Published in final edited form as:

Epidemiology. 2015 March ; 26(2): 238–241. doi:10.1097/EDE.0000000000000241.

The complex role of the prescribing physician for confounder adjustment in comparative effectiveness research methods

Jessica M Franklin*, Sebastian Schneeweiss, Krista F Huybrechts, and Robert J Glynn

Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA

Abstract

In nonrandomized studies of comparative effectiveness of medications, the prescriber may be the most important determinant of treatment assignment, yet the majority of analyses ignore the prescriber. Via Monte Carlo simulation, we evaluated the bias of 3 approaches that utilize the prescriber in analysis compared against the default approach that ignores the prescriber. Prescriber preference instrumental variable (IV) analyses were unbiased when IV criteria were met, which required no clustering of unmeasured patient characteristics within prescriber. In all other scenarios, IV analyses were highly biased, and stratification on the prescriber reduced confounding bias at the patient or prescriber levels. Including a prescriber random intercept in the propensity score model reversed the direction of confounding from measured patient factors and resulted in unpredictable changes in bias. Therefore, we recommend caution when using the IV approach, particularly when the instrument is weak. Stratification on the prescriber may be more robust and warrants additional research.

Keywords

comparative effectiveness; epidemiologic methods; instrumental variable; prescriber; propensity score; random effects; simulation

Introduction

In nonrandomized studies of comparative effectiveness of medications, the prescriber may be the most important determinant of treatment assignment,¹ yet the majority of analyses ignore the prescriber. If the prescriber influences the choice of therapy, but has no other association with patient outcomes, then the prescriber meets the requirements of a valid instrumental variable (IV).² In this case, an IV analysis utilizing prescriber treatment preference as the instrument can yield an unbiased estimate of a homogeneous treatment effect, even in the presence of unmeasured confounding.^{3–13} However, IV estimates can be very imprecise, especially when the IV-exposure association is weak,¹⁴ and the IV assumptions for prescriber preference may be questionable in routine care.

*Address correspondence to Dr. Jessica Franklin, 1620 Tremont St., Suite 3030, Boston, MA 02120; JMFranklin@partners.org; ph: 617-278-0675; fax:617-232-8602.

If IV assumptions are violated, the prescriber may instead be a confounder or a surrogate for confounders of the exposure-outcome association; recent work has suggested several potential strategies for accounting for confounding in multi-level data.^{15–17} Stratifying the analysis on prescriber would mimic the analysis of randomized trials, where randomization is often stratified on centers. Including a random intercept for prescriber in the propensity score model that predicts the probability of treatment given measured covariates¹⁸ could also remove confounding associated with the prescriber. However, if IV criteria are met, these approaches may worsen balance on unmeasured factors associated with treatment choice¹⁹ and amplify residual bias.^{20–22} In addition, with few patients per prescriber, stratification may be impossible, and random effects models may perform poorly.²³

The objective of this study was to evaluate the relative bias of analytic approaches that utilize the prescriber across a range of plausible causal structures and to compare these approaches with the usual practice of ignoring the prescriber.

Methods

We simulated data based on the causal diagram shown in Figure 1 such that the number of patients per prescriber varied across prescribers (details in the Web Appendix). The primary association of interest is the causal effect of drug choice on patient outcomes, specified as a homogeneous risk ratio (RR). We considered 13 simulation scenarios, which varied the true causal structure, the strength of correlation in treatment choice within prescribers, the average number of patients per prescriber, and the study size (Table 1). Treatment effect is confounded by patient characteristics in all scenarios and by the prescriber characteristic in scenarios 6–7, although we assume throughout that it is unobserved. We reran each scenario 4 times, varying the true treatment effect ($RR = 1.0$ or 0.37) and the presence of unmeasured patient characteristics for a total of 52 unique simulation settings with 1000 datasets each.

In each simulated dataset, we estimated treatment effects using five potential analysis approaches: 1) the crude RR, 2) ordinary propensity score matching using measured covariates (PS_1), 3) matching on a propensity score that included measured covariates and a prescriber random intercept (PS_2), 4) matching within prescriber on PS_1 , and 5) instrumental variable analysis that adjusts for measured variables and uses the “last prescription” IV, defined as the treatment for the most recent prior patient within prescriber.^{24,25} For simplicity, we utilized 1-to-1 matching with a caliper in the propensity score analyses, but in practice investigators may prefer other methods that retain more patients in the analysis. In all analyses, we focus on estimating the same estimand, the homogenous RR.

