Lifetime Data Analysis (2020) 26:158–182
https://doi.org/10.1007/s10985-019-09467-z

An extended proportional hazards model for interval-censored data subject to instantaneous failures

Prabhashi W. Withana Gamage[1] · Monica Chaudari[2] · Christopher S. McMahan[3] · Edwin H. Kim[4] · Michael R. Kosorok[2]

Received: 18 May 2018 / Accepted: 11 February 2019 / Published online: 23 February 2019

Abstract

The proportional hazards (PH) model is arguably one of the most popular models used to analyze time to event data arising from clinical trials and longitudinal studies. In many such studies, the event time is not directly observed but is known relative to periodic examination times; i.e., practitioners observe either current status or interval-censored data. The analysis of data of this structure is often fraught with many difficulties since the event time of interest is unobserved. Further exacerbating this issue, in some such studies the observed data also consists of instantaneous failures; i.e., the event times for several study units coincide exactly with the time at which the study begins. In light of these difficulties, this work focuses on developing a mixture model, under the PH assumptions, which can be used to analyze interval-censored data subject to instantaneous failures. To allow for modeling flexibility, two methods of estimating the unknown cumulative baseline hazard function are proposed; a fully parametric and a monotone spline representation are considered. Through a novel data augmentation procedure involving latent Poisson random variables, an expectation–maximization (EM) algorithm is developed to complete model fitting. The resulting EM algorithm is easy to implement and is computationally efficient. Moreover, through extensive simulation studies the proposed approach is shown to provide both reliable estimation and inference. The motivation for this work arises from a randomized clinical trial aimed at assessing the effectiveness of a new peanut allergen treatment in attaining sustained unresponsiveness in children.

Keywords EM algorithm · Instantaneous failure data · Interval-censored data · Monotone splines · Proportional hazards model

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10985-019-09467-z) contains supplementary material, which is available to authorized users.

✉ Christopher S. McMahan
mcmaha2@clemson.edu

Extended author information available on the last page of the article

1 Introduction

Interval-censored data commonly arise in many clinical trials and longitudinal studies, and is characterized by the fact that the event time is not directly observable, but rather is known relative to observation times. As a special case, current status data (or case-1 interval censoring) arise when there exists exactly one observation time per study unit; i.e., at the observation time one discovers whether or not the event has occurred. Data of this structure often occurs in resource limited environments or due to destructive testing. Alternatively, general interval-censored data (or case-2 interval censoring) arise when multiple observation times are available for each study unit, and the event time can be ascertained relative to these observation times; i.e., the event time is said to be left-censored (right-censored) if it occurred before the first (after the last) observation time and interval-censored if it occurred between two observation times. It is well known that ignoring the structure of interval-censored data during an analysis can lead to biased estimation and inaccurate inference; see Odell et al. (1992) and Dorey et al. (1993). Further exasperating this issue, some studies are subject to the occurrence of instantaneous failures; i.e., the event time of interest for a number of the study units occurs at time zero. This feature can occur as an artifact of the study design or may arise during an intent-to-treat analysis (Matsuzaki et al. 2005; Lamborn et al. 2008; Liu et al. 2016). For example, Chen et al. (2015) describes a registry based study of end-stage renal disease patients, with the time of enrollment corresponding to the time at which the patient first received dialysis. In this study, several of the patient expire during the first dialysis treatment, leading to the occurrence of an instantaneous failure. Similarly, Liem et al. (1997) describes an intent-to-treat clinical trial comparing conventional anterior surgery and laparoscopic surgery for repairing inguinal hernia. In this study, various patients not receiving the allocated intervention, were inadequately excluded from the analysis to overcome issues such as consent withdrawal, procedure misfit etc. that would rightly attribute to instantaneous failures. Survival data with instantaneous events is not uncommon in epidemiological and clinical studies, and for this reason, herein a general methodology under the proportional hazards (PH) model is developed for the analysis of interval-censored data subject to instantaneous failures.

Originally proposed by Cox et al. (1972), the PH model has (arguably) become one of the most popular regression models for analyzing time-to-event data. For analyzing interval-censored data under the PH model, several notable contributions have been made in the recent years; e.g., see Finkelstein (1986), Groeneboom and Wellner (1992), Satten (1996), Goggins et al. (1998), Pan (1999), Goetghebeur and Ryan (2000), Pan (2000), Betensky et al. (2002), Cai and Betensky (2003), Sun (2007), Zhang et al. (2010), Zhang and Sun (2010), and Li and Ma (2013). More recently, Wang et al. (2016) developed a methodology under the PH model which can be used to accurately and reliably analyze interval-censored data. In particular, this approach makes use of a monotone spline representation to approximate the cumulative baseline hazard function. In doing so, an expectation–maximization (EM) algorithm is developed through a data augmentation scheme involving latent Poisson random variables which can be used to complete model fitting. It is worthwhile to note, that none of the aforementioned techniques were designed to account for the effects associ-

ated with instantaneous failures. The phenomenon of instantaneous (or early) failures occur in many lifetime experiments; to include, but not limited to, reliability studies and clinical trials. In reliability studies, instantaneous failures may be attributable to inferior quality or faulty manufacturing, where as in clinical trials these events may manifest due to adverse reactions to treatments or clinical definitions of outcomes. When the failure times are exactly observed, as is the case in reliability studies, it is common to incorporate instantaneous failures through a mixture of parametric models, with one being degenerate at time zero; e.g., see Muralidharan (1999), Kale and Muralidharan (2002), Murthy et al. (2004), Muralidharan and Lathika (2006), Pham and Lai (2007), and Knopik (2011). In the case of interval-censored data, more common among epidemiological studies and clinical trials, accounting for instantaneous failures becomes a more tenuous task, with practically no guidance available among the existing literature. Arguably, in the context of interval-censored data, one could account for instantaneous failures by introducing an arbitrarily small constant for each as an observation time, and subsequently treat the instantaneous failures as left-censored observations. In doing so, methods for interval-censored data, such as those discussed above, could be employed. While this approach may seem enticing, in the case of a relatively large number of instantaneous failures it has several pitfalls. In particular, through numerical studies (results not shown) it has been determined that this approach when used in conjunction with modeling techniques such as those proposed in Pan (1999) and Wang et al. (2016) may lead to inaccurate estimation of the survival curves and/or the covariate effects. Further, after an extensive literature review, it does not appear that any methodology has previously been developed to specifically address data of this structure. For these reasons, herein a general methodology under the PH model is developed for the analysis of interval-censored data subject to instantaneous failures.

For the analysis of interval-censored data subject to instantaneous failures a new mixture model is proposed, which is a generalization of the semi-parametric PH model studied in Wang et al. (2016). The proposed PH model is developed under the standard PH assumption; i.e., the covariates provide for a multiplicative effect on the baseline risk of both experiencing a failure at time zero and thereafter. Two separate techniques are developed for the purposes of estimating the cumulative baseline hazard function. The first allows a practitioner to specify a parametric form (up to a collection of unknown coefficients) for the unknown function, while the second provides for more modeling flexibility through the use of the monotone splines of Ramsay (1988). Under either formulation, a two-stage data augmentation scheme involving latent Poisson random variables is used to develop an efficient EM algorithm which can be used to estimate all of the unknown parameters. Through extensive simulation studies the proposed methodology is shown to provide reliable estimation and inference with respect to the covariate effects, cumulative baseline hazard function, and baseline probability of experiencing an instantaneous failure. This work is primarily motivated by a randomized clinical trial supported by both the National Institutes of Allergy and Infectious Diseases (NIAID) and the Wallace Foundation being conducted at the University of North Carolina at Chapel Hill, and is aimed at developing, assessing, and validating the proposed approach as a viable tool which can be used to analyze the data resulting from this trial.

The remainder of this article is organized as follows. Section 2 presents the development of the proposed model, the derivation of the EM algorithm, and outlines uncertainty quantification. The finite sample performance of the proposed approach is evaluated through extensive numerical studies, the features and results of which are provided in Sect. 3. Section 4 presents the analysis of the motivating data. Section 5 concludes with a summary discussion. Further, code which implements the proposed methodology has been added to the existing R software package `ICsurv` and is freely available from the CRAN (i.e., http://cran.us.rproject.org/).

2 Model and methodology

Let T denote the failure time of interest. Under the PH model, the survival function can be generally written as

$$S(t|\mathbf{x}) = S_0(t)^{e^{\mathbf{x}'\boldsymbol{\beta}}} \tag{1}$$

where $\mathbf{x}$ is a $(r \times 1)$-dimensional vector of covariates, $\boldsymbol{\beta}$ is the corresponding vector of regression coefficients, and $S_0(t)$ is the baseline survival function. Under the phenomenon of interest, there is a baseline risk (probability) of experiencing an instantaneous failure; i.e., $S(0|\mathbf{x} = \mathbf{0}_r) = S_0(0) = 1-p$, where $p \in [0, 1]$ is the baseline risk and $\mathbf{0}_r$ is a $(r \times 1)$-dimensional vector of zeros. Thus, under the PH assumptions, the probability of experiencing an instantaneous failure, given the covariate information contained in $\mathbf{x}$, can be ascertained from (1) as

$$\begin{aligned} P(T = 0|\mathbf{x}) &= 1 - S(0|\mathbf{x}) \\ &= 1 - (1 - p)^{e^{\mathbf{x}'\boldsymbol{\beta}}}. \end{aligned}$$

Similarly, given that an instantaneous failure does not occur, it is assumed that the failure time conditionally follows the standard PH model; i.e.,

$$P(T > t|\mathbf{x}, T > 0) = 1 - F(t|\mathbf{x}),$$

where $F(t|\mathbf{x}) = 1 - \exp\{-\Lambda_0(t)\exp(\mathbf{x}'\boldsymbol{\beta})\}$. Given that an instantaneous failure does not occur, $\Lambda_0(\cdot)$ can be viewed as the usual cumulative baseline hazard function and for the ease of exposition will henceforth be referred to as such. Note, in order for $F(\cdot|\mathbf{x})$ to be a proper cumulative distribution function, $\Lambda_0(\cdot)$ should be a monotone increasing function with $\Lambda_0(0) = 0$. Thus, through an application of the Law of Total Probability, one has that

$$\begin{aligned} P(T > t|\mathbf{x}) &= P(T > t|\mathbf{x}, T > 0)P(T > 0|\mathbf{x}) \\ &= \{1 - F(t|\mathbf{x})\}(1 - p)^{e^{\mathbf{x}'\boldsymbol{\beta}}}, \end{aligned}$$

for $t > 0$. Based on these assumptions, the cumulative distribution function of T can be expressed as the following mixture model,

$$H(t|\mathbf{x}) = \begin{cases} 1 - e^{-\alpha e^{\mathbf{x}'\boldsymbol{\beta}}}, & \text{for } t = 0, \\ 1 - e^{-\alpha e^{\mathbf{x}'\boldsymbol{\beta}}}\{1 - F(t|\mathbf{x})\}, & \text{for } t > 0, \end{cases}$$

where, for reasons that will shortly become apparent, $1 - p$ is re-parametrized as $\exp(-\alpha)$, for $\alpha > 0$; i.e., $\alpha = -\log(1 - p)$, where p is the baseline risk of experiencing an instantaneous failure. Under this model, the regression coefficients may be interpreted in the usual manner given the nonoccurrence of an instantaneous failure; i.e., it is easy to show that each regression coefficient represents the change in the log of the hazard ratio relative to a one unit change in the corresponding covariate for $t > 0$.

