Published in final edited form as:

Epidemiology. 2012 May ; 23(3): 443–445. doi:10.1097/EDE.0b013e31824e2d4e.

Theorems, proofs, examples, and rules in the practice of epidemiology

Tyler J. VanderWeele and Elizabeth L. Ogburn

Program on Causal Inference, Harvard School of Public Health, Boston, MA 02115

Tyler J. VanderWeele: tvanderw@hsph.harvard.edu

The practice of epidemiology relies on general rules of thumb to guide analysis. Such rules include: 'control for all common causes of the exposure and the outcome;' 'make sure there are at least ten events (and non-events) per covariate in a logistic regression model;' 'a covariate can be discarded if omitting it does not change the regression coefficient of the exposure by more than 10%;' 'if an unmeasured confounder affects both the exposure and the outcome in the same direction then the bias will be positive; if in opposite directions, negative;' 'bias from an uncontrolled common cause of an exposure and an outcome generally exceeds the bias resulting from conditioning on a pre-exposure 'collider' variable;' 'nondifferential misclassification of an exposure biases effect estimates towards the null;' 'control for a nondifferentially misclassified confounder will give estimates between the crude and true effects.' Some of these rules of thumb are on firmer theoretical footing than others. Some are informed by theoretical results, some by simulations, some by years of tried-and-true practice, and many of them by a combination of all of the above. Without these rules, the teaching and practice of epidemiology would indeed be difficult.

As noted by Greenland in his commentary,¹ theoretical results or theorems have the advantage of clearly laying out the conditions under which they apply, but they can often be of limited use in practice because the complexities generally encountered in any analysis of real data render their conditions inapplicable. Epidemiologists therefore instead rely on general rules of thumb to guide analysis. Although theorems themselves will generally be applicable only in simple settings, they can give rise to rules of thumb, the scope of which is sometimes much broader than the confines of the original theorems. For example, Bross² showed that nondifferential misclassification of a binary outcome (or, by symmetry, a binary exposure) will give valid tests for an exposure-outcome association but may reduce their power. Today many epidemiologists use the rule of thumb 'nondifferential misclassification of an exposure biases effects estimates towards the null.'

From the standpoint of epidemiology as an academic discipline, it is of course important to distinguish among proven results, conjectures, the results of simulations, and rules of thumb; among what we know, what we think might be the case, and what seems to be 'best practice.' All are important in informing practice and in advancing methodology but each plays a different role. Simulations, because they can explore complex settings more easily

than theorems, are in general very important in informing practice and shaping our rules of thumb. Counterexamples, on the other hand, are useful precisely because they can demonstrate that a rule is not universal. In some cases, a counterexample may not ultimately overturn the use of the rule to which it is the exception. For example, it was demonstrated some time ago³ that the rule inspired by Bross, ‘nondifferential misclassification of an exposure biases effects estimates towards the null,’ has exceptions if the exposure is not binary. The warning has been reiterated on numerous occasions. And yet, the general rule – without the ‘binary’ qualification – is often cited in applied work. Perhaps this general rule has not been overturned because in many cases the exposure is binary. Perhaps it has not been overturned because of its simplicity. Perhaps it has not been overturned because there have not been sufficiently dramatic examples of exceptions in actual epidemiologic applications. Perhaps it has not been overturned because ignoring nondifferential exposure ‘works’ fairly well insofar as it preserves, in many common settings though not always, the direction of a trend.^{4,5} Even when counterexamples and qualifications do not alter a rule of thumb, they can serve to alert epidemiologists to cases in which further consideration may be warranted.

As we note in our paper,⁶ numerous textbooks^{7–9} cite a ‘partial control result’ for binary confounders: that for a nondifferentially misclassified binary confounder, the effect estimate adjusted for the misclassified confounder will lie between the true and the crude effect measures. These sources in general state this not as a rule of thumb but as an established result. As we have shown,⁶ this ‘result’, stated in its unqualified form, does not in fact hold, even for a binary confounder, though it will hold if either (i) the binary confounder affects the outcome in the same direction for both the exposed and unexposed or if (ii) the effects are computed only for the exposed, or only for the unexposed, rather than for the entire sample.⁶ Thus the ‘partial control result,’ even for a binary confounder, is subject to caveats.

As Greenland notes, rightly, in our view, the caveats may be of limited relevance in practice. Greenland points out that the risk ratios used in our counterexample are of an order of magnitude that is rarely, if ever, encountered in epidemiology. In the appendix here we give another counterexample in which the risk ratios are more realistic, but neither does this counterexample undermine the fact that, in most settings, it is probably a good idea to adjust for a binary confounder subject to nondifferential misclassification. In the vast majority of settings, the confounder will likely affect the outcome in the same direction for both exposure groups; when it does not, although the theorem does not apply, the ‘partial control’ pattern will often still hold; when even this fails, the additional resulting bias will likely, in most cases, be small. At most, then, our results flag a case in which analysts may want to give further consideration to the possible implications of misclassification. Our counterexamples probably will not, and perhaps even should not, overturn the ‘partial control result’ taken as a rule of thumb. Indeed, the operative rule of thumb used in practice probably encompasses confounders that are not binary and even multiple confounders, despite the fact that it has long been established¹⁰ that counterexamples exist in these cases. Do we not often – almost always – control for confounders of all sorts subject to measurement error in actual practice? As rules of thumb go, there can be exceptions, but, overall, the ‘partial control’ rule is probably a pretty good one. As suggested by Greenland,

the conditions needed for its failure warrant further exploration, but the general rule in many cases works well.

