Apples and Oranges? Interpretations of Risk Adjustment and Instrumental Variable Estimates of Intended Treatment Effects Using Observational Data

Gang Fang*, John M. Brooks, and Elizabeth A. Chrischilles

*Correspondence to Dr. Gang Fang, Division of Pharmaceutical Outcomes and Policy, Eshelman School of Pharmacy, 2202 Kerr Hall, University of North Carolina at Chapel Hill, Chapel Hill, NC27599-7573 (e-mail: gang_fang@unc.edu).

Initially submitted February 12, 2011; accepted for publication July 21, 2011.

Instrumental variable (IV) and risk adjustment (RA) estimators, including propensity score adjustments, are both used to alleviate confounding problems in nonexperimental studies on treatment effects, but it is not clear how estimates based on these 2 approaches compare. Methodological considerations have shown that IV and RA estimators yield estimates of distinct types of causal treatment effects regardless of confounding problems. Many investigators have neglected these distinctions. In this paper, the authors use 3 schematic models to explain visually the relations between IV and RA estimates of intended treatment effects as demonstrated in the methodological studies. When treatment effects are homogeneous across a study population or when treatment effects are heterogeneous across the study population but treatment decisions are unrelated to the treatment effects, RA and IV estimates should be equivalent when the respective assumptions are met. In contrast, when treatment effects are heterogeneous and treatment decisions are related to the treatment effects, RA estimates of treatment effect can asymptotically differ from IV estimates, but both are correct even when the respective assumptions are met. Appropriate interpretations of IV or RA estimates can be facilitated by developing conceptual models related to treatment choice and treatment effect heterogeneity prior to analyses.

Confounding is a particularly vexing problem in clinical epidemiology studies designed to investigate treatment effects using a nonexperimental study design. In clinical practice, treatment is not simply an exposure but also a clinical decision and a choice. Treatment decisions are strongly influenced by confounders such as symptoms, severity, prognosis, and frailty of the patient. Analytical methods used to control confounding include instrumental variable (IV) analysis and risk adjustment (RA) methods, including conventional multivariable regression analyses and propensity score adjustments (1–4). RA methods yield unbiased estimates by adjusting for measured confounders with the assumption that all confounders are measured or that unmeasured confounders are “ignorable” when measured confounders are controlled (5–9). IV methods yield unbiased estimates with the requirement that the specified IVs or “instruments” are associated with treatment choice and the assumptions that they are unrelated to outcome directly and to unmeasured confounders. Thus, instruments essentially serve as natural experiments regarding treatment choice (10–13). In studies of treatment effectiveness using observational data, investigators are increasingly applying both RA and IV estimators, with the goal of alleviating confounding (3, 14–24).

However, estimates of treatment effects derived from RA and IV estimators often differ in the same study, and discussion of the differences often focuses only on the potential of each estimator to alleviate confounding. Although the discrepancy in the IV and RA estimators’ estimates can sometimes be attributed to violation of the key IV or RA assumptions, the differences can also caused by differences in the measures of association they use and by differences in the types of causal treatment effects estimated from IV and RA estimators,
respectively. While RA methods are well known in epidemiology and IV methods were introduced to epidemiologists over a decade ago (12, 13), investigators have not always appreciated that RA and IV methods yield estimates of distinct causal treatment effects, regardless of confounding (25–34). Therefore, it is possible for RA and IV estimates both to differ and to be correct. This distinction is especially germane when treatment decisions are influenced by the anticipated response of patients to treatment in the presence of treatment effect heterogeneity (34–38).

Simulation models have been used to describe these relations of the treatment effect concepts (27). Our purpose in this paper is to use 3 schematic models to describe visually the distinct causal treatment effect concepts produced by RA and IV estimators and their relation when the treatment effect is defined as the intended clinical benefit of treatment.

DISTINCT CAUSAL TREATMENT EFFECT CONCEPTS

The Rubin causal model shows that it is not possible to estimate treatment effects for individual patients, because the counterfactual treatment outcome for each patient is not observed. As a result, researchers must estimate average treatment effects (the difference between the treated and untreated outcomes) across groups of patients (7, 8, 39, 40). The average treatment effect concepts often discussed include average treatment effect in the population (ATE), average treatment effect across treated (ATT) and untreated (ATU) patients, and local average treatment effect for the “marginal patients” (LATE) (7, 8, 25–33, 39, 40). Marginal patients are the subset of patients whose treatment choices were affected by variation in specific factors which are “instruments” in IV analysis. Methodological studies have shown that RA estimators yield estimates of the ATT or ATE with the inverse probability of treatment weighting estimator, whereas IV estimators yield estimates of the LATE for the instruments specified in the analysis (7, 8, 25–33, 39–42).

