Journal of Clinical Epidemiology 68 (2015) 132−143

Journal of Clinical Epidemiology

Regression discontinuity designs are underutilized in medicine, epidemiology, and public health: a review of current and best practice

Ellen Moscoe[a,*], Jacob Bor[b,c], Till Bärnighausen[a,c]

[a]*Department of Global Health and Population, Harvard School of Public Health, 665 Huntington Avenue Building 1, room 1104, Boston, MA, USA*
[b]*Department of Global Health, Boston University School of Public Health, 801 Massachusetts Ave, 3rd Floor, Boston, MA, USA*
[c]*Africa Centre for Health and Population Studies, University of KwaZulu-Natal, P.O. Box 198, Mtubatuba, 3935, South Africa*

Accepted 1 June 2014

Abstract

Objectives: Regression discontinuity (RD) designs allow for rigorous causal inference when patients receive a treatment based on scoring above or below a cutoff point on a continuously measured variable. We provide an introduction to the theory of RD and a systematic review and assessment of the RD literature in medicine, epidemiology, and public health.

Study Design and Setting: We review the necessary conditions for valid RD results, provide a practical guide to RD implementation, compare RD to other methodologies, and conduct a systematic review of the RD literature in PubMed.

Results: We describe five key elements of analysis all RD studies should report, including tests of validity conditions and robustness checks. Thirty two empirical RD studies in PubMed met our selection criteria. Most of the 32 RD articles analyzed the effectiveness of social policies or mental health interventions, with only two evaluating clinical interventions to improve physical health. Seven out of the 32 studies reported on all the five key elements.

Conclusion: Increased use of RD provides an exciting opportunity for obtaining unbiased causal effect estimates when experiments are not feasible or when we want to evaluate programs under ''real-life'' conditions. Although treatment eligibility in medicine, epidemiology, and public health is commonly determined by threshold rules, use of RD in these fields has been very limited until now.

Keywords: Regression discontinuity; Causal inference; Quasi-experimental methods, Systematic review; Natural experiments; Observational studies; Confounding

1. Introduction

Regression discontinuity (RD) designs are a rigorous quasi-experimental method for estimating causal effects of treatments on outcomes. Whenever a decision rule assigns treatment, such as antihypertensive or antiretroviral therapies, to patients who score higher (or lower) than a particular cutoff value on a continuously measured variable, such as blood pressure or CD4 count, RD can be used to estimate the causal effect of the treatment on health and other outcomes. Like randomization, RD can solve problems of confounding by unobserved factors, generating unbiased estimates of the causal effects of a treatment. RD is a particularly useful research design for medicine, epidemiology, and public health because of the ubiquity of treatments assigned based on a cutoff rule [1]. Physicians prescribe statins to those with high cholesterol above a certain cutoff value, use a size cutoff as a guideline for mole excision, determine treatment for hypertension based on blood pressure cutoffs, and recommend surgery for scoliosis when spinal curvature exceeds some threshold of severity. In addition, RD has desirable practical characteristics. When a treatment has already become the standard of care, it may not be possible to conduct a randomized controlled trial (RCT), but RD can provide strong causal evidence on treatment effectiveness in cases where there is little or no experimental evidence or where the existing evidence is of questionable internal or external validity [2]. Additionally, RD may be less costly than experimental methods because it can be implemented using data that is commonly collected in patient files and administrative data. Cohort studies that collect information on a continuous

Conflict of interest: None.

* Corresponding author. Department of Global Health and Population, Harvard School of Public Health, 665 Huntington Avenue, Bldg.1, Room 1104, Boston, MA 02115, USA. Tel.: 503-780-7724; fax: 617-432-6733.
E-mail address: eem571@mail.harvard.edu (E. Moscoe).

http://dx.doi.org/10.1016/j.jclinepi.2014.06.021

What is new?

- Regression discontinuity (RD) is a quasi-experimental study design that is well suited for medical, epidemiologic, and public health research. RD identifies causal effects by exploiting a treatment assignment practice that is common in these fields: the assignment of treatment based on whether a patient scores above or below a cutoff point on a continuously measured variable, such as blood pressure, cholesterol, or CD4 count.
- RD has several advantages over randomized controlled trials (RCTs). In particular, it can be used to evaluate interventions that have become standard practice without preceding RCTs or when there is doubt that trial-based evidence can be generalized to routine health care in particular contexts. In this article, we present the underlying theory and compare RD to randomized trials and traditional cohort studies.
- To date, RD has been underutilized in medicine, epidemiology, and public health. We identified 32 studies in PubMed, 13 of which are published in economics or health economics journals. Very few articles in our systematic review use RD to study the effect of clinical interventions on health. The studies have been of overall good quality, but further improvements are possible. Guidelines for implementing and presenting RD studies can help encourage utilization of this study design in medicine, epidemiology, and public health. In this paper, we provide guidance: in addition to showing the relationship between the assignment variable and the outcome, high quality RD studies should include a discussion of the treatment assignment rule, a histogram of the assignment variable, a discussion of how a particular study meets the conditions necessary for valid RD estimation, covariate balance tests, and robustness checks of the RD estimation approach.
- There is significant potential for RD to generate strong causal evidence using existing clinical, administrative, and programmatic cohort data. Data collection guidelines for clinical and epidemiologic cohort studies and administrative data in public health should be updated to make RD analysis feasible whenever possible, for example, by retaining data on patients not yet eligible for treatment.

diagnostic criterion, the treatment patients receive, and the outcomes in both treated and nontreated groups will have the data necessary to implement RD analyses. A further advantage of RD is that it can be easily graphically presented, allowing results to be shared widely with policy makers and implementing organizations.

RD was first used in the field of educational psychology by Thistlewaite and Campbell [3] in 1960. The design was introduced to statistics by Rubin [4]. Berk and Rauma [5] extended the model to dichotomous variables using logistic models. In a recent paper Bor et al. [1] extended RD to the case of survival analysis. RD has become widely used in economics since the 1990s [6−8]. Studies of the impact of incumbency on electoral outcomes [9], the effects of military conscription on earnings [10], and the relationship between class size and student performance [11] showed that RD could generate important results in a broad range of settings. A number of important advances in the theory of RD have come out of the recent economics literature [12,14]. Economists have also used RD designs to address questions that are of interest to epidemiologists and public health researchers. For example, Almond et al. [15] estimated the causal effect of intensified medical treatment given to very low-birth-weight babies (weighing less than 1,500 g) on 1-year mortality. Using the cutoff age of 21 for legal alcohol purchases, Carpenter and Dobkin [16] evaluated the effect of alcohol consumption on mortality.

