**Annals of Internal Medicine** | RESEARCH AND REPORTING METHODS

# Potential Bias of Instrumental Variable Analyses for Observational Comparative Effectiveness Research

Laura Faden Garabedian, PhD; Paula Chu, MS; Sengwee Toh, ScD; Alan M. Zaslavsky, PhD; and Stephen B. Soumerai, ScD

Instrumental variable analysis is an increasingly popular method in comparative effectiveness research (CER). In theory, the instrument controls for unobserved and observed patient characteristics that affect the outcome. However, the results of instrumental variable analyses in observational settings may be biased if the instrument and outcome are related through an unadjusted third variable: an "instrument–outcome confounder."

The authors identified published CER studies that used instrumental variable analysis and searched the literature for potential confounders of the most common instrument–outcome pairs. Of the 187 studies identified, 114 used 1 or more of the 4 most common instrument categories: distance to facility, regional variation, facility variation, and physician variation. Of these, 65 used mortality as an outcome. Potential unadjusted instrument–outcome confounders were observed in all studies, including patient race, socioeconomic status, clinical risk factors, health status, and urban or rural residency; facility and procedure volume; and co-occurring treatments. Only 4 (6%) instrumental variable CER studies considered potential instrument–outcome confounders outside the study data. Many effect estimates may be biased by the failure to adjust for instrument–outcome confounding. The authors caution against overreliance on instrumental variable studies for CER.

*Ann Intern Med.* 2014;161:131-138. doi:10.7326/M13-1887 **www.annals.org**
For author affiliations, see end of text.

Patients, providers, and payers are increasingly relying on comparative effectiveness research (CER), which compares the benefits and risks of alternative clinical and health care delivery methods (1) to inform evidence-based health care decision making. Because CER is intended to improve patient care and guide health care resource allocation, its validity is crucial. Randomized, controlled trials (RCTs) are the gold standard for identifying the causal impact of a treatment or policy because random treatment assignment usually ensures that study and comparison groups are equivalent with respect to variables that affect the outcome (that is, confounders). However, because RCTs are not always feasible or generalizable, CER relies heavily on observational studies (2, 3), which are susceptible to confounding bias and other threats to validity (4).

Instrumental variable analysis is recommended as a method to establish causal conclusions from observational CER studies (2, 5–7). As an example of this method, several studies have used relative distance to hospitals as an instrument in analyses aimed at estimating the effects on mortality of treatment with invasive cardiac procedures, specifically cardiac catheterization, after myocardial infarction (8–14). Researchers classify each hospital in the study region as a catheterization or noncatheterization hospital on the basis of the presence of a catheterization laboratory or the overall intensity (or volume) of catheterizations. Patients are assigned a value of the binary instrument based on whether they live closer to a catheterization hospital (making them more likely to receive the procedure) or a noncatheterization hospital. This instrumental variable analysis assumes that, similar to random assignment, the relative distance between a patient's residence and a cardiac catheterization hospital predicts treatment choice independently of all characteristics (such as age, socioeconomic status, health status, or use of life-saving medications) that usually confound the relationship between treatment choice and outcome.

In theory, instruments exploit variation in treatment assignment that allows causal inferences similar to those from RCTs. Random treatment assignment is an ideal instrument. In instrumental variable analysis done in observational settings, an instrument shares with experimental randomization certain characteristics that theoretically yield a causal inference. First, potential outcomes (15) for each patient (that is, the outcomes the patient would have under treatment and control conditions) are unrelated to the treatment status of other patients (the stable unit treatment value assumption). Second, the instrument affects receipt of the treatment of interest. Third, this effect is always in the same direction (monotonicity). Fourth, the instrument assigns treatment randomly, meaning that unobserved and observed patient characteristics that affect the outcome are similar in the treatment and comparison groups (ignorable treatment assignment). Finally, the instrument has an effect on the outcome only through the treatment assignment (the exclusion restriction) (16, 17).

Although instrumental variable analysis is mathematically valid under these 5 assumptions (5, 16–25), it is difficult to implement in practice (17, 26). There is a consensus that more research on the validity of instruments in observational CER is needed, particularly concerning violations of the ignorable treatment assignment and exclusion restriction assumptions (2, 5, 27–29). At least one of these assumptions is violated if the instrument is related to the outcome through an unadjusted third variable, which we call an "instrument–outcome confounder" (**Figure 1**).

*See also:*

**Web-Only**
Supplements

**Key Summary Points**

Instrumental variable analysis is an increasingly popular method to establish causal conclusions from observational comparative effectiveness research (CER) studies.

The instruments most commonly used in these studies are distance to facility, regional variation, facility variation, and physician variation.

Instrumental variable analysis relies on several assumptions, some of which are empirically unverifiable and often suspect. The results of instrumental variable analyses may be biased substantially if the instrument and outcome are related through an unadjusted third variable: an "instrument–outcome confounder."

Evidence of potential instrument–outcome confounders was found for all 65 CER studies that used the 4 most common instruments and a mortality outcome.