Results

Figure 2 presents the bias of the log-RR estimates across varying causal structures with a null treatment effect. As expected, when instrument criteria were met, the instrumental variable analysis was approximately unbiased. When the instrument was confounded at the patient or prescriber level, the IV analyses were biased. Conversely, stratification on prescriber increased bias slightly when IV criteria were met and patient-level confounders were unmeasured, but reduced bias when IV criteria were not met or when all patient-level

factors were measured. Matching on PS_2 , which included the prescriber random intercept, moved treatment effect estimates downward from the default propensity score analysis, which sometimes reduced bias and sometimes increased it. Results with varying treatment effect, study size, or strength of prescriber preference were similar (see Web Appendix, S Figures 1–3).

To shed light on the mechanisms of observed bias, figure 3 presents balance on patient and prescriber factors across exposure or IV groups. Results are reported for all variables, even though one patient risk factor, C_2 , and the prescriber characteristic, Z , were “unobserved” in these settings and would not in practice be available for diagnostics. Matching on PS_1 , whether stratified on prescriber or not, created balance on measured patient factors. However, stratification on prescriber additionally balanced the prescriber characteristic. Matching on PS_2 similarly removed imbalance on the prescriber characteristic, but it reversed the direction of imbalance on measured patient characteristics. The magnitude of this imbalance was greater in scenarios with few patients per prescriber (scenarios 9 and 12, shown in SFigures 5–6). Both stratification on prescriber and including the prescriber in the propensity score model increased imbalance on unmeasured patient factors when prescriber characteristics were not associated with patient characteristics (scenarios 1, 2, and 6), and decreased imbalance when true prescriber treatment preference was associated with patient characteristics (scenarios 4 and 7).

Balance across groups defined by instrument status was identical for both measured and unmeasured patient factors because these factors were simulated to have identical associations with exposure, outcome, and prescriber characteristics. Thus, balance on the measured patient covariate correctly indicated that the IV criteria were met in scenarios 1–2, but it missed a violation of IV criteria in scenario 6 because confounding of the IV was due to prescriber factors rather than patient factors. In scenarios 5–7, the IV reduced imbalance on both measured and unmeasured patient characteristics, but still resulted in increased bias over the crude analysis, demonstrating the well-known bias amplification⁵.

Discussion

We evaluated three approaches for incorporating information on the prescribing physician in comparative effectiveness studies of medications. We focused on the case of a homogeneous treatment effect so that the estimand of interest was identical in each analytic approach. If treatment effects are heterogeneous, these approaches would estimate treatment effect in different populations (see Swanson and Hernán²⁶ and Stürmer et al.²⁷ for overviews), and the choice of analytic approach may depend in part on the population of interest.

We found that stratifying on prescriber reduced bias in all simulation settings considered except those with unmeasured patient-level confounders and a valid prescriber preference instrument. These results are consistent with prior research that suggested that controlling for an instrumental variable could increase bias due to unmeasured patient characteristics.^{20, 22, 28, 29} In our study, stratification on prescriber was achieved through 1-to-1 matching, which removed many patients from analysis and led to highly variable estimates. In practice, more research is needed to evaluate methods for conditioning on the

prescriber that retain more patients in the analysis. Other analytic approaches, including those that do not utilize the propensity score for covariate adjustment, would be expected to similarly benefit from stratification on the prescriber. However, all stratification methods may result in large estimator variance in studies with few patients per prescriber on average or with very strong clustering of treatments within prescriber.

Matching on the propensity score that included a prescriber random intercept reversed the direction of imbalance on measured patient confounders. In cases where measured patient characteristics confounded in the same direction (and at approximately the same magnitude) as unmeasured characteristics, this resulted in cancellation of biases, but this behavior cannot be expected in general, as shown in scenarios with all confounders measured. The reversed imbalance on measured confounders was especially large when there were on average only 2 patients per prescriber, suggesting that, as in prior research,^{15–17} this effect may be eliminated when there are many patients per cluster, a scenario not commonly encountered when the clustering unit is the prescriber.

The instrumental variable approach could eliminate bias when instrument criteria were met, but was badly biased when criteria were not met, as expected. In real data, investigators are unlikely to know whether or not prescriber preference is a valid IV. Thus, investigators must weigh the likelihood that the IV criteria are met (or nearly met) versus the potential for bias if their assumptions are violated. In the primary simulation scenarios (1–7), the magnitude of the IV-exposure association was taken from a real study where the IV approach was used successfully,³⁰ but the IV approach more than tripled the bias from the default propensity score analysis in the most extreme scenario (7). When the IV-exposure association was weakened (scenarios 8 and 11), the magnitude of the IV-exposure association was taken from another real example study, and bias was worse. Conversely, the bias amplification caused by stratifying on the prescriber was generally small and was lessened when the IV-exposure association was weakened, as observed in prior work.²⁰

Therefore, we recommend caution in using the prescriber preference IV approach, particularly if the IV-exposure association is weak. In general, the choice of analytic approach will depend on the relative magnitude of the IV-exposure association compared with the likely strength of associations between prescriber characteristics and patient characteristics or outcomes. If instrument criteria are questionable, stratifying on the prescriber or ordinary confounder adjustment ignoring the prescriber may provide estimates with the least bias.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Conflicts of Interest and Source of Funding: This project was funded under Contract No. 290-05-00161 from the Agency for Healthcare Research and Quality, US Department of Health and Human Services as part of the Developing Evidence to Inform Decisions about Effectiveness (DEcIDE) program. The authors of this report are responsible for its content. Statements in the report should not be construed as endorsement by the Agency for Healthcare Research and Quality or the US Department of Health and Human Services.

