Published in final edited form as:

Biometrics. 2013 December ; 69(4): . doi:10.1111/biom.12076.

Doubly-robust Estimators of Treatment-specific Survival Distributions in Observational Studies with Stratified Sampling

Xiaofei Bai^{1,*}, Anastasios A. Tsiatis^{1,**}, and Sean M. O'Brien^{2,***}

¹Department of Statistics, North Carolina State University, Raleigh, North Carolina 27695, U.S.A.

²Department of Biostatistics and Bioinformatics, Duke University Medical Center, Durham, North Carolina 27710, U.S.A.

Abstract

Summary—Observational studies are frequently conducted to compare the effects of two treatments on survival. For such studies we must be concerned about confounding; that is, there are covariates that affect both the treatment assignment and the survival distribution. With confounding the usual treatment-specific Kaplan-Meier estimator might be a biased estimator of the underlying treatment-specific survival distribution. This paper has two aims. In the first aim we use semiparametric theory to derive a doubly robust estimator of the treatment-specific survival distribution in cases where it is believed that all the potential confounders are captured. In cases where not all potential confounders have been captured one may conduct a substudy using a stratified sampling scheme to capture additional covariates that may account for confounding. The second aim is to derive a doubly-robust estimator for the treatment-specific survival distributions and its variance estimator with such a stratified sampling scheme. Simulation studies are conducted to show consistency and double robustness. These estimators are then applied to the data from the ASCERT study that motivated this research.

Keywords

Cox proportional hazard model; Double robustness; Observational study; Stratified sampling; Survival analysis

1. Introduction

This research is motivated by an observational study funded by the National Heart Lung and Blood Institute to compare two widely used treatment options for patients with coronary artery disease (Weintraub et al., 2012). The study, known as ASCERT, was a retrospective analysis of records from 189,793 patients at 644 hospitals who had either two-vessel or three-vessel coronary artery disease and were treated either by surgical revascularization (coronary artery bypass grafting; CABG) or catheter-based revascularization (percutaneous coronary intervention; PCI). Not all patients were followed until the endpoint of interest (all-cause mortality); accordingly, survival outcomes for these patients were censored.

As is well known, in such an observational study, treatments (PCI and CABG here) are determined according to clinician/patient preference and thus may be subject to confounding; that is, patients receiving the two treatments may be prognostically different.

*xbai3@ncsu.edu. **tsiatis@ncsu.edu. ***sean.m.obrien@duke.edu.

Supplementary Materials

The Web Appendix referenced in Section 2, 3, 4 and 5 is available with this paper at the Biometrics website on Wiley Online Library.

Under these conditions, the Kaplan-Meier estimator (Kaplan and Meier, 1958) for a treatment-specific survival distribution will in general be an inconsistent estimator for the true treatment-specific survival distribution, formally defined with the use of potential outcomes in Section 2. Estimators that take account of confounding have been proposed by, e.g., Makuch (1982) and Lee et al. (1992) using a form of regression-based adjustment and Cole and Hernán (2004) using propensity score adjustment. The former methods require that the analyst posit a model for the survival distribution as a function of potential confounding variables while the latter depends on a model for the propensity score, the probability of receiving a treatment given covariates (confounders). Although these estimators lead to valid inference if these models are correctly specified under the assumption that all potential confounding variables are recorded in the database (no unmeasured confounders), if the models do not capture the true relationships between outcome or treatment assignment and covariates, respectively, the estimators may be inconsistent.

To address this limitation and to achieve improved efficiency, we propose to use the theory of Robins and Rotnitzky (1992) and Hubbard et al. (2000) for deriving locally efficient augmented inverse probability weighted complete case (AIPWCC) estimators to derive such estimators for treatment-specific distributions in the observational data context. Under the no unmeasured confounders assumption, such estimators will not only yield greater efficiency, we will show they are doubly robust in that they will be consistent for the true survival distribution if either of the required postulated model for the survival distribution as a function of covariates or those for the propensity score and censoring distribution are correctly specified, even if the other is not.

Moreover, although the ASCERT database is a rich resource, the investigators recognized that assessment of coronary anatomy was limited to a few relatively crude variables and hence acknowledged the possibility of residual confounding. For example, patients with certain high-risk coronary features may have been preferentially selected for CABG over PCI. To address this limitation, a sub-study was conducted in which detailed data on coronary anatomy were collected on a stratified random sample of patients at 54 hospitals in the main ASCERT study. According to the protocol, approximately 500 patient records were to be randomly selected without replacement from each of four strata formed by all combinations of the two treatments and whether or not the patients had two- or three-vessel disease. For each randomly selected patient, X-ray films of coronary angiograms taken prior to the patient's revascularization procedure were retrieved from storage and sent to a central laboratory for expert interpretation and analysis. More details of this study are given in Section 5.

The challenge is that of acknowledging this design in the comparison of treatment-specific survival distributions. There is a vast literature on methods for incorporating data from a subsample into such analyses (e.g., Breslow and Cain, 1988; Flanders and Greenland, 1991; Zhao and Lipsitz, 1992; Mark and Katki, 2006; Stürmer et al., 2007, and references therein). However, to our knowledge, there is no work applicable to a setting like that of ASCERT to develop estimators for survival distributions that take full advantage of the data collected on the entire cohort to achieve the greatest efficiency possible and that possess the double robustness property to protect against model misspecification. Currently, theory for AIPWCC estimators is based on the assumption that observations across subjects are independent and identically distributed (iid), which is clearly not the case for data collected according to a stratified sampling scheme.

In the next section, we derive doubly robust estimators (and their asymptotic variance) of the treatment-specific survival distributions and the difference in the treatment-specific survival distributions in the case when the collected covariates are enough to adjust for all

confounding in the main study. In Section 3, we develop methods to deal with the case when we further conduct a stratified sampling substudy to collect additional potential confounding covariates. The performance of the proposed estimators is demonstrated by simulation studies in Section 4. We then apply our method to the motivating example of ASCERT in Section 5. The paper ends with discussion and conclusion in Section 6.

2. Estimators for Main Study when all Potential Confounders are Captured

As in the motivating example of the ASCERT study, we are interested in estimating treatment-specific survival distributions and their difference in an observational study. We assume here that there are two treatments under consideration which we label 1 and 0. We define potential outcomes for the N individuals in our study as $(X_i, T_i^{(0)}, T_i^{(1)}, C_i^{(0)}, C_i^{(1)})$, $i=1, \dots, N$, assumed identically and independently distributed, where for individual i in our sample X_i denotes a vector of baseline covariates, $T_i^{(j)}$ denotes the potential survival time and $C_i^{(j)}$ denotes the potential censoring time if individual i were given treatment (possibly contrary to fact) j , for $j = 0, 1$. The treatment-specific survival distribution is then defined as $S^{(j)}(u) = P(T^{(j)} \geq u)$ for $j = 0, 1$ and will be the main focus of inference in this paper.

In the ASCERT trial all patients were followed from the date they entered the study until the time of data analysis and their survival status was fully ascertained during this period. Consequently, the potential censoring time was the time from entry into study until time of analysis which would be the same under both treatments. Therefore in the ASCERT trial $C_i^{(1)} = C_i^{(0)} = C_i$ for $i=1, \dots, N$.

