Time Dependent Hazard Ratio Estimation Using Instrumental Variables Without Conditioning on an Omitted Covariate

Todd MacKenzie (todd.mackenzie@dartmouth.edu)
Dartmouth Medical School
Pablo Martinez-Camblor
Dartmouth College Geisel School of Medicine
James O'Malley
Dartmouth College Geisel School of Medicine

Research article

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

Posted Date: November 22nd, 2019

DOI: https://doi.org/10.21203/rs.2.17724/v1

Version of Record: A version of this preprint was published at BMC Medical Research Methodology on March 20th, 2021. See the published version at https://doi.org/10.1186/s12874-021-01245-6.

Time Dependent Hazard Ratio Estimation Using Instrumental Variables

Without Conditioning on an Omitted Covariate

Todd A. MacKenzie (1,2), Pablo Martinez-Camblor (1,2), A. James O'Malley (1,2)

1. Department of Biomedical Data Science, Geisel School of Medicine, Dartmouth College
2. The Dartmouth Institute for Health Policy and Clinical Practice, Geisel School of Medicine, Dartmouth College

Corresponding Author: Todd A. MacKenzie. E-mail: Todd.A.MacKenzie@dartmouth.edu}

Time Dependent Hazard Ratio Estimation Using Instrumental Variables Without Conditioning on an Omitted Covariate

Abstract

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 an 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. We extend IV based estimation of Cox's models 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. Due to the non-collapsibility of the Cox's models these two estimands are not identical. We report the results of simulations that includes 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 demonstrate the method to estimate the time dependent hazard ratio for two vascular interventions.

Keywords: Causal inference; Censoring; Semi-parametric model; Marginal model.

1 Introduction

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 for 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*). Estimators based on instrumental variables are consistent for linear models of the outcome.

Estimation of treatment effects 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 the IV has 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-Camblor 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.

In this paper we extend IV based estimation of the HR beyond proportionality of hazards. We propose a method of estimating the HR as a time-dependent function. We consider 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 times-to-event dependent on omitted covariates that have a marginal structural time dependent hazard ratio semi-parametric model with respect to the exposure of primary interest. We demonstrate this method for time-dependent hazard ratio estimation to compare the effect of two vascular interventions on survival.

2 Derivation of the Estimator

2.1 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, for x in the range of X, e.g. 0 (no treatment) or 1 (treatment). 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

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

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 [7], the corresponding marginal model describing how X impacts the hazard 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: When randomized trials are reported there is never conditioning on covariates that were not measured in the trial, and furthermore, the convention has been not to condition on measured characteristics either [8].

2.2 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 [9], regression splines [10] or penalized regression splines [11] among other techniques.

2.3 Time-to-event, exposure and instrument

Suppose the observed data consists of $\{\Delta_i, T_i, X_i, W_i\}_{i=1}^n$ where Δ_i is the event indicator for subject i, T_i is the time-to-event if $\Delta_i = 1$ and otherwise the censoring time, X_i is the exposure level (either binary, ordinal or continuous) and W_i is the instrument ($1 \leqslant i \leqslant n$).

An instrument is defined by the following properties:

1. It is associated with the exposure ($W \not\perp\!\!\!\perp X$).

2. (a) There is no effect on the outcome except through its effect on the exposure, i.e. the potential outcomes $\{T(x)\}_{\{x=0,1\}} \perp\!\!\!\perp W|U$

   (b) There are no confounders of the instrument and the outcome ($W \perp\!\!\!\perp U$).

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 and U the outcome and instrument are independent. The instrument, W, may be continuous, ordinal or binary.

For each subject there is one *potential* stochastic process $\{N_i(t;x), R_i(t;x)\}_{t\geqslant 0,x}$ for each possible level of x where $R_i(t;x)$ is the *at risk* indicator for the event $T_i < t$, and $N_i(t;x)$ is

the *counting process* equal to the indicator for the event $T_i \geqslant t$. Let $R_i(t) = R_i(t; X)$ and $N_i(t) = N_i(t; X)$.