The goals of this article are (1) to provide an introduction to the theory of RD and a guide for implementation and "best practice" in the context of medicine, epidemiology, and public health and (2) to systematically review and evaluate the use of RD in these fields of research, that is, the "current practice." We further discuss potential applications and limitations of RD in epidemiology and public health.

2. Fundamentals of RD designs

RD can be used when clinical practice or public health programs use a cutoff point on a continuous variable as the decision rule to assign treatment or program eligibility. Treatment assignment following such a rule can be either deterministic (every patient on the one side of the cutoff value receives the treatment and every patient on the other side does not) or probabilistic (the probability of receiving the treatment is higher on the one side of the cutoff value than on the other side). The first case is called "sharp" RD and the second "fuzzy" RD. We present both cases in the following paragraph.

Like a RCT, RD is more than a method of data analysis: it is a description of the data-generating process when a continuously measured variable has a cutoff point that determines treatment status. Under certain conditions, it is possible to infer that a difference in outcomes is the causal result of the assignment variable's cutoff point. Researchers have invoked different assumptions to identify causal effects in RD designs [17]. Early discussions of RD emphasized global average treatment effects and required very strong functional form assumptions [4]. Most recent RD

literature—and this article—focuses on local treatment effects "at the threshold", for which the key assumption is continuity in potential outcomes, i.e. that there are no unobserved confounders at the threshold. As we approach the cutoff value from above and below, the patients in the two groups become more and more alike, on both observable and unobservable characteristics; in a small area around the threshold, the only difference is in treatment assignment. In some settings, it is remarkable how easily the continuity assumption is met. If there is random noise in measurements of the assignment variable and it cannot be precisely manipulated, then continuity in potential outcomes is guaranteed (in expectation) [13]. Because of the random noise in an individual's value of the assignment variable, whether a person near the cutoff falls above or below the cutoff is essentially random. We can thus interpret the difference in outcomes between the people just above the cutoff and those just below it as a true causal effect of the treatment. Although not all RD designs have this "local randomization" interpretation, it is often justified in clinical settings, where blood glucose levels, CD4 counts, and blood pressure are measured with substantial error [17].

2.1. Assumptions and validity conditions

The conditions and assumptions necessary for causal inference in RD are relatively weak compared with other quasi-experimental methods. Furthermore, when the focus is on local treatment effects at the threshold, the critical assumptions can be supported using the available data, something that is not possible with most other quasi-experimental approaches. The three conditions for a valid RD are as follows:

2.1.1. The decision rule and cutoff value are known

Researchers must know the cutoff value of the variable used to assign treatment, known as the assignment variable. Throughout this section, we will use the letter Z to designate the assignment variable. Researchers must also know whether treatment is assigned when Z is below or above the cutoff. It is also helpful to know whether other factors (eg, clinical judgment in addition to a laboratory measure representing Z) contribute to the decision to treat. In such cases, the "fuzzy" variant of RD must be used, in which intent-to-treat effects are estimated and scaled by the level of compliance with the threshold rule to obtain complier average causal effects (CACE) for those receiving the treatment. In both the "sharp" and the "fuzzy" variant of RD, we estimate a causal effect local to the population close to the cutoff value.

2.1.2. The assignment variable is continuous near the cutoff value

The assignment variable Z may be any continuous variable that is measured before treatment, is not affected by the treatment, and determines treatment at some cutoff point. Contrary to other quasi-experimental methods that attempt to control for unobserved confounders (eg, difference-in-difference analysis), there is no area of overlap in sharp RD where observations with differing treatment status have the same values of Z. Hahn et al. [12] show that without this area of overlap, continuity in Z near the cutoff is sufficient to obtain unbiased estimates of the TE. Visual inspection of the data can confirm that Z is continuous at the cutoff.

2.1.3. Potential outcomes are continuous at the threshold

For causal effects to be identified, patients must be similar just above and below the threshold. This is necessary to ensure that their potential outcomes (outcomes if all were or were not treated) would be similar immediately on either side of the threshold. More formally, the conditional distributions of potential outcomes with respect to Z are continuous at the threshold.

The continuity assumption would be violated if the specific cutoff point was determined because of an underlying discontinuity in the relationship between Z and the outcome. For example, if the cutoff for treatment assignment to antihypertensive medication was determined because a physiological phenomenon that is correlated with the outcome of interest, for example, cardiovascular mortality, occurred precisely at the cutoff, then reverse causality could confound the analysis. Similarly, there must be no unobserved confounders that are discontinuously associated with the outcome at the cut-off. For example, when estimating the effect of differences in cigarette taxes on smoking behavior using distance from a state line as the assignment variable, the analysis may be confounded by other aspects of the state policy environment. Plots of other covariates around the cutoff and knowledge about how the cutoff rule is established can help to confirm that the discontinuity in Y is caused only by the cutoff and not by another factor.

The most compelling RD design occurs when there is random noise in measurements of the assignment variable (such as CD4 counts, but not distance to an administrative boundary). This is common in clinical applications. In this case, the assumption of continuity in potential outcomes is trivially satisfied (in expectation) under a much more straightforward and testable assumption: that patients have only imperfect control over the value of assignment variable and cannot precisely manipulate its value. In the simplest case, patients have no control over Z (eg, birth date) and therefore cannot manipulate their treatment status. However, RD can still be applied in scenarios where patients have some degree of control over Z, as long as this control is incomplete, as in a case where adherence to medication is correlated with, but does not perfectly determine, Z. Clinical and public health practice in particular settings will affect the degree to which patients and providers influence their measured value of Z. The presence

of manipulation can be identified in the data by assessing the presence of bunching in the density of Z at the cutoff.

2.2. Inference and estimation in sharp RD

When treatment is a deterministic function of the assignment variable, we will have the simplest form of RD design. Let T be the treatment, Y be the outcome, and Z be the forcing variable. The fundamental problem of causal inference is that we do not observe Y_i when $T_i = 0$ and $T_i = 1$, denoted $Y_i(1)$, $Y_i(0)$. RCTs solve this by estimating an average casual effect (ACE) for the population. In the sharp RD design, we observe $T = 1$ on the one side of the cutoff and $T = 0$ on the other side. Near the cutoff, we examine the TE as Z approaches the cutoff C from either side:

$$\mathrm{ACE_{SRD}} = \lim_{z\uparrow c} E[Y_i(1)|Z_i = z] - \lim_{z\downarrow c} E[Y_i(0)|Z_i = z].$$

This is equal to

$$\mathrm{ACE_{SRD}} = E[Y_i(1) - Y_i(0)|Z_i = c].$$

Under the assumptions aforementioned, the outcome Y would be continuous in the absence of treatment. Instead, c introduces randomness by assigning people just above (or below) the cutoff to the treatment. The direction and magnitude of the difference in Y above and below c gives the ACE. In RD analysis, the $\mathrm{ACE_{SRD}}$ is a local effect, meaning that it is valid near the cutoff, because the identification is due to the local randomness resulting from the cutoff point. To estimate the $\mathrm{ACE_{SRD}}$ in practice, regression models allow for different slopes on either side of the cutoff:

$$Y_i = \beta_0 + \beta_1 T_i + \beta_2(Z_i - c) + \beta_3 T_i(Z_i - c) + \varepsilon_i$$

where $T = 1$ when the individual is treated and β_1 is the $\mathrm{ACE_{SRD}}$ (the ''jump'' at the cutoff point). β_2 is the slope in absence of treatment, and the interaction term's coefficient, β_3, allows for a different relationship between Z and Y on the other side of c. Higher order terms and other covariates can be included in the model and the model should be estimate for a wide range of bandwidths (windows of data) around the threshold. In the next section, we outline the steps for implementing RD.