References

1. Solomon DH, Schneeweiss S, Glynn RJ, Levin R, Avorn J. Determinants of selective cyclooxygenase-2 inhibitor prescribing: are patient or physician characteristics more important? *The American Journal of Medicine*. 2003; 115(9):715–720. [PubMed: 14693324]
2. Brookhart MA, Wang P, Solomon DH, Schneeweiss S. Evaluating short-term drug effects using a physician-specific prescribing preference as an instrumental variable. *Epidemiology*. 2006; 17(3): 268–275. [PubMed: 16617275]
3. Glymour, MM. Natural experiments and instrumental variable analyses in social epidemiology. In: Oakes, JM.; Kaufman, JS., editors. *Methods in Social Epidemiology*. San Francisco: Jossey-Bass; 2006. p. 429–460.
4. Grootendorst P. A review of instrumental variables estimation of treatment effects in the applied health sciences. *Health Services and Outcomes Research Methodology*. 2007; 7(3):159–179.
5. Hernán MA, Robins JM. Instruments for causal inference: an epidemiologist's dream? *Epidemiology*. 2006; 17(4):360. [PubMed: 16755261]
6. Martens EP, Pestman WR, De Boer A, Belitser SV, Klungel OH. Instrumental variables: application and limitations. *Epidemiology*. 2006; 17(3):260. [PubMed: 16617274]
7. Brookhart MA, Rassen JA, Schneeweiss S. Instrumental variable methods in comparative safety and effectiveness research. *Pharmacoepidemiology and Drug Safety*. 2010; 19(6):537–554. [PubMed: 20354968]
8. Chen Y, Briesacher BA. Use of instrumental variables in prescription drug research with observational data: a systematic review. *Journal of Clinical Epidemiology*. 2010
9. Rassen JA, Brookhart MA, Glynn RJ, Mittleman MA, Schneeweiss S. Instrumental variables I: instrumental variables exploit natural variation in nonexperimental data to estimate causal relationships. *Journal of Clinical Epidemiology*. 2009; 62(12):1226–1232. [PubMed: 19356901]
10. Rassen JA, Brookhart MA, Glynn RJ, Mittleman MA, Schneeweiss S. Instrumental variables II: instrumental variable application--in 25 variations, the physician prescribing preference generally was strong and reduced covariate imbalance. *Journal of Clinical Epidemiology*. 2009; 62(12): 1233–1241. [PubMed: 19345561]
11. Abrahamowicz M, Beauchamp M-E, Ionescu-Ittu R, Delaney JA, Pilote L. Reducing the variance of the prescribing preference-based instrumental variable estimates of the treatment effect. *American journal of epidemiology*. 2011; 174(4):494–502. [PubMed: 21765169]
12. Hennessy S, Leonard CE, Palumbo CM, Shi X, Ten Have TR. Instantaneous preference was a stronger instrumental variable than 3-and 6-month prescribing preference for NSAIDs. *Journal of clinical epidemiology*. 2008; 61(12):1285–1288. [PubMed: 18495427]
13. Pratt N, Roughead EE, Ryan P, Salter A. Antipsychotics and the risk of death in the elderly: an instrumental variable analysis using two preference based instruments. *Pharmacoepidemiology and drug safety*. 2010; 19(7):699–707. [PubMed: 20583208]
14. Ionescu-Ittu R, Delaney JAC, Abrahamowicz M. Bias-variance trade-off in pharmacoepidemiological studies using physician-preference-based instrumental variables: a simulation study. *Pharmacoepidemiology and Drug Safety*. 2009; 18(7):562–571. [PubMed: 19437424]
15. Arpino B, Mealli F. The specification of the propensity score in multilevel observational studies. *Computational Statistics & Data Analysis*. 2011; 55(4):1770–1780.
16. Gayat E, Thabut G, Christie JD, Mebazaa A, Mary J-Y, Porcher RI. Within-center matching performed better when using propensity score matching to analyze multicenter survival data: empirical and Monte Carlo studies. *Journal of clinical epidemiology*. 2013
17. Li F, Zaslavsky AM, Landrum MB. Propensity score weighting with multilevel data. *Statistics in medicine*. 2013; 32(9):3373–3387. [PubMed: 23526267]
18. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika*. 1983; 70(1):41–55.
19. Walker AM. Matching on provider is risky. *Journal of clinical epidemiology*. 2013; 66(8):S65–S68. [PubMed: 23849156]

