Published in final edited form as:

J Am Stat Assoc. 2017 ; 112(518): 745–756. doi:10.1080/01621459.2016.1166113.

Evaluating Utility Measurement from Recurrent Marker Processes in the Presence of Competing Terminal Events

Yifei Sun [Postdoctoral Fellow] and

Department of Biostatistics, School of Public Health, Johns Hopkins University, Baltimore, MD 21205 (ysun26@jhu.edu).

Mei-Cheng Wang [Professor]

Department of Biostatistics, School of Public Health, Johns Hopkins University, Baltimore, MD 21205 (mcwang@jhu.edu).

Abstract

In follow-up studies, utility marker measurements are usually collected upon the occurrence of recurrent events until a terminal event such as death takes place. In this article, we define the recurrent marker process to characterize utility accumulation over time. For example, with medical cost and repeated hospitalizations being treated as marker and recurrent events respectively, the recurrent marker process is the trajectory of cumulative cost, which stops to increase after death. In many applications, competing risks arise as subjects are at risk of more than one mutually exclusive terminal event, such as death from different causes, and modeling the recurrent marker process for each failure type is often of interest. However, censoring creates challenges in the methodological development, because for censored subjects, both failure type and recurrent marker process after censoring are unobserved. To circumvent this problem, we propose a nonparametric framework for recurrent marker process with competing terminal events. In the presence of competing risks, we start with an estimator by using marker information from uncensored subjects. As a result, the estimator can be inefficient under heavy censoring. To improve efficiency, we propose a second estimator by combining the first estimator with auxiliary information from the estimate under non-competing risks model. The large sample properties and optimality of the second estimator is established. Simulation studies and an application to the SEER-Medicare linked data are presented to illustrate the proposed methods. Supplemental materials are available online.

Keywords

Competing risks; Cumulative incidence function; Marked recurrent events; Nonparametric estimation

Supplementary Materials

Proofs and Simulation Results: The file contains proof of Theorem 3.1, Theorem 3.2, Corollary 3.3, Theorem 4.1 and Theorem 5.1. The file also contains the additional simulation results not presented in the paper. (websupp.pdf)

R code: R code to perform the methods described in the article. (recmp.R)

1 Introduction

In many follow-up studies, longitudinal data are collected at pre-specified or random sampling times that do not have specific biological or medical implications. In contrast, recurrent marker data are a type of repeated measurements, where the sampling times are recurrent event times, and a marker measurement is observed conditioning on the occurrence of a recurrent event. In this article, we are interested in the case where the marker is a utility or cost measurement, and the marker does not exist unless a recurrent event occurs. For example, with repeated hospitalizations being treated as recurrent events, the marker can be medical cost associated with each hospitalization. In reality, the occurrence of recurrent events can be terminated by a failure event such as death, and competing terminal events are present when subjects can experience any one of several mutually exclusive terminal events (Kalbeisch and Prentice, 2002), such as death from different causes. To characterize real life utility consumption over time, we thus focus on the three components: recurrent events, markers and terminal events. The pattern of recurrent events and markers is usually related to the terminal event type. Besides analyzing recurrent marker data without distinguishing the terminal event type, investigators are also interested in modeling recurrent markers over time for subjects with a specific type of terminal event. This article presents a framework for nonparametric estimation of recurrent marker data in the presence of competing terminal events.

In the absence of competing risks, right censoring creates challenges for methodological developments, and numerous methods have been developed for modeling the total amount of utility or cost at a pre-specified time horizon. Huang and Louis (1998) studied nonparametric estimation of the joint distribution of a survival time and a mark variable, where an important example of mark variable is the lifetime utility or cost. Lin et al. (1997), Bang and Tsiatis (2000), Strawderman (2000), Zhao and Tian (2001) proposed nonparametric estimators of the mean of the mark variable. Huang (2002) and Sun et al. (2009) developed semiparametric models for inference on mark variable and survival outcome. Moreover, Lin (2000a,b) proposed semiparametric models for medical cost data. Another related area is analysis of recurrent event process with a terminal event, where Cook and Lawless (1997), Ghosh and Lin (2000), Schaubel and Zhang (2010) considered marginal models, and Huang and Wang (2004), Liu et al. (2004), Ye et al. (2007), Sinha et al. (2008), Zeng and Lin (2009) studied models where a frailty term is shared among recurrent and terminal events. In the literature, nonparametric estimation of recurrent marker process with competing terminal events has not been considered. In the presence of competing risks, because the failure type is typically unknown for censored subjects, existing methods in the aforementioned papers may not be directly extended to estimate the mean function of cumulative utility attributed to a specific failure type. To solve this problem, we start with an estimator that can be inefficient under heavy censoring, and construct a second estimator with improved efficiency by combining auxiliary information from the estimation ignoring failure types. The proposed methods are also relevant to the quality-of-life research. Specifically, the methodology can be extended to quality-adjusted survival analysis (Glasziou et al., 1990) when the terminal event occurs with competing risks.

It is worthwhile to point out that the focus of our work is not to evaluate treatment effect through markers or survival, but to estimate utility measurements among survivors who have not yet advanced to the terminal event; for example, evaluation of medical cost before death. This type of estimation result would be useful from the perspective of providers or payers, such as government or health care organizations.

The content of this article is organized as follows: In Section 2, a point process for recurrent events is generalized to a recurrent marker process by accounting for additional information from markers as well as risk types. A mean function (MF) is defined with or without competing risks specification. In Section 3, we consider different nonparametric approaches for estimating the MF with or without competing risks specification. In Section 4, an improved estimator of the MF under competing risks model is proposed for an efficiency gain. The adopted technique can be regarded as a nonparametric approach of using auxiliary information to improve estimation efficiency, in the spirit of generalized method of moments (Hansen, 1982). In Section 5, to provide further insight into evaluating utility, we propose nonparametric estimation of MF with adjustment for survival. In Section 6, simulation studies and an analysis of SEER-Medicare linked data are presented for illustration of the proposed approaches. A discussion follows in Section 7 to conclude the paper.

2 Recurrent Marker Processes

Let $\{N(t), t \geq 0\}$ be a recurrent event process with $N(t)$ representing the total number of recurrent events occurring at or prior to t . Suppose the occurrence of a recurrent event at t , i.e., $dN(t) = 1$, is marked by a measurement $Y(t)$. For ease of discussion, we assume that $Y(t)$ is nonnegative. Then, $Y(t)$ can be considered as a marker measurement for the recurrent event occurring at t , and we represent the marked recurrent event process by $\{N(t), Y(t)|_{dN(t)=1}; t \geq 0\}$. Consider the case where a terminal event is present, and the marked recurrent event process $\{N(t), Y(t)|_{dN(t)=1}; t \geq 0\}$ vanishes after the terminal event. Specifically, in situations where the marker $Y(\cdot)$ is a utility measurement such as medical cost or length of stay in hospital, the main interest would naturally be the utility consumed by survivors in the population. Let the time to the terminal event be represented by D , which is possibly correlated with $\{N(t), Y(t)|_{dN(t)=1}; t \geq 0\}$, then the stochastic process $\{Y(t)dN(t), 0 \leq t \leq D\}$ is frequently of interest.

In the presence of a terminal event without competing risks, the recurrent marker process is defined as

$$M^{\text{total}}(t) = \int_0^t Y(u)I(D \geq u)dN(u), \quad (2.1)$$

and we use the superscript “total” to distinguish (2.1) from recurrent marker processes under competing risks model, which is defined later in this section. The mean function (MF) is defined as

$$\Phi(t) = E\{M^{\text{total}}(t)\} = E\left\{\int_0^t Y(u)I(D \geq u)dN(u)\right\}.$$

When $Y(\cdot)|_{dN(\cdot)=1}$ is a utility measurement, the MF $\Phi(t)$ corresponds to the average of cumulative utility consumed before the terminal event during time interval $[0, t]$, which is the utility of real life in case death is the terminal event. Also note that by setting $Y(\cdot) = 1$, the recurrent marker process reduces to a recurrent event process and $M^{\text{total}}(t) = N(\min(D, t))$ (Cook and Lawless, 1997; Ghosh and Lin, 2000).

We further consider recurrent marker process in the presence of a terminal event with competing risks. Suppose the occurrence of the terminal event is caused by one of J different types of risks, where the risk-type indicator is denoted by $\Pi \in \{1, \dots, J\}$. Let τ be a pre-specified constant, due to the identifiability issue in nonparametric estimation, subjects with $D \leq \tau$ are classified according to the original risk-type of terminal event $\Pi = j$ ($j = 1, \dots, J$), and subjects with $D > \tau$ are classified into the last category of risk-type, $J+1$. The value of τ can be determined based on scientific interest as well as the length of the study, and it should be a large enough value to cover the time period of interest. In real applications, subjects of risk-type $J+1$ can be thought of as those “long-term survivors” or “cured cases”.

For $0 \leq t \leq \tau$ and $j = 1, 2, \dots, J$, the recurrent marker process with type- j terminal event is defined as

$$M_j(t) = \int_0^t Y(u)I(\Pi=j, u \leq D \leq \tau)dN(u),$$

and the recurrent marker process with type- $(J+1)$ terminal event is defined as

$$M_{J+1}(t) = \int_0^t Y(u)I(D > \tau)dN(u).$$

Taking expectation of $M_j(t)$, the type- j MF at time t is

$$\Phi_j(t) = E\{M_j(t)\}, \quad j=1, 2, \dots, J+1.$$

For $j = 1, 2, \dots, J+1$, the MF $\Phi_j(t)$ is the expectation of cumulative recurrent marker prior to time t attributed to type- j risk. To connect the function $\Phi(t)$ with $\Phi_j(t)$, it is clear that for each $t \in [0, \tau]$, we have

$$\Phi(t) = \sum_{j=1}^{J+1} \Phi_j(t). \quad (2.2)$$

In practice, another quantity of interest is the conditional mean function of recurrent marker process given risk-type j , that is,

$$\Phi_j^c(t) = \begin{cases} E\{M^{\text{total}}(t)|\Pi=j, D \leq \tau\}, & j=1, \dots, J, \\ E\{M^{\text{total}}(t)|D>\tau\}, & j=J+1. \end{cases}$$

Note that $\Phi_j^c(t)$ is the expectation of cumulative recurrent marker up to time t of subpopulation with type- j risk. We consider nonparametric estimation of $\Phi(t)$, $\Phi_j(t)$ and $\Phi_j^c(t)$ ($j=1, \dots, J+1$, $0 < t \leq \tau$) in later sections.