2.1 Observed data likelihood

In scenarios where interval-censored data arise, one has that the failure time (T) is not directly observed, but rather is known relative to two observation times, say $L < R$; i.e., one has that $L < T < R$. In general, the four different outcomes considered here can be represented through the values of L and R; i.e., an instantaneous failure ($L = R = 0$), T is left-censored ($0 = L < R < \infty$), T is interval-censored ($0 < L < R < \infty$), and T is right-censored ($0 < L < R = \infty$). For notational convenience, let ψ be an indicator denoting the event that T is not an instantaneous failure, and δ_1, δ_2, and δ_3 be censoring indicators denoting left-, interval-, and right-censoring, respectively; i.e., $\psi = I(T > 0)$, $\delta_1 = I(0 = L < R < \infty)$, $\delta_2 = I(0 < L < R < \infty)$, and $\delta_3 = (0 < L < R = \infty)$.

In order to derive the observed data likelihood, it is assumed throughout that the individuals are independent, and that conditional on the covariates, the failure time for an individual is independent of the observational process. This assumption is common among the survival literature; see, e.g., Liu and Shen (2009), Zhang and Sun (2010), and the references therein. The observed data collected on n individuals is given by $\mathcal{D} = \{(L_i, R_i, \mathbf{x}_i, \psi_i, \delta_{i1}, \delta_{i2}, \delta_{i3}); i = 1, 2, \ldots, n\}$, which constitutes n independent realization of $\{(L, R, \mathbf{x}, \psi, \delta_1, \delta_2, \delta_3)$. Thus, under the aforementioned assumptions, the observed data likelihood is given by

$$L_{obs}(\boldsymbol{\theta}) = \prod_{i=1}^{n} \Big[F(R_i|\mathbf{x}_i)^{\delta_{i1}} \{F(R_i|\mathbf{x}_i) - F(L_i|\mathbf{x}_i)\}^{\delta_{i2}} \{1 - F(L_i|\mathbf{x}_i)\}^{\delta_{i3}} \Big]^{\psi_i} \left\{ e^{-\alpha e^{\mathbf{x}'\boldsymbol{\beta}}} \right\}^{\psi_i} \left\{ 1 - e^{-\alpha e^{\mathbf{x}'\boldsymbol{\beta}}} \right\}^{1-\psi_i}, \quad (2)$$

where $\boldsymbol{\theta}$ represents the set of unknown parameters which are to be estimated.

2.2 Representations of $\Lambda_0(\cdot)$

The unknown parameters in the observed likelihood involve the regression parameters $\boldsymbol{\beta}$, α, and $\Lambda_0(\cdot)$. Herein, two techniques for modeling $\Lambda_0(\cdot)$ are discussed. The first approach considers the use of a fully parametric model, which is known up to a set of unknown coefficients. For example, a linear, quadratic, or logarithmic parametric model can be specified by setting $\Lambda_0(t) = \gamma_1 t$, $\Lambda_0(t) = \gamma_1 t + \gamma_2 t^2$, and

$\Lambda_0(t) = \gamma_1 \log(1+t)$, respectively. Note, all of these models obey the constraints placed on $\Lambda_0(\cdot)$, as long as the $\gamma_l > 0$, for $l = 1, 2$. In general, a parametric form can be specified as

$$\Lambda_0(t) = \sum_{l=1}^{k} \gamma_l b_l(t), \tag{3}$$

where $b_l(\cdot)$ is a monotone increasing function, $b_l(0) = 0$, and $\gamma_l > 0$, for $l = 1, \ldots, k$. Under these mild conditions, it is easily verified that $\Lambda_0(\cdot)$ inherits the same properties, and therefore adheres to the aforementioned constraints.

The second approach, which is inspired by the works Lin and Wang (2010), Cai et al. (2011), Wang and Dunson (2011), McMahan et al. (2013), Wang et al. (2015), and Wang et al. (2016), views $\Lambda_0(\cdot)$ as an unknown function and hence an infinite dimensional parameter. To reduce the dimensionality of the problem, the monotone splines of Ramsay (1988) are used approximate $\Lambda_0(\cdot)$. Structurally, this representation is identical to that of (3) with the exception that $b_l(\cdot)$ is a spline basis function and γ_l is an unknown spline coefficient, for $l = 1, \ldots, k$. Again, it is required that $\gamma_l > 0$, for all l, to ensure that $\Lambda_0(\cdot)$ is a monotone increasing function. Briefly, the spline basis functions are piece-wise polynomial functions and are fully determined once a knot sequence and the degree are specified. The shape of the basis splines are predominantly determined by the placement of the knots while the degree controls the smoothness (Cai et al. 2011). For instance, specifying the degree to take values 1, 2 or 3 correspond to the use of linear, quadratic or cubic polynomials respectively. Given the m knots and degree, the k ($k = m+\text{degree}-2$) basis functions are fully determined. In general, both the knot sequence and degree have the potential to impact the shape and flexibility of the spline model, with the former tending to be more influential. To make these specifications, one could fit multiple candidate models using different knot sequences and degrees and then use a model selection criterion (e.g., BIC) to identify the "best" from among the candidate models; for further discussion see Ramsay (1988), McMahan et al. (2013), and Wang et al. (2016). This approach is illustrated in the motivating data application.

In general, choosing the correct parametric form for $\Lambda_0(\cdot)$ can be relatively challenging. Thus, to avoid this challenge and the potential for model misspecification, it is generally suggested that the second approach be adopted. Of course, an alternate strategy could involve fitting the spline model and using the results to guide the choice of the parametric model, if a reasonable parametric model can be identified.

2.3 Data augmentation

Under either of the representations of $\Lambda_0(\cdot)$ proposed in Sect. 2.2, the unknown parameters in the observed data likelihood consist of $\boldsymbol{\theta} = (\boldsymbol{\beta}', \boldsymbol{\gamma}', \alpha)'$, where $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_k)'$. Since the observed data likelihood exists in closed-form, the maximum likelihood estimator MLE of $\boldsymbol{\theta}$ could be obtained by directly maximizing (2) with respect to $\boldsymbol{\theta}$; i.e., one could obtain $\hat{\boldsymbol{\theta}}$, the MLE of $\boldsymbol{\theta}$, as $\hat{\boldsymbol{\theta}} = \text{argmax}_{\boldsymbol{\theta}} L_{obs}(\boldsymbol{\theta})$. It is worthwhile to point out that the numerical process of directly maximizing (2), with respect to $\boldsymbol{\theta}$, is often unstable and rarely performs well (Wang et al. 2015).

To circumvent these numerical instabilities, an EM algorithm was derived for the purposes of identifying the MLE. This algorithm was developed based on a two-stage data augmentation process, where carefully structured latent Poisson random variables are introduced as missing data. The first stage relates both the instantaneous failure indicator and the censoring indicators to latent Poisson random variables; i.e., the Z_i, W_i, and Y_i are introduced such that

$$\begin{aligned} Z_i &\sim \text{Poisson}\{\Lambda_0(t_{i1})\exp(\mathbf{x}_i'\boldsymbol{\beta})\}, \\ W_i &\sim \text{Poisson}[\{\Lambda_0(t_{i2}) - \Lambda_0(t_{i1})\}\exp(\mathbf{x}_i'\boldsymbol{\beta})], \\ Y_i &\sim \text{Poisson}\{\alpha\exp(\mathbf{x}_i'\boldsymbol{\beta})\}, \end{aligned}$$

subject to the following constraints: $\delta_{i1} = I(Z_i > 0), \delta_{i2} = I(Z_i = 0, W_i > 0), \delta_{i3} = I(Z_i = 0, W_i = 0)$, and $\psi_i = I(Y_i = 0)$ where $t_{i1} = R_i I(\delta_{i1} = 1) + L_i I(\delta_{i1} = 0)$, and $t_{i2} = R_i I(\delta_{i2} = 1) + L_i I(\delta_{i3} = 1)$. At this stage of the data augmentation, the conditional likelihood is

$$L_A(\boldsymbol{\theta}) = \prod_{i=1}^{n} \left\{ P_{Z_i}(Z_i) P_{W_i}(W_i)^{\delta_{i2}+\delta_{i3}} C_i \right\}^{\psi_i} P_{Y_i}(Y_i) I(Y_i = 0)^{\psi_i} I(Y_i > 0)^{(1-\psi_i)}, \tag{4}$$

where $C_i = \delta_{i1} I(Z_i > 0) + \delta_{i2} I(Z_i = 0, W_i > 0) + \delta_{i3} I(Z_i = 0, W_i = 0)$ and $P_A(\cdot)$ is the probability mass function of the random variable A. In the second and final stage, the Z_i and W_i are separately decomposed into the sum of k independent latent Poisson random variables; i.e., $Z_i = \sum_{l=1}^{k} Z_{il}$ and $W_i = \sum_{l=1}^{k} W_{il}$, where

$$\begin{aligned} Z_{il} &\sim \text{Poisson}\{\gamma_l b_l(t_{i1})\exp(\mathbf{x}_i'\boldsymbol{\beta})\} \text{ and} \\ W_{il} &\sim \text{Poisson}[\{\gamma_l b_l(t_{i2}) - \gamma_l b_l(t_{i1})\}\exp(\mathbf{x}_i'\boldsymbol{\beta})]. \end{aligned}$$

At this stage of the augmented data likelihood is

$$\begin{aligned} L_C(\boldsymbol{\theta}) = \prod_{i=1}^{n}\prod_{l=1}^{k} \Big[P_{Z_{il}}(Z_{il}) I(Z_i = Z_{i\cdot}) \{P_{W_{il}}(W_{il}) I(W_i = W_{i\cdot})\}^{\delta_{i2}+\delta_{i3}} C_i \Big]^{\psi_i} \\ P_{Y_i}(Y_i) I(Y_i = 0)^{\psi_i} I(Y_i > 0)^{(1-\psi_i)}, \end{aligned} \tag{5}$$

where $Z_{i\cdot} = \sum_{l=1}^{k} Z_{il}$ and $W_{i\cdot} = \sum_{l=1}^{k} W_{il}$. It is relatively easy to show that by integrating (5) over the latent random variables one will obtain the observed data likelihood depicted in (2).

