1-\Delta_i)}, \end{aligned} \tag{2.3}$$

where $f(\cdot;\eta)$ denotes the density function of b_i', which is assumed to be known up to η.

To estimate $\boldsymbol{\theta}$, one natural method is to directly maximize the aforementioned likelihood function. However, this may be difficult and unreliable because of the complex structure of the likelihood function. To address these problems, we first introduce a sieve approximation, and then develop an efficient EM algorithm by incorporating gamma-Poisson latent variables in the maximization process.

3. Sieve Maximum Likelihood Estimation

In this section, we estimate the regression parameters $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$, among others. As discussed earlier, it may not be easy to deal with unknown functions $\Lambda_1(t)$ and $\Lambda_2(t)$ based on the likelihood function $L_n(\boldsymbol{\theta}|\boldsymbol{O})$. Here, following Ramsay (1988) and others, we employ a sieve approach to approximate $\Lambda_1(t)$ by using I-splines. One can use the same method for the function $\Lambda_2(t)$, but it is usually much easier to estimate it directly. Let $\boldsymbol{\Theta}$ denote the parameter space of $\boldsymbol{\theta}$, and define the sieve space

$$\boldsymbol{\Theta}_n = \left\{ \boldsymbol{\theta}_n = \left((\boldsymbol{\beta}_1', \boldsymbol{\beta}_2', \eta)', \Lambda_{1n}, \Lambda_2 \right) \right\} = \mathcal{B} \otimes \mathcal{M}_n^1 \otimes \mathcal{M}^2,$$

where $\mathcal{B}$ is compact subset of R^{2p+1},

$$\mathcal{M}_n^1 = \left\{ \Lambda_{1n}(t) = \sum_{l=1}^{K_n} \gamma_l I_l(t), M_1 \geq \gamma_l \geq 0,\ l = 1, \ldots, K_n \right\},$$
$$\mathcal{M}^2 = \left\{ \Lambda_2(t) : \frac{1}{M_2} \leq \Lambda_2(t) \leq M_2 \right\},$$

with M_1 and M_2 being some positive constants. In the above, I_l' are nondecreasing integrated spline basis functions, each ranging from zero to one, and $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_{K_n})$ are nonnegative coefficients that ensure the monotonicity of $\Lambda_{1n}(t)$. The number of the spline basis functions, K_n, is equal to the number of interior knots plus the degree.

As mentioned above, even with the approximation, maximizing the likelihood function is still not easy, owing to its complex form and the involvement of the latent variables b_i'. To deal with this, a typical approach is to develop a standard EM algorithm, as in Chen et al. (2012), but it may not be stable or efficient. To overcome these issues, following McMahan, Wang and Tebbs (2013) and Li et al. (2017), we use the gamma-Poisson data augmentation, which greatly simplifies the computional burden and yields stable and efficient estimators. Before we present the proposed EM algorithm, note that if b_i' were known, we would have the following pseudo-complete data likelihood function:

$$L(\boldsymbol{\theta}|\boldsymbol{O}, \boldsymbol{b}) = \prod_{i=1}^{n} \{1 - S_T(\tilde{C}_i)\}^{\delta_i} S_T(\tilde{C}_i)^{1-\delta_i} \{ d\Lambda_2(\tilde{C}_i) \exp\left(X_i' \boldsymbol{\beta}_2\right) b_i \}^{\Delta_i} S_C(\tilde{C}_i) f\left(b_i; \eta\right),$$

conditional on $\boldsymbol{b} = (b_1, \ldots, b_n)$.

Let ϕ be a random variable following the gamma distribution $\Gamma(1/r, r)$, with $r > 0$. Then, the survival function of T_i can be rewritten as

$$S_T\left(t\right) = \left\{ 1 + r\Lambda_1(t) e^{X_i{}' \boldsymbol{\beta}_1} b_i \right\}^{-1/r} = \int_0^\infty \exp\{ -\phi \Lambda_1(t) e^{X_i{}' \boldsymbol{\beta}_1} b_i \} f(\phi; r) d\phi.$$

Note that $\lim_{r\to 0}\{1 + r\Lambda_1(t)e^{X_i'\boldsymbol{\beta}_1}b_i\}^{-1/r} = \exp\{-\Lambda_1(t)e^{X_i'\boldsymbol{\beta}_1}b_i\}$, meaning that ϕ degenerates to a constant one for $r = 0$. Thus, we can consider the new pseudo-complete data likelihood function

$$\begin{aligned} L_1(\boldsymbol{\theta}|\boldsymbol{O}, \mathbf{b}, \boldsymbol{\phi}, r) = \prod_{i=1}^{n} &\left[1 - \exp\left\{-\phi_i\Lambda_1(\tilde{C}_i)\exp\left(X_i'\boldsymbol{\beta}_1\right)b_i\right\}\right]^{\delta_i} \\ &\times \left[\exp\left\{-\phi_i\Lambda_1(\tilde{C}_i)\exp\left(X_i'\boldsymbol{\beta}_1\right)b_i\right\}\right]^{1-\delta_i} \\ &\times \left\{d\Lambda_2(\tilde{C}_i)\exp\left(X_i'\boldsymbol{\beta}_2\right)b_i\right\}^{\Delta_i} \\ &\times \exp\left\{-\Lambda_2(\tilde{C}_i)\exp\left(X_i'\boldsymbol{\beta}_2\right)b_i\right\} f\left(b_i;\eta\right) f(\phi_i;r), \end{aligned}$$

based on $\boldsymbol{b}$ and $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_n)$, where $\phi_i \overset{\text{i.i.d.}}{\sim} \Gamma(1/r, r)$, for $i = 1, 2, \ldots, n$.

Let $\{Z_i, i = 1, \ldots, n\}$ be a set of independent Poisson random variables with means $\phi_i\Lambda_1(\tilde{C}_i)\exp\left(X_i'\boldsymbol{\beta}_1\right)b_i$, given ϕ_i and b_i. It is easy to show that the likelihood function $L_1(\boldsymbol{\theta}|\boldsymbol{O}, \mathbf{b}, \boldsymbol{\phi}, r)$ can be equivalently written as

$$\begin{aligned} L_1(\boldsymbol{\theta}|\boldsymbol{O}, \mathbf{b}, \boldsymbol{\phi}, r) = \prod_{i=1}^{n} &P(Z_i > 0)^{\delta_i}P(Z_i = 0)^{1-\delta_i}\left\{d\Lambda_2(\tilde{C}_i)\exp\left(X_i'\boldsymbol{\beta}_2\right)b_i\right\}^{\Delta_i} \\ &\times \exp\left\{-\Lambda_2(\tilde{C}_i)\exp\left(X_i'\boldsymbol{\beta}_2\right)b_i\right\}. \end{aligned}$$

This suggests that we should consider the pseudo-complete data likelihood function

$$\begin{aligned} L_2(\boldsymbol{\theta}|\boldsymbol{O}, \mathbf{b}, \boldsymbol{\phi}, \boldsymbol{Z}, r) = \prod_{i=1}^{n} &\left[\frac{1}{Z_i!}\left\{\phi_i\Lambda_1(\tilde{C}_i)\exp(X_i'\boldsymbol{\beta}_1)b_i\right\}^{Z_i} e^{-\phi_i\Lambda_1(\tilde{C}_i)\exp(X_i'\boldsymbol{\beta}_1)b_i}\right] \\ &\times \left\{d\Lambda_2(\tilde{C}_i)\exp\left(X_i'\boldsymbol{\beta}_2\right)b_i\right\}^{\Delta_i}\exp\left\{-\Lambda_2(\tilde{C}_i)\exp\left(X_i'\boldsymbol{\beta}_2\right)b_i\right\} \\ &f\left(b_i;\eta\right) f(\phi_i;r), \end{aligned}$$

based on $\boldsymbol{b}, \boldsymbol{\phi}$, and $\boldsymbol{Z} = (Z_1, \ldots, Z_n)$, and with the constraint $\delta_i I\left(Z_i > 0\right) + (1 - \delta_i)I\left(Z_i = 0\right) = 1$, for $i = 1, \ldots, n$.

