Statistica Sinica **31** (2021), 673-699
doi:https://doi.org/10.5705/ss.202019.0096

CAUSAL PROPORTIONAL HAZARDS ESTIMATION WITH A BINARY INSTRUMENTAL VARIABLE

Behzad Kianian[1], Jung In Kim[2], Jason P. Fine[2] and Limin Peng[1]

[1]*Emory University and* [2]*University of North Carolina at Chapel Hill*

Abstract: Instrumental variables (IVs) are useful for estimating causal effects in the presence of unmeasured confounding. IV methods are well developed for uncensored outcomes, particularly for structural linear equation models, where simple two-stage estimation schemes are available. Extending these methods to survival settings is challenging, partly because of the nonlinearity of the popular survival regression models, and partly because of the complexity of right censoring and other survival features. Motivated by the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial, we develop a simple causal hazard ratio estimator in a proportional hazards model with right-censored data. The method exploits a special characterization of IVs that enables the use of an intuitive inverse weighting scheme that is generally applicable to more complex survival settings with left truncation, competing risks, or recurrent events. We rigorously establish the asymptotic properties of the estimators, and provide plug-in variance estimators. The proposed method can be implemented in standard software, and is evaluated through extensive simulation studies. We apply the proposed IV method to a data set from the PLCO Cancer Screening Trial to identify the causal effect of flexible sigmoidoscopy screening on colorectal cancer survival, which may be confounded by informative noncompliance with the assigned screening regimen.

Key words and phrases: Causal treatment effect, Cox proportional hazards model, instrumental variable, noncompliance.

1. Introduction

Many recent studies have focused on understanding the causal effect of a treatment or exposure on an outcome of interest (Holland (1986)). In observational studies, unmeasured confounding is a major obstacle to estimating the causal effect of a nonrandomized exposure on disease etiology. Such a challenge also arises in well-designed randomized clinical trials. When there are issues of noncompliance in the treatment arms, the treatment decision may be based on latent (unobserved) factors that strongly correlate with clinical outcomes. This

Corresponding author: Limin Peng, Department of Biostatistics and Bioinformatics, Emory University, Atlanta, GA, 30322, U.S.A. E-mail: lpeng@emory.edu.

would result in bias from unmeasured confounding and, hence, complicate the task of estimating the "efficacy" of the treatment.

Instrumental variables (IVs) are useful for estimating causal treatment or exposure effects in such settings (Imbens and Angrist (1994); Angrist, Imbens and Rubin (1996); Loeys and Goetghebeur (2003); Li and Lu (2015); Li and Gray (2016); MacKenzie, Løberg and O'Malley (2016)). Informally, IVs are independent of unmeasured confounders, related to the treatment, and related to the outcome only through the treatment (Baiocchi, Cheng and Small (2014)). In observational studies, numerous instruments are available for the estimation of treatment or exposure causal effects (Baiocchi, Cheng and Small (2014)). In randomized clinical trials with noncompliance, the treatment assignment mechanism can serve as an IV.

The motivating example for this work is the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial, which is a multi-center randomized trial designed to evaluate the effectiveness of screening using flexible sigmoidoscopy relative to that of usual care. In this study, 77,449 subjects were assigned randomly to the intervention group, but only 85% complied with the assigned sigmoidoscopy protocol. Such noncompliance may be outcome-related. For example, relatively healthy individuals may be more likely to skip the screening. In the presence of unmeasured confounding, neither an intent-to-treat (ITT) analysis nor an "as-treated" analysis would adequately assess the causal benefit of the treatment (i.e., flexible sigmoidoscopy screening). A possible remedy is an IV analysis that properly adjusts for the selection bias induced by subjects' post-randomization care selection. The assigned treatment in a randomized trial serves a natural IV in this analysis.

The IV methodology focuses primarily on linear models and continuous outcomes in contexts without censoring. More recently, interest has grown in using the IV methodology for time-to-event data with right censoring. For example, Baker (1998) developed an IV method for randomized trials using all-or-none compliance and discrete-time survival data that extended the method of latent class IVs (also known as the local average treatment effect method or complier average causal effect method) (Baker and Lindeman (1994); Imbens and Angrist (1994)). Building on the work of Baker (1998), Nie, Cheng and Small (2011) developed an estimation method with improved efficiency. Robins and Tsiatis (1991) considered a structural accelerated failure time model, and developed estimators for the causal treatment effect in a context of noncompliance and administrative censoring only. Joffe (2001) provides a detailed discussion of this

general approach. Imposing parametric distributional assumptions, Li and Lu (2015) developed a Bayesian approach for an IV analysis using censored time-to-event outcomes under a two-stage linear model. Li, Fine and Brookhart (2015) and Tchetgen et al. (2015) developed IV-based methods for additive hazards modeling of time-to-event data. Martinussen et al. (2017) studied a structural cumulative survival model that assesses the time-varying exposure effect directly on the scale of the survival function.

The proportional hazards model is a popular formulation for the effects of treatment and covariates in time-to-event analyses. Several IV approaches have been developed for proportional hazards modeling. For example, for the special case of all- or none- compliance without covariates, Loeys and Goetghebeur (2003) proposed an estimate for the complier proportional hazards effect of the treatment. They derive a properly imputed partial likelihood that recovers the unobserved information on the treatable subgroup in the control arm. Working in a noncompliance setting, Cuzick et al. (2007) constructed a Mantel–Haenszel-type estimator for the case without covariates, and a partial likelihood-based estimator when covariates are present and independent of the types of compliance. A full-likelihood-based approach has been explored for situations in which the covariates are correlated with the types of compliance. Li and Gray (2016) proposed an EM algorithm for this full likelihood-based estimation. Yu et al. (2015) estimate causal estimands, including the complier average causal effect, complier survival probability, and complier quantile causal effect, under a semiparametric transformation model. They adapted the nonparametric likelihood estimation technique of Zeng and Lin (2007) to develop an EM algorithm for implementing the proposed estimation, as well as providing theoretical justifications. Although likelihood-based strategies accommodate both censoring and covariates when estimating the causal treatment effect using censored time-to-event data, the resulting estimation and inference procedures are, in general, very complicated. The resulting computational complexity and stability may become unmanageable when the sample size is large, as in the PLCO Cancer Screening Trial. Furthermore, they require that we specify causal models for all latent compliance classes, not just for the class of interest, which may limit the robustness of these methods to potential model misspecification.

In this work, we develop a new IV approach for estimating causal treatment effects under proportional hazards modeling of time-to-event outcomes that are subject to independent right censoring. The causal estimand of interest is defined within the latent subgroup of compliers, and differs from the population

causal hazards ratio considered in other recent IV methods (e.g., Martinussen, Nørbo Sørensen and Vansteelandt (2019); Wang et al. (2018)). Notably, our method does not need to impose regression models for the latent compliance classes other than on the complier subgroup. Our key strategy adapts the seminal work of Abadie (2003), which provides a simple link between the unconditional moment of the observed data and the conditional moment, given the latent complier group. Abadie (2003) developed a simple weighting strategy that is easily applied to estimating equations that are sums of independent terms. However, an analogous application to the proportional hazards regression is not straightforward. This is because the partial likelihood does not yield an estimating function of the simple form as a sum of independent terms, as with the least squares criterion for linear regressions. To circumvent this difficulty, we incorporate the weighting idea of Abadie (2003) in the asymptotic influence functions of the partial likelihood score equation. We further devise the weighting scheme such that the calculations of the parameter estimate can be easily and stably implemented using existing software for weighted proportional hazards regressions. Compared with currently available weighting methods, such as the time-dependent weighted estimator proposed by Li and Gray (2016) and the estimator based on principal stratification weighting (MacKenzie, Løberg and O'Malley (2016)), our weighting strategy is relatively simple. However, it can be applied readily to more complex survival settings, for example, in the presence of left truncation, competing risks, or recurrent events; see Section S3 of the Supplementary Material for more information. Such a broad applicability appears lacking in existing IV approaches for proportional hazards models. Finally, we establish the large-sample properties of the proposed parameter estimators, including those of consistency and asymptotic normality.

In Section 2, we introduce the potential outcomes framework, including the latent compliance groups, IV assumptions, and setup of the causal proportional hazards regression. We next describe the proposed estimation procedure using randomly censored data, discuss the computational considerations, and present a modification of the proposed method that exhibits improved computational features. Here, we present a rigorous examination of the consistency and asymptotic normality of the estimators. The results include a closed form for the asymptotic variance of the estimator and a consistent plug-in variance estimator. Bootstrap variance estimates are also provided. The results from extensive simulations are reported Section 3, and demonstrate that the methods perform well for realistic sample sizes. In Section 4, we apply our methods to the data from the PLCO

Cancer Screening Trial. Section 5 concludes the paper.

2. Weighted Partial Likelihood Estimation for Causal Proportional Hazards Models

2.1. Potential outcomes framework

We introduce the potential outcomes framework and notation commonly employed in the causal inference literature. Consider potential survival times T_1 and T_0 based on receiving ($D = 1$) and not receiving the treatment ($D = 0$), respectively. Define V as a binary IV, and define the potential treatment D_v such that D_1 denotes the treatment received when $V = 1$, and D_0 denotes the treatment received when $V = 0$. Following the terminology of Abadie (2003), subjects can be classified into four latent compliance groups, based on the potential treatment indicators: *compliers* ($D_1 > D_0$), *always-takers* ($D_1 = D_0 = 1$), *never-takers* ($D_1 = D_0 = 0$), and *defiers* ($D_1 < D_0$). In the PLCO Cancer Screening Trial, compliers are those individuals assigned to the intervention group who underwent flexible sigmoidoscopy screening. Always-takers (or never-takers) always (or never) take the flexible sigmoidoscopy screening. Defiers would take the flexible sigmoidoscopy screening if assigned to the usual care group, but not if assigned to the intervention group. Because D_1 and D_0 cannot be observed simultaneously, we are not able to determine the latent compliance group membership of any individual based on the observed data alone.