2.3. Inference and estimation in fuzzy RD

The fuzzy RD design is similar to a randomized experiment with imperfect treatment compliance. The treatment assignment indicator, $Z > c$, determines treatment status but only probabilistically. Some patients would always take up the treatment, regardless of treatment assignment; other patients would never take up the treatment even if assigned. Recall the TE from a sharp RD, $\mathrm{ACE_{SRD}}$. In the fuzzy RD design, $\mathrm{ACE_{SRD}}$ is equivalent to the intent-to-treat (ITT) effect, that is, the causal effect of treatment assignment on the outcome, which we will call $\mathrm{ITT_{FRD}}$. As in a randomized experiment, the causal effect of treatment status itself can be estimated for the subpopulation of patients whose treatment decision was determined by their treatment assignment (so-called compliers). To obtain this complier average causal effect ($\mathrm{CACE_{FRD}}$), we scale this by the probability of treatment on each side of the cutoff:

$$\mathrm{CACE_{FRD}} = \frac{\lim_{z\uparrow c} E[Y_i(1)|Z_i = z] - \lim_{z\downarrow c} E[Y_i(0)|Z_i = z]}{\lim_{z\uparrow c} P[T_i = 1|Z_i = z] - \lim_{z\downarrow c} P[T_i = 1|Z_i = z]}$$

Note that when the denominator equals 1, treatment is completely determined by the cutoff point and the fuzzy RD estimator equals the sharp RD estimator. When there is no discontinuity in probability of treatment, that is, the denominator equals zero, $\mathrm{CACE_{FRD}}$ is undefined.

In practice, we can use the same regression model as with the sharp analysis, with one modification. We use treatment assignment (Z_i above c) as an instrumental variable (IV) to predict actual treatment status T_i [12]. This instrument will not have the usual weaknesses of IV analysis, namely reliance on the untestable assumption that the instrument is as good as randomly assigned. As with any IV, careful attention must be paid to the assumption of excludability (ie, the instrument only affects the outcome through its effect on treatment status).

2.4. RD in practice: a step-by-step guide to implementation

To structure the description of the practical process of implementing RD, we have divided the RD analysis into four steps.

1. Determine feasibility of RD design: In considering a RD strategy, researchers should determine that it is feasible given the available data and indicate which type (''sharp'' or ''fuzzy'') is appropriate. The fuzzy implementation of RD allows estimation even when the treatment assignment is not deterministically governed by the assignment variable. We recommend that researchers verify and demonstrate that the data contain the following:
    i. Continuous eligibility measure: The assignment variable is measured and reported continuously.
    ii. Universal outcome assessment: The outcomes must be observed for all patients, independent of whether they were assigned the treatment or not, similar to other prospective studies.
    iii. Treatment assignment rule: Precise information on how treatment is assigned to patients (either above or below and either probabilistically or deterministically) is needed to

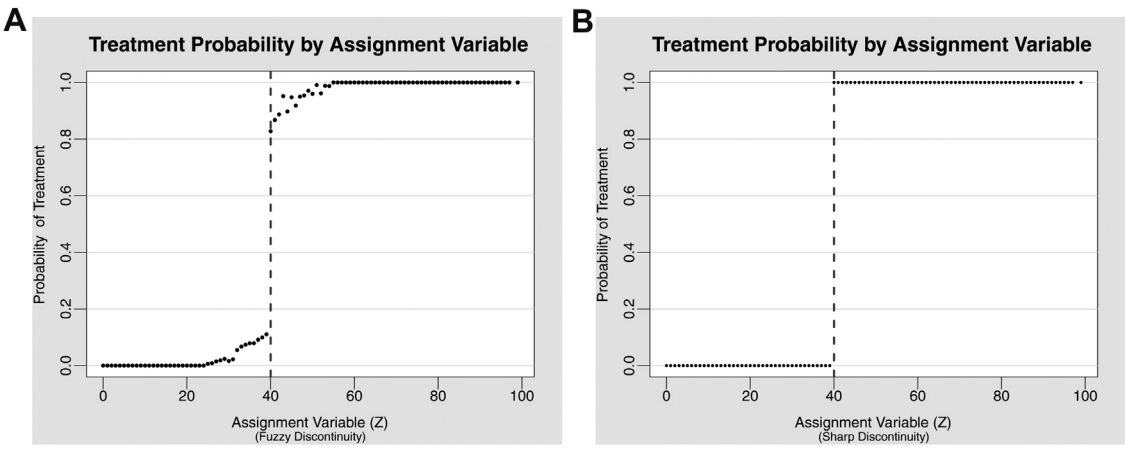

Fig. 1. Probability of receiving treatment for fuzzy (A) and sharp (B) designs.

determine whether the design is "sharp" or "fuzzy." Presenting a plot of the assignment variable against treatment status, which displays the distribution of treatment assignment, demonstrates this for the readers (see Fig. 1).

2. Consider covariate balance and possible manipulation of treatment status: Using a histogram of the assignment variable, *Z*, researchers can confirm that there is no "bunching," (Fig. 2) which would indicate manipulation of treatment status [18]. To confirm that the groups on either side of the cutoff are comparable and that other observed factors are not discontinuous at the cutoff, it is useful to report covariate balance tests (or scatter plots) of *Z* on several pretreatment covariates [6]. RD designs rely on the local randomness resulting from cutoff-based treatment allocation. To support this, it is important to demonstrate that individuals on each side of the cutoff are similar with respect to pretreatment covariates (age, income, and educational attainment, for example) by testing for balance on these variables around the cutoff point. Smooth distributions of covariates also help to establish that there is no manipulation of treatment status similar to an RCT.
3. Visually check for a treatment effect: A plot of *Y* on the assignment variable visually confirms the discontinuity for the reader (see Fig. 3). Visual inspection that reveals a visible jump at the cutoff indicates a nonzero treatment effect, whereas continuity at the cutoff indicates a null effect. When outcomes are discrete, proportions can be estimated in bins. The shape of the scatter plot can indicate whether interaction terms and higher order terms should be included in the regression model.
4. Fit the regression models to estimate the TE: It is possible to either estimate a local linear regression using only the data near the cutoff or estimate a regression model using the full data set, as in Fig. 4. This decision can be made based on data availability and other factors, but in practice, it may be best to estimate multiple specifications of both methods. When using local linear regression, use specifications with data spanning different distances from the cutoff (bandwidths) to establish robustness of the effect size estimate [14]. When using the full data, flexible models with higher order polynomial terms may be compared with linear models. It is also possible to include interaction terms in the regression equation to allow for differences in the relationship between *Z* and *Y* above and below the cutoff, which may occur if the treatment changes the relationship.

