Instrumental variables estimation in the Cox Proportional Hazard regression model

James O'Malley, Ph.D.
Department of Biomedical Data Science
The Dartmouth Institute for Health Policy and Clinical Practice
Lebanon, NH
James.OMalley@Dartmouth.edu

Acknowledgements:
P. Martinez-Camblor, T.A. MacKenzie, D.O. Staiger and P.P. Goodney
Patient-Centered Outcomes Research Institute (PCORI) Award ME-1503-28261

Motivation: Comparative Effectiveness in Vascular Surgery

- Condition: Carotid artery disease, risk factor for stroke and death
- Treatments: Carotid endarterectomy (CEA) vs. carotid stenting (CAS)
- Outcome: Time until death (any cause)
- Setting: Daily clinical practice
- Modeling requirement: Use Cox proportional hazards regression model due to the limited assumptions it makes and its popularity

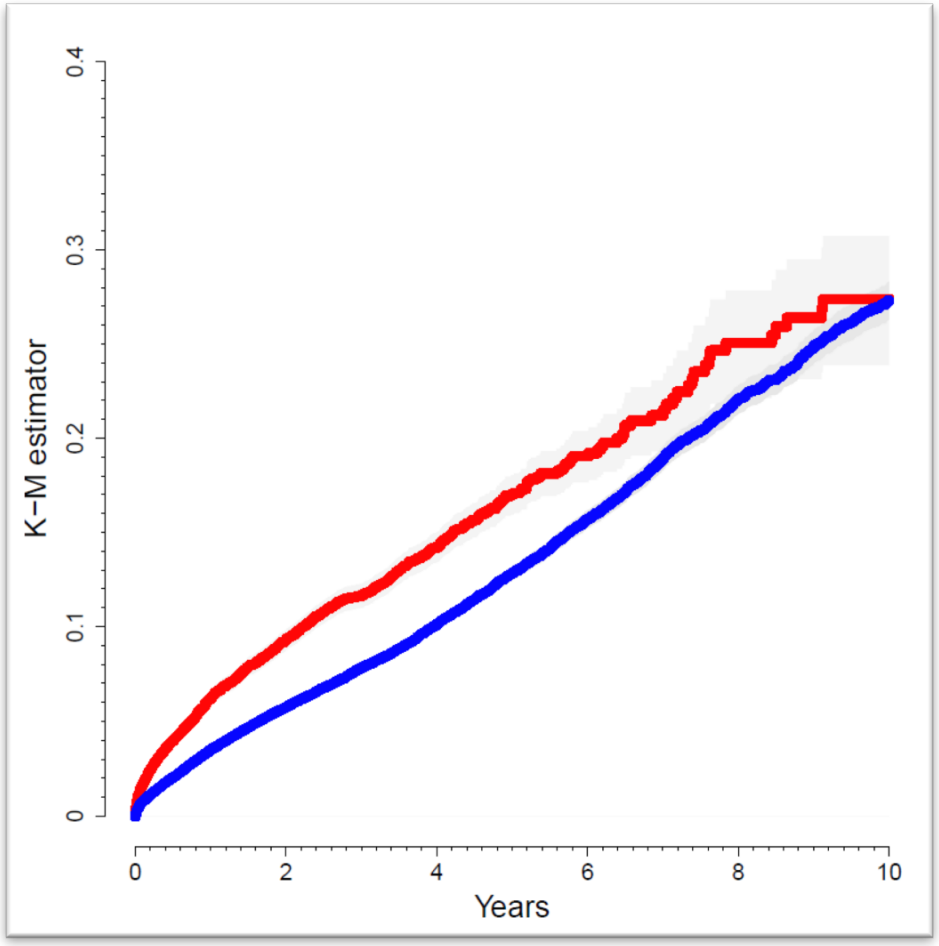
Observational Study

- Detailed clinical data from the Vascular Quality Initiative (VQI) linked to long-term outcome data from Medicare
- www.vascularqualityinitiative.org

- *A total of 86,017 patients (between 2013-2016) contributing* 259,700.2 patient-years (follow-up 3.02±2.36)

	CEA N=73,312	CAS N=12,705
Age	70.3±9.4	69.1±10.4
Male (%)	44,191 (60.4)	8,117 (63.9)

Raw Outcome Data: Kaplan Meier Estimates



		Events/Exposure	*Hazard-ratio of CEA to CAS (95% CI)*
Overall			
	CEA	*6,600/ 73,312*	*0.67 (0.64-0.71)*
	CAS	*1,405/12,705*	
Symptomatic			
	CEA	*2,559/28,689*	*0.61 (0.56-0.66)*
	CAS	*786/6,825*	
Asymptomatic			
	CEA	*4,017/44,395*	*0.76 (0.70-0.83)*
	CAS	*607/5,809*	

Can we infer causality?

