Collider-stratification Bias Due to Censoring in Prospective Cohort Studies

To the Editor:

Collider stratification was described and illustrated using directed acyclic graphs in Greenland et al.,[1] and subsequently proposed as the structural basis of selection bias in Hernán et al.[2] Directed acyclic graphs representing structural biases do not provide insights regarding magnitude.[3] Recently, magnitudes of bias due to collider stratification have been considered.[3,4] Greenland[5] evaluated bias from statistical adjustment when one is unsure about the role of a variable in a causal system. Results of this theoretical work suggested small biases from collider-stratification—smaller than those expected by omitting a confounder.[5] Recent empirical assessment of conditioning on colliders reached similar conclusions, with moderate bias present only under extreme scenarios linking the collider to exposure and outcome.[6]

The notion of collider stratification as the root of selection bias raised a provocative notion that associations estimated in cohort studies could be biased due to factors related to selection and censoring. Pizzi et al.[7] used simulations to evaluate potential magnitude of collider-stratification bias from selection in cohort studies, when stratification occurs because only a single stratum is available for analysis. Under conditions representing a range of selection processes, very limited biases were observed.

In this letter, we consider selection bias due to censoring in cohort studies. We use structural equations to generate conditional probabilities for a range of scenarios to explore expected bias due to censoring unaffected by sampling variability as would occur with simulations. These equations allow specification of interactions (e.g., when exposure effects on censoring probability depend on an unmeasured risk factor).

Specifically, we consider dichotomous censoring that may be affected by exposure (E) fixed at baseline, an unmeasured risk factor (U), and an interaction of E and U, with an additive probability,

$$\Pr(C) = \alpha_0 + [\alpha_E \times \Pr(E)] + [\alpha_U \times \Pr(U)] + [\alpha_{EU} \times \Pr(E) \times \Pr(U)]$$

and an additive outcome probability,

$$\Pr(Y) = \beta_0 + [\beta_U \times \Pr(U)] + [\beta_E \times \Pr(E)]$$

where parameters represent differences in probability. Particularly with regard to interactions, additive models may better reflect biological effects.[8] Given no confounding, the causal effect is expected to be equal to the unconditional association; however, the causal effect may be unrecoverable when estimation is based only on uncensored data.

The Table presents scenarios where the exposure effect on outcome is null. Scenarios not meeting the constant risk difference (RD) assumption (i.e., outside bounds 0, 1) were dropped. Details on calculations and data generation are available in the eAppendix (http://links.lww.com/EDE/B3). Unsurprisingly, biases increased with increasing total censoring, although in the absence of an interaction of E and U on censoring probability these biases were very small, especially to the RD. We observed larger biases when interactions exist such that exposure effects on censoring depend upon unmeasured risk factors, especially to ratio measures; biases may result even when the proportion censored is only minimally different comparing exposed and unexposed. Although not shown here, if censoring/selection is impacted directly by outcome, as may occur in case–control studies, and collider stratification occurs even absent unmeasured risk factors, larger biases have been described.[5]

It is well established that bias will occur in only one stratum under what has been called a "multiplicative survival model."[2] Specifically, if

$$\frac{\Pr(C=0|E=1,U=1}{\Pr(C=0|E=0,U=1} = \frac{\Pr(C=0|E=1,U=0}{\Pr(C=0|E=0,U=0}$$

bias will occur only in the censored and observed study results will be unaffected. Whereas if

$$\frac{\Pr(C=1|E=1,U=1}{\Pr(C=1|E=0,U=1} = \frac{\Pr(C=1|E=1,U=0}{\Pr(C=1|E=0,U=0}$$

bias occurs in the uncensored. Given a censoring model with an interaction on the additive scale, specific scenarios where interaction is present will correspond to the multiplicative survival model such that no bias is expected in one stratum as described above and shown in the table. Additional scenarios can be explored in the macro provided.

Our findings here are consistent with prior studies observing minimal bias from collider stratification related to adjustment and to selection. However, we note that larger biases are possible in the presence of interactions. Notably, we have considered censoring without regard to timing; more complex causal scenarios than those considered here are possible. The SAS macro provided in the eAppendix (http://links.lww.com/EDE/B3) admits additional effects, allowing researchers engaged in cohort studies to perform their own sensitivity analyses when dropout is observed to occur differentially between exposed and unexposed, and helps provide context to results with respect to selection bias of this nature.

Brian W. Whitcomb
Department of Biostatistics
and Epidemiology

Copyright © 2015 Wolters Kluwer Health, Inc. All rights reserved.
ISSN: 1044-3983/16/2702-00e4
DOI: 10.1097/EDE.0000000000000432

Copyright © 2015 Wolters Kluwer Health, Inc. Unauthorized reproduction of this article is prohibited.