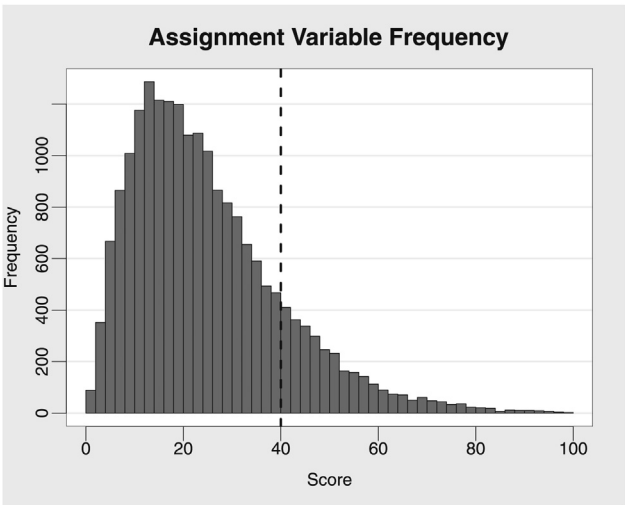

Fig. 2. Histogram of assignment variable.

As with any regression model, misspecification of the functional form can cause incorrect estimation. However, this is a practical rather than a theoretical problem for RD. With enough data, a simple difference in means near the cutoff would give the TE without requiring any assumptions of functional form. The use of local regressions with small

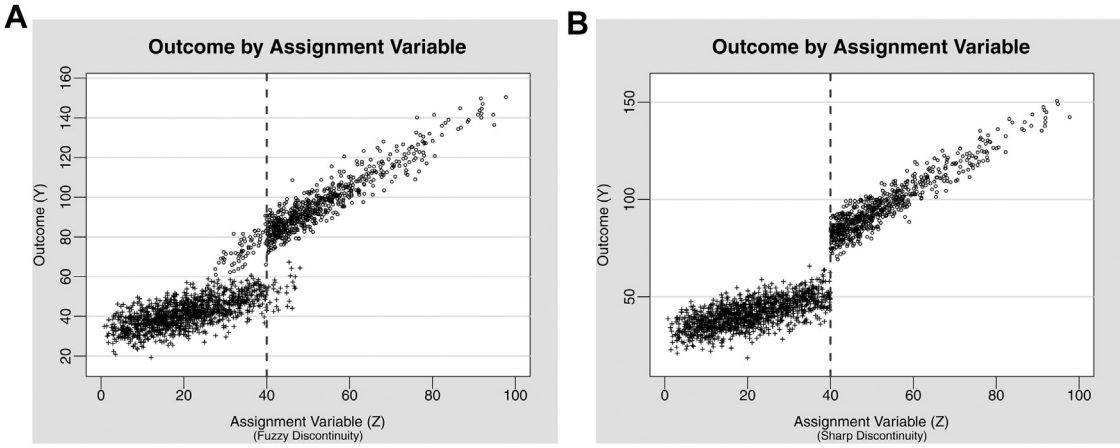

Fig. 3. Plot of outcomes for fuzzy (A) and sharp (B) designs.

bandwidths mitigates the potential problem of incorrect functional form assumptions and acts as a robustness check to global models, and better data collection procedures (oversampling near the cutoff) can improve power at narrower bandwidths. The local nature of RD estimates means that extrapolation to observations far from the cutoff may not be valid, but sufficient robustness checks can confirm that the local estimates are not artifacts of model specification.

3. Systematic review of the literature in PubMed

To establish the frequency and quality of RD use in the epidemiology and public health literature, we performed a systematic review of the PubMed literature. Our aim was to identify empirical publications that use RD to estimate the causal effect of an exposure on a health outcome. To identify other terminology that may refer to RD-type designs, we examined the MeSH terms for study design and method and checked the MeSH terms associated with RD studies. No other label or terminology was found. We therefore chose the broadest search algorithm that would return RD studies: ''regression discontinuity'' OR (''regression'' AND ''discontinuity''). This search term returned 193 unique records in PubMed as of December 8, 2014, catalogued in PubMed between May 1981 and December 8, 2014. We selected records for final review that met the following criteria. First, we rejected articles if they were systematic reviews, case studies, or method articles without an empirical application, or any other empirical strategy. A total of 134 articles were excluded based on these criteria. Second, we rejected articles that did not examine effects on health outcomes. We defined health outcomes broadly to include mortality, morbidity, and diagnostic markers, for both mental and physical health. Based on this criterion, we excluded a further 27 articles, yielding a final database of 32 articles indexed in PubMed that used RD as an empirical approach to assess the effect of an exposure on a health outcome.

Next, we reviewed the articles to determine whether they included descriptions of the elements that are important for successful implementation of RD design, based on the validity conditions and the implementation guide described in section 2.4. We report whether these elements are present or absent as an assessment of study quality. In particular, we examined whether the studies included the following

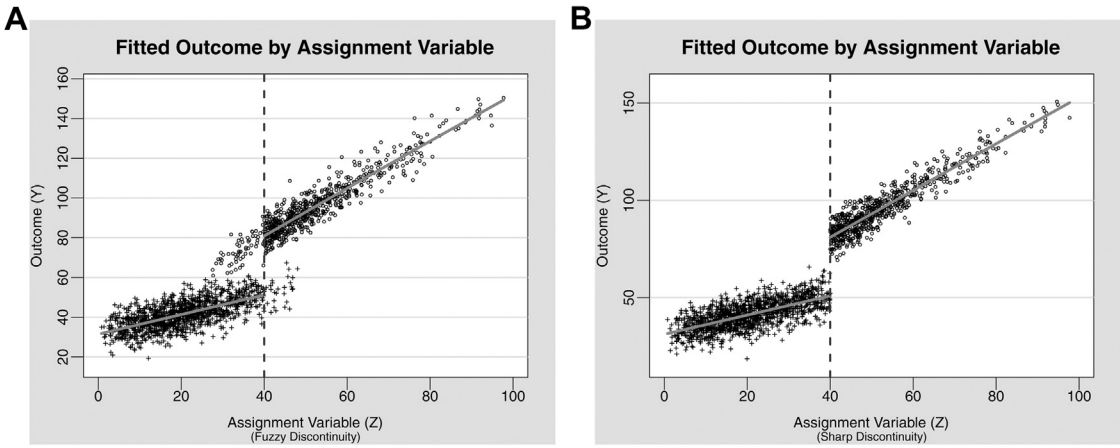

Fig. 4. Plot of outcomes with fitted regression lines for fuzzy (A) and sharp (B) designs.

