Regression discontinuity designs in healthcare research

Atheendar S Venkataramani,[1] Jacob Bor,[2] Anupam B Jena[3]

[1]Division of General Internal Medicine, Massachusetts General Hospital, Boston, MA, USA; and Harvard Center for Population and Development Studies, Cambridge, MA, USA

[2]Department of Global Health, Boston University School of Public Health, Boston, MA, USA

[3]Department of Health Care Policy, Harvard Medical School, Boston, MA, USA; Department of Medicine, Massachusetts General Hospital; and National Bureau of Economic Research, Cambridge, MA, USA

Correspondence to: A B Jena, Department of Health Care Policy, Harvard Medical School, Boston, MA 02115, USA jena@hcp.med.harvard.edu

Clinical decisions are often driven by decision rules premised around specific thresholds. Specific laboratory measurements, dates, or policy eligibility criteria create cut-offs at which people become eligible for certain treatments or health services. The regression discontinuity design is a statistical approach that utilizes threshold based decision making to derive compelling causal estimates of different interventions. In this review, we argue that regression discontinuity is underutilized in healthcare research despite the ubiquity of threshold based decision making as well as the design's simplicity and transparency. Moreover, regression discontinuity provides evidence of "real world" therapeutic and policy effects, circumventing a major limitation of randomized controlled trials. We discuss the implementation, strengths, and weaknesses of regression discontinuity and review several examples from clinical medicine, public health, and health policy. We conclude by discussing the wide array of open research questions for which regression discontinuity stands to provide meaningful insights to clinicians and policymakers.

Randomized controlled trials have long considered the ideal means of assessing causal relations in medicine and health policy. However, not all clinical or policy questions can effectively be approached with experimental methods, owing to financial or ethical constraints or concerns about real world generalizability. As a result, investigators have increasingly relied on quasi-experimental study designs.[1 2]

The regression discontinuity design is one such quasi-experimental method. Regression discontinuity takes advantage of clinical or policy decision rules in which people are differentially assigned to a treatment or intervention if they fall above or below an arbitrary cut-off for a continuous variable.[3 4] Such situations abound in healthcare. For example, decisions to initiate antihypertensive treatment, transfuse blood products, or screen for colorectal cancer are often driven by decision rules premised on specific thresholds (some of which are controversial and in need of evaluation).[5 6] Similarly, eligibility for public sector health insurance in the United States is contingent on specific age and/or income based cut-offs.

Though both the ubiquity of clinical and policy decision rules and the high degree of internal validity of regression discontinuity makes the method "tailor-made" for healthcare research, it has been underutilized.[4] In addition to a lack of familiarity with regression discontinuity, underutilization may be driven by an unclear sense of the kinds of questions to which regression discontinuity may be applied as well as the types of cut-offs that can be exploited for causal inference.

In this article, we describe regression discontinuity and its strengths and weaknesses and characterize the diversity of research questions to which this design applies (in the domains of clinical medicine, public health and prevention, and health policy) and the numerous types of thresholds and decision rules that can be employed. Our goal is to inspire investigators to assess more habitually, whether a regression discontinuity approach can strengthen the causal likelihood of an observed empirical association by illustrating the versatility of the method in answering a wide array of research questions in clinical and health policy research.

Treatment assignment and analysis

Assignment to treatment by an arbitrary cut-off in regression discontinuity can either be "sharp" or be "fuzzy."[3 4 7-9] In sharp regression discontinuity, the probability of receiving treatment jumps deterministically from 0 to 1 at the cut-off. All people on one side of the cut-off receive treatment and those on the other side not. The treatment effect in the sharp regression discontinuity is estimated by comparing outcomes for those "just above" the cut-off to those

SUMMARY POINTS

Treatment decisions and interventions are often based on specific measurement, laboratory, timing, or eligibility cut-offs

The regression discontinuity design is a statistical approach that utilizes threshold based decision making to estimate causal estimates of different interventions

Regression discontinuity is relatively simple to implement, transparent, and provides "real world" effects of treatments and policies