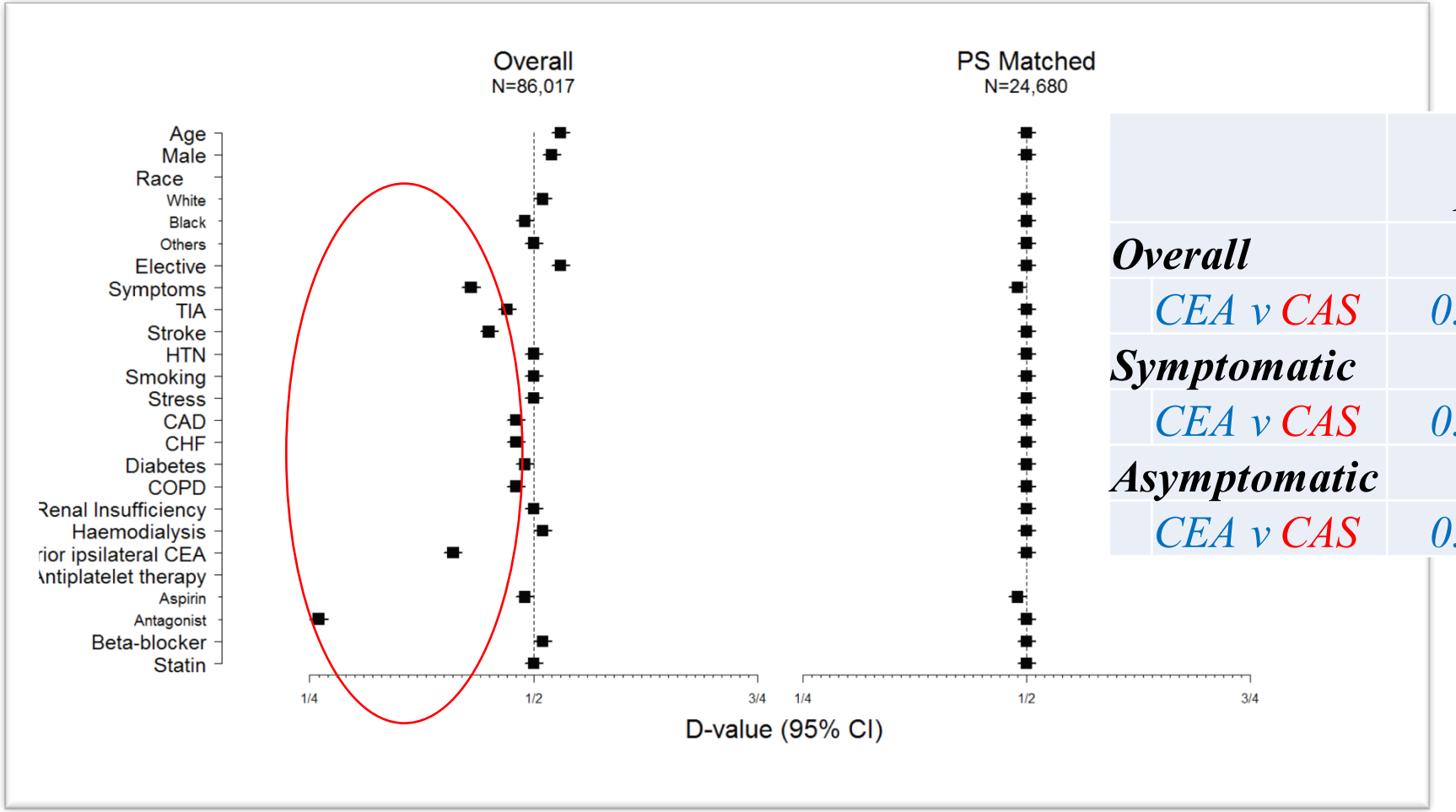
Was treatment decided at random?

