

RESEARCH ARTICLE

Open Access



Time dependent hazard ratio estimation using instrumental variables without conditioning on an omitted covariate

Todd A. MacKenzie^{1,2*}, Pablo Martinez-Cambor^{1,2} and A. James O'Malley^{1,2}

Abstract

Background: Estimation that employs instrumental variables (IV) can reduce or eliminate bias due to confounding. In observational studies, instruments result from natural experiments such as the effect of clinician preference or geographic distance on treatment selection. In randomized studies the randomization indicator is typically a valid instrument, especially if the study is blinded, e.g. no placebo effect. Estimation via instruments is a highly developed field for linear models but the use of instruments in time-to-event analysis is far from established. Various IV-based estimators of the hazard ratio (HR) from Cox's regression models have been proposed.

Methods: We extend IV based estimation of Cox's model beyond proportionality of hazards, and address estimation of a log-linear time dependent hazard ratio and a piecewise constant HR. We estimate the marginal time-dependent hazard ratio unlike other approaches that estimate the hazard ratio conditional on the omitted covariates. We use estimating equations motivated by Martingale representations that resemble the partial likelihood score statistic. We conducted simulations that include the use of copulas to generate potential times-to-event that have a given marginal structural time dependent hazard ratio but are dependent on omitted covariates. We compare our approach to the partial likelihood estimator, and two other IV based approaches. We apply it to estimation of the time dependent hazard ratio for two vascular interventions.

Results: The method performs well in simulations of a stepwise time-dependent hazard ratio, but illustrates some bias that increases as the hazard ratio moves away from unity (the value that typically underlies the null hypothesis). It compares well to other approaches when the hazard ratio is stepwise constant. It also performs well for estimation of a log-linear hazard ratio where no other instrumental variable approaches exist.

Conclusion: The estimating equations we propose for estimating a time-dependent hazard ratio using an IV perform well in simulations. We encourage the use of our procedure for time-dependent hazard ratio estimation when unmeasured confounding is a concern and a suitable instrumental variable exists.

Keywords: Causal inference, Censoring, Semi-parametric model, Marginal model

*Correspondence: todd.a.mackenzie@dartmouth.edu

¹Department of Biomedical Data Science, Dartmouth College, New Hampshire, USA

²The Dartmouth Institute for Health Policy and Clinical Practice, Dartmouth College New Hampshire, USA



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

Confounding is a threat to all observational studies. The factors that influence the outcome of interest may also influence the selection of treatment. In randomized studies, intention-to-treat estimators are generally consistent estimators of the intention-to-treat effect but are not consistent estimators of the treatment effect. Instrumental variables (IV) influence treatment choice but otherwise have no effect on the outcome (*exclusion restriction*) and are independent of any other causes of the outcome (*randomization assumption*).

Estimation of treatment effects using instrumental variables for outcomes subject to right censoring has received the attention of many applied and several theoretical studies. Stukel et al. [1] proposed an ad-hoc estimator of the hazard ratio (HR) based on a linear model. MacKenzie et al. [2] proposed a hazard ratio estimator that assumes omitted covariates have an additive effect on the hazard. Tchetgen et al. [3] proposed an estimator of an additive hazards model based on two stage residual inclusion. They also argue that if the survival curve is close to unity (e.g. above 80%) for most of the followup that their additive hazards approach can be used as a good approximation of the multiplicative hazards approach. Li et al. [4] proposed a consistent estimator of a treatment satisfying an additive hazard model. Martinussen et al. [5] derived a consistent estimator for a structural Cox model. Martínez-Cambor et al. [6] identified the role of frailties in estimation of the hazard ratio via the two-stage residual inclusion algorithm if the treatment and omitted covariate jointly satisfy a Cox model. Wang et al. [7] derived an estimator of the marginal hazard ratio using a binary instrument.

The proportionality of hazards is a key assumption in estimation of the hazard ratio. Many test statistics have been proposed to test this assumption and estimate the hazard ratio as a function of time. The most simple approach to moving beyond proportionality of hazards is a piecewise constant function. The next most simple is a linear function of time. In this paper we extend the IV based estimator of the HR we proposed in 2014 MacKenzie et al. [2] beyond proportionality of hazards. We propose a method of estimating the HR as a time-dependent function. We address two cases, (a) a piecewise constant HR and (b) a log-linear time dependent hazard ratio. We conduct Monte Carlo simulations that include the use of copulas to generate potential times-to-event. We demonstrate this method for time-dependent hazard ratio estimation to compare the effect of two vascular interventions on survival.

Methods

Marginal structural Cox proportional hazards model

MacKenzie et al. [2] implemented a Cox proportional hazards model for the effect of treatment on the

time-to-event that is marginal with respect to any *omitted covariate*. The term marginal is used to mean that if the treatment is applied independently of the omitted covariate, the distribution of the time-to-event conditioning only on the treatment satisfies a Cox PH model, which we now elaborate on.

Let $\{S_x\}$ be the survival curve of the potential outcomes (time-to-event) if subjects receive level x of the exposure. In what follows we shall assume X is binary, e.g. 0 (no treatment) or 1 (treatment), although the methods described below are applicable to a continuously valued exposure. The structural Cox model is given by $S_x(t) = S_0(t)^{\exp\{\beta_X \cdot x\}}$ and it is usually written in terms of the hazard function,

$$\lambda(t; x) = \lambda_0(t) \cdot \exp\{\beta_X \cdot x\},$$

where $\lambda_0(t)$ is the *baseline hazard* function. For any covariate, U , which also affects the time-to-event, let $S_x(\cdot; u)$ be the survival curve for the potential time-to-event if a subject receives level x conditional on $U = u$. The marginal structural Cox proportional hazards model supposes that

$$\begin{aligned} \mathbb{E}_U [S_x(t; u)] &= \int S_x(t; u) dF_U(u) = S_x(t) \\ &= S_0(t)^{\exp\{\beta_X \cdot x\}} \text{ for all } t \geq 0 \\ &\text{and exposure levels } x, \end{aligned}$$

where F_U stands for the distribution function of U on the studied population. The alternative to the marginal model is a conditional model that specifies a structural form for the conditional distribution given both the exposure X and omitted covariate U (e.g. multivariable Cox PH model). Due to non-collapsibility of Cox's model [8], the marginal model that corresponds to this multivariable Cox model is not a Cox PH model.

Specifying that the marginal model is a Cox PH model has the following advantages. First, it makes no parametric assumption about how the omitted covariate U affects the distribution of the time-to-event: We allow for the possibility that the omitted covariate has a causal effect on the outcome without specifying a parametric model. Second, it does not require interpreting the treatment effect as conditional on a variable that is unknown (the omitted covariate). Third, this is the convention used in the reporting of randomized trials: It is unusual to analyze randomized trials using models that control for omitted covariates (such as Cox's model with frailty), and the convention has been not to condition on measured characteristics either [9].