Despite the ubiquity of threshold based decision making in healthcare, regression discontinuity is underutilized

"just below" it. Statistically, this involves either fitting models where the outcome is regressed on a binary indicator denoting being above or below the assignment threshold and a linear or polynomial for the underlying assignment variable, or using non-parametric methods to compare outcomes on either side of the threshold within an optimally selected bandwidth.[3 10]

In contrast, in fuzzy regression discontinuity, the probability of treatment increases at the cut-off but is not deterministic—that is, some people on both sides of the cut-off still receive treatment, though the probability of treatment increases at the cut-off. The fuzzy design thus allows for crossover into or out of treatment (relative to the treatment decision implied by assignment based on threshold) and/or non-adherence. This design is thus more germane to the bulk of clinical and policy questions in the sphere of healthcare, where additional considerations (beyond one's position relative to a specific cut-off) often contribute to decision making about treatment, uptake of treatment or programs, and adherence. Similar methods to analyzing sharp regression discontinuity can be applied to fuzzy regression discontinuity. However, because treatment is not deterministic, the comparison captures the effect of assignment to treatment by the threshold rule—not receiving the treatment itself. Instrumental variable methods can then be used to assess the effect of treatment among compliers.[4 9] Several recent reviews have described the statistical and practical issues around conducting regression discontinuity studies in the area of health.[4 7 8]

Figure 1 illustrates treatment assignment in fuzzy regression discontinuity using real clinical data on the initiation of HIV treatment in South Africa.[7] At the time the data were collected, guidelines recommended treatment for all people with initial CD4 counts of less than 200 cells/mm³. As the illustration shows, the probability of receiving treatment was noticeably higher for patients presenting for care with CD4 counts below this threshold. However, several people below the threshold did not receive treatment and several above the threshold did. The presence of such "crossover" highlights the importance of clinical attrition before initiation of treatment, other indications for treatment regardless of CD4 count, and patient and provider preferences in motivating treatment choice in many clinical settings.

Strengths and limitations
The main strength of regression discontinuity is that it provides a transparent method to estimate causal effects of treatments or policies, particularly when randomized controlled trials are not possible. Causal inference comes from the assumption that, aside from differential use of treatment, those on either side—yet close to the cut-off—are otherwise similar. Put differently, whether someone lies immediately above or immediately below a cut-off is considered effectively random, leading to quasi-random treatment
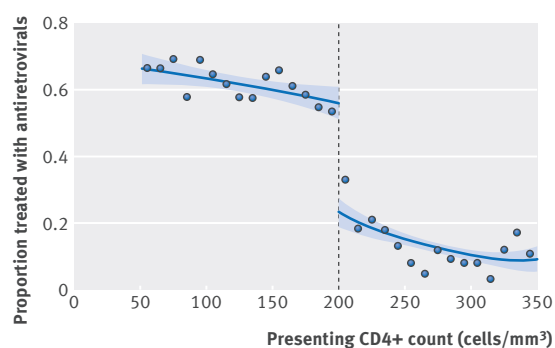
Fig 1 | Fuzzy regression discontinuity. Reanalysis of South African data published in Bor et al.[7] Figure plots the proportion of patients who received early antiretroviral treatment within three months of initial presentation. Circles denote the proportion of patients treated in each CD4 count bin. The solid line is the non-parametric best fit (95% confidence interval is shaded). At the time of data collection, South Africa national guidelines recommended initiating early antiretroviral treatment for all patients with CD4 counts of <200 cells/mm³. Consistent with this, the proportion receiving treatment declined noticeably at this threshold. However, not all patients below the threshold were treated and some patients above the threshold were treated, likely because of other clinical symptoms. This crossover (relative to the assignment threshold) motivates the use of "fuzzy" regression discontinuity. In contrast, in "sharp" regression discontinuity, the probability of receiving early antiretroviral treatment would have been exactly 1 above the threshold and 0 below

assignment for those close to the threshold. The validity of this randomization assumption can be evaluated as in a clinical trial by showing that patients on either side of the threshold have comparable characteristics. Other strengths of regression discontinuity are that non-random attrition and crossover, which create difficulties in randomized trials, can either be avoided or be easily incorporated into the design to support causal inference. Moreover, regression discontinuity studies may better reflect "real world" effects of treatments and policies, as they incorporate important behavioral phenomena such as poor adherence and loss to follow up, which randomized controlled trials seek to minimize.

Limitations of regression discontinuity include concerns around external validity, particularly as the estimates are only interpretable as causal effects for those near the specific cut-off of interest. However, cut-offs, such as those delineated by lipid lowering guidelines, are often of specific policy interest themselves. (With certain assumptions, regression discontinuity estimates can provide information about causal effects further from the threshold, although local randomization is not preserved[11]). Additionally, causal inference in regression discontinuity may be threatened by non-random manipulation around the treatment cut-off. However, in many situations manipulation is either not possible or unlikely to generate bias. For example, random measurement error in laboratory testing may place some people above and some below the treatment threshold in a manner that

cannot easily be manipulated. Moreover, manipulation is visually and statistically testable.[12] Lastly, regression discontinuity studies may require large datasets to generate precise estimates. However, the growing availability of "big data," such as electronic medical information systems, administrative databases, national registries, insurance claims, population health databases, social media, internet searches, and other routinely collected data may loosen this constraint.[13]

Examples of regression discontinuity studies
Table 1 provides examples of regression discontinuity studies in healthcare. These studies, which employ fuzzy designs, are drawn from the clinical, public health, and health policy literatures to illustrate the breadth of research questions investigated using regression discontinuity. Collectively, the studies also show the diversity of usable cut-offs.

Clinical research
Treatment decisions based on therapeutic goals or cut-offs are ubiquitous in clinical medicine. Despite this, few published studies have used the regression discontinuity design to evaluate clinical questions (table 1).[7 14-16] In the example in figure 1, Bor et al show the power of the regression discontinuity approach by examining the role of early antiretroviral treatment on mortality, retention, and immune recovery in South Africa.[7 17] Exploiting a national policy of initiating treatment at CD4 counts less than 200 cells/mm³ to estimate a fuzzy regression discontinuity, they found that people with CD4 counts just below this threshold had a 35% lower hazard of death than those with CD4 counts just above it. Clinical threshold rules based on continuously measured biomarkers (such as CD4 counts) are a particularly strong application of regression discontinuity because assignment to treatment around the threshold is effectively driven by random measurement error.[18 19] Consistent with this, Bor et al showed that patients just above and just below the relevant threshold were statistically identical across baseline characteristics. Similarly, to estimate the impact of statin use on low density lipoprotein cholesterol levels, Geneletti et al. exploited guideline driven differences in statin prescribing for those with just less or just greater than 20% 10 year risk of a cardiovascular event (as determined by a standard risk calculator).[16]

Though the benefits of early antiretroviral treatment and statins are both well established in the medical literature, regression discontinuity studies provide a better sense of what could happen in the real world, which is free of biases to external validity introduced

Table 1 | Examples of regression discontinuity designs in clinical, prevention and public health, and health policy studies

Studies	Exposures	Type of assignment variables	Threshold rules	Outcomes
Clinical:				
Bor et al 2014, 2015[7 16]	HIV antiretroviral therapy	Therapeutic	Patients with CD4 counts <200 cells/mm³ eligible to initiate early antiretroviral treatment in South Africa	Mortality, immune recovery, retention in clinical care
Almond et al 2010[13]	Neonatal intensive care	Therapeutic	Infants with birth weights <1500 g (designated very low birth weight) recommended for intensive care in United States	Infant mortality
Bharadwaj et al 2013[14]	Neonatal intensive care	Therapeutic	Infants with birth weights <1500 g (designated very low birth weight) recommended for intensive care in Chile and Norway	Child cognitive development, academic achievement
Geneletti et al 2015[15]	Statins	Therapeutic	10 year cardiovascular risk >20% as a guideline to initiate statins	Low density lipoprotein cholesterol
Jensen and Wust 2015[19]	Caesarean section	Calendar time	Changes in information and guidelines about efficacy of caesarean section starting 21 October 2000 in Denmark	Apgar score, physician visits, hospital admissions for neonates
Prevention and public health:				
Smith et al 2015[20]	HPV vaccine	Calendar time	Vaccines were available for select age groups after 1 January 1994	Cervical dysplasia and anogenital warts
Callaghan 2013[21]	Minimum drinking age	Age	Adults aged 21 or older can legally purchase and consume alcohol in United States	Mortality
Ludwig and Miller 2007[22]	Head Start program	Program eligibility	Counties ranked <300 based on historic poverty rates were eligible to receive federal Head Start	Child mortality
Chen et al 2013[23]	Air pollution	Geographic	Households north of China's Huai River received subsidies for high emission coal to heat homes	Mortality
Health policy:				
De La Mata et al 2012[24]	Health insurance	Program eligibility	People in households below a specific income threshold were eligible for Medicaid	Healthcare utilization among children
Wherry et al 2015[25]	Health insurance	Calendar time	People born after 1 October 1983 were eligible for more years of Medicaid coverage owing to rule changes	Hospital admissions among adolescents
Sood et al 2014[26]	Health insurance	Geographic	People living in predesignated districts were eligible to receive insurance, whereas those in neighboring districts were not	Mortality
Almond et al 2011[27]	Length of hospital stay	Clock time	Patients admitted after 12 am were allowed longer hospital stays based on insurance rules	Maternal and newborn health

by a trial environment. For instance, the large survival benefit of early antiretroviral treatment in the study by Bor et al may be driven in part by late (or loss to) follow-up among those presenting for care with CD4 counts just above 200 cells/mm³. This is a relevant and important phenomenon that may be missed in randomized controlled trials, where attempts are made to explicitly minimize attrition in both treatment and control groups.[17]

Clinical thresholds can also be used to study the effectiveness of bundles of medical services. For example, to examine the effects of neonatal intensive care on infant mortality, Almond et al used the arbitrary designation of infants weighing less than 1500 g at birth as very low birth weight along with guideline based recommendations on clinical management for these infants.[14] Like CD4 counts in the previous example, treatment assignment was driven by random variation in the measurement of birth weight. Almond et al find that infants just below this cut-off were more likely to receive neonatal intensive care and were 18% less likely to die by age 1 year than those just below the cut-off. Bharadwaj et al used the same cut-off to evaluate the effects of neonatal intensive care on cognitive development and academic achievement later in childhood.[15]

Non-clinical cut-offs can be used to evaluate clinical interventions. Jensen and Wust focused on health outcomes for infants born just before and just after the arrival of information from a major clinical trial in Denmark, which found that caesarean sections conferred better health outcomes than vaginal delivery for term breech infants.[20] Jensen and Wust showed that caesarean rates increased markedly just after the announcement of the trial results, and that infants born just after the announcement had higher Apgar scores and lower rates of hospital admissions in the first year of life than those born just before the announcement (who were more likely to be delivered vaginally).

Prevention and public health research

Creative regression discontinuity approaches have been applied in preventive medicine and public health as well (table 1). For example, in Canada Smith et al explored the effects of the HPV vaccine on cervical dysplasia and anogenital warts.[21] Because of the timing of vaccine introduction as well as age eligibility guidelines, the authors showed an increase in exposure to vaccine across a specific set of birth quarters, which served as the cut-off variable. They found that cervical dysplasia and anogenital warts declined noticeably just after this cut-off as well. Like the clinical studies discussed previously, the study provides a population level insight into the real world effects of HPV vaccination, thereby tackling concerns about reduced efficacy owing to behavioral disinhibition. Age eligibility cut-offs have also been employed in the public health literature. In a prominent example, Callaghan examined the effects of alcohol on mortality among young adults,

examining death rates just above and just below the legal drinking age.[22]

Regression discontinuity approaches have also yielded substantive contributions to the literature on the social determinants of health. Ludwig and Miller showed large reductions in child mortality after the initiation of the Head Start program in the 1970s, which included health and nutritional services in addition to educational programs.[23] They exploited sharp differences in federal support for Head Start provided to counties falling below a poverty threshold, and compared health outcomes in counties on either side of that threshold. (Poverty had been measured using census data before the advent of the program and was thus free of manipulation.)

Geographic boundaries have also been exploited in regression discontinuity studies. Causal inference requires that socioeconomic, demographic, and administrative factors aside from the intervention of interest are similar on either side of the boundary. In a landmark study, Chen et al estimated the impact of air pollution on health using data from China.[24] The authors exploited variation in pollution generated by a national policy, which provided free coal based winter heating for those living north but not south of the Huai River. They found statistically significantly higher levels of ambient pollution just north of the river versus just south, and estimated that as a result of pollution life expectancies were 5.5 years lower north of the river.

Health policy research

Several studies have applied regression discontinuity to assess the impacts of health insurance on health outcomes, a key question in health policy (table 1). De La Mata examined utilization and health outcomes for children in families just above and just below income eligibility thresholds for US Medicaid, the country's largest public insurance program for those on low incomes.[25] To examine the impacts of childhood insurance coverage on adolescent health outcomes, Wherry et al utilized a Medicaid eligibility rule change that increased length of exposure for birth cohorts born just after 1 October 1983.[26] Sood et al used a geographical cut-off, using the fact that insurance was made available in some districts of the Indian state of Karnataka but not in others.[27] They found that mortality rates were lower and utilization rates were higher for those families just over the border where insurance was available compared with those just on the other side. Almond and Doyle examined a specific component of insurance coverage in the United States—namely, rules around minimum length of stay after childbirth.[28] They utilized a specific set of rules around the crude counting of calendar days in the hospital: according to insurance, children born just after midnight are permitted a full day longer in the hospital than those born just before midnight (whose day of birth is counted as a single day in the hospital). They found that longer stays did not change health

Table 2 | Potential clinical questions that can be addressed by regression discontinuity designs and proposed approach

Clinical questions	Threshold rules	Outcomes
What are optimal blood pressure goals for specific subpopulations, such as adults with diabetes?	Compare outcomes for people above and below systolic and diastolic blood pressure for different guideline based thresholds (which have changed over time)	Cardiovascular events, mortality
What is the optimal fluid resuscitation strategy in patients with sepsis?	Compare outcomes across different central venous pressure thresholds (which may vary across time, institutions)	Length of stay on intensive care unit stay, mortality
Are steroids appropriate in patients with alcoholic hepatitis and infections?	Compare outcomes for patients with both alcoholic hepatitis and infections just above and just below the Maddrey discriminant cut-off of 32 (above which steroids are indicated)	Mortality
Is liberal or conservative transfusion more appropriate in acute coronary syndromes?	Compare outcomes across different transfusion thresholds (such as hemoglobin <8 mg/L), which may vary across institutions and over time	Length of stay, mortality
What are the real world effects of cardiac resynchronization therapy?	Compare outcomes just above and just below a QRS complex duration of 120 ms (indication for cardiac resynchronization therapy placement)	Hospital admissions for heart failure, quality of life

outcomes for the mother or the child, but resulted in substantially greater costs. Collectively, these studies, as well as others using regression discontinuity to illustrate wide ranging impacts of insurance,[29][30] exemplify how different types of regression discontinuity can be used to evaluate similar questions.

Future opportunities

The approaches discussed in this article highlight the diversity of questions that can be addressed using regression discontinuity designs. They also illustrate how the vast range of decision thresholds in healthcare can be utilized to bolster the causal likelihood of observed empirical associations. Because of the potential breadth of options available, investigators focusing on clinical questions for which experiments are not ethically or financially feasible or for which the external validity is uncertain should assess whether regression discontinuity approaches are possible and credible before turning to other observational approaches. They should do so keeping in mind that factors as diverse as decision rules based on clinical variables; temporal changes in guidelines or reimbursement policies; income, age, and other eligibility thresholds; or geographic location may all be valid cut-offs for analysis.

Several future opportunities remain for the application of regression discontinuity approaches. Table 2 highlights several clinical areas where there are active controversies around a specific clinical guideline or little evidence of the real world effects of a particular treatment.[31] For example, blood pressure targets for some patient populations (eg, adults with diabetes) continue to be based on expert opinion. Regression discontinuity studies using specific blood pressure cut-offs can be used to strengthen evidence based guidelines in these areas.[6] Another example is cardiac resynchronization treatment, which has been shown to be effective in clinical trials of patients with left ventricular systolic dysfunction (ejection fraction <35%) and QRS prolongation (>120 ms) on resting electrocardiograms. However, real world evidence in specific populations may be bolstered by comparing implantation rates with cardiac resynchronization treatment, hospital admissions for heart failure, and all cause morbidity and mortality for patients with duration of the QRS complex just above versus just below a cut-off of 120 ms.

Conclusion

The preponderance of clinical decision thresholds and changes in policy guidelines provide a rich basis for the application of regression discontinuity. The regression discontinuity design has a number of distinct advantages, even over randomized trials. It can be deployed where clinical trials are not feasible and, unlike randomized controlled trials, allows for an assessment of real world impacts. Though regression discontinuity requires large sample sizes, this design would be sufficiently powered for several existing clinical and administrative databases. The time is opportune for growth of regression discontinuity based studies in healthcare.

1   Rockers PC, Røttingen J-A, Shemilt I, Tugwell P, Bärnighausen T. Inclusion of quasi-experimental studies in systematic reviews of health systems research. *Health Policy* 2015;119:511-21. doi:10.1016/j.healthpol.2014.10.006.

2   Bor J, Geldsetzer P, Venkataramani A, Bärnighausen T. Quasi-experiments to establish causal effects of HIV care and treatment and to improve the cascade of care. *Curr Opin HIV AIDS* 2015;10:495-501. doi:10.1097/COH.0000000000000191.

3   Lee DS, Lemieux T. Regression discontinuity designs in economics. *J Econ Lit* 2010;48:281-355doi:10.1257/jel.48.2.281.

4   Moscoe E, Bor J, Bärnighausen T. Regression discontinuity designs are underutilized in medicine, epidemiology, and public health: a review of current and best practice. *J Clin Epidemiol* 2015;68:122-33. doi:10.1016/j.jclinepi.2014.06.021.

5   Martin SS, Blumenthal RS. Concepts and controversies: the 2013 American College of Cardiology/American Heart Association risk assessment and cholesterol treatment guidelines. *Ann Intern Med* 2014;160:356-8. doi:10.7326/M13-2805.

6   Peterson ED, Gaziano JM, Greenland P. Recommendations for treating hypertension: what are the right goals and purposes? *JAMA* 2014;311:474-6. doi:10.1001/jama.2013.284430.

7   Bor J, Moscoe E, Mutevedzi P, Newell M-L, Bärnighausen T. Regression discontinuity designs in epidemiology: causal inference without randomized trials. *Epidemiology* 2014;25:729-37. doi:10.1097/EDE.0000000000000138.

8   O'Keeffe AG, Geneletti S, Baio G, Sharples LD, Nazareth I, Petersen I. Regression discontinuity designs: an approach to the evaluation of treatment efficacy in primary care using observational data. *BMJ* 2014;349:g5293. doi:10.1136/bmj.g5293.

9   Angrist J, Pishke J-S. *Mastering Metrics.*Princeton University Press, 2015.

10 Gelman A, Imbens G. Why high-order polynomials should not be used in regression discontinuity designs. NBER Working Paper No 20405 2014.

11 Angrist J, Rokkanen M. Wanna get away? Regression discontinuity estimation of exam school effects away from the cutoff. *J Am Stat Assoc* 2015. doi:10.1080/01621459.2015.1012259.

12 McCrary J. Manipulation of the running variable in the regression discontinuity design: A density test. *J Econom* 2008;142:698-714doi:10.1016/j.jeconom.2007.05.005.

13 Schneeweiss S. Learning from big health care data. *N Engl J Med* 2014;370:2161-3. doi:10.1056/NEJMp1401111.

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

15 Bharadwaj P, Loken KV, Neilson C. Early life health interventions and academic achievement. *Am Econ Rev* 2013;103:1862-91doi:10.1257/aer.103.5.1862.

16 Geneletti S, O'Keeffe AG, Sharples LD, Richardson S, Baio G. Bayesian regression discontinuity designs: incorporating clinical knowledge in the causal analysis of primary care data. *Stat Med* 2015;34:2334-52. doi:10.1002/sim.6486.

17 Bor J, Moscoe E, Barninghausen T. When to start HIV treatment: Evidence from a regression discontinuity study in South Africa. In: Population Association of America Annual Meeting. San Diego, CA USA; 2015.

18 Bor J, Moscoe E, Barninghausen T. The "natural experiment" in regression discontinuity design: Randomization without controlled trials. *BMJ* 2014;349:g5293.25199521.

19 Bor J, Moscoe E, Bärnighausen T. Three approaches to causal inference in regression discontinuity designs. *Epidemiology* 2015;26:e28-30, discussion e30. doi:10.1097/EDE.0000000000000256.

20 Jensen VM, Wüst M. Can Caesarean section improve child and maternal health? The case of breech babies. *J Health Econ* 2015;39:289-302. doi:10.1016/j.jhealeco.2014.07.004.

21 Smith LM, Kaufman JS, Strumpf EC, Lévesque LE. Effect of human papillomavirus (HPV) vaccination on clinical indicators of sexual behaviour among adolescent girls: the Ontario Grade 8 HPV Vaccine Cohort Study. *CMAJ* 2015;187:E74-81. doi:10.1503/cmaj.140900.

22 Callaghan RC, Sanches M, Gatley JM, Cunningham JK. Effects of the minimum legal drinking age on alcohol-related health service use in hospital settings in Ontario: a regression-discontinuity approach. *Am J Public Health* 2013;103:2284-91. doi:10.2105/AJPH.2013.301320.

23 Ludwig J, Miller DL. Does Head Start improve children's life chances? Evidence from a regression discontinuity design. *Q J Econ* 2007;122:159-208doi:10.1162/qjec.122.1.159.

24 Chen Y, Ebenstein A, Greenstone M, Li H. Evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River policy. *Proc Natl Acad Sci U S A* 2013;110:12936-41. doi:10.1073/pnas.1300018110.

25 De La Mata D. The effect of Medicaid eligibility on coverage, utilization, and children's health. *Health Econ* 2012;21:1061-79. doi:10.1002/hec.2857.

26 Wherry LR, Miller S, Kaestner R, Meyer BD. Childhood medicaid coverage and later life health care utilization. NBER Working Paper No 20929 2015.

27 Sood N, Bendavid E, Mukherji A, Wagner Z, Nagpal S, Mullen P. Government health insurance for people below poverty line in India: quasi-experimental evaluation of insurance and health outcomes. *BMJ* 2014;349:g5114. doi:10.1136/bmj.g5114.

28 Almond D, Doyle JJ. After midnight: a regression discontinuity design in length of postpartum hospital stays. *Am Econ J Econ Policy* 2011;3:1-34doi:10.1257/pol.3.3.1.

29 Card D, Dobkin C, Maestas N. Does Medicare save lives?*Q J Econ* 2009;124:597-636. doi:10.1162/qjec.2009.124.2.597.

30 Chay KY, Daeho K, Shailender S. *Medicare, hospital utilization, and mortality: Evidence from the program's origins. Mimeo.*Brown University, 2011.

31 Walkey AJ, Bor J. Risk-based heterogeneity of treatment effect in trials and implications for surveillance of clinical effectiveness using regression discontinuity designs. *Am J Respir Crit Care Med* 2015;192:1399. doi:10.1164/rccm.201508-1533LE.