2.4 Estimating Equations for a Time Dependent Hazard Ratio

Let $\lambda_x(t)$ be the hazard function corresponding to survival in the population of subjects were all subjects 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 causal effect of having been exposed to exposure level x is an increase 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

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

and also known as a Martingale residual (Kalbfleisch and Prentice [12]), 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

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

where τ is a suitably large point of time (e.g. $\max_{1 \leqslant i \leqslant n}\{t_i\}$), K is the dimension of the parameter $\theta = (\theta_1, \cdots, \theta_K)$ 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 also unknown it can be solved using the estimating equations

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

for all t which is solved by the Breslow estimator (**?**])

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

Substitution of (4) into (2) yields the score equations for the partial likelihood (Cox [13]) for a time-dependent hazard ratio (Abrahamowicz et al. [10]). 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

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

and therefore the estimating equations (2) are biased. We postulate that if W is an instrumental variable then it is uncorrelated with the counting process because the IV is independent of any omitted confounders. A heuristic justification starts by by writing the causal hazard function conditional on an omitted covariate U that affects both the treatment and the time-to-event. Next we approximate the causal hazard function by a linear Taylor expansion, $\mathrm{E}[dN_i(t)|R_i(t) = 1] = \exp\{\beta_X(v;\theta) \cdot X_i\} d\Lambda_0(v) + \phi(t)[U - \mu_U(t)]$ 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,

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

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

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

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].

2.5 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 $0 = \tau_0 < \tau_1 < \ldots < \tau_K = \tau$ and within each window the hazard ratio is assumed to be constant, $\beta_X(t;\theta) = \theta_i$ for $\tau_{i-1} \leqslant t < \tau_i$. 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 $(1 \leqslant i \leqslant K)$. 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/toddamackenzie/Instrumental-Variable-Hazard-Ratio-Estimation`).

2.6 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

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

and

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

respectively.

3 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. For the log-linear function of time we consider a range of intercepts and a range of slopes. Specific details follow.

3.1 Simulation Methods

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 with respect to the treatment is a Cox model. The steps of the simulation are:

1. Randomly generate a continuous omitted covariate, U, from the standard normal distribution.

2. Randomly generate a continuous instrument, W, from the standard normal distribution.

3. Randomly generate a binary exposure indicator X using a logistic link with intercept of zero and odds ratios, o_W and o_U, for the association of W and U respectively.

4. Generate a random bivariate (B_0, B_1) from a Gaussian copula for which the correlation is through the omitted covariate U, i.e., $B_i = \rho \cdot U + \sqrt{1-\rho^2} \cdot E_i$ for $i = 0, 1$ where E_0, E_1 are independent standard normal deviates.

5. Obtain potential time-to-events $T(0)$ and $T(1)$ corresponding to non exposure and exposure respectively, by transforming B_0 and B_1 by the functions $\Lambda_0^{-1}(-\log(\cdot))$ and $\Lambda_1^{-1}(-\log(\cdot))$ respectively where $\Lambda_1(t) = \int_0^t \exp\{\beta(u)\} d\Lambda_0(u)$. We use $\Lambda_0(t) = t$.

6. Right censor the times-to-event using a uniform distribution whose scale is set to achieve the specified censoring rate.

In this Monte Carlo simulation U has a causal effect on the time-to-event but we are not interested in estimating its effect. Our interest lies in the marginal distribution of the time-to-event for the two exposure levels, and in particular the time-dependent hazard ratio relating these two distributions. The strength of the association between the omitted covariate and the potential time-to-events is controlled by the parameter ρ of step 4 which in these simulations is set at 0.5 so that as U increases the hazard increases. The level of confounding by the omitted covariate was controlled by the odds ratio o_U relating U to X. It was set at (i) 5 for positive confounding, (ii) 1 for no confounding and (iii) 1/5 for negative confounding.

The stepwise constant time-dependent hazard ratio took values from the range, $\{1/3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1, 1.1, 1.2, 1.5, 1.7, 2, 2.5, 3\}$. 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 using the range $\{-1.0, -0.9, -0.8, \ldots, 0.8, 0.9, 1.0\}$ for the intercept and the range $\{-0.50, -0.45, -0.40, \ldots, 0.40, 0.45, 0.5\}$ for the slope. For each simulated data set we randomly sampled with replacement from the intercept and slope ranges.

