Sensitivity of the discrete-time Kaplan–Meier estimate to nonignorable censoring: Application in a clinical trial

Tao Liu[a] and Daniel F. Heitjan[b]*†

Untestable assumptions about the association between survival and censoring times can affect the validity of estimates of the survival distribution including the Kaplan–Meier (KM) nonparametric maximum likelihood estimate (MLE). This paper explores the sensitivity of the KM curve to nonignorable censoring by extending the index of local sensitivity to nonignorability (ISNI; Troxel *et al.*, *Statistica Sinica*, 14, 1221–1237, 2004; Zhang and Heitjan, *Biometrics*, 62, 1260–1268, 2006) to the case of a nonparametric survival model. The method involves, first, specifying a coarse-data selection model to describe the association between the failure and censoring processes and then evaluating the slope of the nonparametric survival MLE ordinate with respect to a nonignorability parameter in the neighborhood of the ignorable model. We define the nonparametric MLE of the survival curve for a fixed value of the nonignorability parameter and show in a simulation that ISNI analysis effectively captures local sensitivity to nonignorability. The method measures sensitivity in the sense of identifying functionals of the nonparametric MLE that nonignorability, if present, can affect substantially. We demonstrate the method with an application to a trial comparing mechanical assistance to optimal medical management in the treatment of end-stage heart failure. Copyright © 2012 John Wiley & Sons, Ltd.

Keywords: coarse-data model; ignorability; informative censoring; ISNI; sensitivity analysis

1. Introduction

Right censoring is a type of data coarsening [1] commonly encountered in survivorship studies. Although it is well known that association between failure and censoring times can distort likelihood-based inferences, data analysts often accept without scrutiny an assumption of ignorable censoring. This is at least partly because it is well known that there is little information in most data sets to permit a robust assessment of the censoring mechanism.

Random censorship gives rise to data of a competing risks structure and consequently partakes of the identifiability issues associated with such data [2, 3]. Peterson [4] proposed an approach to evaluating the sensitivity to nonignorable censoring, deriving bounds for the survivor function across the range of possible censoring mechanisms. These bounds evaluate the impact of nonignorable censoring on survival estimates in extreme cases and therefore can be regarded as a form of *global* sensitivity analysis. Although sharp, they are unfortunately too wide to be of practical value in many applications.

If one is willing to restrict the set of survival and censoring distributions somewhat, one can derive proper inferences by jointly estimating the survival and censoring distributions by using either likelihood methods or estimating equations. Fitting such models is not trivial, but the approach is sound when one has confidence that the assumptions are valid [5].

[a]*Department of Biostatistics, Brown University, Providence, RI, U.S.A.*
[b]*Department of Biostatistics and Epidemiology, University of Pennsylvania, Philadelphia, PA, U.S.A.*
Correspondence to: Daniel F. Heitjan, Department of Biostatistics and Epidemiology, University of Pennsylvania, Philadelphia, PA 19104-6021, U.S.A.
†*E-mail: dheitjan@upenn.edu*

A less ambitious approach is to execute a sensitivity analysis by assuming a parametric model for the censoring mechanism and evaluating the limits of possible inferences for the survival model as the degree of nonignorability varies across its range. Scharfstein and colleagues [6–9] developed methods for estimating survival curves under informative censoring by using weighted estimating equations. To determine the weights, they proposed a 'censoring bias function' that contains a nonidentifiable/sensitivity parameter to capture the dependence of failure time on censoring after adjusting for measured factors. Evaluating the dependence over a range of plausible values for the sensitivity parameter, their methods allow for a *point-wise* sensitivity analysis of nonignorable censoring. Shardell *et al.* [10] described an approach that involves the solicitation of medical expert opinion for determining the plausible range of the sensitivity parameter.

Another approach is to evaluate the sensitivity of inferences to small departures from ignorability, that is, to conduct a *local* sensitivity analysis. The method of Siannis *et al.* [11, 12] assumes a parametric survival/censoring model in which the distribution of the censoring time, given the survival time, depends on the parameter of the survival model and a nonignorability parameter. Under this model, likelihood estimation must account for the censoring mechanism even if no data are censored. An alternative approach is that of Zhang and Heitjan [13] who derived an *index of local sensitivity to nonignorability (ISNI)* [14] for a general parametric model of survival and censoring. This method is also likelihood based and yields a true index of sensitivity that declines to zero together with the fraction of censored observations. Yet the need to specify a parametric survival model limits the applicability of these methods to situations where such an assumption is tenable.

The Kaplan–Meier (KM) estimate of the survival curve is valid without parametric assumptions on the underlying distribution but not without an assumption of ignorability of the censoring mechanism. Taking advantage of its interpretation as a nonparametric maximum likelihood estimate (MLE), however, one can apply the same local sensitivity analysis previously developed for parametric models. In this paper, we develop such an approach. Our method has several advantages over previous methods. First, it avoids specification of a parametric survival model. Second, unlike point-wise sensitivity analysis, it does not require repeatedly estimating a nonignorable model and so potentially avoids arduous computations. Third, it provides a valid index of sensitivity that, in the spirit of likelihood inference and of the Peterson bounds, vanishes when censoring is absent. And fourth, although we do require a model for the censoring process, it is possible to specify such a model with only weak assumptions.

Nonignorable censoring is a concern in clinical trials, where failure and censoring times may be correlated as a consequence of time trends in treatment effects or association of enrollment time with the health of entering patients [13, 15]. Although censored and missing data are both special cases of the coarse-data model, there are important differences between them with respect to sensitivity analysis. With a trial that is subject to potentially nonignorably missing observations, if the nonignorable mechanism and the fraction of missing observations are equal across arms, then comparisons between arms are insensitive to nonignorability [16]. This is not the case with nonignorable censoring, because the incompleteness of the data is a consequence of both the censoring mechanism and the survival distribution. The latter may well differ between arms even if the former does not.

The REMATCH trial [17], which compared mechanical assistance with optimal medical therapy in the treatment of advanced heart failure, provides an instructive example. Investigators had grounds to anticipate differential trends in survival between arms, as the device arm involved novel surgical and medical interventions that were difficult to master and still evolving as the trial began, whereas the medical arm involved only the administration of commonly used drugs. Thus, one could anticipate some correlation of survival and censoring times in the device arm. As expected, there was a strong treatment effect, with the device arm having much better survival and therefore a larger fraction of censoring. This would induce sensitivity even if there was a common nonignorable censoring mechanism that affected both arms equally. A previous analysis based on parametric survival distributions found moderate sensitivity, but results varied to some extent by the model being estimated [15]. Thus, in this paper, we seek to evaluate the sensitivity of inferences that are less reliant on parametric assumptions.

We organize the paper as follows. In Section 2, we review the coarse-data model and the ISNI. In Section 3, we derive the ISNI for KM and for functions of the survival curve such as moments and quantiles. In Section 4, we illustrate the method with simulated data, and in Section 5, we apply it to REMATCH. Section 6 concludes with further discussion.