2.4 EM algorithm

In general, the EM algorithm consists of two steps: the expectation step (E-step) and the maximization step (M-step). The E-step in this algorithm involves taking the expectation of $\log\{L_c(\boldsymbol{\theta})\}$ with respect to all latent variables conditional on the current parameter value $\boldsymbol{\theta}^{(d)} = (\boldsymbol{\beta}^{(d)'}, \boldsymbol{\gamma}^{(d)'}, \alpha^{(d)})'$ and the observed data $\mathcal{D}$. This results in obtaining the $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(d)})$ function, where $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(d)}) = E[\log\{L_c(\boldsymbol{\theta})\}|\mathcal{D}, \boldsymbol{\theta}^{(d)}]$. The

M-step then finds $\boldsymbol{\theta}^{(d+1)} = \text{argmax}_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(d)})$. This process is repeated in turn until convergence of the algorithm is attained. In this particular setting, the E-step yields $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(d)})$ as

$$\begin{aligned} Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(d)}) = \sum_{i=1}^{n}\sum_{l=1}^{k} \psi_i \Big[&\{E(Z_{il}) + (\delta_{i2} + \delta_{i3})E(W_{il})\}\{\log(\gamma_l) + \mathbf{x}_i'\boldsymbol{\beta}\} \\ &- \gamma_l e^{\mathbf{x}_i'\boldsymbol{\beta}}\{(\delta_{i2} + \delta_{i1})b_l(R_i) + \delta_{i3} b_l(L_i)\}\Big] \\ + \sum_{i=1}^{n} &E(Y_i)\log(\alpha e^{\mathbf{x}_i'\boldsymbol{\beta}}) - \alpha e^{\mathbf{x}_i'\boldsymbol{\beta}} + H(\boldsymbol{\theta}^{(d)}), \end{aligned}$$

where $H(\boldsymbol{\theta}^{(d)})$ is a function of $\boldsymbol{\theta}^{(d)}$ but is free of $\boldsymbol{\theta}$. Notice that in $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(d)})$ the dependence of the expectations on the observed data and $\boldsymbol{\theta}^{(d)}$ is suppressed for notational convenience; i.e., from henceforth it should be understood that $E(\cdot) = E(\cdot|\boldsymbol{\mathcal{D}}, \boldsymbol{\theta}^{(d)})$.

An enticing feature, which makes the proposed approach computationally efficient, is that all of the expectations in $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(d)})$ can be expressed in closed-form, and moreover can be computed via simple matrix and vector operations. In particular, from (4) it can be ascertained that if $\delta_{i1} = 1$ and $\psi_i = 1$ then Z_i conditionally, given $\boldsymbol{\theta}^{(d)}$ and $\boldsymbol{\mathcal{D}}$, follows a zero-truncated Poisson distribution, and it follows a degenerate distribution at 0 for any other values of δ_{i1} and ψ_i. Thus, the conditional expectation of Z_i, given $\boldsymbol{\theta}^{(d)}$ and $\boldsymbol{\mathcal{D}}$, can be expressed as

$$E(Z_i) = \delta_{i1}\psi_i \Lambda_0^{(d)}(t_{i1}) \exp(\mathbf{x}_i'\boldsymbol{\beta}^{(d)}) \left[1 - \exp\{-\Lambda_0^{(d)}(t_{i1}) \exp(\mathbf{x}_i'\boldsymbol{\beta}^{(d)})\}\right]^{-1},$$

where $\Lambda_0^{(d)}(t) = \sum_{l=1}^{k} \gamma_l^{(d)} b_l(t)$. Through a similar set of arguments one can obtain the necessary conditional expectations of W_i and Y_i as

$$\begin{aligned} E(W_i) &= \delta_{i2}\psi_i\{\Lambda_0^{(d)}(t_{i2}) - \Lambda_0^{(d)}(t_{i1})\} \exp(\mathbf{x}_i'\boldsymbol{\beta}^{(d)}) \\ &\quad \times \left(1 - \exp[-\{\Lambda_0^{(d)}(t_{i2}) - \Lambda_0^{(d)}(t_{i1})\} \exp(\mathbf{x}_i'\boldsymbol{\beta}^{(d)})]\right)^{-1}, \\ E(Y_i) &= (1 - \psi_i)\alpha^{(d)} \exp(\mathbf{x}_i'\boldsymbol{\beta}^{(d)}) \left[1 - \exp\{-\alpha^{(d)} \exp(\mathbf{x}_i'\boldsymbol{\beta}^{(d)})\}\right]^{-1}, \end{aligned}$$

respectively. Further, from (5) it can be ascertained that if $\delta_{i1} = 1$ and $\psi_i = 1$ then Z_{il} conditionally, given Z_i, $\boldsymbol{\mathcal{D}}$ and $\boldsymbol{\theta}^{(d)}$, follows a binomial distribution with Z_i being the number of trials and $\gamma_l^{(d)} b_l(t_{i1})\{\Lambda_0^{(d)}(t_{i1})\}^{-1}$ being the success probability, and it follows a degenerate distribution at 0 for any other values of δ_{i1} and ψ_i. Thus, through an application of the Law of Iterated Expectations, the conditional expectation of Z_{il}, given $\boldsymbol{\theta}^{(d)}$ and $\boldsymbol{\mathcal{D}}$, can be expressed as

$$E(Z_{il}) = E(Z_i)\gamma_l^{(d)} b_l(t_{i1})\{\Lambda_0^{(d)}(t_{i1})\}^{-1}.$$

Through a similar set of arguments one can obtain the necessary conditional expectation of W_{il} as

$$E(W_{il}) = E(W_i)\gamma_l^{(d)}\{b_l(t_{i2}) - b_l(t_{i1})\}\{\Lambda_0^{(d)}(t_{i2}) - \Lambda_0^{(d)}(t_{i1})\}^{-1}.$$

Note, in the expressions of the expectations of Z_{il} and W_{il} the dependence on δ_{i1}, δ_{i2}, and ψ_i are suppressed with the properties associated with these variables being inherited from the expectations associated with Z_i and W_i, respectively.

The M-step of the algorithm then finds $\boldsymbol{\theta}^{(d+1)} = \text{argmax}_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(d)})$. To this end, consider the partial derivatives of $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(d)})$ with respect to $\boldsymbol{\theta}$ which are given by

$$\frac{\partial Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(d)})}{\partial \gamma_l} = \sum_{i=1}^{n} \psi_i \big[\gamma_l^{-1}\{E(Z_{il}) + (\delta_{i2} + \delta_{i3})E(W_{il})\} - e^{\mathbf{x}_i'\boldsymbol{\beta}}\{(\delta_{i2} + \delta_{i1})b_l(R_i) + \delta_{i3}b_l(L_i)\}\big], \tag{6}$$

$$\frac{\partial Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(d)})}{\partial \alpha} = \sum_{i=1}^{n} -e^{\mathbf{x}_i'\boldsymbol{\beta}} + \alpha^{-1}E(Y_i), \tag{7}$$

$$\frac{\partial Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(d)})}{\partial \boldsymbol{\beta}} = \sum_{i=1}^{n} \big[\psi_i\{E(Z_i) + \delta_{i2}E(W_i)\} - \psi_i\{(\delta_{i1} + \delta_{i2})\Lambda_0(R_i) + \delta_{i3}\Lambda_0(L_i)\}e^{\mathbf{x}_i'\boldsymbol{\beta}} - \alpha e^{\mathbf{x}_i'\boldsymbol{\beta}} + E(Y_i)\big]\mathbf{x}_i. \tag{8}$$

By setting (6) equal to zero and solving for γ_l, one can obtain

$$\gamma_l^*(\boldsymbol{\beta}) = \frac{\sum_{i=1}^{n} \psi_i\{E(Z_{il}) + \delta_{i2}E(W_{il})\}}{\sum_{i=1}^{n} \psi_i\{(\delta_{i2} + \delta_{i1})b_l(R_i) + \delta_{i3}b_l(L_i)\}e^{\mathbf{x}_i'\boldsymbol{\beta}}}, \tag{9}$$

for $l = 1, \ldots, k$. Similarly, by setting (7) equal to zero and solving for α, one can obtain

$$\alpha^*(\boldsymbol{\beta}) = \frac{\sum_{i=1}^{n} E(Y_i)}{\sum_{i=1}^{n} e^{\mathbf{x}_i'\boldsymbol{\beta}}}. \tag{10}$$

Notice that both $\gamma_l^*(\boldsymbol{\beta})$ and $\alpha^*(\boldsymbol{\beta})$ are non-negative since all quantities in these ratios are greater than or equal to zero for all values of $\boldsymbol{\beta}$, thus the updates for these parameters implicitly adhere to the constraints placed on them. Based on these expressions, one can obtain $\boldsymbol{\beta}^{(d+1)}$ by setting (8) equal to zero and solving the resulting system of equations for $\boldsymbol{\beta}$, after replacing γ_l and α by $\gamma_l^*(\boldsymbol{\beta})$ and $\alpha^*(\boldsymbol{\beta})$, respectively. Note, the aforementioned system of equations can easily be solved using a standard Newton Raphson approach. Finally, one obtains $\gamma_l^{(d+1)}$ and $\alpha^{(d+1)}$ as $\gamma_l^*(\boldsymbol{\beta}^{(d+1)})$ and $\alpha^*(\boldsymbol{\beta}^{(d+1)})$, respectively. It can be shown that $\boldsymbol{\theta}^{(d+1)}$ obtained using this process is the unique maximizer of $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(d)})$; the Web Appendix provides a proof of this assertion.

The proposed EM algorithm is now succinctly stated. First, initialize $\boldsymbol{\theta}^{(0)}$ and repeat the following steps until converges.

1. Calculate $\boldsymbol{\beta}^{(d+1)}$ by solving the following system of equations

$$\sum_{i=1}^{n}\left[\psi_i\{E(Z_i)+\delta_{i2}E(W_i)\}-\alpha^*(\boldsymbol{\beta})e^{\mathbf{x}_i'\boldsymbol{\beta}}+E(Y_i)\right]\mathbf{x}_i$$
$$=\sum_{i=1}^{n}\sum_{l=1}^{k}\psi_i\{(\delta_{i1}+\delta_{i2})b_l(R_i)+\delta_{i3}b_l(L_i)\}\gamma_l^*(\boldsymbol{\beta})e^{\mathbf{x}_i'\boldsymbol{\beta}}\mathbf{x}_i,$$

where $\gamma_l^*(\boldsymbol{\beta})$ and $\alpha^*(\boldsymbol{\beta})$ are defined above.

2. Calculate $\gamma_l^{(d+1)}=\gamma_l^*\left(\boldsymbol{\beta}^{(d+1)}\right)$ for $l=1,\ldots,k$ and $\alpha^{(d+1)}=\alpha^*\left(\boldsymbol{\beta}^{(d+1)}\right)$.
3. Update $d=d+1$.