Time dependent hazard ratio

Proportionality of hazards is equivalent to a hazard ratio that does not vary over time. A constant hazard ratio

can be a simple and good approximation, just as a linear model may be chosen when there is evidence of some non-linearity. On the other hand there are times when the hazard ratio varies considerably and attention is focused on its time dependence. For instance, surgery may be associated with increased risks up front compared to endovascular procedures. There is a large literature on estimation of the hazard ratio as a function of time including smoothing splines [10], regression splines [11] or penalized regression splines [12] among other techniques.

Time-to-event, exposure and instrument

Suppose the observed data consists of independent and identically distributed observations, $\{W_i, X_i, \Delta_i, T_i\}_{i=1}^n$ where W_i is the instrument ($1 \leq i \leq n$) and X_i is the exposure level (either binary, ordinal or continuous). $\Delta_i = 1_{T_i^0 < C_i}$ and $T_i = \min(T_i^0, C_i)$ are the event indicator and observed followup time respectively where T_i^0 is the focus of interest, the time-to-event, and C_i is the censoring time. We assume that the potential censoring times $\{C_i(x)\}_{x=0,1}$ are independent of the potential time-to-events $\{T_i^0(x)\}_{x=0,1}$.

An instrument is defined by the following properties:

1. It is associated with the exposure ($W \not\perp X$).
2. It satisfies the following two restrictions:
 - (a) There is no effect of the instrument on the outcome except through its effect on the exposure.
 - (b) There are no confounders of the instrument and the outcome.

The latter two statements can be summarized as the potential outcomes are independent of the instrumental variable, $\{T^0(x)\}_{x=0,1} \perp W$.

To satisfy Assumption 1 the instrument may either be causal (i.e. it affects treatment choice) or non-causal (it is an effect of a variable that also affects exposure). Assumptions 2(a) and 2(b) are often combined by stating that conditional on X , the outcome and instrument are independent. The instrument, W , may be continuous, ordinal or binary.

For each subject there is one *potential* stochastic process $\{N_i(t; x), R_i(t; x)\}_{t \geq 0, x}$ for each possible level of x where $R_i(t; x)$ is the *at risk* indicator $1_{T_i(x) \geq t}$, and $N_i(t; x)$ is the *counting process* $1_{T_i \leq t}$. Let $R_i(t) = R_i(t; X_i)$ and $N_i(t) = N_i(t; X_i)$.

In studies of instrumental variables another assumption that is typically introduced is homogeneity of the treatment effect and monotonicity of the effect of the instrument. In usual applications homogeneity of the treatment effect is the assumption that the effect is same for different levels of the omitted covariate, i.e. there is no interaction

of the effect of the treatment and the omitted covariate. In the marginal model an estimate of the treatment effect will depend on the value of the omitted covariate. We do not specify the effect of the omitted covariate; we avoid this parameterization as the method we propose below relies on a first order linear approximation of the hazard in terms of the omitted covariate.

Another frequently addressed assumption in studies of instrumental variables is monotonicity of the effect of the instrument on the treatment. It assumes that in all subjects the instrument is positively associated with the treatment; for instance, the derivation of the complier average causal effect is dependent on it. In this study we make the monotonicity assumption, $Pr[X \geq x | W = w_1] > Pr[X \geq x | W = w_0]$ if $w_1 > w_0$ for all x .

Estimating equations for a time dependent hazard ratio

Let $\lambda_x(t)$ be the hazard function corresponding to survival in the population were all subjects in that population exposed to level x of the exposure, i.e., $\lambda_x(t) = -dS_x(t)/S_x(t)$ which can also be written $\mathbb{E}[dN_i(t; x) | R_i(t; x) = 1]$. We suppose that, at time t , the marginal (e.g. population) causal effect of having been exposed to level 1 relative to level 0 of the time-varying binary treatment $X(t)$ is a change in the log-hazard of $\beta_X(t)$:

$$\lambda_1(t) = \exp\{\beta_X(t)\} \cdot \lambda_0(t). \tag{1}$$

Further, suppose that the time dependent log-hazard ratio is parameterized by $\beta_X(t; \theta)$. For instance, one simple parameterization is $\beta_X(t; \theta) = \theta_0 + \theta_1 \cdot t$.

If there is no selection bias, i.e. X is independent of any omitted covariate, U , the difference

$$N_i(t) - \int_0^t R_i(v) \cdot \exp\{\beta_X(v) \cdot X_i\} d\Lambda_0(v),$$

which is known as a Martingale residual (Kalb eisch and Prentice [13]), has an expected value of zero and is independent of X ($\Lambda_0(v) = \int_0^v \lambda_0(s) ds$). This implies that

$$\mathbb{E} \left[X_i \cdot \left(N_i(t) - \int_0^t R_i(v) \cdot \exp\{\beta_X(v; \theta) \cdot X_i\} d\Lambda_0(v) \right) \right] = 0 \quad \forall t > 0.$$

Moreover,

$$\mathbb{E} \left[\int_0^t \phi(v) \cdot X_i \cdot (dN_i(v) - R_i(v) \cdot \exp\{\beta_X(v; \theta) \cdot X_i\} d\Lambda_0(v)) \right] = 0$$

for any real function $\phi(\cdot)$.

This equation suggests that an estimator of $\beta_X(t; \theta)$ is that $\hat{\beta}_X(t; \theta)$ for which

$$0 = \sum_{i=1}^n \int_0^\tau \phi_k(v) \cdot X_i \cdot [dN_i(v) - R_i(v) \cdot \exp\{\hat{\beta}_X(v; \theta) \cdot X_i\} d\Lambda_0(v)] \tag{2}$$

where τ is a suitably large point of time (e.g. $\max_{1 \leq i \leq n} \{t_i\}$), K is the dimension of the parameter $\theta = (\theta_1, \dots, \theta_K)$ and $\{\phi_k(t)\}_{k=1}^K$ are suitably selected real functions. These functions should be chosen to minimize the variance of the resulting estimator which motivates the choices $\phi_k(t) = \partial \beta_X(t; \theta) / \partial \theta_k$.

As $\Lambda_0(t)$ is unknown we propose to solve for it using the estimating equations

$$0 = \sum_{i=1}^n \left[N_i(t) - \int_0^t R_i(v) \cdot \exp \left\{ \hat{\beta}_X(v; \theta) \cdot X_i \right\} d\hat{\Lambda}_0(v) \right] \quad (3)$$

for all t which is solved by the Breslow estimator (Lin [14])

$$d\hat{\Lambda}_0(v) = \frac{\sum_{i=1}^n dN_i(v)}{\sum_{i=1}^n R_i(v) \cdot \exp \left\{ \hat{\beta}_X(v; \theta) \cdot X_i \right\}}. \quad (4)$$

Substitution of (4) into (2) yields the score equations for the partial likelihood (Cox [15]) for a time-dependent hazard ratio (Abrahamowicz et al. [11]). That is, the estimating equations (2) are equivalent to the method of maximum partial likelihood estimation of a time-dependent hazard ratio. If X is endogenous (confounded), then X_i is not necessarily independent of the Martingale residual

$$N_i(t) - \int_0^t R_i(v) \cdot \exp \left\{ \beta_X(v; \theta) \cdot X_i \right\} d\Lambda_0(v)$$

and therefore the estimating Eq. (2) are biased.