five elements of good RD practice. The first four elements demonstrate that the data and context are appropriate for a valid RD study (validity conditions), and the last one addresses the practical problem of sensitivity to the functional form specification:

1. A discussion of the RD validity conditions in the context of the particular study.
2. A clear presentation or discussion of the assignment rule.
3. Covariate balance tests for treated and nontreated groups showing that there is no discontinuity in pretreatment variables.
4. A histogram of the assignment variable that shows no bunching of the data around the cutoff to demonstrate no manipulation of treatment status.
5. Multiple RD estimation specifications to check for robustness, including alternative functional forms, nonparametric regression, and local linear regression with varying bandwidths.

Table 1 reports the 32 articles (of 193) that met the inclusion criteria [1,15,16,23–51]. Of the 32 studies, 13 were published in economics or health economics journals, five were published in psychology or psychiatry journals, and the remaining 14 were published in health journals (i.e., in epidemiology, public health, or medical journals). Nine articles were published prior to 2012, six in 2012, six in 2013, and eleven in 2014. The publications in economics and health economics journals addressed questions of relevance for epidemiology, such as the effect of education on body mass index [25] and mortality [23], the impact of teacher training on students' fertility and high-risk sex behaviors [26], and the effects of alcohol consumption on mortality [29–31]. Table 2 lists how well the published studies adhered to the elements of good RD practice. To summarize the findings, the final column of the table displays the total number of included elements for each article, assigning one point for each element (of five possible). Of the 32 studies, seven included all five elements. All studies

Table 1. PubMed articles with health outcomes using regression discontinuity designs

Authors	Year	Journal	Study topic
Albouy and Lequien [23]	2009	Journal Health Economics	Effect of education on mortality
Almond et al. [15]	2010	Quarterly Journal of Economics	Returns to treatment of low-birth-weight newborns
Andalón [24]	2011	Health Economics	Effect of Oportunidades on obesity
Anderson et al. [25]	2011	Journal of Health Economics	Effect of schooling on children's BMI
Arcand and Wouabe [26]	2010	Health Economics	Effect of teacher training on HIV prevention
Banks and Mazzonna [27]	2012	Economics Journal	Effect of education on old-age cognitive ability
Behrman [28]	2014	Social Science and Medicine	Effect of primary schooling on HIV status
Bor et al. [1]	2014	Epidemiology	Effect of early vs. deferred HIV treatment on mortality
Callaghan et al. [29]	2014	Drug and Alcohol Dependence	Effect of legal drinking age on mortality
Callaghan et al. [30]	2013	American Journal of Public Health	Effect of legal drinking age on alcohol-related morbidity
Callaghan et al. [31]	2013	Addiction	Effect of legal drinking age on inpatient morbidity
Carpenter and Dobkin [16]	2009	AEJ: Applied Economics	Effect of alcohol consumption on mortality
Carpenter and Dobkin [32]	2011	Journal of Economic Perspectives	Minimum legal drinking age and public health
Chen et al. [33]	2013	PNAS	Effect of air pollution on mortality
Conover and Scrimgeour [34]	2013	Journal of Health Economics	Health effects of minimum legal drinking age
De La Mata [35]	2012	Health Economics	Effect of Medicaid eligibility on coverage, utilization, and health
Deza [36]	2014	Health Economics	Effect of alcohol use on drug consumption
Flam-Zalcman et al. [37]	2012	Intl J Psych Research	Effect of criterion-based increase in alcohol treatment
Fletcher [38]	2014	Biodemography and Social Biology	Effect of genetics on stress response
Glance et al. [39]	2014	JAMA Surgery	Effect of hospital report cards on mortality
Gormley et al. [40]	2005	Developmental Psychology	Effect of universal pre-kindergarten on cognitive development
Huang and Zhou [41]	2013	Social Science and Medicine	Effect of education of cognition
Jensen and Wust [42]	2014	Journal of Health Economics	Effect of Caesarean section on maternal and child health
McFarlane et al. [43]	2014	Schizophrenia Bulletin	Effect of treatment program on psychosis onset
Miller et al. [44]	2013	AEJ: Applied Economics	Effect of insurance on health spending, utilization, and health
Nishi et al. [45]	2012	Bulletin of the WHO	Health effects of patient cost-sharing
Pierce et al. [46]	2012	Pers Soc Psych Bulletin	Effect of income disparity in marriage
Sloan and Hanrahan [47]	2014	JAMA Ophthalmology	Effect of new therapies on vision loss among elderly patients
Smith et al. [48]	2014	Canadian Medical Association Journal	Effect of HPV vaccine on sexual behavior
Sood et al. [49]	2014	BMJ	Effect of health insurance on mortality
Weaver et al. [50]	2010	Journal of Traumatic Stress	Effect of cognitive-behavioral therapy on trauma symptoms
Yörük and Yörük [51]	2012	Social Science and Medicine	Effect of alcohol on psychological well-being

Abbreviations: BMI, body mass index; HIV, human immunodeficiency virus; AEJ, American Economic Journal; PNAS, Proceedings of the National Academy of Sciences; Intl J Psych Research, International Journal of Methods in Psychiatric Research; JAMA, Journal of the American Medical Association; WHO, World Health Organization; Pers Soc Psych Bulletin, Personality and Social Psychology Bulletin; HPV, human papillomavirus; BMJ, British Medical Journal.

See references for full citation information: [1,15,16,23–51].