3 Nonparametric Estimation of Mean Functions

3.1 Non-competing risks model

We first consider nonparametric estimation of $\Phi(t)$ without the complication of competing risks. In practice, the terminal event time D is usually subject to right censoring due to study end or premature dropout, and the recurrent marker process cannot be observed after censoring. We denote the censoring time by C and assume that C is independent of $\{D, N(t), Y(t)dN(t); 0 \leq t \leq D\}$. Let $S_D(u) = \Pr(D \geq u)$ and $S_C(u) = \Pr(C \geq u)$, and assume S_D and S_C are continuous. The observed terminal event time is $X = \min(D, C)$ with a censoring indicator $\delta = I(D \leq C)$. The observed data $\{X_i, \delta_i, Y_i(t)dN_i(t), N_i(t); 0 \leq t \leq X_i, i=1, \dots, n\}$ are assumed to be independent replicates of $\{X, \delta, Y(t)dN(t), N(t); 0 \leq t \leq X\}$.

Under the independent censoring assumption, subjects in the risk set at time u are a representative sample of survivors at time u in the target population, and it is easy to see that

$$d\Phi(u) = S_D(u) \cdot E\{Y(u)dN(u)|D \geq u\} = S_D(u) \cdot E\{Y(u)dN(u)|X \geq u\}.$$

Thus, for $0 < t \leq \tau$, we have

$$\Phi(t) = \int_0^t S_D(u) \cdot E\{Y(u)dN(u)|X \geq u\}. \quad (3.1)$$

To estimate $\Phi(t)$, one can use the moment estimator of $E\{Y(u)dN(u)|X \geq u\}$ based on subjects in the risk set $\{i: X_i \geq u\}$ and estimate $S_D(u)$ by the Kaplan-Meier estimator $\hat{S}_D(u)$, then a nonparametric estimator of $\Phi(t)$ can be constructed as

$$\tilde{\Phi}(t) = \int_0^t \hat{S}_D(u) \cdot \frac{\sum_{i=1}^n Y_i(u)I(X_i \geq u)dN_i(u)}{\sum_{i=1}^n I(X_i \geq u)}. \quad (3.2)$$

The estimator in (3.2) can be viewed as an extension of the nonparametric estimator of mean frequency function in Cook and Lawless (1997) and Ghosh and Lin (2000), where the special case $Y(\cdot) = 1$ was considered. To study large sample properties of $\tilde{\Phi}(t)$, the following notations are introduced. Let $S_X(u) = \Pr(X \geq u)$ and let Λ_C be the cumulative hazard function of the censoring time C . We then define $\mathcal{N}_i^C(t) = I(C_i \leq t, \Delta_i = 0)$ and

$\mathcal{M}_i^C(t) = \mathcal{N}_i^C(t) - \int_0^t I(X_i \geq u) d\Lambda_C(u)$. Theorem 3.1 summarizes the asymptotic property of $\tilde{\Phi}(t)$. The proof is detailed in Section S.1 of the Supplementary Material.

Theorem 3.1—Under Assumption (A1) and (A2) in Appendix, for $t \in (0, \tau]$, the stochastic process $n^{1/2}\{\tilde{\Phi}(t) - \Phi(t)\}$ has an asymptotically i.i.d. representation

$n^{1/2}\{\tilde{\Phi}(t) - \Phi(t)\} = n^{-1/2} \sum_{i=1}^n a_i(t) + o_p(1)$, where

$$a_i(t) = \int_0^t S_C(u)^{-1} Y_i(u) I(X_i \geq u) dN_i(u) - \Phi(t) + \Phi(0) - \int_0^t S_X(u)^{-1} d\mathcal{M}_i^C(u) - \int_0^t \Phi(u) S_X(u)^{-1} d\mathcal{M}_i^C(u).$$

Moreover, as $n \rightarrow \infty$, $n^{1/2}\{\tilde{\Phi}(t) - \Phi(t)\}$ ($0 < t \leq \tau$) converges weakly to a zero-mean tight Gaussian process with covariance function $E\{a_1(t_1)a_1(t_2)\}$ for $t_1, t_2 \in (0, \tau]$.

3.2 Competing risks model

When the terminal event occurs with competing risks, nonparametric estimation of the MF will need to take into account the data structure that the risk-type indicator, Π , is available only when the terminal event is observed. Under competing risks model, the censoring time C is assumed to be independent of $\{D, \Pi, N(t), Y(t)dN(t); 0 \leq t \leq D\}$. The observed data $\{X_i, \Pi_i, Y_i(t)dN_i(t), N_i(t); 0 \leq t \leq X_i, i = 1, \dots, n\}$ are assumed to be independent replicates of $\{X, \Pi, Y(t)dN(t), N(t); 0 \leq t \leq X\}$.

Similar to the formula in (3.1), for $j = 1, \dots, J$ and $0 < t \leq \tau$, one derives

$$\begin{aligned} \Phi_j(t) &= \int_0^t E\{Y(u)I(\Pi=j, u \leq D \leq \tau)dN(u)\} \\ &= \int_0^t S_D(u)E\{I(\Pi=j)Y(u)dN(u)|X \geq u\} - S_D(\tau)E\{I(\Pi=j)M^{\text{total}}(t)|X > \tau\}. \end{aligned}$$

Therefore, for $j = 1, \dots, J$, along the same line of the estimator $\tilde{\Phi}(t)$ in (3.2), a hypothetical estimator of $\Phi_j(t)$ that utilizes the recurrent marker history data from all the subjects can be obtained as,

$$\begin{aligned} \tilde{\Phi}_j^H(t) &= \int_0^t \hat{S}_D(u) \cdot \frac{\sum_{i=1}^n I(X_i \geq u, \Pi_i=j) Y_i(u) dN_i(u)}{\sum_{i=1}^n I(X_i \geq u)} - \hat{S}_D(\tau) \frac{\sum_{i=1}^n I(X_i > \tau, \Pi_i=j) M_i^{\text{total}}(t)}{\sum_{i=1}^n I(X_i > \tau)}. \end{aligned} \quad (3.3)$$

With subscript i , $M_i^{\text{total}}(t)$ represents the process $M^{\text{total}}(t)$ for the i th subject. For risk type $J+1$, an estimator of $\Phi_{J+1}(t) = S_D(\tau)E\{M^{\text{total}}(t) | X > \tau\}$ can be constructed as

$$\tilde{\Phi}_{J+1}^H(t) = \hat{S}_D(\tau) \frac{\sum_{i=1}^n I(X_i > \tau) M_i^{\text{total}}(t)}{\sum_{i=1}^n I(X_i > \tau)}.$$

Clearly, the estimator $\tilde{\Phi}_j^H(t)$ ($j = 1, \dots, J$) depends on data information of Π from all subjects in the risk set at the observed recurrent event times. In reality, however, knowledge of Π is rarely available from subjects whose terminal events are censored ($\delta = 0$), therefore the hypothetical estimator in (3.3) fails to serve as a proper estimator for most of the applications. In what follows, we propose an estimation approach that is useful for commonly encountered recurrent marker data with competing terminal events.

For different risk types $j = 1, \dots, J$, we define $H_j(t, m, u) = \Pr(\Pi = j, M^{\text{total}}(t) \leq m, D \leq u)$. With straightforward algebra, the mean function for type- j risk is

$$\Phi_j(t) = \int_0^\infty m H_j(t, dm, \tau).$$

Note that for $u \in [0, \tau]$, $H_j(t, \infty, u) = \Pr(\Pi = j, D \leq u)$ is the cumulative incidence function in standard competing risks model (Prentice et al., 1978; Gray, 1988; Fine and Gray, 1999). Define $U_j(t, m, u) = \Pr(\Pi = j, M^{\text{total}}(t) \leq m, \delta = 1, X \leq u)$. Under the assumption that C is independent of $\{D, \Pi, Y(t) dN(t); 0 \leq t \leq D\}$, one derives

$$H_j(t, m, u) = \int_0^u S_D(v) \frac{H_j(t, m, dv)}{S_D(v)} = \int_0^u S_D(v) \frac{U_j(t, m, dv)}{S_X(v)},$$

where $S_X(u) = \Pr(X \leq u)$. Note that when a terminal event is uncensored, the risk type is observed and therefore $U_j(t, m, u)$ can be estimated by its empirical average, that is,

$\hat{U}_j(t, m, u) = n^{-1} \sum_{i=1}^n I(\Pi_i = j, M_i^{\text{total}}(t) \leq m, \Delta_i = 1, X_i \leq u)$. By plugging into the Kaplan-Meier estimate $\hat{S}_D$ and the empirical estimates $(\hat{S}_X, \hat{U}_j)$, $H_j(t, m, u)$ can be nonparametrically estimated by

$$\hat{H}_j(t, m, u) = \int_0^u \hat{S}_D(v) \frac{\hat{U}_j(t, m, dv)}{\hat{S}_X(v)}. \quad (3.4)$$

Thus, an estimator of $\Phi_j(t)$ ($j = 1, \dots, J$) can be constructed as

$$\hat{\Phi}_j(t) = \int_0^\infty m \hat{H}_j(t, dm, \tau). \quad (3.5)$$

and $\Phi_{J+1}(t)$ is still estimated by

$$\hat{\Phi}_{J+1}(t) \equiv \tilde{\Phi}_{J+1}^H(t) = \hat{S}_D(\tau) \frac{\sum_{i=1}^n I(X_i > \tau) M_i^{\text{total}}(t)}{\sum_{i=1}^n I(X_i > \tau)}.$$

Theorem 3.2 summarizes the large sample properties of $\hat{\Phi}_j(t)$, $j = 1, \dots, J+1$. The proof is detailed in Section S.2 of the Supplementary Material.

