Statistics and Its Interface Volume 15 (2022) 487–501

Semiparametric transformation models of survival-out-of-hospital

Xiaowei Sun, Cheng Zeng, and Liuquan Sun*

Recurrent event data with a terminal event commonly arise in biomedical studies, and the survival-out-of-hospital process is a useful alternative framework for the analysis of recurrent/terminal event data with non-negligible event duration. In this article, we propose a class of semiparametric transformation models for the survival-out-of-hospital process, and the proposed models offer great flexibility in formulating covariate effects on the probability of survival-out-of-hospital. Estimating equation approaches are developed for the model parameters, and the asymptotic properties of the resulting estimators are established. The finite sample performance of the proposed estimators is examined through simulation studies. An application to a Centers for Medicare and Medicaid Services study is provided.

AMS 2000 subject classifications: Primary 62N01, 62N02; secondary 62F12.
Keywords and phrases: Estimating equations, Multiple imputation, Recurrent event, Terminal event, Transformation model, Survival-out-of-hospital.

1. INTRODUCTION

Recurrent event data are frequently encountered in clinical and observational studies, where each subject may potentially experience a particular event repeatedly over time [2, 16]. In many applications, there may exist a terminal event such as death that stops the follow-up, and both the recurrent and terminal events are of interest. A number of methods have been developed in the literature for analyzing recurrent/terminal event data, including marginal models and joint models. Marginal models focus on the marginal rates of the recurrent and terminal events, and the association between the recurrent and terminal events is left unspecified [3, 9, 23, 26, 36]. Joint models use frailties or random effects to account for the dependence between the recurrent and terminal events [1, 12, 14, 18, 28, 31, 32].

A frequently arising example of the recurrent/terminal event data involves hospitalization representing the recurrent event and death serving as a terminal event. Since a patient being hospitalized at some time is not at risk for a further hospitalization at the same time, one needs to adjust for hospital duration in the analysis of this type of data. However, there is limited discussion about the analysis of recurrent event data with non-negligible event duration. [35] considered semiparametric and nonparametric analysis of recurrent events with observation gaps. [11] suggested a generalized Cox regression model to accommodates non-negligible event duration using adjusted risk sets. However, these two methods did not take a terminal event into consideration.

For the analysis of recurrent/terminal event data with non-negligible event duration, we are interested in the joint event of being out-of-hospital and being alive, and consider the survival-out-of-hospital process, defined as a temporal process (indicator function) taking the value 1 when the patient is currently alive and not hospitalized, and 0 otherwise [33]. The survival-out-of-hospital process is a useful alternative framework for the analysis of recurrent/terminal event data with non-negligible event duration. In particular, this process takes quality of life information into account, and leave the dependence structure between the temporal indicator process and terminal event completely unspecified. In addition, survival-out-of-hospital may be viewed as a refinement of survival time in the study of chronic diseases, such as in a study of end-stage liver disease, where the refinement is the incorporation of the patient's hospital admission and length of stay information.

Since the survival-out-of-hospital process is framed as a temporal process by a continuous follow-up time, it seems natural to use the existing temporal process regression methods for the analysis [5, 7, 29, 30]. For example, [7] proposed functional generalized linear models, in which covariate effects are completely unspecified. [30] developed a partly functional temporal process regression model, where some covariate effects are constant while others are time-varying. In addition, [20] suggested martingale-based estimation equations to directly modeling the survival function for right censored data. However, none of these approaches can be applied directly to our setting of the survival-out-of-hospital process, along with the assumed data structure and model of interest.

Recently, [33] proposed a semiparametric multiplicative model for the survival-out-of-hospital process, in which the covariates have multiplicative effects on an unspecified baseline probability of being alive-and-out-of-hospital. However,

*Corresponding author.

this multiplicative model cannot guarantee that the estimated baseline probability is not bounded by 1. In this article, we propose a class of semiparametric transformation models for the survival-out-of-hospital process, which avoids the issue that the baseline probability is bounded by 1 by choosing the appropriate transformation function. Specifically, we do not directly estimate the baseline probability, but focus on the estimation of an unspecified baseline function (as shown in model (1) below) without requiring any constraints. The proposed transformation models are flexible and robust in that the baseline probability does not need to be specified, and define a very rich family of models through the transformation function, including the multiplicative model as a special case.

The remainder of the article is organized as follows. Section 2 describes the proposed model, and presents estimation procedures for the model parameters. The asymptotic properties of the resulting estimators are established in Section 3. Section 4 reports some simulation results for evaluating the proposed method. An application to a Centers for Medicare and Medicaid Services study is provided in Section 5, and some concluding remarks are given in Section 6. All proofs are relegated to the Appendix.

2. MODEL AND ESTIMATION PROCEDURES

Let D be the death (terminal event) time and C be the censoring time. Write $T = D \wedge C$ and $\delta = I(D \leq C)$, where $a \wedge b = \min(a, b)$ and $I(\cdot)$ is the indicator function. Let $Z(t)$ be the $p \times 1$ vector of external time-dependent covariates [13]. Let $O(t) = 1$ if a subject is in the hospital at time t, and 0 if out of hospital. The probability of interest is the probability that a subject is alive and out-of-hospital at time t, that is,

$$\pi(t) = P\{O(t) = 0, D > t | Z(t)\}.$$

Our proposed transformation models take the form

$$g\{\pi(t)\} = \pi_0(t) + \beta_0' Z(t), \tag{1}$$

where $g(\cdot)$ is pre-specified and assumed to be twice continuously differentiable and strictly monotonic, β_0 is a p-dimensional vector, and $\pi_0(t)$ is an unspecified baseline function. Clearly, model (1) defines a very rich family of models through the link function g. The obvious choices are the log-log transformation $g(x) = \log(-\log(x))$ and the logit transformation $g(x) = \log(x/(1-x))$. It also encompasses the Box–Cox transformations, in which g is given by $g(x) = [(x+1)^\rho - 1]/\rho$ $(\rho \geq 0)$, where $\rho = 0$ means that $g(x) = \log(x+1)$. When $g(x) = \log(x)$, model (1) reduces to the multiplicative model studied by [33]. If $g(x) = x$, then model (1) becomes a semiparametric additive model, in which the covariates have additive effects on the baseline probability of being alive-and-out-of-hospital.

Define $A(t) = I(D > t)$ as the survival indicator, and $A^0(t) = I(O(t) = 0, D > t)$ as the survival-out-of-hospital indicator. Let τ be the follow-up time, where τ is a prespecified constant such that $P(C \geq \tau) > 0$. In practice, τ could be chosen as the maximum of observation time T. For a random sample of n subjects, the observed data consist of $\{O_i(t), T_i, \delta_i, Z_i(s);\ 0 \leq t \leq T_i,\ 0 \leq s \leq \tau,\ i = 1, ..., n\}$, where $T_i = D_i \wedge C_i$ and $\delta_i = I(D_i \leq C_i)$.

2.1 Known censoring

We first consider the case where the censoring time C is always known for all subjects. This would be the case in a clinical trial that censoring results only from administrative loss-to-follow up, that is, patients have not failed by the time the data are analyzed. In such cases, the censoring time is always observed, even on subjects who die prior to the time of analysis [6]. Although this is inconsistent with most observational studies, it is a useful starting point for our estimation procedures. In what follows, we assume that given $Z(t)$, the censoring time C is independent of $A^0(t)$.

Define

$$M_i(t) = I(C_i \geq t)\big[A_i^0(t) - g^{-1}\{\pi_0(t) + \beta_0' Z_i(t)\}\big],\ \ i = 1, ..., n,$$

where $A_i^0(t) = I(O_i(t) = 0, D_i > t)$, and $g^{-1}(\cdot)$ is the inverse function of $g(\cdot)$. Under model (1), $M_i(t)$ is zero-mean stochastic process. Thus, for a given β, a reasonable estimator for $\pi_0(t)$ is the solution to

$$\begin{aligned}\sum_{i=1}^n I(C_i \geq t)\big[A_i^0(t) - g^{-1}\{\pi_0(t) + \beta' Z_i(t)\}\big]\\ = 0, \quad 0 \leq t \leq \tau,\end{aligned} \tag{2}$$

where τ is a prespecified constant such that $P(T_i \geq \tau) > 0$. Denote this estimator by $\hat{\pi}_0(t; \beta)$. To estimate β_0, using the generalized estimating equation approach [15], we propose the following class of estimating equations for β_0:

$$\begin{aligned}\sum_{i=1}^n \int_0^\tau I(C_i \geq t) Z_i(t)\big[A_i^0(t) - g^{-1}\{\hat{\pi}_0(t; \beta)\\ + \beta' Z_i(t)\}\big] dH(t) = 0,\end{aligned} \tag{3}$$

where $H(t)$ is an increasing and known weight function on $[0, \tau]$. Let $\hat{\beta}$ denote the solution to (3) and $\hat{\pi}_0(t) \equiv \hat{\pi}_0(t; \hat{\beta})$ the corresponding estimator of $\pi_0(t)$.

Write $\dot{g}^{-1}(x) = dg^{-1}(x)/dx$. To solve the estimating equations (2) and (3) simultaneously, we use a Taylor expansion of $g^{-1}\{\pi_0(t) + \beta' Z_i(t)\}$ around the current value of estimates $\pi_0^{(k)}(t)$ and $\beta^{(k)}$ to get approximated estimating equations

$$\begin{aligned}\sum_{i=1}^n I(C_i \geq t)\Big[A_i^0(t) - g^{-1}\{\pi_0^{(k)}(t) + \beta^{(k)'} Z_i(t)\}\\ - \dot{g}^{-1}\{\pi_0^{(k)}(t) + \beta^{(k)'} Z_i(t)\}\{\pi_0(t) - \pi_0^{(k)}(t)\}\\ - \dot{g}^{-1}\{\pi_0^{(k)}(t) + \beta^{(k)'} Z_i(t)\} Z_i(t)'\{\beta - \beta^{(k)}\}\Big] = 0,\end{aligned} \tag{4}$$

and

$$(5)\quad \sum_{i=1}^{n}\int_0^\tau I(C_i\geq t)Z_i(t)\Big[A_i^0(t)-g^{-1}\{\pi_0^{(k)}(t)+\beta^{(k)\prime}\times Z_i(t)\}-\dot g^{-1}\{\pi_0^{(k)}(t)+\beta^{(k)\prime}Z_i(t)\}\{\pi_0(t)-\pi_0^{(k)}(t)\}-\dot g^{-1}\{\pi_0^{(k)}(t)+\beta^{(k)\prime}Z_i(t)\}Z_i(t)'\{\beta-\beta^{(k)}\}\Big]dH(t)=0.$$