Table 2. Inclusion of key elements in regression discontinuity publications

Author	Year	Discussion of RD validity conditions	Discussion of assignment rule	Covariate balance tests	Histogram of assignment variable	Robustness checks	Total (0-5)
Albouy and Lequien [23]	2009	✓	✓	✗	✗	✓	3
Almond et al. [15]	2010	✓	✓	✓	✓	✓	5
Andalón [24]	2011	✓	✓	✓	✓	✓	5
Anderson et al. [25]	2011	✓	✓	✗	✗	✗	2
Arcand and Wouabe [26]	2010	✓	✓	✓	✗	✓	4
Banks and Mazzonna [27]	2012	✓	✓	✓	✗	✓	4
Behrman [28]	2014	✓	✓	✗	✗	✓	3
Bor et al. [1]	2014	✓	✓	✓	✓	✓	5
Callaghan et al. [29]	2014	✓	✓	✗	✗	✓	3
Callaghan et al. [30]	2013	✓	✓	✓	✗	✓	4
Callaghan et al. [31]	2013	✗	✓	✗	✗	✓	2
Carpenter and Dobkin [16]	2009	✓	✓	✓	✗	✓	4
Carpenter and Dobkin [32]	2011	✓	✓	✗	✗	✗	2
Chen et al. [33]	2013	✓	✓	✓	✓	✓	5
Conover and Scrimgeour [34]	2013	✗	✓	✗	✗	✓	2
De La Mata [35]	2012	✓	✓	✓	✗	✓	4
Deza [36]	2014	✓	✓	✓	✗	✓	4
Flam-Zalcman et al. [37]	2012	✓	✓	✗	✗	✗	2
Fletcher [38]	2014	✓	✓	✓	✓	✓	5
Glance et al. [39]	2014	✗	✓	✓	✗	✓	3
Gormley et al. [40]	2005	✓	✓	✓	✗	✓	4
Huang and Zhou [41]	2013	✓	✓	✗	✗	✓	3
Jensen and Wust [42]	2014	✓	✓	✓	✗	✓	4
McFarlane et al. [43]	2014	✓	✓	✓	✗	✓	3
Miller et al. [44]	2013	✓	✓	✓	✓	✓	5
Nishi et al. [45]	2012	✗	✓	✗	✗	✗	2
Pierce et al. [46]	2012	✓	✓	✓	✗	✓	5
Sloan and Hanrahan [47]	2014	✗	✓	✓	✗	✗	2
Smith et al. [48]	2014	✓	✓	✓	✗	✓	4
Sood et al. [49]	2014	✓	✓	✓	✗	✓	4
Weaver et al. [50]	2010	✗	✓	✓	✗	✗	3
Yörük and Yörük [51]	2012	✓	✓	✓	✗	✓	4

Abbreviation: RD, regression discontinuity.
See references for full citation information: [1,15,16,23–51].

included a discussion of the assignment rule and most included a discussion of the RD validity conditions. The most commonly omitted element was a histogram (or description of the distribution) of the assignment variable (25 of 32 omitted).

4. Discussion

There are many opportunities to implement RD designs in medicine, epidemiology, and public health, where treatments are often assigned based on threshold rules. Clinical cohort studies often include the three data elements required for implementation of this technique: information on a continuous diagnostic criterion (Z) used with a cutoff rule to assign treatment, information on whether treatment was received, and outcomes for both treated and untreated individuals. RD is a powerful quasi-experimental method with several advantages over other commonly used approaches for causal inference from observational data, as illustrated by the examples in Box 1. RD's greatest strength is its ability to achieve balance on unobserved factors—much like an RCT. In contrast, the methods based on regression-adjustment and matching that are commonly used to analyze clinical cohorts must rely on the strong assumption that there are no unmeasured confounders. Further, inference using RD relies on weaker assumptions than most other quasi-experimental designs, such as IVs and difference-in-difference approaches [6].

RD even has several advantages over randomized controlled experiments, and in some cases might be considered as an alternative to the clinical trial: First, RD offers an opportunity to exploit preexisting data and therefore is less costly than experimental methods. As discussed above, the data required are often already collected in cohort studies or could be collected or through relatively simple and inexpensive modifications to current collection procedures. Alternatively, cohort data can be linked to existing administrative data, e.g. national registries. Second, the local nature of the RD estimator has advantages in interpretability over a population average effect because it reflects the TE on the marginal unit near the cutoff point. In cases where we are interested in the impact of the cutoff itself, and in optimizing threshold rules, this is precisely the quantity we want to estimate. For example, it would be ideal for whether persons on the margins of eligibility for a supplementary feeding program are benefiting from the program

Box 1 Example study question where RD may be useful

What is the survival impact of early vs. deferred antiretroviral therapy (ART) for HIV patients in sub-Saharan Africa?

Study design	Description	Strengths and weaknesses
Randomized controlled trial [19,20]	Randomly assign HIV patients to immediate ($200 \leq CD4 < 350$) vs. deferred ($CD4 < 200$) ART. Collect survival data for treated and controls; compare using Kaplan–Meier estimator and hazard regression models. Monitor CD4 counts of control subjects and initiate them on ART when eligible to determine efficacy of early vs. deferred ART.	**Strengths** 1. Randomization guarantees balance on both observed and unobserved covariates, in expectation. 2. Valid counterfactual; RCTs can estimate intent-to-treat (ITT) and complier average causal effects (CACE). **Weaknesses** 1. Often a treatment is protective, but the effect size or its generalizability across settings is unknown. If clinical equipoise cannot be met, then an RCT would be unethical. 2. Conducted in controlled settings to assess efficacy but may not be informative of real-world effectiveness. 3. RCTs are expensive, logistically difficult. 4. Because of screening criteria and opt-in consent, study subjects may not be representative of population of interest.
Traditional cohort study [21,22]	Nest additional data collection in existing HIV treatment program. Compare survival among HIV patients who initiate ART at different CD4 counts. Control for available baseline predictors in hazard regression models to reduce confounding.	**Strengths** 1. Relatively easy to implement and inexpensive. 2. Large, representative samples can be obtained. 3. Can evaluate treatment outcomes in real world settings **Weaknesses** 1. Cannot control for unobserved confounders, correlated with both ART delay and survival. 2. Studies typically exclude HIV patients who seek care but do not initiate ART, leading to selection bias. 3. No valid counterfactual, no causal effect.
Regression discontinuity design [1]	Nest additional data collection in existing HIV treatment program. Assess survival among patients with different CD4 counts at initial presentation. Follow-up all patients. Exploit threshold rule (start ART if $CD4 < 200$) to obtain causal effect. Compare predicted survival for patients immediately above vs. below the treatment threshold using hazard regression models. Assess effects at different thresholds that have been implemented in different settings, in order to optimize treatment guidelines.	**Strengths** 1. Relatively easy to implement and inexpensive. 2. Large, representative samples can be obtained. 3. Informative of real-world effectiveness. 4. Random variability in measured CD4 counts yields local randomization at threshold; balance is achieved in both observed and unobserved factors, in expectation. 5. Includes all patients who seek care and have a CD4 test. 6. Valid counterfactual; estimates ''local'' ITT and CACE. **Weaknesses** 1. Local causal effects identified at the threshold so may not be generalizable to CD4 counts far from the threshold. But local effect of marginal change in treatment threshold is often of interest.