T=1 [CEA]

T=0 [CAS]

Adjusted Analyses of VQI data

- Well known methods: Regression adjustment and propensity score analysis



	Adjusted HR (95% CI)	*PS matched HR (95% CI)*
Overall		
CEA v CAS	*0.69 (0.65-0.74)*	*0.71 (0.65-0.77)*
Symptomatic		
CEA v CAS	*0.61 (0.56-0.67)*	*0.59 (0.53-0.66)*
Asymptomatic		
CEA v CAS	*0.79 (0.72-0.87)*	*0.79 (0.71-0.90)*

Unmeasured Confounding!

Estimates based on regression and propensity score methods still only yield ***ASSOCIATIONS***

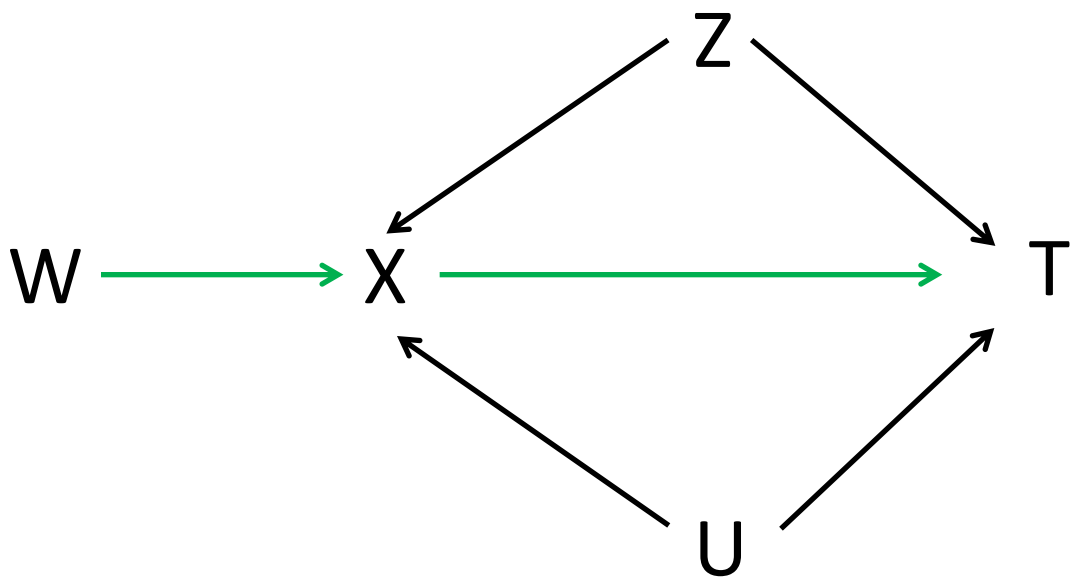
Because,

What happen with the UNMEASURED confounders ???!!

Accounting for Unmeasured Confounding: Instrumental Variables

- Key analytic variables:
  - T = survival time
  - X = treatment (=1 if CEA, 0 if CAS)
  - Z = observed confounders
  - U = unobserved confounders
- W is an instrumental variable (IV) if:
  1. W predicts X conditional on (Z, U)
  2. W is independent of U conditional on Z
  3. W is independent of T conditional on (X, Z, U)
- Use Directed Acyclic Graphs (DAGs) to test IV conditions (Brito and Pearl 2002)

Directed Acyclic Graph (DAG) for Instrumental variable analysis with no censoring



T = (Transformed) survival time

X = Treatment

W = Instrumental variable (IV)