Solving (4) for $\pi_0(t)$ and inserting it into (5), we obtain the $(k+1)$th iterative estimator for β_0:

$$(6)\quad \begin{aligned}\beta^{(k+1)} &= \beta^{(k)}+\Big[\int_0^\tau\Big\{S^{(2)}(t;\pi_0^{(k)},\beta^{(k)})\\ &- \frac{S^{(1)}(t;\pi_0^{(k)},\beta^{(k)})^{\otimes 2}}{S^{(0)}(t;\pi_0^{(k)},\beta^{(k)})}\Big\}dH(t)\Big]^{-1}\\ &\times \int_0^\tau I(C_i\geq t)\big\{Z_i(t)-\bar Z(t;\pi_0^{(k)},\beta^{(k)})\big\}\\ &\times \big[A_i^0(t)-g^{-1}\{\pi_0^{(k)}(t)+\beta^{(k)\prime}\\ &\times Z_i(t)\}\big]dH(t),\end{aligned}$$

where $\bar Z(t;\pi_0,\beta)=S^{(1)}(t;\pi_0,\beta)/S^{(0)}(t;\pi_0,\beta)$, and

$$\begin{aligned}S^{(k)}(t;\pi_0,\beta) &= \sum_{i=1}^{n}\dot g^{-1}\{\pi_0(t)+\beta' Z_i(t)\}I(C_i\geq t)Z_i(t)^{\otimes k},\\ k &= 0,1,2.\end{aligned}$$

Here $a^{\otimes 0}=1$, $a^{\otimes 1}=a$ and $a^{\otimes 2}=aa'$ for any vector a. Applying the updated version $\beta^{(k+1)}$ and solving (5) for $\pi_0(t)$, we can get

$$(7)\quad \begin{aligned}\pi_0^{(k+1)}(t) &= \pi_0^{(k)}(t)+S^{(0)}(t;\pi_0^{(k)},\beta^{(k)})^{-1}\\ &\times \Big\{\sum_{i=1}^{n}I(C_i\geq t)\big[A_i^0(t)\\ &- g^{-1}\{\pi_0^{(k)}(t)+\beta^{(k)\prime}Z_i(t)\}\big]\\ &- S^{(1)}(t;\pi_0^{(k)},\beta^{(k)})'\{\beta^{(k+1)}-\beta^{(k)}\}\Big\}.\end{aligned}$$

This iteration is continued until convergence, and $\hat\beta$ and $\hat\pi_0(t)$ are obtained at convergence. For the convergence, several criteria can be applied and in the simulation studies below, we used the absolute differences between the iterative estimates of the parameters ($\leq 10^{-4}$). Our proposed iterative algorithms are similar to the Newton–Raphson method. If the iterative algorithms (6) and (7) converge, then they will converge to $\hat\beta$ and $\hat\pi$. In order to guarantee the convergence of the iterative algorithms, we need the condition that the matrix A is nonsingular, which is given in condition (C5) below. That is, under this condition, the solutions from the iterative algorithms are asymptotically equivalent to the solutions of (2) and (3).

The integral of the baseline survival-out-of-hospital probability is defined as

$$\Pi_0(t)=\int_0^t \pi_0(u)du,$$

which can be interpreted as the expected length of being alive and out-of-hospital up to time t for a subject with the baseline value of each covariate. We can estimate $\Pi_0(t)$ by $\hat\Pi_0(t)=\int_0^t\hat\pi_0(u)du$. In addition, as shown in model (1), the baseline probability can be expressed as $g^{-1}\{\pi_0(t)\}$, which can be estimated by $g^{-1}\{\hat\pi_0(t)\}/\max_{0\leq t\leq\tau}g^{-1}\{\hat\pi_0(t)\}$. This modification ensures that the estimated baseline probability is bounded by 1.

2.2 Random censoring

We now consider a more realistic scenario where the censoring time C is random, implying that C is unknown when the terminal event time D occurs first. In this case, one cannot use the estimating equations (2) and (3), due to C_i being missing for dead subjects. A naive method is to set the censoring time as D_i or the maximum follow-up time, which may lead to substantial bias in estimating β_0 and $\pi_0(t)$. In addition, one could use weighting techniques to recover missing censoring times, such as the inverse probability of censoring weighting and the inverse probability of survival weighting [9, 19]. However, it may be tedious to carry out weighted versions of (2) and (3), due to the time line being continuous. For this, we use an imputation approach which is easy to implement [23, 33].

We specify the proportional hazards model for the censoring time as

$$(8)\qquad \lambda^C(t|Z)=\lambda_0^C(t)\exp\{\gamma_0' Z(t)\},$$

where γ_0 is a vector of unknown regression parameters, and $\lambda_0^C(t)$ is an unspecified baseline hazard function [4]. Let $\hat\gamma$ be the maximum partial likelihood estimator of γ_0, and $\hat\Lambda_0^C(t)$ be the Breslow estimator of the baseline cumulative hazard function $\Lambda_0^C(t)=\int_0^t\lambda_0^C(u)du$ [8].

We will create M imputed datasets, and in the mth imputed data set, for a subject with $C_i\leq D_i$, we set the imputed censoring time as the known censoring time. For a subject with $C_i>D_i$, we impute $C_i^{\langle m\rangle}$ from the estimated conditional survival function:

$$\hat G_i(t;\hat\gamma)=I(D_i\leq t)\exp\big\{\hat\Lambda_i^C(D_i;\hat\gamma)-\hat\Lambda_i^C(t;\hat\gamma)\big\},$$

where $d\hat\Lambda_i^C(t;\hat\gamma)=\exp\{\hat\gamma' Z(t)\}d\hat\Lambda_0^C(t)$. For the mth imputed data set, by replacing C_i with $C_i^{\langle m\rangle}$ in the estimating equations (2) and (3), we can obtain the estimators of β_0 and $\pi_0(t)$, denoted by $\hat\beta^{\langle m\rangle}$ and $\hat\pi_0^{\langle m\rangle}(t)$, respectively ($m=1,2,...,M$). In what follows, we include $\hat\gamma$ in the uncensored indicator $I(C_i^{\langle m\rangle}\geq t;\hat\gamma)$ to emphasize that $C_i^{\langle m\rangle}$

depends on the estimator $\hat{\gamma}$. Similarly to (6), we can get the $(k+1)$th iterative estimator for β_0:

$$\begin{aligned}\beta^{(k+1)\langle m\rangle} &= \beta^{(k)\langle m\rangle} + \Big[\int_0^\tau \Big\{S^{(2)\langle m\rangle}(t;\pi_0^{(k)\langle m\rangle},\beta^{(k)\langle m\rangle},\hat{\gamma}) \\ &- \frac{S^{(1)\langle m\rangle}(t;\pi_0^{(k)\langle m\rangle},\beta^{(k)\langle m\rangle},\hat{\gamma})^{\otimes 2}}{S^{(0)\langle m\rangle}(t;\pi_0^{(k)\langle m\rangle},\beta^{(k)\langle m\rangle},\hat{\gamma})}\Big\}dH(t)\Big]^{-1} \\ &\times \int_0^\tau I(C_i^{\langle m\rangle} \geq t;\hat{\gamma})\{Z_i(t) \\ &- \bar{Z}^{\langle m\rangle}(t;\pi_0^{(k)\langle m\rangle},\beta^{(k)\langle m\rangle},\hat{\gamma})\} \\ &\times \big[A_i^0(t) - g^{-1}\{\pi_0^{(k)\langle m\rangle}(t) + \beta^{(k)\langle m\rangle\prime} \\ &\times Z_i(t)\}\big]dH(t),\end{aligned}$$

where

$$\begin{aligned}S^{(k)\langle m\rangle}(t;\pi_0,\beta,\gamma) &= \sum_{i=1}^n \dot{g}^{-1}\{\pi_0(t) + \beta' Z_i(t)\} \\ &\times I(C_i^{\langle m\rangle} \geq t;\gamma)Z_i(t)^{\otimes k},\end{aligned}$$

and

$$\bar{Z}^{\langle m\rangle}(t;\pi_0,\beta,\gamma) = \frac{S^{(1)\langle m\rangle}(t;\pi_0,\beta,\gamma)}{S^{(0)\langle m\rangle}(t;\pi_0,\beta,\gamma)}.$$

In a similar manner, we can also obtain

$$\begin{aligned}\pi_0^{(k+1)\langle m\rangle}(t) &= \pi_0^{(k)\langle m\rangle}(t) + S^{(0)\langle m\rangle}(t;\pi_0^{(k)\langle m\rangle},\beta^{(k)\langle m\rangle},\hat{\gamma})^{-1} \\ &\times \Big\{\sum_{i=1}^n I(C_i^{\langle m\rangle} \geq t;\hat{\gamma})\big[A_i^0(t) \\ &- g^{-1}\{\pi_0^{(k)\langle m\rangle}(t) + \beta^{(k)\langle m\rangle\prime} Z_i(t)\}\big] \\ &- S^{(1)\langle m\rangle}(t;\pi_0^{(k)\langle m\rangle},\beta^{(k)\langle m\rangle},\hat{\gamma})' \\ &\times \{\beta^{(k+1)\langle m\rangle} - \beta^{(k)\langle m\rangle}\}\Big\}.\end{aligned}$$

The iteration is continued until convergence, and $\hat{\beta}^{\langle m\rangle}$ and $\hat{\pi}_0^{\langle m\rangle}(t)$ are obtained at convergence for the mth imputed data set $(m = 1, 2, ..., M)$. Then β_0 and $\pi_0(t)$ can be estimated by the following pooled estimators:

$$\hat{\beta}^M = M^{-1}\sum_{m=1}^M \hat{\beta}^{\langle m\rangle}, \tag{9}$$

and

$$\hat{\pi}_0^M(t) = M^{-1}\sum_{m=1}^M \hat{\pi}_0^{\langle m\rangle}(t). \tag{10}$$

An imputation version of $\hat{\Pi}_0(t)$, denoted by $\hat{\Pi}_0^M(t)$, is obtained by integrating (10). In practice, $M = 5$ or 10 would suffice, and increasing M will generally increase precision, albeit with diminishing returns [33, 34]. The proposed method is also valid for $M = 1$, although this may be less efficient.

3. ASYMPTOTIC PROPERTIES

In order to study the asymptotic properties of the proposed estimators, we need the following regularity conditions:

(C1) $\{O_i(\cdot),\ T_i,\ \delta_i,\ Z_i(\cdot)\}$ $(i = 1, ..., n)$ are independent and identically distributed.
(C2) $P\{C_i \geq \tau\} > 0$.
(C3) $Z_i(t)$ is almost surely of bounded variation on $[0, \tau]$.
(C4) $H(t)$ converges almost surely to a nonrandom and bounded function $\tilde{H}(t)$ uniformly in $t \in [0, \tau]$.
(C5) The matrix A is nonsingular, where

$$A = \int_0^\tau \Big[s^{(2)}(t) - \frac{s^{(1)}(t)^{\otimes 2}}{s^{(0)}(t)}\Big]d\tilde{H}(t),$$

and $s^{(k)}(t)$ are the limits of $S^{(k)}(t;\pi_0,\beta)$ $(k = 0, 1, 2)$.

The asymptotic properties of the proposed estimators are summarized in the following theorems with the proof given in the Appendix. We begin by describing the asymptotic properties of $\hat{\beta}$ and $\hat{\pi}_0(t)$.