Theorem 3.2—Under Assumption (A1') and (A2) in Appendix, for $t \in (0, \tau]$ and $j = 1, \dots, J+1$, the stochastic process $n^{1/2}\{\hat{\Phi}_j(t) - \Phi_j(t)\}$ has an asymptotically i.i.d representation $n^{1/2}\{\hat{\Phi}_j(t) - \Phi_j(t)\} = n^{-1/2} \sum_{i=1}^n b_{ij}(t) + o_p(1)$, where for $j = 1, \dots, J$,

$$b_{ji}(t) = M_i^{\text{total}}(t) I(\Pi_i = j, D_i \leq \tau) \Delta_i S_C(D_i)^{-1} - \Phi_j(t) + \Phi_j(t) \int_0^\tau S_X(u)^{-1} d\mathcal{M}_i^C(u) \\ - \int_0^\tau E\{I(\Pi = j, D \leq u) M^{\text{total}}(t)\} S_X(u)^{-1} d\mathcal{M}_i^C(u),$$

and for $j = J+1$,

$$b_{ji}(t) = I(X_i > \tau) M_i^{\text{total}}(t) S_C(\tau)^{-1} - \Phi_{J+1}(t) + \Phi_{J+1}(t) \int_0^\tau S_X(u)^{-1} d\mathcal{M}_i^C(u).$$

Moreover, as $n \rightarrow \infty$, $n^{1/2}\{\hat{\Phi}_j(t) - \Phi_j(t)\}$ ($0 < t \leq \tau$) converges weakly to a zero-mean tight Gaussian process with covariance function $E\{b_{j1}(t_1)b_{j1}(t_2)\}$ for $t_1, t_2 \in (0, \tau]$.

In real data applications, it is also of interest to estimate the mean recurrent marker process within each risk group, $\Phi_j^c(t)$. For $j = 1, \dots, J$, since the equation $\Phi_j^c(t) = \Phi_j(t) / \Pr(\Pi = j, D \leq \tau)$ holds, we propose the following estimator,

$$\hat{\Phi}_j^c(t) = \frac{\hat{\Phi}_j(t)}{\hat{H}_j(0, \infty, \tau)},$$

where $\hat{H}_j(0, \infty, \tau)$ is defined in equation (3.4). For risk-type $j = J+1$, since we have $\Phi_{J+1}^c(t) = \Phi_{J+1}(t) / \Pr(D > \tau)$, we propose the estimator

$$\hat{\Phi}_{J+1}^c(t) = \frac{\hat{\Phi}_{J+1}(t)}{\hat{S}_D(\tau)}.$$

Corollary 3.3 summarizes the large sample properties of $\hat{\Phi}_j^c(t)$. The proof is detailed in Section S.3 of the Supplementary Material.

Corollary 3.3—Under Assumption (A1') and (A2) in Appendix, for $j = 1, \dots, J+1$, the stochastic process $n^{1/2}\{\hat{\Phi}_j^c(t) - \Phi_j^c(t)\}$ ($0 < t \leq \tau$) converges weakly to a zero-mean tight Gaussian process.

4 Improved Estimation in Competing Risks Model

As the estimator $\hat{\Phi}_j(t)$ ($j = 1, \dots, J$) utilizes recurrent marker information only from uncensored subjects, the estimation may be inefficient when censoring is heavy; in contrast, for the estimation of $\Phi(t)$, the estimator $\tilde{\Phi}(t)$ in Section 3.1 utilizes recurrent marker history data from both censored and uncensored subjects. Thus, it is of no surprise that $\tilde{\Phi}(t)$ is more

efficient than $\sum_{j=1}^{J+1} \hat{\Phi}_j(t)$ for estimating $\Phi(t)$. When formulating estimators for type- j MF $\Phi_j(t)$, a question arises as to whether it is possible to borrow information from $\tilde{\Phi}(t)$ to improve the estimation of $\Phi_j(t)$.

Note that, based on (2.2), we can construct an alternative estimator of $\Phi_j(t)$ as

$$\tilde{\Phi}_j(t) = \tilde{\Phi}(t) - \sum_{k \neq j} \hat{\Phi}_k(t).$$

In general, $\tilde{\Phi}_j(t)$ may or may not be more efficient than $\hat{\Phi}_j(t)$, even though the former estimator involves recurrent marker history from both censored and uncensored subjects and the latter only uses recurrent marker information from uncensored ones. In what follows, we propose an estimator that is more efficient than $\tilde{\Phi}_j(t)$ and $\hat{\Phi}_j(t)$. We consider a class of linearly combined estimators $\mathcal{W}_{jt} = \{w_{jt}\hat{\Phi}_j(t) + (1 - w_{jt})\tilde{\Phi}_j(t) : w_{jt} \in \mathbb{R}\}$ and propose the use of the most efficient estimator from this class. Let $\tilde{\Phi}_j(t) = w_{jt}\hat{\Phi}_j(t) + (1 - w_{jt})\tilde{\Phi}_j(t)$ be a weighted average of $\hat{\Phi}_j(t)$ and $\tilde{\Phi}_j(t)$. Clearly, $\hat{\Phi}_j(t)$ and $\tilde{\Phi}_j(t)$ both belong to $\mathcal{W}_{jt}$. Using results from Theorem 3.1 and 3.2, estimators in $\mathcal{W}_{jt}$ are consistent and asymptotically normal. Thus, a question of interest is to identify the estimator in $\mathcal{W}_{jt}$ with minimal asymptotic variance.

To study the asymptotic variance of the estimators in $\mathcal{W}_{jt}$, we define $\Sigma_j(t_1, t_2) = \{\sigma_{jpp}(t_1, t_2)\}_{2 \times 2}$ for $t_1, t_2 \in [0, \tau]$, where $\sigma_{j11}(t_1, t_2)$, $\sigma_{j12}(t_1, t_2)$, $\sigma_{j21}(t_1, t_2)$ and $\sigma_{j22}(t_1, t_2)$ are the asymptotic covariance of

$$\begin{aligned} & [\sqrt{n}\{\hat{\Phi}_j(t_1) - \Phi_j(t_1)\}, \sqrt{n}\{\hat{\Phi}_j(t_2) - \Phi_j(t_2)\}], [\sqrt{n}\{\tilde{\Phi}_j(t_1) - \Phi_j(t_1)\}, \sqrt{n}\{\tilde{\Phi}_j(t_2) - \Phi_j(t_2)\}], \\ & [\sqrt{n}\{\hat{\Phi}_j(t_1) - \Phi_j(t_1)\}, \sqrt{n}\{\tilde{\Phi}_j(t_2) - \Phi_j(t_2)\}], [\sqrt{n}\{\tilde{\Phi}_j(t_1) - \Phi_j(t_1)\}, \sqrt{n}\{\hat{\Phi}_j(t_2) - \Phi_j(t_2)\}], \end{aligned}$$

and

$[\sqrt{n}\{\tilde{\Phi}_j(t_1) - \Phi_j(t_1)\}, \sqrt{n}\{\tilde{\Phi}_j(t_2) - \Phi_j(t_2)\}]$, respectively. We further assume that $\Sigma_j(t, t)$ is a nonsingular matrix for $t \in (0, \tau]$. Define the vector $w_{jt} = (w_{jt}, 1 - w_{jt})^\top$, then the asymptotic variance of $\sqrt{n}\{\tilde{\Phi}_j(t) - \Phi_j(t)\}$ is $w_{jt}^\top \Sigma_j(t, t) w_{jt}$. We consider the optimization of asymptotic variance to obtain the most efficient estimator in the class of $\mathcal{W}_{jt}$:

$$\underset{w_{jt}}{\text{minimize}} \quad w_{jt}^\top \Sigma_j(t, t) w_{jt}.$$

By straightforward algebra, the optimal weight can be derived as

$$w_{jt}^* = \frac{\Sigma_j(t, t)^{-1} \mathbf{e}}{\mathbf{e}^\top \Sigma_j(t, t)^{-1} \mathbf{e}},$$

where we define $\mathbf{e} = (1, 1)^\top$. And using this optimal weight w_{jt}^* , the asymptotic variance of $\sqrt{n}\{\tilde{\Phi}_j(t) - \Phi_j(t)\}$ is $\{\mathbf{e}^\top \Sigma_j(t, t)^{-1} \mathbf{e}\}^{-1}$. By the Cauchy-Schwartz inequality, for any weight w_{jt} we have

$$\{w_{jt}^\top \Sigma_j(t, t) w_{jt}\} \{e^\top \Sigma_j(t, t)^{-1} e\} \geq (w_{jt}^\top e)^2 = 1,$$

or equivalently,

$$w_{jt}^\top \Sigma_j(t, t) w_{jt} \geq \{e^\top \Sigma_j(t, t)^{-1} e\}^{-1}. \quad (4.1)$$

Thus, in the class $\mathcal{W}_{jt}$ using the weight w_{jt}^* results in the estimator with smallest asymptotic variance. In real applications, the optimal w_{jt}^* involves $\Sigma_j(t, t)$, which is unknown and needs to be estimated from data. By applying the results of Theorem 3.1 and 3.2, we can consistently estimate $\Sigma_j(t, t)$ by $\hat{\Sigma}_j(t, t)$, with details given in S.4 of Supplementary Material. We propose the following improved estimator for $\Phi_j(t)$,

$$\hat{\Phi}_j^{\text{imp}}(t) = \hat{w}_{jt}^* \hat{\Phi}_j(t) + (1 - \hat{w}_{jt}^*) \tilde{\Phi}_j(t), \quad (4.2)$$

where $\hat{w}_{jt}^* = (\hat{w}_{jt}^*, 1 - \hat{w}_{jt}^*)^\top = \hat{\Sigma}_j(t, t)^{-1} e / e^\top \hat{\Sigma}_j(t, t)^{-1} e$. Here we indicate that the estimator using the estimated weight $\hat{w}_{jt}^*$ possesses the same efficiency as the estimator with w_{jt}^* as the weight. Theorem 4.1 summarizes the large sample properties of $\hat{\Phi}_j^{\text{imp}}(t)$. The proof is detailed in Section S.4 of the Supplementary Material.

