RTI Author Awards Series

A Review of the Application of Propensity Score Methods Yielded Increasing Use, Advantages in Specific Settings, But Not Substantially Different Estimates Compared With Conventional Multivariable Methods

Stürmer, T., Joshi, M., Glynn, R.J., Avorn, J., **Rothman, K.J.**, Schneeweiss, S. (2006). A review of the application of propensity score methods yielded increasing use, advantages in specific settings, but not substantially different estimates compared with conventional multivariable methods. *Journal of Clinical Epidemiology*, 59 (5):437-447.


Kenneth Rothman

Propensity scoring (estimating the probability that someone is exposed) is one of two methods of obtaining a summary confounder score that can be used as a proxy for all measured confounding variables in a study. The other method is disease risk scoring. Propensity score (PS) modeling has been used to address confounding by indication, which is a bias in pharmacoepidemiologic studies in which the effect of a treatment is confounded with the indications for getting that treatment. The indications can be as straightforward as demographic factors such as age and sex, but can also include medical history, family history, and other details of the patient's condition.

At the time that we wrote this paper, there was little empirical evidence that PS analyses achieved any better control of confounding by indication than did conventional multivariate outcome modeling. In this review, we assessed the use of propensity scores over time and critically evaluated studies published through 2003.

We documented that the use of propensity scores had increased from a total of 8 papers in the literature before 1998 to 71 papers in 2003. (The increase has continued.) Most of the published studies reviewed assessed the effect of medications or surgical interventions, mainly in cardiology and cardiac surgery. Whether PS methods or conventional outcome regression models were used to control for confounding had little effect on results in those studies in which such comparison was possible. Only 9 out of 69 studies (13%) had an effect estimate that differed by more than 20% from that obtained with a conventional outcome model. Thus, despite the increase in use of propensity score methods, there is little evidence that these methods yield substantially different estimates from more conventional multivariable methods.

Link: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1448214/pdf/nihms8835.pdf