Finally note that we can decompose Z_i as the sum of K_n independent latent variables Z_{il}', with Z_{il} following a Poisson distribution with mean $\phi_i\gamma_l I_l(\tilde{C}_i)$ $\exp\left(X_i'\boldsymbol{\beta}_1\right)b_i$, conditional on ϕ_i and b_i, for $l = 1, \ldots, K_n$. Thus, we propose to basing our EM algorithm on the pseudo-complete data likelihood function

$$\begin{aligned} L_c(\boldsymbol{\theta}) = \prod_{i=1}^{n} &\left[\prod_{l=1}^{K_n}\frac{1}{Z_{il}!}\left\{\phi_i\gamma_l I_l(\tilde{C}_i)\exp(X_i'\boldsymbol{\beta}_1)b_i\right\}^{Z_{il}} e^{-\phi_i\gamma_l I_l(\tilde{C}_i)\exp(X_i'\boldsymbol{\beta}_1)b_i}\right] \\ &\times\left\{d\Lambda_2(\tilde{C}_i)\exp\left(X_i'\boldsymbol{\beta}_2\right)b_i\right\}^{\Delta_i}\exp\left\{-\Lambda_2(\tilde{C}_i)\exp\left(X_i'\boldsymbol{\beta}_2\right)b_i\right\} f\left(b_i;\eta\right) f(\phi_i;r), \end{aligned}$$

based on $\{\boldsymbol{b}, \boldsymbol{\phi}, \boldsymbol{Z}_l = \left(Z_{1l}, \ldots, Z_{nl}\right), l = 1, \ldots, K_n\}$, and with the constraint

$\delta_i I(\sum_{l=1}^{K_n} Z_{il} > 0) + (1-\delta_i) I(\sum_{l=1}^{K_n} Z_{il} = 0) = 1$, for $i = 1, \ldots, n$. Note that by integrating over $\boldsymbol{b}, \boldsymbol{\phi}$, and $\boldsymbol{Z}_l'$, $L_c(\boldsymbol{\theta})$ reduces to the observed data likelihood function given in (2.3).

Let $l_c(\boldsymbol{\theta}) = \log L_c(\boldsymbol{\theta})$ and define $\boldsymbol{\theta} = (\boldsymbol{\theta}_1', \boldsymbol{\theta}_2', \eta)'$, with $\boldsymbol{\theta}_1 = (\boldsymbol{\beta}_1', \Lambda_{1n})'$ and $\boldsymbol{\theta}_2 = (\boldsymbol{\beta}_2', \Lambda_2)'$. In the E-step of the EM algorithm, we need to determine the conditional expectation of $l_c(\boldsymbol{\theta})$ with respect to all latent variables. Specifically, at the mth iteration, the conditional expectation of $l_c(\boldsymbol{\theta})$ can be expressed as the summation of four parts

$$Q\left(\boldsymbol{\theta}; \boldsymbol{\theta}^{(m)}\right) = Q_1\left(\boldsymbol{\theta}_1; \boldsymbol{\theta}^{(m)}\right) + Q_2\left(\boldsymbol{\theta}_2; \boldsymbol{\theta}^{(m)}\right) + Q_3\left(\eta; \boldsymbol{\theta}^{(m)}\right) + Q_4\left(\boldsymbol{\theta}^{(m)}\right),$$

where

$$Q_1\left(\boldsymbol{\theta}_1; \boldsymbol{\theta}^{(m)}\right) = \sum_{i=1}^{n}\sum_{l=1}^{K_n} \left\{\log(\gamma_l) + X_i'\boldsymbol{\beta}_1\right\} \hat{E}(Z_{il}) - \gamma_l I_l(\tilde{C}_i) \exp(X_i'\boldsymbol{\beta}_1)\, \hat{E}(\phi_i b_i),$$

$$Q_2\left(\boldsymbol{\theta}_2; \boldsymbol{\theta}^{(m)}\right) = \sum_{i=1}^{n} \Delta_i (\log d\Lambda_2(\tilde{C}_i) + X_i'\boldsymbol{\beta}_2) - \Lambda_2(\tilde{C}_i) \exp(X_i'\boldsymbol{\beta}_2)\, \hat{E}(b_i),$$

$$Q_3\left(\eta; \boldsymbol{\theta}^{(m)}\right) = n\eta^{-1}\log\eta^{-1} - n\log\Gamma\left(\eta^{-1}\right) + \eta^{-1}\sum_{i=1}^{n}[\hat{E}\{\log(b_i)\} - \hat{E}(b_i)],$$

and $Q_4(\boldsymbol{\theta}^{(m)})$ denotes a function of $\boldsymbol{\theta}^{(m)}$ free of $\boldsymbol{\theta}$. In the above, $\hat{E}(Z_{il})$, $\hat{E}(\phi_i b_i)$, $\hat{E}(b_i)$, and $\hat{E}(\log b_i)$ denote the conditional expectations with respect to all latent variables.

In the calculation of the conditional expectations above, note that

$$\hat{E}(Z_i) = \Delta_i \Lambda_{1n}(\tilde{C}_i) \exp(X_i'\boldsymbol{\beta}_1) E_{b_i}\left\{\frac{b_i}{1 - S_T(\tilde{C}_i)}\right\}.$$

Using $Z_{il} \mid Z_i \sim \text{Binomial}[Z_i, \phi_i\gamma_l I_l(\tilde{C}_i)\exp(X_i'\boldsymbol{\beta}_1)b_i / \{\phi_i \Lambda_{1n}(\tilde{C}_i)\exp(X_i'\boldsymbol{\beta}_1)b_i\}]$ and applying the iterated rule of expectations, we have that

$$\hat{E}(Z_{il}) = \frac{\gamma_l I_l(\tilde{C}_i)}{\Lambda_{1n}(\tilde{C}_i)}\hat{E}(Z_i) = \Delta_i\gamma_l I_l(\tilde{C}_i)\exp(X_i'\boldsymbol{\beta}_1) E_{b_i}\left\{\frac{b_i}{1 - S_T(\tilde{C}_i)}\right\},$$

and

$$\hat{E}(\phi_i b_i) = E_{b_i}\left[b_i S_T(\tilde{C}_i)^{r(1-\delta_i)}\left\{\frac{1 - S_T(\tilde{C}_i)^{r+1}}{1 - S_T(\tilde{C}_i)}\right\}^{\delta_i}\right].$$