TABLE. Structural Selection Bias Due to Censoring: Proportion Censored, Risk of Outcome, and Observed Measures of Association Under Varying Effects of Exposure (*E*), Unmeasured Risk Factor (*U*), and *E* + *U* Interaction on Censoring

Effects on Censoring[a]			Censoring Probabilities by *E*, *U*				Exposure–Outcome Effect Measures[b]			
			U = 1		*U = 0*		Observed		Censored	
E = 1	*U = 1*	*(E × U) = 1*	*E = 1*	*E = 0*	*E = 1*	*E = 0*	RD	RR	RD	RR
0.05	0.05	0	0.15	0.10	0.10	0.05	<0.001	0.999	−0.003	0.923
	0.30		0.40	0.35	0.10	0.05	<0.001	0.990	−0.004	0.930
0.30	0.05		0.40	0.10	0.35	0.05	<0.001	0.991	−0.007	0.846
	0.30		0.65	0.35	0.35	0.05	−0.003	0.907	−0.011	0.791
0.05	0.05	0.05	0.20	0.10	0.10	0.05	−0.001	0.977	0.000[c]	1.000[c]
	0.30		0.45	0.35	0.10	0.05	−0.001	0.956	−0.003	0.947
0.30	0.05		0.45	0.10	0.35	0.05	−0.001	0.959	−0.005	0.880
	0.30		0.70	0.35	0.35	0.05	−0.005	0.851	−0.010	0.806
0.05	0.05	0.20	0.35	0.10	0.10	0.05	−0.003	0.902	0.006	1.128
	0.30		0.60	0.35	0.10	0.05	−0.005	0.837	−0.001	0.983
0.30	0.05		0.60	0.10	0.35	0.05	−0.005	0.846	−0.002	0.960
	0.30		0.85	0.35	0.35	0.05	−0.011	0.639	−0.008	0.845

Exposure has no true effect on outcome. Full scenario: prevalence of $E = 0.5$; prevalence of $U = 0.2$; baseline probability of censoring = 0.05; baseline risk = 0.01 effect of U on risk is a 5% increase, i.e., $\Pr(Y = 1|U = 1) - \Pr(Y = 1|U = 0) = 0.05$; as exposure has no true effect on risk, all non-null measures of association represent an expected bias.

[a]Effects on censoring expressed as the increase in probability of censoring with the factor.

[b]RD = $\Pr(Y = 1|E = 1, C = 0) - \Pr(Y = 1|E = 1, C = 0)$; RR (relative risk) = $[\Pr(Y = 1|E = 1, C = 0)]/[\Pr(Y = 1|E = 0, C = 0)]$.

[c]In this scenario, the censoring ratio (i.e., $\Pr(C = 1|E = 1, U = u)/\Pr(C = 1|E = 0, U = u)$) is equal for $U = 1$ and $U = 0$, so that the multiplicative model holds and no bias is expected in the $C = 1$ stratum.

University of Massachusetts
Amherst, MA
bwhitcomb@schoolph.umass.edu

Patrick F. McArdle
Division of Endocrinology,
Diabetes, and Nutrition
University of Maryland
School of Medicine
Baltimore, MD

REFERENCES

1. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology*. 1999;10:37–48.
2. Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. *Epidemiology*. 2004;15:615–625.
3. Pearce N, Richiardi L. Commentary: three worlds collide: Berkson's bias, selection bias and collider bias. *Int J Epidemiol*. 2014;43:521–524.
4. Snoep JD, Morabia A, Hernández-Díaz S, Hernán MA, Vandenbroucke JP. Commentary: a structural approach to Berkson's fallacy and a guide to a history of opinions about it. *Int J Epidemiol*. 2014;43:515–521.
5. Greenland S. Quantifying biases in causal models: classical confounding vs collider-stratification bias. *Epidemiology*. 2003;14:300–306.
6. Liu W, Brookhart MA, Schneeweiss S, Mi X, Setoguchi S. Implications of M bias in epidemiologic studies: a simulation study. *Am J Epidemiol*. 2012;176:938–948.
7. Pizzi C, De Stavola B, Merletti F, et al. Sample selection and validity of exposure-disease association estimates in cohort studies. *J Epidemiol Community Health*. 2011;65:407–411.
8. Greenland S. Additive risk versus additive relative risk models. *Epidemiology*. 1993;4:32–36.

A Sensitivity Analysis to Assess Bias Due to Selecting Subjects Based on Treatment Received

To the Editor:

Instrumental variables (IVs) are widely used to avoid bias caused by unmeasured confounders.[1,2] In studies comparing the effect of treatments A and B, a common approach is to select patients who have received either of those treatments but not others. Swanson et al.[3] showed that selecting patients based on their received treatment in IV analyses produces biased results, and described the structure of this bias using causal diagrams. Ertefaie et al.[4] further formalized this problem and showed that the local average treatment effect can be identified as a function of a four-dimensional sensitivity analysis.

We utilized this sensitivity analysis to study the effect on body mass index (BMI) of sulfonylureas versus dipeptidyl peptidase-4 (DPP-4) inhibitors as second-line therapy added to metformin. Rapid adoption of DPP-4s prompted us to use calendar time as an IV (IV = DPP4 2008–2009; IV=Sulf 2004–2007).[5] We included patients who were followed for at least 180 days on metformin, had a glycosylated hemoglobin (HbA1c) of ≥7%, then started on a second antidiabetic agent added to metformin. The outcome was the first

Supported by National Science Foundation Grant SES-1260782.
The authors report no conflicts of interest.
SDC Supplemental digital content is available through direct URL citations in the HTML and PDF versions of this article (www.epidem.com).

Copyright © 2015 Wolters Kluwer Health, Inc. All rights reserved.
ISSN: 1044-3983/16/2702-00e5
DOI: 10.1097/EDE.0000000000000430

Copyright © 2015 Wolters Kluwer Health, Inc. Unauthorized reproduction of this article is prohibited.