and whether eligibility should be contracted or expanded. However, the local nature of the RD estimator does mean that it may not be generalizable to observations far from the cutoff. Third, when ethical, political, or technical reasons preclude random allocation of treatment, RD analyses may be used to establish causal effect sizes if the treatment in question is assigned via a cutoff rule. In particular, such an opportunity for RD analysis will be valuable when a lack of equipoise, existing clinical practice, or political constraints make randomized trials infeasible [52]. For example, a study in our systematic review by Weaver et al. [50] uses RD to evaluate a program that assigns treatment based on severity of symptoms and cites the ethical constraints of conducting an experiment with a vulnerable population suffering from posttraumatic stress as a motivation for the use of a RD design. Fourth, RD is often better for estimating ''real-life'' effectiveness as opposed to clinical efficacy in the highly controlled settings of a randomized experiment and can be used for evaluation after a policy is already implemented. For impact evaluations of large-scale public health programs, governments and policy makers may be more concerned with the de facto effects as opposed to the effects under ideal conditions. In most of these cases, the fuzzy form of RD will be used because the cutoff will only determine treatment status for some patients. Studies comparing the performance of RD to randomized trials find that it performs well [53,54].

Given its advantages vis-a-vis other study designs, it is surprising that RD has not been more widely used in medicine, epidemiology, or public health. With only two exceptions [1,15], all of the 32 RD studies in this systematic review aimed to establish the effectiveness of social policies or mental health interventions, despite the potential of RD to be used to answer many research questions about clinical interventions to improve physical health. One of the only two studies applying RD to study an intervention to improve physical health was published in an economics journal [15]. Almond et al. [15] use the low birth weight threshold rule to determine eligibility of medical interventions for low birth-weight babies on infant mortality. The other study applying RD to study an intervention to improve physical health was published in an epidemiology journal. Bor et al. [1] exploit the CD4 count threshold rule used to determine eligibility for HIV treatment to estimate the causal effect of immediate vs. deferred antiretroviral therapy on mortality.

The review results demonstrate increasing interest in RD in recent years. Almost one third of all studies using RD included in our systematic review were published in 2014. However, almost all of the recent studies, including those in health journals, aimed to establish the effectiveness of social policies, rather than the effectiveness of clinical interventions. The potential of RD to generate insights on the effectiveness of a wide range of clinical interventions aiming to improve physical health yet is to be realized. Although our search for RD studies yielded 193 records in PubMed, a search for the MeSH term ''randomized controlled trial'' yielded 490,086 records over the same period. Yet, there are many circumstances where RD is preferred to randomized trials, either for theoretical reasons (the local nature of the RD estimator) or practical reasons that preclude usage of RCTs. RDs can answer questions that experiments cannot and that other quasi-experimental methods cannot answer as convincingly, such as the local effect of changing diagnostic guidelines or impact evaluations of programs as they are truly implemented in the context of health system imperfections.

One possible explanation for the underutilization of RD designs in epidemiology, medicine, and public health is a lack of agreement on the underlying principles and terminology, as Cook [8] argues and our results in Table 2 suggest. In 1996, the American Journal of Public Health published two articles and a commentary that urged the use of a risk-based allocation design [55−57], which is a subtype of RD where treatments are allocated based on clinically measured risk. The principle underpinning this design is very similar to RD, but the relationship with the broader class of RD studies is not made explicit.

Interestingly, a similar design, interrupted time series (ITS), has been widely used in epidemiology to estimate the effects of policy changes. In the most rigorous ITS scenario—when a policy is implemented very rapidly and short-term outcomes are assessed at frequent intervals—ITS can be interpreted as a sub-type of the RD design, in which calendar time is the assignment variable Z and the cutoff occurs when a new policy is implemented. The acceptance of ITS studies indicates that the basic concepts of RD are already familiar to researchers in epidemiology and public health but that the intuition behind ITS has not been generalized to the full range of applications that can be analyzed with an RD design, in which threshold rules on other continuous baseline characteristics influence treatment assignment.

The results of our systematic review suggest that despite an overall good quality of RD articles in medicine, epidemiology, and public health, there are several elements that are frequently not presented. We identified validity conditions for RD designs and guidelines on which elements should always be presented to support the validity of the design. Without these, the articles cannot establish as convincingly that their data and analysis meet the conditions for a valid RD study. Currently, these conditions and their corresponding empirical recommendations have not been codified in medicine, epidemiology, or public health in contrast to the STROBE Statement for reporting of observational studies and the CONSORT Statement for reporting RCTs. Standardization of the presentation of RD studies in these fields would enable readers to more readily discern the quality of these studies and would elevate the level of quality of RD analysis and their acceptability in the literature. The elements we suggest here are a starting point for this work. We used a sum of the included

elements to quickly summarize the findings presented in Table 2, but this does not provide information about the relative importance of each element, which should be incorporated into a quality score. Future studies should adhere to these guidelines to ensure transparency in reporting and to increase confidence in study validity.

5. Conclusion

In light of RD's many potential health applications, its advantages vis-a-vis other study designs, the availability of required data, and the proliferation of RD in other fields, RD is currently underutilized in medicine, epidemiology, and public health. Increased use of RD designs in these fields could lead to a wide range of novel insights into causal effects, when randomized controlled experiments are not feasible or cannot generate the answers to questions about the effectiveness of interventions implemented under the imperfect conditions that are pervasive in real-life health systems. Institutions involved in synthesizing evidence for medical practice and health policy, such as the Cochrane Collaboration and the GRADE system, should consider adding RD as a study design that can generate results of strong causal validity, and not just another observational study design. Most importantly, researchers in medicine, epidemiology, and public health should increasingly take advantage of existing cohort data and threshold treatment assignment rules to generate insights that are new and, in many cases, will be unlikely to be generated through other approaches.