2. The index of local sensitivity to nonignorability in the coarse-data model

2.1. Notation and the coarse-data model for right censoring

Let T be a discrete failure time and C its corresponding right censoring time, both taking values on a common, finite support $\mathbb{T} = \{t_{(1)}, \ldots, t_{(m)}\}$ with $0 < t_{(1)} < \cdots < t_{(m)}$. If the supports for T and C differ, we let $\mathbb{T}$ be their union. Let p_θ^T be the probability mass function of T, where θ is the governing parameter. Adopting a nonparametric perspective, we can think of θ as a vector $\theta = (\theta_1, \ldots, \theta_m)^\top$, where $\theta_k \geqslant 0$ is the probability mass on $t_{(k)} \in \mathbb{T}$ and $\sum_{k=1}^m \theta_k = 1$. We denote the survival and hazard functions by $S(t) = \Pr(T \geqslant t)$ and $\lambda(t) = S(t)^{-1}\Pr(T = t)$. Our interest is in estimating $S(t)$ or $\lambda(t)$, but we are concerned that a nonignorable right censoring mechanism may distort our estimates. Note that the survival function is a step function with jumps only at $\{S(t), t \in \mathbb{T}\}$, and the hazard function is 0 except for taking values $\lambda(t)$ at $t \in \mathbb{T}$. We henceforth represent the two functions by using the coordinates $S = (S_1, \ldots, S_m)^\top$ and $\lambda = (\lambda_1, \ldots, \lambda_m)^\top$ where $S_k = S(t_{(k)})$ and $\lambda_k = \lambda(t_{(k)})$.

In the case of random censorship, where we do not know the censoring time except when a subject's event time is actually censored, the right censoring is a data-coarsening process on failure time, and similarly, the mortality is a data-coarsening process on the censoring time [1]. Mathematically, the two simultaneous coarsening processes can be expressed by

$$Y(T, C) = \begin{cases} \{T\} & T \leqslant C \\ \mathbb{T} \cap (C, \infty) & \text{otherwise,} \end{cases}$$

and

$$G(C, T) = \begin{cases} \{C\} & T > C \\ \mathbb{T} \cap [T, \infty) & \text{otherwise,} \end{cases}$$

where (Y, G) are observable set outcomes. For data collected from n independent subjects, the (Y_i, G_i) pairs are independent copies of the observed, coarsened version of the (T_i, C_i) pairs for $i = 1, \ldots, n$.

Under general assumptions, the distribution of (Y, G) does not uniquely determine the failure time distribution p_θ^T; this is the so-called nonidentifiability problem. Thus, except in certain special cases, to obtain a correct estimate of θ, we must take into account the censoring mechanism. Throughout this paper, we use a selection model to describe the censoring mechanism

$$C \mid T \sim p_\gamma^{C|T},$$

where $p_\gamma^{C|T}$ denotes the conditional distribution of censoring time C for a given failure time T and γ is assumed to be distinct from θ in the sense of Rubin [18]. We further assume that γ can be decomposed as $\gamma = (\gamma_0, \gamma_1)^\top$, where γ_0 is a nuisance parameter (possibly a vector) and γ_1 is a sensitivity parameter that satisfies the following three conditions:

- γ_1 controls the magnitude of the nonignorability in the censoring mechanism;
- it is not identified by the data; and
- supplying a value of γ_1 allows the rest of the parameters to be identified.

In the sequel, we refer to γ_1 as the *nonignorability parameter*. Moreover, we assume that when $\gamma_1 = 0$, the censoring mechanism is ignorable in the sense that the random censoring times can be treated as if they are fixed prior to the study and independent of the failure times.

Given a censoring model, the likelihood of θ and γ conditional on the observable $(Y = y, G = g)$ is [19]

$$L(\theta, \gamma) = \sum_{t \in y} \left\{ p_\theta^T(t) \sum_{c \in g} p_\gamma^{C|T}(c|t) \right\}. \tag{1}$$

When $\gamma_1 = 0$, the likelihood inference for θ based on (1) is exactly the same as that based on

$$L_{\text{ign}}(\theta) = \sum_{t \in y} p_\theta^T(t). \tag{2}$$

This *ignorable likelihood* is the likelihood that one uses in applications that disregard the stochastic nature of the censoring mechanism.

2.2. Review of the index of local sensitivity to nonignorability

The ISNI [13, 14] is a measure of the extent to which small deviations in the nonignorability parameter around the ignorable value $\gamma_1 = 0$ affect likelihood-based inferences on the parameter of interest. We give here a brief review of the method. Throughout the paper, we will use ˜ (tilde) to denote the MLE under the complete-data likelihood (1) for a given γ_1 and ˆ (hat) to denote the MLE under the ignorable likelihood (2).

For ease of illustration, assume for the moment that γ_1 is a scalar. When γ_1 is sufficiently small, the log of the correct likelihood (1) $\ell(\theta, \gamma_0, \gamma_1) = \log L(\theta, \gamma_0, \gamma_1)$ can be approximated using the Taylor expansion

$$
\begin{aligned}
\ell(\theta, \gamma_0, \gamma_1) \approx{}& \ell(\theta, \gamma_0, 0) \\
&+ (\theta - \hat{\theta}, \gamma_0 - \hat{\gamma}_0, \gamma_1) \frac{\partial \ell}{\partial(\theta, \gamma_0, \gamma_1)^{\top}} \Big|_{(\hat{\theta}, \hat{\gamma}_0, 0)} \\
&+ \frac{1}{2}(\theta - \hat{\theta}, \gamma_0 - \hat{\gamma}_0, \gamma_1) \frac{\partial^2 \ell}{\partial(\theta, \gamma_0, \gamma_1)\partial(\theta, \gamma_0, \gamma_1)^{\top}} \Big|_{(\hat{\theta}, \hat{\gamma}_0, 0)} (\theta - \hat{\theta}, \gamma_0 - \hat{\gamma}_0, \gamma_1)^{\top}
\end{aligned}
$$

where $\hat{\theta}$ and $\hat{\gamma}_0$ are the MLEs assuming the ignorable model. We differentiate the right-hand side with respect to θ for fixed γ_1 and then solve for θ and, after some manipulation, obtain

$$
\tilde{\theta}(\gamma_1) \approx \hat{\theta} + \mathcal{I}_{\text{ign}}^{-1} \frac{\partial^2 \ell}{\partial \theta \partial \gamma_1} \Big|_{(\hat{\theta}, \hat{\gamma}_0, 0)} \gamma_1,
$$

where $\tilde{\theta}(\gamma_1)$ is the MLE of θ given the data and assuming fixed γ_1 and

$$
\mathcal{I}_{\text{ign}} = -\frac{\partial^2 \log L_{\text{ign}}}{\partial \theta \partial \theta^{\top}} \Big|_{\hat{\theta}}
$$

is the observed information for the ignorable likelihood. This expression motivates the definition of ISNI [14] as

$$
\text{ISNI}(\hat{\theta}) = \mathcal{I}_{\text{ign}}^{-1} \frac{\partial^2 \ell}{\partial \theta \partial \gamma_1} \Big|_{(\hat{\theta}, \hat{\gamma}_0, 0)} . \tag{3}
$$