It is clear that none of the above conditional expectations have closed forms; thus, we can approximate them using the Monte Carlo method. Specifically, the conditional expectation of any arbitrary function $h(b_i)$ and sufficiently large L can be calculated using

$$E_{b_i}\{h\left(b_i\right) \mid O_i, \hat{\boldsymbol{\theta}}^{(m)}\} = \frac{\int_{-\infty}^{\infty} h\left(b_i\right) \Psi_i(b_i; \hat{\boldsymbol{\theta}}^{(m)}) f(b_i; \hat{\eta}^{(m)}) db_i}{\int_{-\infty}^{\infty} \Psi_i(b_i; \hat{\boldsymbol{\theta}}^{(m)}) f(b_i; \hat{\eta}^{(m)}) db_i} \approx \frac{L^{-1} \sum_{i=1}^{L} h\left(b_{il}\right) \Psi_i(b_{il}; \hat{\boldsymbol{\theta}}^{(m)})}{L^{-1} \sum_{i=1}^{L} \Psi_i(b_{il}; \hat{\boldsymbol{\theta}}^{(m)})},$$

where

$$\Psi_i(b_i; \boldsymbol{\theta}) = \{1 - S_T(\tilde{C}_i)\}^{\delta_i} S_T(\tilde{C}_i)^{1-\delta_i} \{d\Lambda_2(\tilde{C}_i) \exp\left(X_i' \boldsymbol{\beta}_2\right) b_i\}^{\Delta_i} S_C(\tilde{C}_i),$$

and $\{b_{il}, l = 1, \ldots, L\}$ are generated form the density function $f(b_i; \eta^{(m)})$.

In the M-step of the EM algorithm, by setting $\partial Q_1(\boldsymbol{\theta}_1; \boldsymbol{\theta}^{(m)})/\partial \gamma_l = 0$, we calculate γ_l using the following closed-form expression:

$$\gamma_l = \frac{\sum_{i=1}^{n} \hat{E}\left(Z_{il}\right)}{\sum_{i=1}^{n} I_l(\tilde{C}_i) \exp\left(X_i' \boldsymbol{\beta}_1\right) \hat{E}\left(\phi_i b_i\right)}, \quad l = 1, \ldots, K_n. \tag{3.1}$$

By substituting the estimators above into $Q_1(\boldsymbol{\theta}_1; \boldsymbol{\theta}^{(m)})$, we obtain the score functions for $\boldsymbol{\beta}_1$ as

$$\sum_{i=1}^{n} \sum_{l=1}^{K_n} \hat{E}\left(Z_{il}\right) \left\{ X_i - \frac{\sum_{i=1}^{n} I_l(\tilde{C}_i) \exp\left(X_i' \boldsymbol{\beta}_1\right) \hat{E}\left(\phi_i b_i\right) X_i}{\sum_{i=1}^{n} I_l(\tilde{C}_i) \exp\left(X_i' \boldsymbol{\beta}_1\right) \hat{E}\left(\phi_i b_i\right)} \right\} = 0. \tag{3.2}$$

To determine the updated estimators of $\boldsymbol{\beta}_2$ and $\Lambda_2(t)$, by treating $\Lambda_2(t)$ as a piecewise constant function between the uncensored observation times, we have the following score function:

$$\frac{\partial Q_2(\boldsymbol{\theta}_2, \boldsymbol{\theta}^{(m)})}{\partial \boldsymbol{\beta}_2} = \sum_{i=1}^{n} \int_0^{\infty} \left\{ X_i - \bar{X}\left(t, \boldsymbol{\beta}_2\right) \right\} dN_i(t) = 0, \tag{3.3}$$

where $N_i(t) = I(\tilde{C}_i \leq t, \Delta_i = 1)$ and

$$\bar{X}\left(t, \boldsymbol{\beta}_2\right) = \frac{\sum_{i=1}^{n} Y_i(t) \exp\left(X_i' \boldsymbol{\beta}_2\right) \hat{E}\left(b_i\right) X_i}{\sum_{i=1}^{n} Y_i(t) \exp\left(X_i' \boldsymbol{\beta}_2\right) \hat{E}\left(b_i\right)},$$

with $Y_i(t) = I(\tilde{C}_i \geq t)$. Thus, the estimator $\hat{\Lambda}_2^{(m+1)}$ can be updated using the Breslow-type estimator

$$\hat{\Lambda}_2^{(m+1)}(t) = \sum_{i=1}^{n} \int_0^{t} \frac{dN_i(s)}{\sum_{i=1}^{n} Y_i(s) \exp(X_i' \boldsymbol{\beta}_2^{(m+1)}) \hat{E}\left(b_i\right)}. \tag{3.4}$$

Finally, we obtain the updated estimator of η by solving the score equation

$$\frac{\partial Q_3(\eta; \boldsymbol{\theta}^{(m)})}{\partial \eta} = 0.$$

The following EM algorithm combines the above steps.

Step 0. Select an initial estimate of $\boldsymbol{\theta}^{(0)}$, such as $\boldsymbol{\beta}_1^{(0)} = \boldsymbol{\beta}_2^{(0)} = \mathbf{0}, \boldsymbol{\gamma}^{(0)} = \mathbf{1}$, $\eta = 1$, and a prespecified value of r;

Step 1. At the $(m+1)$th iteration, generate the random sample $\{b_{il}, l = 1, \ldots, L, i = 1, \ldots, n\}$ from the density function $f(b_i; \eta^{(m)})$;

Step 2. Calculate the conditional expectations $\hat{E}(Z_{il})$, $\hat{E}(\phi_i b_i)$, $\hat{E}(b_i)$, and $\hat{E}\{\log(f(b_i; \eta))\}$ based on $\boldsymbol{\theta}^{(m)}$;

Step 3. Update $\hat{\boldsymbol{\beta}}_1^{(m+1)}$ based on (3.2) by using the Newton−Raphson method, and then determine $\gamma_l^{(m+1)}$ using (3.1) for $l = 1, \ldots, K_n$;

Step 4. Update $\hat{\boldsymbol{\beta}}_2^{(m+1)}$ based on (3.3) by using the Newton−Raphson method, and then determine $\hat{\Lambda}_2^{(m+1)}$ using (3.4);

Step 5. Calculate $\eta^{(m+1)}$ by solving $\partial Q_3(\eta; \boldsymbol{\theta}^{(m)})/\partial\eta = 0$;

Step 6. Repeat Steps 1-5 until the difference between the eatimates of two consecutive iterations is less than a prespecified constant, such as 0.0001.

To implement this EM algorithm, we need to choose the degree of the splines and the number of interior knots. Here, a popular approach is to arrange the knots so that thay are equally spaced or based on the quartiles, and to set the degree of the splines to two or three. An alternative is to try different values, and then apply Akaike's information criterion (AIC), defined as

$$AIC = -2l_n(\hat{\boldsymbol{\theta}}_n) + 2df_k,$$

to choose the optimal model. In the above, $\hat{\boldsymbol{\theta}}_n = (\hat{\boldsymbol{\beta}}_1', \hat{\boldsymbol{\beta}}_2', \hat{\eta}, \hat{\Lambda}_{1n}, \hat{\Lambda}_2)'$ denotes the estimator, $l_n(\hat{\boldsymbol{\theta}}_n) = \log L_n(\hat{\boldsymbol{\theta}}_n)$, and df_k denotes the number of parameters to be estimated. The same AIC can be used to choose the nonnegative argument r, which we assume to be known, although this may not be true in practice. Specifically, we can perform a grid search by considering a sequence of values for r and choosing the value that gives the smallest AIC value.