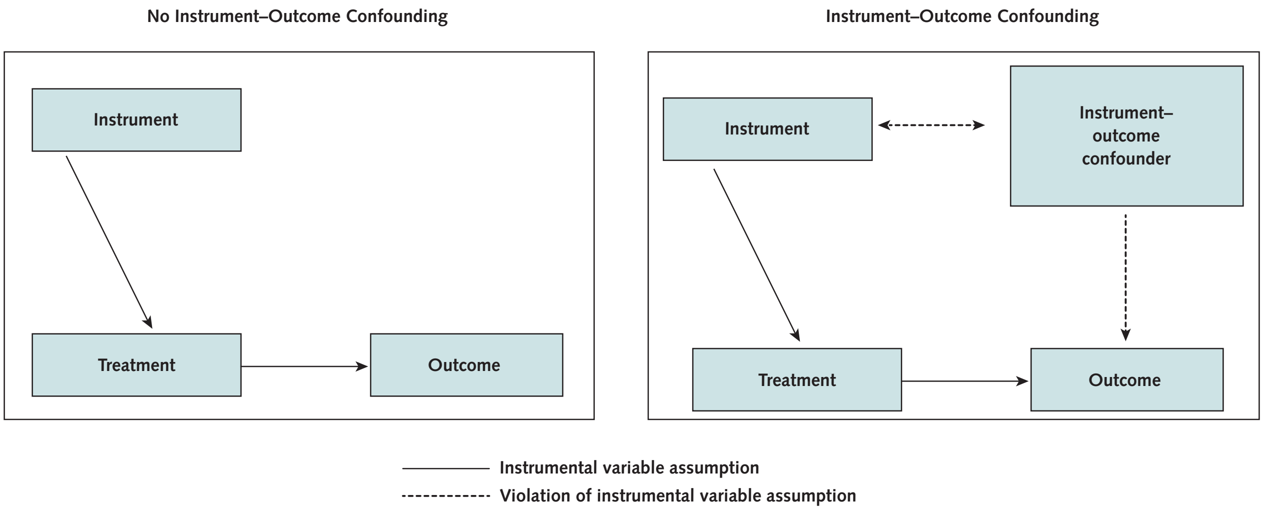
Findings from CER studies using instrumental variables should be evaluated critically for possible confounding.

Instrumental variable analysis estimates of causal effects may be biased if an instrument–outcome confounder has an effect on both the instrument and the outcome (violating the ignorable treatment assumption) or mediates an effect of the instrument on the outcome (violating the exclusion restriction). Although these assumptions are technically unverifiable, the identification of instrument–outcome confounders through other sources provides evidence that an instrumental variable estimate may be biased. However, few researchers have searched for evidence of instrument–outcome confounders outside their own, often limited, data (27, 29).

The instrumental variable analysis in the aforementioned example assumes that the association between relative distance to the hospital (the instrument) and mortality (the outcome) is due only to the effect of relative distance on treatment assignment after control for observed variables (**Figure 1**). A plausible instrument–outcome confounder is rural residence. Patients living in rural areas are less likely to live close to a catheterization hospital; thus, the instrument is associated with rurality (8). Furthermore, ample evidence shows that rural residence is associated with several risk factors for mortality (30–32). Therefore, an instrumental variable analysis would probably overstate the effect of catheterization because patients in the comparison group are, on average, sicker and more likely to die.

In this study, we review relevant literature to identify instruments in CER, evaluate trends in the use of instruments in published CER studies, examine whether instrumental variable CER studies clearly state and attempt to address the assumption of no instrument–outcome confounding, and identify the potential existence and effect of instrument–outcome confounders for commonly used instruments. We list instrument–outcome confounders that may compromise the validity of CER studies that use some of the most common instruments. We conclude by assessing the limitations and potential of instrumental variable analysis in CER.

*Figure 1.* **Instrumental variable assumptions.**




The instrumental variable method substitutes actual random assignment to treatment with an instrument, a variable that predicts treatment assignment but is not related to all other factors that influence the outcome. This method relies on 5 critical assumptions (see text). Instrumental variable estimates of causal effects may be biased if a third variable, an instrument–outcome confounder, has an effect on both the instrument and the outcome (violating the ignorable treatment assumption) or mediates an effect of the instrument on the outcome (violating the exclusion restriction).

*Table 1.* **Description of the 4 Most Commonly Used Instrument Categories in Comparative Effectiveness Research**

| Instrument Category | Studies, *n (%)** | How Instrument Assigns to Treatment Status† | How Instrument Defines Treatment Status† | Example of Instrument Treatment Assignment† |
|---|---|---|---|---|
| Distance to facility | 38 (20) | Distance‡ from patient's residence to facility of interest§ | Facility level: high vs. low treatment rate, existence of specialty provider or unit, special designation (e.g., trauma, teaching) | Instrument assigns patient who resides closer to a hospital with a high cardiac catheterization rate as "treated" |
| Regional variation | 49 (26) | Treatment patterns (e.g., local practice styles) in region where patient lives or is treated | Regional level: high vs. low treatment rate, policies that affect practices in region, provider supply or market share | Instrument assigns patient who resides in an area with a high cardiac catheterization rate as "treated" |
| Facility variation | 22 (12) | Treatment patterns (e.g., local practice styles) in facility where patient is treated§ | Facility level: high vs. low treatment rate, procedure volume, provision of specific services, indicator of quality measures | Instrument assigns patient who is treated in hospital with a high cardiac catheterization rate as "treated" |
| Physician variation | 14 (8) | Treatment patterns (e.g., preference) of treating physician | Physician level: high vs. low treatment rate, most recent prescription in therapeutic area to a new patient | Instrument assigns patient who is treated by physician with a high cardiac catheterization rate as "treated" |

* Studies that used instruments from multiple categories were counted more than once. Percentages are based on 187 total instrumental variable comparative effectiveness research studies.
† "Treatment" refers to instrument assignment to treatment group rather than actual receipt of treatment.
‡ Can be measured in absolute or relative terms using various methods (e.g., straight-line/Euclidean distance or travel time). Slightly more than half (21 studies [55.3%]) of distance instruments used relative, or differential, distance to predict treatment (e.g., distance from patient's home to high–procedure rate hospital minus distance from patient's home to low–procedure rate hospital). The rest of the studies used absolute distance (e.g., distance from patient's home to high–procedure rate hospital).
§ Often a hospital.

