

Simpson's Paradox and Confounding

Vance W. Berger, PhD
National Cancer Institute and University of Maryland Baltimore County
Biometry Research Group, National Cancer Institute
Executive Plaza North, Suite 3131
6130 Executive Boulevard, MSC 7354
Bethesda, MD 20892-7354
(301) 435-5303 (voice), (301) 402-0816 (fax), vb78c@nih.gov

Editor: Ameringer et al. (2009) refer to Simpson's paradox as "a case of extreme confounding", which of course it is, but unfortunately it is not the only confounding that one finds in this article. Nor, for that matter, is the "overlooked confounding variable" the only key item that has been overlooked. The authors propose blocking and minimization as competing methods for preventing this problem, but what was described as blocking is actually stratification. Gender can be a stratification factor, and often is. Blocking refers instead to stratification on time of entry into the trial. This form of blocking, but not the one described in the paper, truly is a competitor to minimization; either can be used within the strata formed by gender or any other stratification factor(s). Of course, there are other considerations that make both blocking and minimization inappropriate for clinical trials. Most notable for blocking is the ability of investigators to predict future allocations and thereby induce confounding by selecting healthier patients to enroll when one treatment is due to be allocated, and sicker patients when the other one is (Berger, 2005A). Minimization in its pure form is not randomization at all, but rather is deterministic, with only ties broken by randomization. It, too, can be subverted. Currently, the maximal procedure appears to be the best method of randomization (Berger, 2005A).

So much for confounding blocking and stratification. What was omitted in the article was a discussion of when the stratified analysis is likely to mislead, and when the unstratified one is. In fact, the case is not always as simple as the authors make it seem to be. If the

treatment can influence the predictor, then the unstratified analysis is the more appropriate one (Berger, 2004), so another implication is that covariates in randomized trials must not be measured subsequent to randomization. In observational studies, tremendous care is required to arrive at the truth, or even to arrive at what the data appear to be presenting as the truth. Finally, one must consider both the extent of imbalance across treatment groups and the predictive ability of the predictor when choosing predictors to include in the model, so as to best prevent Simpson's paradox (Berger, 2005B).

References:

- Ameringer, S, Serlin, RC, Ward, S (2009). "Simpson's Paradox and Experimental Research", *Nursing Research* **58**, 2, 123-127.
- Berger, VW (2004). "Valid Adjustment for Binary Covariates of Randomized Binary Comparisons", *Biometrical Journal* **46**, 5, 589-594.
- Berger, VW (2005A). "Selection Bias and Covariate Imbalances in Randomized Clinical Trials", John Wiley & Sons, Chichester.
- Berger, VW (2005B). "A Novel Criterion for Selecting Covariates", *Drug Information Journal* **39**, 233-241.