4. Asymptotic Properties

In this section, we discuss the asymptotic properties of the proposed estimator. Let $\boldsymbol{\theta}_0 = (\boldsymbol{\beta}_{10}, \boldsymbol{\beta}_{20}, \eta_0, \Lambda_{10}, \Lambda_{20})$ denote the true value of $\boldsymbol{\theta}$, and for two different $\check{\boldsymbol{\theta}}_n$ and $\tilde{\boldsymbol{\theta}}_n \in \boldsymbol{\Theta}_n$, define the norm

$$\begin{aligned} &d(\check{\boldsymbol{\theta}}_n, \tilde{\boldsymbol{\theta}}_n) \\ &= \left(\|\check{\Lambda}_{1n} - \tilde{\Lambda}_{1n}\|_2^2 + \|\check{\Lambda}_2 - \tilde{\Lambda}_2\|_2^2 + \|\check{\boldsymbol{\beta}}_1 - \tilde{\boldsymbol{\beta}}_1\|_E^2 + \|\check{\boldsymbol{\beta}}_2 - \tilde{\boldsymbol{\beta}}_2\|_E^2 + \|\check{\eta} - \tilde{\eta}\|_E^2\right)^{1/2}, \end{aligned}$$

where $\|\cdot\|_2$ and $\|\cdot\|_E$ denote the L^2 and Euclidean norms, respectively. For the asymptotic properties, we need the following regularity conditions.

(C1) The maximum spacing of the knots satisfies $\tilde{\Delta} = \max_{2\le j\le k_n+2} |t_j - t_{j-1}| = O(n^{-\nu})$, with $v \in (0, 0.5)$. Moreover, there exists a constant $M > 0$ such that $\tilde{\Delta}/\tilde{\delta} \le M$ uniformly in n, where $\tilde{\delta} = \min_{2\le j\le k_n+2} |t_j - t_{j-1}|$ and k_n is the number of interior knots.

(C2) The covariates X_i' have a bounded support in R^p.

(C3) The true cumulative hazard functions, $\Lambda_{10}(t)$ and $\Lambda_{20}(t)$, are increasing functions. In addition, the κth derivative of $\Lambda_{10}(t)$ is bounded and continuous, and $\Lambda_{20}(t)$ satisfies $1/M_2 \le \Lambda_{20}(t) \le M_2$.

(C4) Define $\boldsymbol{\vartheta} = (\boldsymbol{\beta}_1', \boldsymbol{\beta}_2', \eta)'$ and

$$L_{(\boldsymbol{\vartheta},y_1,y_2)}(\delta,\Delta,X) = \int_0^\infty \{1 - S_T(\tilde{c};\boldsymbol{\vartheta},y_1)\}^\delta S_T(\tilde{c};\boldsymbol{\vartheta},y_1)^{1-\delta} f_C(\tilde{c};\boldsymbol{\vartheta},y_2)^\Delta \times S_C(\tilde{c};\boldsymbol{\vartheta},y_2)^{1-\Delta} f(b;\eta)\mathrm{d}b.$$

Suppose that there exists t^* in the support of the conditional distribution of $\tilde{C}$ given $X = x$ for which there are $2p+3$ different values (δ,Δ,X), such that if

$$\left(a_1'\frac{\partial}{\partial\boldsymbol{\vartheta}} + a_2\frac{\partial}{\partial y_1} + a_3\frac{\partial}{\partial y_2}\right)\log L_{(\boldsymbol{\vartheta},y_1,y_2)}(\delta,\Delta,X)\Big|_{(\boldsymbol{\vartheta},y_1,y_2)=(\boldsymbol{\vartheta}_0,\Lambda_1(t^*),\Lambda_2(t^*))} = 0,$$

then $a_1 = \mathbf{0}$ and $a_2 = a_3 = 0$, where $(\delta,\Delta) = (0,0),(1,0),(0,1)$, or $(1,1)$, $S_C(t;\boldsymbol{\vartheta},y_2(t)) = \exp\{-y_2(t)\exp(X'\boldsymbol{\beta}_2)\,b\}$, $f_C\{t;\boldsymbol{\vartheta},y_2(t)\} = \mathrm{d}y_2(t)\exp(X'\boldsymbol{\beta}_2)\,bS_C\{t;\vartheta,y_2(t)\}$, and

$$S_T(t;\boldsymbol{\vartheta},y_1(t)) = \begin{cases} \exp\left\{-y_1(t)e^{X'\boldsymbol{\beta}_1}b\right\} & r=0 \\ \left\{1+ry_1(t)e^{X'\boldsymbol{\beta}_1}b\right\}^{-1/r} & r>0 \end{cases}.$$

(C5) The matrix Σ, defined in the Supplementary Material is finite and positive definite.

Note that the aforementioned conditions are mild and usually satisfied in practice. In particular, Condition (**C1**) is the same as Condition 1 of Lu, Zhang and Huang (2007), and is necessary for the construction of the spline sieve space and the consistency of the spline likelihood-based estimators. The bounded support assumption stated in Condition (**C2**) for covariates is commonly used in the literature on current-status data, and is needed for the uniform convergence; see Van der Vaart and Wellner (1996) for a detailed discussion. Condition (**C3**), the smoothness assumption of the cumulative hazard functions, is standard in

the nonparametric smoothing literature. Condition (**C4**) is a sufficient condition for the identifiability of the parameters, and is common in the existing literature (Li, Taylor and Sy (2001); Chang, Wen and Wu (2007)). The following theorems give the asymptotic properties.

Theorem 1. Assume that Conditions (**C1**) – (**C4**) hold. Then, as $n \to \infty$, we have that

$$d(\hat{\boldsymbol{\theta}}_n, \boldsymbol{\theta}_0) = O_p(n^{-(1-v)/2} + n^{-\kappa v}).$$

Theorem 2. Assume that Conditions (**C1**) – (**C5**) hold. Then, as $n \to \infty$, we have that

$$n^{1/2}\boldsymbol{\alpha}' \left((\hat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_{10})', (\hat{\boldsymbol{\beta}}_2 - \boldsymbol{\beta}_{20})', (\hat{\eta} - \eta_0) \right)' + \int_0^{\tau_c} g(t) d \left\{ \hat{\Lambda}_2(t) - \Lambda_{20}(t) \right\}$$
$$\xrightarrow{D} N(0, \Sigma),$$

where $\boldsymbol{\alpha}$ is any $(2p+1)$-dimensional vector with $\|\boldsymbol{\alpha}\|_E \leq 1$, τ_c denotes the longest follow-up time, g is a function with bounded variation on $[0, \tau_c]$, and Σ is the semiparametric efficiency bound defined in the Supplementary Material.