At the point of convergence, define $\boldsymbol{\theta}^{(d)}=(\boldsymbol{\beta}^{(d)'},\boldsymbol{\gamma}^{(d)'},\alpha^{(d)})'$ to be the proposed estimator $\hat{\boldsymbol{\theta}}=(\hat{\boldsymbol{\beta}}',\hat{\boldsymbol{\gamma}}',\hat{\alpha})'$, which is the MLE of $\boldsymbol{\theta}$. As is generally advisable for numerical optimization algorithms, it is suggested that the proposed EM algorithm be implemented using multiple points of initialization in an effort to assess whether the algorithm is converging to the correct value. That said, numerical studies (see Sect. 3) have been conducted and tend to suggest that the proposed EM algorithm is relatively robust with respect to initialization.

2.5 Variance estimation

For the purposes of uncertainty quantification, several variance estimators were considered and evaluated; e.g., the inverse of the observed Fisher information matrix, the Huber sandwich estimator, and the outer product of gradients (OPG) estimator. After extensive numerical studies (results not shown), it was determined that the most reliable variance estimator, among those considered, was that of the OPG estimator. In general, the OPG estimator is obtained as

$$\widehat{V}(\widehat{\boldsymbol{\theta}})=\left[\frac{1}{n}\sum_{i=1}^{n}\dot{l}_i(\widehat{\boldsymbol{\theta}})\dot{l}_i'(\widehat{\boldsymbol{\theta}})\right]^{-1}$$

where $\dot{l}_i(\widehat{\boldsymbol{\theta}})=\partial l_i(\boldsymbol{\theta})/\partial\boldsymbol{\theta}|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}$ and $l_i(\boldsymbol{\theta})$ is the log-likelihood contribution of the ith individual. Using this estimator, one can conduct standard Wald type inference.

3 Simulation study

In order to investigate the finite sample performance of the proposed methodology, the following simulation study was conducted. The true distribution of the failure times was specified to be

$$H(t|\mathbf{x})=\begin{cases}1-e^{-\alpha e^{\mathbf{x}'\boldsymbol{\beta}}}, & \text{for } t=0,\\ 1-e^{-\alpha e^{\mathbf{x}'\boldsymbol{\beta}}}\{1-F(t|\mathbf{x})\}, & \text{for } t>0,\end{cases}$$

where $p=0.3$ (i.e., $\alpha=-\log(0.7)$), $\mathbf{x}=(\mathbf{x}_1,\mathbf{x}_2)'$, $\mathbf{x}_1\sim\mathrm{N}(0,1)$, $\mathbf{x}_2\sim$ Bernoulli(0.5), and $\boldsymbol{\beta}=(\beta_1,\beta_2)'$, where β_1 and β_2 take on values of -0.5 and

0.5 resulting in four regression parameter configurations. Additionally, these studies consider two cumulative baseline hazard functions; i.e., a logarithmic $\Lambda_0(t) = \log(t+1)/\log(11)$ and a linear $\Lambda_0(t) = 0.1t$. These choices were made so that these functions have similar scale but different shapes over a majority of the support of the observational process. In total, these specifications lead to eight separate data generating models for the failure times. Two generating processes were considered for the observation times: an exponential distribution with a mean of 10 and a discrete uniform over the interval [1, 17]. In both cases, a single observation time, O, was generated for each failure time which was not instantaneous (i.e., $T > 0$), and the intervals were created such that $L = 0\,(R = \infty)$ and $R = O\,(L = O)$ if T was smaller (greater) than O. A few comments are warranted on the selection of the observation processes. First, the latter process is actually indicative of the observation process considered in the motivating clinical trial, while the former attempts to match the baseline characteristics of the failure time distribution. Second, note that the specification of the two observation processes result in case-1 interval-censored (i.e., current status) data. This was done because data of this nature is collected as a part of the motivating clinical trial. In total, these data generating steps lead to sixteen generating mechanisms, and each were used to create 500 independent data sets consisting of n observations, where $n \in \{50, 100\}$.

In order to examine the performance of the proposed approach across a broad spectrum of characteristics, several different models were used to analyze each data set. First, following the development presented in Sect. 2.2, three different parametric forms were considered for the cumulative baseline hazard function: $\Lambda_{0_1}(t) = \gamma_1 \log(t+1)$, $\Lambda_{0_2}(t) = \gamma_1 t$, and $\Lambda_{0_3}(t) = \gamma_1 t + \gamma_2 t^2$, which are henceforth referred to as models M1, M2, and M3, respectively. Note, these specifications allow one to examine the performance of the proposed approach when the cumulative baseline hazard function is correctly specified (e.g., M2 when $\Lambda_0(t) = 0.1t$), over specified (e.g., M3 when $\Lambda_0(t) = 0.1t$), and misspecified (e.g., M1 when $\Lambda_0(t) = 0.1t$). Further, for each data set a model (M4) was fit using the monotone spline representation for the cumulative baseline hazard function developed in Sect. 2.2. In order to specify the basis functions, the degree was specified to be two, in order to provide adequate smoothness, and one interior knot was placed at the median of the observed finite nonzero interval end points, with boundary knots being placed at the minimum and maximum of the same. The EM algorithm derived in Sect. 2.4 was used to complete model fitting for M1-M4. The starting value for all implementations was set to be $\boldsymbol{\theta}^{(0)} = (\boldsymbol{\beta}^{(0)'}, \boldsymbol{\gamma}^{(0)'}, \alpha^{(0)}) = (\mathbf{0}_2', \mathbf{1}_k', 0.1)$, where $\mathbf{0}_k(\mathbf{1}_k)$ is a $(k \times 1)$-dimensional vector of zeros (ones). Convergence was declared when the maximum absolute differences between the parameter updates were less than the specified tolerance of 1×10^{-5}. Other studies (results not shown) suggest that the proposed algorithm is relatively robust to the point of initialization; i.e., in these studies the same point of convergence was attained from multiple points of initialization by the proposed algorithm.

Table 1 summarizes the estimates of the regression coefficients and the baseline instantaneous failure probability for all considered simulation configurations and models, when the observation times were drawn from a exponential distribution. This summary includes the empirical bias, the sample standard deviation of the 500 point estimates, the average of the 500 standard error estimates, and the empirical coverage probabilities associated with 95% Wald confidence intervals. Table 2 provides

Table 1 Simulation results summarizing the estimates of the regression coefficients and the baseline instantaneous failure probability obtained from M1–M4 across all simulation settings, when the observation times were sampled from an exponential distribution

	True $\Lambda_0(t) = \log(t+1)/\log(11)$ and $n = 50$															
	M1(True)				M2(Misspecified)				M3(Misspecified)				M4(Spline)			
Parameter	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95
$\beta_1 = -0.5$	−0.05	0.25	0.23	0.93	−0.10	0.28	0.23	0.89	−0.10	0.28	0.24	0.90	−0.07	0.26	0.24	0.93
$\beta_2 = -0.5$	−0.04	0.42	0.42	0.95	−0.09	0.46	0.41	0.92	−0.09	0.47	0.42	0.92	−0.07	0.44	0.44	0.95
$p = 0.3$	0.00	0.08	0.08	0.95	0.00	0.09	0.08	0.94	0.00	0.09	0.08	0.94	0.00	0.08	0.08	0.95
$\beta_1 = -0.5$	−0.03	0.21	0.21	0.96	−0.08	0.23	0.21	0.92	−0.08	0.23	0.22	0.93	−0.05	0.22	0.23	0.95
$\beta_2 = 0.5$	0.04	0.36	0.38	0.96	0.10	0.43	0.38	0.91	0.10	0.43	0.39	0.92	0.07	0.39	0.41	0.95
$p = 0.3$	−0.01	0.08	0.08	0.96	−0.02	0.08	0.08	0.93	−0.02	0.08	0.08	0.94	−0.01	0.08	0.08	0.95
$\beta_1 = 0.5$	0.05	0.23	0.23	0.94	0.08	0.25	0.23	0.90	0.09	0.25	0.23	0.91	0.07	0.24	0.24	0.94
$\beta_2 = -0.5$	−0.07	0.43	0.42	0.94	−0.12	0.48	0.40	0.89	−0.12	0.48	0.42	0.91	−0.09	0.45	0.44	0.94
$p = 0.3$	0.00	0.08	0.08	0.96	0.00	0.09	0.08	0.94	0.00	0.09	0.08	0.95	0.00	0.08	0.08	0.96
$\beta_1 = 0.5$	0.04	0.22	0.22	0.95	0.09	0.24	0.22	0.90	0.10	0.24	0.23	0.91	0.07	0.24	0.23	0.93
$\beta_2 = 0.5$	0.01	0.41	0.38	0.94	0.06	0.46	0.38	0.90	0.07	0.47	0.40	0.91	0.03	0.33	0.41	0.93
$p = 0.3$	0.00	0.08	0.08	0.95	−0.01	0.09	0.08	0.92	−0.01	0.09	0.08	0.93	−0.01	0.08	0.08	0.94
	True $\Lambda_0(t) = 0.1t$ and $n = 50$															
	M1(Misspecified)				M2(True)				M3(Over specified)				M4(Spline)			
Parameter	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95
$\beta_1 = -0.5$	0.01	0.24	0.23	0.94	−0.03	0.26	0.24	0.93	−0.06	0.27	0.25	0.92	−0.06	0.27	0.26	0.93
$\beta_2 = -0.5$	0.01	0.41	0.43	0.95	−0.03	0.43	0.44	0.95	−0.06	0.47	0.46	0.94	−0.06	0.47	0.48	0.94
$p = 0.3$	0.00	0.08	0.08	0.94	0.00	0.08	0.08	0.94	0.00	0.09	0.08	0.95	0.00	0.08	0.09	0.95