In the neighborhood of the ignorable model (i.e., where $\gamma_1 = 0$), $\tilde{\theta}(\gamma_1) \approx \hat{\theta} + \gamma_1 \text{ISNI}$, so ISNI is the rate of change of the MLE $\tilde{\theta}$ as γ_1 varies around the ignorable model. Thus, ISNI measures the local sensitivity of the parameter estimate to departures from ignorability. A large value of ISNI implies that the ignorable likelihood estimator is sensitive to modest departures of the censoring model from ignorability, whereas a small value suggests that the estimator is locally robust. ISNI is typically easy to compute and avoids the requirement to repeatedly maximize $L(\theta, \gamma)$ under a range of nonignorable models.

When γ_1 is a q-dimensional vector, Xie and Heitjan [16] proposed an alternative definition of ISNI as

$$
\text{ISNI}(\hat{\theta}) = \sqrt{q} \left\| \frac{\partial \tilde{\theta}}{\partial \gamma_1^{\top}} \Big|_{(\hat{\theta}, \hat{\gamma}_0, 0)} \right\|^{1/2},
$$

which approximates the maximal sensitivity when the vector-valued nonignorability parameter lies on a ball of radius $\sqrt{q}$ around the origin; that is, $\text{ISNI} \approx \max_{\|\gamma_1\| = \sqrt{q}}(\tilde{\theta} - \hat{\theta})$.

2.3. Local sensitivity assessment

Because ISNI refers to the derivative of the MLE with respect to the nonignorability parameter and the nonignorability parameter depends on the measurement scale of T, the interpretation of ISNI also depends on the measurement scale. To avoid such dependence, Troxel *et al.* [14] defined the statistic

$$
\gamma_1^* = \left| \frac{\text{SE}(\hat{\theta})}{\text{ISNI}(\hat{\theta})} \right|, \tag{4}
$$

where the numerator is the standard error (SE) of $\hat{\theta}$ under the ignorable model. This γ_1^* is approximately the value of γ_1 that can cause a one-SE change in $\hat{\theta}$. If γ_1^* is implausibly large, then we would require an implausibly large amount of nonignorability to change inferences substantially, and therefore, we deem the estimation insensitive to nonignorability. If γ_1^* is a plausible value of the nonignorability parameter for the data at hand, then we conclude that $\hat{\theta}$ is sensitive to nonignorable censoring.

3. Evaluating sensitivity of Kaplan–Meier

To derive ISNI for the KM survival estimate, we use the fact that KM is also an MLE, albeit a nonparametric MLE in a model with a parameter of potentially high dimension. In the following, we commit a slight abuse of our notation $\mathbb{T}$ by letting it also denote the set of unique ordered failure and censoring times that are observed. We first derive ISNI for the KM hazard estimates, which turns out to be straightforward, and then derive ISNI for the KM survival estimates by using the one-to-one mapping between survival and hazard.

3.1. The index of local sensitivity to nonignorability for the hazard estimate

Let d_k and r_k be the numbers of subjects dead and at risk, respectively, at time $t_{(k)} \in \mathbb{T}$. The KM hazard estimate at $t_{(k)}$ is $\hat{\lambda}_k = d_k / r_k$, which is also the nonparametric MLE under the ignorable likelihood

$$L_{\text{ign}}(\theta) \propto \prod_{i=1}^{n} \left(\sum_{t_{(k)} \in y_i} \theta_k \right) = \prod_{k=0}^{m} \lambda_k^{d_k} (1 - \lambda_k)^{r_k - d_k}.$$

The first factor in the definition of ISNI (3) is the inverse information under the ignorable likelihood, which for the hazard estimate, is

$$\mathcal{I}_{\text{ign}}^{-1} = \text{Diag}\left\{ r_0^{-1} \hat{\lambda}_0 (1 - \hat{\lambda}_0), \dots, r_m^{-1} \hat{\lambda}_m (1 - \hat{\lambda}_m) \right\}. \tag{5}$$

For $d_j = 0$, we adopt the convention $0 \cdot \infty = 1$ to avoid ambiguities in calculating the matrix inverse.

To derive the second factor in (3), we write out the log of the correct likelihood (1) as the sum of log likelihoods for observed and censored death times

$$\ell(\theta, \gamma) = \sum_{\delta_i = 0} \left[\log \left\{ p_\theta^T(t_i) \sum_{c \in g_i} p_\gamma^{C|T}(c|t_i) \right\} \right] + \sum_{\delta_i = 1} \left[\log \left\{ \sum_{t \in y_i} p_\theta^T(t) p_\gamma^{C|T}(c_i|t) \right\} \right],$$

where $\delta_i = \mathbf{1}(t_i > c_i)$ indicates whether a right censoring occurs. When taking the second-order derivative $\frac{\partial^2}{\partial\theta\partial\gamma_1}$, the first term vanishes with the assumption that θ and γ are distinct [18]. Consistent with our intuition, for local sensitivity, what really matters is the sum of the log likelihood for censored failures. With some algebra, we find that

$$\frac{\partial^2 \ell}{\partial \lambda_k \partial \gamma_1} \bigg|_{\gamma_1 = 0, \lambda = \hat{\lambda}} = - \sum_{c_i < t_{(k)}} \left[\frac{\delta_i A_{\gamma_0}(c_i)}{p_{\gamma_0}^C(c_i)} \sum_{j \geq k} \frac{(t_{(j+1)} - t_{(j)}) \prod_{c_i < t_{(l)} \leq t_{(j)}} (1 - \hat{\lambda}_l)}{1 - \hat{\lambda}_k} \right], \tag{6}$$

where $A_{\gamma_0}(c) = \frac{\partial p_\gamma^{C|T}}{\partial(\gamma_1 t)} \big|_{\gamma_1 = 0}$ and $p_{\gamma_0}^C(c) = p_\gamma^{C|T}(c \mid t; \gamma_1 = 0)$, which does not depend on t. Now, bringing (5) and (6) into (3), we obtain the ISNI for the KM hazard estimates as

$$\text{ISNI}(\hat{\lambda}) = (\text{ISNI}(\hat{\lambda}_1), \dots, \text{ISNI}(\hat{\lambda}_m))^\top,$$

where

$$\text{ISNI}(\hat{\lambda}_k) = \frac{\hat{\lambda}_k}{r_k} \sum_{c_i < t_{(k)}} \left[\frac{\delta_i A_{\hat{\gamma}_0}(c_i)}{p_{\hat{\gamma}_0}^C(c_i)} \sum_{j \geq k} \{(t_{(j+1)} - t_{(j)}) \prod_{c_i < t_{(l)} \leq t_{(j)}} (1 - \hat{\lambda}_l)\} \right].$$