Theorem 1. *Under the regularity conditions (C1)–(C5), $\hat{\beta}$ and $\hat{\pi}_0(t)$ exist and are unique. Moreover, $\hat{\beta}$ is strongly consistent to β_0, and $\hat{\pi}_0(t)$ converges almost surely to $\pi_0(t)$ uniformly in $t \in [0, \tau]$.*

Theorem 2. *Under the regularity conditions (C1)–(C5), we have*

(i) $n^{1/2}(\hat{\beta} - \beta_0)$ converges in distribution to a zero-mean normal random vector with a covariance matrix that can be consistently estimated by $\hat{A}^{-1}\hat{\Sigma}\hat{A}^{-1}$, where $\hat{\Sigma} = n^{-1}\sum_{i=1}^n \hat{\xi}_i^{\otimes 2}$,

$$\begin{aligned}\hat{\xi}_i &= \int_0^\tau I(C_i \geq t)\{Z_i(t) - \bar{Z}(t;\hat{\pi}_0,\hat{\beta})\}\big[A_i^0(t) \\ &- g^{-1}\{\hat{\pi}_0(t) + \hat{\beta}' Z_i(t)\}\big]dH(t),\end{aligned}$$

and

$$\hat{A} = \int_0^\tau \Big[S^{(2)}(t;\hat{\pi}_0,\hat{\beta}) - \frac{S^{(1)}(t;\hat{\pi}_0,\hat{\beta})^{\otimes 2}}{S^{(0)}(t;\hat{\pi}_0,\hat{\beta})}\Big]dH(t).$$

(ii) $n^{1/2}\{\hat{\pi}_0(t) - \pi_0(t)\}$ converges weakly on $[0, \tau]$ to a zero-mean Gaussian process whose covariance function at (t, s) can be consistently estimated by $\hat{\Gamma}(t, s) = n^{-1}\sum_{i=1}^n \hat{\phi}_i(t)\hat{\phi}_i(s)$, where

$$\begin{aligned}\hat{\phi}_i(t) &= S^{(0)}(t;\hat{\pi}_0,\hat{\beta})^{-1}I(C_i \geq t)\big[A_i^0(t) \\ &- g^{-1}\{\hat{\pi}_0(t) + \hat{\beta}' Z_i(t)\}\big] - \bar{Z}(t;\hat{\pi}_0,\hat{\beta})'\hat{A}^{-1}\hat{\xi}_i.\end{aligned}$$

The asymptotic properties of $\hat{\beta}^M$ and $\hat{\pi}_0^M(t)$ are given in the next two theorems.

Theorem 3. *Under the regularity conditions (C1)–(C5), $\hat{\beta}^M$ and $\hat{\pi}_0^M(t)$ exist and are unique. Moreover, $\hat{\beta}^M$ is strongly consistent to β_0, and $\hat{\pi}_0^M(t)$ converges almost surely to $\pi_0(t)$ uniformly in $t \in [0, \tau]$.*

Theorem 4. *Under the regularity conditions (C1)–(C5), we have*

(i) $n^{1/2}(\hat{\beta}^M - \beta_0)$ *converges in distribution to a zero-mean normal random vector with a variance matrix that can be consistently estimated by* $(\hat{A}^M)^{-1}\hat{\Sigma}^M(\hat{A}^M)^{-1}$, *where* $\hat{\Sigma}^M = n^{-1}\sum_{i=1}^n \hat{\xi}_i^{M\otimes 2}$,

$$
\begin{aligned}
\hat{\xi}_i^M &= M^{-1}\sum_{m=1}^M \int_0^\tau I(C_i^{\langle m\rangle} \geq t; \hat{\gamma})\big\{Z_i(t) \\
&- \bar{Z}^{\langle m\rangle}(t; \hat{\pi}_0^{\langle m\rangle}, \hat{\beta}^{\langle m\rangle}, \hat{\gamma})\big\}\big[A_i^0(t) \\
&- g^{-1}\{\hat{\pi}_0^{\langle m\rangle}(t) + \hat{\beta}^{\langle m\rangle\prime} Z_i(t)\}\big]\, dH(t),
\end{aligned}
$$

and

$$
\begin{aligned}
\hat{A}^M &= M^{-1}\sum_{m=1}^M \int_0^\tau \Big[S^{(2)\langle m\rangle}(t; \hat{\pi}_0^{\langle m\rangle}, \hat{\beta}^{\langle m\rangle}, \hat{\gamma}) \\
&- \frac{S^{(1)\langle m\rangle}(t; \hat{\pi}_0^{\langle m\rangle}, \hat{\beta}^{\langle m\rangle}, \hat{\gamma})^{\otimes 2}}{S^{(0)\langle m\rangle}(t; \hat{\pi}_0^{\langle m\rangle}, \hat{\beta}^{\langle m\rangle}, \hat{\gamma})}\Big]\, dH(t).
\end{aligned}
$$

(ii) $n^{1/2}\{\hat{\pi}_0^M(t) - \pi_0(t)\}$ *converges weakly on* $[0, \tau]$ *to a zero-mean Gaussian process whose covariance function at* (t, s) *can be consistently estimated by* $\hat{\Gamma}^M(t, s) = n^{-1}\sum_{i=1}^n \hat{\phi}_i^M(t)\hat{\phi}_i^M(s)$, *where*

$$
\begin{aligned}
\hat{\phi}_i^M(t) &= M^{-1}\sum_{m=1}^M S^{(0)\langle m\rangle}(t; \hat{\pi}_0^{\langle m\rangle}, \hat{\beta}^{\langle m\rangle}, \hat{\gamma})^{-1} \\
&\times I(C_i^{\langle m\rangle} \geq t; \hat{\gamma})\big[A_i^0(t) - g^{-1}\{\hat{\pi}_0^{\langle m\rangle}(t) \\
&+ \hat{\beta}^{\langle m\rangle\prime} Z_i(t)\}\big] - M^{-1} \\
&\times \sum_{m=1}^M \bar{Z}^{\langle m\rangle}(t; \hat{\pi}_0^{\langle m\rangle}, \hat{\beta}^{\langle m\rangle}, \hat{\gamma})'(\hat{A}^M)^{-1}\hat{\xi}_i^M.
\end{aligned}
$$

4. SIMULATION STUDIES

We conducted simulation studies to examine the finite sample properties of the estimators. In the study, the covariate Z was generated from a Bernoulli distribution with success probability 0.5. The model for the survival-out-of-hospital process was taken as $g\{\pi(t)\} = \pi_0(t) + \beta_0' Z$, where $\pi_0(t) = 0.5 - \log(t)$ or $\pi_0(t) = 0.5 - 0.01t$, and $\beta_0 = 0$ or 0.5. The death time D was generated by the hazard function $\lambda^D(t) = \lambda_0^D(t)\exp\{\theta_0 Z\}$, where $\lambda_0^D(t) = 0.3$ and $\theta_0 = 0.3$. The censoring time C was generated from the hazard function $\lambda^C(t) = \lambda_0^C(t)\exp\{\gamma_0 Z\}$, where $\lambda_0^C(t) = 0.5$ and $\gamma_0 = 0.3$. We considered two choices for g: the log-log transformation $g_1(x) = \log(-\log(x))$ and the logit transformation $g_2(x) = \log(x/(1-x))$. In all simulations, we considered two choices for the weight functions: $H_1(t) = t$ and $H_2(t) = n^{-1}\sum_{i=1}^n A_i^0(t)$. Under the preceding settings, the censoring rate is about 60%. In addition to the known censoring case, we evaluate the performance of the proposed imputation method with $M = 1$ and 5. The results presented below are based on 1000 replications with sample sizes $n = 100$ and 200, and final estimates were reached at convergence.

Tables 1–3 present the simulation results on the estimate of β_0 under $g_1(x)$ and $g_2(x)$ with $M = 1$ and 5, respectively. In these tables, Bias is the sample mean of the estimate minus the true value, SE is the sampling standard error of the estimate, SEE is the sample mean of the standard error estimate, and CP is the 95% empirical coverage probability based on the normal approximation. It can be seen from Tables 1 and 2 that the proposed estimators were nearly unbiased, indicating that our estimators were consistent even based on $M = 1$. Also, there was a good agreement between the estimated and the empirical standard errors, and the 95% empirical coverage probabilities were reasonable. The results given in Table 3 show that the proposed method worked efficiently based on $M = 5$, and multiple imputation was more efficient than single imputation. In general, increasing M will decrease the estimated standard errors, although with diminishing returns. This was also demonstrated empirically in [33, 34]. In addition, for the above situations, the results are similar for the two weight functions $H_1(t)$ and $H_2(t)$, and the results become better when the sample size increases from 100 to 200.

Tables 4–7 present the simulation results on estimation of $\pi_0(t)$ at time points $t = 1, 2,$ and 3 under $g_1(x)$ and $g_2(x)$ with $M = 1$. These simulation results suggest that the proposed method still performed well and essentially provided unbiased estimates for $\pi_0(t)$. The asymptotic standard errors present a reasonable description of the variability for the proposed estimators, and the coverage probabilities of the 95% confidence intervals were close to the nominal level. We also considered other setups including $M = 5$, and the results were similar to those given above.

5. AN APPLICATION

In this section, we applied the proposed method to a dataset from the Centers for Medicare and Medicaid Services (CMS) [10, 24, 25]. CMS is the agency within United States Department of Health and Human Services that manages the major health care plans in the United States. CMS supervises programs including Medicare, Medicaid, the Children's Health Insurance Program and the state and federal health insurance marketplaces. CMS collects and analyzes data, generates research reports, and works to eliminate instances of fraud and abuse within the healthcare system. The CMS dataset used in this study were collected between

Table 1. *Simulation results for the estimation of β_0 under $g_1(x)$ with $M = 1$*

Censoring	n	$\pi_0(t)$	β_0	Bias	SE	SEE	CP
			$H_1(t)$				
C known	100	$0.5 - \log(t)$	0	-0.0015	0.0923	0.0886	0.944
			0.5	0.0246	0.0887	0.0822	0.925
		$0.5 - 0.01t$	0	-0.0033	0.0620	0.0616	0.938
			0.5	0.0363	0.0692	0.0684	0.928
	200	$0.5 - \log(t)$	0	0.0002	0.0674	0.0587	0.948
			0.5	0.0169	0.0636	0.0605	0.932
		$0.5 - 0.01t$	0	0.0012	0.0473	0.0459	0.950
			0.5	0.0185	0.0496	0.0497	0.934
C random	100	$0.5 - \log(t)$	0	-0.0019	0.0791	0.0779	0.937
			0.5	0.0237	0.0908	0.0838	0.930
		$0.5 - 0.01t$	0	0.0013	0.0638	0.0619	0.943
			0.5	0.0350	0.0693	0.0684	0.926
	200	$0.5 - \log(t)$	0	-0.0010	0.0648	0.0585	0.937
			0.5	0.0091	0.0631	0.0611	0.938
		$0.5 - 0.01t$	0	-0.0002	0.0468	0.0459	0.938
			0.5	0.0171	0.0501	0.0494	0.930
			$H_2(t)$				
C known	100	$0.5 - \log(t)$	0	-0.0012	0.0934	0.0871	0.936
			0.5	0.0334	0.0937	0.0910	0.927
		$0.5 - 0.01t$	0	0.0005	0.0674	0.0641	0.936
			0.5	0.0386	0.0776	0.0717	0.933
	200	$0.5 - \log(t)$	0	-0.0013	0.0701	0.0655	0.936
			0.5	0.0106	0.0702	0.0678	0.941
		$0.5 - 0.01t$	0	0.0006	0.0481	0.0472	0.948
			0.5	0.0239	0.0540	0.0525	0.930
C random	100	$0.5 - \log(t)$	0	0.0011	0.0935	0.0865	0.930
			0.5	0.0219	0.0942	0.0903	0.931
		$0.5 - 0.01t$	0	0.0021	0.0692	0.0637	0.930
			0.5	0.0385	0.0726	0.0717	0.928
	200	$0.5 - \log(t)$	0	0.0027	0.0708	0.0641	0.931
			0.5	0.0079	0.0709	0.0682	0.952
		$0.5 - 0.01t$	0	0.0013	0.0463	0.0476	0.950
			0.5	0.0216	0.0554	0.0524	0.930

2008 and 2010, and our research interests include identifying clinical variables associated with the survival-out-of-hospital and describing their effects on the survival-out-of-hospital process.