## Methods

### Study Selection

We conducted a systematic review in PubMed, EconLit, PsycINFO, Social Services Abstracts, Social Sciences Citation Index, and Web of Science to identify instrumental variable CER studies that were published in an English-language, peer-reviewed journal through 31 December 2011 and conducted in the United States and other industrialized countries. Specific search terms are provided in **Table 1** of **Supplement 1** (available at www.annals.org).

We used the Institute of Medicine's broad definition of CER, which includes both patient-level clinical interventions and system-level health care policies (1). We included noninterventional studies (for example, a study on the association between school junk food exposure and obesity) if the topic was amenable to clinical interventions or policy changes and the study included health-related outcomes. We excluded studies that were purely methodological, used only simulated data, or applied instrumental variable methods in an RCT. We also excluded studies that used Mendelian randomization as an instrument to elucidate biological mechanisms of disease (33) and studies that used an instrument to adjust for the effects of measurement error (34).

### Analysis of Instrumental Variable CER Studies

We created a database of instrumental variable CER studies and catalogued them by year of publication, country, type of intervention, study population, type of instrument, strength of instrument, and outcome. We measured the trend in use of instruments in published CER studies by year and identified the most commonly used instrument–outcome pairs. For each article that used one of these pairs, we also gathered information on the type of data set used and whether instrumental variable analysis was the sole type of analysis used. Finally, we assessed whether each instrumental variable CER article stated the assumption of no instrument–outcome confounding and attempted to show, via additional analyses or discussion, whether the assumption was met.

### Identification of Instrument–Outcome Confounders

Instrument–outcome confounders are variables that are related to both the instrument and the outcome of interest, conditional on measured covariates. They violate the causal inference assumption that the instrument is independent of potential outcomes (15) and suggest that the instrument is not equivalent to random assignment (18). For the purposes of this paper, we included as instrument–outcome confounders variables that have an effect on both the instrument and the outcome (such as rurality) or that mediate an effect of the instrument on the outcome (such as receipt of another life-saving treatment).

We used a structured but not exhaustive search in PubMed and other databases to identify peer-reviewed articles that provided evidence of confounding for the most commonly used instrument–outcome pairs. Specific search terms and strategies are provided in **Table 2** of **Supplement 1**. We rereviewed the instrumental variable CER articles to determine which studies controlled for the potential instrument–outcome confounders we identified.

This study was reviewed by the Harvard Pilgrim Health Care Institutional Review Board and was deemed not to be human subjects research.

## Results

### Systematic Review

A total of 1024 studies were reviewed, and 187 met our eligibility criteria (**Figure** of **Supplement 1**). Use of

*Table 2.* **Potential Instrument–Outcome Confounders for the Most Commonly Used Instruments and a Mortality Outcome**

| Confounder Category*, by Instrument Category | Potential Confounders | Reference |
|---|---|---|
| **Distance to facility instrument†** | | |
| Geographic location | Urban/rural, U.S. region, absolute distance (for relative difference) | 35, 36 |
| Patient characteristics | Race, education, income, age, insurance status, health status/comorbid conditions, health behaviors | 36–38 |
| Treatment characteristics | Receipt of other treatments, time to treatment, transfer status | 39, 40 |
| Facility characteristics | Procedure volume, facility volume, clinical services offered, departments, teaching status, profit status, trauma designation, delivery system type, practice type | – |
| **Regional variation instrument** | | |
| Geographic location | Urban/rural, U.S. region | 35, 41 |
| Patient characteristics | Race, education, income, age, insurance status, health status/comorbid conditions, health behaviors | 35–40, 42 |
| Provider supply | Number of hospital beds or nursing homes | 41, 43 |
| Technology adoption and utilization | Invasive cardiac procedures, radical prostatectomies, prescribing behavior, practice patterns | 44, 45 |
| Treatment characteristics | Receipt of other treatments, time to treatment, transfer status | 36–40 |
| Facility characteristics | Procedure volume, facility volume, clinical services offered, departments, teaching status, profit status, trauma designation, delivery system type, practice type | – |
| **Facility variation instrument** | | |
| Geographic location | Urban/rural | 38, 46 |
| Patient characteristics | Race, education, income, age, insurance status, health status/comorbid conditions, health behaviors | 35–40, 42 |
| Facility characteristics | Procedure volume, facility volume, clinical services offered, departments, teaching status, profit status, trauma designation, delivery system type, practice type | 47–49 |
| Treatment characteristics | Receipt of other treatments | 36–40 |
| **Physician variation instrument** | | |
| Physician characteristics | Age, sex, specialty, board certification, physician volume | 50, 51 |
| Patient characteristics | Race, education, income, age, sex, insurance status, health status/comorbid conditions, health behaviors | 52, 53 |
| Treatment characteristics | Receipt of other treatments | 54, 55 |
| Health system characteristics | Reimbursement policies and regional variation and facility variation confounders, such as geographic location, provider supply, technology adoption and utilization, and facility characteristics | 56, 57 |

* We found evidence in the literature for all of these potential instrument–outcome confounders. The references listed are illustrative. **Supplement 2** provides a more comprehensive list.
† Applicable to both absolute and relative distance instruments unless otherwise noted.

instrumental variable methods in CER has accelerated since the early 1990s (**Figure 2**), with a large spike in U.S.-based studies in 2010 and 2011, possibly due to increased federal funding for CER as part of the economic stimulus in 2009 (3).

More than half (114 of 187 [61%]) of the instrumental variable CER studies used at least 1 of the 4 most common instrument categories: distance to facility (20%), regional variation (26%), facility variation (12%), and physician variation (8%) (**Table 1**). Of these, 65 assessed mortality, the most common outcome. We focus on these articles in the remainder of this section.