3.2. The index of local sensitivity to nonignorability for Kaplan–Meier

The KM ordinates are given by $\hat{S}_j = \prod_{k=0}^{j}(1 - \hat{\lambda}_k)$. To simplify the derivation, consider logs of the survival function:

$$\log S_k = \sum_{j=1}^{k} \log(1 - \lambda_j) \Rightarrow \frac{\partial \log S_k}{\partial \lambda_l} = \begin{cases} -\frac{1}{1-\lambda_l}, & l \leqslant k; \\ 0, & \text{o.w.} \end{cases}$$

By the chain rule, the ISNI of KM is

$$\text{ISNI}(\hat{S}_k) = \left\{ \frac{\partial(\log S_k)}{\partial S_k} \Big|_{S_k = \hat{S}_k} \right\}^{-1} \left\{ \frac{\partial(\log S_k)}{\partial \lambda} \Big|_{\lambda = \hat{\lambda}} \right\}^{\top} \text{ISNI}(\hat{\lambda}) \tag{7}$$

$$= \hat{S}_k \sum_{l=1}^{k} \left\{ \frac{d_l}{r_l^2} \sum_{c_i < t_{(l)}} \left[\frac{\delta_i A_{\hat{\gamma}_0}(c_i)}{p_{\hat{\gamma}_0}^{C}(c_i)} \sum_{j \geqslant l} \frac{(t_{(j+1)} - t_{(j)}) \prod_{c_i < t_{(z)} \leqslant t_{(j)}} (1 - d_z/r_z)}{1 - d_l/r_l} \right] \right\}.$$

This expression suggests that the KM curve is insensitive up to the time of the first censoring event. After that, its sensitivity depends on the fraction of data censored, and as KM approaches 0, its sensitivity declines to 0 again because the upper and lower bounds of the survival function limit the effect of nonignorability.

3.3. Example for a specified $p_{\gamma}^{C|T}$

In the preceding section, we derived ISNI for KM with a general nonignorable model. Here, we show that the nonignorable model $p_{\gamma}^{C|T}$ can have a semiparametric form such that we only need to explicitly specify the part that involves the interaction of γ_1 and the survival time.

Our example assumes that the censoring time, conditional on the failure time, follows a proportional hazards model $\lambda^{C|T}(c|t) = \lambda_0^{C}(c) \exp(\gamma_1 t)$. With this setup, the nuisance parameter is the baseline hazard of censoring $\gamma_0 = \{\lambda_0^{C}(c), c \in \mathbb{T}\}$, which can be left unspecified. Further, we can derive the selection model as $p_{\gamma}^{C|T}(c \mid t) = \lambda_0^{C}(c) \exp(\gamma_1 t) \prod_{t_{(j)}<c} [1 - \lambda_0^{C}(t_{(j)}) \exp(\gamma_1 t)]$ and calculate the following term of (7):

$$\frac{A_{\gamma_0}(c)}{p_{\gamma_0}^{C}(c)} = \frac{\partial \left\{ \log p_{\gamma}^{C|T}(c|t) \right\}}{\partial(\gamma_1 t)} \Big|_{\gamma_1 = 0} = 1 - \sum_{t_{(j)}<c} \frac{\lambda_0^{C}(t_{(j)})}{1 - \lambda_0^{C}(t_{(j)})}. \tag{8}$$

We can estimate the nuisance parameter $\{\lambda_0^{C}(c), c \in \mathbb{T}\}$ by the KM hazard estimator, treating censoring as an 'event' that can be censored by an event.

3.4. The index of local sensitivity to nonignorability for statistics derived from Kaplan–Meier

Because ISNI is simply a derivative, once one has computed it for a particular parameter, one can apply the chain rule to compute it for any function of that parameter. In this subsection, we derive ISNI for three scalar functions of the survival curve that are relevant in applications.

3.4.1. Mean survival time.
An estimate of mean survival is the area under the survival curve, $\hat{\mu}_T = \sum_{j=1}^{m} \hat{S}_{j-1}(t_{(j)} - t_{(j-1)})$. It follows immediately that

$$\text{ISNI}(\hat{\mu}_T) = \sum_{j=1}^{m} \text{ISNI}(\hat{S}_{j-1})(t_{(j)} - t_{(j-1)});$$

that is, ISNI for mean survival is just a linear combination of the ISNI values computed at each survival time.

3.4.2. Restricted mean survival. When late censoring is substantial or failure time is highly skewed, it may be preferable to consider τ-restricted mean survival, defined as the mean lifetime among a subpopulation that experiences failure before τ. Karrison [20] proposed the estimator $\hat{\mu}(\tau) = \sum_{t_{(j)} \leqslant \tau} \hat{S}_j \phi_j$, where $\phi_j = (\phi_j^* + \phi_{j-1}^*)/2$ with $\phi_0^* = 0$ and $\phi_j^* = t_{(j+1)} - t_{(j)}$. The ISNI for the τ-restricted mean life is therefore

$$\text{ISNI}(\hat{\mu}(\tau)) = \sum_{t_{(j)} \leqslant \tau} \text{ISNI}(\hat{S}_j)\phi_j.$$

3.4.3. The index of local sensitivity to nonignorability for a quantile. Another parameter of potential interest is the median survival, denoted by $T_{0.5}$, or more generally the wth quantile of survival T_w. In Appendix A, we show how to calculate the ISNI for the estimated wth quantile when failure time T is continuous. When $\mathbb{T}$ is discrete, there is ambiguity in the definition of the wth quantile. Henceforth, we take the following definition:

$$T_w = \frac{t_{(j*+1)}(S_{j*} - w) + t_{(j*)}(w - S_{j*+1})}{S_{j*} - S_{j*+1}}, \tag{9}$$

where $j^* = j : \text{Pr}(T < t_{(j)}) \leqslant w < \text{Pr}(T < t_{(j+1)})$. This is equivalent to regarding the survival function as continuous with linear interpolations between successive distinct survival probabilities over $\mathbb{T}$,

$$\bar{S}(t) = \sum_j \mathbf{1}(t \in [t_{(j-1)}, t_{(j)})) \frac{(t_{(j)} - t)S_{j-1} + (t - t_{(j-1)})S_j}{t_{(j)} - t_{(j-1)}}. \tag{10}$$

Note that definition (10) implies a probability density function

$$\bar{f}(t) = \sum_j \mathbf{1}(t \in [t_{(j-1)}, t_{(j)})) \frac{S_{j-1} - S_j}{t_{(j)} - t_{(j-1)}},$$

and it is easy to check that the definition in (9) can be expressed as $T_w = \bar{S}^{-1}(w)$.