If W is an instrumental variable then it is approximately independent of the counting process conditional on X , or equivalently, the IV is independent of the Martingale residual. A heuristic justification starts by writing the causal hazard function conditional on an omitted covariate U that affects both the treatment and the time-to-event. We approximate the causal hazard function by a linear Taylor expansion, $E[dN_i(t) | R_i(t) = 1] = \exp\{\beta_X(v; \theta) \cdot X_i\} d\Lambda_0(v) + \phi(t) [U - \mu_U(t)]$ where $\mu_U(t)$ is the expected value of the omitted covariate among people at risk at time t . Then equating to zero the correlation of the instrument with the counting process leads to the following estimating equation θ ,

$$0 = \sum_{i=1}^n \int_0^\tau W_i \cdot \frac{\partial}{\partial \theta_k} \beta_X(v; \theta) \cdot \left[dN_i(v) - R_i(v) \cdot \exp \left\{ \hat{\beta}_X(v; \theta) \cdot X_i \right\} d\Lambda_0(v) \right]. \quad (5)$$

If we substitute the Breslow estimator for $\Lambda_0(v)$ into the latter equation it results in the estimating equation

$$0 = \sum_{i=1}^n \int_0^\tau \phi_k(v) \cdot \left[W_i - \frac{\sum_{j=1}^n W_j \cdot R_j(v) \cdot \exp \left\{ \beta_X(v; \theta) \cdot X_j \right\}}{\sum_{j=1}^n R_j(v) \cdot \exp \left\{ \beta_X(v; \theta) \cdot X_j \right\}} \right] dN_i(v). \quad (6)$$

If the time-dependent function is constant, i.e. $\beta_X(v; \theta)$ equals a scalar θ this is the same estimating equation proposed in MacKenzie et al. [2].

IV estimation for piecewise proportional hazards

In applications, categorization of follow-up time is favored as an approach to assessing time-dependence of the hazard ratio because of its simplicity. For instance, one can make statements such as, in the first month the hazard ratio was 1.5 but after one month was 0.5. The time-frame is categorized into intervals $\{\tau_{i-1}, \tau_i\}_{i=1}^K$ for $0 = \tau_0 < \tau_1 < \dots < \tau_K = \tau$ and within each window the hazard ratio is assumed to be constant, $\beta_X(t; \theta) = \theta_i$ for $\tau_{i-1} \leq t < \tau_i$. The partial likelihood estimator of a step function time dependent hazard ratio is equivalent to applying the partial likelihood estimator of the proportional hazards model within each time window. For instance, to estimate the hazard ratio in the interval $[\tau_{i-1}, \tau_i)$ exclude all subjects eliminated from risk before time τ_{i-1} and censor any events after time τ_i ($1 \leq i \leq K$). For this piecewise constant hazard ratio the IV based estimator proposed in (6) is equivalent to applying the IV based method of MacKenzie et al. [2]. The **R** function supplied in that paper can be used accordingly to implement this approach for each time window (<https://github.com/toddmackenzie/Instrumental-Variable-Hazard-Ratio-Estimation>).

IV estimation for linear time dependent log hazard ratio

Any parameterization of the hazard ratio could be implemented with the instrumental variable estimating equations we propose. We chose to illustrate the approach using a log-linear time-dependent model of the hazard ratio, $\beta_X(t; \theta) = \theta_0 + \theta_1 \cdot t$. In this parameterization $\exp\{\theta_0\}$ is the hazard ratio at inception and $\exp\{\theta_1\}$ is the multiplicative change in the hazard ratio per unit time. The two parameters θ_0 and θ_1 can be estimated using the equations

$$0 = \sum_{i=1}^n \int_0^\tau \left[W_i - \frac{\sum_{j=1}^n W_j \cdot R_j(v) \cdot \exp \left\{ \beta_X(v; \theta) \cdot X_j \right\}}{\sum_{j=1}^n R_j(v) \cdot \exp \left\{ \beta_X(v; \theta) \cdot X_j \right\}} \right] dN_i(v), \quad (7)$$

and

$$0 = \sum_{i=1}^n \int_0^\tau v \cdot \left[W_i - \frac{\sum_{j=1}^n W_j \cdot R_j(v) \cdot \exp \left\{ \beta_X(v; \theta) \cdot X_j \right\}}{\sum_{j=1}^n R_j(v) \cdot \exp \left\{ \beta_X(v; \theta) \cdot X_j \right\}} \right] dN_i(v), \quad (8)$$

respectively.

Monte Carlo simulations

We evaluated the behavior of the estimating equations we propose in (6) under two scenarios for the marginal time-dependent hazard ratio; i) a three piece constant hazard

ratio, and *ii*) a log linear time-dependent hazard ratio. In the first scenario, the estimation method is equivalent to applying the approach of MacKenzie et al. [2] in early, middle and late follow-up settings, and that is how we implement and report the results. These simulations are novel in that the 2014 paper did not consider estimation of the hazard ratio later in follow-up, such as where the survival curve dips below 75%. Because a much greater proportion of the study will have either experienced the outcome event or been censored by that time, it is reasonable to expect that the performance of the method could differ over this time-period compared to the early follow-up period. For the log-linear function of time we consider a range of intercepts and a range of slopes. Specific details follow.

Comparators

In addition to evaluating the bias of our estimator for the stepwise constant and log-linear hazard ratio models, we evaluate the performance of the partial likelihood estimator. For the stepwise hazard model, because it is a proportional hazards model in each of the three periods, we also evaluate the performance of two other estimators designed for proportional hazards. First, we have evaluated the approach of Wang et al.[7] which estimates the hazard ratio for a marginal structural model like ours. Because Wang et al's approach is designed for a binary instrument, we dichotomize the continuous instrument of our simulations at the median. Second, we have evaluated the performance of the estimator of Martínez-Camblor et al. [6] which estimates the hazard ratios of a Cox model for the multivariable effect of the treatment and the omitted covariate; that is, it conditions on the omitted covariate.

Simulation methods

We have conducted extensive simulations to evaluate the bias of the estimator we propose. Our simulation uses a copula to generate times-to-event that depend on an omitted covariate and treatment in such a way that the marginal distribution of the potential outcomes, treated and untreated, satisfy a model with the specified time-dependent hazard ratio. The settings and parameter values for the simulation are listed in Tables 1 and 2.