In all simulations we set the sample size at 1000 and censoring frequency at 50%. Complete `R` code is available at (`https://github.com/toddamackenzie/Instrumental-Variable-Hazard-R`) for users interested in conducting the same simulations.

3.2 Simulation Results for Early, Mid and Late Estimators of the HR

Figure 1 shows the results of the simulations. 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 blue point represents the median of over 1000 estimates using the instrument and estimating equations we propose. The instrument 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.

3.3 Simulation Results for Log-Linear Time Dependent HR

Figure 2 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.

4 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 Initative, 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 consider by Columbo et al. [14]. Figure 3 (top) shows the Kaplan-Meier estimates for 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. The validity of this instrument is argued in Martínez-Camblor et al. [15]. Figure 3, bottom, depicts the instrument distribution (histograms at left and box-plot at right) in both groups.

Figure 4 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 and 6 months. In the first month the partial likelihood 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 estimator based on the instrumental variable 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 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 augmenting the observed effect of the treatment.

5 Summary

We have proposed a method for using instruments in the estimation of time-dependent hazard ratio. 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 effect in randomized studies, the model we employ does not explicitly condition on the omitted covariate. 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 disadvange of the conditional model is that it conditions on intangible characteristics, that is unmeasured characteristics, that somehow affected the treatment selection.

Simulations indicated the estimators have 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 to multi-variable models without much additional complexity. That is, measured covariates can be incorporated into the model.

6 Declarations

-

Ethics approval and consent to participate: The data utilized was deidentified.

Consent for publication: The co-aothors consent to publication.

Competing interests: There are no conflicts of interest to report.

Funding:

Authors' contributions: All three authors contibuted to the concept, writing and statistical analysis.

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

References

[1] Thérèse A Stukel, Elliott S Fisher, David E Wennberg, David A Alter, Daniel J Gottlieb, and Marian J Vermeulen. 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*, 297(3):278–285, 2007.

[2] Todd A MacKenzie, Tor D Tosteson, Nancy E Morden, Therese A Stukel, and A James

O'Malley. Using instrumental variables to estimate a Cox's proportional hazards regression subject to additive confounding. *Health Services and Outcomes Research Methodology*, 14(1-2):54–68, 2014.

[3] Eric J Tchetgen Tchetgen, Stefan Walter, Stijn Vansteelandt, Torben Martinussen, and Maria Glymour. Instrumental variable estimation in a survival context. *Epidemiology (Cambridge, Mass.)*, 26(3):402, 2015.

[4] Jialiang Li, Jason Fine, and Alan Brookhart. Instrumental variable additive hazards models. *Biometrics*, 71(1):122–130, 2015.

[5] Torben Martinussen, Stijn Vansteelandt, Eric J Tchetgen Tchetgen, and David M Zucker. Instrumental variables estimation of exposure effects on a time-to-event endpoint using structural cumulative survival models. *Biometrics*, 73(4):1140–1149, 2017.

[6] Pablo Martínez-Camblor, Todd A MacKenzie, Douglas O Staiger, Phillip P Goodney, and A James O'Malley. An instrumental variable procedure for estimating Cox models with non-proportional hazards in the presence of unmeasured confounding. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 2019.

[7] Torben Martinussen and Stijn Vansteelandt. On collapsibility and confounding bias in Cox and Aalen regression models. *Lifetime data analysis*, 19(3):279–296, 2013.

[8] Brennan C Kahan, Vipul Jairath, Caroline J Doré, and Tim P Morris. The risks and rewards of covariate adjustment in randomized trials: an assessment of 12 outcomes from 8 studies. *Trials*, 15(1):139, 2014.

[9] Trevor Hastie and Robert Tibshirani. Varying-coefficient models. *Journal of the Royal Statistical Society: Series B (Methodological)*, 55(4):757–779, 1993.

[10] Michal Abrahamowicz, Todd Mackenzie, and John M Esdaile. Time-dependent hazard ratio: modeling and hypothesis testing with application in lupus nephritis. *Journal of the American Statistical Association*, 91(436):1432–1439, 1996.