We can estimate $\hat{T}_w$ by bringing the KM estimates into (9). Applying (15) (from Appendix A), we obtain that when $\hat{T}_w \notin \mathbb{T}$,

$$\text{ISNI}(\hat{T}_w) \approx \frac{(\hat{T}_w^+ - \hat{T}_w)\text{ISNI}(\hat{S}_w^-) + (\hat{T}_w - \hat{T}_w^-)\text{ISNI}(\hat{S}_w^+)}{\hat{S}_w^- - \hat{S}_w^+}, \tag{11}$$

and when $\hat{T}_w \in \mathbb{T}$,

$$\text{ISNI}(\hat{T}_w) \approx \begin{cases} \frac{(\hat{T}_w - \hat{T}_w^-)\text{ISNI}(\hat{S}_w)}{\hat{S}_w^- - \hat{S}_w}, & \text{ISNI}(\hat{S}_w)\gamma_1 < 0; \\ \frac{(\hat{T}_w^+ - \hat{T}_w)\text{ISNI}(\hat{S}_w)}{\hat{S}_w - \hat{S}_w^+}, & \text{ISNI}(\hat{S}_w)\gamma_1 > 0 \end{cases} \tag{12}$$

where $\hat{T}_w^-$ ($\hat{T}_w^+$) is the largest (smallest) value in $\mathbb{T}$ less (greater) than $\hat{T}_w$ and $\hat{S}_w^-$ ($\hat{S}_w^+$) is the KM ordinate at $\hat{T}_w^-$ ($\hat{T}_w^+$).

4. Simulation

We conducted a simulation to evaluate the ability of ISNI to approximate the impact of small departures from ignorability on the MLE. We generated 5000 independent failure times from the exponential distribution with mean 10 and discretized the data by taking the ceiling ($\lceil \cdot \rceil$) of each outcome; that is,

$$U \sim \text{EXP}(10) \quad \text{and} \quad T = \lceil U \rceil.$$

We simulated the corresponding censoring times from the nonignorable censoring mechanism

$$V|T \sim \text{EXP}(10\exp(\gamma_1 T)) \quad \text{and} \quad C = \lceil V \rceil.$$

We generated the data, assuming a value of the nonignorability parameter $\gamma_1 = 0.04$.

Results appear in Figures 1–3. In Figure 1, the KM curve for the simulated data is the thin black line, and the true survival curve is the gray line. With $\gamma_1 > 0$, people with long survival times have long censoring times, and those with short survival times have short censoring times. Thus, long survival times are over-represented in the data set, and the KM curve overestimate survival, as can be seen in every panel of Figure 1.

Also, in Figure 1, we present the self-consistent (SC; Appendix A) estimate of the survivor function as the dark solid lines along the KM and true survival curves, for the range of values of the nonignorability parameter $\gamma_1 = -0.01, 0, 0.01, 0.02, 0.04, 0.06$. That is, the thick black lines represent estimates of the survival curve assuming fixed values of the nonignorability parameter. Therefore, when $\gamma_1 = 0$, we are assuming ignorability, and the SC curve equals the KM curve. When we assume that $\gamma_1 = 0.04$, the nonignorable model is correctly specified, and the SC curve closely matches the true survival curve.

Figure 2 plots the SC survival probability estimates (shown as circles) against the values of γ_1 at times $1, 3, \ldots, 11$. The dashed lines have intercepts equal to the survival probabilities evaluated at $\gamma_1 = 0$ (the KM estimates) and slopes equal to ISNI, indicating that ISNI successfully approximates the impact of γ_1 on survival estimates in the neighborhood of $\gamma_1 = 0$.

Figure 3 plots the ISNI values for the KM ordinates together with the KM curve and the estimated Peterson bounds. The ISNI 'curve' begins at 0 and is right continuous, as suggested in (7). The absolute value of ISNI has a maximum around $t = 20$ and then decreases to 0 for larger times, showing that the middle part of the survival curve is more sensitive to nonignorability. Prior to time 2 (the observed event
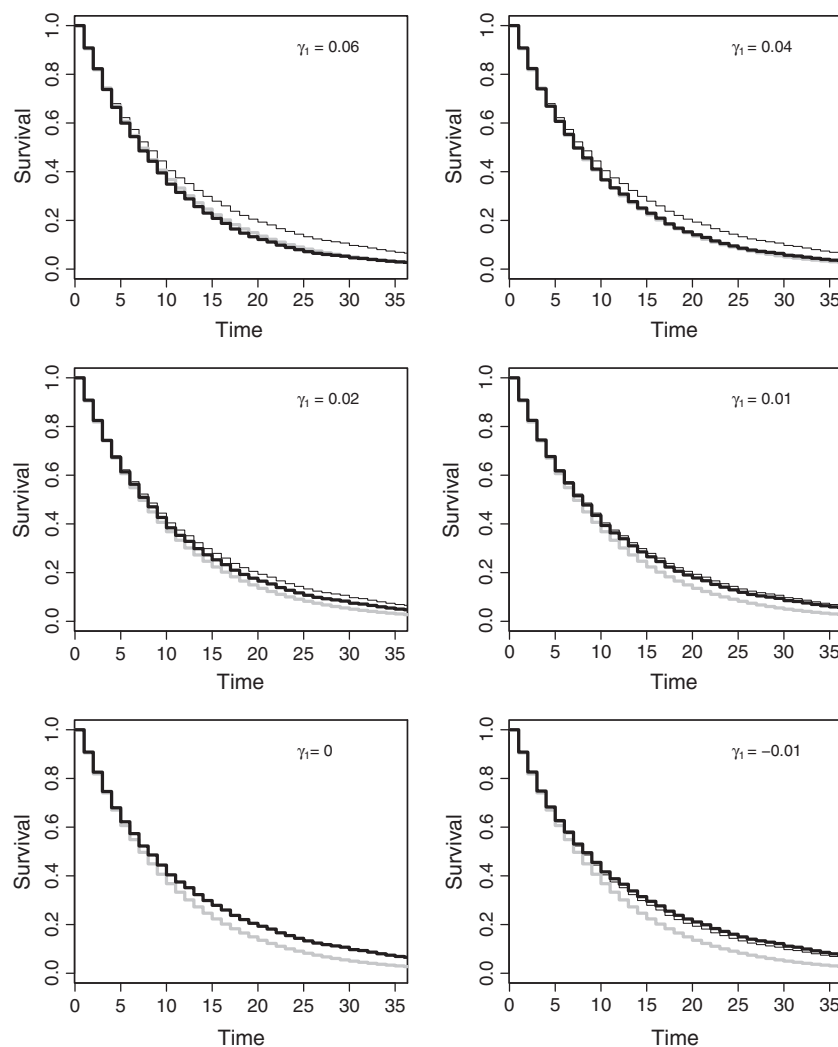