The steps of the simulation are:

1. Randomly generate the bivariate random variable, (Y_0, Y_1) , from a bivariate copula., i.e. Y_0 and Y_1 are uniformly distributed but not independent.
2. Let the omitted covariate be the standard normal deviate obtained as $U = \Phi^{-1}(Y_0)$.
3. Randomly generate a continuous instrument, W , from the standard normal distribution.
4. Randomly generate a binary exposure indicator X using a logistic link with intercept of zero (yields

Table 1 Simulation Variables: $\Phi(\cdot)$ is the cumulative distribution functions for the standard normal distribution and F_1 is the cumulative distribution function for the time-to-event that has log-hazard of $\beta_1, \beta_2, \beta_3$ in periods 1,2 and 3 respectively

Variable	Symbol	Generation
Correlated Uniform Marginals	(Y_0, Y_1)	$\sim \text{copula}(\rho)$
Omitted Covariate	U	$\Phi^{-1}(Y_0)$
Instrumental Variable	W	$\sim N(0,1)$
Potential Outcome if not Exposed	T_0	$-\log(Y_0)$
Potential Outcome if Exposed	T_1	$F_1^{-1}(Y_1)$
Exposure	X	$\Pr[X = 1] = \text{logistic}(\alpha_{UX}U + \alpha_{WX}W)$

equal proportions of treated and untreated) and log odds ratios, α_W and α_U , for the association of W and U respectively.

5. Obtain potential time-to-events $T(0) = -\log(Y_0)$ (unit exponential distribution) and $T(1) = F_1^{-1}(Y_1)$ where F_1 is the cumulative distribution function whose hazard equals the specified time-dependent hazard ratio
6. Right censor the times-to-event using a uniform distribution whose scale is set to achieve the specified censoring rate.

In our Monte Carlo simulation the confounding (endogeneity) is created by (i) setting the omitted covariate as a monotonic function of the potential outcome under no treatment, which is correlated with the potential outcome

Table 2 Parameters are constant within a dataset but vary between datasets

Element	Value Set
log Odds Ratio U vs X: α_{UX}	$(-\log(10), \log(10))$
log Odds Ratio W vs X: α_{WX}	$(\log(2), \log(50))$
ρ : correlation of copula	0.50, 0.99
Copula family	Gaussian, Clayton, Gumbel
Stepwise log Hazard Ratio Period 1 β_1	$[-\ln(3), -\ln(3) + \log(6)/24, \dots, \ln(3)]$
Stepwise log Hazard Ratio Period 2 β_2	$[-\ln(3), -\ln(3) + \log(6)/24, \dots, \ln(3)]$
log-linear HR intercept	$[-1, -0.8, \dots, 1]$
log-linear HR slope over time	$[-0.5, -0.4, \dots, 0.5]$
Sample size of each dataset	1000
Censoring frequency	0.50
Number of events	500

Before each dataset is constructed the parameters above are randomly drawn with equal probability from the value set

under treatment (as controlled by the bivariate copula using either Pearson or Kendal correlations of 0.50 or 0.99) and (ii) randomly generating the treatment indicator based on the omitted covariate (as well as the instrument). Therefore, the level of confounding by the omitted covariate is controlled by a single parameter, the odds ratio, $\exp(\alpha_{UX})$ relating U (the untreated time-to-event transformed to have a standard normal distribution) to X . For instance, if $\alpha_{UX} = 0$ there is no confounding. Before the generation of each dataset it was randomly drawn from a Uniform distribution on the interval $\log(1/10), \log(10)$. Values above $\log(5)$ were considered positive confounding, and values below $-\log(5)$ were considered negative confounding.

The stepwise constant time-dependent hazard ratio took values as specified in Table 2. For each simulated dataset we randomly sampled with replacement from these grids for the early, mid and late follow-up hazard ratios. The early corresponded to approximately that time period up to which survival curve was above 90%, the

middle 90% down to 75%, and late survival of 75% or less. The log-linear time-dependent hazard ratio was generated based on the randomly drawn intercept and slope as specified in Table 2.

Complete R code is available at: <https://github.com/toddmackenzie/Instrumental-Variable-Hazard-Ratio-Estimation>, for users interested in conducting the same simulations.

Simulation results for early, mid and late estimators of the HR

Figure 1 shows the results of the simulations when the instrument is strong, which we defined as an odds ratio between the treatment and the instrument that exceeds 5. Each red point represents the median of over 1000 maximum partial likelihood estimates of the stepwise hazard ratio. The partial likelihood estimators are clearly subject to bias from confounding, which decreases as follow-up increases. Each black point represents the median of over 1000 estimates using the instrument and

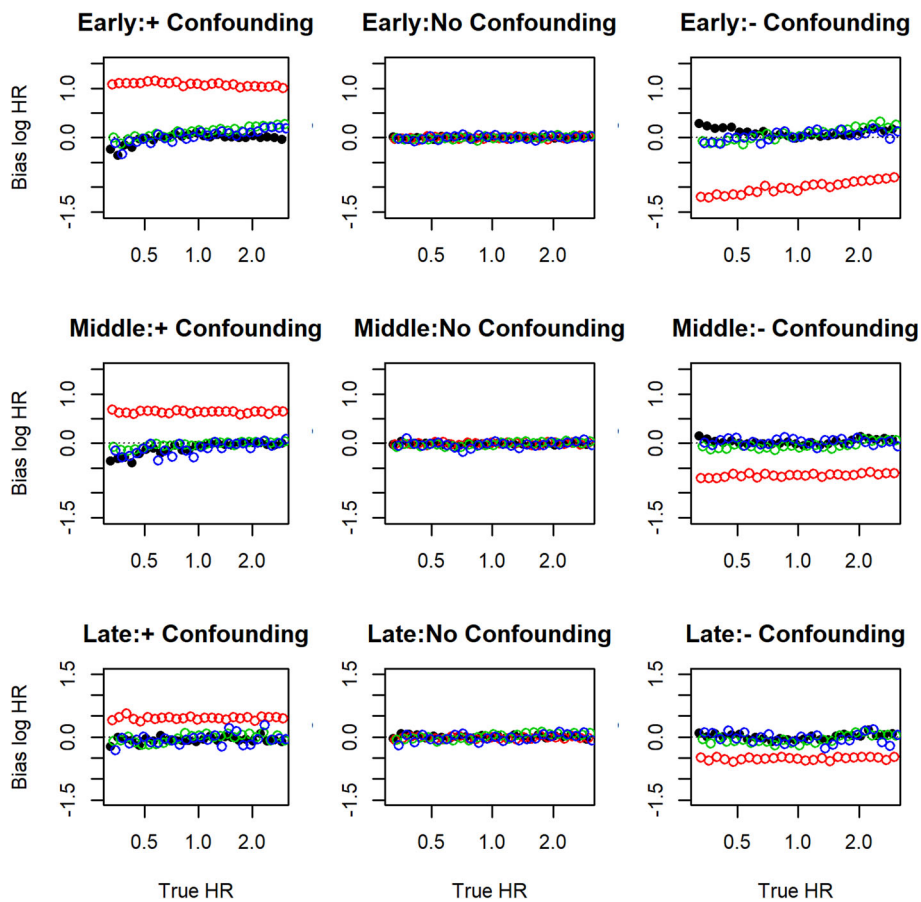


Fig. 1 Simulation Results for estimating three-piece time-dependent hazard ratio when the instrument is strong: The nine panes represent each combination of early, mid and late follow-up, with (i) strong positive confounding, (ii) no confounding and (iii) strong negative confounding. Each pane overlays scatterplots of the median estimated bias versus the true hazard ratio, our IV based estimator in black, the standard Cox MLE in red, the estimator of Wang in green and the estimator of Martinez-Cambor et al. in blue. Each point represents the median of over 1000 estimates obtained from that many randomly generated datasets