Proofs of these theorems are provided in the Supplementary Material. Note that Theorem 1 states that the proposed estimator is consistent and achieves the optimal convergence rate. Theorem 2 states that the proposed estimators $\hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_2$, and $\hat{\eta}$ are asymptotically efficient. In particular, by setting $g = 0$ in Theorem 2, we have that

$$n^{1/2}\boldsymbol{\alpha}'((\hat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_{10})', (\hat{\boldsymbol{\beta}}_2 - \boldsymbol{\beta}_{20})', (\hat{\eta} - \eta_0))' \xrightarrow{D} N(0, \Sigma_{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \eta}).$$

In orderto perform an inference on $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$, and η, we need to estimate the asymptotic covariance matrix of the corresponding estimators. Because it is difficult to derive a consistent estimator of $\Sigma_{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \eta}$, we suggest using the nonparametric bootstrap method (Efron (1981)), as follows. We first draw new data sets of sample size n, with replacement, from the original observed data $\boldsymbol{O}$ repeatedly Q times, where Q is a prespecified positive integer. The newly resampled data sets are denoted by $\boldsymbol{O}^{(q)}$, for $q = 1, \ldots, Q$. Let $\hat{\boldsymbol{\beta}}_1^{(q)}$, $\hat{\boldsymbol{\beta}}_2^{(q)}$, and $\hat{\eta}^{(q)}$ denote the maximum likelihood estimators of $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$, and η, respectively, based on the bootstrap sample $\boldsymbol{O}^{(q)}$. Then, we can estimate the covariance matrix or variance of $\hat{\boldsymbol{\beta}}_1$, $\hat{\boldsymbol{\beta}}_2$, and $\hat{\eta}$ using the empirical covariance matrix or variance of $\hat{\boldsymbol{\beta}}_1^{(q)}$, $\hat{\boldsymbol{\beta}}_2^{(q)}$, and $\hat{\eta}^{(q)}$, respectively. The results of our simulation study, discussed in the next section, suggest that this method works well for practical situations.

Table 1. Simulation results for regression parameters based on simulated data with the same baseline functions.

r	Par.	Bias	SSE	SEE	CP	Bias	SSE	SEE	CP
$(\beta_{11},\beta_{12},\beta_{21},\beta_{22})=(0,0,0,0)$						$(\beta_{11},\beta_{12},\beta_{21},\beta_{22})=(0.2,0.2,0.2,0.2)$			
0	β_{11}	0.007	0.274	0.280	95.2	0.027	0.278	0.278	95.4
	β_{12}	0.029	0.481	0.497	94.8	0.033	0.496	0.480	94.8
	β_{21}	0.006	0.264	0.262	93.0	0.062	0.258	0.255	93.4
	β_{22}	0.017	0.504	0.506	94.4	0.032	0.477	0.470	95.4
	η	0.005	0.128	0.151	96.4	0.017	0.126	0.116	93.8
0.5	β_{11}	-0.009	0.309	0.320	95.6	0.014	0.329	0.318	93.4
	β_{12}	0.019	0.562	0.565	93.8	0.072	0.560	0.551	94.2
	β_{21}	-0.009	0.259	0.257	94.4	0.044	0.279	0.251	91.6
	β_{22}	-0.014	0.482	0.494	93.0	0.043	0.492	0.469	93.6
	η	-0.018	0.112	0.138	95.4	-0.006	0.122	0.113	92.8
1	β_{11}	0.004	0.347	0.353	94.8	0.013	0.360	0.360	93.4
	β_{12}	-0.005	0.624	0.630	95.8	0.038	0.607	0.620	93.4
	β_{21}	-0.004	0.253	0.249	93.6	0.045	0.258	0.248	93.4
	β_{22}	0.009	0.478	0.481	94.2	0.031	0.484	0.462	94.0
	η	-0.034	0.114	0.129	93.2	-0.016	0.115	0.111	92.6
2	β_{11}	-0.001	0.442	0.431	93.2	0.026	0.431	0.432	93.6
	β_{12}	-0.031	0.782	0.751	93.2	0.051	0.783	0.750	93.2
	β_{21}	-0.009	0.253	0.242	92.8	0.054	0.274	0.243	92.2
	β_{22}	0.015	0.483	0.457	92.0	0.035	0.486	0.452	92.6
	η	-0.060	0.110	0.121	91.2	-0.049	0.102	0.106	92.4

5. A Simulation Study

We conducted an extensive simulation study to assess the finite-sample performance of the proposed estimation procedure. In the study, we considered two covariates, X_{i1} and X_{i2}. The first covariate follows a Bernoulli distribution with a success probability of 0.5, and the second covariate follows the uniform distribution over $(0, 1)$. The latent variables b_i' were generated from a gamma distribution with mean one and variance $\eta = 0.4$. To generate the observed data, we assumed that T_i and C_i follow models (2.1) and (2.2), respectively, with $\Lambda_1(t) = \Lambda_2(t) = 0.05t^2$ or $\Lambda_1(t) = 0.5\log(1+t)$, and $\Lambda_2(t) = 0.05t^2$. The independent observation times C_i^c were set to the constant τ_c, which was chosen to yield the desired censoring percentage. The results given below are based on 500 replications, with $Q = 50$, $n = 200$, and $r = 0, 0.5, 1$, or 2.

Tables 1 and 2 present the results for the estimation of the parameters by the proposed method with the same baseline functions and various true values. The results include the estimated bias (Bias), given by the average of the estimates minus the true value, the sample standard error (SSE) of the estimates, the average of the estimated standard errors (SEE), and the 95% empirical coverage

Table 2. Simulation results for regression parameters based on simulated data with same baseline functions.

r	Par.	Bias	SSE	SEE	CP	Bias	SSE	SEE	CP
$(\beta_{11},\beta_{12},\beta_{21},\beta_{22})$=(0.2,0.2,-0.2,-0.2)						$(\beta_{11},\beta_{12},\beta_{21},\beta_{22})$=(-0.2,-0.2,-0.2,-0.2)			
0	β_{11}	0.005	0.279	0.268	94.4	-0.033	0.278	0.280	93.4
	β_{12}	0.025	0.453	0.470	95.8	-0.005	0.472	0.497	95.6
	β_{21}	-0.053	0.240	0.261	96.8	-0.063	0.250	0.267	96.4
	β_{22}	-0.026	0.510	0.499	94.2	-0.024	0.485	0.498	95.4
	η	-0.029	0.085	0.119	95.0	-0.032	0.137	0.165	93.8
0.5	β_{11}	-0.009	0.309	0.320	95.6	-0.006	0.323	0.326	96.0
	β_{12}	0.019	0.562	0.565	93.8	-0.009	0.556	0.570	95.6
	β_{21}	-0.009	0.259	0.257	94.4	-0.057	0.258	0.263	95.0
	β_{22}	-0.014	0.482	0.494	93.0	-0.042	0.480	0.495	95.2
	η	-0.018	0.112	0.138	95.4	-0.032	0.125	0.159	92.6
1	β_{11}	-0.010	0.357	0.350	93.6	-0.003	0.362	0.361	92.4
	β_{12}	-0.029	0.605	0.612	95.0	-0.006	0.627	0.630	95.2
	β_{21}	-0.037	0.260	0.254	94.2	-0.027	0.253	0.255	95.0
	β_{22}	-0.003	0.497	0.482	92.8	-0.012	0.513	0.481	93.2
	η	-0.053	0.091	0.119	93.2	-0.051	0.117	0.141	92.2
2	β_{11}	0.022	0.442	0.424	94.0	-0.002	0.418	0.433	96.4
	β_{12}	0.022	0.746	0.746	93.6	-0.011	0.742	0.758	94.2
	β_{21}	-0.032	0.248	0.247	95.0	-0.019	0.252	0.245	94.8
	β_{22}	-0.068	0.491	0.462	93.6	-0.047	0.474	0.460	92.2
	η	-0.059	0.090	0.114	90.8	-0.069	0.102	0.130	90.0

probability (CP). Here, for the monotone splines approximation, following Wang et al. (2016), we used five equally spaced knots in terms of percentiles, and an order of three for the monotone splines.