Figure 1. Estimated survival curves for the simulated data. The thin black lines are the KM curves, and the thick gray lines are the underlying true survival functions. These are the same for all six plots. The thick black lines are the estimated SC curves assuming different values of γ_1.
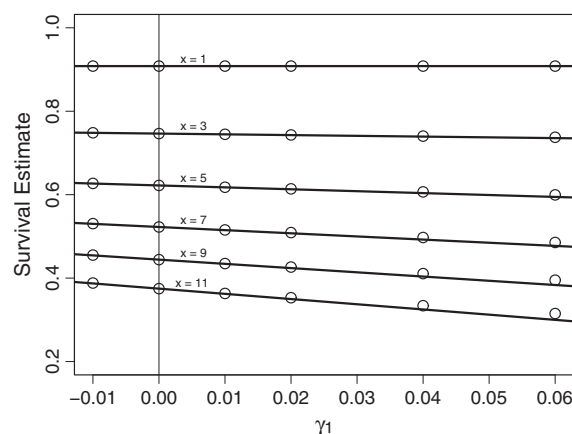
Figure 2. The index of local sensitivity to nonignorability (ISNI) for the simulated data. The circles connected by the solid lines are the self-consistent ordinates assuming $\gamma_1 = -0.01, 0, \ldots, 0.04$ at times $1, 3, \ldots, 11$. The circles at $\gamma_1 = 0.04$ are the true survival probabilities. The solid lines have intercepts equal to the Kaplan-Meier survival estimates and slopes equal to the corresponding ISNI.
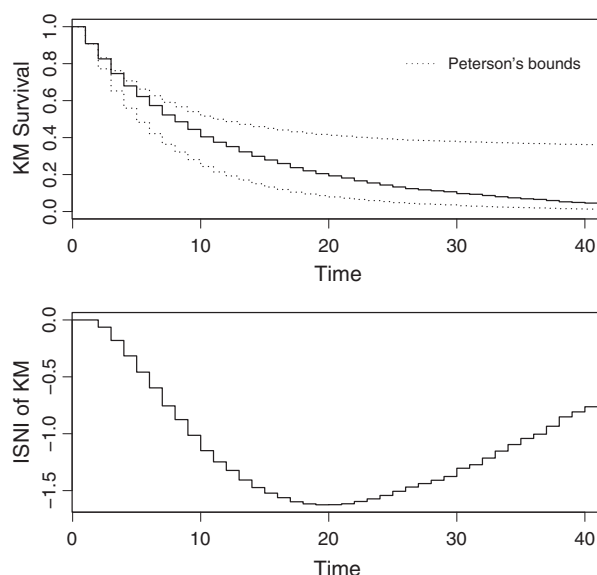

Figure 3. The index of local sensitivity to nonignorability (ISNI) of Kaplan-Meier (KM) across time for the simulated data: (top) the KM curve with the Peterson bounds and (bottom) ISNI over time.

time after the first censoring), overlapping of the upper and lower Peterson bounds suggests that KM is completely insensitive. This well agrees with ISNI, which is also 0 before time 2. As time increases, both the lower Peterson bound and the survival curve converge to 0, suggesting that later survival estimates are insensitive to nonignorability, if only in one direction. ISNI also converges to 0, suggesting a complete lack of sensitivity for large t.

5. Application to REMATCH

The REMATCH trial [17] evaluated the long-term survival benefit of a surgically implanted left ventricular assist device (LVAD) in heart failure patients who were ineligible for cardiac transplantation. The study enrolled 129 patients and randomized them to either LVAD ($n = 68$) or optimal medical management (OMM; $n = 61$). Five patients on OMM and 16 on LVAD were censored as of the closing date for our data set. To assess the impact of right censoring on survival estimates in REMATCH, we considered the following model for each arm, allowing the parameters γ_0 and γ_1 to differ between arms: $V|T \sim \text{EXP}(\exp(\gamma_0 + \gamma_1 T))$ and $C = \lceil V \rceil$.

Figure 4 shows the KM srvival estimates, 95% confidence limits, and KM ISNI curves for the OMM and LVAD arms of REMATCH. The KM curves suggest that LVAD has a substantially better survival rate. Only the later portions (after the smallest censoring times) of the KM curves are subject to the impact of possible nonignorable censoring. Local sensitivity for the KM estimate is much greater for the LVAD arm than for the OMM arm, because the proportion of censored events for the LVAD arm is higher.

Table I summarizes the mean survival (assuming that after the last observation, the survival curve declines to 0), the restricted mean survival for $\tau = 18$ months, the median survival, and four landmark survival probabilities, together with their ISNI assessment. For comparison of the two arms, we calculated ISNI using the method of Xie and Heitjan [16] for vector-valued γ_1, assuming that the two arms have distinct values of this parameter. By the definition of γ_1^*, $\log(12\gamma_1^*)$ is the hazard ratio of staying in the trial, for two people whose survival times differ by 1 year, which can cause a one-SE change
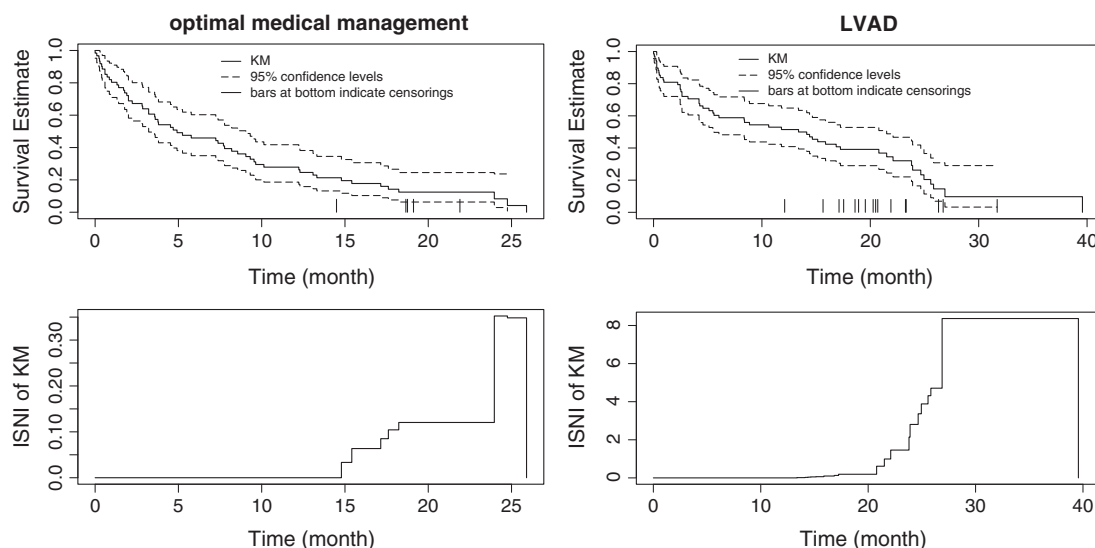

Figure 4. The Kaplan-Meier (KM) curve with 95% confidence levels and the index of local sensitivity to nonignorability (ISNI) for the REMATCH data. LVAD, left ventricular assist device.

Table I. Sensitivity analysis for the mean, restricted mean ($\tau = 18$ months), median, and survival probabilities at 6, 12, 18, and 24 months in the REMATCH data.