Define the potential outcome for each subject as T_{vd}, which represents the survival time T if $V = v$ and $D = d$. Let $\boldsymbol{X}$ represent the covariate vector. We re-state several key assumptions from Abadie (2003) about the IV, V. Let $T_{vd}, \boldsymbol{X}, V, D, D_v$ be defined as above. Then, we have the following:

(A1) *Independence of the instrument:*

$$(T_{00}, T_{01}, T_{10}, T_{11}, D_0, D_1) \perp V | \boldsymbol{X}.$$

(A2) *Exclusion of the instrument:* $P(T_{1d} = T_{0d} | \boldsymbol{X}) = 1$, for $d = 0, 1$.

(A3) *First stage:* $0 < P(V = 1 | \boldsymbol{X}) < 1$ and $P(D_1 = 1 | \boldsymbol{X}) > P(D_0 = 1 | \boldsymbol{X})$.

(A4) *Monotonicity:* $P(D_1 \geq D_0 | \boldsymbol{X}) = 1$.

Assumption (A1) states that the instrument V is as good as random conditional on the covariates $\boldsymbol{X}$, or equivalently, that V is independent of unmeasured

confounders conditional on $\boldsymbol{X}$. Assumption (A2) states that the instrument V influences the outcome T only through its effect on the treatment D. Assumption (A3) states that every subject has some chance of receiving the instrument V, conditional on the covariate $\boldsymbol{X}$, and that conditional on $\boldsymbol{X}$, V has an effect on the treatment received. Finally, assumption (A4) states that, with probability one, defiers do not exist.

2.2. Model formulation

Our focus is to estimate and make inferences about the treatment effect for the latent group of compliers. Specifically, we adopt Cox's proportional hazards regression model to formulate the effects of the treatment and covariates for compliers:

$$h(t; D, \boldsymbol{X}) = h_0(t) \exp\{\beta_d D + \boldsymbol{\beta}_x^T \boldsymbol{X}\}, \tag{2.1}$$

where $h(t; D, \boldsymbol{X})$ is the hazard function for compliers, defined as

$$h(t; D, \boldsymbol{X}) = \lim_{\Delta t \to 0} \frac{\Pr(t \le T < t + \Delta t | T \ge t, D_1 > D_0, D, \boldsymbol{X})}{\Delta t},$$

and $h_0(t)$ is an unspecified baseline hazard at time t. In model (2.1), β_d is the causal estimand of the primary interest. This is because, by assumption (A.1), it is easy to show that $\Pr(t \le T < t + \Delta t | T \ge t, D_1 > D_0, D = d, \boldsymbol{X}) = \Pr(t \le T^{(d)} < t + \Delta t | T^{(d)} \ge t, V = d, D_1 > D_0, \boldsymbol{X}) = \Pr(t \le T^{(d)} < t + \Delta t | T^{(d)} \ge t, D_1 > D_0, \boldsymbol{X})$. Thus, β_d can be interpreted as the causal treatment effect for compliers, after adjusting for the covariate effects captured by $\boldsymbol{\beta}_x$ (Abadie (2003)). Such quantities are well studied in the literature (e.g., Loeys and Goetghebeur (2003); Cuzick et al. (2007); Yu et al. (2015)). Note that the proportional hazards model (2.1) is assumed for compliers only. In contrast, likelihood-based approaches (e.g., Cuzick et al. (2007); Yu et al. (2015); Li and Gray (2016)) typically require distributional modeling for the other compliance subgroups (e.g., *always takers*, *never-takers*), and may be biased under a misspecification of these models.

2.3. Estimation

In practice, T is often subject to right censoring by C; thus, we observe $W = min(T, C)$ and $\delta = I(T \le C)$, instead of T. We adopt the standard random censoring assumptions that C is independent of T, conditional on $(V, D, \boldsymbol{X})$. We further assume that C is independent of V, given $\boldsymbol{X}$. Define $\boldsymbol{O} = (W, \delta, D, \boldsymbol{X}, V)$. The observed data consist of n independently and identically distributed (i.i.d.)

replicates of $\boldsymbol{O}$, denoted by $\{\boldsymbol{O}_i\}_{i=1}^n = \{(W_i, \delta_i, D_i, \boldsymbol{X}_i, V_i)\}_{i=1}^n$. Define $Y_i(t) = I(W_i \geq t)$ and $N_i(t) = I(W_i \leq t, \delta_i = 1)$, which represent the at-risk process and the observed event-counting process, respectively, for subject i. We further assume that there are no ties (i.e., $dN_i(t) \leq 1$). In what follows, we use the subscript i to differentiate between the population quantities and their sample analogues.

Let $\boldsymbol{\beta_0} = (\beta_d, \boldsymbol{\beta}_x)$ and $\boldsymbol{Z} = (D, \boldsymbol{X})$. When all subjects are known to be compliers, we can estimate $\boldsymbol{\beta_0}$ using the standard Cox regression analysis (Andersen and Gill (1982)). This is because, in this case, the hazard function for the overall study population, $\lambda(t|\boldsymbol{Z}) \equiv \lim_{\Delta t \to 0} \Pr(t \leq T \leq t + \Delta t | T \geq t, D, \boldsymbol{X})/\Delta t$, is equal to that for the latent complier subgroup, $\exp(\boldsymbol{\beta}^T \boldsymbol{Z})\ h_0(t)$. Then, $M(t) \equiv N(t) - \int_0^t Y(s) \exp(\boldsymbol{\beta}_0^T \boldsymbol{Z}) h_0(s) ds$ is a martingale, and, thus, we can obtain a consistent estimator of $\boldsymbol{\beta}_0$ as the solution to the partial likelihood score equation,

$$\boldsymbol{U}_n(\boldsymbol{\beta}) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \int_0^\infty \left\{ \mathbf{Z}_i - \frac{\boldsymbol{S}_n^{(1)}(\boldsymbol{\beta}, s)}{S_n^{(0)}(\boldsymbol{\beta}, s)} \right\} dN_i(s), \tag{2.2}$$

where $\boldsymbol{S}_n^{(j)}(\boldsymbol{\beta}, s) = \sum_{l=1}^n Y_l(s) \mathbf{Z}_l^{\otimes j} e^{\boldsymbol{\beta}^T \mathbf{Z}_l}$, for $j = 0, 1, 2$. Here, and in what follows, for a vector $\boldsymbol{v}$, $\boldsymbol{v}^{\otimes 0} = 1$, $\boldsymbol{v}^{\otimes 1} = \boldsymbol{v}$, and $\boldsymbol{v}^{\otimes 2} = \boldsymbol{v}\boldsymbol{v}^T$.

Next, we consider the more realistic case in which the study population consists of both compliers and noncompliers. In this case, $\lambda(t|\boldsymbol{Z})$ usually deviates from the hazard function assumed for the complier group, $h_0(t) \exp(\boldsymbol{\beta}_0^T \boldsymbol{Z})$. As a result, $M(t)$ is no longer a martingale for the overall study population, and equation (2.2) fails to provide a valid estimate for $\boldsymbol{\beta}_0$.

To construct an appropriate estimating equation for $\boldsymbol{\beta_0}$, we use the fact that $M(t)$ remains a martingale for the complier group. Thus, we can show that $\boldsymbol{\mu}_c(\boldsymbol{\beta}_0) = 0$ under model (2.1), where $\boldsymbol{s}_c^{(j)}(\boldsymbol{\beta}, s) = E(Y(s) \mathbf{Z}^{\otimes j} e^{\boldsymbol{\beta}^T \mathbf{Z}} | D_1 > D_0)$ $(j = 0, 1, 2)$ and

$$\boldsymbol{\mu}_c(\boldsymbol{\beta}) = E\left[\int_0^\infty \left\{ \boldsymbol{Z} - \frac{\mathbf{s_c^{(1)}}(\boldsymbol{\beta}, \mathbf{s})}{s_c^{(0)}(\boldsymbol{\beta}, s)} \right\} dM(s) \middle| D_1 > D_0 \right].$$

However, $\boldsymbol{\mu}_c(\boldsymbol{\beta})$ cannot be used directly to estimate $\boldsymbol{\beta}_0$ because the latent complier group, $\{D_1 > D_0\}$, is not observed. To address this difficulty, we adopt the strategy of Abadie (2003), who established a simple link between the unconditional moment of the observed data and the conditional moment of the data within the complier group. A simple weighting approach may be employed

to identify the regression parameters associated with the complier group. More specifically, let $g(\cdot)$ be a measurable real function of $(T, D, \boldsymbol{X}, C)$, such that $E|g(T, D, \boldsymbol{X}, C)| < \infty$. Under assumptions (A1)–(A4), and given that C is independent of V given $\boldsymbol{X}$, an application of Theorem 3.1 of Abadie (2003) immediately implies that

$$E\{g(T, D, \boldsymbol{X}, C)|D_1 > D_0\} = \frac{1}{\Pr(D_1 > D_0)} E\{\kappa \cdot g(T, D, \boldsymbol{X}, C)\}, \tag{2.3}$$

where

$$\kappa = 1 - \frac{D(1-V)}{\Pr(V=0|\boldsymbol{X}, C)} - \frac{(1-D)V}{\Pr(V=1|\boldsymbol{X}, C)} = 1 - \frac{D(1-V)}{\Pr(V=0|\boldsymbol{X})} - \frac{(1-D)V}{\Pr(V=1|\boldsymbol{X})}. \tag{2.4}$$

This result suggests that a weighting scheme involving κ can lead to the identification of moment-type statistics for compliers. Note that κ can take both positive and negative values. This differs from standard weighting procedures based on a probability weighting, where the weights are always positive because the probabilities are nonnegative. This creates nonstandard computational challenges, which are discussed further below.