estimating equations we propose. The green and blue points represent the median of the distribution of the estimators of Wang et al. [7] and Martínez-Cambor et al. [6]. The instrumental variable based estimator we propose is unbiased as an estimator of the marginal hazard ratio in each of the three periods, with a slight tendency toward bias by confounding for large and small hazard ratios; that is, the magnitude of the bias increases as the HR moves away from null and in the direction opposite of the confounding. The estimators of Wang et al. [7] and Martínez-Cambor et al. [6] demonstrate similarity and slightly better accuracy.

The bias was affected by the strength of the instrument; for each doubling of the odds ratio between the treatment and the instrument, the bias reduced by approximately 5%. Figure 2 shows the results restricted to instruments for which the odds ratio between treatment and instrument was less than 5. The results are similar to those reported for the scenario of a strong instrument. The bias of our estimator did not vary with respect to choice of copula, or the correlation between the treated and untreated potential outcomes (results not reported).

Simulation results for log-Linear time dependent HR

Figure 3 shows the results of the simulations for estimation of the log-linear hazard ratio. The top row demonstrates that the intercept term of the log-linear hazard ratio is estimated with very little bias using the instrumental variable estimating equation we propose (blue points represent median of over 1000 estimates from that many simulated datasets), unlike the maximum partial likelihood estimators (points in red). Results from the estimation of slope parameter of the log-linear hazard ratio are found in the bottom row, which again show that the IV based estimating equation we propose yields unbiased estimators (blue points), unlike the partial likelihood estimator (in red) with the exception that the IV based estimator of the slope is biased when the true slope is zero. The estimators of Wang et al. [7] and Martínez-Cambor et al.[6] are not designed for log-linear estimators or any time-dependent hazard ratio beyond piecewise constant.

The simulation findings did not change with respect to the choice of the copula, the correlation between the treated and untreated potential outcomes, or the strength of the instrument.

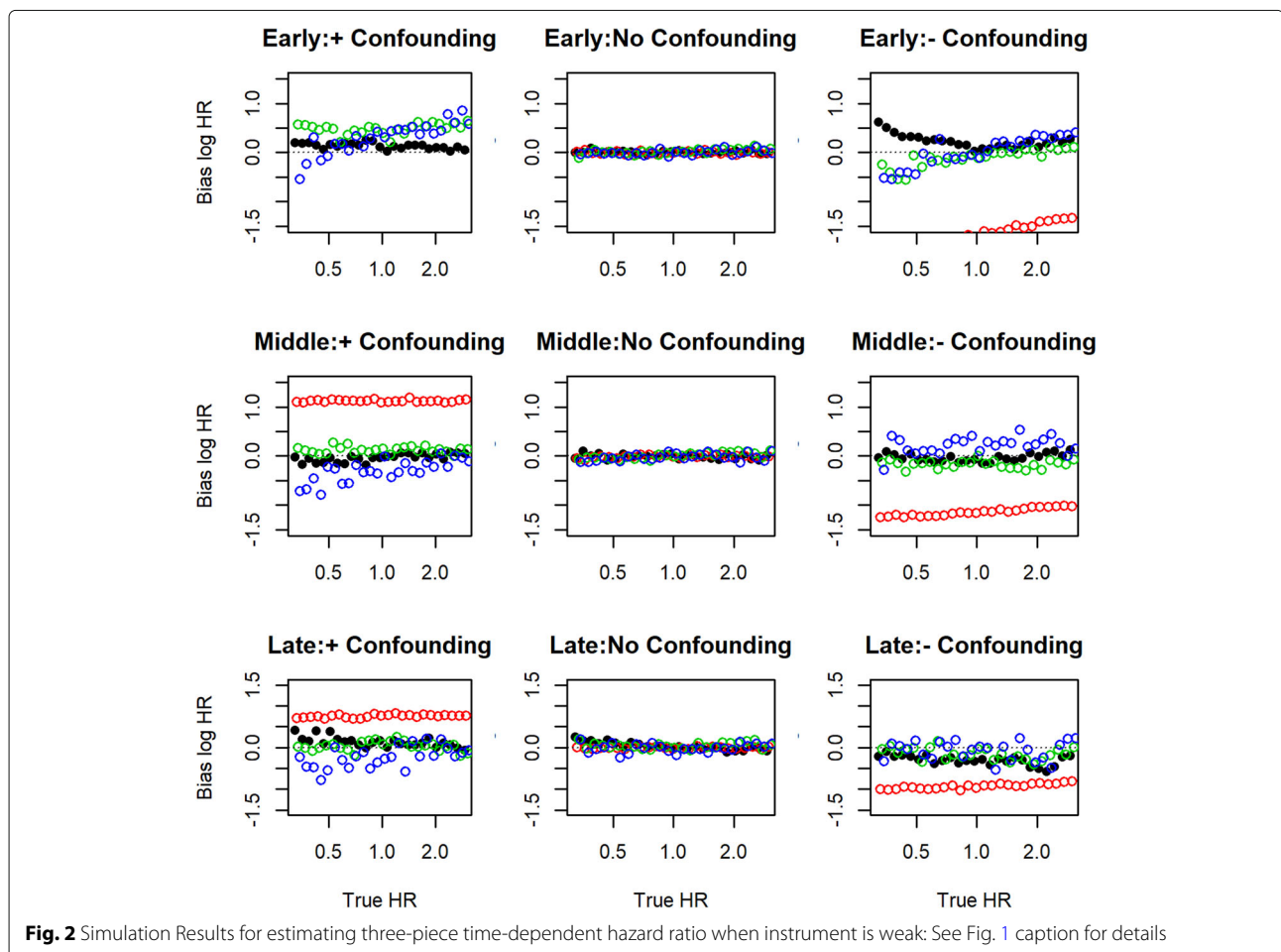


Fig. 2 Simulation Results for estimating three-piece time-dependent hazard ratio when instrument is weak: See Fig. 1 caption for details