Table 1 continued

	True $\Lambda_0(t) = 0.1t$ and $n = 50$															
	M1(Misspecified)				M2(True)				M3(Over specified)				M4(Spline)			
Parameter	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95
$\beta_1 = -0.5$	0.02	0.20	0.22	0.97	−0.03	0.22	0.22	0.95	−0.05	0.23	0.23	0.94	−0.05	0.24	0.24	0.93
$\beta_2 = 0.5$	−0.01	0.34	0.39	0.97	0.04	0.37	0.40	0.96	0.07	0.40	0.42	0.95	0.07	0.40	0.43	0.94
$p = 0.3$	0.00	0.08	0.08	0.97	−0.01	0.08	0.08	0.96	−0.01	0.08	0.08	0.96	−0.01	0.08	0.08	0.95
$\beta_1 = 0.5$	0.00	0.22	0.23	0.95	0.04	0.23	0.24	0.95	0.07	0.25	0.25	0.93	0.07	0.25	0.26	0.94
$\beta_2 = -0.5$	−0.01	0.41	0.43	0.97	−0.05	0.44	0.44	0.95	−0.09	0.48	0.46	0.93	−0.09	0.48	0.48	0.94
$p = 0.3$	0.00	0.08	0.08	0.97	0.00	0.08	0.08	0.96	0.00	0.09	0.08	0.96	0.00	0.08	0.09	0.96
$\beta_1 = 0.5$	0.00	0.22	0.22	0.95	0.04	0.23	0.23	0.95	0.07	0.25	0.24	0.95	0.07	0.25	0.25	0.92
$\beta_2 = 0.5$	−0.03	0.40	0.40	0.94	0.02	0.43	0.41	0.93	0.04	0.46	0.42	0.93	0.04	0.46	0.43	0.91
$p = 0.3$	0.01	0.08	0.08	0.95	0.00	0.08	0.08	0.95	−0.01	0.09	0.08	0.94	−0.01	0.09	0.08	0.93
	True $\Lambda_0(t) = \log(t+1)/\log(11)$ and $n = 100$															
	M1(True)				M2(Misspecified)				M3(Misspecified)				M4(Spline)			
Parameter	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95
$\beta_1 = -0.5$	−0.03	0.15	0.15	0.94	−0.06	0.17	0.15	0.89	−0.06	0.17	0.15	0.90	−0.04	0.16	0.16	0.93
$\beta_2 = -0.5$	−0.01	0.29	0.28	0.94	−0.05	0.33	0.27	0.89	−0.05	0.33	0.28	0.90	−0.02	0.30	0.29	0.93
$p = 0.3$	0.00	0.06	0.06	0.95	0.00	0.06	0.06	0.94	0.00	0.06	0.06	0.94	0.00	0.06	0.06	0.95
$\beta_1 = -0.5$	−0.03	0.15	0.14	0.94	−0.07	0.17	0.14	0.87	−0.07	0.17	0.14	0.88	−0.04	0.15	0.15	0.94
$\beta_2 = 0.5$	0.01	0.26	0.26	0.94	0.06	0.30	0.25	0.89	0.06	0.30	0.26	0.89	0.03	0.27	0.27	0.94
$p = 0.3$	0.00	0.06	0.06	0.94	−0.01	0.06	0.05	0.92	−0.01	0.06	0.06	0.93	−0.01	0.06	0.06	0.94
$\beta_1 = 0.5$	0.01	0.15	0.15	0.95	0.05	0.16	0.15	0.91	0.05	0.16	0.15	0.92	0.02	0.15	0.15	0.95
$\beta_2 = -0.5$	0.01	0.28	0.28	0.94	−0.03	0.32	0.27	0.90	−0.03	0.32	0.27	0.91	0.01	0.29	0.29	0.94
$p = 0.3$	0.00	0.05	0.06	0.97	0.00	0.06	0.06	0.96	0.00	0.06	0.06	0.96	0.00	0.05	0.06	0.96

Table 1 continued

	True $\Lambda_0(t) = \log(t+1)/\log(11)$ and $n = 100$															
	M1(True)				M2(Misspecified)				M3(Misspecified)				M4(Spline)			
Parameter	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95
$\beta_1 = 0.5$	0.02	0.14	0.14	0.95	0.07	0.15	0.14	0.89	0.07	0.15	0.14	0.90	0.03	0.14	0.15	0.95
$\beta_2 = 0.5$	0.02	0.28	0.26	0.94	0.06	0.32	0.25	0.86	0.06	0.32	0.26	0.87	0.03	0.28	0.27	0.93
$p = 0.3$	0.00	0.06	0.06	0.93	−0.01	0.06	0.05	0.90	−0.01	0.06	0.06	0.91	0.00	0.06	0.06	0.93
	True $\Lambda_0(t) = 0.1t$ and $n = 100$															
	M1(Misspecified)				M2(True)				M3(Over specified)				M4(Spline)			
Parameter	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95
$\beta_1 = -0.5$	0.01	0.15	0.16	0.96	−0.03	0.16	0.16	0.94	−0.05	0.17	0.16	0.94	−0.04	0.16	0.17	0.94
$\beta_2 = -0.5$	0.04	0.28	0.29	0.95	0.00	0.30	0.30	0.95	−0.03	0.32	0.30	0.94	−0.02	0.32	0.31	0.95
$p = 0.3$	0.00	0.06	0.06	0.95	0.00	0.06	0.06	0.95	0.00	0.06	0.06	0.95	0.00	0.06	0.06	0.95
$\beta_1 = -0.5$	0.02	0.15	0.15	0.94	−0.02	0.15	0.15	0.94	−0.04	0.16	0.15	0.93	−0.04	0.16	0.16	0.94
$\beta_2 = 0.5$	−0.02	0.26	0.27	0.96	0.02	0.28	0.27	0.93	0.03	0.29	0.28	0.93	0.03	0.29	0.29	0.95
$p = 0.3$	0.01	0.06	0.06	0.94	0.00	0.06	0.06	0.94	−0.01	0.06	0.06	0.94	−0.01	0.06	0.06	0.94
$\beta_1 = 0.5$	−0.02	0.14	0.16	0.96	0.02	0.15	0.16	0.96	0.03	0.16	0.16	0.95	0.03	0.16	0.17	0.95
$\beta_2 = -0.5$	0.06	0.29	0.29	0.94	0.02	0.30	0.29	0.93	0.01	0.31	0.30	0.93	0.01	0.31	0.31	0.94
$p = 0.3$	0.00	0.05	0.06	0.97	0.00	0.05	0.06	0.96	0.00	0.05	0.06	0.96	0.00	0.06	0.06	0.97
$\beta_1 = 0.5$	−0.02	0.14	0.15	0.95	0.02	0.15	0.15	0.95	0.04	0.15	0.15	0.95	0.04	0.15	0.16	0.95
$\beta_2 = 0.5$	−0.02	0.27	0.27	0.94	0.02	0.29	0.27	0.94	0.04	0.30	0.28	0.93	0.04	0.30	0.29	0.93
$p = 0.3$	0.01	0.06	0.06	0.96	0.00	0.06	0.06	0.95	−0.01	0.06	0.06	0.94	−0.01	0.06	0.06	0.95

This summary include the average of the 500 point estimates minus the true value (Bias), the sample standard deviation of the 500 point estimates (SD), the average of the estimated standard errors (ESE), and empirical coverage probabilities associated with 95% Wald confidence intervals (CP95). Note, when $\Lambda_0(t) = \log(t+1)/\log(11)$ then M1 is the true parametric model and when $\Lambda_0(t) = 0.1t$ then M2 is the true parametric model

the analogous results for the case in which the observation times are sampled from a discrete uniform distribution. From these results, one will first note that across all considered simulation settings the proposed approach performs very well for M4 and the correct parametric model (i.e., M1 when $\Lambda_0(t) = \log(t+1)/\log(11)$ and M2 when $\Lambda_0(t) = 0.1t$); i.e., the parameter estimates exhibit very little bias, the sample standard deviation of the 500 point estimates are in agreement with the average of the standard error estimates, and the empirical coverage probabilities are at their nominal level. In summary, these findings tend to suggest that the proposed methodology can be used to reliably estimate the covariate effects, the instantaneous failure probability, and quantify the uncertainty in the same. Additionally, these findings generally continue to persist for the case in which the parametric model is over specified (e.g., M3 when $\Lambda_0(t) = 0.1t$), with the resulting estimates in some cases exhibiting a slightly larger bias and a bit more variability, as one would expect. Further, from these results one will also note that when the parametric model is misspecified (e.g., M2 and M3 when $\Lambda_0(t) = \log(t+1)/\log(11)$) the estimates tend to exhibit more bias and less reliable inference, which is expected under the misspecification of the cumulative baseline hazard function. Finally, the estimates obtained under M4 (i.e., the model which makes use of the monotone splines) exhibit little if any difference when compared to the estimates resulting from the correct parametric model. In summary, based on these findings it is generally suggested that the approach which makes use of the spline representation to approximate $\Lambda_0(t)$ should be used, since it avoids the potential of model misspecification and it obtains estimators of the unknown parameters, as well as their standard errors, that are equivalent to those estimators obtained under the true parametric model, the form of which is generally not known.

Figures 1 and 2 summarize the estimates of the baseline survival function (i.e., $S_0(t) = S(t|\mathbf{x} = \mathbf{0}_r)$) obtained from M1-M4 when $\Lambda_0(t) = \log(t+1)/\log(11)$ and $\Lambda_0(t) = 0.1t$, respectively. These results were obtained across all considered regression parameter configurations with the observation times being sampled from the exponential distribution and are based on a sample of when $n = 100$ observations. Web Figures 1–6 provide an analogous summary for the other simulation configurations. In particular, these figures present the true baseline survival functions, the average of the point-wise estimates, and the point-wise 2.5th and 97.5th percentiles of the estimates. These figures reinforce the main findings discussed above; i.e., M4 and the correctly specified parametric model again provide reliable estimates of the baseline survival function, and hence the cumulative baseline hazard function, across all simulation configurations. Similarly the over specified model also provides reliable estimates, but the same can not be said for the misspecified models. It is worthwhile to point out that the baseline survival curves do not extend to unity as time goes toward the origin, this is due to the fact that the baseline instantaneous failure probability is $p = 0.3$. Again, these findings support the recommendation that the spline based representation of the cumulative baseline hazard function should be adopted in lieu of a parametric model, thus obviating the possible ramifications associated with misspecification.

In summary, this simulation study illustrates that proposed methodology can be used to analyze current status data which is subject to instantaneous failures, and moreover that the monotone spline approach discussed in Sect. 2.2 should be adopted for approximating the unknown cumulative baseline hazard function. A few additional details

Table 2 Simulation results summarizing the estimates of the regression coefficients and the baseline instantaneous failure probability obtained from M1–M4 across all simulation settings, when the observation times were sampled from a discrete uniform distribution