We also denote by Z_i the binary treatment assignment for patient i , where $Z_i = 1, 0$ and make the strong ignorability assumption (Rubin, 1978), or no unmeasured confounders assumption, that Z is conditionally independent of $T^{(j)}$ given X , denoted by $Z \perp\!\!\!\perp T^{(j)} | X$. In an observational study, if we collected sufficient covariates X that are related to survival and that we believe were used in the treatment decision, then an assumption of no unmeasured confounders may be reasonable.

We also make the usual assumption of non-informative censoring; namely, that $C \perp\!\!\!\perp T^{(j)} | X, Z$. Together these assumptions imply that

$$(Z, C) \perp\!\!\!\perp T^{(j)} | X. \quad (1)$$

We use this assumption in the remainder of this paper and refer to it as assumption (1).

In contrast to the potential outcomes, some of which may not be observed, the data that are observable can be summarized as $(Z_i, X_i, U_i, \Delta_i)$, $i = 1, \dots, N$, where, in addition to (Z_i, X_i) , already defined, we also observe for individual i their time to death or censoring $U_i = \min(T_i, C_i)$ and the failure indicator $\Delta_i = I(T_i \leq C_i)$ where $T_i = Z_i T_i^{(1)} + (1 - Z_i) T_i^{(0)}$; i.e., the time to death on the assigned treatment.

2.1 Inference for Treatment-specific Survival Distributions

In this section, we mainly focus on estimating the underlying survival distribution on treatment 1 while obvious analogs could be used for treatment 0. Thus, the primary goal is to estimate the function $S^{(1)}(u) = P(T^{(1)} \geq u)$, $u \geq 0$ using the observed data $(Z_i, X_i, U_i,$

Δ_i , $i = 1, \dots, N$. Of course, we only get to observe $T^{(1)}$ in the data when $Z = 1$ and $\Delta = 1$ ($C > T$); otherwise $T^{(1)}$ is missing or coarsened. As indicated by Tsiatis (2006), section 8.6, this can be viewed as monotone coarsening, where we only get to observe the data on treatment 1 when $Z = 1$ and only data that were observed through censoring time C and the variable of interest $T^{(1)}$ is observed (complete-case) when $Z=1, C \geq T^{(1)}$. Because $(Z, C) \perp\!\!\!\perp T^{(1)} | X$, this implies that we have monotone coarsening at random and thus the probability of not being coarsened; i.e., observing a complete case, is given by $\pi(X) K_c^{(1)}(T^{(1)}, X)$, where $\pi(X) = P(Z = 1|X)$ is the so called propensity score, and $K_c^{(1)}(r, X) = P(C \geq r | X, Z=1)$ is the conditional survival function of the treatment specific censoring time given X .

If the model for the coarsening, defined through the propensity score and the treatment specific censoring distribution given the covariates are correctly specified, then without making any additional assumption on the survival distribution, as shown by Robins and Rotnitzky (1992) and Tsiatis (2006), that all estimators of the survival distribution can be written as AIPWCC estimators with corresponding estimating functions:

$$\frac{Z\Delta}{\pi(X) K_c^{(1)}(T^{(1)}, X)} \left\{ I(T^{(1)} \geq u) - S^{(1)}(u) \right\} - \left\{ \frac{Z - \pi(X)}{\pi(X)} \right\} h_1(u, X) + \int_0^\infty \frac{Z}{\pi(X)} \frac{dM_c^{(1)}(r, X)}{K_c^{(1)}(r, X)} h_2(u, r, X), \quad (2)$$

where $dM_c^{(1)}(r, X)$ is the martingale increment for the censoring distribution; namely,

$$dM_c^{(1)}(r, X) = \frac{dN_c(r) - \lambda_c^{(1)}(r, X) Y(r)}{K_c^{(1)}(r, X)}, \quad N_c(r) = I(U \leq r, \Delta=0), \quad Y(r) = I(U \geq r) \text{ and}$$

$\lambda^{(1)}(r, X) = \frac{-d \log K_c^{(1)}(r, X)}{dr}$ is the conditional hazard function for C given $Z = 1$ and X , $h_1(u, X)$ and $h_2(u, r, X)$ are arbitrary functions. We note that $\left\{ I(T^{(1)} \geq u) - S^{(1)}(u) \right\}$ denotes the full-data estimating function that would be used if $T^{(1)}$ were observed on all subjects. The first term in (2) is the inverse probability weighted complete case term, and the latter two terms are augmentation terms.

For monotonely coarsened data, we can use the results from Theorem 10.4 of Tsiatis (2006) to show that the optimal choice for $h_1(u, X)$ and $h_2(u, r, X)$, i.e., the estimator with the smallest asymptotic variance is given by $h_1(u, X) = E \left\{ I(T^{(1)} \geq u) - S^{(1)}(u) | X \right\}$ and $h_2(u, r, X) = E \left\{ I(T^{(1)} \geq u) - S^{(1)}(u) | T^{(1)} \geq r, X \right\}$.

Let us define $P(T^{(1)} \geq r | X) = H^{(1)}(r, X)$. We also note that $P(T \geq r | Z=1, X) = P(T^{(1)} \geq r | Z=1, X) = P(T^{(1)} \geq r | X)$, the last equality following from assumption (1). Consequently, the optimal estimating function is given as

$$\begin{aligned} & \frac{Z\Delta}{\pi(X) K_c^{(1)}(U, X)} \left\{ I(U \geq u) - S^{(1)}(u) \right\} \\ & - \left\{ \frac{Z - \pi(X)}{\pi(X)} \right\} \left\{ H^{(1)}(u, X) - S^{(1)}(u) \right\} \\ & + \int_0^\infty \frac{Z}{\pi(X)} \frac{dM_c^{(1)}(r, X)}{K_c^{(1)}(r, X)} \left[I(r < u) \left\{ \frac{H^{(1)}(u, X)}{H^{(1)}(r, X)} - S^{(1)}(u) \right\} + I(r \geq u) \left\{ 1 - S^{(1)}(u) \right\} \right]. \end{aligned}$$

As shown in the supplementary material (Web Appendix A), the above estimating function can be simplified as

$$\frac{ZI(U \geq u)}{\pi(X) K_c^{(1)}(u, X)} - \left\{ \frac{Z - \pi(X)}{\pi(X)} \right\} H^{(1)}(u, X) + \int_0^u \frac{Z}{\pi(X)} \frac{dM_c^{(1)}(r, X)}{K_c^{(1)}(r, X)} \left\{ \frac{H^{(1)}(u, X)}{H^{(1)}(r, X)} \right\} - S^{(1)}(u). \quad (3)$$

If we knew the propensity score $\pi(X)$, the conditional censoring distribution $K_c^{(1)}(r, X)$, and the conditional survival distribution $H^{(1)}(r, X)$, then this estimating function could be used as a basis to obtain the optimal AIPWCC estimator for $S^{(1)}(u)$; namely,

$$\hat{S}^{(1)}(u) = N^{-1} \sum_{i=1}^N \left[\frac{Z_i I(U_i \geq u)}{\pi(X_i) K_c^{(1)}(u, X_i)} - \left\{ \frac{Z_i - \pi(X_i)}{\pi(X_i)} \right\} H^{(1)}(u, X_i) + \int_0^u \frac{Z_i}{\pi(X_i)} \frac{dM_{c,i}^{(1)}(r, X_i)}{K_c^{(1)}(r, X_i)} \left\{ \frac{H^{(1)}(u, X_i)}{H^{(1)}(r, X_i)} \right\} - \hat{S}^{(1)}(u) \right], \quad (4)$$

and the variance of the estimator for $\hat{S}^{(1)}(u)$ can be estimated by using the standard sandwich variance estimator; namely,

$$N^{-1} \sum_{i=1}^N \left[\frac{Z_i I(U_i \geq u)}{\pi(X_i) K_c^{(1)}(u, X_i)} - \left\{ \frac{Z_i - \pi(X_i)}{\pi(X_i)} \right\} H^{(1)}(u, X_i) + \int_0^u \frac{Z_i}{\pi(X_i)} \frac{dM_{c,i}^{(1)}(r, X_i)}{K_c^{(1)}(r, X_i)} \left\{ \frac{H^{(1)}(u, X_i)}{H^{(1)}(r, X_i)} \right\} - \hat{S}^{(1)}(u) \right]^2. \quad (5)$$