We restricted the sample for the CMS dataset to $n = 298$ patients with multiple hospitalization records, and the average number of hospitalization per patient ranged from 2 to 3. The covariates included gender, as well as the following list of chronic conditions: Alzheimer's or related disorders or senility, heart failure, cancer, chronic obstructive pulmonary disease (COPD) and diabetes. Since the hospitalization and death times were recorded in days, t represents days (that is, days after entering the CMS) in our analysis. The average hospitalization days of patients was 12.1 days, about 10% of the subjects died during the observation (that is, for them $C_i > D_i$). We applied the proposed method to estimate the regression parameters with the weight function $H(t) = t$ and $M = 1$ and 5, and the analysis results are summarized in Table 8, where the link functions $g_1(x)$ and $g_2(x)$ are given in Section 4.

From Table 8, it can be seen that the covariate effects had the opposite sign under $g_1(x)$ and $g_2(x)$. This was because for $x \in (0, 1)$, $g_1(x)$ is a decreasing function, while $g_2(x)$ an increasing function. These results showed that Alzheimer's, heart failure and COPD had significant negative effects on the survival-out-of-hospital process. That is, the patients with Alzheimer's, heart failure and COPD tended to have smaller survival-out-of-hospital probability. It is only worth noting that the parameters have different interpretations for different $g(x)$, and the covariate effects have different sizes for various choices of $g(x)$. In addition, as in the simulation studies, under the same link function, the standard error estimates of the covariate effects based on $M = 5$ were smaller than those based on $M = 1$, while the estimates of the covariate effects were comparable.

In order to examine which link function fit the data better, we used the Akaike's information criterion (AIC)

Table 2. Simulation results for the estimation of β_0 under $g_2(x)$ with $M = 1$

Censoring	n	$\pi_0(t)$	β_0	Bias	SE	SEE	CP
			$H_1(t)$				
C known	100	$0.5 - \log(t)$	0	-0.0058	0.1120	0.1091	0.933
			0.5	0.0203	0.1187	0.1123	0.930
		$0.5 - 0.01t$	0	0.0004	0.1063	0.1020	0.933
			0.5	0.0123	0.1102	0.1069	0.943
	200	$0.5 - \log(t)$	0	-0.0022	0.0819	0.0826	0.948
			0.5	0.0108	0.0826	0.0844	0.956
		$0.5 - 0.01t$	0	-0.0031	0.0776	0.0751	0.944
			0.5	0.0137	0.0820	0.0792	0.939
C random	100	$0.5 - \log(t)$	0	0.0027	0.1102	0.1101	0.953
			0.5	0.0261	0.1182	0.1125	0.930
		$0.5 - 0.01t$	0	-0.0016	0.1054	0.1010	0.943
			0.5	0.0167	0.1134	0.1056	0.933
	200	$0.5 - \log(t)$	0	-0.0004	0.0863	0.0823	0.937
			0.5	0.0083	0.0867	0.0842	0.934
		$0.5 - 0.01t$	0	0.0014	0.0786	0.0753	0.937
			0.5	0.0057	0.0798	0.0791	0.947
			$H_2(t)$				
C known	100	$0.5 - \log(t)$	0	0.0088	0.1158	0.1103	0.950
			0.5	0.0272	0.1192	0.1166	0.930
		$0.5 - 0.01t$	0	0.0020	0.1091	0.1031	0.939
			0.5	0.0129	0.1167	0.1098	0.937
	200	$0.5 - \log(t)$	0	0.0002	0.0817	0.0839	0.950
			0.5	0.0081	0.0902	0.0877	0.934
		$0.5 - 0.01t$	0	0.0033	0.0809	0.0783	0.954
			0.5	0.0124	0.0825	0.0818	0.932
C random	100	$0.5 - \log(t)$	0	0.0055	0.1172	0.1118	0.933
			0.5	0.0248	0.1161	0.1156	0.931
		$0.5 - 0.01t$	0	-0.0031	0.1102	0.1041	0.937
			0.5	0.0119	0.1182	0.1096	0.931
	200	$0.5 - \log(t)$	0	-0.0010	0.0883	0.0835	0.931
			0.5	0.0070	0.0909	0.0856	0.940
		$0.5 - 0.01t$	0	-0.0011	0.0772	0.0778	0.952
			0.5	0.0093	0.0897	0.0818	0.930

and the Bayesian information criterion (BIC) to compare the performances of the models with the two link functions, where $\text{AIC} = 2m/n + \log(\text{RSS}/n)$, $\text{BIC} = m\log(n)/n + \log(\text{RSS}/n)$, m is the number of parameters, $\text{RSS} = \sum_{i=1}^n r_i^2$, and r_i is the residual for each subject i with

$$\begin{aligned} r_i &= M^{-1}\sum_{m=1}^{M}\int_0^\tau I(C_i^{\langle m\rangle} \geq t)\big[A_i^0(t) \\ &- g^{-1}\{\hat{\pi}_0^{\langle m\rangle}(t) + \hat{\beta}^{\langle m\rangle\prime} Z_i(t)\}\big]dH(t). \end{aligned}$$

The results are presented in Table 9. It can be seen that under AIC and BIC, the model with the link function $g_1(x)$ fits the data better for $M = 1$ or 5.

6. DISCUSSION

In this article, we proposed a class of semiparametric transformation models for the survival-out-of-hospital process, which includes the multiplicative model as a special case. The proposed transformation models offer great flexibility in formulating covariate effects on the probability of survival-out-of-hospital, and can avoid the issue that the baseline probability is bounded by 1. Estimating equation approaches were developed to obtain consistent and asymptotically normal estimators, in which multiple imputation was used to recover missing censoring times for dead subjects. The simulation studies suggested that the proposed method performs well. An application to the CMS dataset was provided to illustrate our method.

For the case of random censoring, the proposed approach requires modeling the censoring distribution, and we have used the proportional hazards model for the censoring time. Other competing models, such as the additive hazards model, the accelerated failure time model, and the linear transformation model may be used as well. It would be worthwhile to investigate the potential bias due to misspecification for each of these models both analytically and

Table 3. Simulation results for the estimation of β_0 with $M = 5$

g	n	$\pi_0(t)$	β_0	Bias	SE	SEE	CP
				$H_1(t)$			
$g_1(x)$	100	$0.5 - \log(t)$	0	-0.0037	0.0736	0.0697	0.932
			0.5	0.0223	0.0786	0.0720	0.930
		$0.5 - 0.01t$	0	0.0028	0.0606	0.0575	0.948
			0.5	0.0337	0.0639	0.0634	0.933
	200	$0.5 - \log(t)$	0	0.0042	0.0560	0.0526	0.930
			0.5	0.0113	0.0596	0.0534	0.935
		$0.5 - 0.01t$	0	0.0012	0.0435	0.0425	0.941
			0.5	0.0160	0.0445	0.0458	0.940
$g_2(x)$	100	$0.5 - \log(t)$	0	-0.0023	0.1046	0.1022	0.950
			0.5	0.0210	0.1102	0.1058	0.932
		$0.5 - 0.01t$	0	0.0011	0.0965	0.0939	0.936
			0.5	0.0204	0.1068	0.1003	0.935
	200	$0.5 - \log(t)$	0	0.0011	0.0797	0.0758	0.933
			0.5	0.0127	0.0808	0.0781	0.935
		$0.5 - 0.01t$	0	-0.0005	0.0737	0.0698	0.943
			0.5	0.0084	0.0776	0.0742	0.934
				$H_2(t)$			
$g_1(x)$	100	$0.5 - \log(t)$	0	-0.0080	0.0813	0.0762	0.938
			0.5	0.0264	0.0831	0.0798	0.930
		$0.5 - 0.01t$	0	-0.0026	0.0614	0.0596	0.931
			0.5	0.0436	0.0704	0.0674	0.929
	200	$0.5 - \log(t)$	0	0.0001	0.0583	0.0571	0.954
			0.5	0.0090	0.0624	0.0579	0.930
		$0.5 - 0.01t$	0	-0.0002	0.0452	0.0443	0.952
			0.5	0.0223	0.0475	0.0458	0.930
$g_2(x)$	100	$0.5 - \log(t)$	0	-0.0047	0.1107	0.1037	0.934
			0.5	0.0270	0.1066	0.1053	0.935
		$0.5 - 0.01t$	0	-0.0026	0.1010	0.0967	0.931
			0.5	0.0063	0.1018	0.1015	0.962
	200	$0.5 - \log(t)$	0	0.0011	0.0809	0.0771	0.934
			0.5	0.0105	0.0774	0.0763	0.965
		$0.5 - 0.01t$	0	-0.0064	0.0768	0.0727	0.936
			0.5	0.0117	0.0815	0.0766	0.940

numerically. In addition, the proposed method assumed independent censoring, and a temporal process may be subject to dependent censoring [34]. It would be desirable to extend the proposed method to allow for both independent and dependent censoring.

In practice, the choice of an appropriate link function g may be based on prior data or the desiring interpretation of the regression parameters [17]. Note that the magnitudes of the parameter estimates are quite different for various choices of g. This is because the parameters have different interpretations for different g. In order to examine which g fits the data best in application, we may use the Akaike's information criterion and the Bayesian information criterion to compare the performances of the models with different link functions. However, it is difficult to find a data-driven method to estimate the link function. This is a challenging problem and requires further research efforts.

In the estimating equation approach, an interesting issue is the effect of the weight $H(t)$. Ideally, we would choose $H(t)$ to minimize the variances of the proposed estimators. However, it does not appear possible to derive an optimal weight without specification of the dependence structure on the increments of the survival-out-of-hospital process. The choice of an optimal weight function is usually a complicated problem [17], and developing a simple but more efficient estimation procedure requires further research efforts.