	True $\Lambda_0(t) = \log(t+1)/\log(11)$ and $n = 50$															
	M1(True)				M2(Misspecified)				M3(Misspecified)				M4(Spline)			
Parameter	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95
$\beta_1 = -0.5$	−0.05	0.24	0.23	0.93	−0.07	0.25	0.23	0.91	−0.07	0.25	0.23	0.91	−0.06	0.24	0.24	0.92
$\beta_2 = -0.5$	−0.02	0.38	0.40	0.96	−0.04	0.41	0.40	0.94	−0.04	0.41	0.41	0.95	−0.03	0.40	0.42	0.96
$p = 0.3$	−0.01	0.08	0.08	0.96	−0.01	0.08	0.08	0.95	−0.01	0.08	0.08	0.95	−0.01	0.08	0.08	0.95
$\beta_1 = -0.5$	−0.06	0.22	0.22	0.95	−0.08	0.23	0.22	0.92	−0.09	0.23	0.22	0.93	−0.08	0.22	0.23	0.95
$\beta_2 = 0.5$	0.02	0.39	0.38	0.93	0.05	0.42	0.37	0.91	0.05	0.43	0.38	0.92	0.04	0.41	0.40	0.94
$p = 0.3$	0.00	0.08	0.08	0.94	−0.01	0.08	0.08	0.92	−0.01	0.09	0.08	0.92	−0.01	0.08	0.08	0.95
$\beta_1 = 0.5$	0.04	0.24	0.22	0.94	0.06	0.25	0.22	0.92	0.06	0.25	0.23	0.92	0.05	0.24	0.23	0.93
$\beta_2 = -0.5$	−0.02	0.42	0.40	0.94	−0.04	0.44	0.40	0.93	−0.04	0.45	0.41	0.93	−0.04	0.44	0.42	0.95
$p = 0.3$	0.00	0.09	0.08	0.96	0.00	0.09	0.08	0.94	0.00	0.09	0.08	0.95	0.00	0.09	0.08	0.96
$\beta_1 = 0.5$	0.04	0.22	0.21	0.94	0.07	0.23	0.21	0.92	0.07	0.23	0.22	0.92	0.06	0.23	0.22	0.93
$\beta_2 = 0.5$	0.05	0.39	0.38	0.94	0.08	0.42	0.37	0.92	0.08	0.42	0.39	0.92	0.07	0.40	0.40	0.94
$p = 0.3$	−0.01	0.08	0.08	0.95	−0.02	0.08	0.08	0.93	−0.02	0.08	0.08	.0.94	−0.01	0.08	0.08	0.95
	True $\Lambda_0(t) = 0.1t$ and $n = 50$															
	M1(Misspecified)				M2(True)				M3(Over specified)				M4(Spline)			
Parameter	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95
$\beta_1 = -0.5$	−0.02	0.23	0.23	0.94	−0.04	0.24	0.23	0.94	−0.06	0.25	0.24	0.93	−0.06	0.25	0.25	0.93
$\beta_2 = -0.5$	0.00	0.39	0.41	0.96	−0.02	0.40	0.42	0.95	−0.03	0.41	0.43	0.95	−0.03	0.42	0.44	0.96
$p = 0.3$	−0.01	0.08	0.08	0.95	−0.01	0.08	0.08	0.95	−0.01	0.08	0.08	0.95	−0.01	0.08	0.08	0.94

Table 2 continued

	True $\Lambda_0(t) = 0.1t$ and $n = 50$															
	M1(Misspecified)				M2(True)				M3(Over specified)				M4(Spline)			
Parameter	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95
$\beta_1 = -0.5$	−0.04	0.22	0.22	0.94	−0.06	0.23	0.22	0.93	−0.08	0.23	0.23	0.93	−0.08	0.23	0.24	0.94
$\beta_2 = 0.5$	0.01	0.39	0.38	0.94	0.04	0.40	0.39	0.93	0.05	0.41	0.40	0.93	0.06	0.42	0.42	0.94
$p = 0.3$	0.00	0.08	0.08	0.94	−0.01	0.08	0.08	0.93	−0.01	0.08	0.08	0.93	−0.01	0.08	0.08	0.93
$\beta_1 = 0.5$	0.02	0.24	0.23	0.94	0.04	0.25	0.23	0.94	0.05	0.25	0.24	0.93	0.05	0.26	0.24	0.93
$\beta_2 = -0.5$	0.00	0.41	0.41	0.95	−0.02	0.43	0.42	0.95	−0.03	0.45	0.43	0.94	−0.04	0.45	0.44	0.94
$p = 0.3$	0.00	0.08	0.08	0.95	0.00	0.09	0.08	0.95	0.00	0.09	0.08	0.95	0.00	0.09	0.08	0.95
$\beta_1 = 0.5$	0.02	0.21	0.22	0.95	0.04	0.22	0.22	0.94	0.06	0.23	0.23	0.93	0.06	0.23	0.23	0.94
$\beta_2 = 0.5$	0.02	0.38	0.38	0.95	0.05	0.40	0.39	0.94	0.07	0.42	0.40	0.94	0.07	0.42	0.42	0.94
$p = 0.3$	0.00	0.08	0.08	0.95	−0.01	0.08	0.08	0.94	−0.01	0.08	0.08	0.94	−0.01	0.08	0.08	0.94
	True $\Lambda_0(t) = \log(t+1)/\log(11)$ and $n = 100$															
	M1(True)				M2(Misspecified)				M3(Misspecified)				M4(Spline)			
Parameter	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95
$\beta_1 = -0.5$	−0.02	0.14	0.15	0.95	−0.03	0.15	0.15	0.92	−0.03	0.15	0.15	0.93	−0.03	0.15	0.15	0.94
$\beta_2 = -0.5$	0.00	0.27	0.27	0.96	−0.02	0.29	0.27	0.94	−0.02	0.29	0.27	0.94	−0.01	0.28	0.28	0.96
$p = 0.3$	−0.01	0.06	0.06	0.95	−0.01	0.06	0.06	0.93	−0.01	0.06	0.06	0.94	−0.01	0.06	0.06	0.94
$\beta_1 = -0.5$	−0.03	0.14	0.14	0.94	−0.05	0.15	0.14	0.92	−0.05	0.15	0.14	0.92	−0.03	0.14	0.15	0.94
$\beta_2 = 0.5$	0.00	0.26	0.25	0.94	0.03	0.28	0.25	0.92	0.03	0.28	0.25	0.93	0.01	0.26	0.26	0.94
$p = 0.3$	0.00	0.06	0.06	0.95	0.00	0.06	0.06	0.94	0.00	0.06	0.06	0.94	0.00	0.06	0.06	0.96
$\beta_1 = 0.5$	0.02	0.16	0.15	0.94	0.03	0.16	0.15	0.94	0.03	0.16	0.15	0.94	0.02	0.16	0.15	0.94
$\beta_2 = -0.5$	0.00	0.28	0.27	0.95	−0.01	0.29	0.27	0.93	−0.01	0.29	0.27	0.93	0.00	0.28	0.28	0.95
$p = 0.3$	0.00	0.06	0.06	0.95	0.00	0.06	0.06	0.95	0.00	0.06	0.06	0.95	0.00	0.06	0.06	0.95

Table 2 continued

	True $\Lambda_0(t) = \log(t+1)/\log(11)$ and $n = 100$															
	M1(True)				M2(Misspecified)				M3(Misspecified)				M4(Spline)			
Parameter	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95
$\beta_1 = 0.5$	0.01	0.14	0.14	0.95	0.04	0.15	0.14	0.92	0.04	0.15	0.14	0.93	0.02	0.14	0.14	0.95
$\beta_2 = 0.5$	0.03	0.27	0.25	0.94	0.05	0.29	0.25	0.91	0.05	0.29	0.25	0.91	0.04	0.28	0.26	0.93
$p = 0.3$	0.00	0.06	0.06	0.95	−0.01	0.06	0.05	0.94	−0.01	0.06	0.06	0.94	0.00	0.06	0.06	0.95
	True $\Lambda_0(t) = 0.1t$ and $n = 100$															
	M1(Misspecified)				M2(True)				M3(Over specified)				M4(Spline)			
Parameter	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95
$\beta_1 = -0.5$	0.00	0.15	0.15	0.95	−0.02	0.15	0.15	0.93	−0.03	0.16	0.16	0.93	−0.03	0.16	0.16	0.94
$\beta_2 = -0.5$	0.02	0.27	0.28	0.96	0.00	0.28	0.28	0.96	−0.01	0.29	0.29	0.96	−0.01	0.29	0.29	0.97
$p = 0.3$	−0.01	0.06	0.06	0.95	−0.01	0.06	0.06	0.95	−0.01	0.06	0.06	0.94	−0.01	0.06	0.06	0.95
$\beta_1 = -0.5$	0.00	0.14	0.14	0.95	−0.02	0.14	0.14	0.94	−0.03	0.15	0.15	0.94	−0.03	0.15	0.15	0.93
$\beta_2 = 0.5$	−0.03	0.25	0.26	0.95	0.00	0.27	0.26	0.94	0.01	0.27	0.27	0.94	0.01	0.27	0.27	0.94
$p = 0.3$	0.01	0.06	0.06	0.96	0.00	0.06	0.06	0.95	0.00	0.06	0.06	0.95	0.00	0.06	0.06	0.95
$\beta_1 = 0.5$	0.00	0.15	0.15	0.94	0.02	0.16	0.15	0.94	0.02	0.16	0.16	0.94	0.02	0.16	0.16	0.95
$\beta_2 = -0.5$	0.01	0.28	0.28	0.95	−0.01	0.29	0.28	0.95	−0.01	0.30	0.28	0.94	−0.01	0.30	0.29	0.94
$p = 0.3$	0.00	0.06	0.06	0.96	0.00	0.06	0.06	0.96	0.00	0.06	0.06	0.96	0.00	0.06	0.06	0.96
$\beta_1 = 0.5$	−0.02	0.13	0.14	0.96	0.01	0.14	0.14	0.95	0.02	0.14	0.15	0.95	0.02	0.14	0.15	0.96
$\beta_2 = 0.5$	0.00	0.26	0.26	0.96	0.03	0.27	0.26	0.95	0.04	0.28	0.27	0.95	0.04	0.28	0.27	0.95
$p = 0.3$	0.00	0.06	0.06	0.96	0.00	0.06	0.06	0.95	0.00	0.06	0.06	0.95	0.00	0.06	0.06	0.95

This summary include the average of the 500 point estimates minus the true value (Bias), the sample standard deviation of the 500 point estimates (SD), the average of the estimated standard errors (ESE), and empirical coverage probabilities associated with 95% Wald confidence intervals (CP95). Note, when $\Lambda_0(t) = \log(t+1)/\log(11)$ then M1 is the true parametric model and when $\Lambda_0(t) = 0.1t$ then M2 is the true parametric model