The estimator given above is similar to the estimator proposed by Hubbard et al. (2000), using the same principle of developing the locally efficient estimator, however, we demonstrate in Web Appendix B that this estimator is doubly-robust; that is, if either the estimator for the coarsening defined through the estimator for $P(Z = 1|X) = \pi(X)$ and the estimator for $P(C \geq r|Z=1, X) = K_c^{(1)}(r, X)$ are both consistent or the estimator for $P(T^{(1)} \geq r|X) = H^{(1)}(r, X)$ is consistent, the estimator (4) will be consistent. Moreover, we use this estimator as a springboard for the AIPWCC estimator in the case of stratified sampling in Section 3.

Of course, we do not know $\pi(X)$, $K_c^{(1)}(r, X)$, or $H^{(1)}(r, X)$; hence, each of these must be estimated from the data. A logistic regression model is commonly used to estimate the propensity score $P(Z = 1|X)$ with data $\{(Z_i, X_i), i = 1, \dots, n\}$. For $K_c^{(1)}(r, X)$ a treatment-specific proportional hazards model $\lambda_c^{(1)}(r, X) = \lambda_c^{(1)}(r) \exp\{\beta_c^T X\}$ is often used with data $\{(U_i, (1 - \Delta_i), X_i), i : Z_i = 1\}$. For $H^{(1)}(r, X)$ a treatment-specific proportional hazards model $\lambda_t^{(1)}(r, X) = \lambda_t^{(1)}(r) \exp\{\beta_t^T X\}$ with data $\{(U_i, \Delta_i, X_i), i : Z_i = 1\}$ is often used.

We must however be careful with regards to the variance estimator (5) because there may be an effect on the variance due to estimating the functions $\pi(X)$, $K_c^{(1)}(r, X)$ and $H^{(1)}(r, X)$ and their possible misspecification. What we know from the theory of AIPWCC estimators is that if all these functions were consistently estimated, then the asymptotic variance of the corresponding estimator would be unaffected in which case we could estimate the variance of $\hat{S}^{(1)}(u)$ consistently by substituting the estimators $\hat{\pi}(X)$, $\hat{K}_c^{(1)}(r, X)$ and $\hat{H}^{(1)}(r, X)$ into (5). If, however, the propensity score $\pi(X)$ and the conditional censoring distribution $K_c^{(1)}(r, X)$ were consistently estimated whereas, the conditional survival distribution $H^{(1)}(r, X)$ was not, then the variance estimator would be conservative. In practice it has been shown that a reasonable attempt for estimating $H^{(1)}(r, X)$ will often result in negligible conservativeness even with misspecification for this model. This will be investigated further in our simulation studies in Section 4. When the estimator for $H^{(1)}(r, X)$ is consistent and

either of the estimators for $\pi(X)$ or $K_c^{(1)}(r, X)$ are not, then even though the resulting estimator for $S^{(1)}(u)$ is consistent (part of the double robustness property), the resulting estimator for the variance may be biased and there is no theory that we know of that suggests the direction or extent of the bias. In such cases, we may want to use a bootstrap estimator of the asymptotic variance. This too will be investigated in the simulation studies of Section 4.

2.2 Inference on Difference of Treatments

As well as estimating the individual treatment specific survival distributions, we are also interested in testing the null hypothesis on whether the survival distributions of the two treatments are the same. That is, estimating the difference in the treatment-specific survival distribution $S_{diff}(u) = S^{(1)}(u) - S^{(0)}(u) = P(T^{(1)} \geq u) - P(T^{(0)} \geq u)$, $u \geq 0$. In Section 2.1, we developed the estimating function for $S^{(1)}(u)$ and by analogy we can easily obtain the estimating function for $S^{(0)}(u)$. It is then straightforward that the estimating function for $S_{diff}(u)$ is just the difference between the estimating function of $S^{(1)}(u)$ and of $S^{(0)}(u)$. For ease of notation we denote for patient i and treatment $j = 0, 1$

$$\phi_j^*(V_i) = \left[\frac{(Z_i)^j (1 - Z_i)^{1-j} I(U_i \geq u)}{\{\pi(X_i)\}^j \{1 - \pi(X_i)\}^{1-j} K_c^{(j)}(u, X_i)} - \frac{(2j-1) \{Z_i - \pi(X_i)\}}{\{\pi(X_i)\}^j \{1 - \pi(X_i)\}^{1-j}} H^{(j)}(u, X_i) + \int_0^u \frac{(Z_i)^j (1 - Z_i)^{1-j}}{\{\pi(X_i)\}^j \{1 - \pi(X_i)\}^{1-j}} \frac{dM_c^{(j)}(r)}{K_c^{(j)}(r)} \right]$$

where $V_i = (Z_i, X_i, U_i, \Delta_i)$ are the data observed for the i -th individual, $K_c^{(1)}(r, X) = P(C \geq r | X, Z=j)$, and $H^{(j)}(r, X) = P(T \geq r | X, Z=j)$. We also denote by $\hat{\phi}_j^*(V_i)$ the same as in (6) after substituting the estimators for $\pi(X_i)$, $K_c^{(j)}(r, X_i)$, and $H^{(j)}(r, X_i)$. Hence, the doubly robust estimator for $S_{diff}(u)$ is given by

$\hat{S}_{diff}(u) = N^{-1} \sum_{i=1}^N \{\hat{\phi}_1^*(V_i) - \hat{\phi}_0^*(V_i)\}$, and for the variance estimator of $\hat{S}_{diff}(u)$ we propose using the sandwich variance

$$\widehat{\text{Var}}\{\hat{S}_{diff}(u)\} = N^{-1} \sum_{i=1}^N \{\hat{\phi}_1^*(V_i) - \hat{\phi}_0^*(V_i) - \hat{S}_{diff}(u)\}^2.$$

The same comments regarding the appropriateness of this estimator when some of the models are misspecified apply as in the estimation of $S^{(1)}(u)$ discussed previously.

3. Stratified Sampling with Additional Information from a Subsample

Let $V_i = (U_i, \Delta_i, Z_i, X_{1i})$, $i = 1, \dots, N$ denote the data available on all subjects in the main study, where now X_1 represents covariates collected on these subjects. The key requirement for using the previously proposed method is the strong ignorability assumption, which here would take the form $Z \perp\!\!\!\perp T^{(j)} | X_1$. If, however, there is concern that this assumption is untenable based only on X_1 , then a subset of patients may be sampled from the main study to collect additional covariates X_2 . As in the ASCERT study the subsample may be obtained using a stratified sampling design in which K strata are identified based on (Z, X_1) , and

letting ε denote the stratum indicator taking values $1, \dots, K$ such that $N_k = \sum_{i=1}^N I(\varepsilon_i = k)$ is the number of subjects in stratum k , $k = 1, \dots, K$, then from stratum k , a fixed (by design) number of subjects n_k are sampled from the N_k subjects at random without replacement. For each subject so included in the substudy, additional covariates X_2 are collected, where,

letting $X = (X_1^T, X_2^T)^T$, it is believed that the strong ignorability assumption $Z \perp\!\!\!\perp T^{(j)} | X$ holds. Let $R_i = 1$ if subject $i = 1, \dots, N$, was selected into the subsample and 0 otherwise.

