Published in final edited form as:

JAMA. 2011 February 23; 305(8): 822–823. doi:10.1001/jama.2011.163.

Index event bias: an explanation for the paradoxes of recurrence risk research

Issa J. Dahabreh, MD and David M. Kent, MD, MS

Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, Boston, MA

Keywords

recurrence risk; collider bias; index event bias

A review of the recurrence risk literature uncovers several paradoxical observations: thrombophilias predispose for a first episode of deep venous thrombosis (DVT) but not for recurrence (the thrombophilia paradox¹); aspirin takers are at lower risk for cardiac events, but higher risk of recurrence (the aspirin paradox²); and while obesity is an established risk factor for coronary artery disease, it appears to protect against recurrent coronary events (the obesity paradox³). While several hypotheses have been proposed to explain these paradoxes, it is not generally appreciated that all recurrent risk analyses are prone to a particular bias that may induce such paradoxical results. We call it “index event bias” because it arises in studies that select patients based on the occurrence of an index event.

We demonstrate how “index event bias” emerges and that – because of the general congruence between risk factors for the index and recurrent events – it will generally tend to bias studies toward the null, causing their contribution to be substantially underestimated or even reversed. We use the association between patent foramen ovale (PFO) and cryptogenic stroke as a prototypical example, and describe several other instances where it may have affected study results.

PFO and the risk of stroke

Although the prevalence of PFO in the general population is 25%, it is approximately double among cryptogenic stroke patients.⁴ This association is attributed to paradoxical embolism, i.e. a thromboembolus from the venous system gaining access to the arterial circulation via a right-to-left shunt in the heart. The consistent association of PFO with cryptogenic stroke (suggesting a pathogenic role for this otherwise benign congenital anomaly) has been accompanied by a surprising yet equally consistent finding: PFO is not a risk factor for recurrent stroke.⁴ The discrepancy can be reconciled by the following observation: among patients with a cryptogenic stroke, those with a PFO tend to be younger, have approximately half the rate of diabetes and hypertension and are less likely to have hypercholesterolemia or smoke.⁴ Thus, the similarity in recurrence rates in patients with cryptogenic stroke with or

without PFO suggests that PFO is an important risk factor for stroke recurrence, since it “compensates” for a profound shortfall in conventional stroke risk factors.

Index event bias

When multiple risk factors contribute to the risk of an outcome, conditioning on the outcome induces dependence between the risk factors, even when they are independently distributed in the general population (e.g. there is no association between PFO and diabetes, hypertension, hypercholesterolemia in the general population, but there is a negative association in the cryptogenic stroke population). Heuristically, the association may be thought to arise because patients with PFOs do not require the same burden of stroke risk factors to have the index event; thus the presence of a PFO “protects” cryptogenic stroke patients from other risk factors (both known and unknown), see Table.⁵⁻⁶ Because risk factors often have congruent effects on the index and recurrent events, this negative association will tend to bias any estimation of the PFO effect on recurrence risk towards the null, unless there is a thorough accounting for all shared risk factors⁷ (for example, in a recent study propensity score matched analysis eliminated the obesity paradox⁸). Obviously, bias will always be a concern since unknown risk factors are often operating to cause a disease. More generally, the biases we describe are special cases of “collider bias”, which is well described in the epidemiologic literature.⁹

A threat for causal inference in recurrence risk research

“Index event bias” is often underappreciated and should be considered as a possible contributing factor to many of the paradoxes enumerated in the introduction. As with PFO, thrombophilias presumably contribute to the risk of a second DVT, but they also “protect” against other DVT risk factors among patients with a first event. Similarly, not taking aspirin will “protect” against other heart attack risk factors among patients with a first heart attack. While the assumptions required to be mathematically certain that one risk factor will consistently “protect” against the presence of others are very stringent and not typically met in the real world,⁷ dependence between risk factors will occur in all studies that select participants based on the occurrence of a first event because the selection influences the distribution of risk factors in the patients that are included in the study and affects the association of these risk factors with the outcome of interest (the recurrent event).⁹ However, in such cases, the distribution of the risk factors and the behavior of the bias in estimating causal effects are unpredictable.

In addition to being a ubiquitous issue with studies examining risk factors of recurrence, index event bias may affect research that examines disease progression and severity, when there are multiple risk factors for progression or for severity that are also risk factors for having the disease in the first place. For example, the so-called “smokers paradox” describes the better prognosis of smokers compared to non-smokers among patients with myocardial infarction.¹⁰ Also, among patients with heart failure, better survival has been associated not only with obesity, but with higher blood pressure and higher levels of low-density lipoprotein and higher total cholesterol – both established risk factors for the development of heart failure.