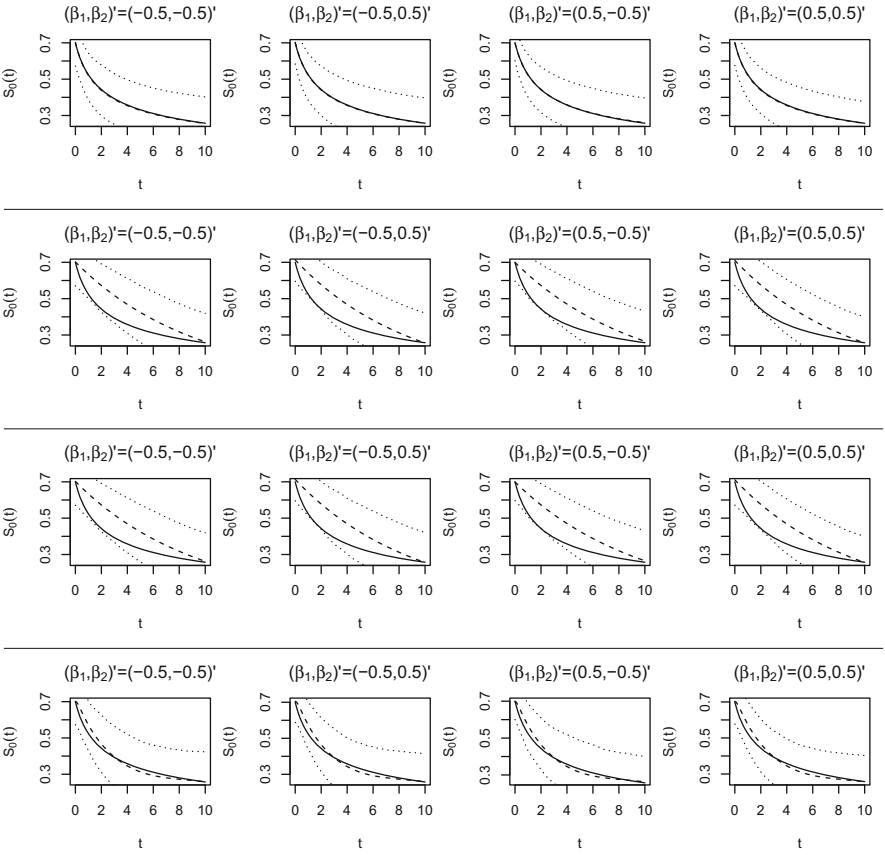

Fig. 1 Simulation results summarizing the estimates of the baseline survival function obtained by the proposed approach under M1 (first row), M2 (second row), M3 (third row), and M4 (fourth row) when $\Lambda_0(t) = \log(t+1)/\log(11)$, $n = 100$, and observation times were drawn from an exponential distribution. The solid line provides the true value, dashed line represents the average estimated value, and the dotted lines indicate the 2.5% and 97.5% quantiles, of the point-wise estimates. Note, M2 and M3 are misspecified models in this setting

about the numerical aspects of the approach follow. First, the average time required to complete model fitting was approximately 0.4, and 0.8 seconds for $n = 50$ and 100, respectively, supporting the claim that the proposed approach is computationally efficient. Second, for the reasons of complete transparency, when $n = 50$ and 100, there were 41 and 1 data sets, respectively, that experienced numerical issues among 8000. These numerical issues were resolved by simply changing the placement of the interior knot.

3.1 Simulation study II

To demonstrate the performance of the proposed methodology with respect to analyzing interval-censored data subject to instantaneous failures, an additional simulation

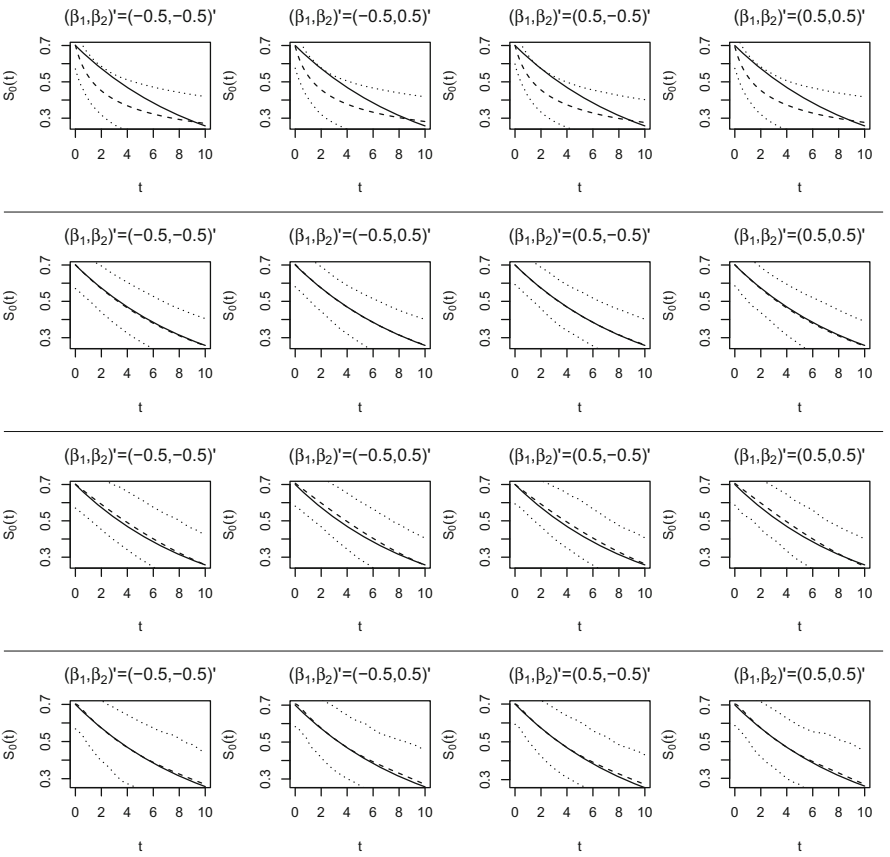

Fig. 2 Simulation results summarizing the estimates of the baseline survival function obtained by the proposed approach under M1 (first row), M2 (second row), M3 (third row), and M4 (fourth row) when $\Lambda_0(t) = 0.1t$, $n = 100$, and observation times were drawn from an exponential distribution. The solid line provides the true value, dashed line represents the average estimated value, and the dotted lines indicate the 2.5% and 97.5% quantiles, of the point-wise estimates. Note, M1 and M3 are misspecified and over specified models, respectively, in this setting

study was conducted. This study considers the exact same data generating process described above with the exception that multiple observation times were generated through an independent observational process. In particular, the number of observation times were determined as one plus a Poisson random variable having mean parameter three, which assured that each individual had at least one observation time while allowing for a different number of observation times across individuals. The waiting times between adjacent observation times were generated according to an exponential distribution having mean of 2.5. For non-instantaneous failure times, the interval endpoints (i.e., L and R) were determined by examining which of the two observation times bounded the failure time, with the convention that if T was smaller (greater) than the smallest (largest) observation time then $L = 0$ ($R = \infty$). Using this data genrating mechanism, 500 independent data sets were generated and analyzed

in the exact same fashion as described above. Web Table 1 provides the summary of the regression parameter estimates and Web Figures 7–10 summarize the estimates of $\Lambda_0(t)$ that were obtained as a part of this study. These results reinforce all the findings discussed above; i.e., the proposed model can be used to accurately and efficiently analyze interval-censored data subject to instantaneous failures. Moreover, the average model fitting time per data set was approximately 0.3, and 0.4 seconds for $n = 50$ and 100, respectively. Demonstrating that the proposed approach continues to be computationally efficient even when tasked with the analysis of interval-censored data.

4 Data application

Supported by both the National Institutes of Allergy and Infectious Diseases and the Wallace Foundation and conducted at the University of North Carolina at Chapel Hill, the Sublingual Immunotheray for Peanut Allergy and Induction of Tolerance (SLIT-TLC) clinical trial (NCT01373242 2017) was initiated in 2011 to assess the effectiveness of peanut SLIT to induce clinical tolerance. In response to growing evidence that tolerance may not be achievable through SLIT, the protocol was revised in 2016 with an altered study design to assess time to loss of desensitization among subjects completing a 48 month course of SLIT for peanut allergy. Participants include children between 1-11 years of age at the time of enrollment. The revised study consisted of a build-up/maintenance phase (approximately, 48 months), wherein SLIT therapy was incrementally increased during the initial 6 months and maintained thereafter, and an avoidance phase (17 weeks), wherein the therapy was discontinued. Throughout the study, Double-Blind Placebo Control Food Challenges (DBPCFC) were used to assess participants' reactive thresholds to peanut allergen. For further study details see Chaudhari et al. (2018).

This study focuses on the avoidance phase beginning from the 48-month challenge until the administration of the post-avoidance challenge with the primary endpoint defined as the time to loss of desensitization to a targeted dose of peanut allergen. In particular, at the time of the 48-month challenge each participant was given a series of incrementally increasing doses up to 5000mg. Subsequently, any participant safely ingesting over 2900mg were then reassessed up to this level after avoidance to determine whether they remained desensitized. To assess time to loss of desensitization, the time of the final DBPCFC for each participant was randomly (in a uniform fashion) selected during the 17 week avoidance phase. It is important to note that the endpoint is not observable in this study but rather is known relative to the time of the final DBPCFC resulting in current status data. Further, some participants failed the DBPCFC prior to the avoidance phase (i.e., at the time of the 48-month challenge) and thus experienced an instantaneous failure at the initiation of the avoidance phase. The primary focus of this analysis involves assessing the association of risk factors with the loss of desensitization, as well as estimating the baseline survival function and baseline probability of an instantaneous event.

Of the fifty-four participants who enrolled in this study, seven withdrew before the avoidance phase and one did not complete the final DBPCFC. Thus, complete data were

Table 3 SLIT data analysis: estimated regression coefficients, estimated standard errors (ESE), and p values obtained by the proposed method

Covariate	Targeted threshold 2900 mg		
	Estimate	ESE	p Value
Gender	− 0.2488	0.4019	0.5360
Age (centered at mean)	− 0.0060	0.0816	0.9413
IgE	0.2245	0.1189	0.0590

available on $n = 46$ participants. Among these participants, twenty-four experienced instantaneous failures, seven were left-censored, and fifteen were right-censored. The available covariates include gender (1 = male), age (in years, centered), and the logarithm of the baseline (i.e., at the initiation of the avoidance phase) Immunoglobulin E (IgE) levels (kU/L), with each entering the proposed model as a first order term. The EM algorithm was used to fit the proposed model. To provide for modeling flexibility, the monotone spline representation was used to approximate the cumulative baseline hazard function. To specify the spline model in a data driven manner several candidate models which made use of different knot sequences and degree values were considered; for the specific configurations see Web Table 2. The proposed approach was then used to fit each of the candidate models with the final model being selected based on the BIC criterion; Web Table 2 provides the estimated regression coefficients and BIC values for each of the candidate models. The final model used a degree value of two and placed a single interior knot at the first quartile of the observation times.

Table 3 presents the estimated regression coefficients along with their estimated standard errors. These results indicate that only one of the considered covariates is significantly associated (at the 0.10 level) with the loss of desensitization. In particular, it appears that elevated baseline IgE (on the log-scale) levels correspond to an increased hazard of experiencing a loss of desensitization. This finding is reasonable since IgE is an antibody produced by the immune system when it overreacts to an allergen, which causes cells to release chemicals that in turn cause an allergic reaction. Thus, it is reasonable to believe that participants with elevated IgE levels at the initiation of the study were at a higher risk of experiencing loss of desensitization; i.e., of having an allergic reaction to the targeted dose. In addition to this finding, the proposed method also provides an estimate of the baseline survival function (Fig. 3) and the baseline probability of experiencing an instantaneous failure (0.3879). Note, in Fig. 3, the magnitude of the drop in the baseline survival function at time zero is expected given the estimated baseline probability of experiencing an instantaneous failure.