		Restricted mean			Landmark survival (months)			
	Mean	($\tau = 18$)	Median	6	12	18	24
OMM							
Estimate	8.13	7.12	4.80	0.46	0.28	0.14	0.08
ISNI	0.35	0.07	0	0	0	0.10	0.35
SE/ISNI (γ_1^*)	2.82	11.6	∞	∞	∞	0.44	0.13
LVAD							
Estimate	14.5	10.4	13.3	0.60	0.51	0.39	0.26
ISNI	14.7	0.12	12.3	0	0	0.19	2.81
SE/ISNI (γ_1^*)	0.11	7.41	0.28	∞	∞	0.31	0.02
Difference: (LVAD − OMM)							
Estimate	6.38	3.30	8.53	0.14	0.24	0.25	0.18
ISNI	20.8	0.19	17.4	0	0	0.31	4.00
SE/ISNI (γ_1^*)	0.09	6.07	0.21	∞	∞	0.24	0.02

We calculated the standard errors (SEs) for the mean, restricted mean, and median by using a bootstrap with 1000 resamples. OMM, optimal medical management; ISNI, index of local sensitivity to nonignorability; LVAD, left ventricular assist device.

in the estimate of the parameter of interest. In REMATCH, we think that this hazard ratio is plausibly no more than 2. Transformed back to the scale of γ_1, the critical figure is 0.06. By this criterion, the mean, restricted mean, median, and their comparisons between arms are insensitive to nonignorability. Compared with mean survival, the restricted mean and median survival are more locally robust. The survival estimate in the LVAD arm at 24 months and the difference between the two arms at 24 months are potentially sensitive to nonignorability, whereas the survival estimates at the other landmark times are robust. These results agree well with previous analyses by Zhang and Heitjan [15] who assumed parametric distributions for the survival times.

6. Discussion

Nonignorability of the random censoring mechanism has the potential to spoil estimates of the survival curve. In this paper, we describe a local sensitivity method to assess the impact of nonignorable censoring on the discrete-time KM survival estimate. Using the SC estimates for a simulated data set, we show that our method performs satisfactorily as a measure of local sensitivity. Previous authors have considered sensitivity analyses that assume parametric models for survival and censoring [11, 12, 15]. In this paper, we apply the local sensitivity method of Troxel *et al.* [14] to a model that entirely avoids parametric assumptions about the survival distribution.

Our method has several advantages over other available procedures. First, because sensitivity assessment via ISNI does not require estimation of the nonignorable censoring mechanism, it allows sensitivity analyses to be performed with little effort under a range of nonignorable models. As we showed in Section 3.3, one can apply a general censoring model such as a semiparametric proportional hazards model, reducing concerns about model misspecification. Moreover, we can readily extend analysis from the KM curve itself to derived quantities such as the mean, restricted mean, median, and other quantiles, not all of which may be equally sensitive to nonignorability.

We emphasize that the purpose of our method is to measure the sensitivity of the KM curve, in its guise as a nonparametric MLE, to potential nonignorability in the censoring mechanism. The underlying idea is that when faced with random coarsening, the preferred option is always to assume that the coarsening is ignorable. This is because the standard statistical approach of nesting the ignorable model within a nonignorable model, then testing for adequacy of fit, can be impracticable even when the coarsening mechanism is assumed 'known' up to parameters. Thus, sensitivity measures such as ISNI do not attempt to estimate the nonignorability but to evaluate whether modest departures from ignorability would substantially affect estimates. Moreover, ISNI is Bayesian in spirit in that it applies to the data at hand rather than to long-run averages including other data sets that might have been observed but were not. So, for example, when substantial censoring is present only in the far right tail of the distribution, ISNI will show that sensitivity is modest for KM ordinates at short survival times. In this sense, ISNI diagnostics are similar to regression diagnostics that identify parameters whose estimation may be sensitive to departures from the assumed model at one or more points in the observed data.

A potential limitation of the method is its continued reliance on model assumptions, however weak, regarding the censoring mechanism. Because we typically have little prior information about the censoring mechanism, it is advisable to conduct sensitivity analyses under a plausibly broad range of assumptions. Evidently, the only alternative to making some kind of assumption is to resort to Peterson's bounds, which are unacceptably wide in many applications.

The derivation of ISNI for KM takes advantage of the fact that right censoring falls under the rubric of the Heitjan–Rubin coarse-data model [1, 19]. In a strict sense, this model requires that the sample space be discrete, because the likelihood function is the integral of the underlying density over the set of possible values of the complete data that could have given rise to the observed, coarsened data. With a continuous model, a perfectly observed data point (in REMATCH, an observed death) therefore contributes a factor of 0 to the likelihood. It is possible to extend the coarse-data model to underlying continuous sample spaces [21], but the analysis is complex. In any event, the discrete model more accurately represents real-world data, which we can never measure with infinite precision. We argue that models assuming continuous random variables are merely approximations that one should eschew when they add no value to the analysis.

A next obvious step in the research is to develop measures for sensitivity of parameters in nonparametric or semiparametric survival regression models such as the Cox model.

Appendix A

Nonparametric survival maximum likelihood estimate with informative censoring

In our simulation study (Section 4), we proposed an SC procedure to estimate the survival curve for a fixed value of γ_1. To derive the SC procedure, we first write out the complete-data likelihood (1) using survival coordinates

$$L(\theta, \gamma) \propto \prod_{j=1}^{m} \left[\left\{ (S_{j-1} - S_j) \sum_{k \geq j} p_\gamma^{C|T}(t_{(k)}|t_{(j)}) \right\}^{d_j} \left\{ \sum_{k > j} (S_{k-1} - S_k) p_\gamma^{C|T}(t_j|t_k) \right\}^{r_j - d_j} \right].$$

Solving the score equations for a fixed γ_1 and requiring that $1 = S_0 \geq S_1 \geq \cdots \geq S_m$ lead to the following equations:

$$\tilde{S}_j = \frac{\sum_{k>j} d'_k}{\sum_{k \geq j} d'_k} \tilde{S}_{j-1}, \text{ for } j = 1, 2, \ldots, m, \tag{13}$$

where $d'_j = d_j + \mathbf{1}(j > 1) \sum_{k<j} d^*_{jk}$ and $d^*_{jk} = \mathbf{1}(j > k) \frac{(r_k - d_k)(\tilde{S}_{j-1} - \tilde{S}_j) p_\gamma^{C|T}(t_{(k)}|t_{(j)})}{\sum_{l>k} (\tilde{S}_{l-1} - \tilde{S}_l) p_\gamma^{C|T}(t_{(k)}|t_{(l)})} + \mathbf{1}(j \leq k) \frac{d_j p_\gamma^{C|T}(t_{(k)}|t_{(j)})}{\sum_{l \geq j} p_\gamma^{C|T}(t_{(l)}|t_{(j)})}$, and the nuisance parameter $\tilde{\gamma}_0$ must satisfy

$$\frac{\partial \left\{ \sum_j \sum_k d^*_{jk} \log p_\gamma^{C|T}(t_{(k)}|t_{(j)}) \right\}}{\partial \gamma_0} \bigg|_{\gamma_0 = \tilde{\gamma}_0} = 0. \tag{14}$$

Simultaneously solving (13) and (14) for $\{\tilde{S}_j\}$ and $\tilde{\gamma}_0$ is not straightforward, but one can readily solve the equations with a self-consistency algorithm [22], which in this case is equivalent to an EM algorithm [23]. The idea is based on Turnbull [24, 25] and is similar to that of Link [26].