Tables 1 and 2 show that the proposed maximum likelihood estimators seem to be unbiased, and that the bootstrap variance estimates are appropriate. In addition, the normal approximation to the distribution of the estimators appears to be reasonable, because all empirical coverage probabilities are close to the nominal value. The results given in Tables 3 and 4 were obtained using different baseline functions; the other settings were the same as those in Tables 1 and 2 on the estimation of the regression parameters. The results shown in 3 and 4 are similar to those presented in Tables 1 and 2, again suggesting that the proposed method performs well in practice.

Note that the proposed estimation procedure assumes that the distribution of the latent variables b_i' is known up to the parameter η; however, this may not be true in practice. Thus, a question of interest is the robustness of the proposed estimation procedure to the distribution. To assess this, we repeated the study reported in Table 1, except that we generated b_i' from a log-normal distribution

Table 3. Simulation results on regression parameters based on the simulated data with different baseline functions.

r	Par.	Bias	SSE	SEE	CP	Bias	SSE	SEE	CP
$(\beta_{11},\beta_{12},\beta_{21},\beta_{22})=(0,0,0,0)$						$(\beta_{11},\beta_{12},\beta_{21},\beta_{22})=(0.2,0.2,0.2,0.2)$			
0	β_{11}	-0.007	0.244	0.259	95.4	0.042	0.245	0.255	94.1
	β_{12}	0.032	0.438	0.456	95.2	0.076	0.423	0.443	95.5
	β_{21}	0.006	0.254	0.259	95.0	0.062	0.262	0.255	92.0
	β_{22}	-0.010	0.453	0.486	94.8	0.042	0.501	0.460	91.4
	η	0.012	0.174	0.213	95.2	0.034	0.230	0.202	90.8
0.5	β_{11}	-0.011	0.293	0.301	94.2	0.046	0.300	0.303	95.3
	β_{12}	-0.018	0.509	0.527	94.4	0.033	0.514	0.525	93.6
	β_{21}	0.019	0.264	0.258	94.6	0.049	0.267	0.260	93.6
	β_{22}	-0.005	0.471	0.483	95.0	0.034	0.513	0.471	91.2
	η	-0.006	0.187	0.204	94.0	-0.049	0.214	0.195	91.6
1	β_{11}	-0.016	0.327	0.338	94.6	0.030	0.338	0.341	93.7
	β_{12}	-0.024	0.579	0.596	95.2	-0.007	0.558	0.598	96.7
	β_{21}	-0.008	0.263	0.257	95.0	0.050	0.262	0.255	93.1
	β_{22}	0.009	0.470	0.476	93.8	0.042	0.456	0.462	93.9
	η	-0.011	0.194	0.200	92.6	0.024	0.208	0.193	92.7
2	β_{11}	-0.007	0.431	0.414	93.8	0.059	0.417	0.424	94.3
	β_{12}	0.046	0.749	0.730	93.2	-0.057	0.639	0.723	96.6
	β_{21}	-0.014	0.246	0.253	95.2	0.062	0.260	0.257	94.5
	β_{22}	0.010	0.475	0.468	94.6	0.040	0.455	0.464	94.1
	η	-0.028	0.183	0.182	90.8	-0.011	0.218	0.187	90.0

with mean one and variance $\eta = 0.4$. The estimation was still based on the gamma distribution. The results are presented in Table 5, showing that the proposed estimators perform well, and that the estimation procedure is robust to a misspecification of the distribution of the latent variable. We also considered other setups, including different sample sizes and different numbers of knots and orders for the monotone splines. As expected, the performance improved when the sample size increased, and the results were stable with respect to different numbers of knots and orders.

6. An Application

In this section, we apply the proposed methodology to data from the tumorigenicity experiment described in Sun (2006), among others. The data set contains data on 144 RFM mice, and current-status data are available only for the time to tumor onset, the variable of interest, because the status of the tumor was examined only at death. The experiment involves two treatments, namely, a germ-free environment (48 mice) and a conventional environment (96 mice). Among the animals in the two groups, 35 and 27 mice, respectivelt, were

Table 4. Simulation results for regression parameters based on simulated data with different baseline functions.

r	Par.	Bias	SSE	SEE	CP	Bias	SSE	SEE	CP
$(\beta_{11},\beta_{12},\beta_{21},\beta_{22})$=(0.2,0.2,-0.2,-0.2)						$(\beta_{11},\beta_{12},\beta_{21},\beta_{22})$=(-0.2,-0.2,-0.2,-0.2)			
0	β_{11}	-0.016	0.228	0.260	96.0	-0.031	0.264	0.260	92.7
	β_{12}	-0.010	0.439	0.460	94.8	-0.041	0.438	0.453	94.2
	β_{21}	-0.030	0.262	0.276	94.4	-0.044	0.247	0.250	93.6
	β_{22}	-0.061	0.531	0.504	92.0	-0.035	0.452	0.446	93.1
	η	0.042	0.227	0.246	95.2	-0.067	0.167	0.208	91.8
0.5	β_{11}	-0.042	0.290	0.305	95.2	-0.023	0.293	0.298	94.0
	β_{12}	-0.003	0.481	0.530	96.2	-0.041	0.527	0.522	94.9
	β_{21}	-0.038	0.272	0.270	93.8	-0.037	0.242	0.247	94.2
	β_{22}	-0.051	0.471	0.498	93.6	-0.009	0.428	0.453	93.7
	η	0.017	0.228	0.230	95.0	-0.064	0.166	0.196	91.5
1	β_{11}	-0.067	0.326	0.343	94.2	-0.016	0.356	0.333	91.6
	β_{12}	-0.016	0.587	0.600	94.0	-0.021	0.583	0.594	94.1
	β_{21}	-0.049	0.252	0.268	96.0	-0.036	0.250	0.251	92.7
	β_{22}	-0.029	0.504	0.496	93.4	-0.021	0.462	0.449	93.2
	η	0.018	0.220	0.224	94.4	-0.080	0.173	0.196	90.7
2	β_{11}	-0.019	0.440	0.426	93.0	-0.014	0.405	0.401	94.8
	β_{12}	-0.029	0.734	0.737	94.0	-0.028	0.709	0.713	94.8
	β_{21}	-0.043	0.242	0.263	95.6	-0.036	0.236	0.244	94.0
	β_{22}	-0.071	0.486	0.484	95.0	-0.037	0.442	0.456	95.5
	η	-0.003	0.214	0.210	94.0	-0.086	0.177	0.196	89.8

observed to have tumors at the time of death, and because lung tumors are between lethal and nonlethal, we have dependent or informative censoring. One objective of the study was to compare the tumor growth rates between the two different environments or treatments.