Using (2.3), we obtain the following key results for deriving an estimating equation for $\boldsymbol{\beta}_0$:

$$\boldsymbol{\mu}_c(\boldsymbol{\beta}) = \frac{1}{Pr(D_1 > D_0)} E\left[\kappa \int_0^\infty \left\{\boldsymbol{Z} - \frac{\mathbf{s_c^{(1)}}(\boldsymbol{\beta}, \mathbf{s})}{s_c^{(0)}(\boldsymbol{\beta}, s)}\right\} dM(s)\right],$$

where

$$\boldsymbol{s}_c^{(j)}(\boldsymbol{\beta}, s) = \frac{E(\kappa Y(s) \mathbf{Z}^{\otimes j} e^{\boldsymbol{\beta}^T \mathbf{Z}})}{\Pr(D_1 > D_0)}, \quad j = 0, 1, 2.$$

Suppose κ_i is known for each subject i. One may construct a weighted estimating equation for $\boldsymbol{\beta}_0$, $\boldsymbol{U}_{n,\kappa}(\boldsymbol{\beta}) = 0$, where

$$\boldsymbol{U}_{n,\kappa}(\boldsymbol{\beta}) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \int_0^\infty \kappa_i \left(\mathbf{Z}_i - \left\{\frac{\boldsymbol{S}_{n,\kappa}^{(1)}(\boldsymbol{\beta}, s)}{S_{n,\kappa}^{(0)}(\boldsymbol{\beta}, s)}\right\}\right) dN_i(s),$$

with $\boldsymbol{S}_{n,\kappa}^{(j)}(\boldsymbol{\beta}, s) = \sum_{l=1}^n \kappa_l Y_l(s) \mathbf{Z}_l^{\otimes j} e^{\boldsymbol{\beta}^T \mathbf{Z}_l}$. Note that $\boldsymbol{U}_{n,\kappa}(\boldsymbol{\beta})$ remains the same if $dN_i(s)$ is replaced by $dM_i(s)$; hence, $\boldsymbol{U}_{n,\kappa}(\boldsymbol{\beta})$ is proportional to an empirical counterpart of $\boldsymbol{\mu}_c(\boldsymbol{\beta})$. This justifies using $\boldsymbol{U}_{n,\kappa}(\boldsymbol{\beta})$ to construct the estimating equation for $\boldsymbol{\beta}_0$.

In general, κ_i is known a priori, for example, from external information. In

practice, we propose estimating κ_i by imposing additional modeling assumptions for $\Pr(V = 1|\boldsymbol{X})$. Specifically, we assume a logistic regression model for V:

$$P(V = 1|\boldsymbol{X}) \equiv \psi(\boldsymbol{\alpha}_0, \boldsymbol{X}) = \frac{\exp(\alpha_{01} + \boldsymbol{\alpha}_{02}^T \boldsymbol{X})}{1 + \exp(\alpha_{01} + \boldsymbol{\alpha}_{02}^T \boldsymbol{X})}, \tag{2.5}$$

with $\boldsymbol{\alpha}_0 = (\alpha_{01}, \boldsymbol{\alpha}_{02}^T)^T$. Let $\hat{\boldsymbol{\alpha}}$ be the maximum likelihood estimator of $\boldsymbol{\alpha}_0$ (Gourieroux and Monfort (1981); Agresti (2013)), and define

$$\hat{\kappa}_i = 1 - \frac{D_i(1 - V_i)}{1 - \psi(\hat{\boldsymbol{\alpha}}, \boldsymbol{X}_i)} - \frac{(1 - D_i)V_i}{\psi(\hat{\boldsymbol{\alpha}}, \boldsymbol{X}_i)}. \tag{2.6}$$

Replacing κ_i in $\boldsymbol{U}_{n,\kappa}(\boldsymbol{\beta})$ with $\hat{\kappa}_i$ leads to the proposed estimating equation:

$$\boldsymbol{U}_{n,\hat{\kappa}}(\boldsymbol{\beta}) = 0, \tag{2.7}$$

where

$$\boldsymbol{U}_{n,\hat{\kappa}}(\boldsymbol{\beta}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_0^\infty \hat{\kappa}_i \left(\mathbf{Z}_i - \left\{ \frac{\boldsymbol{S}_{n,\hat{\kappa}}^{(1)}(\boldsymbol{\beta}, s)}{S_{n,\hat{\kappa}}^{(0)}(\boldsymbol{\beta}, s)} \right\} \right) dN_i(s). \tag{2.8}$$

Denote the solution to equation (2.7) by $\hat{\boldsymbol{\beta}}$. The computational algorithm for obtaining $\hat{\boldsymbol{\beta}}$ is discussed in the next subsection, along with related issues and remedies.

2.4. The computational algorithm

The form of the proposed estimation equation (2.7) closely resembles that of a weighted Cox proportional hazards regression. However, an important distinction is that $\hat{\kappa}_i$ in (2.7) can take negative values. As a result, $\boldsymbol{U}_{n,\hat{\kappa}}(\boldsymbol{\beta})$ can have a highly irregular surface, with multiple zero-crossings. To address this complication, we propose locating $\hat{\boldsymbol{\beta}}$ by maximizing a properly designed objective function. Specifically, instead of solving $\boldsymbol{U}_{n,\hat{\kappa}}(\boldsymbol{\beta}) = 0$ directly, we propose obtaining $\hat{\boldsymbol{\beta}}$ as the maximizer of the following objective function:

$$\bar{C}_{n,\hat{\kappa}}(\boldsymbol{\beta}) = \frac{1}{n} \sum_{i=1}^{n} \hat{\kappa}_i \delta_i \left[\boldsymbol{\beta}^T \mathbf{Z}_i - \log\{\tilde{S}_{n,\hat{\kappa}}^{(0)}(\boldsymbol{\beta}, W_i)\} \right], \tag{2.9}$$

where $\tilde{S}_{n,\hat{\kappa}}^{(0)}(\boldsymbol{\beta}, t) = \max(S_{n,\hat{\kappa}}^{(0)}(\boldsymbol{\beta}, t), \nu)$, and ν is a prespecified small positive value. The justification for doing so is that $\partial \bar{C}_{n,\hat{\kappa}}(\boldsymbol{\beta})/\partial \boldsymbol{\beta}^T$ is nearly the same as $n^{-1/2} \boldsymbol{U}_{n,\hat{\kappa}}(\boldsymbol{\beta})$, because ν can be arbitrarily small. Truncating $S_{n,\hat{\kappa}}^{(0)}(\boldsymbol{\beta}, t)$ below by ν ensures the positiveness of the resulting quantity. In theory, the asymptotic

limit of $S^{(0)}_{n,\hat{\kappa}}(\boldsymbol{\beta}, t)$ is strictly positive under mild regularity conditions. Therefore, such a truncation should have negligible impact on the finite-sample performance of $\hat{\boldsymbol{\beta}}$ when n is reasonably large. In our numerical studies, we choose $\nu = 10^{-4}$.

The procedure for obtaining $\hat{\boldsymbol{\beta}}$ is as follows:

1. Fit the logistic regression model (2.5) to $\{(V_i, \boldsymbol{X}_i)\}_{i=1}^n$, and obtain $\hat{\boldsymbol{\alpha}}$.
2. Calculate $\hat{\kappa}_i$ using formula (2.6).
3. Find the maximizer of the objective function $\bar{C}_{n,\hat{\kappa}}(\boldsymbol{\beta})$ in (2.9) using an optimization routine, such as the *optim()* function in R (R Core Team (2017)).

2.5. A modified weighting scheme

In principle, the objective function $\bar{C}_{n,\hat{\kappa}}(\boldsymbol{\beta})$ approaches a limit that is concave, and standard optimization routines are expected to work well when the sample size is large. However, the presence of negative weights κ_i can lead to highly irregular surfaces for $\bar{C}_{n,\hat{\kappa}}(\boldsymbol{\beta})$ and $U_{n,\hat{\kappa}}(\boldsymbol{\beta})$ (see the figures in Section S4 of the Supplementary Material) and result in numerical instability in the estimate of $\boldsymbol{\beta}_0$. To address this problem, we propose a modified weighting scheme that avoids negative weights and obtains $\hat{\boldsymbol{\beta}}$ using standard computational routines for a weighted proportional hazards regression, such as the *coxph()* function in R (Therneau (2015)).

Let $\boldsymbol{U} = (W, \delta, D, \boldsymbol{X})$. We define the modified weight by projecting the original weight κ, as follows:

$$\kappa_v = E(\kappa|\boldsymbol{U}) = 1 - \frac{D(1 - v_0(\boldsymbol{U}))}{P(V = 0|\boldsymbol{X})} - \frac{(1 - D)v_0(\boldsymbol{U})}{P(V = 1|\boldsymbol{X})}, \tag{2.10}$$

where $v_0(\boldsymbol{U}) = E(V|\boldsymbol{U}) = P(V = 1|W, \delta, D, \boldsymbol{X})$. Adapting the arguments of Abadie, Angrist and Imbens (2002), we can show that $\kappa_v = P(D_1 > D_0|\boldsymbol{U})$, and that κ_v plays the same role as κ in equation (2.3) (see Section S1 of the Supplementary Material). This result indicates that κ_v is a probability; thus, it is always nonnegative and can be regarded as a proper weight. Adopting the weighting scheme based on κ_v avoids the potential numerical issues posed by using κ. We propose estimating κ_v as follows:

1. Stratify the data by the censoring and treatment status: $\{(\delta = c, D = d)\}$, $c = 0, 1$, $d = 0, 1$.
2. Within each stratum, fit a nonparametric or parametric regression model for

V, given covariates $(W, \boldsymbol{X})$. This provides an estimate for $v_0(\boldsymbol{U})$, denoted by $\hat{v}(\boldsymbol{U})$.