Theorem 4.1

Under Assumption (A1'), (A2) and (A3) in Appendix, for $j = 1, \dots, J+1$,

$\sqrt{n}\{\hat{\Phi}_j^{\text{imp}}(t) - \Phi_j(t)\} (0 < t \leq \tau)$ converges weakly to a zero-mean tight Gaussian process with

$$\frac{e^\top \Sigma_j(t_1, t_1)^{-1} \Sigma_j(t_1, t_2) \Sigma_j(t_2, t_2)^{-1} e}{e^\top \Sigma_j(t_1, t_1)^{-1} e e^\top \Sigma_j(t_2, t_2)^{-1} e}$$

as the covariance function at (t_1, t_2) for $t_1, t_2 \in (0, \tau]$. Moreover, the process

$\sqrt{n}\{\hat{\Phi}_1^{\text{imp}}(t) - \Phi_1(t), \dots, \hat{\Phi}_{j+1}^{\text{imp}}(t) - \Phi_{j+1}(t)\}^\top (0 < t \leq \tau)$ converges weakly to a mean zero multivariate Gaussian process with covariance function $E\{p_{j+1}(t_1) p_{j+1}(t_2)\}$ for j' th and j th process at (t_1, t_2) , and p_{ji} is defined in S.4 of the Supplementary Material.

The result of Theorem 4.1 implies that the asymptotic variance of $\sqrt{n}\{\hat{\Phi}_j^{\text{imp}}(t) - \Phi_j(t)\}$, $\{e^\top \Sigma_j(t, t)^{-1} e\}^{-1}$, is smaller than or equal to the asymptotic variance of $\sqrt{n}\{\hat{\Phi}_j(t) - \Phi_j(t)\}$ and $\sqrt{n}\{\tilde{\Phi}_j(t) - \Phi_j(t)\}$.

Besides the estimator $\hat{\Phi}_j^c(t)$ introduced in Section 3.2, a second estimator of $\Phi_j^c(t)$ can be obtained as

$$\tilde{\Phi}_j^c(t) = \begin{cases} \frac{\tilde{\Phi}_j(t)}{\tilde{H}_j(0, \infty, \tau)}, & j=1, \dots, J, \\ \frac{\tilde{\Phi}_j(t)}{\tilde{S}_D(\tau)}, & j=J+1. \end{cases}$$

An optimization procedure similar to the preceding one can be applied to construct an improved estimator of $\Phi_j^c(t)$ as

$$\hat{\Phi}_j^{c, \text{imp}}(t) = \hat{w}_{jt}^{c*} \hat{\Phi}_j^c(t) + (1 - \hat{w}_{jt}^{c*}) \tilde{\Phi}_j^c(t),$$

where $(\hat{w}_{jt}^{c*}, 1 - \hat{w}_{jt}^{c*})^\top = \hat{\Sigma}_j^c(t, t)^{-1} e / e^\top \hat{\Sigma}_j^c(t, t)^{-1} e$ and $\hat{\Sigma}_j^c(t, t)$ is defined in Section S.4 of the Supplementary Material. The asymptotic properties of $\hat{\Phi}_j^{c, \text{imp}}(t)$ are also derived in Section S.4 of the Supplementary Material.

5 Nonparametric Estimation Adjusting for Survival Time

While the focus of our work is mainly on estimation of utility measurement for survivors, the estimation with adjustment for length of survival time is also of interest and will provide further insight into evaluating the utility measurement. For example, higher lifetime utility value could be associated with either prolonged survival or higher utility cost for each recurrent event. For the purpose of comparison, it is reasonable to compare the utility given subjects have the same survival time. The proposed methodology in Section 3.2 can be extended to evaluate the utility measurement adjusting for the length of survival time. Under the non-competing risk model, the quantity $\mu(s, t) = E\{M^{\text{total}}(s) \mid D = t\}$ is of interest. When competing terminal events are present, along the same line as in Section 2, we consider the functions

$$\nu_j(s, t) = \begin{cases} E\{M^{\text{total}}(s) \mid D = t \leq \tau, \Pi = j\}, & j=1, \dots, J, \\ E\{M^{\text{total}}(s) \mid D = t > \tau\}, & j=J+1. \end{cases}$$

Note that $\nu_{J+1}(s, t) = \mu(s, t)$ if $t > \tau$. In what follows, we propose nonparametric estimation for $\mu(s, t)$ and $\nu_j(s, t)$ ($j = 1, \dots, J$).

With straightforward algebra, for $s, t \in (0, \tau)$, $j = 1, \dots, J$, the conditional mean function $\nu_j(s, t)$ can be derived as

$$\nu_j(s, t) = \frac{E\{M^{\text{total}}(s) I(\Pi = j) dI(D \leq t)\}}{E\{dI(D \leq t, \Pi = j)\}} = \frac{\int_0^\infty m H_j(s, dm, dt)}{H_j(s, \infty, dt)}.$$

The function H_j can be consistently estimated by $\hat{H}_j$ in (3.4). When $0 < s < t < \tau$, $\nu_j(s, t)$ can be estimated by the kernel type estimator

$$\hat{\nu}_j(s, t) = \frac{\int_s^{\tau_C} \int_0^\infty K_{s,t}(\frac{t-u}{h}) m \hat{H}_j(s, dm, du)}{\int_s^{\tau_C} K_{s,t}(\frac{t-u}{h}) \hat{H}_j(s, \infty, du)},$$

where $\tau_C = \sup\{u : \min(S_C(u), S_D(u)) > 0\}$ and we assume $\tau_C > \tau$, h is the bandwidth parameter, and the function $K_{s,t}(x)$ is a second order modified kernel function adjusted for the interval $[s, \tau_C]$ (Müller, 1991). Specifically, $K_{s,t}$ is a bounded symmetric kernel function on $[-1, 1]$ for interior $t \in [s + h, \tau_C - h]$. For left boundary $t \in (s, s + h)$, the support of $K_{s,t}$ is $[-1, (t - s)/h]$; and for right boundary $t \in (\tau_C - h, \tau_C)$, the support of $K_{s,t}$ is $[(\tau_C - t)/h, 1]$.

Moreover, for $t \in (s, \tau_C)$, the kernel $K_{s,t}$ satisfies that $\int_{(\tau_C - t)/h}^{(t-s)/h} K_{s,t}(x) dx = 1$ and $\int_{(\tau_C - t)/h}^{(t-s)/h} x K_{s,t}(x) dx = 0$. An example of the explicit form of $K_{s,t}$ is given in the Supplementary Materials. When $0 < t - s < \tau$, the function $\nu_j(s, t) = \nu_j(t, t)$, and $\nu_j(t, t)$ can be estimated by

$$\hat{\nu}_j(t, t) = \frac{\int_0^{\tau_C} \int_0^\infty K_{0,t}(\frac{t-u}{h}) m \hat{H}_j(u, dm, du)}{\int_0^{\tau_C} K_{0,t}(\frac{t-u}{h}) \hat{H}_j(u, \infty, du)}.$$

As defined above, $K_{0,t}$ is a second order modified kernel function adjusted for the interval $[0, \tau_C]$. Similarly, for $0 < s < t < \tau_C$, the function $\mu(s, t)$ can be estimated by

$$\hat{\mu}(s, t) = \frac{\sum_{j=1}^J \int_s^{\tau_C} \int_0^\infty K_{s,t}(\frac{t-u}{h}) m \hat{H}_j(s, dm, du)}{\sum_{j=1}^J \int_s^{\tau_C} K_{s,t}(\frac{t-u}{h}) \hat{H}_j(s, \infty, du)},$$

and for $0 < t - s < \tau_C$, the function $\mu(s, t)$ can be estimated by

$$\hat{\mu}(t, t) = \frac{\sum_{j=1}^J \int_0^{\tau_C} \int_0^\infty K_{0,t}(\frac{t-u}{h}) m \hat{H}_j(u, dm, du)}{\sum_{j=1}^J \int_0^{\tau_C} K_{0,t}(\frac{t-u}{h}) \hat{H}_j(u, \infty, du)}.$$

Note that when censoring is absent, the estimators $\hat{\nu}_j(s, t)$ and $\hat{\mu}(s, t)$ reduce to the corresponding Nadaraya-Watson estimator (Nadaraya, 1964; Watson, 1964) with boundary kernels. To study large sample properties of the proposed estimators, a few notations are introduced as follows: Denote by f_j the sub-density function of D due to cause j , and define

$f(t) = \sum_{j=1}^J f_j(t)$. The functions $f^{(1)}$ and $f_j^{(1)}$ denotes the derivative of f and f_j respectively.