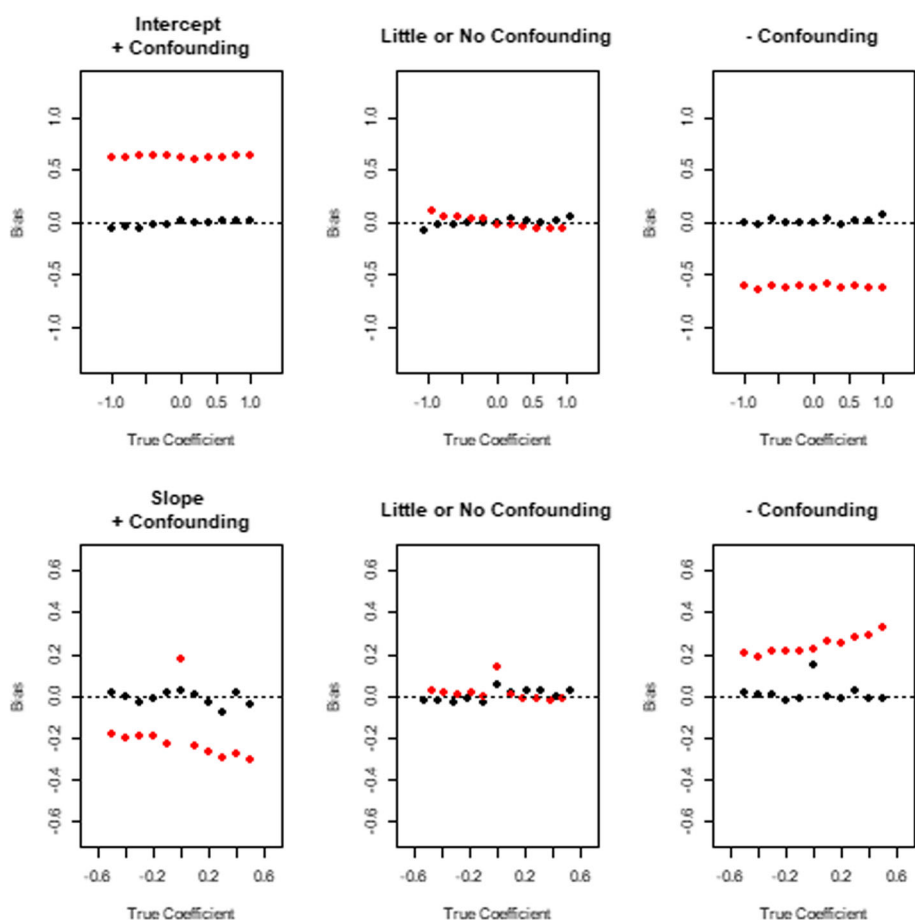


Fig. 3 Simulation Results for Estimating a Log-Linear Time Dependent Hazard Ratio: The top 3 panes represent estimation of the intercept parameter of the log-linear model for (i) strong positive confounding, (ii) no confounding and (iii) strong negative confounding, while the bottom three panes represent results for estimation of the slope parameter. Each panel overlays scatterplots of the median bias (estimated hazard ratio minus true coefficient) versus the true coefficient, our IV based estimator in black, and the Cox MPLE in red. Each point represents the median of over 1000 estimates (from that many simulated datasets)

Real-world application

We illustrate our method to address a comparative effectiveness question in patients with carotid stenosis. We use data from the nationwide Vascular Quality Initiative, VQI, (<http://www.vascularqualityinitiative.org>) to compare the composite outcome of death or stroke following intervention between those undergoing endarterectomy (CEA) and those undergoing carotid stenting (CAS). The data consists of 73,312 patients who received CEA and 12705 who received CA during the years 2003-2016. The number of events was 8,005 of which 730 occurred in the first 30 days, 2498 in months 1 through 12 and 4,777 after the first year. This example has been previously considered by Columbo et al. [16]. Figure 4 (top) shows the Kaplan-Meier estimates for the survival curves on both the CEA and the CAS groups.

The population of patients who undergo CEA may not be the same as those who undergo CAS, and therefore

any estimator of comparative efficacy may be biased by confounding. Therefore we utilize an instrumental variable approach. The instrument we employ is the center level relative frequency of CEA versus CAS procedures over the 12 months prior to the current patient. A value of zero indicates that the patient received the surgery in a facility which during the prior 12 months, has not performed CEA. The validity of this instrument is argued in Martínez-Cambor et al. [17]. Figure 4, bottom, depicts the instrument's distribution (histograms at left and box-plot at right) in both groups.

Figure 5 shows estimates of the hazard ratio as a function of time. The left panel contains the partial likelihood estimators of the stepwise constant hazard ratio and the log-linear hazard ratio. After consulting with physicians and a visual inspection of the Kaplan-Meier survival curve, we chose the cut-offs of 1 month (30 days) and 6 months. In the first month the partial likelihood