3. Calculate the estimated κ_v as

$$\hat{\kappa}_v = 1 - \frac{D(1-\hat{v}(\boldsymbol{U}))}{1-\psi(\hat{\boldsymbol{\alpha}}, \boldsymbol{X}_i)} - \frac{(1-D)\hat{v}(\boldsymbol{U})}{\psi(\hat{\boldsymbol{\alpha}}, \boldsymbol{X}_i)}.$$

In Step 2, we can employ a nonparametric power series (NPPS) regression or a logistic regression for V, given $(W, \boldsymbol{X})$. However, our extensive numerical findings (including those reported and not reported in Section 3), show that a second-order logistic regression model with an interaction between W and $\boldsymbol{X}$ outperforms approaches that estimate $\hat{\kappa}_v$ using an NPPS or a first-order logistic regression. When the dimension of $\boldsymbol{X}$ is large, we recommend using a penalized logistic regression to obtain a reasonable estimate for $v_0(\boldsymbol{U})$. Note that with finite sample sizes, the resulting estimator $\hat{\kappa}_v$ may be negative or greater than one. To circumvent the undesirable numerical properties associated with negative weights, we propose a slightly modified weight, $\hat{\kappa}_{v,tr}$, that truncates $\hat{\kappa}_v$ such that its value lies strictly in an interval $\mathcal{I} \subset (0,1)$, say $[0.01, 0.99]$. Because the true weight κ_v is between zero and one and we can let $\mathcal{I}$ be arbitrarily close to $(0,1)$, there should be negligible asymptotic bias induced by such a truncation. Then, using $\hat{\kappa}_{v,tr}$ in place of κ in (2.7), we easily obtain $\hat{\boldsymbol{\beta}}$ from the R function *coxph()*, with the weight argument properly specified. In Section 3, we examine the performance of the proposed estimator for different choices of weights.

2.6. Large-sample results

We assume the following regularity conditions:

(C1) : The parameter space for $\boldsymbol{\beta}$, $\mathcal{B}$, is compact.

(C2) : $\|\boldsymbol{Z}\| < \infty$ and $|\kappa| < \infty$.

(C3) : $s_c^{(0)}(\boldsymbol{\beta}, t)$ is bounded away from zero uniformly in $\boldsymbol{\beta}$ and t.

(C4) : $\boldsymbol{\Sigma}_0 > 0$, where $\boldsymbol{\Sigma}_0$ is defined in (B.1) of Section S2 of the Supplementary Material.

(C5) : $\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}_0 \to_{a.s.} 0$.

(C6) : There exists an influence function $\boldsymbol{I}_{\boldsymbol{\alpha}}(\cdot)$, such that

$$\left\| n^{1/2}(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha_0}) - n^{-1/2} \sum_{i=1}^{n} I_{\boldsymbol{\alpha}}(\boldsymbol{\alpha_0}, \boldsymbol{O}_i) \right\| = o(1), \ a.s.$$

We establish the consistency and asymptotic normality for the proposed estimator in the following theorems.

Theorem 1. *(Consistency) Under conditions* (C1)–(C5), $\hat{\boldsymbol{\beta}} \to_{a.s.} \boldsymbol{\beta_0}$.

Theorem 2. *(Asymptotic normality) Under conditions* (C1)–(C6), $n^{1/2}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta_0}) \to_d N(0, \Omega)$, *where* Ω *is defined in Section* S2 *of the Supplementary Material (see equation* (B.12)*).*

The regularity conditions (C1)–(C2) impose the boundedness of the parameter space and covariates, which are mild and often met in practice. The boundedness of κ is satisfied when $\Pr(V = 0|\boldsymbol{X})$ is always away from zero and one. Conditions (C3)–(C4) are standard assumptions for Cox proportional hazard regression methods. For example, condition (C4) ensures the identifiability of $\boldsymbol{\beta}_0$. Conditions (C5)–(C6) depict reasonable requirements on the estimator of $\boldsymbol{\alpha}_0$, such as consistency and asymptotic i.i.d. sum representation. Detailed proofs of Theorems 1 and 2 are provided in Appendix B.

Note that the theoretical properties, including consistency and root-n asymptotic normality, can also be established for the proposed estimator based on the modified weighting scheme presented in Section 2.5. The theoretical arguments follow similar lines to the proofs of Theorems 1–2. The main distinction lies in the derivation of the influence function, which needs to account for the additional variability induced by $\hat{v}(\boldsymbol{U})$. The detailed asymptotic results for the estimator with the modified weight are omitted here, but are available upon request.

2.7. Variance estimation

In the proof of Theorem 2, we derive a closed form for the asymptotic variance of $n^{1/2}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta_0})$; see equation (B.12) of Section S2 of the Supplementary Material. A consistent variance estimator for $\hat{\boldsymbol{\beta}}$ (with weight $\hat{\kappa}$) can be obtained using $\hat{\boldsymbol{\Omega}}/n$, where $\hat{\boldsymbol{\Omega}}$ is $\boldsymbol{\Omega}$, with unknown quantities replaced by their empirical counterparts or consistent estimators.

An alternative approach to estimating the asymptotic variance of $\hat{\boldsymbol{\beta}}$ (with weight $\hat{\kappa}$ or $\hat{\kappa}_v$) is to use bootstrapping: Step 1: Resample n observations from the original data set with replacement, $\{\boldsymbol{O}_i^b\}_{i=1}^n$, and add some small amount of

noise (e.g., $N(0, 10^{-10})$) to avoid the presence of ties in the resampled data; Step 2: Calculate $\hat{\boldsymbol{\beta}}^b$ based on $\{\boldsymbol{O}_i^b\}_{i=1}^n$, with weights as described in Section 2 (i.e., $\hat{\kappa}$, $\hat{\kappa}_v$, or $\hat{\kappa}_{v,tr}$); Step 3: Repeat steps 1–2, for $b = 1, \ldots, B$; Step 4: Estimate the asymptotic variance of $\hat{\boldsymbol{\beta}}$ using the empirical variance of $\{\hat{\boldsymbol{\beta}}^b\}_{b=1}^B$.

In the bootstrapping procedure, the computations in Step 2 may fail to converge. In such a case, we carry out Step 3 until there are B convergent estimates. In addition, repeated resampling may occasionally produce outlier estimates that artificially inflate the empirical variance in Step 4. If this occurs, we may estimate the standard deviation of $\hat{\boldsymbol{\beta}}$ using the median absolute deviation, namely, $1.4826 \times MAD$, where $MAD = median(|\hat{\boldsymbol{\beta}}_b - median(\hat{\boldsymbol{\beta}}_b)|)$ (Rousseeuw and Croux (1993)). This alternative approach performs quite well, based on our numerical results.

3. Simulation Study

We conduct extensive simulations to assess the performance of the proposed estimators. We create data under assumptions (A1) to (A4), as follows:

1. Generate $\boldsymbol{X}$ from a bounded distribution.
2. Generate the latent group membership (i.e., *complier*, *always-taker*, or *never-taker*) from a multinomial distribution.
3. Generate $V \sim Bernoulli(P(V = 1|\boldsymbol{X}))$, where $P(V = 1|\boldsymbol{X}) = \exp(\alpha_{01} + \alpha_{02}^T\boldsymbol{X})/(1 + \exp(\alpha_{01} + \alpha_{02}^T\boldsymbol{X}))$, and determine D from V and the latent group membership.
4. For compliers, generate potential survival times $T_{00} = T_{10} = \exp(-\boldsymbol{\beta}_x^T\boldsymbol{X} + \epsilon)$ and $T_{01} = T_{11} = \exp(-\boldsymbol{\beta}_x^T\boldsymbol{X} - \boldsymbol{\beta}_d + \epsilon)$, where ϵ follows the extreme value distribution.
5. For noncompliers, generate T_{00} or T_{01}, given $\boldsymbol{X}$, possibly from a nonCox regression model, and let $T_{00} = T_{10}$ and $T_{01} = T_{11}$.
6. Set $T = T_{vd}$ and draw independent censoring times $C \sim Exponential(0.5)$.

We consider two basic data-generation scenarios with a single covariate X. In scenario 1, for compliers, survival times are generated using $\boldsymbol{\beta_0} = (\beta_d, \beta_x) = (-0.5, -0.2)$. Survival times for noncompliers in scenario 1 are generated according to $T = \exp(-0.02X + \epsilon_1)$, where $\epsilon_1 \sim N(0, 0.01)$ (i.e., no treatment effect). In scenario 2, compliers' survival times are generated using $\boldsymbol{\beta_0} = (-0.3, 0.05)$. Noncompliers' survival times are also generated from a Cox proportional hazards

regression model, where $T = \exp(0.5D - 0.05X + \epsilon_2)$, and ϵ_2 follows the extreme value distribution.

For each scenario, we consider eight cases with different combinations of rate of compliers, sample size, and covariate distribution. Specifically, in cases 1–4, X follows a $Uniform(-1, 1)$ distribution, and in cases 5–8, X follows a $Bernoulli(0.5)$ distribution. The sample size n = 1,000 in cases 1, 2, 5, and 6, and n = 4,000 in cases 3, 4, 7, and 8. The probability of compliers is equal to 1/3 in cases 1, 3, 5, and 7, and 2/3 in cases 2, 4, 6, and 8.

We compare several estimation methods: (1) a *benchmark* estimate, based only on compliers (unknown in a real-data analysis); (2) a *naive* estimate, which assumes the entire sample follows the same Cox model; (3) the proposed $\hat{\kappa}$-weighted estimate; (4) the modified $\hat{\kappa}_v$-weighted estimate; and (5) the estimate based on the truncated modified weights $\hat{\kappa}_{v,tr}$. Hereafter, we refer to these methods as *complier*, *naive*, κ, κ_v, and $\kappa_{v,tr}$.

To estimate $\hat{\kappa}_v$ and $\hat{\kappa}_{v,tr}$, we estimate $v_0(U) = P(V = 1|W, X, D, \delta)$ using the method described in Section 2.5 with a second-order logistic regression, including the interaction between W and X for each of the four partitions by the censoring and treatment status. Note that the estimations using $\hat{\kappa}$ and $\hat{\kappa}_v$ follow the algorithms and caveats laid out in Sections 2.4 and 2.5, where $\hat{\boldsymbol{\beta}}$ is estimated by maximizing the objective function in (2.9). Specifically, we use the `R` function `optim` with the `BFGS` method option (R Core Team (2017)), considering three different starting values (based on the naive estimate, ±0.5) to solve the maximization problem. For the method $\kappa_{v,tr}$, we use the `R` function `coxph` to implement the proposed estimation, as described in Section 2.5. For each method under comparison, we check whether the resulting estimate solves the proposed estimating equation within some tolerance (e.g., 0.05). We record a failure to converge if such an estimate cannot be produced.