### Instrument–Outcome Confounders

We found strong evidence of instrument–outcome confounders that may violate the assumption that the instrument is only related to the outcome through the treatment (**Table 2** and **Supplement 2**, available at www.annals.org). Potential confounders of the 4 most commonly used instruments and a mortality outcome include patient race, socioeconomic status, risk factors for mortality, health status, and urban or rural residency as well as facility and procedure volume. Many other confounders are less researched but well-recognized, such as factors associated with time to treatment (for example, door-to-needle time), receipt of other treatments (for example, life-saving medications, such as thrombolytics), and facility characteristics (for example, teaching hospital status) that are associated with mortality.

### Control for Instrument–Outcome Confounding in the Current Literature

We reviewed the 65 instrumental variable CER studies that used 1 of the 4 most common instruments and a mortality outcome to determine whether the authors discussed or controlled for the potential instrument–outcome confounders. Most (54 of 65 [83%]) stated the assumption of no instrument–outcome confounding (**Supplement 3**, available at www.annals.org). More than half (41 of 65 [63%]) provided additional analyses or discussion to determine whether the assumption was met. However, few (4 of 65 [6%]) considered potential instrument–outcome confounders outside of those measured in the study data (that is, retrieved data from an external database [58] or referenced literature [59–61]).

None of the studies in our review controlled for all potential instrument–outcome confounders we identified in the literature (**Table 3**; **Supplement 3** provides results by individual study). Although most instrumental variable CER studies controlled for at least 1 patient health status variable and race, less than half of the studies in each instrument category controlled for income, education, urban or rural location, and volume (**Table 3**). Studies that used a regional variation instrument performed slightly better: 70% controlled for patient income, and 52% controlled for urban or rural location.

Most (48 of 65 [74%]) of these instrumental variable CER studies used multiple analytic methods—often multivariable regression and propensity score matching—and most (39 of 65 [60%]) used data from only administrative databases, such as electronic medical records and insurer claims (**Supplement 3**). We found that although most (60 of 65 [92%]) studies assessed the strength of the instrument, the reported strength of the instruments varied widely.

## Discussion

Instrumental variable analysis is an increasingly popular method for CER, perhaps in part because, unlike other statistical methods that control for observed confounders (such as multivariable regression and propensity score matching), instruments theoretically control for both observed and unobserved confounders. The identification of instrument–outcome confounders suggests that key instrumental variable assumptions are violated. We found widespread evidence of potential instrument–outcome confounders of the 4 most popular instruments.

The validity of instrumental variable analysis might improve if instrument–outcome confounders are measured and controlled for. However, none of the 65 studies that used 1 of the 4 most commonly used instruments and a mortality outcome controlled for all potential confounders identified in the literature. Furthermore, some confounders (such as race [62]) are difficult to measure or collect. Controlling for inadequately measured confounders will still result in residual bias, a problem common to all observational analyses (63). Also, although most of these studies stated the assumption of no instrument–outcome confounding, only 4 (58–61) went beyond the study data to identify potential confounders and only 1 (58) performed a sensitivity analysis to assess the robustness of study findings in the presence of unmeasured confounding.

Some instrument–outcome confounders have particularly great potential to introduce bias. For example, geography, race, and income are strongly linked to mortality (30–32, 64). In the United States, the gap between race–county combinations with the highest and lowest life expectancies is more than 35 years (30). We found evidence

*Figure 2.* **Instrumental variable comparative effectiveness research studies (*n* = 187), by year of publication.**