- Extracts good variation in X; avoids conditioning on X

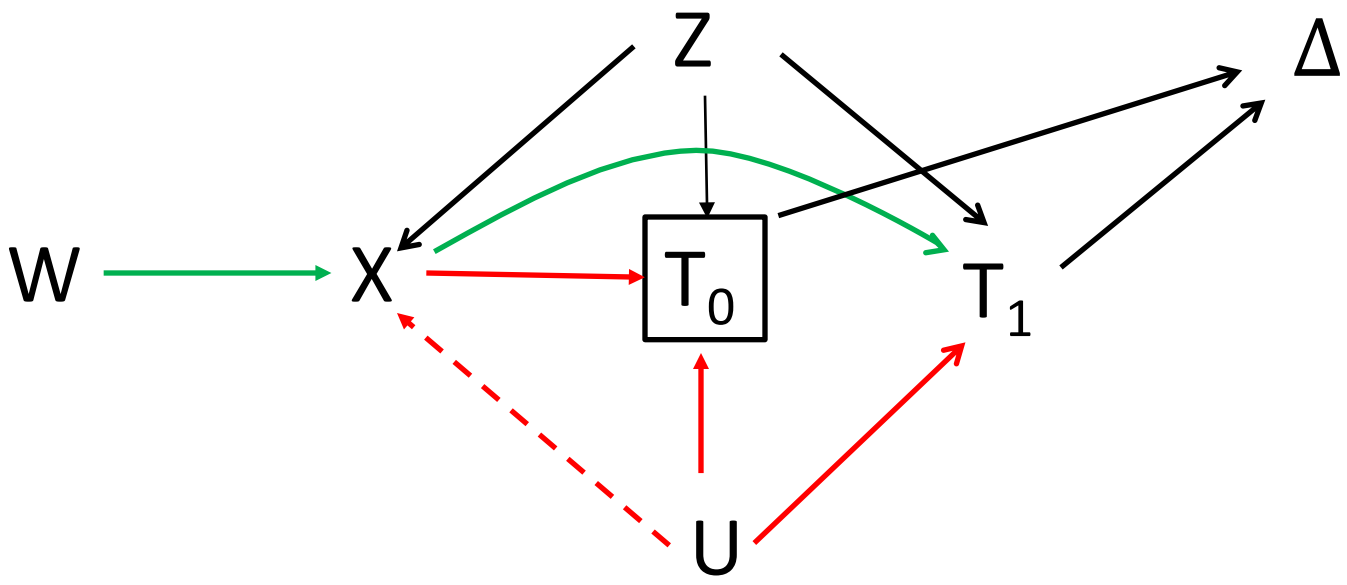
Z = Observed confounders

U = Unobserved confounders

Accounting for censoring and hazard ratios!

- The survival time T may only be known to exceed a given value, denoted C
  - C is referred to as the censoring time
- Observe $T_{obs} = \min(T, C)$ and $\Delta = I(T \leq C)$
- Cannot condition on Δ to account for censoring!
- Furthermore, at-risk sample becomes increasingly selective with follow-up time (Hernan et al 2004)
  - (W, U) associated after conditioning on Z even if X has no effect on T
- We need an IV procedure that works with the Cox regression model

The Cox survival model and its estimation of Hazard-Ratios add further twists!



T = (Transformed) survival time

X = Treatment

Z = Observed confounders

U = Unobserved covariates; not necessarily confounders

W = Instrumental variable

Δ = Whether observe survival time or censoring time

Methodological solutions

- Ignore unmeasured confounding
  - Sensitivity analysis to unmeasured confounder (e.g., Tchetgen Tchetgen and Robins 2012)
- Adaptation of existing IV methods to survival data
  - Two-stage least squares (Stukel et al 2007)
  - Two-stage predictor substitution
  - Two-stage residual inclusion (Terza et al 2008, Cai et al 2011, Gore et al 2011, Palmer 2013)
- Structural equation models (SEM) involving full parametric models (Choi and O'Malley 2015)

Cox Proportional Hazards (CPH) Model with Treatment selection

- The CPH survival time model is specified in terms of the instantaneous change in the survival probability across time:

$$\lambda(t) = -\frac{d}{dt} log(\Pr(T > t|.)),$$

where $\Pr(T > t|.)$ is the probability of surviving to time t

- Survival time model:

$$\lambda(t|X_i, \boldsymbol{Z}_i, U_i) = \lambda_0(t) \exp(\beta_X X_i + \boldsymbol{\beta}_Z^T \boldsymbol{Z}_i + \beta_U U_i)$$

- Treatment selection model:

$$X_i = \alpha_0 + \boldsymbol{\alpha}_W^T \boldsymbol{W}_i + \boldsymbol{\alpha}_Z^T \boldsymbol{Z}_i + \alpha_U U_i + \epsilon_i$$

Proposed IV Procedure for Cox Model

- Use ordinary least squares to compute the fitted values:

$$\widehat{X}_i = \hat{\alpha}_0 + \hat{\alpha}_W^T W_i + \hat{\alpha}_Z^T Z_i$$

and compute

$$\hat{R}_i = X_i - \hat{X}_i$$
$$= (\alpha_0 - \hat{\alpha}_0) + (\alpha_W - \hat{\alpha}_W)^T W_i + (\alpha_Z - \hat{\alpha}_Z)^T Z_i + \alpha_U U_i + \varepsilon_i$$
$$\rightarrow \alpha_U U_i + \varepsilon_i = R_i \text{ in expectation}$$