20. Myers JA, Rassen JA, Gagne JJ, Schneeweiss S, Huybrechts KF, Rothman KJ, Joffe MM, Glynn RJ. The effects of adjusting for instrumental variables on bias and precision of effect estimates. *American Journal of Epidemiology*. 2011; 174(11):1213–1222. [PubMed: 22025356]
21. Brookhart MA, Schneeweiss S, Rothman KJ, Glynn RJ, Avorn J, Stürmer T. Variable selection for propensity score models. *American Journal of Epidemiology*. 2006; 163(12):1149–1156. [PubMed: 16624967]
22. Patrick AR, Schneeweiss S, Brookhart MA, Glynn RJ, Rothman KJ, Avorn J, Stürmer T. The implications of propensity score variable selection strategies in pharmacoepidemiology: an empirical illustration. *Pharmacoepidemiology and Drug Safety*. 2011; 20(6):551–559. [PubMed: 21394812]
23. Austin PC. Estimating multilevel logistic regression models when the number of clusters is low: a comparison of different statistical software procedures. *The International Journal of Biostatistics*. 2010; 6(1):16.
24. Palmer TM, Sterne JAC, Harbord RM, Lawlor DA, Sheehan NA, Meng S, Granell R, Smith GD, Didelez V. Instrumental Variable Estimation of Causal Risk Ratios and Causal Odds Ratios in Mendelian Randomization Analyses. *American Journal of Epidemiology*. 2011; 173(12):1392–1403. [PubMed: 21555716]
25. Rassen JA, Schneeweiss S, Glynn RJ, Mittleman MA, Brookhart MA. Instrumental variable analysis for estimation of treatment effects with dichotomous outcomes. *American Journal of Epidemiology*. 2009; 169(3):273–284. [PubMed: 19033525]
26. Swanson SA, Hernán MA. Commentary: How to Report Instrumental Variable Analyses (Suggestions Welcome). *Epidemiology*. 2013; 24(3):370–374. [PubMed: 23549180]
27. Stürmer T, Rothman KJ, Glynn RJ. Insights into different results from different causal contrasts in the presence of effect?measure modification. *Pharmacoepidemiology and drug safety*. 2006; 15(10):698–709. [PubMed: 16528796]
28. Walker, AM. The deleterious effects of control for provider in CER. Fourth Symposium on Comparative Effectiveness Research Methods; AHRQ Conference Center, Rockville, MD. 2012.
29. Pearl, J. Causality: models, reasoning, and inference. 2nd Edition ed. Cambridge Univ Press; 2009.
30. Schneeweiss S, Solomon DH, Wang PS, Rassen J, Brookhart MA. Simultaneous assessment of short-term gastrointestinal benefits and cardiovascular risks of selective cyclooxygenase 2 inhibitors and nonselective nonsteroidal antiinflammatory drugs: an instrumental variable analysis. *Arthritis & Rheumatism*. 2006; 54(11):3390–3398. [PubMed: 17075817]

Figure 1.
Causal diagram for the simulation study.

Figure 2.

Bias of the log-RR estimates for all analysis approaches across the first 7 simulation scenarios. Results are shown for settings with a true null treatment effect. The solid line displays results from the crude, unadjusted analysis. The regular and stratified propensity score approaches each use PS_1 , which does not account for prescriber. The prescriber RE approach uses PS_2 , which includes a prescriber random intercept.

Figure 3.

Imbalance across exposure groups and across IV groups in the first 7 simulation scenarios.

The left, middle, and right panels display imbalance on C_1 , C_2 , and Z, respectively. The solid line is the crude imbalance across exposure groups before matching. Results are shown for settings where C_2 is unobserved; when C_2 is observed, balance on C_2 and C_1 are identical.

Table 1

Data generation parameters for the simulation study. The values of γ , θ , and β_Z control the association between the binary prescriber characteristic (Z_i) and patient characteristics, prescriber preference, and patient outcomes, respectively. The value of σ^2 determines the heterogeneity in preference (P_i) across prescribers. m and n are the numbers of prescribers and patients, respectively.

Scenario	γ	θ	β_Z	σ^2	m	n
1	0	0	1	1.2	1000	5000
2	0	1	0	1.2	1000	5000
3	1	0	0	1.2	1000	5000
4	1	1	0	1.2	1000	5000
5	1	0	1	1.2	1000	5000
6	0	1	1	1.2	1000	5000
7	1	1	1	1.2	1000	5000
8	0	1	0	0.15	1000	5000
9	0	1	0	1.2	2500	5000
10	0	1	0	1.2	100	500
11	1	1	0	0.15	1000	5000
12	1	1	0	1.2	2500	5000
13	1	1	0	1.2	100	500