Having chosen a starting value (KM for $\{S_j\}$ and any reasonable guess for γ_0 will serve), one solves the equations by using the following steps: (1) compute $\{d^*_{jk}\}$ and $\{d'_j\}$; (2) with $\tilde{S}_0 = 0$, compute $\{\tilde{S}_j\}$ by using (13); (3) compute $\{d^*_{jk}\}$ and $\{d'_j\}$ with the updated $\{\tilde{S}_j\}$ from step 2; and (4) solve for $\tilde{\gamma}_0$ using (14). Cycle through steps 1–4 until convergence.

If censoring is ignorable (i.e., if $\gamma_1 = 0$), then the SC curve equals the KM curve. Although the SC curve can serve as the basis of a sensitivity analysis, the computation is laborious and may be impractical in many situations.

The index of local sensitivity to nonignorability for an estimated quantile when T is continuous

Consider the case of continuous T with estimated survival function $\tilde{S}(t)$ that is differentiable at the wth estimated quantile $\tilde{T}_w$. Taking derivatives on both sides of the equation, $\tilde{S}(\tilde{T}_w) = 1 - w$, with respect to γ_1, we have

$$\frac{\partial \tilde{S}(t)}{\partial \gamma_1} \bigg|_{t = \tilde{T}_w} - \tilde{f}(\tilde{T}_w) \frac{\partial \tilde{T}_w}{\partial \gamma_1} = 0,$$

where $\tilde{f}$ is the estimated density function of T. Rearranging this, we obtain the ISNI for the wth estimated quantile as

$$\text{ISNI}(\hat{T}_w) = \frac{\partial \tilde{T}_w}{\partial \gamma_1} \bigg|_{\gamma_1 = 0, \gamma_0 = \hat{\gamma}_0} = \frac{\text{ISNI}(\hat{S}(t))}{\hat{f}(t)} \bigg|_{t = \hat{T}_w}. \tag{15}$$

Acknowledgements

References

1. Heitjan DF, Rubin DB. Ignorability and coarse data. *Annals of Statistics* 1991; **19**:2244–2253.
2. Tsiatis AA. A nonidentifiability aspect of the problem of competing risks. *Proceedings of the National Academy of Sciences USA* 1975; **72**:20–22.
3. Tsiatis AA. An example of nonidentifiability in competing risks. *Scandinavian Actuarial Journal* 1978; **1978**(4):235–239.
4. Peterson AV, Jr. Bounds for a joint distribution function with fixed sub-distribution functions: application to competing risks. *Proceedings of the National Academy of Sciences USA* 1976; **73**:11–13.
5. Heitjan DF. Ignorability and coarse data: some biomedical examples. *Biometrics* 1993; **49**:1099–1109.
6. Scharfstein DO, Robins JM, Eddings W, Rotnitzky A. Inference in randomized studies with informative censoring and discrete time-to-event endpoints. *Biometrics* 2001; **57**:404–413.
7. Scharfstein DO, Robins JM. Estimation of the failure time distribution in the presence of informative right censoring. *Biometrika* 2002; **89**:617–635.
8. Rotnitzky A, Farall A, Bergesio A, Scharfstein DO. Analysis of failure time data under competing censoring mechanisms. *Journal of the Royal Statistical Society, Series B* 2007; **69**:307–327.
9. Shardell M, Scharfstein DO, Vlahov D, Galai N. Inference for cumulative incidence functions with informatively coarsened discrete event-time data. *Statistics in Medicine* 2008; **27**:5861–5879.
10. Shardell M, Scharfstein DO, Vlahov D, Galai N. Sensitivity analysis using elicited expert information for inference with coarsened data: illustration of censored discrete event times in ALIVE. *American Journal of Epidemiology* 2008; **168**:1460–1469.
11. Siannis F. Applications of a parametric model for informative censoring. *Biometrics* 2004; **60**:704–714.
12. Siannis F, Copas J, Lu G. Sensitivity analysis for informative censoring in parametric survival models. *Biostatistics* 2005; **6**:77–91.
13. Zhang J, Heitjan DF. A simple local sensitivity analysis tool for nonignorable coarsening: application to dependent censoring. *Biometrics* 2006; **62**:1260–1268.
14. Troxel AB, Ma G, Heitjan DF. An index of local sensitivity to nonignorability. *Statistica Sinica* 2004; **14**:1221–1237.
15. Zhang J, Heitjan DF. Nonignorable censoring in randomized clinical trials. *Clinical Trials* 2005; **2**:1–9.
16. Xie H, Heitjan DF. Sensitivity analysis of causal inference in a clinical trial subject to crossover. *Clinical Trials* 2004; **1**:21–30.
17. Rose EA, Gelijns AC, Moskowitz AJ, Heitjan DF, Stevenson LW, Dembitsky W, Long JW, Ascheim DD, Tierney AR, Levitan RG, Watson JT, Meier P, The REMATCH Study Group. Long-term use of a left ventricular assist device for end-stage heart failure. *New England Journal of Medicine* 2001; **345**:1435–1443.
18. Rubin DB. Inference and missing data. *Biometrika* 1976; **63**:581–592.
19. Heitjan DF. Ignorability, sufficiency and ancillarity. *Journal of the Royal Statistical Society, Series B* 1997; **59**:375–381.
20. Karrison T. Restricted mean life with adjustment for covariates. *Journal of the American Statistical Association* 1987; **82**:1169–1176.
21. Jacobsen M, Keiding N. Coarsening at random in general sample spaces and random censoring in continuous time. *The Annals of Statistics* 1995; **23**:774–786.
22. Efron B. The two-sample problem with censored data. In *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 4. University of California Press: Berkeley, 1967; 831–853.
23. Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B* 1977; **39**:1–22.
24. Turnbull BW. Nonparametric estimation of a survivorship function with doubly censored data. *Journal of the American Statistical Association* 1974; **69**:169–173.
25. Turnbull BW. The empirical distribution function with arbitrarily grouped, censored and truncated data. *Journal of the Royal Statistical Society, Series B* 1976; **38**:290–295.
26. Link WA. An identifiable model for informative censoring. *Computer Science and Statistics: Proceedings of the 20th Symposium on the Interface*, Fairfax, Virginia, April 1988; 725–728.