The goal is to estimate $S^{(1)}(u)$ (and $S_{diff}(u)$) using the observed data $O_i = (U_i, \Delta_i, Z_i, X_{1i}, R_i, R_i X_{2i})$.

3.1 Inference for Treatment-specific Survival Distributions

Again, here we use the missing data analogy. We identify the “full data” that would ideally have been collected on all N subjects as $F_i = (U_i, \Delta_i, Z_i, X_{1i}, X_{2i})$, $i = 1, \dots, N$, recognizing that X_2 is collected only on subjects in the substudy. Were there full data available, a consistent, asymptotically normal, doubly robust estimator for $S^{(1)}(u)$ could be found by

solving $\sum_{i=1}^N \phi \{F_i, S^{(1)}(u)\} = 0$, where $\phi\{F, S^{(1)}(u)\}$ is the estimating function (3). We may identify the observed data as full data with X_2 missing for subjects not in the subsample. Accordingly, following Tsiatis (2006), we propose the class of AIPWCC estimators defined as the solution to

$$\sum_{i=1}^N \left[\frac{R_i \phi \{F_i, S^{(1)}(u)\}}{\eta(\epsilon_i)} - \left\{ \frac{R_i - \eta(\epsilon_i)}{\eta(\epsilon_i)} \right\} h(V_i) \right] = 0, \quad (7)$$

where $\eta(\epsilon_i) = P(R_i = 1/\epsilon_i)$, taking values n_k/N_k when $\epsilon_i = k$, the probability of selection into the subsample as a function of the stratum to which subject i belongs; and $h(V_i)$ is an arbitrary function of the data collected on all subjects. Here, if we set $h(V_i) = 0$, then the resulting estimator is a simple inverse probability weighted estimator, which we refer to as the weighted estimator. However, the choice of $h(V)$ will affect the asymptotic variance of the resulting estimator.

The usual theory for AIPWCC estimators applies when the observed data O_i , $i = 1, \dots, N$, are iid. We briefly outline our proposed modification to account for the stratified sampling. Rearranging (7), we write this estimating equation as $T\{\mathbf{O}, S^{(1)}(u)\} = 0$, where $\mathbf{O} = (O_1, \dots, O_n)$ (all the observed data) and

$$T\{\mathbf{O}, S^{(1)}(u)\} = \sum_{k=1}^K \frac{N_k}{N} \left(\frac{\sum_{i:\epsilon_i=k} R_i [\phi \{F_i, S^{(1)}(u)\} - h(V_i)]}{n_k} + \frac{\sum_{i:\epsilon_i=k} h(V_i)}{N_k} \right). \quad (8)$$

Letting $\mu_k(\phi) = E[\phi\{F, S^{(1)}(u)\} | \epsilon = k]$ and $\mu_k(h) = E[h(V) | \epsilon = k]$, it is straightforward to

$$\begin{aligned} E[T\{\mathbf{O}, S^{(1)}(u)\}] &= E\left(E[T\{\mathbf{O}, S^{(1)}(u)\} | N_1, \dots, N_K]\right) \\ &= E\left[\sum_{k=1}^K (N_k/N) \{\mu_k(\phi) - \mu_k(h) + \mu_k(h)\}\right] \\ &= \sum_{k=1}^K p_k \mu_k(\phi) \end{aligned}$$

observe that $= E[\phi\{F, S^{(1)}(u)\}]$, where $p_k = P(\epsilon = k)$. This implies that $T\{\mathbf{O}, S^{(1)}(u)\}$ is an unbiased estimating function. Consequently, the solution to (7) will yield an estimator for $S^{(1)}(u)$ which will converge in limit to the same estimand as

the solution to full data estimating equation given by (3) with $X = (X_1^T, X_2^T)^T$. In particular, this means that the proposed AIPWCC estimator will be doubly robust as described in the previous section. This result is true regardless of the choice for $h(V)$.

To derive the optimal choice of $h(V)$ and to support large sample inference, we must derive the asymptotic variance of estimators solving (7) and an estimator for the asymptotic

variance of $T\{\mathbf{O}, S^{(1)}(u)\}$. Note that $\text{Var}[T\{\mathbf{O}, S^{(1)}(u)\}] = E(\text{Var}[T\{\mathbf{O}, S^{(1)}(u)\} | N_1, \dots, N_K]) + \text{Var}(E[T\{\mathbf{O}, S^{(1)}(u)\} | N_1, \dots, N_K])$. We showed above that

$$E\left[T\left\{\mathbf{O}, S^{(1)}(u)\right\} | N_1, \dots, N_K\right] = \sum_{k=1}^K (N_k/N) \mu_k(\phi), \text{ which, because } (N_1, \dots, N_K) \text{ is multinomial } (N, p_1, \dots, p_K), \text{ implies}$$

$$\text{Var}\left(E\left[T\left\{\mathbf{O}, S^{(1)}(u)\right\} | N_1, \dots, N_K\right]\right) = N^{-1} \left[\sum_{k=1}^K p_k \mu_k^2(\phi) - \left\{ \sum_{k=1}^K p_k \mu_k(\phi) \right\}^2 \right]. \quad (9)$$

Because the data from the n_k and $N_k - n_k$ individuals in stratum k who were selected or not in the subsample are independent realizations from the population in stratum k , and because the data from each stratum are independent, after some algebra (see supplementary material, Web Appendix C), we obtain that $\text{Var}[T\{\mathbf{O}, S^{(1)}(u)\} | N_1, \dots, N_K]$ equals

$$\sum_{k=1}^N \left(\frac{N_k}{N} \right)^2 \left[\{\text{Var}_k(\phi) - 2\text{Cov}_k(\phi, h) + \text{Var}_k(h)\} \left(\frac{1}{n_k} - \frac{1}{N_k} \right) + \text{Var}_k(\phi) \left(\frac{1}{N_k} \right) \right], \quad (10)$$

where $\text{Var}_k(\phi) = \text{Var}[\phi\{F, S^{(1)}(u)\} | \varepsilon = k]$, $\text{Var}_k(h) = \text{Var}\{h(V) | \varepsilon = k\}$, and $\text{Cov}_k(\phi, h) = \text{Cov}[\phi\{F, S^{(1)}(u)\}, h(V) | \varepsilon = k]$. Thus $\text{Var}[T\{\mathbf{O}, S^{(1)}(u)\}]$ is given by (9) plus the expectation of (10), and the function $h(V)$ that minimizes this variance is obtained by minimizing (10). Because $\{\text{Var}_k(\phi) - 2\text{Cov}_k(\phi, h) + \text{Var}_k(h)\} = \text{Var}[\phi\{F, S^{(1)}(u)\} - h(V) | \varepsilon = k]$, it is straightforward to show that the optimal choice is $h^{opt}(V) = E[\phi\{F, S^{(1)}(u)\} | V]$, which, although ε is a function of V , we write as

$h^{opt}(V) = E\left[\phi\left\{F, S^{(1)}(u)\right\} | V, \varepsilon\right] = \sum_{k=1}^K I(\varepsilon=k) E\left[\phi\left\{F, S^{(1)}(u)\right\} | V, \varepsilon=k\right]$ to emphasize that the best choice for $h(V)$ in each summand of (8) is the conditional expectation of $\phi\{F, S^{(1)}(u)\}$ given V within each stratum.