Below, we use 3 schematic models to illustrate the relations between different treatment effect concepts and the interpretations of RA and IV estimates when treatment effects are homogeneous or heterogeneous and when treatment choice is related or unrelated to the treatment effects. To focus on the causal treatment effect concepts in these models, it is assumed that there are no unmeasured confounding variables and that all other assumptions are met. Shading is used in the accompanying figures to display the treatment effects, with solid shading representing homogeneity and graded shading representing heterogeneity (with darker shading signifying a greater treatment benefit). Border colors are used in the figures to indicate the portion of the population treated (red border) or untreated (black border).

MODELS FOR THE RELATION BETWEEN RA ESTIMATES AND IV ESTIMATES

Model 1: Treatment effects are homogeneous across the study population

Figure 1 presents the schematic model of different treatment effect concepts when treatment effects are homogeneous across patients. The rectangular box in the upper part of Figure 1 represents all patients. The shading in the box represents variation in treatment effect across the patients. Because treatment effects are homogeneous in this model, the box is shaded consistently across the population. The portion of the box with a red border represents patients who are treated, while the black-border portion represents patients who are not treated. RA estimators yield estimates of the average treatment effect for the patients within the red border—the average treatment effects in the treated (ATT) or the average treatment effects in the population (ATE)—with the inverse probability of treatment weighting estimator. The lower part of Figure 1 shows the basic model for an IV estimator using a single instrument that affects the probability of a patient’s receiving treatment. The instrument essentially functions as a natural randomizer that can be used to sort patients into groups with different treatment rates (3, 11–14, 19, 32, 43). For example, in Figure 1, patients are divided into 2 groups based on their instrument values. Because the instrument is related to the probability of receiving treatment, treatment rates will vary across groups—a high treatment rate group (IV group 2) and a low treatment rate group (IV group 1). IV estimators would yield a treatment effect estimate that is the average of effects for the patients between the high treatment rate and the low treatment rate (LATE). These are the patients for whom the instrument affected treatment choice—the marginal patients for this instrument. In this model, since the treatment effects are homogeneous across patients (constant color), ATT = ATU = ATE = LATE, and RA estimates will be equal to IV estimates (25–29, 31).

Model 2: Treatment effects are heterogeneous across the study population but the treatment decision is unrelated to treatment effect heterogeneity

Figure 2 presents the scenario termed “nonessential heterogeneity” in the literature (25, 27–29). Nonessential heterogeneity is represented by the color gradient shading, with darker color indicating a higher treatment benefit. Nonetheless, treatment decisions are not based on the differences in the treatment effects across patients. In other words, treatment is not sorted toward patients with a higher intended treatment benefit—effect modification factors are unknown or unrelated to treatment decisions. This scenario occurs when there is no available evidence suggesting who will benefit more from treatment, even though there is true underlying heterogeneity of treatment effects across patients. Under this scenario, ATT = ATU = ATE = LATE, and RA estimates will be equivalent to IV estimates (25–29, 31). Thus, as in model 1, both IV and RA estimate the average treatment effect of the treatment in the study population (ATE).

Model 3: Treatment effects are heterogeneous across the study population and the treatment decision is related to the treatment-effects heterogeneity

Figure 3 presents the schematic model that is termed “essential heterogeneity” in the methodological literature (25, 27–29). Essential heterogeneity is depicted in Figure 3 by the association of the color gradient with treatment decision.
Figure 1. Relation between different concepts of treatment effects when the treatment effects are homogeneous (ATT = ATU = ATE = LATE). ATE, average treatment effects for all patients; ATT, average treatment effects for the treated patients; ATU, average treatment effects for the untreated patients; IV, instrumental variable; LATE, local average treatment effects for the marginal patients.

Figure 2. Relation between different concepts of treatment effects when treatment effects are heterogeneous but the treatment decision is unrelated to treatment effects heterogeneity (ATT = ATU = ATE = LATE). ATE, average treatment effects for all patients; ATT, average treatment effects for the treated patients; ATU, average treatment effects for the untreated patients; IV, instrumental variable; LATE, local average treatment effects for the marginal patients.