The top row of Figure 1 shows the convergence rates for the three proposed estimators. In scenario 1, the convergence rates of $\hat{\kappa}$ and $\hat{\kappa}_v$ are both close to 100% across the eight cases considered. In scenario 2, the convergence rates vary considerably but, in general, increase with n and the proportion of compliers $P(D_1 > D_0)$. Anecdotal examination reveals that the objective and estimating function surfaces for this scenario can be highly irregular. In contrast, because the $\hat{\kappa}_{v,tr}$-weights are always positive, the resultant surfaces are smooth and the convergence rates are always 100%. The second row of Figure 1 shows the empirical bias by comparing the treatment and covariate parameter estimates with the truth. In general, the naive parameter estimators demonstrate large empirical

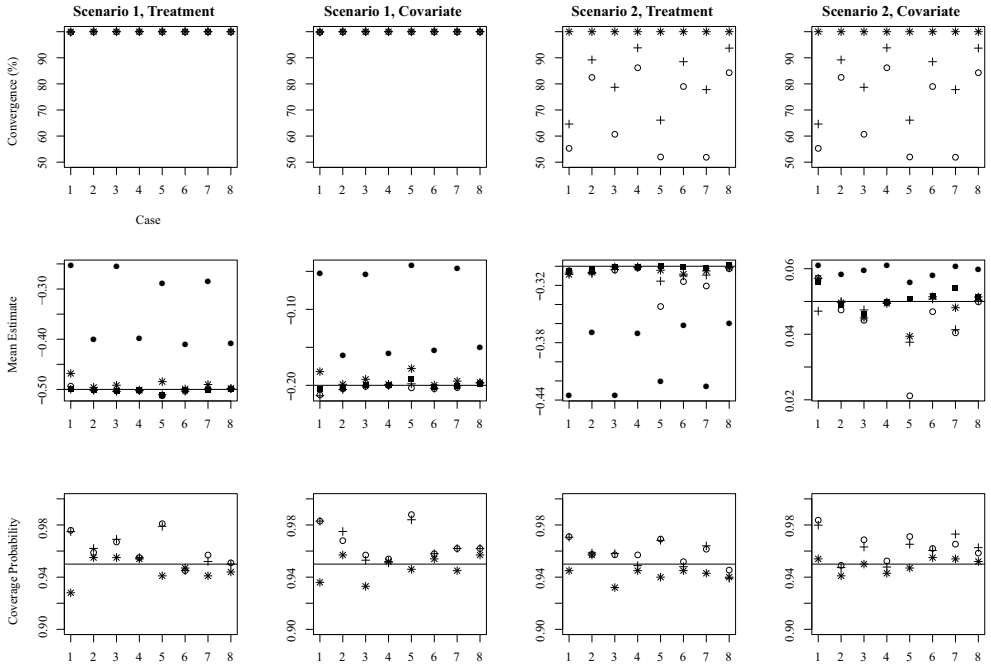

Figure 1. Simulation results: Convergence rates, mean estimates, and empirical coverage probabilities of 95% confidence intervals: Complier (■); Naive (•); κ (∘); κ_v (+); κ_{vtr} (∗).

bias, and the proposed methods reduce the bias considerably.

In Figure 2, we compare various standard error (SE) estimates against the empirical standard deviations (SD) of the proposed estimators. We denote the mean and median estimated SE based on the analytic variance estimation by *Mean SE* and *Median SE*, respectively, and denote the mean and median estimated SE based on the bootstrapping variance estimation by *Mean Bootstrap SE* and *Median Bootstrap SE*, respectively. The empirical standard deviation (SD) is denoted by *Empirical*. For the method κ, we evaluate both the analytic variance estimation and the bootstrapping-based variance estimation. The results show that both *Mean Bootstrap SE* and *Median Bootstrap SE* are close to the corresponding empirical SDs in Scenarios 1 and 2. With regard to the analytic variance estimation, the *Median SE*s are in good agreement with the empirical SDs, while in Scenario 2, many *Mean SE*s depart considerably from the empirical SDs. The latter phenomenon may reflect the unstable performance of the κ-weighted estimator in Scenario 2, which is consistent with the lower con-

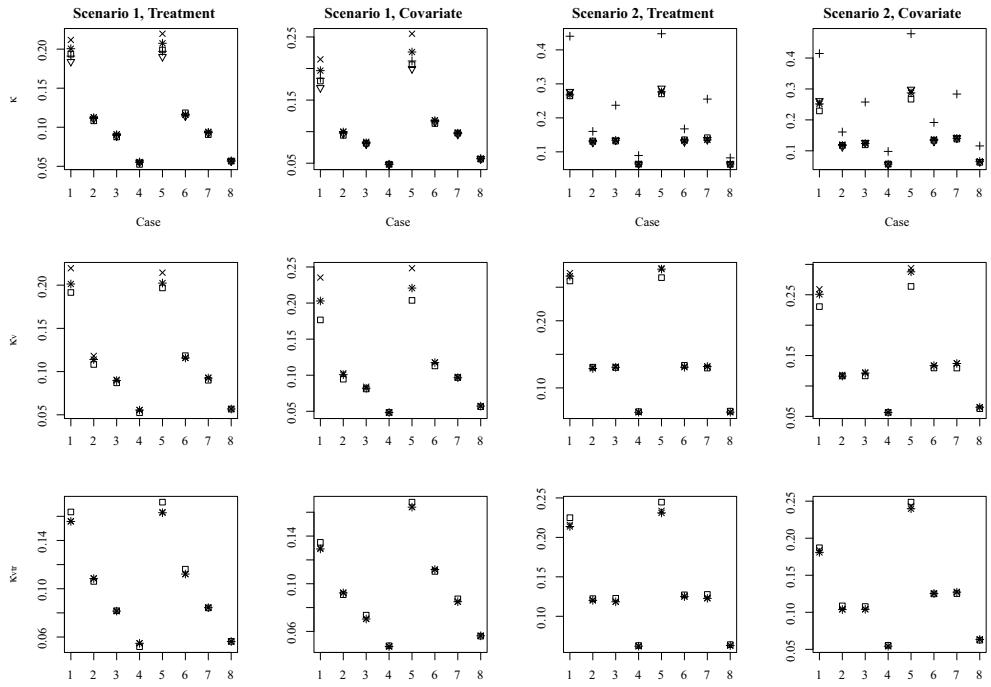

Figure 2. Simulation results: the estimated standard errors and empirical standard deviations of $\hat{\kappa}$, $\hat{\kappa}_v$, $\hat{\kappa}_{v,tr}$ weighted estimators: Empirical (□); Mean SE (+); Median SE ($\triangledown$); Mean Bootstrap SE (×); Median Bootstrap SE ($\ast$).

vergence rates of the method κ in Scenario 2. For the methods κ_v and $\kappa_{v,tr}$, we examine the bootstrapping-based variance estimation only. Two extreme outliers are removed from the calculation of the mean bootstrap SE for the covariate coefficient estimate based on method κ_v in Case 5 of Scenario 1. We observe fairly small discrepancies between *Mean Bootstrap SE*s, *Median Bootstrap SE*s, and the empirical SDs in Scenarios 1 and 2, while the method $\kappa_{v,tr}$ shows slightly better performance.

The bottom row of Figure 1 shows the empirical coverage probabilities of the 95% confidence intervals, constructed as $\hat{\boldsymbol{\beta}} \pm z_{0.975} \times \hat{SE}(\hat{\boldsymbol{\beta}})$, where $\hat{SE}(\hat{\boldsymbol{\beta}})$ denotes the bootstrapping-based SE. The coverage probabilities associated with the method $\kappa_{v,tr}$ are fairly close to the nominal 95% level, decreasing to 93% in a few cases. The methods κ and κ_v have similar and generally more conservative performance in terms of their empirical coverage probabilities. Note that the results presented for these two methods are based only on simulations that produce converged estimates. In Scenario 2, where the convergence rates of κ

and κ_v can be considerably below one, the results in Figure 1 may over-represent the performance of these two methods.

Based on all simulations, the method $\kappa_{v,tr}$ exhibits the best performance of the weighting methods, including good coverage probabilities, low bias, and reliable convergence.

4. Colon Cancer Screening with Flexible Sigmoidoscopy

The PLCO Cancer Screening Trial is a multi-center, two-armed randomized trial, sponsored by the National Cancer Institute, of screening tests for prostate, lung, colorectal, and ovarian cancers. Ten centers across the United States recruited approximately 155,000 participants between November 1993 and July 2001. Data were collected until December 31, 2009. One objective of the trial is to evaluate the effectiveness of screening for colorectal cancer using flexible sigmoidoscopy, in terms of mortality, versus that of usual-care methods. Prorok et al. (2000) reported further details about this trial.

The original data consist of 154,897 individuals aged 55 to 74 years. They were assigned randomly to either the usual-care (control, N = 77,453) group or the screening with flexible sigmoidoscopy (intervention, N = 77,444) group. For the intervention group, subjects were offered the screening at the baseline and three or five years later. After randomization, data were discarded for 187 participants who dropped out, died, were diagnosed with cancer, or had an organ removed before the first screening visit, as were data on four participants with no follow-ups. Thus, we considered 154,706 individuals in our analyses.

Table 1 presents the descriptive statistics for the baseline characteristics of the participants, stratified by the screening assignment (i.e., $V = 0, V = 1$) and the actual screening status (i.e., $D = 0$, $D = 1$). We also consider risk factors, including age (in years), gender, family history of any cancer, family history of colorectal cancer, colorectal polyps, and diabetes. We apply t-tests or chi-squared tests to check the balance of these observed risk factors between the groups, determined by the screen assignment or the actual screening status. Based on the p-values reported in Table 1, there is strong evidence that this trial was well randomized, with small and nonsignificant associations between the screening assignment and the risk factors. However, most of these risk factors are unbalanced by the actual screening status. The summary statistics in Table 1 suggest that older male participants with a family history of any cancer, a family history of colorectal cancer, or diabetes were more likely to take the colon cancer screen-

Table 1. Characteristics of the study participants.