If we write $\phi\{F, S^{(1)}(u)\}$ of equation (3) as $\phi_1^*(F) - S^{(1)}(u)$, where $\phi_1^*(F)$ equals

$$\frac{ZI(U \geq u)}{\pi(X) K_c^{(1)}(u, X)} - \left\{ \frac{Z - \pi(X)}{\pi(X)} \right\} H^{(1)}(u, X) + \int_0^u \frac{Z}{\pi(X)} \frac{dM_c^{(1)}(r, X)}{K_c^{(1)}(r, X)} \left\{ \frac{H^{(1)}(u, X)}{H^{(1)}(r, X)} \right\}, \quad (11)$$

then after some algebra, we can show that the optimal estimator for $S^{(1)}(u)$ is given by

$$\hat{S}_{strat}^{(1)}(u) = N^{-1} \sum_{i=1}^N \left[\frac{R_i \phi_1^*(F_i)}{\eta(\epsilon_i)} - \left\{ \frac{R_i - \eta(\epsilon_i)}{\eta(\epsilon_i)} \right\} E\{\phi_1^*(F_i) | V_i, \epsilon_i\} \right], \quad (12)$$

where $E\{\phi_1^*(F_i) | V_i, \epsilon_i\} = \sum_{k=1}^K I(\epsilon_i=k) E\{\phi_1^*(F_i) | V_i, \epsilon_i=k\}$.

As the $E\{\phi_1^*(F) | V, \varepsilon=k\}$ are not known, we would propose models $w_k(V, \psi_k^{(1)})$ for them in terms of parameters $\psi_k^{(1)}$. The $\psi_k^{(1)}$ may be estimated via least squares as follows using the data on subjects in the subsample from stratum k (where both F and V are collected). For each subject i in the subsample, form an estimator $\hat{\phi}_{1i}^*$, say, for $\phi_1^*(F_i)$ as given by (11) by substituting appropriate estimators for $K_c^{(1)}(r, X_i)$, $H^{(1)}(r, X_i)$ and $\pi(X_i)$. The estimator $\hat{\psi}_k^{(1)}$ is derived by minimizing $\sum_{i: R_i=1, \epsilon_i=k} \left\{ \hat{\phi}_{1i}^* - w_k(V_i, \psi_k^{(1)}) \right\}^2$ in $\psi_k^{(1)}$. The estimator for $S^{(1)}$ is the solution to (12) with $E\{\phi_1^*(F_i) | V_i, \epsilon_i\}$ replaced by $w_k(V_i, \psi_k^{(1)})$ when $\epsilon_i = k$.

Care must be taken in estimating $\pi(X)$, $H^{(1)}(r, X)$ and $K_c^{(1)}(r, X)$, as all variables in X are collected only on subjects in the subsample. Because $H^{(1)}(r, X)$ and $K_c^{(1)}(r, X)$ are probabilities conditional on X and $Z = 1$, and because the strata are defined through X_1 (the subset of X observed for all subjects) and Z , consistent estimators for these functions may be derived by modeling and fitting based on the subsample in a manner analogous to that based the entire sample when there is no stratified sampling. However, this is not the case for the propensity score $\pi(X)$. For the sake of clarity, suppose as in ASCERT that the $K = 2M$ strata are defined by all combinations of M categories derived through components of X_1 and the two treatment levels 0 and 1. Denote the number of subjects in the stratum corresponding to the m -th category for treatments 1 and 0 as N_{1m} and N_{0m} and the corresponding numbers in the subsample as n_{1m} and n_{0m} , respectively, $m = 1, \dots, M$. Then, using redundant notation similar to that above and letting Cat be the indicator for category, it is straightforward to show that if $\text{logit}\{P(Z = 1|X, Cat = m)\} = \zeta(X, m)$, then $\text{logit}\{P(Z = 1|X, Cat = m, R = 1)\} = \zeta(X, m) + \log(n_{1m}/N_{1m}) - \log(n_{0m}/N_{0m})$, where $\text{logit}(\theta) = \log\{\theta/(1 - \theta)\}$. Consequently, writing $\pi(X) = \sum_{m=1}^M I(Cat=m) \pi(X, m)$, where $\pi(X, m) = P(Z = 1|X, Cat = m)$, we propose estimating $\pi(X)$ by fitting postulated models for each $\pi(X, m)$. This may be based on fitting logistic regression models to the subsample data (Z_i, X_i) , $i : R_i = 1$ of the form $\text{logit}\{P(Z = 1|X, Cat = m, R = 1)\} = \zeta^*(X, m)$, so that the estimator for $\pi(X, m)$ is of the form $\text{expit}\{\zeta^*(X, m) - \log(n_{1m}/N_{1m}) + \log(n_{0m}/N_{0m})\}$.

A simple but effective suggestion for modeling $w_k(V, \psi_k^{(1)})$ is based on the following: If the additional covariate information X_2 was not predictive for $\pi(X)$, $K_c^{(1)}(r, X)$ or $H^{(1)}(r, X)$, then $E\{\phi_1^*(F) | V, \epsilon=k\} = \phi_1^*(V)$. Of course, this is generally not the case but this suggests that we might consider a model where we assume that

$E\{\phi_1^*(F) | V, \epsilon=k\} = \psi_{0k}^{(1)} + \psi_{1k}^{(1)} \phi_1^*(V)$. Keep in mind that we will obtain a doubly robust estimator for $S^{(1)}(u)$ regardless of whether or not the model for $w_k(V, \psi_k^{(1)})$ is correctly specified. However, a good approximation for $E\{\phi_1^*(F) | V, \epsilon=k\}$ should result in gains in efficiency. Consequently, the suggested strategy is to derive an estimated $\tilde{\phi}_1^*(V_i)$ for individual i exactly as was suggested in Section 2 for estimating (11) using all the data from the main study with only X_1 . The parameter $\psi_k^{(1)}$ is estimated using least squares by minimizing

$$\sum_{i: R_i=1, \epsilon_i=k} \left\{ \hat{\phi}_{1i}^* - \psi_{0k}^{(1)} - \psi_{1k}^{(1)} \tilde{\phi}_1^*(V_i) \right\}^2 \text{ in } \psi_k^{(1)}.$$

The resulting AIPWCC stratified estimator is given by

$$\hat{S}_{strat}^{(1)}(u) = N^{-1} \sum_{i=1}^N \left[\frac{R_i \hat{\phi}_{1i}^*}{\eta(\epsilon_i)} - \left\{ \frac{R_i - \eta(\epsilon_i)}{\eta(\epsilon_i)} \right\} \left\{ \hat{\psi}_{0k}^{(1)} + \hat{\psi}_{1k}^{(1)} \tilde{\phi}_1^*(V_i) \right\} \right]. \quad (13)$$

As we will demonstrate in the simulations studies in the next section this strategy leads to an estimator that has good performance.