[11] Robert J Gray. Flexible methods for analyzing survival data using splines, with applications to breast cancer prognosis. *Journal of the American Statistical Association*, 87(420):942–951, 1992.

[12] J. D. Kalbfleisch and R. L. Prentice. *The Statistical Analysis of Failure Time Data*. John Wiley & Sons, 2nd edition, 2002.

[13] D. R. Cox. Regression models and life-tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, 34(2):187–220, 1972.

[14] Jesse A Columbo, Pablo Martínez-Camblor, Todd A MacKenzie, Douglas O Staiger, Ravinder Kang, Philip P Goodney, and A James O'Malley. 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 network open*, 1(5):e181676–e181676, 2018.

[15] Pablo Martínez-Camblor, Todd A. Mackenzie, Douglas O. Staiger, Philip P. Goodney, and A. James O'Malley. Adjusting for bias introduced by instrumental variable estimation in the Cox proportional hazards model. *Biostatistics*, 20(1):80–96, 2019.

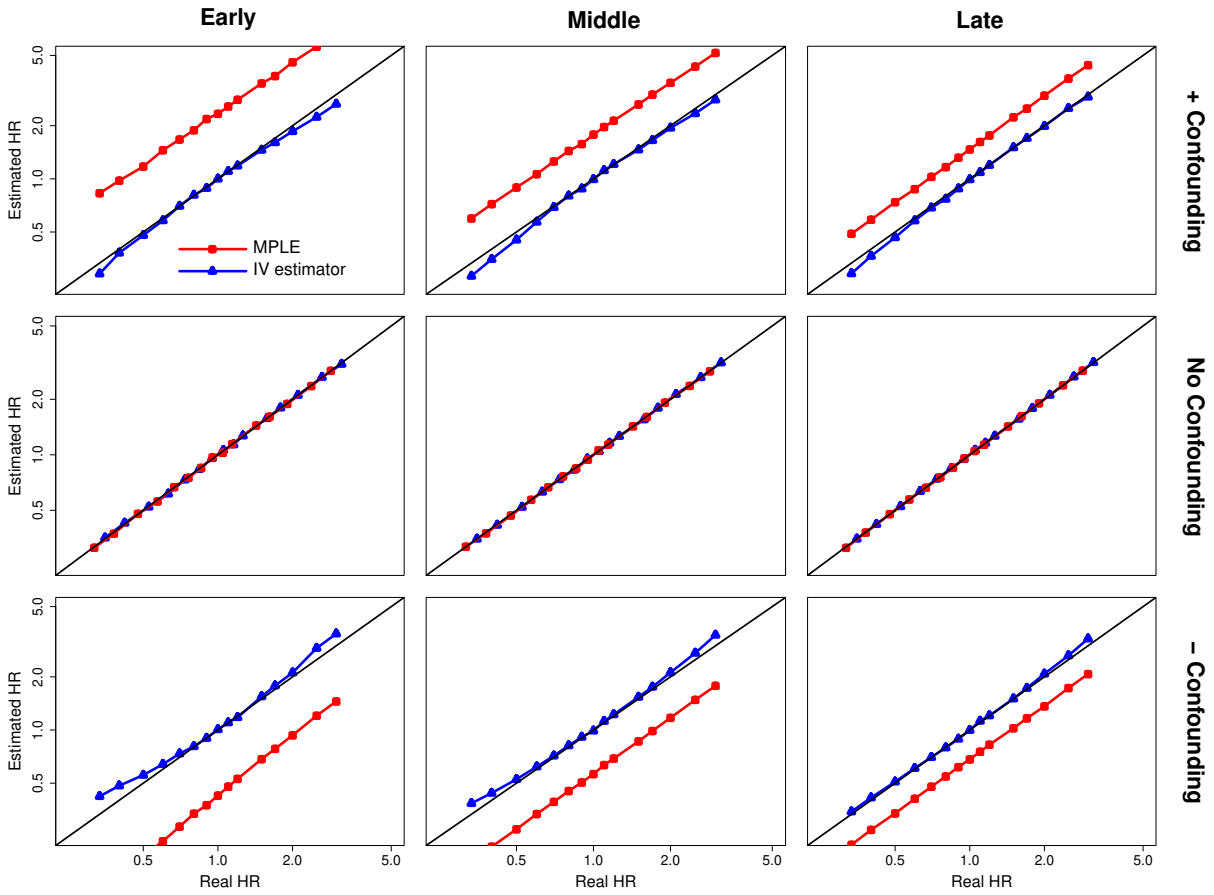