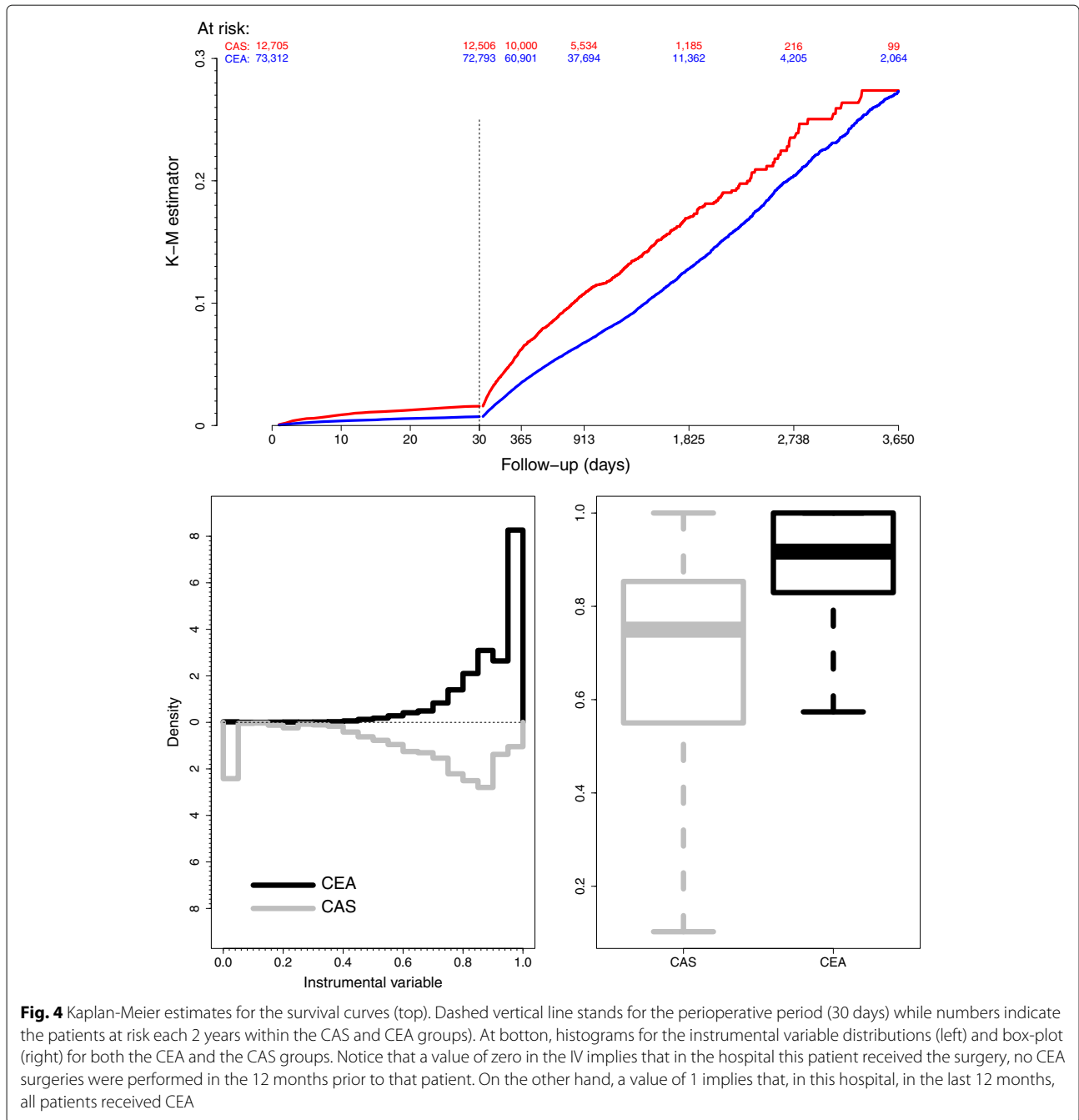


Fig. 4 Kaplan-Meier estimates for the survival curves (top). Dashed vertical line stands for the perioperative period (30 days) while numbers indicate the patients at risk each 2 years within the CAS and CEA groups. At bottom, histograms for the instrumental variable distributions (left) and box-plot (right) for both the CEA and the CAS groups. Notice that a value of zero in the IV implies that in the hospital this patient received the surgery, no CEA surgeries were performed in the 12 months prior to that patient. On the other hand, a value of 1 implies that, in this hospital, in the last 12 months, all patients received CEA

estimate of the hazard ratio comparing CEA to CAS is 0.46 (95%CI: 0.39 to 0.54), while it is 0.55 (95% CI: 0.48 to 0.62) in months 2 through 6, and considerably closer to unity thereafter at 0.75 (95%CI: 0.71 to 0.82). The corresponding estimates based on our instrumental variable approach are similar in the first interval, 0.44 (95%CI: 0.31 to 0.61), although the confidence interval is wider as expected for IV based estimators. The IV based estimator of the hazard ratio in the middle interval is closer to unity, 0.66 (95%CI: 0.50 to 0.88). The

estimator in the final interval is 0.77 (95%CI: 0.65 to 0.91).

The model for the hazard ratio as a log-linear function of time indicates the same pattern of the hazard ratio moving toward unity as time progresses. Both the partial likelihood estimator and the IV based estimator indicate the hazard ratio begins at approximately 0.5 but the IV based estimator reaches unity earlier. The observed results at both early and late follow-up were quite close in the standard and proposed methodologies suggesting

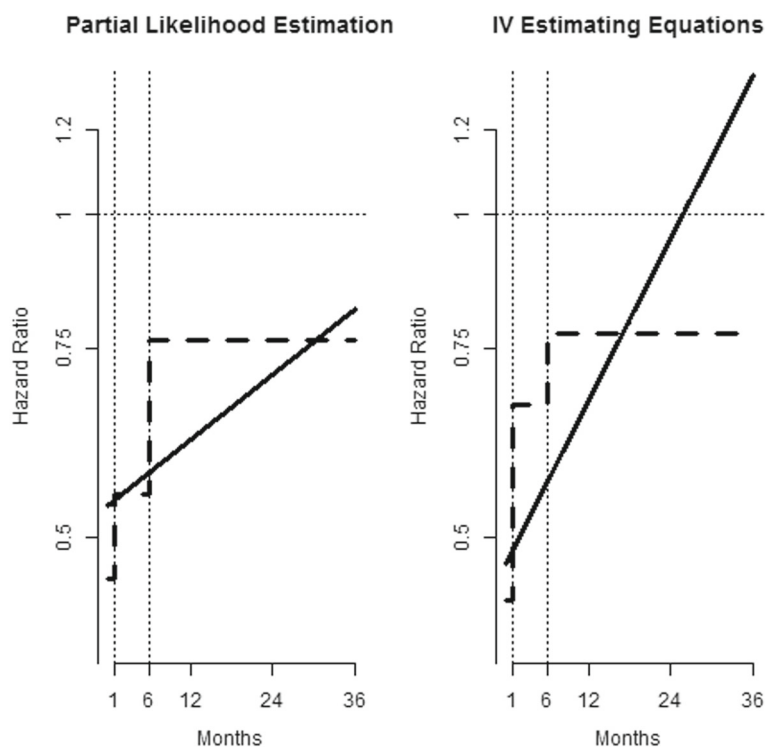


Fig. 5 Hazard Ratios Comparing Risk of Stroke or Death Between CSA and CEA: The left pane shows partial likelihood estimators of a 3 piece time-dependent constant hazard ratio (steps at 1 and 6 months, in a dashed line) and a log-linear time-dependent hazard ratio (solid line). The pane on the right shows the corresponding estimators based on the instrumental variable we proposed

that, in average, the potential covariates affecting the survival have a minor impact on the effect of the treatment. Between the first and six months, the difference is around 20% indicating that in this period some omitted covariates spuriously enhance the observed effect of the treatment.

Discussion

We have proposed a method for using instruments in the estimation of time-dependent hazard ratios. The framework is general enough to accommodate any parameterization for a time-dependent hazard ratio. Like maximum partial likelihood estimation of Cox's proportional hazards model it does not require a parameterization of the baseline hazard. We have illustrated our approach using two forms for a time-dependent hazard ratio, (i) a piecewise constant hazard ratio and (ii) a log linear function of time.

Our approach focuses on time-dependent hazard ratios that are marginal with respect to the omitted covariate. That is, just like estimators of treatment effects in randomized studies, the model we employ does not explicitly condition on the omitted covariate although it is motivated by a linear approximation of a hybrid hazard function in which the dependence on the omitted covariate is additive and the dependence on treatment and

the observed covariates is multiplicative. Our approach is analogous to estimators of population averaged parameters such as generalized estimating equations. We encourage analysts to estimate both marginal models like ours and models that condition on the omitted covariate. A disadvantage of the conditional model is that it conditions on intangible characteristics, that is unmeasured characteristics, that somehow affected the treatment selection.