The variance estimator for the estimator of $S^{(1)}(u)$ would then use (9) plus (10), substituting the natural estimators for p_k , $\mu_k(\phi)$, $\mu_k(h)$, $\text{Var}_k(\phi)$, $\text{Var}_k(h)$ and $\text{Cov}_k(\phi, h)$, where here we take $h(V) = \sum_{k=1}^K I(\epsilon=k) \left\{ \psi_{0k}^{(1)} + \psi_{1k}^{(1)} \phi_1^*(V) - S^{(1)}(u) \right\}$. As in the case when there is no stratified sampling, the double robustness property as well as the adequacy of the proposed variance estimator will depend on whether or not the models for $\pi(X)$ (used for the propensity score), $K_c^{(1)}(r, X)$ and $H^{(1)}(r, X)$ are correctly specified.

3.2 Inference on Difference of Treatments

Since we are also interested in making inference on $S_{diff}(u)$, we also provide an estimator for $S_{diff}(u)$ and its variance estimator when a substudy is conducted. In Section 3.1, we have shown that (13) is a consistent estimator of $S^{(1)}(u)$. Similarly, we can estimate $S_{diff}(u)$ by

$$\hat{S}_{diff, strat}(u) = N^{-1} \sum_{i=1}^N \left(\frac{R_i \hat{\phi}_{diff, i}^*}{\eta(\epsilon_i)} - \left\{ \frac{R_i - \eta(\epsilon_i)}{\eta(\epsilon_i)} \right\} \hat{h}_{diff}^*(V_i) \right),$$

where $\hat{\phi}_{diff, i}^* = \hat{\phi}_{1i}^* - \hat{\phi}_{0i}^*$, for treatment $j = (0, 1)$, $\hat{\phi}_{ji}^*$ is the estimator for $\phi_j^*(F_i)$ defined by

equation (6) with $X_i = (X_{1i}^T, X_{2i}^T)^T$ and

$\hat{h}_{diff}^*(V_i) = \left\{ \hat{\psi}_{0k}^{(1)} + \hat{\psi}_{0k}^{(0)} \tilde{\phi}_1^*(V_i) \right\} - \left\{ \hat{\psi}_{1k}^{(0)} + \hat{\psi}_{1k}^{(0)} \tilde{\phi}_0^*(V_i) \right\}$, where $\hat{\psi}_k^{(j)}$ is obtained using least

squares by minimizing $\sum_{i: R_i=1, \epsilon_i=k} \left\{ \hat{\phi}_{ji}^* - \psi_{0k}^{(j)} - \psi_{1k}^{(j)} \tilde{\phi}_j^*(V_i) \right\}^2$, $j = 0, 1$. The asymptotic

variance can be estimated by substituting the above $\hat{\phi}_{diff}^*$ and $\hat{h}_{diff}(V)$ into (9) and (10).

4. Simulations

Several simulation studies have been carried out to demonstrate the performance of the proposed estimators. Results from the main simulation study are shown here and results from other scenarios can be found in the Web Appendices. The simulation studies consist of two parts. We first demonstrate the double robustness property of the proposed estimator for $S^{(1)}(u)$ using the data only from the main study assuming that all potential confounders were available. In the second set of simulations we show results when a substudy implementing a stratified sampling design to collect additional potential confounding covariates on a subsample of patients is used. We generated 500 replications and within each replicate, $N = 2,000$ observations are generated as follows: $B \sim \text{Bernoulli}(0.5)$, $W \sim N(0, 1)$ and $X_2 \sim N(0, 1)$, mutually independent. Here B is the strata variable with two levels, mimicking the ASCERT data where patients have either two- or three-vessel disease. We denote by $X_1 =$

$(B, W, W^2)^T$, and $X = (X_1^T, X_2)^T$. The treatment assignment propensity is generated by

$P(Z=1|X) = \frac{\exp(0.1B + 0.1W + 0.5W^2 + 0.5X_2)}{1 + \exp(0.1B + 0.1W + 0.5W^2 + 0.5X_2)}$, and the treatment-specific hazard functions are given by $\lambda_{Z=1}(t|X) = \exp(-1 + 0.1B + 0.1W + 0.5W^2 + 0.5X_2)$ and $\lambda_{Z=0}(t|X) = \exp(-0.5 + 0.1B + 0.1W + 0.5W^2 + 0.5X_2)$, respectively. We also generated an independent censoring variable C as $\text{Uniform}(0, 4)$.

We considered four scenarios: the first one uses all the covariates X to fit all three models $\pi(X)$, $K_c^{(1)}(r, X)$ and $H^{(1)}(r, X)$. This scenario is the “correct-correct” model, where $\pi(X)$, $K_c^{(1)}(r, X)$ and $H^{(1)}(r, X)$ are estimated consistently. In the second scenario, we leave out the quadratic term W^2 and use only $(B, W, X_2)^T$ to fit $\pi(X)$, $K_c^{(1)}(r, X)$ and use X to fit $H^{(1)}(r, X)$. We refer to this as the “wrong propensity, correct regression” scenario. The third one, which is referred as the “correct propensity, wrong regression” scenario, is where we use X to fit $\pi(X)$, $K_c^{(1)}(r, X)$ and $(B, W, X_2)^T$ to fit $H^{(1)}(r, X)$. Finally, in the last scenario, we use only $(B, W, X_2)^T$ to fit $\pi(X)$, $K_c^{(1)}(r, X)$ and $H^{(1)}(r, X)$ and this is the “wrong propensity, wrong regression” scenario. Because the distribution of the censoring variable is independent of the covariates, the misspecification of the coarsening model is only through

$\pi(X)$. In Web Appendix I, we explore the scenario where misspecification of the coarsening model occurs because of misspecifying $K_c^{(1)}(r, X)$ as well.

In Table 1, we present results for estimating the treatment-specific survival distribution for treatment 1, $S^{(1)}(u)$. Results regarding the difference between two treatments $S_{diff}(u)$ are similar and shown in the supplementary material, Web Appendix D. We can see in scenarios 1, 2, and 3 that the estimator defined in (4) is indeed consistent as desired. For scenarios 2 and 3, even though parts of the models (propensity models/regression model) are not correctly specified, the resulting estimator is still consistent. This illustrates the doubly-robustness property. In scenario 4, the estimator is biased as is expected because the models $\pi(X)$ and $H^{(1)}(u, X)$ are misspecified. The coverage rate is calculated using the 95% confidence interval derived as the estimator plus/minus 1.96 estimated standard errors using the sandwich variance formulas for the four scenarios. Because in scenario 4, the estimator $\hat{S}^{(1)}(u)$ is not consistent, the coverage rate is far below 0.95. In theory, the estimated variance in scenario 3 may be conservative however, as expected, the coverage rate was close to the nominal rate of 0.95.

We also note that the coverage rate in scenario 2 did not achieve the 0.95 nominal rate. This indeed is the scenario discussed at the end of Section 2.1 where the sandwich variance estimator is not expected to be consistent. For this scenario we also estimated the asymptotic variance using a nonparametric bootstrap using 100 bootstrap replicates (see Web Appendix E). This resulted in a variance estimator that was very close to the Monte-Carlo variance and a confidence interval with coverage very close to the nominal rate.

We have also studied the case when the sample size is relatively small $N = 200$. The simulation results showed that even with this smaller sample size the survival curve estimator showed virtually no bias. The proposed standard error estimator, on the other hand, tends to be unstable especially at the tail of the curve. The bootstrap standard error estimator performs much better and hence is preferable for such a smaller sample size. See Web Appendix F for details.