Define $\kappa_{s,t} = \int_{(\tau_C - t)/h}^{(t-s)/h} x^2 K_{s,t}(x) dx$ and $\|K_{s,t}\|_2^2 = \int_{(\tau_C - t)/h}^{(t-s)/h} K_{s,t}(x)^2 dx$. Moreover, we define

$\eta_j(t) = \nu_j(t, t)$ and $\eta(t) = \mu(t, t)$. Let $\eta_j^{(1)}(t)$ and $\eta_j^{(2)}(t)$ denote the first and second order derivatives of the function $\eta_j(t)$. Similarly, let $\eta^{(1)}(t)$ and $\eta^{(2)}(t)$ denote the first and second order derivatives of the function $\eta(t)$. Theorem 5.1 states the asymptotic properties of $\hat{\nu}_j(s, t)$ and $\hat{\mu}(s, t)$. The proof is detailed in Section S.5 of the Supplementary Material.

Theorem 5.1

(i) Under Assumptions (A1'), (A2), (A4'), (A5) in Appendix, for $j = 1, \dots, J$ and $0 < s < t < \tau$, $(nh)^{1/2}\{\hat{\nu}_j(s, t) - \nu_j(s, t) - b_{\nu_j, h}(s, t)\}/\sigma_{\nu_j}(s, t)$ converges in distribution to a standard normal random variable, where

$$b_{\nu_j, h}(s, t) = \frac{1}{2} \kappa_{s, t} h^2 \left\{ \frac{\partial^2 \nu_j(s, t)}{\partial t^2} + 2 \frac{\partial \nu_j(s, t)}{\partial t} f_j^{(1)}(t) f_j(t)^{-1} \right\},$$

and

$$\sigma_{\nu_j}(s, t)^2 = \frac{\|K_{s, t}\|_2^2}{f_j(t) S_C(t)} [E\{M^{\text{total}}(s)^2 | D=t, \Pi=j\} - \nu_j(s, t)^2].$$

For $j = 1, \dots, J$ and $0 < t < \tau$, $(nh)^{1/2}\{\hat{\nu}_j(t, t) - \nu_j(t, t) - b_{\nu_j, h}(t)\}/\sigma_{\nu_j}(t)$ converges in distribution to a standard normal random variable, where

$$b_{\nu_j, h}(t) = \frac{1}{2} \kappa_{0, t} h^2 \left\{ \eta_j^{(2)}(t) + 2 \eta_j^{(1)}(t) f_j^{(1)}(t) f_j(t)^{-1} \right\},$$

and

$$\sigma_{\nu_j}(t)^2 = \frac{\|K_{0, t}\|_2^2}{f_j(t) S_C(t)} [E\{M^{\text{total}}(t)^2 | D=t, \Pi=j\} - \nu_j(t, t)^2].$$

(ii) Under Assumptions (A1), (A2), (A4), (A5), for $0 < s < t < \tau_C$, $(nh)^{1/2}\{\hat{\mu}(s, t) - \mu(s, t) - b_{\mu, h}(s, t)\}/\sigma_{\mu}(s, t)$ converges in distribution to a standard normal random variable, where

$$b_{\mu, h}(s, t) = \frac{1}{2} \kappa_{s, t} h^2 \left\{ \frac{\partial^2 \mu(s, t)}{\partial t^2} + 2 \frac{\partial \mu(s, t)}{\partial t} f^{(1)}(t) f(t)^{-1} \right\},$$

and

$$\sigma_{\mu}(s, t)^2 = \frac{\|K_{s, t}\|_2^2}{f(t) S_C(t)} [E\{M^{\text{total}}(s)^2 | D=t\} - \mu(s, t)^2].$$

For $0 < t < \tau_C$, $(nh)^{1/2}\{\hat{\mu}(t, t) - \mu(t, t) - b_{\mu, h}(t)\}/\sigma_{\mu}(t)$ converges in distribution to a standard normal random variable, where

$$b_{\mu, h}(t) = \frac{1}{2} \kappa_{0, t} h^2 \left\{ \eta^{(2)}(t) + 2 \eta^{(1)}(t) f^{(1)}(t) f(t)^{-1} \right\},$$

and

$$\sigma_{\eta}(t)^2 = \frac{\|K_{0,t}\|_2^2}{f(t)S_c(t)} [E\{M^{\text{total}}(t)^2 | D=t\} - \mu(t, t)^2].$$

For the estimator $\hat{v}_j(s, t)$ ($0 < s < t < \tau$), the bias function $b_{v_j, h}(s, t)$ is proportional to the second moment of the kernel and a measure of local curvature of $v_j(s, \cdot)$ at t . Moreover, the variance $\sigma_{v_j}(s, t)$ depends on censoring time distribution and increase as censoring becomes heavy. Similar conclusion can also be drawn for the estimators $\hat{v}_j(t, t)$, $\hat{\mu}(t, t)$, and $\hat{\mu}(s, t)$.

Remark 5.1

When $0 < t < s < \tau$, we have $\mu(s, t) = E\{M^{\text{total}}(D) | D = t\}$, and $v_j(s, t) = E\{M_j(D) | D = t, \Pi = j\}$, which are the lifetime utility cost adjusting for survival time. As discussed above, one way to estimate $E\{M^{\text{total}}(D) | D = t\}$ is using $\hat{\mu}(t, t)$. Another estimator of $E\{M^{\text{total}}(D) | D = t\}$ can be constructed as an extension of the methodology of Section 3.1, which is

$$\tilde{\mu}(t, t) = \frac{\int_0^{\tau_C} K_{0,t}(\frac{t-u}{h}) d\tilde{E}\{M^{\text{total}}(D)I(D \leq u)\}}{\sum_{j=1}^J \int_0^{\tau_C} K_{0,t}(\frac{t-u}{h}) \hat{H}_j(t, \infty, du)},$$

where $\tilde{E}\{M^{\text{total}}(D)I(D \leq u)\} = \tilde{\Phi}(u) - \frac{\sum_{i=1}^n M_i^{\text{total}}(u)I(X_i > u)}{\sum_{i=1}^n I(X_i > u)} \cdot \hat{S}_D(u)$ utilizes recurrent marker history information. However, since kernel type estimators are locally smooth and the recurrent marker history information carried in the kernel smoothers is ignorable, it can be proved that $\tilde{\mu}(t, t)$ is less efficient than $\hat{\mu}(t, t)$. Similarly, another estimator of $v_j(t, t)$, based on $\tilde{E}\{M^{\text{total}}(D)I(D \leq u)\}$, can be constructed as

$$\tilde{v}_j(t, t) = \frac{\int_0^{\tau_C} K_{0,t}(\frac{t-u}{h}) d\tilde{E}\{M^{\text{total}}(D)I(D \leq u)\} - \sum_{k \neq j} \int_0^{\tau_C} \int_0^{\infty} K_{0,t}(\frac{t-u}{h}) m \hat{H}_k(u, dm, du)}{\int_0^{\tau_C} K_{0,t}(\frac{t-u}{h}) \hat{H}_j(t, \infty, du)}.$$

Again, $\tilde{v}_j(t, t)$ does not contain additional information and is less efficient than $\hat{v}_j(t, t)$. Therefore, there is no need to apply the methodology in Section 4 to improve the efficiency of estimating $v_j(t, t)$.

6 Numerical Studies

6.1 Simulations

A series of simulation experiments are carried out to examine finite sample performance of the proposed methods. Two types of failure event is considered for the simulation studies. The data are generated from a three level model: (i) The type- j survival time is generated using a shared frailty Weibull model $\lambda_j(t) = Z\lambda_{0j}(t)$, where Z follows gamma distribution with mean 1 and variance 0.1 and $\lambda_{0j}(t) = \alpha_j t^{\beta_j} \beta_j^{\alpha_j-1} / \beta_j$. (ii) Conditioning on $D = t$ and $\Pi = j$, the recurrent event process $N(\cdot)$ is a non-stationary Poisson process with rate function $E\{dN(u) | D = t, \Pi = j\} = h_j(u, t) du$. (iii) Conditioning on $D = t$, $\Pi = j$ and $dN(u) = 1$, the utility associated with the recurrent event follows gamma distribution with mean m_j and

variance v_j . We consider four simulation scenarios. In Scenario I, we set $\alpha_1 = \alpha_2 = 5$, $\beta_1 = \beta_2 = 5$, $h_1(u, t) = h_2(u, t) = 10t^{-1}(1 + e^{-u})$, and $m_1 = m_2 = 2$, $v_1 = v_2 = 0.1$. In Scenario II, we set $\alpha_1 = \alpha_2 = 5$, $\beta_1 = 4.5$, $\beta_2 = 6$, $h_1(u, t) = h_2(u, t) = 10t^{-1}(1 + e^{-u})$, $m_1 = m_2 = 2$, $v_1 = v_2 = 0.1$. In Scenario III, we set $\alpha_1 = \alpha_2 = 5$, $\beta_1 = \beta_2 = 5$, $h_1(u, t) = 10t^{-1}(1 + e^{-u})$, $h_2(u, t) = 5t^{-1}(1 + e^{-u})$, and $m_1 = m_2 = 2$, $v_1 = 0.2$, $v_2 = 0.1$. In Scenario IV, we set $\alpha_1 = \alpha_2 = 5$, $\beta_1 = 4.5$, $\beta_2 = 6$, $h_1(u, t) = 10t^{-1}(1 + e^{-u})$, $h_2(u, t) = 5t^{-1}(1 + e^{-u})$, $m_1 = m_2 = 2$, $v_1 = 0.2$, $v_2 = 0.1$. Moreover, we set $\tau = 4.5$, and subjects with $D > \tau$ belong to the third risk type. The censoring time C is generated from exponential distribution with different rate parameters to produce 25%, 50% and 75% censoring rates.