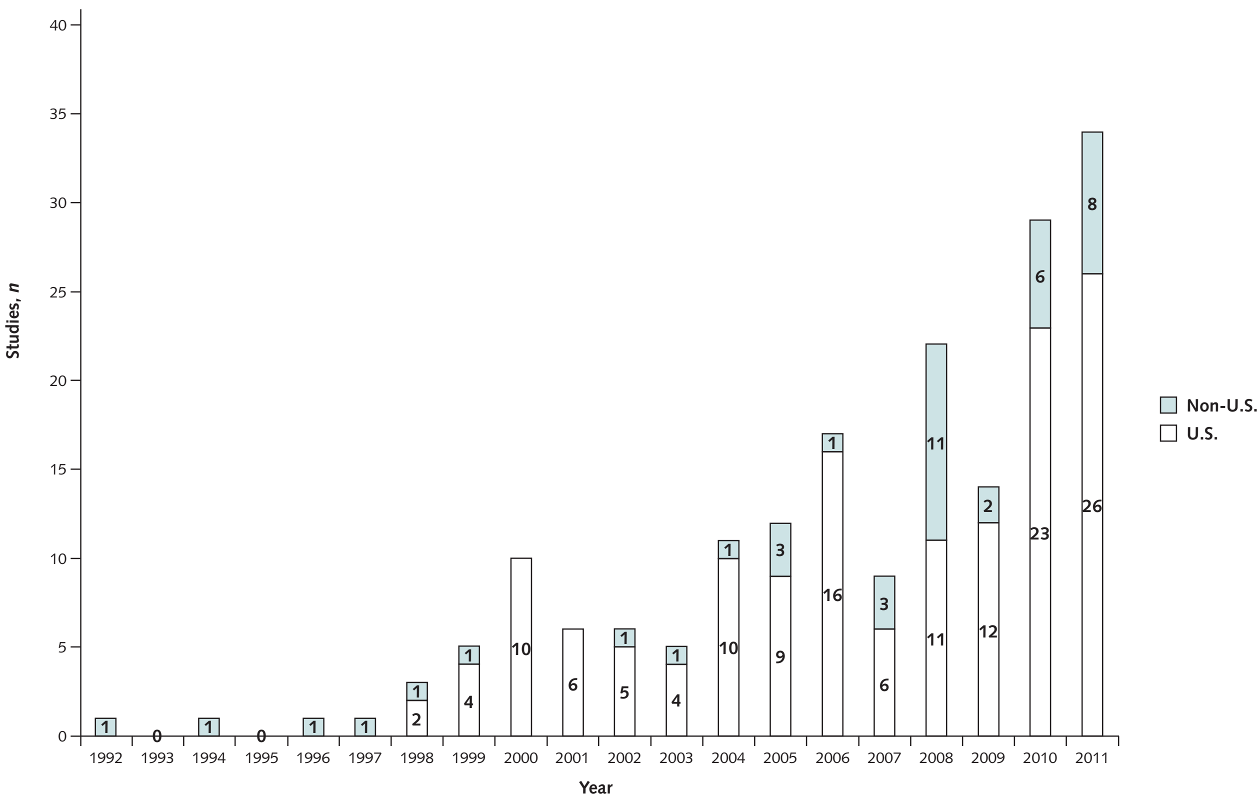


*Table 3.* **Studies That Controlled for Potential Instrument–Outcome Confounders, by Instrument Category**

| Confounder | Studies, %* | | | |
|---|---|---|---|---|
| | Distance to Facility (*n* = 27) | Regional Variation (*n* = 23) | Facility Variation (*n* = 14) | Physician Variation (*n* = 9) |
| Patient race | 78 | 70 | 79 | 67 |
| Patient income | 44 | 70 | 14 | 22 |
| Patient education | 15 | 22 | 14 | 0 |
| Patient comorbid conditions/health status | 100 | 83 | 86 | 100 |
| Urban/rural (patient residence or facility location) | 44 | 52 | 7 | 22 |
| Volume (procedure)† | 4 | 0 | 27 | 11 |
| Volume (facility)† | 41 | 41 | 39 | 11 |

* Analysis was limited to studies that used ≥1 of the 4 most commonly used instruments and a mortality outcome (see **Supplement 3**). Studies that used >1 of the 4 instruments appear in multiple columns.
† Studies that used procedure or facility volume as an instrument or independent variable were removed from the denominator.

suggesting that the most popular instruments in CER may be associated with these variables. If these potential instrument–outcome confounders are not controlled for, they may bias the instrumental variable analysis effect estimate.

Quantitative assessments of bias are possible under some circumstances (65). For some instrument–outcome confounders, the direction of the bias is predictable. The confounders will cause an overestimation of the treatment effect if predictors of reduced mortality are correlated with each other. For example, patients who live close to a hospital of high technical quality and are therefore assigned to the "treatment group" in an instrumental variable analysis are more likely to receive the treatment of interest, such as cardiac catheterization, as well as other high-quality, time-sensitive, life-saving treatments that improve survival (47, 48, 50). These confounders will result in an overestimation of the beneficial effect of the treatment: The confounded instrumental variable estimate incorrectly attributes the positive effect of the other treatments and aspects of care to the treatment of interest. However, in general, the direction of the bias introduced by instrument–outcome confounders is study-dependent and may be hard to predict unless the confounder is measured in the study population.

Confounding is a frequent concern in observational CER studies because of limited availability of adequate information on important confounders, such as sociodemographic and health system characteristics, particularly in administrative health care databases. However, confounding is especially problematic in instrumental variable analyses compared with other observational methods because even minor bias introduced by instrument–outcome confounders is magnified when the instrument is weak (65), the bias may be exacerbated if the instrument becomes weaker after confounders are controlled for, and the relative opaqueness of instrumental variable analysis and the naive claims often made for its ability to correct for confounding (27–29) may make it more likely that researchers will not adjust for confounders.

Researchers are responsible for assessing potential instrument–outcome confounding in specific instrumental variable CER studies (7, 28, 29) by controlling for potential confounders in the study data or using plausible estimates of their effects, acquired from other sources, in sensitivity analysis (or falsification tests [29]) if the confounders are not represented by available variables. The assumption that the instrument is only related to the outcome through the treatment may apply best to specific, focused, plausibly exogenous interventions or events, such as natural experiments (for example, the Oregon Medicaid lottery [66]) or changes in policy or technology (67). Conversely, the instruments we evaluated, such as distance and region, are long-term properties of an area or subgroup that are likely to have causal effects on outcomes of interest through multiple pathways.

This study has several limitations. First, our search for instrument–outcome confounders focused on only the most commonly used instrument categories. Second, our systematic review and search for instrument–outcome confounders included only studies published from January 1992 (the date of the first identified article) to December 2011. Finally, the studies that we cite as evidence for confounders may themselves have limitations, such as lack of control for other confounders.

In conclusion, the use of instrumental variable analyses in CER is often a reaction to limited resources and data availability. Although the instrumental variable method is theoretically sound when all assumptions are met, we found that, in practice, most CER studies that performed this type of analysis were overconfident in asserting that instrument–outcome confounding was not present; in particular, the 4 most commonly used instruments should be used cautiously in CER because their results may be biased. Any instrumental variable analysis that does not control for likely instrument–outcome confounders should be interpreted with caution. Although no observational method can completely eliminate confounding, we recommend against treating instrumental variable analysis as a solution to the inherent biases in observational CER studies.

From Harvard Medical School, Harvard Pilgrim Health Care Institute, and Harvard University, Boston, Massachusetts.

**Note:** This project was awarded first place at the 2012 Association for Public Policy Analysis and Management Fall Research Conference Poster Awards.

**Disclaimer:** The authors of this article are responsible for its content. Statements in the article should not be construed as endorsement by the Agency for Healthcare Research and Quality or the U.S. Department of Health and Human Services.