APPENDIX: PROOFS OF ASYMPTOTIC PROPERTIES

Proof of Theorem 1. Without loss of generality, assume that $g(\cdot)$ is strictly increasing. Because $I(C_i \geq t)[A_i^0(t) - g^{-1}\{\pi_0(t) + s + \beta' Z_i(t)\}]$ $(i = 1, ..., n)$ can be written as sums or products of monotone functions in t, s, and all components of β, it can be shown that the processes are manageable [21]. It then follows from the uniform strong law of

Table 4. Simulation results for the estimation of $\pi_0(t)$ under $g_1(x)$ and with $n = 100$ and $M = 1$

Censoring	$\pi_0(t)$	β_0	t	Bias	SE	SEE	CP	Bias	SE	SEE	CP
				$H_1(t)$				$H_2(t)$			
C known	$0.5 - \log(t)$	0	1	0.0256	0.1839	0.1731	0.945	0.0031	0.1831	0.1738	0.938
			2	-0.0153	0.2727	0.2541	0.940	-0.0092	0.2657	0.2537	0.950
			3	0.0572	0.4163	0.3905	0.953	-0.0042	0.4082	0.3904	0.952
		0.5	1	-0.0221	0.1936	0.1837	0.960	0.0271	0.1901	0.1840	0.952
			2	0.0068	0.2522	0.2457	0.949	0.0163	0.2449	0.2448	0.951
			3	0.0500	0.3814	0.3663	0.943	-0.0572	0.3938	0.3678	0.954
	$0.5 - 0.01t$	0	1	0.0173	0.1711	0.1705	0.960	0.0077	0.1761	0.1713	0.945
			2	0.0142	0.2449	0.2277	0.945	0.0219	0.2377	0.2289	0.957
			3	0.0271	0.3317	0.3030	0.946	0.0376	0.3391	0.3025	0.956
		0.5	1	-0.0013	0.1883	0.1831	0.959	0.0211	0.1965	0.1838	0.951
			2	0.0202	0.2743	0.2416	0.956	0.0263	0.2819	0.2433	0.953
			3	0.0633	0.3221	0.3072	0.938	0.0840	0.3489	0.3069	0.928
C random	$0.5 - \log(t)$	0	1	0.0418	0.1754	0.1724	0.953	0.0587	0.1816	0.1743	0.950
			2	-0.0009	0.2584	0.2529	0.942	-0.0076	0.2640	0.2538	0.939
			3	-0.0266	0.4036	0.3881	0.954	-0.0324	0.4046	0.3931	0.947
		0.5	1	-0.0235	0.2022	0.1841	0.949	0.0309	0.1976	0.1844	0.945
			2	0.0159	0.2495	0.2456	0.943	0.0188	0.2521	0.2455	0.954
			3	-0.0543	0.3838	0.3698	0.955	-0.0078	0.3864	0.3665	0.956
	$0.5 - 0.01t$	0	1	0.0040	0.1767	0.1706	0.944	0.0047	0.1787	0.1714	0.946
			2	-0.0091	0.2351	0.2286	0.950	0.0235	0.2412	0.2290	0.944
			3	0.0315	0.3277	0.3041	0.947	0.0506	0.3493	0.3051	0.944
		0.5	1	0.0014	0.1849	0.1827	0.960	-0.0004	0.1916	0.1843	0.938
			2	0.0242	0.2727	0.2424	0.950	0.0166	0.2673	0.2449	0.956
			3	0.0515	0.3104	0.3087	0.932	0.0643	0.3419	0.3056	0.930

Table 5. Simulation results for the estimation of $\pi_0(t)$ under $g_1(x)$ and with $n = 200$ and $M = 1$

Censoring	$\pi_0(t)$	β_0	t	Bias	SE	SEE	CP	Bias	SE	SEE	CP
				$H_1(t)$				$H_2(t)$			
C known	$0.5 - \log(t)$	0	1	0.0416	0.1311	0.1214	0.940	0.0993	0.1271	0.1223	0.956
			2	-0.0012	0.1843	0.1767	0.943	-0.0014	0.1908	0.1772	0.930
			3	-0.0265	0.2868	0.2708	0.947	0.0116	0.2884	0.2691	0.940
		0.5	1	0.0487	0.1366	0.1290	0.948	-0.0198	0.1217	0.1285	0.957
			2	-0.0052	0.1728	0.1716	0.958	0.0216	0.1712	0.1710	0.954
			3	-0.0431	0.2580	0.2536	0.960	-0.0150	0.2661	0.2599	0.940
	$0.5 - 0.01t$	0	1	-0.0001	0.1219	0.1201	0.948	-0.0034	0.1260	0.1206	0.937
			2	0.0168	0.1688	0.1596	0.932	0.0056	0.1692	0.1603	0.948
			3	0.0150	0.2169	0.2129	0.947	0.0030	0.2342	0.2141	0.941
		0.5	1	0.0031	0.1291	0.1280	0.955	-0.0041	0.1278	0.1213	0.961
			2	0.0068	0.1754	0.1698	0.956	0.0025	0.1779	0.1691	0.940
			3	0.0246	0.2507	0.2349	0.950	0.0344	0.2491	0.2261	0.945
C random	$0.5 - \log(t)$	0	1	-0.0069	0.1279	0.1211	0.938	-0.0230	0.1296	0.1218	0.933
			2	-0.0247	0.1824	0.1769	0.942	0.0114	0.1852	0.1775	0.950
			3	0.0109	0.2773	0.2659	0.934	-0.0273	0.2701	0.2691	0.956
		0.5	1	-0.0286	0.1384	0.1285	0.942	0.0508	0.1380	0.1292	0.950
			2	0.0074	0.1823	0.1718	0.931	-0.0045	0.1725	0.1720	0.948
			3	-0.0271	0.2658	0.2545	0.948	0.0068	0.2496	0.2530	0.953
	$0.5 - 0.01t$	0	1	0.0006	0.1209	0.1202	0.952	0.0059	0.1210	0.1201	0.950
			2	0.0102	0.1707	0.1608	0.932	0.0118	0.1572	0.1591	0.963
			3	0.0179	0.2242	0.2130	0.940	0.0135	0.2186	0.2133	0.955
		0.5	1	0.0007	0.1284	0.1274	0.957	-0.0032	0.1278	0.1280	0.951
			2	0.0136	0.1731	0.1695	0.954	0.0020	0.1711	0.1692	0.948
			3	0.0280	0.2335	0.2250	0.953	0.0307	0.2528	0.2274	0.944

Table 6. Simulation results for the estimation of $\pi_0(t)$ under $g_2(x)$ and $n = 100$ and $M = 1$

Censoring	$\pi_0(t)$	β_0	t	Bias	SE	SEE	CP	Bias	SE	SEE	CP
				$H_1(t)$				$H_2(t)$			
C known	$0.5 - \log(t)$	0	1	0.0388	0.2960	0.2844	0.950	0.0535	0.2952	0.2861	0.952
			2	-0.0220	0.3914	0.3713	0.943	-0.0404	0.3847	0.3732	0.959
			3	-0.0530	0.5499	0.5288	0.960	-0.0364	0.5532	0.5271	0.961
		0.5	1	-0.0371	0.2984	0.2959	0.955	0.0431	0.3205	0.3016	0.942
			2	0.0009	0.3776	0.3716	0.948	-0.0007	0.3977	0.3713	0.954
			3	-0.0744	0.5657	0.5217	0.959	-0.0475	0.5557	0.5216	0.959
	$0.5 - 0.01t$	0	1	0.0023	0.2932	0.2836	0.956	0.0188	0.2794	0.2840	0.955
			2	0.0391	0.3801	0.3809	0.953	0.0006	0.4004	0.3810	0.941
			3	0.0288	0.5201	0.5135	0.960	0.0462	0.5737	0.5385	0.956
		0.5	1	0.0151	0.3028	0.2948	0.945	0.0035	0.2975	0.2942	0.957
			2	0.0142	0.4108	0.3954	0.947	0.0376	0.4197	0.3959	0.954
			3	0.0334	0.5868	0.5481	0.958	0.0829	0.5843	0.5481	0.958
C random	$0.5 - \log(t)$	0	1	0.0299	0.2966	0.2849	0.946	0.0447	0.2809	0.2854	0.958
			2	0.0154	0.3763	0.3688	0.950	-0.0127	0.3766	0.3729	0.948
			3	-0.0427	0.5497	0.5235	0.960	-0.0683	0.5719	0.5422	0.960
		0.5	1	0.0102	0.2958	0.2954	0.949	0.0320	0.3018	0.2967	0.938
			2	-0.0029	0.3747	0.3728	0.951	-0.0056	0.3831	0.3731	0.956
			3	-0.0096	0.5151	0.5162	0.959	-0.0536	0.5579	0.5224	0.956
	$0.5 - 0.01t$	0	1	0.0102	0.2831	0.2836	0.954	0.0220	0.2817	0.2839	0.956
			2	0.0054	0.3876	0.3810	0.953	0.0106	0.3964	0.3809	0.952
			3	0.0373	0.5391	0.5161	0.960	0.0684	0.5796	0.5296	0.954
		0.5	1	0.0208	0.2861	0.2948	0.957	0.0371	0.3023	0.2958	0.953
			2	0.0049	0.4018	0.3942	0.958	0.0040	0.4099	0.3941	0.955
			3	0.0271	0.5583	0.5347	0.955	0.0143	0.5674	0.5322	0.945

Table 7. Simulation results for the estimation of $\pi_0(t)$ under $g_2(x)$ and $n = 200$ and $M = 1$

Censoring	$\pi_0(t)$	β_0	t	Bias	SE	SEE	CP	Bias	SE	SEE	CP
				$H_1(t)$				$H_2(t)$			
C known	$0.5 - \log(t)$	0	1	-0.0286	0.2123	0.2006	0.941	0.0010	0.2106	0.2009	0.934
			2	-0.0414	0.2660	0.2592	0.951	-0.0202	0.2551	0.2589	0.952
			3	-0.0147	0.3805	0.3625	0.942	-0.0010	0.3486	0.3497	0.960
		0.5	1	0.0455	0.2164	0.2072	0.950	0.0445	0.2094	0.2078	0.950
			2	-0.0111	0.2612	0.2601	0.954	0.0062	0.2693	0.2610	0.940
			3	0.0028	0.3645	0.3556	0.948	-0.0048	0.3650	0.3574	0.931
	$0.5 - 0.01t$	0	1	0.0130	0.2006	0.2000	0.946	-0.0175	0.1881	0.1890	0.959
			2	0.0214	0.2699	0.2659	0.956	0.0115	0.2629	0.2647	0.952
			3	0.0161	0.3644	0.3568	0.953	-0.0002	0.3439	0.3448	0.963
		0.5	1	0.0007	0.2121	0.2072	0.952	0.0177	0.2157	0.2080	0.950
			2	0.0127	0.2780	0.2747	0.953	0.0162	0.2814	0.2749	0.940
			3	0.0170	0.3653	0.3659	0.960	0.0326	0.3850	0.3672	0.944
C random	$0.5 - \log(t)$	0	1	0.0453	0.1991	0.2005	0.951	0.0390	0.2185	0.2001	0.942
			2	-0.0206	0.2617	0.2601	0.958	-0.0117	0.2686	0.2585	0.950
			3	-0.0018	0.3746	0.3615	0.956	-0.0183	0.3721	0.3615	0.963
		0.5	1	0.0150	0.2166	0.2074	0.946	0.0367	0.2242	0.2076	0.942
			2	0.0027	0.2683	0.2600	0.951	-0.0281	0.2603	0.2593	0.956
			3	-0.0456	0.3624	0.3563	0.953	-0.0085	0.3696	0.3547	0.960
	$0.5 - 0.01t$	0	1	0.0027	0.2018	0.1993	0.951	0.0244	0.2011	0.2003	0.956
			2	0.0045	0.2663	0.2651	0.948	0.0273	0.2691	0.2656	0.954
			3	0.0071	0.3507	0.3543	0.959	0.0465	0.3776	0.3558	0.942
		0.5	1	0.0128	0.2071	0.2069	0.954	0.0184	0.2083	0.2073	0.960
			2	0.0146	0.2731	0.2743	0.958	0.0187	0.2984	0.2757	0.936
			3	0.0070	0.3722	0.3654	0.952	0.0197	0.3881	0.3695	0.942

Table 8. Analysis results for the CMS data: covariate effects on survival-out-of-hospital

M	Covariate	Est	SE	p-value	Est	SE	p-value
			$g_1(x)$			$g_2(x)$	
$M=1$	Gender	-0.5780	0.4048	0.1533	0.1921	0.2090	0.3580
	Alzheimer's	1.1442	0.4150	0.0058	-0.4175	0.2759	0.1302
	Heart failure	1.8619	0.6545	0.0044	-0.7518	0.2325	0.0012
	Cancer	0.2835	0.4878	0.5611	-0.1096	0.2591	0.6723
	COPD	2.1958	0.4784	< 0.0001	-0.8267	0.2130	< 0.0001
	Diabetes	-0.5879	0.5491	0.2843	0.1743	0.3155	0.5806
$M=5$	Gender	-0.5055	0.3878	0.1924	0.1915	0.1950	0.3261
	Alzheimer's	1.0019	0.3959	0.0114	-0.4144	0.2539	0.1027
	Heart failure	1.6782	0.6227	0.0070	-0.7496	0.2166	0.0005
	Cancer	0.2403	0.4680	0.6076	-0.1084	0.2423	0.6546
	COPD	1.9569	0.4535	< 0.0001	-0.8210	0.1984	< 0.0001
	Diabetes	-0.5135	0.5276	0.3304	0.1734	0.2939	0.5552

Note: Est is the estimate of the parameter, and SE is the standard error estimate

Table 9. Model checking for the CMS data under different link functions

g	M	AIC	BIC
$g_1(x)$	$M=1$	3.1500	3.1819
	$M=5$	3.2071	3.2391
$g_2(x)$	$M=1$	3.2245	3.2563
	$M=5$	3.2816	3.3135

large numbers [21] that for any $\epsilon > 0$ and $\kappa > 0$,

$$n^{-1}\sum_{i=1}^{n} I(C_i \geq t)\big[A_i^0(t) - g^{-1}\{\pi_0(t) + \eta + \beta' Z_i(t)\}\big] \rightarrow E[I(C_i \geq t)(g^{-1}\{\pi_0(t) + \beta_0' Z_i(t)\} - g^{-1}\{\pi_0(t) + \eta + \beta' Z_i(t)\})]$$

almost surely and uniformly in $t \in [0, \tau], \eta \in [0, \kappa]$ and $\beta \in \mathcal{B} = \{\beta : \|\beta - \beta_0\| \leq \epsilon\}$. Thus, except for a null set, for all large $n, t \in [0, \tau], \beta \in \mathcal{B}$, and sufficiently large η,

(11)
$$n^{-1}\sum_{i=1}^{n} I(C_i \geq t)\big[A_i^0(t) - g^{-1}\{\pi_0(t) + \eta + \beta' Z_i(t)\}\big] < 0.$$

Likewise,

(12)
$$n^{-1}\sum_{i=1}^{n} I(C_i \geq t)\big[A_i^0(t) - g^{-1}\{\pi_0(t) - \eta + \beta' Z_i(t)\}\big] > 0.$$

By (11) and (12), and the monotonicity and continuity of g, for any $t \in [0, \tau], \beta \in \mathcal{B}$, there exists a unique $\hat{\pi}_0(t; \beta)$ that satisfies

$$\sum_{i=1}^{n} I(C_i \geq t)\big[A_i^0(t) - g^{-1}\{\hat{\pi}_0(t; \beta) + \beta' Z_i(t)\}\big] = 0. \tag{13}$$

Hence, to prove the existence and uniqueness of $\hat{\beta}$ and $\hat{\pi}_0(t)$, it suffices to show that there exists a unique solution to $U(\beta) = 0$, where

$$U(\beta) = \sum_{i=1}^{n}\int_0^{\tau} I(C_i \geq t) Z_i(t)\big[A_i^0(t) - g^{-1}\{\hat{\pi}_0(t; \beta) + \beta' Z_i(t)\}\big] dH(t) = 0.$$

Differentiation of (13) with respect to β gives

$$\frac{\partial \hat{\pi}_0(t; \beta)}{\partial \beta} = -\bar{Z}(t; \beta), \tag{14}$$

where

$$\bar{Z}(t; \beta) = \frac{\sum_{i=1}^{n} \dot{g}^{-1}\{\hat{\pi}_0(t; \beta) + \beta' Z_i(t)\} I(C_i \geq t) Z_i(t)}{\sum_{i=1}^{n} \dot{g}^{-1}\{\hat{\pi}_0(t; \beta) + \beta' Z_i(t)\} I(C_i \geq t)}.$$

Let $\hat{A}(\beta) = -n^{-1}\partial U(\beta)/\partial \beta'$. It follows from (14) and some simple algebra that

$$\hat{A}(\beta) = n^{-1}\sum_{i=1}^{n}\int_0^{\tau} \dot{g}^{-1}\{\hat{\pi}_0(t; \beta) + \beta' Z_i(t)\} I(C_i \geq t) \times \{Z_i(t) - \bar{Z}(t, \beta)\}^{\otimes 2} dH(t),$$

which is always nonnegative definite.

By the uniform strong law of large numbers, we have that $\hat{\pi}_0(t; \beta)$ and $\bar{Z}(t; \beta)$ converge almost surely to nonrandom functions $\pi_0(t; \beta)$ and $\bar{z}(t; \beta)$ uniformly in t and β. Thus, $\hat{A}(\beta)$ converges almost surely to a nonrandom function $A(\beta)$

uniformly in β, and $A(\beta_0) = A$, where

$$A(\beta) = E\Big[\int_0^\tau \dot{g}^{-1}\{\pi_0(t;\beta) + \beta' Z(t)\} I(C \geq t)\{Z(t) - \bar{z}(t;\beta)\}^{\otimes 2} d\tilde{H}(t)\Big].$$

It can be checked that $n^{-1}U(\beta_0) \to 0$ almost surely, and A is nonsingular by condition (C5). Then the uniform convergence of $\hat{A}(\beta)$ and the continuity of $A(\beta)$ imply that there exists a small neighborhood of β_0 in which $\hat{A}(\beta)$ is nonsingular when n is large enough. Thus, it follows from the inverse function theorem [22] that within a small neighborhood of β_0, there exists a unique solution $\hat{\beta}$ to $U(\beta) = 0$ for all sufficiently large n. The nonnegative definiteness of $\hat{A}(\beta)$ in the entire domain of β also implies the global uniqueness of $\hat{\beta}$. Hence there exist unique estimators $\hat{\beta}$ and $\hat{\pi}_0(t)$ $(0 \leq t \leq \tau)$. The above proof also implies that $\hat{\beta}$ is strongly consistent. Then by the uniform convergence of $\hat{\pi}_0(t;\beta)$, we have that $\hat{\pi}_0(t;\hat{\beta})$ converges almost surely to $\pi_0(t;\beta_0) \equiv \pi_0(t)$ uniformly in $t \in [0,\tau]$.

Proof of Theorem 2. (i) Taking the linear expansion of $g^{-1}(\cdot)$ in $U(\beta_0)$, we have

$$U(\beta_0) = \sum_{i=1}^n \int_0^\tau I(C_i \geq t) Z_i(t)\big[A_i^0(t) - g^{-1}\{\pi_0(t) + \beta_0' Z_i(t)\} - \dot{g}^{-1}\{\pi^*(t) + \beta_0' Z_i(t)\}\{\hat{\pi}_0(t;\beta_0) - \pi_0(t)\}\big] dH(t), \tag{15}$$

where $\pi^*(t)$ lies between $\hat{\pi}_0(t;\beta_0)$ and $\pi_0(t)$. Similarly, the linear expansion of (13) with $\beta = \beta_0$ yields

$$\hat{\pi}_0(t;\beta_0) - \pi_0(t) = \frac{\sum_{i=1}^n M_i(t)}{\sum_{i=1}^n \dot{g}^{-1}\{\pi^{**}(t) + \beta_0' Z_i(t)\} I(C_i \geq t)}, \tag{16}$$

where $\pi^{**}(t)$ also lies between $\hat{\pi}_0(t;\beta_0)$ and $\pi_0(t)$. Note that $\hat{\pi}_0(t;\beta_0) - \pi_0(t) = o_p(1)$ and $\sum_{i=1}^n M_i(t) = O_p(n^{1/2})$. Then it follows from (15) and (16) that

$$\begin{aligned} U(\beta_0) &= \sum_{i=1}^n \int_0^\tau M_i(t)\Big[Z_i(t) - \frac{\sum_{j=1}^n \dot{g}^{-1}\{\pi^*(t) + \beta_0' Z_j(t)\} I(C_j \geq t) Z_j(t)}{\sum_{j=1}^n \dot{g}^{-1}\{\pi^{**}(t) + \beta_0' Z_j(t)\} I(C_j \geq t)}\Big] \times dH(t) \\ &= \sum_{i=1}^n \int_0^\tau M_i(t)\{Z_i(t) - \bar{z}(t)\} d\tilde{H}(t) + o_p(n^{1/2}). \end{aligned} \tag{17}$$

Utilizing the multivariate central limit theorem, $n^{-1/2}U(\beta_0)$ converges in distribution to a zero-mean normal vector with covariance matrix

$$\Sigma = E\Big[\int_0^\tau M_i(t)\{Z_i(t) - \bar{z}(t)\} d\tilde{H}(t)\Big]^{\otimes 2}.$$

By the Taylor expansion of $U(\hat{\beta})$ at β_0,

$$n^{1/2}(\hat{\beta} - \beta_0) = A^{-1} n^{-1/2} U(\beta_0) + o_p(1). \tag{18}$$

This implies that $n^{1/2}(\hat{\beta} - \beta_0)$ converges in distribution to a zero-mean normal vector with covariance matrix $A^{-1}\Sigma A^{-1}$. By replacing all the unknown quantities in A and Σ with their empirical counterparts, the covariance matrix can be consistently estimated by $\hat{A}^{-1}\hat{\Sigma}\hat{A}^{-1}$ defined in Theorem 2(i).