5 Discussion

This work proposed a new PH model which can be used to analyze interval-censored data subject to instantaneous failures. Through the model development, two techniques for approximating the unknown cumulative baseline hazard function are illustrated. To complete model fitting, a novel EM algorithm is developed through a two-stage data augmentation process. The resulting algorithm is easy to implement and is computationally efficient. These features are likely attributable to the fact that the carefully

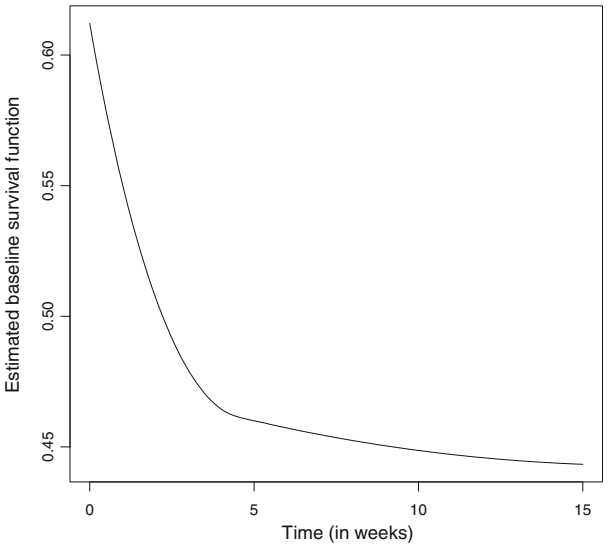

Fig. 3 Estimated baseline survival function for time to loss of SU to a targeted dose level of 2900 mg obtained using the proposed method

structured data augmentation steps lead to closed-form expressions for all necessary quantities in the E-step of the algorithm. Moreover, in the M-step the regression coefficients are updated through solving a low-dimensional system of equations, while all other unknown parameters are updated in closed-form. The finite sample performance of the proposed approach was exhibited through an extensive numerical study. This study suggests that the use of the monotone spline representation of the cumulative baseline hazard function would in general be preferable, in order to circumvent the possibility of model misspecification. Further, the proposed approach was applied to current status data collected on children as a part of the Sublingual Immunotheray for Peanut Allergy and Induction of Tolerance clinical trial. To further disseminate this work, code, written in R, has been prepared and is available upon request from the corresponding author.

The large sample properties (i.e., asymptotic normality of the regression coefficient estimates and consistency of the spline estimator) are expected but are not rigorously established here. Following the works of Zhang et al. (2010), Lu and Li (2017), and Lu and McMahan (2018) it should be possible to formally establish these results, under similar regularity conditions. This conjecture is supported by findings from Sect. 3. That being said, a topic for future research could be targeted at formally establishing these results. In so doing, one could also endeavor to develop goodness-of-fit tests which could be used to assess the reasonableness of the proposed extended proportional hazards model.

Acknowledgements This research was partially supported by the Wallace Foundation and National Institutes of Health grants AI121351 and UL1 TR001111.

References

Betensky RA, Lindsey JC, Ryan LM, Wand M (2002) A local likelihood proportional hazards model for interval censored data. Stat Med 21(2):263–275. https://doi.org/10.1002/sim.993

Cai B, Lin X, Wang L (2011) Bayesian proportional hazards model for current status data with monotone splines. Comput Stat Data Anal 55(9):2644–2651. https://doi.org/10.1016/j.csda.2011.03.013

Cai T, Betensky RA (2003) Hazard regression for interval-censored data with penalized spline. Biometrics 59(3):570–579. https://doi.org/10.1111/1541-0420.00067

Chaudhari M, Kim EH, Withana Gamage PW, McMahan CS, Kosorok MR (2018) Study design with staggered sampling times for evaluating sustained unresponsiveness to peanut sublingual immunotherapy. Stat Med 37:3944–3958

Chen CM, Lai CC, Cheng KC, Weng SF, Liu WL, Shen HN (2015) Effect of end-stage renal disease on long-term survival after a first-ever mechanical ventilation: a population-based study. Crit Care 19(1):354. https://doi.org/10.1186/s13054-015-1071-x

Cox R et al (1972) Regression models and life tables (with discussion). J R Stat Soc 34:187–220

Dorey FJ, Little RJ, Schenker N (1993) Multiple imputation for threshold-crossing data with interval censoring. Stat Med 12(17):1589–1603

Finkelstein DM (1986) A proportional hazards model for interval-censored failure time data. Biometrics 42:845–854

Goetghebeur E, Ryan L (2000) Semiparametric regression analysis of interval-censored data. Biometrics 56(4):1139–1144. https://doi.org/10.1111/j.0006-341X.2000.01139.x

Goggins WB, Finkelstein DM, Schoenfeld DA, Zaslavsky AM (1998) A markov chain monte carlo em algorithm for analyzing interval-censored data under the cox proportional hazards model. Biometrics 54:1498–1507

Groeneboom P, Wellner JA (1992) Information bounds and nonparametric maximum likelihood estimation, vol 19. Springer, Berlin

Kale B, Muralidharan K (2002) Optimal estimating equations in mixture distributions accommodating instantaneous or early failures. Qual Control Appl Stat 47(6):677–680

Knopik L (2011) Model for instantaneous failures. Sci Probl Mach Oper Maint 46(2):37–45

Lamborn KR, Yung WA, Chang SM, Wen PY, Cloughesy TF, DeAngelis LM, Robins HI, Lieberman FS, Fine HA, Fink KL et al (2008) Progression-free survival: an important end point in evaluating therapy for recurrent high-grade gliomas. Neuro-oncology 10(2):162–170. https://doi.org/10.1215/15228517-2007-062

Li J, Ma S (2013) Survival analysis in medicine and genetics. CRC Press, Boca Raton

Liem MS, van der Graaf Y, van Steensel CJ, Boelhouwer RU, Clevers GJ, Meijer WS, Stassen LP, Vente JP, Weidema WF, Schrijvers AJ et al (1997) Comparison of conventional anterior surgery and laparoscopic surgery for inguinal-hernia repair. N Engl J Med 336(22):1541–1547. https://doi.org/10.1056/NEJM199705293362201

Lin X, Wang L (2010) A semiparametric probit model for case 2 interval-censored failure time data. Stat Med 29(9):972–981. https://doi.org/10.1002/sim.3832

Liu C, Yang W, Devidas M, Cheng C, Pei D, Smith C, Carroll WL, Raetz EA, Bowman WP, Larsen EC et al (2016) Clinical and genetic risk factors for acute pancreatitis in patients with acute lymphoblastic leukemia. J Clin Oncol 34(18):2133–2140. https://doi.org/10.1200/JCO.2015.64.5812

Liu H, Shen Y (2009) A semiparametric regression cure model for interval-censored data. J Am Stat Assoc 104(487):1168–1178. https://doi.org/10.1198/jasa.2009.tm07494

Lu M, Li CS (2017) Penalized estimation for proportional hazards models with current status data. Stat Med 36(30):4893–4907

Lu M, McMahan CS (2018) A partially linear proportional hazards model for current status data. Biometrics 74:1240–1249

Matsuzaki A, Nagatoshi Y, Inada H, Nakayama H, Yanai F, Ayukawa H, Kawakami K, Moritake H, Suminoe A, Okamura J (2005) Prognostic factors for relapsed childhood acute lymphoblastic leukemia: impact of allogeneic stem cell transplantation-a report from the kyushu-yamaguchi children's cancer study group. Pediatric Blood Cancer 45(2):111–120. https://doi.org/10.1002/pbc.20363

McMahan CS, Wang L, Tebbs JM (2013) Regression analysis for current status data using the em algorithm. Stat Med 32(25):4452–4466. https://doi.org/10.1002/sim.5863

Muralidharan K (1999) Tests for the mixing proportion in the mixture of a degene-rate and exponential distribution. J Ind Stat Assoc 37:105–119

Muralidharan K, Lathika P (2006) Analysis of instantaneous and early failures in weibull distribution. Metrika 64(3):305–316. https://doi.org/10.1007/s00184-006-0050-2
Murthy DP, Xie M, Jiang R (2004) Weibull models, vol 505. Wiley, Hoboken
NCT01373242 (2017) Sublingual immunotherapy for peanut allergy and induction of tolerance (slit-tlc): Nct01373242. http://clinicaltrials.gov/show/NCT01373242NLMIdentifier:NCT01373242
Odell P, Anderson K, Agostino R (1992) Maximum likelihood estimation for interval-censored data using a weibull-based accelerated failure time model. Biometrics. https://doi.org/10.2307/2532360
Pan W (1999) Extending the iterative convex minorant algorithm to the cox model for interval-censored data. J Comput Graph Stat 8(1):109–120
Pan W (2000) A multiple imputation approach to cox regression with interval-censored data. Biometrics 56(1):199–203. https://doi.org/10.1111/j.0006-341X.2000.00199.x
Pham H, Lai CD (2007) On recent generalizations of the weibull distribution. IEEE Trans Reliab 56(3):454–458. https://doi.org/10.1109/TR.2007.903352
Ramsay JO (1988) Monotone regression splines in action. Stat Sci. https://doi.org/10.1214/ss/1177012761
Satten GA (1996) Rank-based inference in the proportional hazards model for interval censored data. Biometrika 83(2):355–370. https://doi.org/10.1093/biomet/83.2.355
Sun J (2007) The statistical analysis of interval-censored failure time data. Springer, Berlin
Wang L, Dunson DB (2011) Semiparametric bayes' proportional odds models for current status data with underreporting. Biometrics 67(3):1111–1118. https://doi.org/10.1111/j.1541-0420.2010.01532.x
Wang L, McMahan CS, Hudgens MG, Qureshi ZP (2016) A flexible, computationally efficient method for fitting the proportional hazards model to interval-censored data. Biometrics 72(1):222–231. https://doi.org/10.1111/biom.12389
Wang N, Wang L, McMahan CS (2015) Regression analysis of bivariate current status data under the gamma-frailty proportional hazards model using the em algorithm. Comput Stat Data Anal 83:140–150. https://doi.org/10.1016/j.csda.2014.10.013
Zhang Y, Hua L, Huang J (2010) A spline-based semiparametric maximum likelihood estimation method for the cox model with interval-censored data. Scand J Stat 37(2):338–354. https://doi.org/10.1111/j.1467-9469.2009.00680.x
Zhang Z, Sun J (2010) Interval censoring. Stat Methods Med Res 19(1):53–70. https://doi.org/10.1177/0962280209105023

Affiliations

Prabhashi W. Withana Gamage[1] · Monica Chaudari[2] · Christopher S. McMahan[3] · Edwin H. Kim[4] · Michael R. Kosorok[2]

1 Department of Mathematics & Statistics, James Madison University, Harrisonburg, VA 22807, USA

2 Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

3 School of Mathematical and Statistical Sciences, Clemson University, Clemson, SC 29634, USA

4 Division of Rheumatology, Allergy and Immunology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA