

## JAMA Guide to Statistics and Methods

## Regression Discontinuity Design

Matthew L. Maciejewski, PhD; Anirban Basu, PhD

**In the May 3, 2019, issue** of *JAMA Network Open*, Sukul and colleagues<sup>1</sup> used a regression discontinuity design (RDD) study to determine if a 2011 change in Medicare policy, expanding the number of secondary diagnostic codes allowed in Medicare billing from 9 to 24, was associated with a change in the observed disease severity of Medicare patients, independent of any real change in these patients' medical condition.

The apparent disease severity, based on secondary diagnostic codes, might increase because of real changes in the overall disease burden over time, as an artifact from the listing of additional codes due to the greater number allowed, or some combination of these factors. The RDD is a method for separating the contribution of the abrupt change in policy—discontinuity—to observed increases in outcomes from the contribution of other factors that may produce more gradual changes.<sup>2</sup>

**Use of the RDD Method****Why Is an RDD Study Used?**

The RDD method is used to study the outcomes related to an abrupt change when it is not possible to randomly assign patients to the conditions before and after the change. A goal of an RDD is to minimize the effect of confounding on the estimated effect of a policy or treatment change.<sup>2</sup> Unobserved confounding is particularly concerning in nonrandomized studies because this bias cannot be completely removed using conventional statistical methods, such as regression or propensity score-based analysis.<sup>3</sup> An RDD attempts to minimize the risk for unobserved confounding when generating the association between an exposure and the change in the outcome of interest.<sup>4</sup>

**Description of the RDD Method**

The RDD approach relies on having a continuous variable, the "running variable," for which the levels above a certain cutoff abruptly change the probability of receiving one treatment or the other, resulting in a "natural experiment." The 2011 change in Medicare policy for secondary diagnosis codes represents just such a natural experiment, and calendar time represents the running variable. Because the presence of potentially confounding factors should be the same immediately around the date when the policy change occurred, estimates of the effect of the change at the cutoff point should not be substantially confounded.<sup>5</sup>

To evaluate the relationship of the change with the outcome of interest, the outcome is first estimated for patients *above* the cutoff as the value of the running variable approaches the cutoff from above. The same outcome is then estimated for patients *below* the cutoff as the value of the running variable approaches the cutoff from below. The difference in these estimates, with both estimates extrapolated to hypothetical patients right at the cutoff, quantifies the association between the change and the outcome at the discontinuity.<sup>2,5</sup>

An RDD requires a continuous running variable that separates patients into 2 groups; the running variable is not always time. For example, because of practice guidelines, newborns just below 1500 g of birth weight are much more likely to be admitted to the neonatal intensive care unit (NICU) than those just above. Infants born at 1499 g and 1501 g should be essentially identical. Thus, an RDD can be used to understand the association between NICU care and birth outcomes.<sup>6</sup>

To estimate the difference between the outcomes of infants of close to 1500-g weight attributable to NICU care, separate regression analyses of infants below and above the cutoff would be conducted. One regression estimates the outcomes of hypothetical 1500-g infants, based on infants with weights greater than 1500 g (including some substantially heavier infants).<sup>6</sup> A second regression estimates the outcomes of hypothetical 1500-g infants, but based on infants with weights less than 1500 g (including some substantially lighter infants). In each case, although the regression is based on data from infants with weights different than 1500 g, the regression is used to predict the outcomes for a 1500-g infant. The difference in outcomes for 1500-g infants predicted using the greater-weight and lesser-weight NICU cohorts is the RDD estimate of the association of NICU care with outcomes.

The RDD approach can be extended to more complex situations, for example, cases in which the relationship ascribed to the running variable is not absolute (eg, fuzzy), meaning there may be some patients with values near the cutoff who may receive either treatment, or cases in which patients are followed up over time and some of the same patients contribute results both before and after the time cutoff. The first case might occur when examining the association between blood transfusion and outcomes, with the hemoglobin value used as the running variable, because patients with hemoglobin values near the transfusion threshold may or may not be transfused based on other clinical consideration. The second case might occur when studying a change in a practice guideline for the care of chronic pain on opioid use, because some patients may receive care and be followed up both before and after the change in practice guideline.

**Limitations of the RDD Method**

A number of assumptions must be satisfied for an RDD to be valid. For example, while there should be an abrupt change in the treatment (or policy) received below and above the threshold, an assumption of the RDD approach is that there are no abrupt changes in the relationship between the running variable and the treatment or outcome except at the discontinuity threshold. This helps ensure the validity of the regressions conducted using data from each group. For example, the relationship between NICU admission and birth weight should be smooth (likely decreasing with increasing birth weight), with no abrupt changes except at 1500 g. A similar consideration would apply to clinical outcomes and birth weight.

A second assumption is that no one in the capacity of allocating the treatment or influencing outcomes can manipulate the running variable. For example, a practitioner might alter the recorded birth weight to influence the admission of an infant to the NICU. To test this assumption, patterns of treatments may be examined to determine whether there are any unexplained patterns in the running variable near the threshold that might suggest external manipulation.

The generalizability of estimates of treatment effects based on RDDs may be limited because valid estimates can only be generated for patients close to the running variable's threshold. Thus, the relationship between Medicare coverage and outcomes using an RDD with age as the running variable with age 65 years as the cutoff may be generalizable to 64-year-olds in a study comparing 64- and 66-year-olds, but that estimate would not generalize to 40-year-olds.

### How Was the RDD Method Used?

The study by Sukul and colleagues<sup>1</sup> examined whether the number of available secondary diagnosis coding positions (treatment) was associated with the number of condition categories coded (outcome) for all hospital discharges of Medicare fee-for-service beneficiaries between 2008 and 2015. Because clinicians were allowed to use more secondary diagnosis coding positions starting

in 2011, a simple comparison of the average number of secondary diagnoses before and after 2011 might be biased because of real changes in the patient population over time—such as an increasing burden of disease—that would influence the number of diagnoses made. Thus, the authors exploited a natural experiment using an RDD with date of the hospital discharge as the running variable.

### How Should the Analysis Based on the RDD Method Be Interpreted?

Sukul and colleagues<sup>1</sup> found that the unadjusted mean number of condition categories coded increased from 1.7 in 2008 to 2.7 in 2015. This observed increase likely represents a combination of secular changes and the relationship with the change in Medicare coding rules. Based on the RDD used, adjusting for time, there was an increase of 0.4 diagnostic codes at the discontinuity of the 2011 coding policy change. The authors concluded that the 2011 expansion in the number of secondary diagnosis coding positions was associated with an increase in the estimated severity of illness of hospitalized Medicare beneficiaries. The RDD allowed the authors to separate the association of secular changes from the association of the change in the Medicare rule with the number of diagnoses recorded. This result aids in the quantitative comparison of illness severity in the patient populations treated before and after 2011.

#### ARTICLE INFORMATION

**Author Affiliations:** Center for Health Services Research in Primary Care, Durham Veterans Affairs Medical Center, Durham, North Carolina (Maciejewski); Department of Population Health Sciences, Duke University School of Medicine, Durham, North Carolina (Maciejewski); Division of General Internal Medicine, Department of Medicine, Duke University School of Medicine, Durham, North Carolina (Maciejewski); The Comparative Health Outcomes, Policy, and Economics (CHOICE) Institute, Departments of Pharmacy, Health Services, and Economics, University of Washington, Seattle (Basu).

**Corresponding Author:** Matthew L. Maciejewski, PhD, Center for Health Services Research in Primary Care, Durham Veterans Affairs Medical Center and Duke University Medical Center, 411 W Chapel Hill St, Ste 600, Durham, NC 27705 ([Matthew.Maciejewski@va.gov](mailto:Matthew.Maciejewski@va.gov)).

**Section Editors:** Roger J. Lewis, MD, PhD, Department of Emergency Medicine, Harbor-UCLA Medical Center and David Geffen School of Medicine at UCLA; and Edward H. Livingston, MD, Deputy Editor, *JAMA*.

**Published Online:** July 2, 2020.  
doi:10.1001/jama.2020.3822

**Conflict of Interest Disclosures:** Dr Maciejewski reported ownership of Amgen stock through his spouse's employment. Dr Basu reported receiving personal fees from Salutis Consulting LLC.

**Funding/Support:** This work was supported by the Office of Research and Development, Health Services Research and Development Service, Department of Veterans Affairs (CIN 13-410). Dr Maciejewski was supported by a Research Career Scientist award from the VA (RCS 10-391).

**Role of the Funder/Sponsor:** The Department of Veterans Affairs had no role in the preparation, review, or approval of the manuscript or the decision to submit the manuscript for publication.

**Disclaimer:** The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs, Duke University, or the University of Washington.

#### REFERENCES

1. Sukul D, Hoffman GJ, Nuliyalu U, et al. Association between Medicare policy reforms and

changes in hospitalized Medicare beneficiaries' severity of illness. *JAMA Netw Open*. 2019;2(5):e193290. doi:10.1001/jamanetworkopen.2019.3290

2. Thistlethwaite D, Campbell D. Regression-discontinuity analysis: an alternative to the ex-post facto experiment. *J Educ Psychol*. 1960;51(6):309-317. doi:10.1037/h0044319

3. Haukoos JS, Lewis RJ. The propensity score. *JAMA*. 2015;314(15):1637-1638. doi:10.1001/jama.2015.13480

4. Imbens G, Lemieux T. Regression discontinuity designs: a guide to practice. *J Econom*. 2008;142(2):615-635. doi:10.1016/j.jeconom.2007.05.001

5. Shadish WR, Galindo R, Wong VC, Steiner PM, Cook TD. A randomized experiment comparing random and cutoff-based assignment. *Psychol Methods*. 2011;16(2):179-191. doi:10.1037/a0023345

6. Almond D, Doyle JJ, Kowalski AE, Williams H. Estimating marginal returns to medical care: evidence from at-risk newborns. *Q J Econ*. 2010;125(2):591-634. doi:10.1162/qjec.2010.125.2.591