- Hints:
  - Want to control for $\hat{R}_i$
  - Might better control for U_i if can separately account for ε_i

Theory

- Under the Cox model

$$Pr(T_i \geq t|X_i, \boldsymbol{Z}_i, U_i) = exp\{-\Lambda_0(t)\, exp(\beta_X X_i + \boldsymbol{\beta}_Z^T \boldsymbol{Z}_i + \beta_U U_i)\}$$

  where $\Lambda_0(t) = \int_0^t \lambda_0(t)dt$

- But if $R_i = \alpha_U U_i + \varepsilon_i$ it follows that $U_i = \alpha_U^{-1}(R_i - \varepsilon_i)$ and

$$Pr(T_i \geq t|X_i, R_i, \varepsilon_i) = exp\{-\Lambda_0(t)\phi_i\, exp(\beta_X X_i + \boldsymbol{\beta}_Z^T \boldsymbol{Z}_i + \beta_U \alpha_U^{-1} R_i)\}$$

  where $\phi_i = \exp(-\beta_U \alpha_U^{-1} \varepsilon_i)$

- Control for R_i (via $\hat{R}_i$) and separately account for ε_i by including an individual frailty with an unrestricted scale to absorb the impact of $\exp(-\beta_U \alpha_U^{-1} \varepsilon_i)$

Second-stage procedure

- After computing $\hat{R}_i$ estimate the Cox survival time model with an individual frailty given by:

$$\lambda\left(t\middle|X_i, \boldsymbol{Z}_i, \hat{R}_i\right) = \lambda_0(t)\phi_i \exp\left(\beta_X X_i + \boldsymbol{\beta}_Z^T \boldsymbol{Z}_i + \beta_R \hat{R}_i\right)$$

- Under standard regularity conditions this two-stage procedure yields a consistent estimator of β_X
- The frailty makes the adjustment for $\hat{R}_i$ more closely aligned with adjusting for U_i
- Procedure generalizes the control function approach known as two-stage residual inclusion (2SRI) popularized by Terza et al (2008)
  - Acronym for procedure is 2SRI-F (F for Frailty)
- But need to specify the distribution of the individual frailty $\phi_i \sim f(.)$
  - Most commonly $f(.)$ is assumed to be a common distribution such as the log-normal or the gamma with unknown scale parameter

Illustrative Monte Carlo Simulations

Time: Weibull(1,2)

Expected censorship: 20%

Treatment model:

$$X = W + Z + U + \epsilon$$

$W \sim N(0,1)$

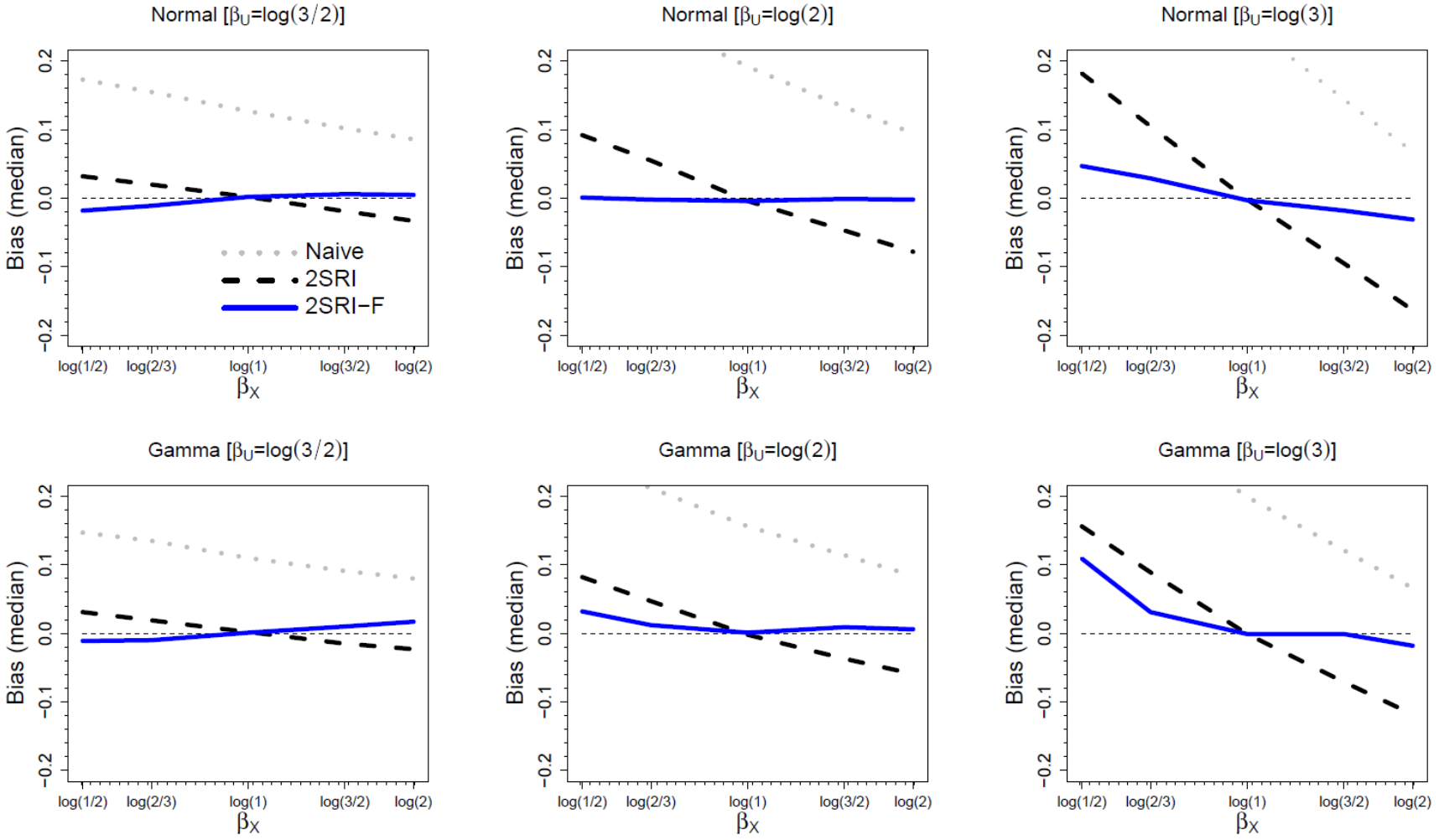
Z [measured] $\sim N(0,1)$

ϵ [unknown] $\sim N(0,1)$

U [unknown] $\sim N(0,1), \Gamma(1,1) - 1$

Survival model (risk):

$$\lambda(t) \cdot \exp\{\beta_X \cdot X + Z + \beta_U \cdot U\}$$



Analysis of the VQI data

Estimate the **causal effect** of endarterectomy (CEA) vs. carotid stenting (CAS) on the time to death (any reason) of patients suffering from carotid artery disease who are treated in regular clinical practice

Instrumental Variable for VQI Data

Our IV (W) is the proportion of CEA procedures out the total number of surgeries **[CEA/(CEA+CAS)]** performed in the same hospital over the 12-months prior to the current surgery

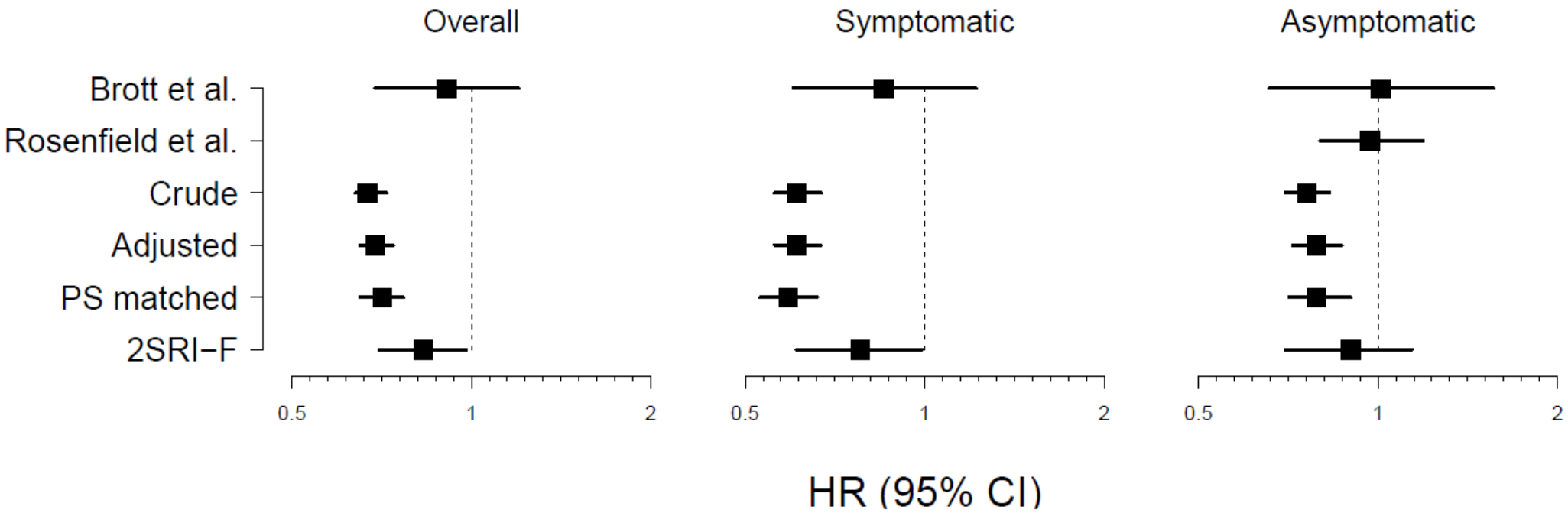
Justification?