**Acknowledgment:** The authors thank Sebastian Schneeweiss, MD, ScD, and John Seeger, PharmD, DrPH, both from the Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, for their insightful feedback on the study design and manuscript and Elizabeth Garry, MPH, from the Division of Pharmacoepidemiology and Pharmacoeconomics, for providing helpful administrative support throughout the project.

**Financial Support:** This work was supported as a subcontract from the Brigham and Women's Hospital DEcIDE Method Center under contract 290-05-0016I from the Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services, as part of the Developing Evidence to Inform Decisions about Effectiveness (DEcIDE) program. Dr. Soumerai received grant support from the Centers for Disease Control and Prevention's Natural Experiments for Translation in Diabetes.

**Disclosures:** Disclosures can be viewed at www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNum=M13-1887.

**Requests for Single Reprints:** Laura Faden Garabedian, PhD, Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, 133 Brookline Avenue, 6th Floor, Boston, MA 02215; e-mail, laura.garabedian@post.harvard.edu.

Current author addresses and author contributions are available at www.annals.org.

## References

1. **Institute of Medicine.** Initial National Priorities for Comparative Effectiveness Research. Washington, DC: Institute of Medicine; 2009. Accessed at www.iom.edu/Reports/2009/ComparativeEffectivenessResearchPriorities.aspx on 31 January 2011.
2. **Tunis SR, Benner J, McClellan M.** Comparative effectiveness research: Policy context, methods development and research infrastructure. Stat Med. 2010;29:1963-76. [PMID: 20564311]
3. **Garber AM.** How the Patient-Centered Outcomes Research Institute can best influence real-world health care decision making. Health Aff (Millwood). 2011;30:2243-51. [PMID: 22147851]
4. **Shadish WR, Cook TD, Campbell DT.** Experimental and Quasi-experimental Designs for Generalized Causal Inference. Boston: Houghton Mifflin; 2002:33-63.
5. **Brookhart MA, Rassen JA, Schneeweiss S.** Instrumental variable methods in comparative safety and effectiveness research. Pharmacoepidemiol Drug Saf. 2010;19:537-54. [PMID: 20354968]
6. **Johnson ML, Crown W, Martin BC, Dormuth CR, Siebert U.** Good research practices for comparative effectiveness research: analytic methods to improve causal inference from nonrandomized studies of treatment effects using secondary data sources: the ISPOR Good Research Practices for Retrospective Database Analysis Task Force Report—Part III. Value Health. 2009;12:1062-73. [PMID: 19793071]
7. **Patient-Centered Outcomes Research Institute.** PCORI Methodology Standards. Washington, DC: Patient-Centered Outcomes Research Institute; 2013. Accessed at www.pcori.org/assets/PCORI-Methodology-Standards.pdf on 14 February 2013.
8. **McClellan M, McNeil BJ, Newhouse JP.** Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality? Analysis using instrumental variables. JAMA. 1994;272:859-66. [PMID: 8078163]
9. **McClellan M.** Are the returns to technological change in health care declining? Proc Natl Acad Sci U S A. 1996;93:12701-8. [PMID: 8917482]
10. **Brooks JM, McClellan M, Wong HS.** The marginal benefits of invasive treatments for acute myocardial infarction: does insurance coverage matter? Inquiry. 2000;37:75-90. [PMID: 10892359]
11. **Glickman ME, Normand ST.** The derivation of a latent threshold instrumental variables model. Stat Sin. 2000;10:517-44.
12. **Beck CA, Penrod J, Gyorkos TW, Shapiro S, Pilote L.** Does aggressive care following acute myocardial infarction reduce mortality? Analysis with instrumental variables to compare effectiveness in Canadian and United States patient populations. Health Serv Res. 2003;38:1423-40. [PMID: 14727781]
13. **Cutler DM.** The lifetime costs and benefits of medical technology. J Health Econ. 2007;26:1081-100. [PMID: 18023900]
14. **Pilote L, Beck CA, Eisenberg MJ, Humphries K, Joseph L, Penrod JR, et al.** Comparing invasive and noninvasive management strategies for acute myocardial infarction using administrative databases. Am Heart J. 2008;155:42-8. [PMID: 18082487]
15. **Little RJ, Rubin DB.** Causal effects in clinical and epidemiological studies via potential outcomes: concepts and analytical approaches. Annu Rev Public Health. 2000;21:121-45. [PMID: 10884949]
16. **Angrist JD, Imbens GW, Rubin DB.** Identification of causal effects using instrumental variables. J Am Stat Assoc. 1996;91:444-55.
17. **Landrum MB, Ayanian JZ.** Causal effect of ambulatory specialty care on mortality following myocardial infarction: a comparison of propensity score and instrumental variable analyses. Health Serv Outcomes Res Methodol. 2001;2:221-45.
18. **Angrist JD, Pischke JS.** Mostly Harmless Econometrics: An Empiricist's Companion. Princeton: Princeton Univ Pr; 2009.
19. **Wooldridge JM.** Introductory Econometrics: A Modern Approach. 3rd ed. Mason, OH: Thomson; 2006.
20. **Newhouse JP, McClellan M.** Econometrics in outcomes research: the use of instrumental variables. Annu Rev Public Health. 1998;19:17-34. [PMID: 9611610]
21. **Greenland S.** An introduction to instrumental variables for epidemiologists. Int J Epidemiol. 2000;29:722-9. [PMID: 10922351]
22. **Martens EP, Pestman WR, de Boer A, Belitser SV, Klungel OH.** Instrumental variables: application and limitations. Epidemiology. 2006;17:260-7. [PMID: 16617274]
23. **Glymour MM.** Natural experiments and instrumental variable analyses in social epidemiology. In: Oakes JM, Kaufman JS, eds. Methods in Social Epidemiology. San Francisco: J Wiley; 2006:429-68.
24. **Hernán MA, Robins JM.** Instruments for causal inference: an epidemiologist's dream? Epidemiology. 2006;17:360-72. [PMID: 16755261]
25. **Grootendorst P.** A review of instrumental variables estimation of treatment effects in the applied health sciences. Health Serv Outcomes Res Methodol. 2007;10:159-79.
26. **Rosenbaum PR.** Design of Observational Studies. New York: Springer; 2010:139-40.
27. **Chen Y, Briesacher BA.** Use of instrumental variable in prescription drug research with observational data: a systematic review. J Clin Epidemiol. 2011;64:687-700. [PMID: 21163621]
28. **Davies NM, Smith GD, Windmeijer F, Martin RM.** Issues in the reporting and conduct of instrumental variable studies: a systematic review. Epidemiology. 2013;24:363-9. [PMID: 23532055]
29. **Swanson SA, Hernán MA.** Commentary: how to report instrumental variable analyses (suggestions welcome). Epidemiology. 2013;24:370-4. [PMID: 23549180]
30. **Murray CJ, Kulkarni SC, Michaud C, Tomijima N, Bulzacchelli MT, Iandiorio TJ, et al.** Eight Americas: investigating mortality disparities across races, counties, and race-counties in the United States. PLoS Med. 2006;3:e260. [PMID: 16968116]
31. **Cosby AG, Neaves TT, Cossman RE, Cossman JS, James WL, Feierabend N, et al.** Preliminary evidence for an emerging nonmetropolitan mortality penalty in the United States. Am J Public Health. 2008;98:1470-2. [PMID: 18556611]
32. **Cossman JS, James WL, Cosby AG, Cossman RE.** Underlying causes of the emerging nonmetropolitan mortality penalty. Am J Public Health. 2010;100:1417-9. [PMID: 20558803]
33. **Didelez V, Sheehan N.** Mendelian randomization as an instrumental variable approach to causal inference. Stat Methods Med Res. 2007;16:309-30. [PMID: 17715159]
34. **Hu Y, Schennach SM.** Instrumental variable treatment of nonclassical measurement error models. Econometrica. 2008;76:195-216.