In the case illustrated here, patients with a higher treatment benefit (darker color) are more likely to be selected for treatment. This scenario applies when there is clinical evidence or common knowledge in clinical practice suggesting that certain subgroups of patients are more apt to benefit from a particular treatment and physicians sort patients with a greater treatment benefit toward treatment—effect modification factors are known and are related to treatment decisions. Under this scenario, ATT > LATE > ATU, and RA estimates of ATT will be greater than IV estimates of LATE (25–29, 31).

The directional association between ATT and LATE in model 3 is not general. Alternative scenarios may exist in which treatment is sorted toward patients with a lower intended treatment benefit. This can occur when the harm associated with unintended adverse side effects is positively correlated with the intended treatment benefit and net treatment benefit (the intended treatment benefit offset by the risk of unintended harm) is inversely associated with the intended treatment benefit. For example, elderly patients with the most potential survival gain from a treatment may be more fragile, with a higher number of comorbid conditions, and as a result may have a greater risk of major unintended adverse side effects (35, 44). Younger patients may have less to gain from the treatment but little risk of side effects. This alternative scenario can also occur when patients with a higher expected treatment benefit are less likely to receive treatment because of disparities in access to health care and quality of care. In these cases, it is possible that the true relation among treatment effect averages will be ATT < LATE < ATU, and RA estimates of ATT will be less than IV estimates of LATE of the intended treatment effects (25–29, 31). The alternative scenario can be illustrated in the model by reversing the direction of the shading.

**DISCUSSION**

We intend for these models to serve as a useful tool with which to discriminate between 4 different treatment effect concepts that are of interest and show how estimators from 2 different analytic traditions (RA and IV) estimate these distinct treatment effects. Potential limitations of these illustrations are that treatment decisions can be related to multiple outcomes of a treatment and the evaluation of the net benefit of all outcomes. Physicians can sort treatment toward patients in some systematic way; however, the sorting is not based on the heterogeneity in solely one particular treatment outcome but the net benefit of all outcomes. The relation between RA and IV estimates of one particular treatment outcome will then also rely on the relation between the net benefit of all outcomes and the particular treatment outcome. This complexity cannot be illustrated easily in the diagrams. However, if treatment effects are operationalized as the value of the net benefit of all outcomes, the 3 schematic models still hold theoretically.
In summary, methodological studies have shown that estimators produce estimates of distinct treatment effect concepts. When the respective assumptions are met, the relation between RA estimates and IV estimates is determined by treatment effect heterogeneity and whether treatment choice is related to heterogeneity. For example, when treatment effects are homogeneous or treatment effects are heterogeneous but treatment decisions are unrelated to treatment effect heterogeneity, the investigator should expect RA and IV estimates to be similar. When treatment effects are heterogeneous and treatment is sorted toward patients with a higher intended treatment benefit, one may expect RA estimates of treatment benefit to be greater than IV estimates of the intended treatment effects. To interpret estimates of treatment effects based on IV or RA methods in clinical epidemiology, it is necessary to have a conceptual model for the relation between the heterogeneous responses to treatment and the decision to treat. Researchers may need to include such explication in the discussion of their findings. Prior to analyses, considerations of whether treatment effects are heterogeneous across patients and how treatment heterogeneity and other factors may influence treatment decisions will facilitate appropriate interpretations of both RA and IV estimates.

ACKNOWLEDGMENTS

Author affiliations: Division of Pharmaceutical Outcomes and Policy, Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina (Gang Fang); Department of Epidemiology, College of Public Health, University of Iowa, Iowa City, Iowa (Gang Fang, Elizabeth A. Chrischilles); and Department of Pharmacy Practice and Science, College of Pharmacy, University of Iowa, Iowa City, Iowa (John M. Brooks).

Dr. Gang Fang was supported by the American Heart Association (AHA) National Clinical Research Program (grant CR2610053). This study was also supported in part by an Agency for Healthcare Research and Quality (AHRQ) Centers for Education and Research on Therapeutics cooperative agreement (cooperative agreement no. 5, grant U18 HSO16094).

The AHA and AHRQ played no role in the design and conduct of the study; the collection, analysis, and interpretation of the data; or the preparation, review, or approval of the manuscript for publication. The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the AHA or AHRQ.

Conflict of interest: none declared.

REFERENCES