1. W is strongly associated with treatment
2. W is independent of observed patient characteristics so might also be independent of unobserved confounders (?)
3. The relationship of W with time to death is only through the received procedure if any institutional learning/improvement is general as opposed to procedure specific (?)

Results: VQI data analysis

	Raw (unadjusted) HR (95% CI)	*Adjusted HR (95% CI)*	*PS matched HR (95% CI)*	*2SRI-F HR (95% CI)*
Overall				
CEA vs CAS	*0.67 (0.64-0.71)*	*0.69 (0.65-0.74)*	*0.71 (0.65-0.77)*	*0.83 (0.70-0.98)*
Symptomatic				
CEA vs CAS	*0.61 (0.56-0.66)*	*0.61 (0.56-0.67)*	*0.59 (0.53-0.66)*	*0.78 (0.61-0.99)*
Asymptomatic				
CEA vs CAS	*0.76 (0.70-0.83)*	*0.79 (0.72-0.87)*	*0.79 (0.71-0.90)*	*0.90 (0.70-1.14)*

Results continued: Stratified Analyses



Forest-plots for the HRs obtained for two RCTs and the different models considered for the VQI data analyses

Conclusions

1. 2SRI-F: Consistent under theoretical assumptions
2. Frailty term improves the results but does not remove "all" the bias in finite samples
   - R Cox-Frailty model routines are approximate
3. Easy to compute
4. Robust with respect the frailty distribution
5. VQI analysis shows coherence between 2SRI-F and RCTs
6. Methodology has been extended to allow:
   - An interaction effect involving the treatment
   - Non-proportional hazards

References

- Martinez-Camblor P, MacKenzie TA, Staiger D, Goodney PP, and O'Malley AJ. Adjusting for bias introduced by instrumental variable estimation in the Cox Proportional Hazards Model. Biostatistics, 2017. doi: 10.1093/biostatistics/kxx062. PMID: 29267847
- Choi J, O'Malley AJ. Estimating the Causal Effect of Treatment in Observational Studies with Survival Time Endpoints and Unmeasured Confounding. Journal of the Royal Statistical Society, Series C (Applied Statistics), 2017, 66, 159-185
- Columbo JA, Martinez-Camblor P, MacKenzie TA, Staiger DO, Kang R, Goodney PP, O'Malley AJ. 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. In Press: JAMA Network Open
- Martinez-Camblor P, MacKenzie TA, Staiger D, Goodney PP, and O'Malley AJ. An instrumental variable procedure for estimating Cox models with non-proportional hazards in the presence of unmeasured confounding. Under revision

Monte Carlo simulations: Binary treatment

Time: Weibull(1,2)

Expected censorship: 20%

Treatment model (binary):

$$X = I(W + Z + U + \varepsilon > 0)$$

$W \sim N(0,1)$

Z [measured] $\sim N(0,1)$

ϵ [unknown] $\sim N(0,1)$

U [unknown] $\sim N(0,1), \Gamma(1,1) - 1$

V [unknown] $\sim N(0,1), \Gamma(1,1) - 1$

$\mathrm{Cov}(U,V) = 0.5$

Survival model (risk):

$$\lambda(t) \cdot \exp\{\beta_X \cdot X + Z + \beta_U \cdot U\}$$