Characteristics	Control (V = 0) N = 77,449		Intervention (V = 1) N = 77,257			Not Screened (D = 0) N = 90,056		Screened (D = 1) N = 64,650		
	Number of Participants (%)		*Number of Participants* (%)		p-value	*Number of Participants* (%)		*Number of Participants* (%)		p-value
Age ⋆										
	62.60(5.37)		62.59(5.39)		0.8274	62.65(5.39)		62.52(5.33)		<0.0001
Age Level										
55-59 yr	25,838	(33.36)	25,789	(33.38)		29,902	(33.20)	21,725	(33.60)	
60-64 yr	23,767	(30.69)	23,736	(30.72)		27,451	(30.48)	20,052	(31.02)	
65-69 yr	17,473	(22.56)	17,402	(22.52)		20,352	(22.60)	14,523	(22.46)	
70-74 yr	10,371	(13.39)	10,330	(13.37)	0.9967	12,351	(13.71)	8,350	(12.92)	<0.0001
Sex										
Male	38,340	(49.50)	38,229	(49.48)		43,529	(48.34)	33,040	(51.11)	
Female	39,109	(50.50)	39,028	(50.52)	0.9393	46,527	(51.66)	31,610	(48.89)	<0.0001
Family History of Any Cancer										
No	32,742	(42.28)	33,327	(43.14)		37,798	(41.97)	28,271	(43.73)	
Yes	41,305	(53.33)	41,971	(54.33)	0.8735§	47,137	(52.34)	36,139	(55.90)	0.0190§
Unknown	3,402	(4.39)	1,959	(2.54)	<0.0001	5,121	(5.69)	240	(0.37)	<0.0001
Family History of Colorectal Cancer										
No	64,504	(83.29)	65,203	(84.40)		73,997	(82.17)	55,710	(86.17)	
Yes †	7,320	(9.45)	7,627	(9.87)	0.0809§	8,331	(9.25)	6,616	(10.23)	0.0022§
Possibly ‡/Unknown	5,625	(7.26)	4,427	(5.73)	<0.0001	7,728	(8.58)	2,324	(3.59)	<0.0001
Colorectal Polyps										
No	68,690	(88.69)	69,910	(90.49)		78,705	(87.40)	59,895	(92.65)	
Yes	4,947	(6.39)	5,185	(6.71)	0.1565§	5,739	(6.37)	4,393	(6.80)	0.7865§
Unknown	3,812	(4.92)	2,162	(2.80)	<0.0001	5,612	(6.23)	362	(0.56)	<0.0001
Diabetes										
No	68,028	(87.84)	6,9371	(89.79)		77,773	(86.36)	59,626	(92.23)	
Yes	5,699	(7.36)	5,810	(7.52)	0.9971§	6,776	(7.52)	4,733	(7.32)	<0.0001§
Unknown	3,722	(4.81)	2,076	(2.69)	<0.0001	5,507	(6.12)	291	(0.45)	<0.0001

⋆ denotes a continuous variable. Mean and standard deviation are reported.

† indicates colorectal cancer family history in immediate family member.

‡ indicates colorectal cancer family history in relatives or unclear cancer type.

§ indicates p-value without considering missing category.

ing when it was assigned. Thus, there is some evidence to suggest that the study participants' post-randomization care selections and their potential survival outcomes are dependent. Hence, the traditional ITT or "as-treated" analyses may be problematic for evaluating the causal effect of flexible sigmoidoscopy screening on colorectal cancer mortality.

To address this issue, we employ the proposed IV methods, with the survival outcome of interest (T) defined as the time from trial entry (i.e., randomization) to death from colorectal cancer (in years), and the IV chosen as the screening assignment (V). In our data set, 351 and 249 colorectal cancer deaths were observed in the control group (n = 77,098) and the intervention group (n = 77,098), respectively; 409 and 191 colorectal cancer deaths were observed in the

group without screening (n = 89,647) and the group with screening (n = 64,459), respectively. In our analysis, deaths due to other causes are competing risks for death from colon cancer. As discussed in Section S3 of the Supplementary Material, naively treating such competing events as censoring events leads to a valid IV proportional hazards analysis of the cause-specific hazard function for colon cancer death. Our IV is justified as follows: (i) the screening assignment is highly informative of the actual screening status (D) (i.e., screened vs. not screened); (ii) the screen assignment is random and, hence, is expected to be independent of unmeasured confounders (given the observed risk factors); and (iii) it is reasonable to expect that the impact of the screening assignment on the survival outcome is only through its influence on the actual screening status.

We first assess the unadjusted causal effect of the flexible sigmoidoscopy screening by fitting model (2.1) without $\boldsymbol{X}$ to the full data set and stratifying the analysis by each risk factor. For comparison purposes, we also perform the "as-treated" counterparts (i.e., fitting a Cox model for T, with D as the only covariate) and the ITT counterparts (i.e., fitting a Cox model for T, with V as the only covariate) of these IV analyses. For the IV analyses, we implement the three methods κ, κ_v, and $\kappa_{v,tr}$ in the same way as in our simulation studies (see Section 3), except that we use a simple logistic regression model stratified by (δ, D) to estimate $v_0(U)$ in (2.10). Table 2 reports the parameter estimates and the associated standard errors. For the IV methods, we present the bootstrap-based standard errors. Table 2 also reports the rate of compliance in the intervention group (i.e., the proportion of screened participants in the intervention group), p_c.

From Table 2, we observe that the estimates for the causal effect of screening are very similar among the three IV methods. The conclusions on the survival impact of screening are, in general, consistent across the IV analyses, as-treated analyses, and ITT analyses, except for the sub-cohort with a baseline age between 70 and 74 years, and the sub-cohort with diabetes. In these two cases, significant benefits of screening are suggested by the as-treated analyses, but not by the ITT or IV analyses. Such discrepancies may be explained by the relatively high noncompliance rates ($\approx$ 19%) observed in the intervention group. That is, study participants who refused assigned screening are likely to be less health-conscious, which may be associated with worse potential survival outcomes. When the nonscreened group includes a large proportion of such participants, the as-treated analyses would tend to over-estimate the benefit of screening, because they ignore the survival impact of the unmeasured confounder

Table 2. Analyses for the unadjusted screening effect based on the whole data set or stratified by each risk factor.

Data	N	p_c	As-Treated	ITT	κ	κ_v	$\kappa_{v,tr}$
(Subgroup)			Parameter Estimates (Standard Errors)				
Total	154,706	0.84	-0.442*	-0.343*	-0.427*	-0.427*	-0.427*
			(0.088)	(0.083)	(0.099)	(0.097)	(0.101)
Age Level							
55-59 yr	51,627	0.84	-0.572*	-0.380*	-0.496*	-0.496*	-0.496*
			(0.198)	(0.184)	(0.229)	(0.240)	(0.248)
60-64 yr	47,503	0.84	-0.313	-0.130	-0.169	-0.169	-0.169
			(0.160)	(0.153)	(0.201)	(0.198)	(0.193)
65-69 yr	34,875	0.83	-0.475*	-0.590*	-0.654*	-0.655*	-0.655*
			(0.164)	(0.158)	(0.178)	(0.164)	(0.182)
70-74 yr	20,701	0.81	-0.420*	-0.264	-0.351	-0.350	-0.350
			(0.188)	(0.176)	(0.228)	(0.213)	(0.218)
Sex							
Male	76,569	0.86	-0.549*	-0.445*	-0.536*	-0.536*	-0.536*
			(0.115)	(0.109)	(0.124)	(0.123)	(0.123)
Female	78,137	0.81	-0.319*	-0.200	-0.262	-0.262	-0.262
			(0.135)	(0.128)	(0.156)	(0.172)	(0.166)
Family History of Any Cancer							
Yes	83,276	0.86	-0.237*	-0.258*	-0.294*	-0.294*	-0.294*
			(0.114)	(0.111)	(0.120)	(0.124)	(0.127)
No	66,069	0.85	-0.704*	-0.492*	-0.639*	-0.639*	-0.639*
			(0.144)	(0.132)	(0.158)	(0.179)	(0.162)
Family History of Colorectal Cancer							
Yes	14,947	0.87	-0.010	-0.097	-0.105	-0.106	-0.106
			(0.241)	(0.239)	(0.251)	(0.271)	(0.254)
No	129,707	0.85	-0.457*	-0.391*	-0.469*	-0.469*	-0.469*
			(0.099)	(0.094)	(0.117)	(0.113)	(0.104)
Colorectal Polyps							
Yes	10,132	0.85	0.315	0.288	0.335	0.336	0.336
			(0.305)	(0.309)	(0.388)	(0.405)	(0.389)
No	138,600	0.86	-0.490*	-0.401*	-0.490*	-0.485*	-0.485*
			(0.093)	(0.089)	(0.111)	(0.112)	(0.110)
Diabetes							
Yes	11,509	0.81	-1.036*	-0.335	-0.606	-0.603	-0.603
			(0.311)	(0.253)	(0.451)	(0.454)	(0.438)
No	137,399	0.86	-0.355*	-0.349*	-0.404*	-0.404*	-0.404*
			(0.093)	(0.090)	(0.095)	(0.099)	(0.092)

* indicates p-value ≤ 0.05

related to health-consciousness. Therefore, in these two cases, it is more plausible to conclude that flexible sigmoidoscopy screening offers little in the sense of survival benefits for participants aged between 70 and 74 years and for participants with diabetes. Overall, the unadjusted stratified analyses support the benefit of flexible sigmoidoscopy in reducing colorectal mortality, with the greatest benefit evidenced in subpopulations with relatively low mortality risk, for example, the age group 55–59 years, and subjects without a family history of colorectal cancer.