Nonetheless it seems that what was ultimately only a general pattern or a rule of thumb was taken, for a number of years, as a formal result. It is not entirely clear why or at exactly what point this came to be. The original work of Greenland¹¹ in 1980, which is generally cited for the ‘partial control result’, arguably does not make this general claim. However, certainly by the time of Brenner’s 1993 paper,¹⁰ it was widely assumed that the result had in fact been demonstrated, and textbooks by and large followed suit. The implications of this confusion are certainly not catastrophic, but the sequence of events serves as an interesting case study in the history of epidemiologic methods and the nature of intuition and formal argument. Sharper distinction between examples, conjectures and rules of thumb on the one hand, and established results on the other, may have circumvented the mix-up in this case. Simulations, numerical examples, intuitions, and rules of thumb play an important role in epidemiology but, in the absence of a formal analytic argument, should not be interpreted as having established a formal result. The primary intellectual contribution of our paper was to articulate conditions under which the ‘partial control result’ does hold and to develop proofs. The practical consequences of our results may indeed be limited since, in this case, the theorems and proofs came after, not before, the general rule of thumb.

References

1. Greenland S. Intuitions, simulations, theorems: the role and limits of methodology. *Epidemiology*. in press.
2. Bross I. Misclassification in 2 X 2 tables. *Biometrics*. 1954; 10(4):478–486.
3. Dosemeci M, Wacholder S, Lubin JH. Does nondifferential misclassification of exposure always bias a true effect towards the null value? *Am J Epidemiol*. 1990; 132(4):746–748. [PubMed: 2403115]
4. Dosemeci M, Wacholder S, Lubin JH. Does nondifferential misclassification of exposure always bias a true effect towards the null value? *Am J Epidemiol*. 1990; 132(4):746–748. [PubMed: 2403115]
5. Weinberg CA, Umbach DM, Greenland S. When will nondifferential misclassification of an exposure preserve the direction of a trend? *Am J Epidemiol*. 1994; 140(6):565–571. [PubMed: 8067350]
6. Ogburn EL, VanderWeele TJ. On the nondifferential misclassification of a binary confounder. *Epidemiology*. in press.
7. Lash, TL.; Fox, MP.; Fink, AK. *Applying Quantitative Bias Analysis to Epidemiologic Data*. New York, NY: Springer Verlag; 2009.
8. Szklo, M.; Nieto, FJ. *Epidemiology: Beyond the Basics*. Sudbury, MA: Jones and Bartlett Publishers; 2004.
9. Rothman, KJ.; Greenland, S.; Lash, TL. *Modern Epidemiology*. 3. Philadelphia, PA: Lippincott Williams and Wilkins; 2008.
10. Brenner H. Bias due to non-differential misclassification of polytomous confounders. *J Clin Epidemiol*. 1993; 46:57–63. [PubMed: 8433115]
11. Greenland S. The effect of misclassification in the presence of covariates. *Am J Epidemiol*. 1980; 112:564–569. [PubMed: 7424903]

Appendix

An important and yet-unresolved question raised by our findings is how common violations of the partial control result are in real-world settings. We do not at this time have an answer to that question, but we do know that counterexamples exist for many different data configurations. In Table 1 we present a counterexample that does not have the highly discrepant stratum-specific risk ratios which were a feature of the counterexample in our paper. In this example, the risk ratios for the effect of exposure on outcome within levels of the confounder are 1.5 ($C = 1$) and 0.83 ($C = 0$), representing slightly less than a two-fold difference. The risk ratios for the effect of the confounder on the outcome within levels of exposure are also 1.5 ($A = 1$) and 0.83 ($A = 0$). The risk ratio for the effect of the confounder on the exposure is 0.35. The partial control ordering is violated by a true risk ratio of 0.850, a crude risk ratio of 0.849, and an observed adjusted risk ratio of 0.846. We have found counterexamples with more dramatic violations to the partial control ordering, but each has a large risk ratio for at least one of the relationships between the confounder and the outcome conditional on the exposure, between the exposure and the outcome conditional on the confounder, or between the confounder and the exposure.

Table 1

Counterexample to the Partial Control Result

Crude		A = 1	A = 0					RD_{crude}	-0.059
	Y = 1	508	215					RR_{crude}	0.849
	Y = 0	1008	330					OR_{crude}	0.774
True	C = 1	A = 1	A = 0	C = 0	A = 1	A = 0		RD_{true}	-0.059
	Y = 1	8	15	Y = 1	500	200		RR_{true}	0.850
	Y = 0	8	30	Y = 0	1000	300		OR_{true}	0.773
Sensitivity = 1, Specificity = 0.97									
Observed	C' = 1	A = 1	A = 0	C' = 0	A = 1	A = 0		RD_{obs}	-0.061
	Y = 1	23	21	Y = 1	485	194		RR_{obs}	0.846
	Y = 0	38	39	Y = 0	970	291		OR_{obs}	0.768