Figure 1: **Simulation Results** for estimating three-piece Time-dependent hazard ratio: The nine panes represent each combination of early, mid and late follow-up, with (i) strong positive confounding, (ii) no confounding and (iii) strong negative confounding. Each pane overlays scatterplots of the estimated hazard ratio versus the true hazard ratio, our IV based estimator in blue, and the standard Cox MPLE in red. Each point represents the median of over 1000 Monte Carlo iterations.

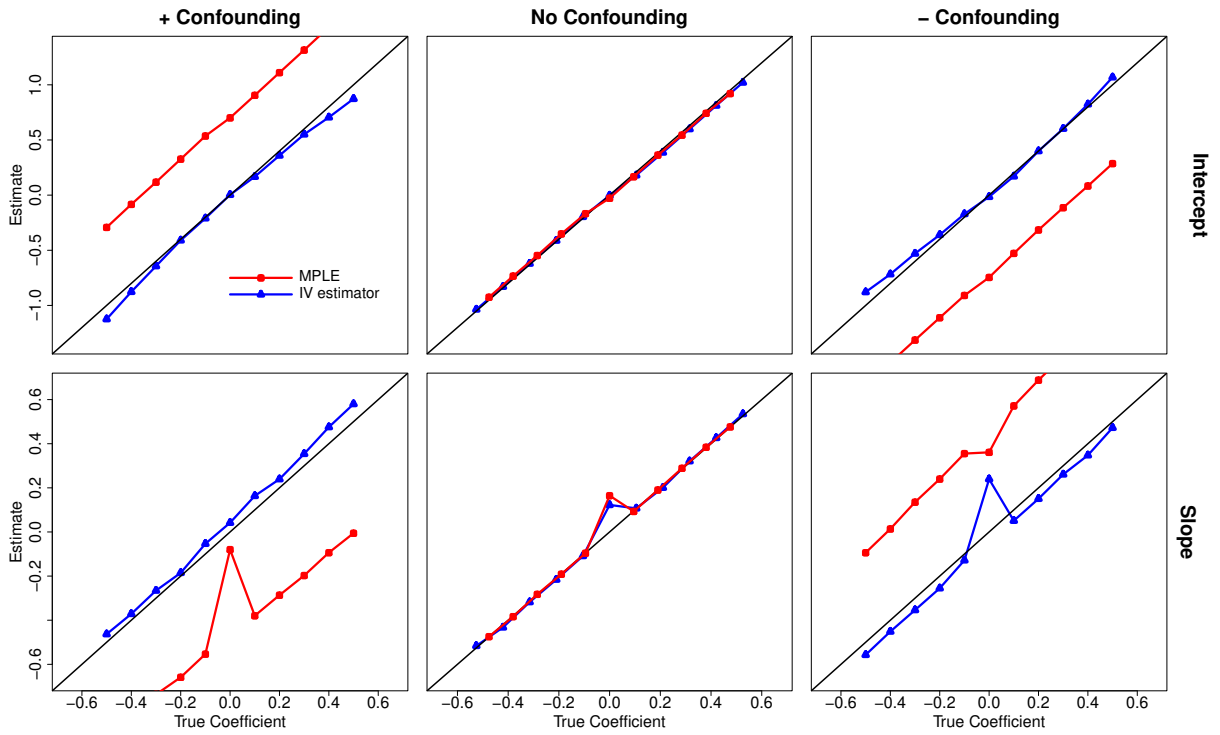

Figure 2: **Simulation Results** for Estimating a Log-Linear Time Dependent Hazard Ratio: The top 3 panes represent estimation of the intercept parameter of the log-linear model for (i) strong positive confounding, (ii) no confounding and (iii) strong negative confounding, while the bottom three panes represent results for estimation of the slope parameter. Each pane overlays scatterplots of the estimated hazard ratio versus the true hazard ratio, 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)

.