Readers should keep in mind that the true causal models underlying the associations discussed in this commentary are unknown. For many of the above examples, other factors no doubt also contribute. Presumably “confounding by indication” also plays a role in the emergence of the aspirin paradox (since those with more risk factors may have preferentially been prescribed aspirin) and some of the “reverse epidemiology” seen in the context of heart failure and other diseases may actually reflect unknown causal pathways. Our purpose is to underscore how the tenuous relationship between association and causation becomes even more so in the context of recurrence risk research employing simple regression models. In this context – because the index and recurrent events have common risk factors – many so-called paradoxes should actually be expected, since they are induced by conditioning the analyses on the occurrence of the index event.

Acknowledgments

This work was partially funded by grants UL1 RR025752 and R01 NS062153, both from the National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funder. The authors have no conflicts of interest to report.

References

1. Baglin T. Unraveling the thrombophilia paradox: from hypercoagulability to the prothrombotic state. *J Thromb Haemost.* Feb; 2010 8(2):228–233. [PubMed: 19943876]
2. Rich JD, Cannon CP, Murphy SA, Qin J, Giugliano RP, Braunwald E. Prior aspirin use and outcomes in acute coronary syndromes. *J Am Coll Cardiol.* Oct 19; 2010 56(17):1376–1385. [PubMed: 20946994]
3. Gruberg L, Weissman NJ, Waksman R, et al. The impact of obesity on the short-term and long-term outcomes after percutaneous coronary intervention: the obesity paradox? *J Am Coll Cardiol.* Feb 20; 2002 39(4):578–584. [PubMed: 11849854]
4. Kent DM, Thaler DE. Is patent foramen ovale a modifiable risk factor for stroke recurrence? *Stroke.* Oct; 2010 41(10 Suppl):S26–30. [PubMed: 20876498]
5. Krasuski RA, Hart SA, Allen D, et al. Prevalence and repair of intraoperatively diagnosed patent foramen ovale and association with perioperative outcomes and long-term survival. *JAMA.* Jul 15; 2009 302(3):290–297. [PubMed: 19602688]
6. Weimar C, Holle DN, Benemann J, et al. Current management and risk of recurrent stroke in cerebrovascular patients with right-to-left cardiac shunt. *Cerebrovasc Dis.* 2009; 28(4):349–356. [PubMed: 19628936]
7. VanderWeele TJ, Hernan MA, Robins JM. Causal directed acyclic graphs and the direction of unmeasured confounding bias. *Epidemiology.* Sep; 2008 19(5):720–728. [PubMed: 18633331]
8. Adamopoulos C, Meyer P, Desai RV, et al. Absence of obesity paradox in patients with chronic heart failure and diabetes mellitus: a propensity-matched study. *Eur J Heart Fail.* Oct 7.2010
9. Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. *Epidemiology.* Sep; 2004 15(5):615–625. [PubMed: 15308962]
10. Barbash GI, Reiner J, White HD, et al. Evaluation of paradoxical beneficial effects of smoking in patients receiving thrombolytic therapy for acute myocardial infarction: mechanism of the “smoker's paradox” from the GUSTO-I trial, with angiographic insights. *Global Utilization of Streptokinase and Tissue-Plasminogen Activator for Occluded Coronary Arteries. J Am Coll Cardiol.* Nov 1; 1995 26(5):1222–1229. [PubMed: 7594035]

Table

Distribution of selected established stroke risk factors in “unselected” individuals and individuals who have had an index stroke event, stratified by PFO status. It is apparent that conditioning on the index stroke radically changes the distribution of risk factors, with patients not having a PFO carrying a much greater burden of other stroke risk factors. Data for “unselected” individuals were extracted from the baseline data from Krasuski et al.⁵, a study of incidentally discovered PFO among patients undergoing heart surgery. Data for patients who have already had an index stroke were extracted from Weimar et al.⁶, a study of individuals who had experienced a cryptogenic stroke and were followed up for stroke recurrence. Percentages have been rounded to the nearest integer.

Risk factors	Unselected population (n=13,815)		Population selected for “index” stroke (n=1126)	
	PFO (n=2277)	No PFO (n=10,815)	PFO (n=404)	No PFO (n=722)
Mean age (years)	63.5	62.9	53.6	60.6
Hypertension (%)	65	67	43	62
Diabetes (%)	23	24	13	19
Smoking (%)	58	58	27	35

PFO = patent foramen ovale.