In the first set of simulation, we evaluate the estimators proposed in Section 3 and 4. Under the above four scenarios, we generate 1000 datasets, each with a sample size of 400. Table 1 presents the summary statistics of $\hat{\Phi}_1^{\text{imp}}$, $\hat{\Phi}_1$ and $\tilde{\Phi}_1$ (The results of risk type 2 and 3 are presented in Supplemental Materials due to limit on space). It can be seen that the proposed estimators $\hat{\Phi}_j$, $\tilde{\Phi}_j$ and $\hat{\Phi}_j^{\text{imp}}$ all perform well with moderate sample sizes. In Scenarios II, III, IV, $\tilde{\Phi}_1$ has smaller standard error than $\hat{\Phi}_1$, while the standard error of $\hat{\Phi}_1$ outperforms that of $\tilde{\Phi}_1$ in Scenario I. In whichever case, the improved estimator $\hat{\Phi}_j^{\text{imp}}$ is either more efficient than or as efficient as $\hat{\Phi}_j$ and $\tilde{\Phi}_j$. As expected, $\hat{\Phi}_j^{\text{imp}}$ gains more efficiency as the censoring rate increases. Table 2 presents the summary statistics of $\hat{\Phi}_1^{c, \text{imp}}$, $\hat{\Phi}_1^c$ and $\tilde{\Phi}_1^c$ (The results of risk type 2 and 3 are presented in Supplemental Materials), similarly, all estimators perform well and the improved estimator $\hat{\Phi}_j^{c, \text{imp}}$ is more efficient or as efficient as estimators $\hat{\Phi}_j^c$ and $\tilde{\Phi}_j^c$. The variance estimation is reasonably accurate. The coverage probabilities of both Wald-type and accelerated bootstrap confidence intervals (Efron, 1987) are presented. The Wald-type confidence interval may not be satisfactory in a few settings (for example, $\hat{\Phi}_2^{c, \text{imp}}$ with 75% heavy censoring). We note that, with 75% heavy censoring rate and sample size of 400, there may not be adequate number of uncensored observations in each risk group to ensure the large sample properties of $\hat{\Phi}_j^c$ and $\hat{\Phi}_j^{c, \text{imp}}$ to hold - for example, in Scenario II and IV, the average number of uncensored observations is around 15 for risk-type 2, which might explain why the Wald-type confidence interval is not satisfactory in such cases. In contrast, the bootstrap confidence interval exhibits notable improvement.

We also evaluate the nonparametric estimators proposed in Section 5. We generate 1000 datasets from scenario IV, each with a sample size of 800. The censoring time is generated from exponential distribution to produce a 25% censoring rate. We use the boundary kernel adjusted on interval $[2, 5.5]$, because the probability that D is not in $[2, 5.5]$ is ignorable. Figure 1 presents the true curves as well as the averages of the estimated curves. Specifically, the left panel contains the true curves and average of estimates of $v_1(2, t)$, $\mu(2, t)$ and $v_2(2, t)$ on the interval $[2, 5.5]$, the middle panel contains true curves and estimates for $v_1(3, t)$, $\mu(3, t)$ and $v_2(3, t)$ on the interval $[3, 5.5]$; the right panel contains true curves and estimates for $v_1(t, t)$, $\mu(t, t)$ and $v_2(t, t)$ on the interval $[2, 5.5]$. The simulation results in the figure confirm the validity of the proposed nonparametric estimators.

6.2 A SEER-Medicare data example

The proposed methods are applied to SEER–Medicare linked database; see Warren et al. (2002) for an overview of the data. For illustration, we assess the medical cost of breast cancer patients diagnosed at age 65+ during 1994–1999 among Medicare enrollees from counties in metro areas. We take the time of first diagnosis of breast cancer to be the time origin, and D is the time from first diagnosis of breast cancer to death. $N(\cdot)$ is the counting process that characterizes the number of inpatient or outpatient cares, and upon each occurrence of inpatient or outpatient cares, $Y(t)$ is the cost charged for medical treatment, adjusted to year 2004 value by the medical care component of the consumer price index, available from the website of the U.S. Department of Labor (<http://www.bls.gov/cpi/>). As a well known fact, cardiovascular disease competes with breast cancer as the leading cause of death for older females diagnosed with breast cancer (Patnaik et al., 2011). In particular, we are interested in three competing terminal events within ten years since diagnosis of breast cancer: (i) breast cancer mortality ($\Pi = 1$), (ii) death from a cardiovascular disease ($\Pi = 2$), (iii) mortality from other causes ($\Pi = 3$). The subjects are divided into two groups by the historic stage determined at diagnosis: 12580 subjects with localized stage and 4777 subjects with regional stages. In the following analysis, the cost accumulation process from first diagnosis to $\tau = 10$ years is of interest, and people who survive more than 10 years are classified to the fourth category (long-term survivors).

We begin with estimating the cumulative incidence in standard competing risks model. For patients diagnosed with localized stage, the 10-year cumulative incidence was 7.7% (SE : 0.3%) for breast cancer (BC) deaths, 13.6% (SE : 0.4%) for cardiovascular disease (CVD) death, and 27.3% (SE : 0.5%) for other cause mortality. For patients diagnosed with regional stage, the 10-year cumulative incidence was 28.5% (SE : 0.8%) for BC deaths, 15.1% (SE : 0.6%) for CVD death, and 24.7% (SE : 0.8%) for other cause mortality. Thus at the end of the tenth year, the regional stage group has larger proportion of patients with BC death than the localized stage group, and the two groups has similar proportions of CVD death and death due to other causes.

We then analyze the medical cost up to a time horizon $\tau = 10$ years with our proposed methods. The results are presented in Table 3. When we do not distinguish the three types of death, the estimator in Section 3.1 is employed to estimate the total medical cost. It can be seen that the regional stage group has higher average ten-year medical cost than the localized stage group. We further take into account the three competing risks, $\Pi = 1, 2, 3$, and the estimates are obtained by using the improved estimators in Section 4. For each of the localized and regional stage groups, the average cost of patients with CVD death ($\Phi_2^c(\tau)$) is much higher than the average cost of patients with BC mortality ($\Phi_1^c(\tau)$). In contrast, for the overall spending of medical cost, $\Phi(\tau)$, patients with other mortality spent the most when compared with BC and CVD mortality, which can be explained by the large proportion of patients of with other cause of death at the tenth year.

For a better understanding of costs associated with BC and CVD mortality, we present plots of functions that characterize medical cost accumulation over time in Figure 2 for localized and regional stage group. The plots in the first row present the cumulative incidence

functions over time. For the localized stage patients, larger proportion of subjects died of CVD, while for regional stage patients, BC was the main cause of death. The plots in the second row present the overall spending of medical cost. For the localized stage patients, cost attributed to CVD is more than cost attributed to BC, as opposed to regional stage patients. From plots in the third row, in both stages, CVD mortality patients has higher cost than BC mortality patients. Thus for the regional group, the higher cost attributed to BC mortality (right panel of the second row) is explained by the larger proportion of patients of with BC mortality at the tenth year. The plots in the fourth row show the relationship between the total amount of 10-year medical cost and survival time. Given the same survival time, for localized stage patients, patients who died of CVD have higher 10-year medical cost than BC mortality patients; for regional stage patients, the trend is similar except that for patients live longer than 8.5 years, BC mortality patients have higher 10-year medical cost than CVD mortality patients given the same survival time.

7 Discussion

This paper proposed nonparametric estimators of the mean recurrent marker process, with a specific focus on competing risks model. In Section 3.1, we considered a nonparametric estimation approach which uses marker history information from both censored and uncensored subjects, but the estimator cannot be generalized to handle problems involving competing risks because the risk type information is unknown for those censored subjects. A consistent and asymptotically normally distributed estimator of type- j mean function is then constructed in Section 3.2 under competing risks model, where the proposed estimator only uses risk type information from uncensored subjects. Furthermore, using auxiliary information from the estimate of the non-competing risks mean function, an optimal estimator among a class of weighted estimators is proposed in Section 4 to improve the estimation efficiency over the two estimators proposed in Section 3.2. Such an improvement can be viewed as a nonparametric approach of using auxiliary information, in the spirit of generalized method of moments (GMM) (Hansen, 1982, 2001).

It is important to distinguish the roles of terminal event in the two different scenarios considered by Wang et al. (2001) and the current work. The former paper focused on marker effects explained by treatment or risk factor, where the presence of terminal event becomes a possible source of informative censoring. As a contrast, the latter work deals with utility marker consumed by survivors who have not yet advanced to the terminal event (for example, evaluation of medical cost before death), where the terminal event naturally ends the consumption of utility and therefore is not treated as a source of informative censoring. The estimation results in the current manuscript would be especially useful from the perspective of providers, such as government or insurance company.

This article considers only one-sample estimation while extension from non-parametric estimation to regression analysis is possible and of interest, and these extensions will be considered as future research topics.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported in part by the National Institute of Health grants R01HL122212 and R01CA193888.

Appendix

The regularity conditions used in the theorems are as follows.

- (A1) The censoring time C is independent of $\{D, N(t), Y(t)dN(t); 0 \leq t \leq D\}$ and $\Pr(X = \tau) > 0$. The distribution of C and D are continuous.
- (A1') The censoring time C is independent of $\{D, N(t), Y(t)dN(t), \Pi; 0 \leq t \leq D\}$ and $\Pr(X = \tau) > 0$. The distribution of C and D are continuous.
- (A2) The stochastic process $N(t)$ and $Y(t)dN(t)$ are bounded for $t \in [0, \tau]$.
- (A3) The covariance matrix $\Sigma_f(t, t)$ is nonsingular for $t \in [0, \tau]$.
- (A4) The second order derivative of $\mu(s, t)$ with respect to t exists for $t \in [s, \tau_C]$, and the second order derivative of $\eta(t) (\equiv \mu(t, t))$ with respect to t exists for $t \in [0, \tau_C]$. The second order derivative of $f(t)$ exists for $t \in [0, \tau_C]$.
- (A4') For $j = 1, \dots, J$, the second order derivative of $v_j(s, t)$ with respect to t exists for $t \in [s, \tau]$, and the second order derivative of $\eta_j(t) (\equiv v_j(t, t))$ with respect to t exists for $t \in [0, \tau]$. The second order derivative of $f_j(t)$ exists for $t \in [0, \tau]$.
- (A5) The bandwidth h satisfies $h = h_0 n^{-1/5}$ for some constant h_0 .