Other competing estimators were also considered: the inverse probability weighted estimator, the outcome regression estimator and the treatment-specific Kaplan-Meier estimator. The numerical results can be found in Web Appendix G. When correctly specified, the first two estimators are consistent. Our proposed AIPWCC estimator is more efficient than the inverse probability weighted estimator and slightly less efficient than the outcome regression estimator. On the other hand, both of these estimators are biased if the models are misspecified. The unadjusted treatment-specific Kaplan-Meier estimator is always biased as it does not account for confounding.

In the next set of simulations we consider the case where we conduct a substudy using stratified sampling and consider the performance of our AIPWCC estimators. Using the same simulation setup as before, we collected X_1 on the main study of $N = 5,000$ observations and randomly sampled $n_{11} = n_{12} = n_{01} = n_{02} = 300$ observations from each of $K = 4$ stratum combination of (Z, B) without replacement and collected additional covariates X_2 on these subsamples. In these simulations we used the correct models for $\pi(X)$, $K_c^{(j)}(r, X)$, and $H^{(j)}(r, X)$, $j = 0, 1$.

We compared the consistency and efficiency of three estimators for $S^{(1)}(u)$ under this setting. The first estimator is the estimator (13). We refer to this estimator as the **AIPWCC stratified estimator**. The second one used the estimator resulting from (7), where the augmentation term $h(V_i) = 0$, referred to as the **weighted estimator**. We also considered the infeasible estimator (4) with covariates X_1, X_2 on the whole dataset as the gold standard.

As expected, all three estimators along with their estimated standard errors are consistent and the confidence intervals have coverage rate close to the nominal rate of 0.95. The relative efficiency is computed by taking the ratio of the Monte-Carlo variance of each estimator over the Monte-Carlo variance of the gold standard estimator. Here we see the large gains in efficiency using the AIPWCC stratified estimator compared to the weighted estimator with little loss of efficiency compared to the gold standard. This illustrates that choosing the augmentation term appropriately can recover a great deal of information from the main study and result in large gains in efficiency.

We also compared the performance of estimator for $S_{diff}(u)$ (see supplementary material, Web Appendix D) and the results are similar to the above estimators of $S^{(1)}(u)$. In addition, we conducted simulation studies where the models for either the propensity score $\pi(X)$ or the conditional survival distribution $H^{(j)}(r, X)$ were misspecified, and the results were as expected. The estimators were doubly robust and the estimated variance and coverage behaved as it did for the first set of simulations where X was collected on all the patients in the main study (see supplementary material, Web Appendix H).

5. Analysis of the ASCERT Study

In this section, we apply the proposed AIPWCC estimator to data from the ASCERT study. The goal of the analysis was to estimate the AIPWCC-adjusted 4-year survival distribution for PCI and CABG using baseline covariate data from the main ASCERT study database augmented with information on coronary anatomy from a subsample of patients in the ASCERT angiographic companion study. The sampling frame consisted of records from 9,800 patients in the ASCERT database who underwent a coronary revascularization procedure (CABG or PCI) at one of 54 hospitals participating in the ASCERT study that agreed to be part of the companion study. For the purpose of this analysis we considered the patients from the 54 hospitals of the companion study to be the main focus of inference. Eighteen covariates were used on the full sample which included demographics (e.g., age, sex), risk factors (e.g., body mass index, smoking), symptoms and history of cardiovascular disease (e.g., chest pain, congestive heart failure), and comorbidities (e.g., diabetes). The subsample includes records from approximately 2,000 patients chosen by design (roughly 500 in each of the four strata determined by all combinations of the two treatments and whether or not the patients had two- or three-vessel disease). Information collected on the subsample includes features of the patient's coronary anatomy (e.g., left-side dominance) and features of each individual blockage (e.g., lesion length, tortuosity, calcification, degree of stenosis).

The original sampling frame comprised 9,800 patients. Subsequently it was determined that some patients were ineligible for analysis. Consequently, the main study participants consisted of 7,393 eligible patients among the 9,800 patients in the 54 hospitals of the companion study and the subsample consisted of 1,554 eligible patients among the 2,049 originally sampled. For details on the sample size before and after eligibility was determined, refer to the supplementary material, Web Appendix J.

Figure 1 presents plots of the estimated AIPWCC-adjusted survival distributions with accompanying 95% pointwise confidence intervals for PCI, CABG and their difference as estimated by the proposed AIPWCC estimator. Overall, results are highly consistent with the published primary ASCERT analysis (Weintraub et al., 2012). Short-term risk appears to be lower with PCI but long-term risk appears to be lower with CABG.

Table 3 presents the estimated AIPWCC-adjusted cumulative incidence (one minus the survival distribution) of mortality at two time points (30 days and 4 years) and compares two

versions of the adjusted analysis together with the unadjusted treatment-specific Kaplan-Meier estimators. The first AIPWCC estimator adjusts only for covariates available in the full sample and includes all 7,393 patients. The second analysis used the AIPWCC stratified estimator and includes the additional covariates on coronary anatomy collected on the subsample and corresponds to the analysis shown in Figure 1. In each case, the 30-day estimate favors PCI whereas the 4-year estimate favors CABG. Both adjusted analyses gave similar results with possibly CABG having slightly better performance compared to PCI in the unadjusted analysis.

Overall all three analyses gave similar results suggesting that the confounding, either with covariates used in the main study or those collected in the substudy, was not substantial.

6. Conclusion and Discussion

Viewing censored survival data from an observational study as monotonely coarsened versions of treatment-specific potential survival times, we were able to use the powerful semiparametric theory for missing data to derive doubly robust estimators for the underlying treatment-specific survival distributions as well as the difference in these treatment-specific distributions. We also derived an estimator for the asymptotic variance of these estimators. Using simulation studies we demonstrated that the large sample properties of these estimators and corresponding confidence intervals worked well in most cases where we expected consistent estimators.

More importantly, we showed that if additional covariate information is necessary to make the assumption of no unmeasured confounders tenable, then we could conduct a substudy using a stratified sampling design to collect such additional covariates and we have proposed a method for obtaining doubly robust estimators with such a design which uses the data from the main study to gain efficiency. The corresponding estimator and standard error estimator was shown in simulation studies to perform well.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by NIH grants 1R01HL118336-01, R37AI031789, P01CA142538 and RC2HL101489. The authors gratefully acknowledge Dr. William Weintraub, Dr. Fred Edwards, and the ASCERT investigator team for assembling the study database.

References

- Breslow N, Cain KC. Logistic regression for two-stage case-control data. *Biometrika*. 1988; 75:11–20.
- Cole SR, Hernán MA. Adjusted survival curves with inverse probability weights. *Computer Methods and Programs in Biomedicine*. 2004; 75:45–49. [PubMed: 15158046]
- Hubbard, AE.; van der Laan, MJ.; Robins, JM. Nonparametric locally efficient estimation of the treatment specific survival distributions with right censored data and covariates in observational studies.. In: Halloran, E.; Berry, D., editors. *Statistical Models in Epidemiology: The Environment and Clinical Trials*. Springer; New York: 1999. p. 134–178.
- Flanders WD, Greenland S. Analytic methods for two-stage case control studies and other stratified designs. *Statistics in Medicine*. 1991; 10:739–747. [PubMed: 2068427]
- Kaplan EL, Meier Paul. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*. 1958; 53:457–481.