35. **Abrams TE, Vaughan-Sarrazin M, Fan VS, Kaboli PJ.** Geographic isolation and the risk for chronic obstructive pulmonary disease-related mortality: a cohort study. Ann Intern Med. 2011;155:80-6. [PMID: 21768581]

36. **Khan JA, Casper M, Asimos AW, Clarkson L, Enright D, Fehrs LJ, et al.** Geographic and sociodemographic disparities in drive times to Joint Commission-certified primary stroke centers in North Carolina, South Carolina, and Georgia. Prev Chronic Dis. 2011;8:A79. [PMID: 21672403]

37. **Lorch SA, Silber JH, Even-Shoshan O, Millman A.** Use of prolonged travel to improve pediatric risk-adjustment models. Health Serv Res. 2009;44:519-41. [PMID: 19207591]

38. **Ambardekar AV, Fonarow GC, Dai D, Peterson ED, Hernandez AF, Cannon CP, et al; Get With The Guidelines Steering Committee and Hospitals.** Quality of care and in-hospital outcomes in patients with coronary heart disease in rural and urban hospitals (from Get With the Guidelines-Coronary Artery Disease Program). Am J Cardiol. 2010;105:139-43. [PMID: 20102907]

39. **Vukmir RB.** Survival from prehospital cardiac arrest is critically dependent upon response time. Resuscitation. 2006;69:229-34. [PMID: 16500015]

40. **Maybury RS, Chang DC, Freischlag JA.** Rural hospitals face a higher burden of ruptured abdominal aortic aneurysm and are more likely to transfer patients for emergent repair. J Am Coll Surg. 2011;212:1061-7. [PMID: 21458304]

41. **Hutt E, Elder SJ, Fish R, Min SJ.** Regional variation in mortality and subsequent hospitalization of nursing residents with heart failure. J Am Med Dir Assoc. 2011;12:595-601. [PMID: 21450206]

42. **Welch HG, Sharp SM, Gottlieb DJ, Skinner JS, Wennberg JE.** Geographic variation in diagnosis frequency and risk of death among Medicare beneficiaries. JAMA. 2011;305:1113-8. [PMID: 21406648]

43. **Fisher ES, Wennberg JE, Stukel TA, Skinner JS, Sharp SM, Freeman JL, et al.** Associations among hospital capacity, utilization, and mortality of US Medicare beneficiaries, controlling for sociodemographic factors. Health Serv Res. 2000;34:1351-62. [PMID: 10654835]

44. **Chen J, Krumholz HM, Wang Y, Curtis JP, Rathore SS, Ross JS, et al.** Differences in patient survival after acute myocardial infarction by hospital capability of performing percutaneous coronary intervention: implications for regionalization. Arch Intern Med. 2010;170:433-9. [PMID: 20212179]

45. **Epstein AJ, Polsky D, Yang F, Yang L, Groeneveld PW.** Geographic variation in implantable cardioverter-defibrillator use and heart failure survival. Med Care. 2012;50:10-7. [PMID: 22167063]

46. **Galandiuk S, Mahid SS, Polk HC Jr, Turina M, Rao M, Lewis JN.** Differences and similarities between rural and urban operations. Surgery. 2006;140:589-96. [PMID: 17011906]

47. **Ayanian JZ, Weissman JS.** Teaching hospitals and quality of care: a review of the literature. Milbank Q. 2002;80:569-93, v. [PMID: 12233250]

48. **Hayanga AJ, Mukherjee D, Chang D, Kaiser H, Lee T, Gearhart S, et al.** Teaching hospital status and operative mortality in the United States: tipping point in the volume-outcome relationship following colon resections? Arch Surg. 2010;145:346-50. [PMID: 20404284]