We conducted extensive simulations to evaluate our estimator of the piecewise hazard ratio and the log-linear hazard ratio. For the former, which reduces to a proportional hazards model for any of the time windows (in which the hazard ratio is constant) we compared our approach to the partial likelihood estimator, the approach of Martínez-Cambor et al. [6] and the approach of Wang et al. [7]. For the log-linear hazard ratio model, we compared our approach to the partial likelihood estimator. The simulations indicated our estimator of a piecewise hazard ratio has some bias toward the null which is comparable to the bias of Martínez-Cambor et al. [6] and Wang et al. [7]. Our simulations indicate our estimator of the log-linear hazard ratio has little bias, especially in comparison to maximum partial likelihood estimation when confounding by omitted covariates is present.

While we do not address it, this method generalizes readily to multivariable models without much additional complexity. That is, measured covariates can be incorporated into the model.

The method we propose has some notable limitations. We do not believe it is a consistent estimator because we derived our estimating equations by substituting for the baseline hazard the Breslow estimator without justification of why this is the ideal estimator using an instrumental variable. In addition, it is based on an approximation of the hazard as an additive function of the omitted covariate.

We believe our estimator can be improved by utilizing inverse probability weighting in our estimating equations. In particular, like the approach of Wang et al. [7] and Huling et al. [18] one should weight the subjects at risk at any time t using the instrumental variable to make those subjects at risk generalizable to the entire population, i.e. those at risk at time $t = 0$, at which time the instrument is orthogonal to the omitted covariates.

Conclusion

The estimating equations we propose for estimating a time-dependent hazard ratio using an IV perform well in simulations. We encourage the use of our procedure for time-dependent hazard ratio estimation when unmeasured confounding is a concern and a suitable instrumental variable exists.

Abbreviations

IV: Instrumental Variable; HR: Hazard Ratio

Acknowledgements

The authors would like to acknowledge Phil Goodney and Jesse Columbo for facilitating access to the data.

Authors' contributions

TM conceived of the method, developed the math, performed the simulations and led the writing of the manuscript. PM-C contributed to the theory, performed the statistical analysis for the example and contributed text and editing of the manuscript. JO contributed to the concept, simulations, example and writing of the manuscript. All three authors have approved the manuscript.

Funding

This work was supported in part by a grant ME-1503-28261 from the USA federally funded agency, Patient-Centered Outcomes Research Institute, to develop statistical methodology for using instrumental variables in survival analysis. PCORI played no roles in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Availability of data and materials

The code for this method and the simulations is available at: <https://github.com/toddmackenzie/Instrumental-Variable-Hazard-Ratio-Estimation>. The data used in the example was accessed from Centers for Medicare and Medicaid Services under contract and is not available for sharing.

Declarations

Ethics approval and consent to participate

All data are collected under the auspices of an Agency for Healthcare Research and Quality designated Patient Safety Organization and were deidentified. Our study was approved by the Center for the Protection of Human Subjects at Dartmouth. All patient personal health information was protected, records,

and outcomes were deidentified, and no testing or procedures were required for this study. Thus, the need for specific consent was waived.

Consent for publication

Not applicable.

Competing interests

There are no conflicts of interest to report.

Received: 31 July 2020 Accepted: 2 March 2021

Published online: 20 March 2021

References

- Stukel T, Fisher E, Wennberg D, Alter D, Gottlieb D, Vermeulen M. Analysis of observational studies in the presence of treatment selection bias: effects of invasive cardiac management on AMI survival using propensity score and instrumental variable methods. *JAMA*. 2007;297(3):278–85.
- MacKenzie T, Tosteson T, Morden N, Stukel T, O'Malley A. Using instrumental variables to estimate a Cox's proportional hazards regression subject to additive confounding. *Health Serv Outcome Res Methodol*. 2014;14(1-2):54–68.
- Tchetgen E, Walter S, Vansteelandt S, Martinussen T, Glymour M. Instrumental variable estimation in a survival context. *Epidemiol Camb Mass*. 2015;26(3):402.
- Li J, Fine J, Brookhart A. Instrumental variable additive hazards models. *Biom*. 2015;71(1):122–30.
- Martinussen T, Vansteelandt S, Tchetgen Tchetgen E, Zucker D. Instrumental variables estimation of exposure effects on a time-to-event endpoint using structural cumulative survival models. *Biom*. 2017;73(4):1140–9.
- Martínez-Cambor P, MacKenzie T, Staiger D, Goodney P, James O'Malley A. An instrumental variable procedure for estimating Cox models with non-proportional hazards in the presence of unmeasured confounding. *J R Stat Soc: Ser C: Appl Stat*. 2019.
- Wang L, Tchetgen E, Martinussen T, Vansteelandt S. Learning Causal Hazard Ratio with Endogeneity. 2018. <http://arxiv.org/abs/1807.05313>.
- Martinussen T, Vansteelandt S. On collapsibility and confounding bias in Cox and Aalen regression models. *Lifetime Data Anal*. 2013;19(3):279–96.
- Kahan B, Jairath V, Doré C, Morris T. The risks and rewards of covariate adjustment in randomized trials: an assessment of 12 outcomes from 8 studies. *Trials*. 2014;15(1):139.
- Hastie T, Tibshirani R. Varying-coefficient models. *J R Stat Soc Ser B Methodol*. 1993;55(4):757–79.
- Abrahamowicz M, Mackenzie T, Esdaile J. Time-dependent hazard ratio: modeling and hypothesis testing with application in lupus nephritis. *J Am Stat Assoc*. 1996;91(436):1432–9.
- Gray R. Flexible methods for analyzing survival data using splines, with applications to breast cancer prognosis. *J Am Stat Assoc*. 1992;87(420):942–51.
- Kalbfleisch J, Prentice R. *The Statistical Analysis of Failure Time Data*, 2nd: John Wiley & Sons; 2002. ISBN 978-0-471-36357-6.
- Lin D. On the Breslow estimator. *Lifetime Data Anal*. 2007;13(4):471–80.
- Cox D. Regression models and life-tables. *J R Stat Soc Ser B Methodol*. 1972;34(2):187–220.
- Columbo J, Martínez-Cambor P, MacKenzie T, Staiger D, Kang R, Goodney P, O'Malley A. Comparing long-term mortality after carotid endarterectomy vs carotid stenting using a novel instrumental variable method for risk adjustment in observational time-to-event data. *JAMA Netw open*. 2018;1(5):181676.
- Martínez-Cambor P, Mackenzie T, Staiger D, Goodney P, O'Malley A. Adjusting for bias introduced by instrumental variable estimation in the Cox proportional hazards model. *Biostat*. 2019;20(1):80–96.
- Huling J, Yu M, O'Malley A. Instrumental variable based estimation under the semiparametric accelerated failure time model. *Biom*. 2019;75(2):516–27.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.