References

- Bang H, Tsiatis AA. Estimating Medical Costs With Censored Data. *Biometrika*. 2000; 87:329–343.
- Cook RJ, Lawless JF. Marginal Analysis of Recurrent Events and a Terminating Event. *Statistics in Medicine*. 1997; 16:911–924. [PubMed: 9160488]
- Efron B. Better Bootstrap Confidence Intervals. *Journal of the American statistical Association*. 1987; 82:171–185.
- Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a Competing Risk. *Journal of the American statistical association*. 1999; 94:496–509.
- Ghosh D, Lin DY. Nonparametric Analysis of Recurrent Events and Death. *Biometrics*. 2000; 56:554–562. [PubMed: 10877316]
- Glasziou PP, Simes RJ, Gelber RD. Quality Adjusted Survival Analysis. *Statistics in Medicine*. 1990; 9:1259–1276. [PubMed: 2277877]
- Gray RJ. A Class of K-sample Tests for Comparing the Cumulative Incidence of a Competing Risk. *Annals of Statistics*. 1988; 16:1141–1154.
- Hansen LP. Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica: Journal of the Econometric Society*. 1982; 50:1029–1054.
- Hansen LP. Generalized Method of Moments Estimation: A Time Series Perspective. *International Encyclopedia of Social and Behavioral Sciences*. 2001
- Huang C-Y, Wang M-C. Joint Modeling and Estimation for Recurrent Event Processes and Failure Time Data. *Journal of the American Statistical Association*. 2004; 99:1153–1165. [PubMed: 24068850]

- Huang Y. Calibration Regression of Censored Lifetime Medical Cost. *Journal of the American Statistical Association*. 2002; 97:318–327.
- Huang Y, Louis TA. Nonparametric Estimation of the Joint Distribution of Survival Time and Mark Variables. *Biometrika*. 1998; 85:785–798.
- Kalbeisch, JD., Prentice, RL. *The Statistical Analysis of Failure Time Data*. 2nd. New York: Willey; 2002.
- Lin D, Feuer EJ, Etzioni R, Wax Y. Estimating Medical Costs from Incomplete Follow-up Data. *Biometrics*. 1997;419–434. [PubMed: 9192444]
- Lin DY. Linear Regression Analysis of Censored Medical Costs. *Biostatistics*. 2000a; 1:35–47. [PubMed: 12933524]
- Lin DY. Proportional Means Regression for Censored Medical Costs. *Biometrics*. 2000b; 56:775–778. [PubMed: 10985215]
- Liu L, Wolfe RA, Huang X. Shared Frailty Models for Recurrent Events and a Terminal Event. *Biometrics*. 2004; 60:747–756. [PubMed: 15339298]
- Müller H-G. Smooth Optimum Kernel Estimators Near Endpoints. *Biometrika*. 1991; 78:521–530.
- Nadaraya EA. On Estimating Regression. *Theory of Probability & Its Applications*. 1964; 9:141–142.
- Patnaik JL, Byers T, DiGiuseppi C, Dabelea D, Denberg TD. Cardiovascular Disease Competes with Breast Cancer as the Leading Cause of Death for Older Females Diagnosed with Breast Cancer: A Retrospective Cohort Study. *Breast Cancer Res*. 2011; 13:R64. [PubMed: 21689398]
- Prentice RL, Kalbeisch JD, Peterson AV Jr, Flournoy N, Farewell VT, Breslow NE. The Analysis of Failure Times in the Presence of Competing Risks. *Biometrics*. 1978:541–554. [PubMed: 373811]
- Schaubel DE, Zhang M. Estimating Treatment Effects on the Marginal Recurrent Event Mean in the Presence of a Terminating Event. *Lifetime Data Analysis*. 2010; 16:451–477. [PubMed: 20063183]
- Sinha D, Maiti T, Ibrahim JG, Ouyang B. Current Methods for Recurrent Events Data With Dependent Termination: A Bayesian Perspective. *Journal of the American Statistical Association*. 2008; 103:866–878. [PubMed: 19169433]
- Strawderman RL. Estimating the Mean of an Increasing Stochastic Process at a Censored Stopping Time. *Journal of the American Statistical Association*. 2000; 95:1192–1208.
- Sun YQ, Gilbert PB, McKeague IW. Proportional Hazards Models with Continuous Marks. *Annals of Statistics*. 2009; 37:394–426. [PubMed: 19838313]
- Wang M-C, Qin J, Chiang C-T. Analyzing Recurrent Event Data with Informative Censoring. *Journal of the American Statistical Association*. 2001; 96:1057–1065.
- Warren JL, Klabunde CN, Schrag D, Bach PB, Riley GF. Overview of the SEER-Medicare Data: Content, Research Applications, and Generalizability to the United States Elderly Population. *Medical care*. 2002; 40:IV–3.
- Watson GS. Smooth Regression Analysis. *Sankhyā : The Indian Journal of Statistics, Series A*. 1964:359–372.
- Ye Y, Kalbeisch JD, Schaubel DE. Semiparametric Analysis of Correlated Recurrent and Terminal Events. *Biometrics*. 2007; 63:78–87. [PubMed: 17447932]
- Zeng D, Lin DY. Semiparametric Transformation Models with Random Effects for Joint Analysis of Recurrent and Terminal Events. *Biometrics*. 2009; 65:746–752. [PubMed: 18945267]
- Zhao H, Tian L. On Estimating Medical Cost and Incremental Cost-Effectiveness Ratios with Censored Data. *Biometrics*. 2001; 57:1002–1008. [PubMed: 11764238]

Figure 1.

The true curves (solid) and estimated curves (dashed) for $\{v_1(s, t), \mu(s, t), v_2(s, t)\}$ (upper, middle and lower curves) with $s = 2$ (left panel), $s = 3$ (middle panel) and $s = t$ (right panel).

Figure 2.

Cost analysis under competing risks model

Note: The left panels are for localized stage patients and the right panels are for regional stage patients. The solid line is for BC ($j = 1$) and the dashed line is for CVD ($j = 2$). The first row shows the cumulative incidence function for BC and CVD mortality with the corresponding pointwise 95% confidence interval. The second row shows the cost ($\hat{\Phi}_j^{\text{imp}}(t)$) attributed to BC and CVD mortality overtime with pointwise 95% confidence interval. The

third row shows the cost accumulation ($\hat{\Phi}_j^{c, \text{imp}}(t)$) of BC and CVD mortality patients over time with pointwise 95% confidence interval. The fourth row shows the total amount of lifetime medical cost adjusting for survival and risk type, specifically, the x-axis denotes the survival time and y-axis is the total amount of lifetime medical cost given the survival time t ($v_j(\tau, t)$).

Table 1

Simulation summary statistics for $\hat{\Phi}_1^{\text{imp}}$, $\hat{\Phi}_1$ and $\tilde{\Phi}_1$

Cen	t	$\widehat{\Phi}_1^{imp}(t)$					$\hat{\Phi}_1(t)$			$\hat{\Phi}_1(t)$		
		Bias	SE	SEE	CP _W	CP _B	Bias	SE	RE	Bias	SE	RE
Scenario I												
75%	1	0.003	0.049	0.049	0.940	0.952	0.003	0.052	1.122	0.004	0.056	1.276
	2	0.004	0.082	0.081	0.943	0.950	0.005	0.086	1.098	0.005	0.087	1.126
	3	0.007	0.106	0.105	0.943	0.952	0.007	0.111	1.091	0.008	0.110	1.078
50%	1	0.001	0.037	0.037	0.957	0.956	0.002	0.038	1.111	0.001	0.039	1.144
	2	0.002	0.060	0.060	0.949	0.949	0.003	0.062	1.080	0.002	0.062	1.060
	3	0.003	0.076	0.077	0.953	0.951	0.004	0.079	1.062	0.003	0.078	1.041
25%	1	0.001	0.032	0.032	0.942	0.948	0.001	0.032	1.047	0.001	0.033	1.072
	2	0.001	0.050	0.052	0.959	0.954	0.001	0.051	1.035	0.001	0.051	1.037
	3	0.001	0.065	0.066	0.947	0.954	0.002	0.066	1.030	0.002	0.066	1.022
Scenario II												
75%	1	3e-04	0.053	0.053	0.949	0.953	0.001	0.060	1.298	0.001	0.055	1.078
	2	0.003	0.086	0.086	0.947	0.951	0.003	0.096	1.224	0.003	0.088	1.040
	3	0.004	0.112	0.111	0.941	0.951	0.004	0.122	1.182	0.005	0.113	1.016
50%	1	0.002	0.039	0.041	0.959	0.955	0.002	0.043	1.187	0.003	0.040	1.035
	2	0.003	0.065	0.065	0.950	0.948	0.004	0.069	1.121	0.004	0.065	1.013
	3	0.005	0.083	0.083	0.943	0.953	0.005	0.087	1.107	0.006	0.083	1.002
25%	1	-0.001	0.035	0.036	0.948	0.948	-0.001	0.037	1.091	-2e-04	0.036	1.013
	2	-0.001	0.056	0.057	0.949	0.952	-0.001	0.058	1.072	-4e-05	0.056	1.000
	3	-0.001	0.071	0.072	0.952	0.951	-0.001	0.073	1.066	1e-04	0.071	0.997
Scenario III												
75%	1	-0.002	0.082	0.082	0.946	0.954	0.004	0.106	1.678	0.001	0.083	1.039
	2	-0.001	0.131	0.135	0.956	0.956	0.006	0.170	1.678	0.002	0.132	1.014
	3	-0.004	0.172	0.177	0.955	0.960	0.005	0.219	1.606	-1e-04	0.172	1.000
50%	1	-0.006	0.064	0.066	0.958	0.962	-0.004	0.074	1.345	-0.002	0.064	1.019
	2	-0.009	0.106	0.108	0.958	0.959	-0.007	0.121	1.295	-0.005	0.106	0.999