49. **Billingsley KG, Morris AM, Dominitz JA, Matthews B, Dobie S, Barlow W, et al.** Surgeon and hospital characteristics as predictors of major adverse outcomes following colon cancer surgery: understanding the volume-outcome relationship. Arch Surg. 2007;142:23-31. [PMID: 17224497]

50. **Willison DJ, Soumerai SB, Palmer RH.** Association of physician and hospital volume with use of aspirin and reperfusion therapy in acute myocardial infarction. Med Care. 2000;38:1092-102. [PMID: 11078050]

51. **Southern WN, Bellin EY, Arnsten JH.** Longer lengths of stay and higher risk of mortality among inpatients of physicians with more years in practice. Am J Med. 2011;124:868-74. [PMID: 21745651]

52. **Weissman J, Meyers BS, Ghosh S, Bruce ML.** Sociodemographic and clinical factors associated with antidepressant type in a national sample of the home health care elderly. Gen Hosp Psychiatry. 2011;33:587-93. [PMID: 21920609]

53. **Groenwold RH, Hak E, Klungel OH, Hoes AW.** Instrumental variables in influenza vaccination studies: mission impossible?!. Value Health. 2010;13:132-7. [PMID: 19695007]

54. **Brookhart MA, Rassen JA, Wang PS, Dormuth C, Mogun H, Schneeweiss S.** Evaluating the validity of an instrumental variable study of neuroleptics: can between-physician differences in prescribing patterns be used to estimate treatment effects? Med Care. 2007;45:S116-22. [PMID: 17909369]

55. **Parsons C, Johnston S, Mathie E, Baron N, Machen I, Amador S, et al.** Potentially inappropriate prescribing in older people with dementia in care homes: a retrospective analysis. Drugs Aging. 2012;29:143-55. [PMID: 22204669]

56. **Shahinian VB, Kuo YF, Gilbert SM.** Reimbursement policy and androgen-deprivation therapy for prostate cancer. N Engl J Med. 2010;363:1822-32. [PMID: 21047226]

57. **Fischer MA, Choudhry NK, Winkelmayer WC.** Impact of Medicaid prior authorization on angiotensin-receptor blockers: can policy promote rational prescribing? Health Aff (Millwood). 2007;26:800-7. [PMID: 17485759]

58. **Tsai AC, Votruba M, Bridges JF, Cebul RD.** Overcoming bias in estimating the volume-outcome relationship. Health Serv Res. 2006;41:252-64. [PMID: 16430610]

59. **Sundmacher L, Busse R.** The impact of physician supply on avoidable cancer deaths in Germany. A spatial analysis. Health Policy. 2011;103:53-62. [PMID: 21962525]

60. **Weiner BJ, Alexander JA, Shortell SM, Baker LC, Becker M, Geppert JJ.** Quality improvement implementation and hospital performance on quality indicators. Health Serv Res. 2006;41:307-34. [PMID: 16584451]

61. **Gillum LA, Johnston SC.** Influence of physician specialty on outcomes after acute ischemic stroke. J Hosp Med. 2008;3:184-92. [PMID: 18570345]

62. **Zaslavsky AM, Ayanian JZ, Zaborski LB.** The validity of race and ethnicity in enrollment data for Medicare beneficiaries. Health Serv Res. 2012;47:1300-21. [PMID: 22515953]

63. **Carroll CJ, Ruppert D, Stefanski LA, Crainiceanu C.** Measurement Error in Nonlinear Models: A Modern Perspective. 2nd ed. Boca Raton, FL: Chapman and Hall/CRC; 2006.

64. **Geronimus AT, Bound J, Colen CG.** Excess black mortality in the United States and in selected black and white high-poverty areas, 1980-2000. Am J Public Health. 2011;101:720-9. [PMID: 21389293]

65. **Brookhart MA, Schneeweiss S.** Preference-based instrumental variable methods for the estimation of treatment effects: assessing validity and interpreting results. Int J Biostat. 2007;3:Article 14. [PMID: 19655038]

66. **Baicker K, Taubman SL, Allen HL, Bernstein M, Gruber JH, Newhouse JP, et al; Oregon Health Study Group.** The Oregon experiment—effects of Medicaid on clinical outcomes. N Engl J Med. 2013;368:1713-22. [PMID: 23635051]

67. **Angrist JD, Krueger AB.** Instrumental variables and the search for identification: From supply and demand to natural experiments. J Econ Perspect. 2001;15:69-85.

**Current Author Addresses:** Drs. Garabedian, Toh, and Soumerai: Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, 133 Brookline Avenue, 6th Floor, Boston, MA 02215.
Ms. Chu: Harvard PhD Program in Health Policy, 14 Story, 4th Floor, Cambridge, MA 02138.
Dr. Zaslavsky: Department of Health Care Policy, Harvard Medical School, 180 Longwood Avenue, Boston, MA 02115.

**Author Contributions:** Conception and design: L.F. Garabedian, S. Toh, A.M. Zaslavsky, S.B. Soumerai.
Analysis and interpretation of the data: L.F. Garabedian, P. Chu, S. Toh, A.M. Zaslavsky, S.B. Soumerai.
Drafting of the article: L.F. Garabedian.
Critical revision of the article for important intellectual content: L.F. Garabedian, P. Chu, S. Toh, A.M. Zaslavsky, S.B. Soumerai.
Final approval of the article: L.F. Garabedian, P. Chu, S. Toh, A.M. Zaslavsky, S.B. Soumerai.
Statistical expertise: L.F. Garabedian, P. Chu, S. Toh, A.M. Zaslavsky, S.B. Soumerai.
Obtaining of funding: S.B. Soumerai.
Collection and assembly of data: L.F. Garabedian, P. Chu.