(ii) To show the weak convergence of $n^{1/2}\{\hat{\pi}_0(\cdot) - \pi_0(\cdot)\}$, first note that uniformly in $t \in [0,\tau]$,

$$\begin{aligned} & n^{1/2}\{\hat{\pi}_0(t) - \pi_0(t)\} \\ &= n^{1/2}\{\hat{\pi}_0(t;\hat{\beta}) - \hat{\pi}_0(t;\beta_0)\} + n^{1/2}\{\hat{\pi}_0(t;\beta_0) - \pi_0(t)\} \\ &= -\bar{z}(t)' n^{1/2}(\hat{\beta} - \beta_0) + n^{1/2}\{\hat{\pi}_0(t;\beta_0) - \pi_0(t)\} + o_p(1), \end{aligned} \tag{19}$$

where the second equality follows from the Taylor expansion of π_0, together with (14) and the convergence of $\bar{Z}$. Thus, it follows from (16)–(19) that uniformly in $t \in [0,\tau]$,

$$n^{1/2}\{\hat{\pi}_0(t) - \pi_0(t)\} = n^{-1/2}\sum_{i=1}^n \phi_i(t) + o_p(1), \tag{20}$$

where

$$\phi_i(t) = s^{(0)}(t)^{-1} M_i(t) - \bar{z}(t)' A^{-1} \int_0^\tau M_i(t)[Z_i(t) - \bar{z}(t)] d\tilde{H}(t).$$

Because $\phi_i(t)$ $(i = 1, ..., n)$ are independent zero-mean random variables for each t, the multivariate central limit theorem implies that $n^{1/2}\{\hat{\pi}_0(t) - \pi_0(t)\}$ converges in finite-dimensional distributions to a zero-mean Gaussian process. Since $\phi_i(t)$ can be written as sums or products of monotone functions of t and are thus tight [27]. Thus, $n^{1/2}\{\hat{\pi}_0(t) - \pi_0(t)\}$ is tight and converges weakly to a zero-mean Gaussian process with the covariance function $\Gamma(t,s) \equiv E\{\phi_1(t)\phi_1(s)\}$ at (t,s), which can be consistently estimated by $\hat{\Gamma}(t,s)$ defined in Theorem 2(ii).

Proof of Theorem 3. First note that when $C_i^{\langle m\rangle}$ is imputed from the true underlying $G(t;\gamma_0)$, $I(C_i^{\langle m\rangle} \geq t;\gamma_0)$ and $A_i^0(t)$ are independent conditional on $Z_i(t)$. Next, consider $I(C_i^{\langle m\rangle} \geq t;\hat{\gamma})$, where $C_i^{\langle m\rangle}$ are imputed from $\hat{G}_i(t;\hat{\gamma})$ if $N_i^D(T_i) = 1$, where $N_i^D(t) = I(D_i \leq t \wedge C_i)$. The conditional survival function of imputed censoring time given $Z_i(t)$ and $N_i^D(T_i) = 1$ is given by

$$E[I(C_i^{\langle m\rangle} \geq t;\hat{\gamma}) | Z_i(t), N_i^D(T_i) = 1] = I(D_i \leq t) \times \exp\{\hat{\Lambda}_i^C(D_i;\hat{\gamma}) - \hat{\Lambda}_i^C(t;\hat{\gamma})\},$$

where $d\hat{\Lambda}_i^C(t;\gamma) = \exp\{\gamma' Z(t)\} d\hat{\Lambda}_0^C(t;\gamma)$, and

$$d\hat{\Lambda}_0^C(t;\gamma) = \frac{\sum_{i=1}^n dN_i^C(t)}{\sum_{i=1}^n I(T_i \geq t)\exp\{\gamma' Z_i(t)\}},$$

with $N_i^C(t) = I(C_i \leq t \wedge D_i)$.

Using the strong consistency of $\hat\gamma$ and the uniform strong law of large numbers, we obtain that $\hat\Lambda_0^C(t;\hat\gamma) - \Lambda_0^C(t)$ converges almost surely (a.s.) to zero uniformly in $t \in [0,\tau]$. By the continuous mapping theorem, we have that uniformly in $t \in [0,\tau]$,

(21)
$$E[I(C_i^{\langle m\rangle} \ge t;\hat\gamma)|Z_i(t), N_i^D(T_i)=1] \xrightarrow{a.s.} E[I(C_i^{\langle m\rangle} \ge t;\gamma_0)|Z_i(t), N_i^D(T_i)=1],$$

which implies that asymptotically, $C_i^{\langle m\rangle}$ are drawn from the true $G(t;\gamma_0)$ for subjects with $N_i^D(T_i)=1$. Note that for subjects with $N_i^D(T_i)=0$, $I(C_i^{\langle m\rangle} \ge t;\gamma_0) = I(C_i^{\langle m\rangle} \ge t;\hat\gamma) = I(C_i \ge t)$. It then follows from (21) that uniformly in $t\in[0,\tau]$,

(22)
$$E[I(C_i^{\langle m\rangle} \ge t;\hat\gamma)|Z_i(t)] = E[I(C_i^{\langle m\rangle} \ge t;\gamma_0)|Z_i(t)] + o(1).$$

Thus, $I(C_i^{\langle m\rangle} \ge t;\hat\gamma)$ and $A_i^0(t)$ are asymptotically independent conditional on $Z_i(t)$. By the strong law of large numbers, we get that $n^{-1}U^{\langle m\rangle}(\beta;\hat\gamma) \xrightarrow{a.s.} U(\beta)$ uniformly in β, where

$$\begin{aligned} U^{\langle m\rangle}(\beta;\gamma) &= \sum_{i=1}^n \int_0^\tau I(C_i^{\langle m\rangle} \ge t;\gamma) Z_i(t)\big[A_i^0(t) \\ &- g^{-1}\{\hat\pi_0(t;\beta)+\beta' Z_i(t)\}\big]dH(t). \end{aligned}$$

Following the arguments in the proof of Theorem 1, we have that $\hat\beta^{\langle m\rangle}$ and $\hat\pi_0^{\langle m\rangle}(t)$ $(0\le t\le\tau)$ exist and are unique. Moreover, we also have that $\|\hat\beta^{\langle m\rangle}-\beta_0\| \xrightarrow{a.s.} 0$ and $\hat\pi_0^{\langle m\rangle}(t;\hat\beta) \xrightarrow{a.s.} \pi_0(t)$ uniformly in $t\in[0,\tau]$. Note that by the triangle inequality, $\|\hat\beta^M-\beta_0\| \le M^{-1}\sum_{m=1}^M \|\hat\beta^{\langle m\rangle}-\beta_0\|$ and $\|\hat\pi^M(t)-\pi_0(t)\| \le M^{-1}\sum_{m=1}^M\|\hat\pi_0^{\langle m\rangle}(t)-\pi_0(t)\|$. Then we obtain that $\hat\beta^M$ is strongly consistent to β_0, and $\hat\pi_0^M(t)$ converges almost surely to $\pi_0(t)$ uniformly in $t\in[0,\tau]$.

Proof of Theorem 4. (i) In view of the assumption of independent censoring and (22), following similar arguments as in the proof Theorem 2(i), we get

$$\begin{aligned} & n^{1/2}(\hat\beta^{\langle m\rangle}-\beta_0) \\ =\ & [\hat A^{\langle m\rangle}(\pi_0,\beta_0,\hat\gamma)]^{-1} n^{-1/2}\sum_{i=1}^n\int_0^\tau M_i^{\langle m\rangle}(t;\pi_0,\beta_0,\hat\gamma) \\ & \times\{Z_i(t)-\bar Z^{\langle m\rangle}(t;\pi_0,\beta_0,\hat\gamma)\}dH(t)+o_p(1) \\ =\ & A^{-1}n^{-1/2}\sum_{i=1}^n\int_0^\tau M_i^{\langle m\rangle}(t;\pi_0,\beta_0,\gamma_0)\big\{Z_i(t) \\ & -\bar z^{\langle m\rangle}(t;\pi_0,\beta_0,\gamma_0)\big\}d\tilde H(t)+o_p(1), \end{aligned}$$

where

$$\begin{aligned} \hat A^{\langle m\rangle}(\pi_0,\beta,\gamma) = \int_0^\tau \Big[& S^{(2)\langle m\rangle}(t;\pi_0,\beta,\gamma) \\ & - \frac{S^{(1)\langle m\rangle}(t;\pi_0,\beta,\gamma)^{\otimes 2}}{S^{(0)\langle m\rangle}(t,\pi_0,\beta,\gamma)}\Big]dH(t), \end{aligned}$$

$$\begin{aligned} M_i^{\langle m\rangle}(t;\pi_0,\beta,\gamma) &= I(C_i^{\langle m\rangle} \ge t;\gamma)\big[A_i^0(t) \\ &\quad - g^{-1}\{\pi_0(t)+\beta' Z_i(t)\}\big], \\ \bar z^{\langle m\rangle}(t;\pi_0,\beta,\gamma) &= \frac{E[\dot g^{-1}\{\pi_0(t)+\beta' Z_i(t)\} I(C_i^{\langle m\rangle}\ge t;\gamma) Z_i(t)]}{E[\dot g^{-1}\{\pi_0(t)+\beta' Z_i(t)\} I(C_i^{\langle m\rangle}\ge t;\gamma)]}. \end{aligned}$$

Note that $\hat\beta^M = M^{-1}\sum_{m=1}^M \hat\beta^{\langle m\rangle}$. Then

$$\begin{aligned} & n^{1/2}(\hat\beta^M-\beta_0) \\ =\ & A^{-1}n^{-1/2}\sum_{i=1}^n\Big[M^{-1}\sum_{m=1}^M\int_0^\tau M_i^{\langle m\rangle}(t;\pi_0,\beta_0,\gamma_0) \\ & \times\big\{Z_i(t)-\bar z^{\langle m\rangle}(t;\pi_0,\beta_0,\gamma_0)\big\}d\tilde H(t)\Big]+o_p(1). \end{aligned}$$

The multivariate central limit theorem implies that $n^{1/2}(\hat\beta^M-\beta_0)$ converges in distribution to a zero-mean normal vector with covariance matrix $A^{-1}\Sigma^M A^{-1}$, where

$$\begin{aligned} \Sigma^M &= E\Big[M^{-1}\sum_{m=1}^M\int_0^\tau M^{\langle m\rangle}(t;\pi_0,\beta_0,\gamma_0)\big\{Z(t) \\ &- \bar z^{\langle m\rangle}(t;\pi_0,\beta_0,\gamma_0)\big\}d\tilde H(t)\Big]^{\otimes 2}. \end{aligned}$$

The covariance matrix $A^{-1}\Sigma^M A^{-1}$ can be consistently estimated by $(\hat A^M)^{-1}\hat\Sigma^M(\hat A^M)^{-1}$ defined in Theorem 4(i).