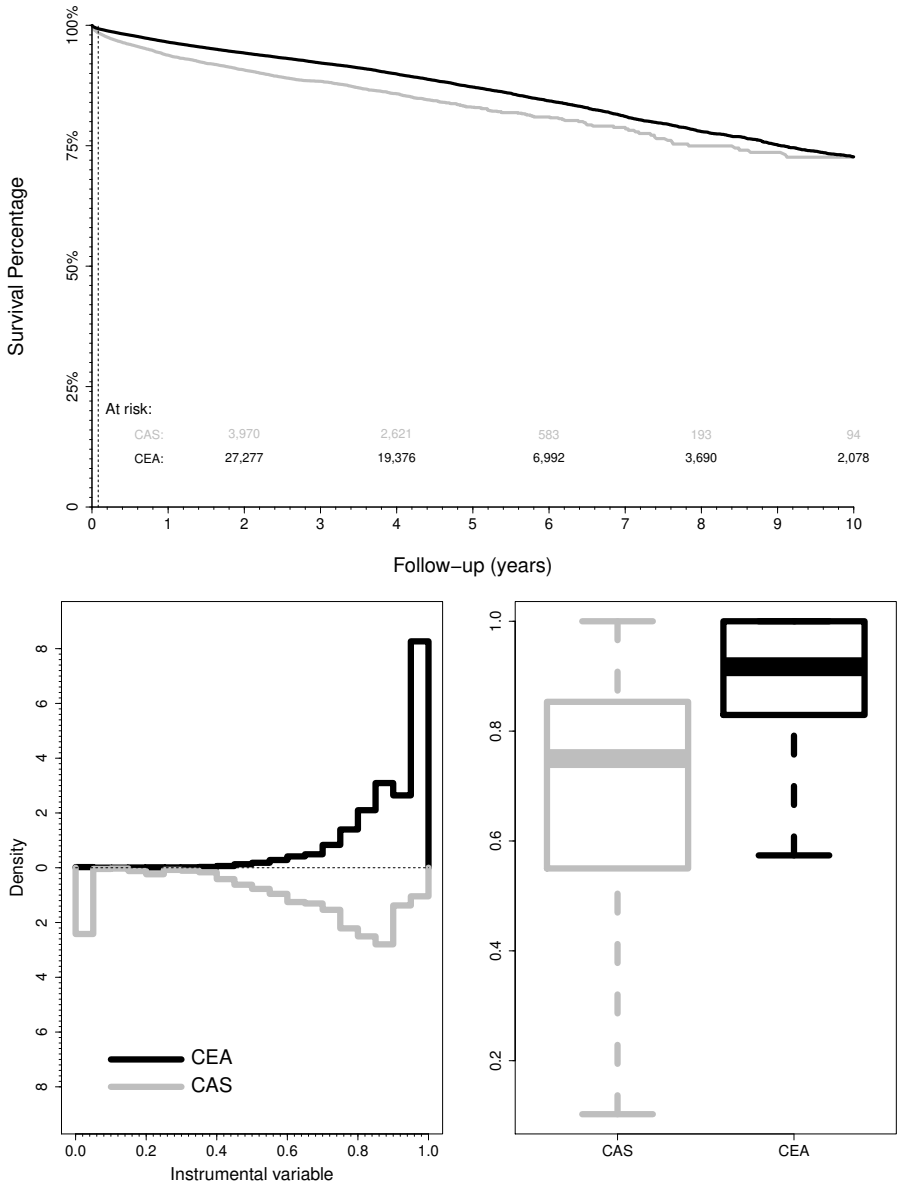

Figure 3: Kaplan-Meier estimates (top) and instrument distributions (bottom) for both the CEA and the CAS groups