We next evaluate the causal effect of screening, while accounting for other risk factors. Specifically, we fit model (2.1), with $\boldsymbol{X}$ capturing gender, family history of any cancer, family history of colorectal cancer, colorectal polyps, and diabetes separately for the four age groups, 55–59 years, 60–64 years, 65–69 years, and 70–74 years. Table 3 provides the summary statistics (i.e., count and percentage) of the risk factors by V and by D within each age group, along with the p-values from testing the association of the risk factors with V or D, based on the chi-squared tests. Similarly to Table 1, within each age group, the risk factors show little association with the screening assignment D, but may vary significantly between participants who were screened versus those who were not screened.

Table 4 presents the parameter estimates and the associated standard errors based on the IV methods, κ, κ_v, and $\kappa_{v,tr}$. The coefficient estimates from the as-treated analysis (i.e., a multivariate Cox model for T, given D and $\boldsymbol{X}$) and the ITT analysis (i.e., a multivariate Cox model for T, given V and $\boldsymbol{X}$) are also presented, along with the corresponding standard errors. From Table 4, we again observe relatively good agreement among the three IV estimates. The IV analyses suggest that the flexible sigmoidoscopy screening has a significant protective effect on colorectal cancer mortality in the older age groups, 65–69 years and 70–74 years, but not in the younger age groups, 55–59 years and 60–64 years, after adjusting for age, gender, family history of any cancer, family history of colorectal cancer, colorectal polyps, and diabetes.

This finding is consistent with that based on the ITT analyses, but moderately disagrees with the results from the as-treated analyses, particularly for the age groups 55–59 years and 60–64 years. The similarity between the ITT analyses and the proposed analyses might be due to the dilution effect commonly seen in screening trials, as discussed in Baker, Kramer and Prorok (2002). To understand the discrepancies between the results of the other analyses and those of the as-treated analyses, note that there is a marked imbalance in the risk factors by actual screening status for the age groups 55–59 years and 60–64 years compared

Table 3. Characteristics of the study participants by age subgroups.

Age Level	55-59 yr						60-64 yr					
Covariates	$V=0$	$V=1$	p-value	$D=0$	$D=1$	p-value	$V=0$	$V=1$	p-value	$D=0$	$D=1$	p-value
Gender												
Male	11,078 (46.8)	11,576 (47.6)		12,403 (46.1)	10,251 (48.7)		10,831 (49.5)	11,145 (50.0)		12,074 (48.4)	9,902 (51.4)	
Female	12,595 (53.2)	12,724 (52.4)	0.0661	14,530 (53.9)	10,789 (51.3)	<0.0001*	11,070 (50.5)	11,164 (50.0)	0.2946	12,886 (51.6)	9,348 (48.6)	<0.0001*
Family History of Any Cancer												
No	11,148 (47.1)	11,545 (47.5)		12,793 (47.5)	9,900 (47.1)		9,979 (45.6)	10,139 (45.4)		11,427 (45.8)	8,691 (45.1)	
Yes	12,525 (52.9)	12,755 (52.5)	0.3633	14,140 (52.5)	11,140 (52.9)	0.3361	11,922 (54.4)	12,170 (54.6)	0.8138	13,533 (54.2)	10,559 (54.9)	0.1882
Family History of Colorectal Cancer												
No	21,485 (90.8)	22,000 (90.5)		24,455 (90.8)	19,030 (90.4)		19,681 (89.9)	19,923 (89.3)		22,455 (90.0)	17,149 (89.1)	
Yes	2,188 (9.2)	2,300 (9.5)	0.4118	2,478 (9.2)	2,010 (9.6)	0.1935	2,220 (10.1)	2,386 (10.7)	0.0565	2,505 (10.0)	2,101 (10.9)	0.0029*
Colorectal Polyps												
No	22,691 (95.9)	23,264 (95.7)		25,806 (95.8)	20,149 (95.8)		20,432 (93.3)	20,747 (93.0)		23,259 (93.2)	17,920 (93.1)	
Yes	982 (4.1)	1,036 (4.3)	0.5448	1,127 (4.2)	891 (4.2)	0.8029	1,469 (6.7)	1,562 (7.0)	0.2282	1,701 (6.8)	1,330 (6.9)	0.7117
Diabetes												
No	22,217 (93.8)	22,888 (94.2)		25,223 (93.7)	19,882 (94.5)		20,243 (92.4)	20,648 (92.6)		23,008 (92.2)	17,883 (92.9)	
Yes	1,456 (6.2)	1,412 (5.8)	0.1211	1,710 (6.3)	1,158 (5.5)	0.0001*	1,658 (7.6)	1,661 (7.4)	0.6308	1,952 (7.8)	1,367 (7.1)	0.0047*
	65-69 yr						70-74 yr					
	$V=0$	$V=1$	p-value	$D=0$	$D=1$	p-value	$V=0$	$V=1$	p-value	$D=0$	$D=1$	p-value
Gender												
Male	8,042 (50.1)	8,192 (50.4)		8,985 (48.7)	7,249 (52.3)		4,579 (48.1)	4,641 (48.5)		5,177 (46.3)	4,043 (51.0)	
Female	8,015 (49.9)	8,073 (49.6)	0.6203	9,483 (51.3)	6,605 (47.7)	<0.0001*	4,949 (51.9)	4,925 (51.5)	0.5368	5,997 (53.7)	3,877 (49.0)	<0.0001*
Family History of Any Cancer												
No	7,252 (45.2)	7,309 (44.9)		8,399 (45.5)	6,162 (44.5)		4,083 (42.9)	4,153 (43.4)		4,844 (43.4)	3,392 (42.8)	
Yes	8,805 (54.8)	8,956 (55.1)	0.6898	10,069 (54.5)	7,692 (55.5)	0.0754	4,083 (57.1)	4,153 (56.6)	0.4421	4,844 (56.6)	3,392 (57.2)	0.4819
Family History of Colorectal Cancer												
No	14,320 (89.2)	14,461 (88.9)		16,476 (89.2)	12,305 (88.8)		4,083 (88.4)	4,153 (88.5)		4,844 (88.6)	3,392 (88.2)	
Yes	1,737 (10.8)	1,804 (11.1)	0.4415	1,992 (10.8)	1,549 (11.2)	0.2686	1,104 (11.6)	1,102 (11.5)	0.9029	1,275 (11.4)	931 (11.8)	0.4771
Colorectal Polyps												
No	1,737 (91.5)	1,804 (91.2)		1,992 (91.5)	1,549 (91.3)		8,561 (89.9)	8,582 (89.7)		10,036 (89.8)	7,107 (89.7)	
Yes	1,360 (8.5)	1,426 (8.8)	0.3509	1,576 (8.5)	1,210 (8.7)	0.5387	967 (10.1)	984 (10.3)	0.7722	1,138 (10.2)	813 (10.3)	0.8750
Diabetes												
No	14,652 (91.2)	14,799 (91.0)		16,798 (91.0)	12,653 (91.3)		967 (90.5)	984 (89.5)		1,138 (90.1)	813 (89.8)	
Yes	1,405 (8.8)	1,466 (9.0)	0.4169	1,670 (9.0)	1,201 (8.7)	0.2506	907 (9.5)	1,007 (10.5)	0.0218*	1,110 (9.9)	804 (10.2)	0.6390

N (Row Percentage, %)

* indicates p-value ≤ 0.05

Table 4. Results of adjusted models within age subgroups.

Age Level		As-Treated		ITT		κ		κ_v		$\kappa_{v,tr}$	
(p_c)	Covariates	Point Estimates (Standard Errors)									
55-59 yr	Screening	-0.474*	(0.207)	-0.296	(0.196)	-0.373	(0.228)	-0.373	(0.242)	-0.373	(0.246)
(0.84)	Female	-0.101	(0.195)	-0.089	(0.195)	-0.003	(0.244)	-0.013	(0.246)	-0.013	(0.232)
	Family History of Any Cancer	0.204	(0.208)	0.201	(0.208)	0.468	(0.272)	0.463	(0.280)	0.465	(0.280)
	Family History of Colorectal Cancer	0.194	(0.313)	0.192	(0.313)	-0.080	(0.471)	-0.071	(0.468)	-0.073	(0.386)
	Colorectal Polyps	0.276	(0.422)	0.277	(0.422)	0.137	(0.736)	0.135	(1.725)	0.131	(1.768)
	Diabetes	0.168	(0.392)	0.179	(0.392)	0.127	(0.606)	0.125	(0.591)	0.126	(0.710)
60-64 yr	Screening	-0.333*	(0.169)	-0.184	(0.163)	-0.228	(0.197)	-0.229	(0.205)	-0.242	(0.181)
(0.84)	Female	-0.419*	(0.167)	-0.409*	(0.166)	-0.579*	(0.214)	-0.585*	(0.206)	-0.563*	(0.189)
	Family History of Any Cancer	-0.182	(0.176)	-0.183	(0.176)	-0.055	(0.234)	-0.054	(0.231)	-0.071	(0.225)
	Family History of Colorectal Cancer	0.396	(0.260)	0.391	(0.260)	0.564*	(0.279)	0.566*	(0.275)	0.556*	(0.276)
	Colorectal Polyps	-0.124	(0.329)	-0.121	(0.329)	-0.141	(0.429)	-0.147	(0.446)	-0.108	(0.351)
	Diabetes	0.520*	(0.258)	0.526*	(0.258)	0.114	(0.458)	0.117	(0.554)	0.206	(0.369)
65-69 yr	Screening	-0.386*	(0.168)	-0.526*	(0.165)	-0.564*	(0.187)	-0.568*	(0.166)	-0.576*	(0.188)
(0.83)	Female	-0.402*	(0.164)	-0.388*	(0.164)	-0.426*	(0.194)	-0.435*	(0.181)	-0.408*	(0.182)
	Family History of Any Cancer	-0.182	(0.176)	-0.185	(0.176)	-0.187	(0.190)	-0.190	(0.196)	-0.198	(0.186)
	Family History of Colorectal Cancer	0.565*	(0.129)	0.563*	(0.139)	0.625*	(0.260)	0.642*	(0.242)	0.627*	(0.253)
	Colorectal Polyps	-0.306	(0.314)	-0.299	(0.314)	-0.226	(0.331)	-0.243	(0.349)	-0.245	(0.328)
	Diabetes	0.370	(0.251)	0.377	(0.251)	0.036	(0.419)	0.045	(0.411)	0.138	(0.313)
70-74 yr	Screening	-0.414*	(0.196)	-0.364	(0.186)	-0.437	(0.225)	-0.439*	(0.223)	-0.439*	(0.223)
(0.81)	Female	-0.486*	(0.189)	-0.467*	(0.189)	-0.472*	(0.228)	-0.484*	(0.246)	-0.486*	(0.228)
	Family History of Any Cancer	0.157	(0.195)	0.152	(0.195)	0.387	(0.250)	0.389	(0.268)	0.388	(0.233)
	Family History of Colorectal Cancer	-0.219	(0.318)	-0.222	(0.318)	-0.344	(0.508)	-0.333	(0.405)	-0.342	(0.384)
	Colorectal Polyps	0.088	(0.286)	0.093	(0.286)	0.137	(0.387)	0.124	(0.396)	0.121	(0.349)
	Diabetes	0.444	(0.270)	0.451	(0.270)	-0.027	(0.583)	-0.031	(0.589)	-0.037	(0.409)