- Lee J, Yoshizawa C, Wilkens L, Lee HP. Covariance adjustment of survival curves based on Cox's proportional hazards regression model. *Computer Applications in the Biosciences*. 1992; 8:23–27. [PubMed: 1568122]
- Makuch RW. Adjusted survival-curve estimation using covariates. *Journal of Chronic Diseases*. 1982; 35:437–443. [PubMed: 7042727]
- Mark SD, Katki HA. Specifying and implementing nonparametric and semiparametric survival estimators in two-stage (nested) cohort studies with missing case data. *Journal of the American Statistical Association*. 2006; 101:460–471.
- Robins, JM.; Rotnitzky, A. Recovery information and adjustment for dependent censoring using surrogate markers.. In: Jewell, N.; Dietz, K.; Farewell, V., editors. *AIDS Epidemiology, Methodological Issues*. Birkhauser; Boston: 1992. p. 297-331.
- Rubin DB. Bayesian inference for causal effects: The role of randomization. *Annals of Statistics*. 1978; 6:34–58.
- Stürmer T, Glynn RJ, Rothman KJ, Avorn J, Schneeweiss S. Adjustments for unmeasured confounders in pharmacoepidemiologic database studies using external information. *Medical Care*. 2007; 45:S157–S165.
- Tsiatis, AA. *Semiparametric Theory and Missing Data*. Springer; New York: 2006.
- Weintraub WS, Grau-Sepulveda MV, Weiss JM, O'Brien SM, Peterson ED, Kolm P, Zhang Z, Klein LW, Shaw RE, McKay C, Ritzenthaler LL, Popma JJ, Messenger JC, Shahian DM, Grover FL, Mayer JE, Shewan CM, Garratt KN, Moussa ID, Dangas GD, Edwards FH. Comparative effectiveness of revascularization strategies. *New England Journal of Medicine*. 2012; 366:1467–1476. [PubMed: 22452338]
- Zhao LP, Lipsitz S. Design and analysis of two-stage studies. *Statistics in Medicine*. 1992; 11:769–782. [PubMed: 1594816]

Figure 1.

Estimated AIPWCC Stratified Survival Estimator for CABG and PCI. The confidence interval is computed as the treatment-specific survival estimator plus/minus 1.96 times the proposed standard error estimator. This figure appears in color in the electronic version of this article.

Table 1

Simulation 1, no stratified sampling, sample size N=2000.

	u	$S^{(1)}(u)$	$\hat{S}^{(1)}(u)$	Bias	Monte-Carlo SE	Estimated SE	Coverage Rate
Scenario 1	0.5	0.693	0.694	0.001	0.013	0.013	0.96
	1.0	0.523	0.524	0.001	0.015	0.015	0.95
	1.5	0.408	0.408	0.000	0.016	0.016	0.94
	2.0	0.324	0.325	0.001	0.017	0.016	0.94
	2.5	0.262	0.262	0.000	0.017	0.017	0.95
	3.0	0.214	0.214	0.000	0.017	0.018	0.96
Scenario 2	0.5	0.693	0.694	0.001	0.013	0.013	0.96
	1.0	0.523	0.524	0.001	0.015	0.015	0.93
	1.5	0.408	0.408	0.000	0.016	0.015	0.94
	2.0	0.324	0.325	0.001	0.016	0.015	0.93
	2.5	0.262	0.261	-0.001	0.016	0.015	0.93
	3.0	0.214	0.214	0.000	0.016	0.016	0.94
Scenario 3	0.5	0.693	0.694	0.001	0.013	0.013	0.96
	1.0	0.523	0.524	0.001	0.015	0.016	0.95
	1.5	0.408	0.408	0.000	0.016	0.016	0.95
	2.0	0.324	0.325	0.001	0.017	0.017	0.95
	2.5	0.262	0.261	-0.001	0.017	0.017	0.95
	3.0	0.214	0.214	0.000	0.017	0.018	0.96
Scenario 4	0.5	0.693	0.661	-0.032	0.014	0.014	0.35
	1.0	0.523	0.489	-0.035	0.015	0.015	0.36
	1.5	0.408	0.374	-0.034	0.016	0.016	0.41
	2.0	0.324	0.294	-0.031	0.016	0.016	0.49
	2.5	0.262	0.234	-0.028	0.016	0.016	0.57
	3.0	0.214	0.190	-0.025	0.016	0.017	0.67

Table 2

Simulation 2, with stratified sampling, sample size N=5000 n=1200.

	u	$S^{(1)}(u)$	$\hat{S}_{strat}^{(1)}(u)$	Bias	Monte-Carlo SE	Estimated SE	Coverage Rate	Relative Efficiency
AIPWCC	0.5	0.693	0.694	0.000	0.009	0.009	0.96	1.20
	1.0	0.523	0.523	0.000	0.011	0.011	0.95	1.31
	1.5	0.408	0.408	0.000	0.012	0.012	0.95	1.36
	2.0	0.324	0.324	0.000	0.012	0.012	0.95	1.34
	2.5	0.262	0.262	0.000	0.013	0.013	0.95	1.37
	3.0	0.214	0.213	-0.001	0.013	0.013	0.94	1.32
Weighted	0.5	0.693	0.694	0.001	0.019	0.019	0.95	5.19
	1.0	0.523	0.524	0.001	0.021	0.021	0.96	5.10
	1.5	0.408	0.410	0.002	0.023	0.023	0.94	5.56
	2.0	0.324	0.325	0.001	0.022	0.023	0.96	4.76
	2.5	0.262	0.262	0.000	0.023	0.024	0.96	4.46
	3.0	0.214	0.214	0.000	0.024	0.025	0.96	4.26
Gold	0.5	0.693	0.693	0.000	0.008	0.008	0.95	1.00
	1.0	0.523	0.523	0.000	0.009	0.010	0.95	1.00
	1.5	0.408	0.407	0.001	0.010	0.010	0.94	1.00
	2.0	0.324	0.324	0.000	0.010	0.010	0.96	1.00
	2.5	0.262	0.261	-0.001	0.011	0.011	0.94	1.00
	3.0	0.214	0.214	0.000	0.012	0.011	0.94	1.00

Table 3

Estimated AIPWCC-Adjusted Cumulative Incidence of Mortality at 30 Days and 4 Years by Treatment Group, Expressed as Percentage With 95% Confidence Interval. The proposed standard error estimator is used to compute the confidence interval.

		30-day (%)	4-year (%)
Unadjusted Kaplan Meier	CABG	1.77 (1.24, 2.30)	14.69 (12.90, 16.49)
	PCI	1.14 (0.84, 1.43)	18.64 (17.25, 20.03)
	Difference	-0.64 (-1.24, -0.03)	3.95 (1.68, 6.22)
AIPWCC (Full)	CABG	2.17 (1.50, 2.83)	15.93 (13.98, 17.88)
	PCI	1.08 (0.79, 1.38)	18.31 (16.95, 19.67)
	Difference	-1.08 (-1.81, -0.36)	2.38 (0.04, 4.73)
AIPWCC (Stratified)	CABG	2.26 (1.39, 3.13)	14.83 (12.00, 17.65)
	PCI	1.21 (0.85, 1.57)	17.98 (16.70, 19.27)
	Difference	-1.05 (-1.99, -0.11)	3.16 (0.07, 6.25)