In the absence of randomized treatment assignment, comparative effectiveness studies of different health interventions are plagued by selection bias. When multiple treatment options are available, the choice of intervention is often influenced by health system, physician, and patient characteristics that are correlated with health outcomes. For instance, Black patients are less likely to receive major surgical procedures than White patients in the US, and it is well-established that Black race (which is correlated with social and structural determinants of health, including systemic racism) is associated with higher morbidity and mortality. In *JAMA Network Open*, Howard and colleagues use instrumental variable (IV) analysis as a method to control for selection bias in the comparative effectiveness of 2 bariatric surgery procedures (laparoscopic sleeve gastrectomy and laparoscopic gastric bypass) among Medicare beneficiaries. But, as the authors note, "although effective in their ability to reduce bias, instrumental variables can be difficult to identify."

**Key Assumptions of Instrumental Variable Analysis**

The goal is to find an instrument (ie, variable) that randomly predicts whether a person receives one treatment option vs another. The instrument must impact the outcome only through the treatment assignment, and it cannot be related to any patient, physician, or system characteristics that are associated with the outcome. While these IV assumptions are often not empirically testable, researchers implementing the IV method (and readers of IV studies) should think carefully about variables that are potentially correlated with both the IV and the study outcome (ie, IV-outcome confounders), which could bias results and lead to incorrect conclusions. Potential IV-outcome confounders can be identified in the study data, or by searching the literature for variables that are correlated with the instrument and exploring whether those variables are also associated with the study outcome.

**Potential Bias in Geographical Variation Instrumental Variables**

Howard et al use state variation in bariatric procedure rates in the prior year to project whether a patient seeking care in the study year receives sleeve gastrectomy (vs gastric bypass). Geographical variation in procedure rates is 1 of the 4 most commonly used IVs in comparative effectiveness research, but prior research has shown that this instrument is likely biased by IV-outcome confounders. Geographical variation is strongly associated with variables that are correlated with health outcomes, such as race and ethnicity, urban vs rural status, and access to other health interventions. In fact, Howard et al found that Black patients undergoing bariatric procedures were more likely to live in areas with higher sleeve gastrectomy rates (and were therefore assigned to the sleeve gastrectomy treatment group by the IV) than White patients. A 2019 *JAMA Surgery* study found that Black patients had higher rates of complications and health care utilization following bariatric surgery than White patients, even when controlling for surgery type. Therefore, race is an IV-outcome confounder—patients who live in areas with higher rates of sleeve gastrectomy are more likely to be Black and fare worse on bariatric surgery outcomes. An IV analysis using geographical variation in bariatric procedure rates that does not adequately measure and control for a patient’s
race would underestimate the superior effectiveness of sleeve gastrectomy vs gastric bypass, since patients assigned to the sleeve gastrectomy group by the IV are, on average, at higher risk of poor health outcomes. Conversely, not controlling for a variable that is related to both higher sleeve gastrectomy rates and improved health would lead to an overestimate of the treatment effect. While Howard et al controlled for race in their models, imbalance across levels of the instrument on observed IV-outcome confounders calls into question whether there is imbalance on unmeasured IV-outcome confounders that could bias their results.

Use of Insurance Coverage Decisions in Comparative Effectiveness Research

The authors ask: “Can the effects of national insurance coverage decisions serve as an instrumental variable in comparative effectiveness research?” However, the insurance coverage decision was not the instrumental variable used in this study. Howard et al\(^3\) were not able to measure insurance coverage, and their data suggest that all regions elected to cover sleeve gastrectomy (ie, there was no variability in coverage). Therefore, the variation they used for the instrument stems from regional differences in the adoption of the procedure following coverage, which is likely related to health system, physician, and patient characteristics that are related to outcomes. In using this approach, they missed an opportunity to use the policy coverage decision itself as the instrument. They hypothesize that “initiation of insurance coverage for sleeve gastrectomy was associated with significant geographic and temporal variation” in bariatric procedure rates, but they do not take advantage of the temporal variation in this cross-sectional analysis.

An ideal insurance coverage policy for quasi-experimental evaluation is one that creates variation in treatment coverage across groups of patients over time (eg, state-level differences in coverage, which would be conducive to a difference-in-difference design [ie, pre-post with control]\(^6\)) or creates a cutoff that induces treatment variation among a group of patients (eg, age of eligibility for Medicare coverage, which would be conducive to a regression discontinuity design\(^7\)). In cases where there is a national policy, an interrupted time series takes advantage of temporal changes over time and can provide important evidence, even in the absence of a control group.\(^8\) These robust quasi-experimental designs, which estimate the effectiveness of the coverage policy on health outcomes (akin to an intent-to-treat estimate in a randomized clinical trial, since not everyone eligible for coverage receives the treatment), could be coupled with an instrumental variable analysis that uses the coverage decision as an instrument to estimate the efficacy of the treatment.\(^9\)\(^10\)

Instrumental variable analysis is a potentially useful tool in comparative effectiveness research, but it must be used wisely and appropriately in order to produce accurate results to inform patient care.

ARTICLE INFORMATION

Published: August 18, 2022. doi:10.1001/jamanetworkopen.2022.25970

Open Access: This is an open access article distributed under the terms of the CC BY License. © 2022 Garabedian LF. JAMA Network Open.

Corresponding Author: Laura F. Garabedian, PhD, MPH, Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, 401 Park Dr, Ste 401, Boston, MA 02215 (laura.garabedian@post.harvard.edu).

Author Affiliation: Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, Massachusetts.

Conflict of Interest Disclosures: None reported.

Additional Contributions: The author would like to acknowledge Stephen Soumerai, ScD, for providing feedback on this commentary. He was not compensated for his contribution.

Downloaded From: https://jamanetwork.com/ on 08/29/2022
REFERENCES


7. Venkataramani AS, Bor J, Jena AB. Regression discontinuity designs in healthcare research. BMJ. 2016;352:i2126. doi:10.1136/bmj.i2126