Cen	t	$\hat{\Phi}_1^{imp}(t)$					$\Phi_1(t)$			$\Phi_1(t)$		
		Bias	SE	SEE	CP _W	CP _B	Bias	SE	RE	Bias	SE	RE
25%	3	-0.010	0.137	0.140	0.955	0.959	-0.008	0.154	1.268	-0.005	0.137	0.999
	1	2e-04	0.061	0.061	0.949	0.945	0.004	0.068	1.218	0.004	0.061	1.002
	2	0.002	0.098	0.099	0.946	0.948	0.007	0.107	1.187	0.006	0.098	1.002
	3	0.003	0.125	0.126	0.951	0.953	0.009	0.136	1.190	0.008	0.125	1.007
Scenario IV												
75%	1	-4e-04	0.096	0.095	0.947	0.950	1e-04	0.122	1.624	0.002	0.096	1.006
	2	0.001	0.157	0.155	0.947	0.954	-0.003	0.191	1.486	0.004	0.156	0.998
	3	-0.001	0.205	0.203	0.947	0.951	-0.006	0.245	1.428	0.001	0.205	0.994
50%	1	-0.004	0.078	0.077	0.945	0.948	-0.003	0.090	1.311	-0.002	0.078	0.992
	2	-0.004	0.126	0.123	0.940	0.943	-0.005	0.141	1.261	-0.002	0.125	0.989
	3	-0.002	0.159	0.158	0.945	0.947	-0.003	0.177	1.239	1e-04	0.158	0.994
25%	1	-0.001	0.075	0.070	0.931	0.932	0.001	0.080	1.131	0.001	0.075	1.001
	2	0.001	0.118	0.111	0.937	0.937	0.002	0.124	1.096	0.003	0.118	0.999
	3	-3e-04	0.149	0.140	0.933	0.938	0.001	0.156	1.103	0.001	0.149	1.002

Note: Cen is the censoring rate; Bias is the empirical bias; SE is the empirical standard error; SEE is the empirical coverage probability of the 95% Wald type confidence interval; CP_B is the empirical coverage probability of the 95% bootstrap confidence interval based on a replication number of 1000. RE is the ratio of empirical mean square error of the estimator $(\Phi_1 \text{ or } \hat{\Phi}_1)$ to empirical mean square error of $\hat{\Phi}_1^{imp}$.

Table 2

Simulation summary statistics for $\hat{\Phi}_1^{c,imp}$, $\hat{\Phi}_1^c$ and $\tilde{\Phi}_1^c$

Cen	t	$\widehat{\Phi}_1^{c,imp}(t)$					$\widehat{\Phi}_1^c(t)$				$\widetilde{\Phi}_1^c(t)$			
		Bias	SE	SEE	CP _w	CP _B	Bias	SE	RE	Bias	SE	RE		
Scenario I														
75%	1	0.002	0.083	0.084	0.951	0.948	0.004	0.093	1.271	0.007	0.123	2.230		
	2	0.001	0.117	0.116	0.944	0.951	0.005	0.129	1.214	0.008	0.170	2.127		
	3	0.004	0.133	0.131	0.946	0.946	0.007	0.144	1.178	0.014	0.192	2.100		
50%	1	-0.002	0.065	0.065	0.945	0.948	-0.001	0.068	1.114	0.001	0.084	1.668		
	2	-0.003	0.090	0.089	0.941	0.942	-0.001	0.094	1.101	1e-04	0.114	1.620		
	3	-0.004	0.102	0.099	0.930	0.932	-0.002	0.106	1.080	0.001	0.130	1.636		
25%	1	-4e-04	0.055	0.056	0.944	0.949	0.001	0.057	1.074	0.002	0.063	1.277		
	2	-0.001	0.076	0.077	0.949	0.947	0.001	0.078	1.059	0.001	0.085	1.256		
	3	-0.001	0.085	0.085	0.953	0.954	0.001	0.087	1.052	0.001	0.095	1.263		
Scenario II														
75%	1	-0.003	0.062	0.060	0.938	0.940	4e-04	0.075	1.475	0.001	0.073	1.396		
	2	-0.002	0.085	0.085	0.954	0.948	0.002	0.100	1.392	0.002	0.100	1.392		
	3	-0.001	0.098	0.096	0.947	0.946	0.004	0.113	1.330	0.003	0.117	1.410		
50%	1	0.001	0.047	0.048	0.952	0.955	0.003	0.053	1.287	0.003	0.051	1.185		
	2	-1e-04	0.068	0.066	0.947	0.939	0.001	0.074	1.204	0.003	0.073	1.141		
	3	-0.002	0.076	0.074	0.938	0.939	-0.001	0.082	1.177	0.001	0.082	1.157		
25%	1	-0.001	0.043	0.043	0.951	0.949	-1e-04	0.045	1.124	-1e-04	0.044	1.062		
	2	5e-04	0.059	0.059	0.954	0.954	0.002	0.062	1.105	0.002	0.060	1.050		
	3	0.002	0.065	0.065	0.948	0.953	0.003	0.068	1.092	0.003	0.066	1.050		
Scenario III														
75%	1	-0.015	0.172	0.167	0.943	0.947	-0.001	0.186	1.165	0.004	0.254	2.160		
	2	-0.018	0.243	0.237	0.941	0.958	-0.003	0.252	1.074	0.011	0.375	2.374		
	3	-0.021	0.273	0.268	0.934	0.947	-0.002	0.281	1.052	0.006	0.443	2.618		
50%	1	-0.010	0.128	0.129	0.943	0.954	-0.002	0.136	1.138	0.001	0.162	1.599		
	2	-0.008	0.179	0.180	0.937	0.949	0.001	0.187	1.092	0.004	0.238	1.776		

Cen	t	$\widehat{\Phi}_1^{c,imp}(t)$					$\widehat{\Phi}_1^c(t)$					$\widetilde{\Phi}_1^c(t)$				
		Bias	SE	SEE	CP _W	CP _B	Bias	SE	RE	Bias	SE	RE	Bias	SE	RE	
25%	3	−0.010	0.198	0.201	0.947	0.953	−0.001	0.204	1.060	0.003	0.278	1.966				
	1	−0.005	0.115	0.113	0.94	0.947	4e-04	0.121	1.090	0.002	0.125	1.175				
	2	−0.004	0.157	0.154	0.951	0.954	0.001	0.160	1.049	0.001	0.176	1.257				
75%	3	−0.005	0.170	0.170	0.952	0.953	3e-04	0.174	1.041	0.001	0.197	1.337				
	Scenario IV															
	1	−0.006	0.124	0.122	0.948	0.955	0.002	0.147	1.415	0.003	0.143	1.339				
50%	2	−0.007	0.174	0.174	0.949	0.955	0.005	0.197	1.280	0.006	0.209	1.437				
	3	−0.011	0.205	0.199	0.936	0.943	0.001	0.225	1.211	0.004	0.253	1.525				
	1	−0.010	0.094	0.097	0.952	0.952	−0.008	0.106	1.273	−0.003	0.103	1.192				
25%	2	−0.009	0.133	0.135	0.953	0.954	−0.007	0.144	1.175	−0.001	0.150	1.261				
	3	−0.009	0.151	0.152	0.943	0.946	−0.007	0.162	1.149	0.001	0.172	1.304				
	1	−3e-04	0.089	0.087	0.941	0.943	0.002	0.094	1.124	0.004	0.092	1.061				
25%	2	0.002	0.121	0.119	0.940	0.944	0.004	0.126	1.077	0.007	0.126	1.083				
	3	0.001	0.135	0.131	0.939	0.941	0.003	0.140	1.085	0.004	0.140	1.084				

Note: Cen is the censoring rate; Bias is the empirical bias; SE is the empirical standard error; SEE is the empirical coverage probability of the 95% Wald type confidence interval; CP_B is the empirical coverage probability of the 95% bootstrap confidence interval based on a replication number of 1000. RE is the ratio of empirical mean square error of the estimator $(\widehat{\Phi}_1^c \text{ or } \widetilde{\Phi}_1^c)$ to empirical mean square error of $\widehat{\Phi}_1^{c,imp}$.

Table 3

Analysis of SEER-Medicare Data

	Localized		Regional		
	Estimate	SE	Estimate	SE	
Total	$\Phi(\tau)$	103.43	1.58	109.21	2.88
BC mortality	$\Phi_1(\tau)$	8.02	0.54	25.72	1.22
	$\Phi_1^C(\tau)$	103.99	5.33	90.35	3.35
CVD mortality	$\Phi_2(\tau)$	19.45	0.88	20.61	1.50
	$\Phi_2^C(\tau)$	143.31	4.96	136.92	7.82
Other mortality	$\Phi_3(\tau)$	42.20	1.47	36.18	2.69
	$\Phi_3^C(\tau)$	154.28	4.37	146.88	9.57
Long-term survivors	$\Phi_4(\tau)$	33.83	1.17	26.22	2.11
	$\Phi_4^C(\tau)$	65.89	2.03	82.38	5.92

NOTE: Estimate is the cost in 1000 US Dollars at $\tau = 10$ year. SE is the standard error estimate. BC is for breast cancer and CVD is for cardiovascular disease.