* indicates p-value ≤ 0.05

to those of the two older age groups. For example, in the age group 60–64 years, participants who were female, had diabetes, or had no family history of colorectal cancer were significantly less likely to comply with the assigned screening assignment than were males who had no diabetes, or who had a family history of colorectal cancer. Such associations may bias the estimation of the causal treatment effect by the as-treated analyses, which may explain the discrepancies observed in Table 4 between the as-treated analyses and the IV analyses. In addition, the IV analyses provide strong evidence for the lower colorectal cancer mortality risk in females (versus males) in all age groups beyond the age of 60 years. The results also suggest some survival disadvantages (related to colorectal cancer mortality) associated with a family history of colorectal cancer.

5. Conclusion

The use of IVs in survival settings with binary treatments has been severely limited by complexities arising from nonlinear model specifications, as with the proportional hazards model. The application of simple two-stage estimation pro-

cedures developed for linear models is challenging, and only valid in special cases. Alternative procedures may include strong modeling assumptions on strata other than that of interest, tend to be complex, both computationally and inferentially, and are not readily implemented using standard software. Our approach based on a special characterization of IVs enables a simple two-stage procedure analogous to propensity score weighting. At the first stage, a binary regression model is fit to the IV, and in the second stage, the fitted regression model from the first stage is used to construct a weight that "debiases" the naive estimating equation for the proportional hazards model. Previous work on this approach (Abadie, Angrist and Imbens (2002); Abadie (2003)) considers only iid estimating equations, with limited attention being given to practical computational issues. The current study demonstrates rigorously the validity of the approach using the partial likelihood score function. Moreover, the proposed estimators can be easily computed using existing software for the proportional hazards model, with the variance estimation based on bootstrapping correctly accounting for the first-stage estimation of the weights. Moreover, the estimators are generally applicable to IV estimations for the proportional hazards model in complex survival scenarios, for example, in the presence of left truncation, competing risks, and recurrent events.

Supplementary Material

The online Supplementary Material provides theoretical justifications and proofs, discussions of generalizations to complex survival settings, and additional figures and tables.

Acknowledgments

The first two authors contributed equally to this work. The authors would like to express special thanks to Jerome Mabie, Tom Riley, Ryan Nobel, and Josh Rathmell, of Information Management Services (IMS) Inc, for supporting and managing the PLCO data. The authors also thank Dr. Stuart G. Baker, National Cancer Institute, for kindly introducing the IMS team for this research. The authors gratefully acknowledge the support from the National Institutes of Health grant R01 HL113548.

References

Abadie, A. (2003). Semiparametric instrumental variable estimation of treatment response models. *Journal of econometrics* **113**, 231–263.

Abadie, A., Angrist, J. and Imbens, G. (2002). Instrumental variables estimates of the effect of subsidized training on the quantiles of trainee earnings. *Econometrica* **70**, 91–117.

Agresti, A. (2013). *Categorical Data Analysis*. 3rd Edition. Wiley, New Jersey.

Andersen, P. K. and Gill, R. D. (1982). Cox's regression model for counting processes: A large sample study. *The Annals of Statistics* **10**, 1100–1120.

Angrist, J. D., Imbens, G. W. and Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association* **91**, 444–455.

Baiocchi, M., Cheng, J. and Small, D. S. (2014). Instrumental variable methods for causal inference. *Statistics in Medicine* **33**, 2297–2340.

Baker, S. G. (1998). Analysis of survival data from a randomized trial with all-or-none compliance: estimating the cost-effectiveness of a cancer screening program. *Journal of the American Statistical Association* **93**, 929–934.

Baker, S. G., Kramer, B. S. and Prorok, P. C. (2002). Statistical issues in randomized trials of cancer screening. *BMC Medical Research Methodology* **2**, 11.

Baker, S. G. and Lindeman, K. S. (1994). The paired availability design: A proposal for evaluating epidural analgesia during labor. *Statistics in Medicine* **13**, 2269–2278.

Cuzick, J., Sasieni, P., Myles, J. and Tyrer, J. (2007). Estimating the effect of treatment in a proportional hazards model in the presence of non-compliance and contamination. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **69**, 565–588.

Gourieroux, C. and Monfort, A. (1981). Asymptotic properties of the maximum likelihood estimator in dichotomous logit models. *Journal of Econometrics* **17**, 83–97.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association* **81**, 945–960.

Imbens, G. and Angrist, J. (1994). Identification and estimation of local average treatment effects. *Econometrica* **62**, 467–476.

Joffe, M. M. (2001). Administrative and artificial censoring in censored regression models. *Statistics in Medicine* **20**, 2287–2304.

Li, G. and Lu, X. (2015). A bayesian approach for instrumental variable analysis with censored time-to-event outcome. *Statistics in Medicine* **34**, 664–684.

Li, J., Fine, J. and Brookhart, A. (2015). Instrumental variable additive hazards models. *Biometrics* **71**, 122–130.

Li, S. and Gray, R. J. (2016). Estimating treatment effect in a proportional hazards model in randomized clinical trials with all-or-nothing compliance. *Biometrics* **3**, 742–750.

Loeys, T. and Goetghebeur, E. (2003). A causal proportional hazards estimator for the effect of treatment actually received in a randomized trial with all-or-nothing compliance. *Biometrics* **59**, 100–105.

MacKenzie, T. A., Løberg, M. and O'Malley, A. J. (2016). Patient centered hazard ratio estimation using principal stratification weights: application to the norccap randomized trial of colorectal cancer screening. *Observational Studies* **2**, 29.

Martinussen, T., Nørbo Sørensen, D. and Vansteelandt, S. (2019). Instrumental variables estimation under a structural cox model. *Biostatistics* **20**, 65-79.

Martinussen, T., Vansteelandt, S., Tchetgen, E. and Zucker, D. M. (2017). Instrumental variables estimation of exposure effects on a time-to-event response using structural cumulative survival models. *Biometrics* **73**, 1140–1149.

Nie, H., Cheng, J. and Small, D. S. (2011). Inference for the effect of treatment on survival probability in randomized trials with noncompliance and administrative censoring. *Biometrics* **67**, 1397–1405.

Prorok, P. C., Andriole, G. L., Bresalier, R. S., Buys, S. S., Chia, D., Crawford, E. D., Fogel, R., Gelmann, E. P., Gilbert, F., Hasson, M. A., Hayes, R. B., Johnson, C. C., Mandel, J. S., Oberman, A., O'Brien, B., Oken, M. M., Rafla, S., Reding, D., Rutt, W., Weissfeld, J. L., Yokochi, L., Gohagan, J. K., Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial Project Team (2000). Design of the prostate, lung, colorectal and ovarian (PLCO) cancer screening trial. *Controlled Clinical Trials* **21**, 273S – 309S.

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Robins, J. M. and Tsiatis, A. A. (1991). Correcting for non-compliance in randomized trials using rank preserving structural failure time models. *Communications in Statistics-Theory and Methods* **20**, 2609–2631.

Rousseeuw, P. J. and Croux, C. (1993). Alternatives to the median absolute deviation. *Journal of the American Statistical Association* **88**, 1273–1283.

Tchetgen, E. J. T., Walter, S., Vansteelandt, S., Martinussen, T. and Glymour, M. (2015). Instrumental variable estimation in a survival context. *Epidemiology (Cambridge, Mass.)* **26**, 402.

Therneau, T. M. (2015). *A Package for Survival Analysis in S*, version 2.38.

Wang, L., Tchetgen, E. T., Martinussen, T. and Vansteelandt, S. (2018). Learning causal hazard ratio with endogeneity. *arXiv preprint arXiv:1807.05313*.

Yu, W., Chen, K., Sobel, M. E. and Ying, Z. (2015). Semiparametric transformation models for causal inference in time-to-event studies with all-or-nothing compliance. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **77**, 397–415.

Zeng, D. and Lin, D. Y. (2007). Maximum likelihood estimation in semiparametric regression models with censored data (with discussion). *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **69**, 507–564.

Behzad Kianian

Department of Biostatistics and Bioinformatics, Emory University, Atlanta, GA, 30322, U.S.A.

E-mail: behzad.kianian@emory.edu

Jung In Kim

Departments of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, U.S.A.

E-mail: jikim@live.unc.edu

Jason P. Fine

Departments of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, U.S.A.

E-mail: jfine@email.unc.edu

Limin Peng

Department of Biostatistics and Bioinformatics, Emory University, Atlanta, GA, 30322, U.S.A.
E-mail: lpeng@emory.edu

(Received March 2019; accepted June 2019)