To apply the proposed estimation procedure, let T_i denote the time to tumor onset, and define $X_i = 1$ if the ith mouse was in a germ-free environment, and zero otherwise. For the analysis, we need to select the degree of the splines, the number of knots of splines, and the nonnegative argument r. To do so, we used a three-dimensional grid search based on the AIC values. Specifically, we considered two, three, and four for the degree of the monotone splines, and varied the number of interior knots from three to eight to provide sufficient model flexibility and less of a computational burden. For r, we considered values from zero to two with increments of 0.1. The optimal frailty model is given by two degrees of splines with four interior knots and $r = 0.4$.

Table 6 shows the results under the optimal model for the estimation of the effects of the covariates, the estimated standard errors, and the p-values for testing no covariate effect. For comparison, we also include results based on three

Table 5. Simulation results for regression parameters based on a misspecified distribution with one covariate $X \sim B(1, 0.5)$.

		$n = 200$				$n = 400$			
r	Par.	Bias	SSE	SEE	CP	Bias	SSE	SEE	CP
$(\beta_1, \beta_2) = (0, 0)$									
0	β_1	0.008	0.261	0.260	93.8	0.004	0.186	0.179	92.6
	β_2	-0.011	0.191	0.184	93.8	-0.006	0.130	0.131	95.0
0.5	β_1	-0.012	0.311	0.309	95.0	-0.002	0.200	0.214	95.8
	β_2	0.008	0.185	0.184	95.2	-0.003	0.137	0.130	93.6
1	β_1	-0.002	0.362	0.357	94.4	0.000	0.238	0.244	95.0
	β_2	0.010	0.192	0.183	93.6	-0.008	0.134	0.131	94.4
2	β_1	0.011	0.435	0.435	93.4	-0.009	0.301	0.307	94.6
	β_2	0.003	0.186	0.184	94.0	-0.005	0.135	0.132	94.2
$(\beta_1, \beta_2) = (0.5, 0.5)$									
0	β_1	0.047	0.283	0.272	94.3	0.033	0.192	0.190	92.4
	β_2	0.013	0.192	0.191	94.7	0.010	0.138	0.137	93.7
0.5	β_1	0.032	0.319	0.315	93.9	0.022	0.217	0.219	94.9
	β_2	0.017	0.200	0.189	94.9	0.011	0.138	0.134	93.5
1	β_1	-0.009	0.359	0.356	94.1	0.001	0.242	0.246	94.4
	β_2	0.008	0.191	0.187	96.1	0.009	0.138	0.133	93.0
2	β_1	-0.017	0.423	0.431	94.9	-0.015	0.289	0.303	94.6
	β_2	0.014	0.183	0.186	94.3	0.012	0.129	0.133	95.6

and four degrees of splines, with the selected r and number of interior knots, as well as the corresponding results obtained under the PH frailty model and the PO frailty model. The results in Table 6 all seem to be consistent, suggesting that the animals in the germ-free environment had significantly longer survival times than those in the conventional environment, but that the animals in the two groups have similar risks of developing tumors. In addition, the results indicate that the tumor onset time and the death time were significantly correlated.

Also for comparison, we reanalyzed the data, imposing a noninformative censoring or independence assumption between the tumor onset time and the death time by setting $b_i = 1$ in models (2.1) and (2.2); the results are shown in Table 6. In this case, the animals in the germ-free environment had a significantly higher rate of tumor growth than those in the conventional environment. In other words, ignoring informative censoring may yield incorrect results.

7. Conclusion

We have proposed a sieve semiparametric maximum likelihood estimation procedure for a regression analysis of current-status data in the presence of informative censoring. The method uses the GOR frailty model to describe the covariate effects and the association between the failure time of interest and the

Table 6. Estimated covariate effects for the lung tumor study.

Degree	Model	$\hat{\beta}_1$	SEE	p-value	$\hat{\beta}_2$	SEE	p-value	$\hat{\eta}$	SEE	p-value
With informative censoring										
2	PH frailty	0.599	0.423	0.157	-2.104	0.289	< 0.001	0.080	0.040	0.043
	PO frailty	1.020	0.592	0.085	-2.101	0.201	< 0.001	0.079	0.039	0.044
	Optimal model	0.691	0.429	0.107	-2.119	0.283	< 0.001	0.090	0.034	0.008
3	PH frailty	0.674	0.448	0.132	-2.091	0.289	< 0.001	0.072	0.033	0.031
	PO frailty	1.061	0.590	0.072	-2.100	0.204	< 0.001	0.078	0.039	0.043
	Optimal model	0.769	0.470	0.102	-2.113	0.247	< 0.001	0.086	0.035	0.013
4	PH frailty	0.604	0.362	0.095	-2.129	0.266	< 0.001	0.096	0.037	0.010
	PO frailty	1.010	0.552	0.067	-2.075	0.247	< 0.001	0.062	0.022	0.006
	Optimal model	0.747	0.444	0.092	-2.117	0.228	< 0.001	0.088	0.043	0.041
Without informative censoring										
2	PH frailty	0.933	0.401	0.020	-1.966	0.258	< 0.001	-	-	-
	PO frailty	1.845	0.430	< 0.001	-1.966	0.268	< 0.001	-	-	-
	Optimal model	1.080	0.463	0.020	-1.966	0.273	< 0.001	-	-	-
3	PH frailty	0.933	0.391	0.017	-1.966	0.229	< 0.001	-	-	-
	PO frailty	1.842	0.424	< 0.001	-1.966	0.242	< 0.001	-	-	-
	Optimal model	1.074	0.452	0.017	-1.966	0.275	< 0.001	-	-	-
4	PH frailty	0.937	0.383	0.014	-1.966	0.305	< 0.001	-	-	-
	PO frailty	2.058	0.471	< 0.001	-1.966	0.260	< 0.001	-	-	-
	Optimal model	1.074	0.501	0.032	-1.966	0.275	< 0.001	-	-	-

observation time. We have also presented a novel EM algorithm with which to implement the methodology. In particular, the observed data are augmented by introducing latent variables and the proposed algorithm is computationally efficient. In addition, the proposed estimators are shown to be consistent and asymptotically normal, and the numerical results suggest that the proposed procedure work well in practice.

To address informative censoring, an alternative to the proposed frailty-based approach is the copula model-based approach. However, there are differences between the two methods. First, the former method gives unique estimators, whereas the latter may have an identifiability issue. In particular, the correlation parameters are only included in the joint distribution function for the latter method, but are included in both the marginal and the joint distribution functions for the former method.

The proposed method uses I-spline functions to approximate the unknown functions and to reduce the estimation of an infinite-dimensional function to one of finite parameters. As an alternative, one can use other smooth functions, such as B-spline functions; the method is similar. We further assume that the observation time C follows the PH frailty model, which helps to reduce the computational burden. Instead, one could use other models, such as the PO frailty model or the GOR frailty model, to model the covariate effects on C. In this case, the proposed approach can still be applied.