(ii) It can be shown that uniformly in $t\in[0,\tau]$,

$$\begin{aligned} & n^{1/2}\{\hat\pi_0^M(t)-\pi_0(t)\} \\ =& n^{1/2}\{\hat\pi_0^M(t;\hat\beta^M)-\hat\pi_0^M(t;\beta_0)\}+n^{1/2}\{\hat\pi_0^M(t;\beta_0)-\pi_0(t)\} \\ =& n^{1/2}M^{-1}\sum_{m=1}^M\{\hat\pi_0^{\langle m\rangle}(t;\hat\beta^M)-\hat\pi_0^{\langle m\rangle}(t;\beta_0)\} \\ & + n^{1/2}M^{-1}\sum_{m=1}^M\{\hat\pi_0^{\langle m\rangle}(t;\beta_0)-\pi_0(t)\} \\ =& n^{1/2}M^{-1}\sum_{m=1}^M\{\hat\pi_0^{\langle m\rangle}(t;\beta_0)-\pi_0(t)\} \\ & - M^{-1}\sum_{m=1}^M \bar z^{\langle m\rangle}(t;\pi_0,\beta_0,\gamma_0)' n^{1/2}(\hat\beta^M-\beta_0)+o_p(1). \end{aligned}$$

As in the proof of Theorem 2(ii), we have that uniformly in $t\in[0,\tau]$,

$$n^{1/2}\{\hat\pi_0^M(t)-\pi_0(t)\} = n^{-1/2}\sum_{i=1}^n\hat\phi_i^M(t)+o_p(1),$$

where

$$\begin{aligned} \phi_i^M(t) = M^{-1}\sum_{m=1}^M & \frac{M_i^{\langle m\rangle}(t;\pi_0,\beta_0,\gamma_0)}{E[\dot g^{-1}\{\pi_0(t)+\beta_0' Z(t)\}I(C^{\langle m\rangle}\ge t;\gamma_0)]} \\ & - \bar z^{\langle 1\rangle}(t;\pi_0,\beta_0,\gamma_0)' A^{-1}M^{-1} \end{aligned}$$

$$\times \sum_{m=1}^{M} \int_0^{\tau} M_i^{\langle m \rangle}(t; \pi_0, \beta_0, \gamma_0)$$
$$\times \left\{ Z_i(t) - \bar{z}^{\langle m \rangle}(t; \pi_0, \beta_0, \gamma_0) \right\} d\tilde{H}(t).$$

Thus, $n^{1/2}\{\hat{\pi}_0^M(t) - \pi_0(t)\}$ converges weakly to a zero-mean Gaussian process with the covariance function $\Gamma^M(t,s) \equiv E\{\phi_1^M(t)\phi_1^M(s)\}$ at (t,s), which can be consistently estimated by $\hat{\Gamma}^M(t,s)$ defined in Theorem 4(ii).

ACKNOWLEDGEMENTS

The authors would like to thank the Editor-in-Chief, Professor Yuedong Wang, an Associate Editor and two referees for their constructive and insightful comments and suggestions that greatly improve the article. This research was supported by the Fundamental Research Funds for the Central Universities (2021JBM045) and the National Natural Science Foundation of China (Grant Nos. 12171463, 11771431 and 11690015).

Received 20 August 2021

REFERENCES

[1] Charles-Nelson, A., Katsahian, S. and Schramm, C. (2019). How to analyze and interpret recurrent events data in the presence of a terminal event: an application on readmission after colorectal cancer surgery. *Statistics in Medicine* **38** 3476–3502. MR3979821

[2] Cook, R. and Lawless, J. (2007). *The Statistical Analysis of Recurrent Events*. Springer, New York. MR3822124

[3] Cook, R., Lawless, J., Lakhal-Chaieb, L. and Lee, K. (2009). Robust estimation of mean functions and treatment effects for recurrent events under event-dependent censoring and termination: application to skeletal complications in cancer metastatic to bone. *Journal of the American Statistical Association* **104** 60–75. MR2502850

[4] Cox, D. (1972). Regression models and life-tables (with discussion). *Journal of the Royal Statistical Society, Series B* **34** 187–200. MR0341758

[5] Estes, J., Nguyen, D., Dalrymple, L., Mu, Y. and Sentürk, D. (2015). Time-varying effect modeling with longitudinal data truncated by death: conditional models, interpretations, and inference. *Statistics in Medicine* **35** 1834–1847. MR3513488

[6] Fine, J. and Gray, R. (1999). A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association* **94** 496–509. MR1702320

[7] Fine, J., Yan, J. and Kosorok, M. (2004). Temporal process regression. *Biometrika* **91** 683–703. MR2090630

[8] Fleming, T. and Harrington, D. (1991). *Counting Processes and Survival Analysis*. Wiley, New York. MR1100924

[9] Ghosh, D. and Lin, D. (2002). Marginal regression models for recurrent and terminal events. *Statistica Sinica* **12** 663–688. MR1929958

[10] Graham, D., Ouellet-Hellstrom, R., MaCurdy, T., Ali, F., Sholley, C., Worrall, C. and Kelman, J. (2010). Risk of acute myocardial infarction, stroke, heart failure, and death in elderly medicare patients treated with rosiglitazone or pioglitazone. *JAMA* **304** 411–418.

[11] Hu, X., Lorenzi, M., Spinelli, J., Ying, S. and McBride, M. (2011). Analysis of recurrent events with non-negligible event duration, with application to assessing hospital utilization. *Lifetime Data Analysis* **17** 215–233. MR2777118

[12] Huang, C. and Wang, M. (2004). Joint modeling and estimation for recurrent event processes and failure time data. *Journal of the American Statistical Association* **99** 1153–1165. MR2109503

[13] Kalbfleisch J. and Prentice, R. (2002). *The Statistical Analysis of Failure Time Data*, 2nd edition. New York: Wiley. MR0570114

[14] Kalbfleisch J., Schaubel, D., Ye, Y. and Gong, Q. (2013). An estimating function approach to the analysis of recurrent and terminal events. *Biometrics* **69** 366–374. MR3071055

[15] Liang, K. and Zeger, S. (1986). Longitudinal data analysis using generalized linear models. *Biometrika* **73** 13–22. MR0836430

[16] Lin, D., Wei, L., Yang, I. and Ying, Z. (2000). Semiparametric regression for the mean and rate function of recurrent events. *Journal of the Royal Statistical Society B* **62** 711–730. MR1796287

[17] Lin, D., Wei, L. and Ying, Z. (2001). Semiparametric transformation models for point processes. *Journal of the American Statistical Association* **96** 620–628. MR1946429

[18] Liu, L., Wolfe, R. and Huang, X. (2004). Shared frailty models for recurrent events and a terminal event. *Biometrics* **60** 747–756. MR2089451

[19] Mao, L. and Lin, D. (2016). Semiparametric regression for the weighted composite endpoint of recurrent and terminal events. *Biostatistics* **17** 390–403. MR3516008

[20] Peng, L. and Huang, Y. (2007). Survival analysis with temporal covariate effects. *Biometrika* **94** 719–733. MR2410019

[21] Pollard, D. (1990). *Empirical Processes: Theory and Applications*, Institute of Mathematical Statistics, Hayward, CA. MR1089429

[22] Rudin, W. (1976). *Principles of Mathematical Analysis*, 3rd Edition. McGraw-Hill, New York. MR0385023

[23] Schaubel, D. and Zhang, M. (2010). Estimating treatment effects on the marginal recurrent event mean in the presence of a terminating event. *Lifetime Data Analysis* **16** 451–477. MR2726219

[24] Schelleman, H., Bilker, W., Brensinger, C., Han, X., Kimmel, S. and Hennessy, S. (2008). Warfarin with fluoroquinolones, sulfonamides, or azole antifungals: interactions and the risk of hospitalization for gastrointestinal bleeding. *Clinical Pharmacology and Therapeutics* **84** 581–588.

[25] Schelleman, H., Bilker, W., Kimmel, S., Daniel, G., Newcomb, C., Guevara, J., Cziraky, M., Strom, B. and Hennessy, S. (2012). Methylphenidate and risk of serious cardiovascular events in adults. *American Journal of Psychiatry* **169** 178–185.

[26] Sun, X. and Sun, L. (2019). Semiparametric estimation of differences in treatment-specific recurrent event means with a terminal event. *Statistics and Its Interface* **12** 1–9. MR3876405

[27] van der Vaart, A. and Wellner, J. (1996). *Weak Convergence and Empirical Processes*. Springer, New York. MR1385671

[28] Xu, G., Chiou, S., Huang, C., Wang, M. and Yan, J. (2017). Joint scale-change models for recurrent events and failure time. *Journal of the American Statistical Association* **112** 794–805. MR3671771

[29] Yan, J. and Fine, J. (2005). Functional association models for multivariate survival processes. *Journal of the American Statistical Association* **100** 184–196. MR2156828

[30] Yan, J. and Huang, J. (2009). Partly functional temporal process regression with semiparametric profile estimating functions. *Biometrics* **65** 431–440. MR2751466

[31] Ye, Y., Kalbfleisch, J. and Schaubel, D. (2007). Semiparametric analysis of correlated recurrent and terminal events. *Biometrics* **63** 78–87. MR2345577

[32] Zeng, D. and Lin, D. (2009). Semiparametric transformation models with random effects for joint analysis of recurrent and terminal events. *Biometrics* **65** 746–752. MR2649847

[33] Zhan, T. and Schaubel, D. (2019). Semiparametric temporal process regression of survival-out-of-hospital. *Lifetime Data Analysis* **25** 322–340. MR3924679

[34] Zhan, T. and Schaubel, D. (2020). Semiparametric regression methods for temporal processes subject to multiple sources of censoring. *Canadian Journal of Statistics* **48** 222–237. MR4095261

[35] Zhao, Q. and Sun, J. (2006). Semiparametric and nonparametric analysis of recurrent events with observation gaps. *Computational Statistics and Data Analysis* **51** 1924–1933. MR2307552

[36] Zhao, X., Zhou, J. and Sun, L. (2011). Semiparametric transformation models with time-varying coefficients for recurrent and terminal events. *Biometrics* **67** 404–414. MR2829009

Xiaowei Sun
Department of Mathematics
School of Science
Beijing Jiaotong University
Beijing 100044
People's Republic of China
E-mail address: sunxiaowei115@mails.ucas.ac.cn

Cheng Zeng
Institute of Applied Mathematics
Academy of Mathematics and Systems Science
Chinese Academy of Sciences
Beijing 100190
People's Republic of China
E-mail address: zengcheng@amss.ac.cn

Liuquan Sun
Institute of Applied Mathematics
Academy of Mathematics and Systems Science
Chinese Academy of Sciences
Beijing 100190
People's Republic of China
University of Chinese Academy of Sciences
Beijing 100049
People's Republic of China
E-mail address: slq@amt.ac.cn