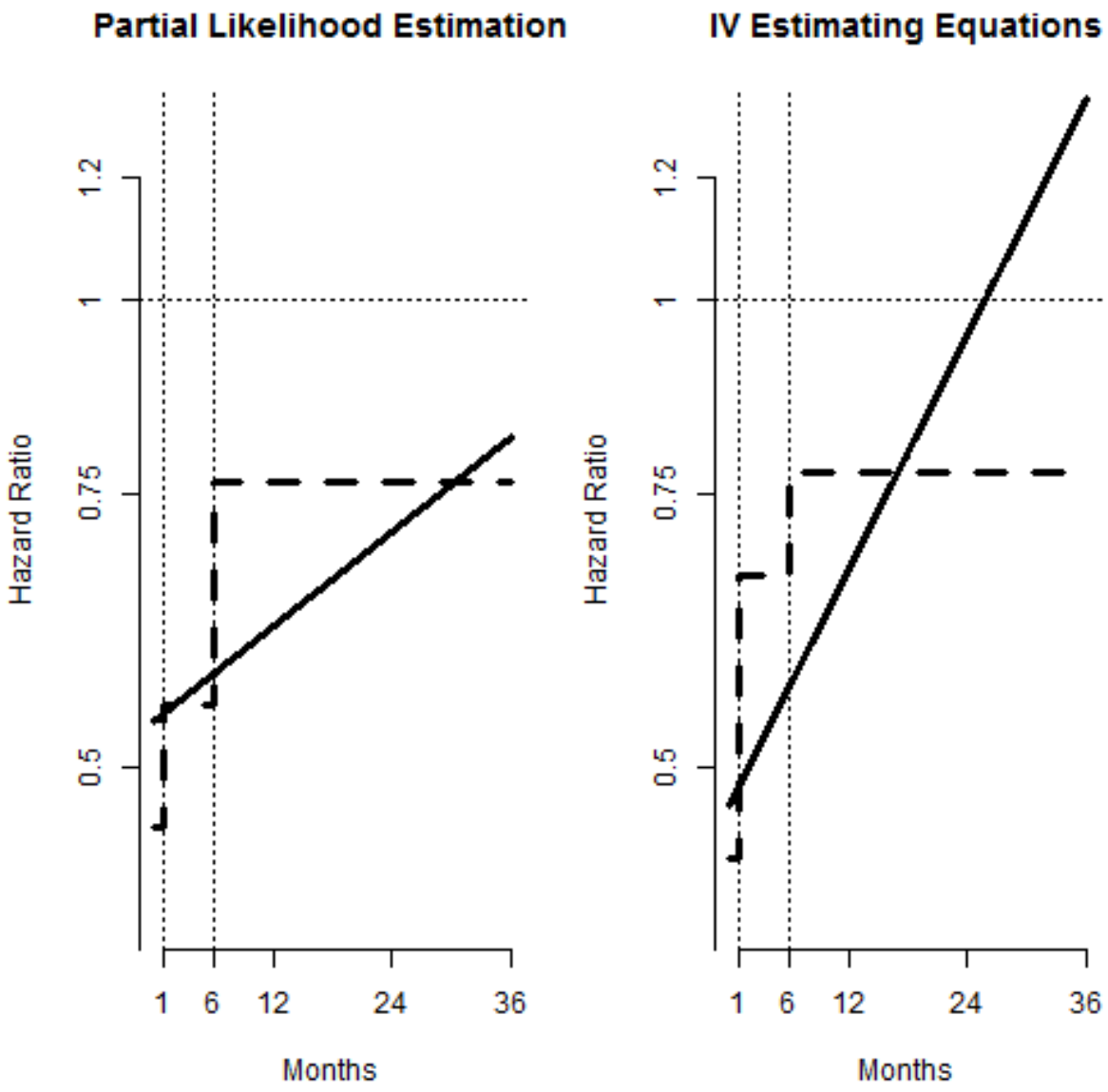

Figure 4: **Hazard Ratios** Comparing Risk of Stroke or Death Between CSA and CEA: The left pane shows partial likelihood estimators of a 3 piece time-dependent constant hazard ratio (steps at 1 and 6 months, in a dashed line) and a log-linear time-dependent hazard ratio (solid line). The pane on the right shows the corresponding estimators based on instrumental variables that we propose.

.