Note that although the proposed approach applies to a general class of GOR frailty models characterized by the nonnegative argument r, we assume that r is known. Thus, it would be useful to develop simultaneous estimation procedures for r. However, this is usually not possible without some extra assumptions or information. An alternative is to choose r based on some selection criterion to indicate the preferred model, as in the real-data analysis. Note that GOR frailty models do not include the additive hazards frailty model as a special case. Furthermore, it remains challenging to develop statistical methods to assess the goodness-of-fit of each model and compare the appropriateness of both models.

With respect to the distribution of the frailty, in the numerical study, we focused on the gamma distribution, because it has a large left tail, and thus is more appropriate for strong late dependence (Hougaard (1995)), which is usually the case for tumorigenicity experiments. Nevertheless, our numerical results suggest that the proposed estimation procedure is robust to a misspecification of the distribution of b_i. Of course, caution is required, because only limited studies were performed.

There are several possible directions for future research. First, we focus on current-status data. However, one may encounter informative interval-censored failure time data, which include current-status data as a special case (Sun (2006)). Thus, it would be useful to generalize the proposed method to the latter situation. Second, it would be useful to generalize the proposed method to allow for multivariate failure times of interest. Model checking is also important for these models, but is yet to be studied.

Supplementary Material

All technical proofs are given in the online Supplementary Material.

Acknowledgments

The authors would like to thank the editor and two referees for their helpful comments and suggestions. This research was funded by the Beijing Natural Science Foundation Z210003 and National Natural Science Foundation of China (NSFC) (Grant Nos: 12171328, 11971064). Shishun Zhao's study was supported by the National Natural Science Foundation of China (NSFC) (12071176). The R code for the proposed method can be downloaded at `https://github.com/xymath/EM.Iter`.

References

Banerjee, T., Chen, M. H., Dey, D. K. and Kim, S. (2007). Bayesian analysis of generalized odds-rate hazards models for survival data. *Lifetime Data Analysis* **13**, 241–260.

Chang, I. S., Wen, C. C. and Wu, Y. J. (2007). A profile likelihood theory for the correlated gamma-frailty model with current status family data. *Statistica Sinica* **17**, 1023–1046.

Chen, C. M., Lu, T. F. C., Chen, M. H. and Hsu, C. M. (2012). Semiparametric transformation models for current status data with informative censoring. *Biometrical Journal* **54**, 641–656.

Chen, M. H., Tong, X. and Sun, J. (2009). A frailty model approach for regression analysis of multivariate current status data. *Statistics in Medicine* **28**, 3424–3436.

Dinse, G. E. and Lagakos, S. W. (1983). Regression analysis of tumour prevalence data. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* **32**, 236–248.

Du, M., Hu, T. and Sun, J. (2019). Semiparametric probit model for informative current status data. *Statistics in Medicine* **38**, 2219–2227.

Efron, B. (1981). Censored data and the bootstrap. *Journal of the American Statistical Association* **76**, 312–319.

Hougaard, P. (1995). Frailty models for survival data. *Lifetime Data Analysis* **1**, 255–273.

Huang, J. (1996). Efficient estimation for the proportional hazards model with interval censoring. *The Annals of Statistics* **24**, 540–568.

Jewell N. P. and Emerson R. (2013). Current status data: An illustration with data on avalanche victims. In *Handbook of Survival Analysis*, 391C412. Chapman & Hall/CRC.

Lagakos, S. W. and Louis, T. A. (1988). Use of tumour lethality to interpret tumorigenicity experiments lacking cause-of-death data. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* **37**, 169–179.

Li, C. S., Taylor, J. M. and Sy, J. P. (2001). Identifiability of cure models. *Statistics and Probability Letters* **54**, 389–395.

Li, S., Hu, T., Wang, P. and Sun, J. (2017). Regression analysis of current status data in the presence of dependent censoring with applications to tumorigenicity experiments. *Computational Statistics and Data Analysis* **110**, 75–86.

Lin, D., Oakes, D. and Ying, Z. (1998). Additive hazards regression with current status data. *Biometrika* **85**, 289–298.

Lu, M., Zhang, Y. and Huang, J. (2007). Estimation of the mean function with panel count data using monotone polynomial splines. *Biometrika* **84**, 705–718.

Ma, L., Hu, T. and Sun, J. (2015). Sieve maximum likelihood regression analysis of dependent current status data. *Biometrika* **102**, 731–738.

McMahan, C. S., Wang, L. and Tebbs, J. M. (2013). Regression analysis for current status data using the EM algorithm. *Statistics in Medicine* **32**, 4452–4466.

Ramsay, J. O. (1988). Monotone regression splines in action. *Statistical Science* **3**, 425–441.

Rossini, A. J. and Tsiatis, A. A. (1996). A semiparametric proportional odds regression model for the analysis of current status data. *Journal of the American Statistical Association* **91**, 713–721.

Scharfstein, D. O., Tsiatis, A. A. and Gilbert, P. B. (1998). Semiparametric efficient estimation in the generalized odds-rate class of regression models for right-censored time-to-event data. *Lifetime Data Analysis* **4**, 355–391.

Shen, X. (1997). On methods of sieves and penalization. *The Annals of Statistics* **25**, 2555–2591.

Shen, X. and Wong, W. H. (1994). Convergence rate of sieve estimates. *The Annals of Statistics* **22**, 580–615.

Shiboski, S. C. (1998). Generalized additive models for current status data. *Lifetime Data Analysis* **4**, 29–50.

Sun J. (2006). *The Statistical Analysis of Interval-Censored Failure Time Data*. Springer.

Van der Vaart, A. W. and Wellner, J. (1996) *Weak Convergence and Empirical Processes*. Springer.

Wang, L., McMahan, C. S., Hudgens, M. G. and Qureshi, Z. P. (2016). A flexible, computationally efficient method for fitting the proportional hazards model to interval-censored data. *Biometrics* **72**, 222–231.

Xu, D., Zhao, S., Hu, T., Yu, M. and Sun, J. (2019). Regression analysis of informative current status data with the semiparametric linear transformation model. *Journal of Applied Statistics* **46**, 187–202.

Xu, D., Zhao, S. and Sun, J. (2020). Regression analysis of dependent current status data with the accelerated failure time model. *Communications in Statistics-Simulation and Computation* **51**, 6188–6196.

Zhang, Z., Sun, J. and Sun, L. (2005). Statistical analysis of current status data with informative observation times. *Statistics in Medicine* **24**, 1399–1407.

Zhao, S., Hu, T., Ma, L., Wang, P. and Sun, J. (2015). Regression analysis of informative current status data with the additive hazards model. *Lifetime Data Analysis* **21**, 241–258.

Zhou, J., Zhang, J. and Lu, W. (2017). An expectation maximization algorithm for fitting the generalized odds-rate model to interval censored data. *Statistics in Medicine* **36**, 1157–1171.

Yang Xu

Jilin University, Changchun City 130012, China.

E-mail: xymath@foxmail.com

Shishun Zhao

Department of Statistics, Jilin University, Changchun City 130012, China.

E-mail: zhaoss@jlu.edu.cn

Tao Hu

School of Mathematical Sciences, Capital Normal University, Haidian District, Beijing 100048, China.

E-mail: hutao@cnu.edu.cn

Jianguo Sun

Department of Statistics, University of Missouri, Columbia, MO 65211, USA.

E-mail: sunj@missouri.edu

(Received November 2021; accepted April 2022)