References

[1] Bor J, Moscoe E, Mutevedzi P, Newell ML, Bärnighausen T. Regression discontinuity designs in epidemiology: causal inference without randomized trials. Epidemiology 2014;25:729−37.
[2] Rothwell PM. External validity of randomized controlled trials: "to whom do the results of this trial apply?". Lancet 2005;365:82−93.
[3] Thistlethwaite DL, Campbell DT. Regression-discontinuity analysis: an alternative to the ex post facto experiment. J Educ Psychol 1960;51(6):309.
[4] Rubin DB. Assignment to treatment group on the basis of a covariate. J Educ Stat 1977;2(1):1−26.
[5] Berk RA, Rauma D. Capitalizing on nonrandom assignment to treatments: a regression-discontinuity evaluation of a crime-control program. J Am Stat Assoc 1983;78:21−7.
[6] Lee DS, Lemieux T. Regression discontinuity designs in economics. J Econ Lit 2010;48(2):281−355.
[7] Imbens GW, Lemieux T. Regression discontinuity designs: a guide to practice. J Econom 2008;142(2):615−35.
[8] Cook TD. "Waiting for life to arrive": a history of the regression-discontinuity design in psychology, statistics and economics. J Econ 2008;142(2):636−54.
[9] Lee DS. The electoral advantage to incumbency and voters' valuation of politicians' experience: a regression discontinuity analysis of elections to the US House No. w8441. Cambridge, MA: National Bureau of Economic Research; 2001.
[10] Imbens G, Van Der Klaauw W. Evaluating the cost of conscription in the Netherlands. J Bus Econ Stat 1995;13(2):207−15.
[11] Angrist JD, Lavy V. Using Maimonides' rule to estimate the effect of class size on scholastic achievement. Q J Econ 1999;114(2):533−75.
[12] Hahn J, Todd P, Van der Klaauw W. Identification and estimation of treatment effects with a regression-discontinuity design. Econometrica 2001;69(1):201−9.
[13] Lee DS. Randomized experiments from non-random selection in US house elections. J Econ 2008;142(2):675−97.
[14] Imbens G, Kalyanaraman K. Optimal bandwidth choice for the regression discontinuity estimator. Rev Econ Stud 2012;79(3):933−59.
[15] Almond D, Doyle JJ Jr, Kowalski AE, Williams H. Estimating marginal returns to medical care: evidence from at-risk newborns. Q J Econ 2010;125(2):591−634.
[16] Carpenter C, Dobkin C. The effect of alcohol consumption on mortality: regression discontinuity evidence from the minimum drinking age. Am Econ J Appl Econ 2009;1(1):164.
[17] Bor J, Moscoe E, Bärnighausen T. Three approaches to causal inference in regression discontinuity designs (Letter). Epidemiology, in press.
[18] McCrary J. Manipulation of the running variable in the regression discontinuity design: a density test. J Econ 2008;142(2):698−714.
[19] Cohen MS, Chen YQ, McCauley M, Gamble T, Hosseinipour MC, Kumarasamy N, et al. Prevention of HIV-1 infection with early antiretroviral therapy. N Engl J Med 2011;365(6):493−505.
[20] Severe P, Jean Juste MA, Ambroise A, Eliacin L, Marchand C, Apollon S, et al. Early versus standard antiretroviral therapy for HIV-infected adults in Haiti. N Engl J Med 2010;363(3):257−65.
[21] Sterne JAC, May M, Costagliola D, de Wolf F, Phillips A, Harris R, et al. Timing of initiation of antiretroviral therapy in AIDS-free HIV-1-infected patients: a collaborative analysis of 18 HIV cohort studies. Lancet 2009;373:1352−63.
[22] Kitahata MM, Gange SJ, Abraham AG, Merriman B, Saag MS, Justice AC, et al. Effect of early versus deferred antiretroviral therapy for HIV on survival. N Engl J Med 2009;360:1815−26.
[23] Albouy V, Lequien L. Does compulsory education lower mortality? J Health Econ 2009;28(1):155−68.
[24] Andalón M. Oportunidades to reduce overweight and obesity in Mexico? Health Econ 2011;20(Suppl 1):1−18.
[25] Anderson PM, Butcher KF, Cascio EU, Schanzenbach DW. Is being in school better? The impact of school on children's BMI when starting age is endogenous. J Health Econ 2011;30(5):977−86.
[26] Arcand J-L, Wouabe ED. Teacher training and HIV/AIDS prevention in West Africa: regression discontinuity design evidence from the Cameroon. Health Econ 2010;19(Suppl 1):36−54.
[27] Banks J, Mazzonna F. The effect of education on old age cognitive abilities: evidence from a regression discontinuity design. Econ J 2012;122(560):418−48.
[28] Behrman JA. The effect of increased primary schooling on adult women's HIV status in Malawi and Uganda: Universal Primary Education as a natural experiment. Soc Sci Med, in press. Published online 21 June, 2014.
[29] Callaghan RC, Sanches M, Gatley JM, Stockwell T. Impacts of drinking-age laws on mortality in Canada, 1980−2009. Drug Alcohol Depend 2014;138:137−45.
[30] 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(12):2284−91.
[31] Callaghan RC, Sanches M, Gatley JM. Impacts of the minimum legal drinking age legislation on in-patient morbidity in Canada, 1997−2007: a regression-discontinuity approach. Addiction 2013;108(9):1590−600.
[32] Carpenter C, Dobkin C. The minimum legal drinking age and public health. J Econ Perspect 2011;25(2):133.
[33] 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(32):12936−41.
[34] Conover E, Scrimgeour D. Health consequences of easier access to alcohol: New Zealand evidence. J Health Econ 2013;32(3):570−85.
[35] De La Mata D. The effect of Medicaid eligibility on coverage, utilization, and children's health. Health Econ 2012;21(9):1061−79.
[36] Deza M. The effects of alcohol on the consumption of hard drugs: regression discontinuity evidence from the National Longitudinal Study of Youth, 1997. Health Econ 2014. http://dx.doi.org/10.1002/hec.3027.
[37] Flam-Zalcman R, Mann RE, Stoduto G, Nochajski TH, Rush BR, Koski-Jännes A, et al. Evidence from regression-discontinuity analyses for beneficial effects of a criterion-based increase in alcohol treatment. Int J Methods Psychiatr Res 2012;22(1):59−70.
[38] Fletcher JM. Enhancing the gene-environment interaction framework through a quasi-experimental research design: evidence from differential responses to September 11. Biodemography Soc Biol 2014;60(1):1−20.
[39] Glance LG, Osler TM, Mukamel DB, Meredith JW, Dick AW. Effectiveness of nonpublic report cards for reducing trauma mortality. JAMA Surg 2014;149(2):137−43.
[40] Gormley WT Jr, Gayer T, Phillips D, Dawson B. The effects of universal pre-K on cognitive development. Dev Psychol 2005;41(6):872.
[41] Huang W, Zhou Y. Effects of education on cognition at older ages: evidence from China's Great Famine. Soc Sci Med 2013;98:54−62.
[42] Jensen VM, Wust M. Can Caesarean section improve child and maternal health? The case of breech babies. working paper available at: http://sole-jole.org/13404.Pdf. J Health Econ 2012. http://dx.doi.org/10.1016/j.jhealeco.2014.07.004. 2014 Aug 13. pii: S0167-6296(14)00096-4. [Epub ahead of print].
[43] McFarlane WR, Levin B, Travis L, Lucas FL, Lynch S, Verdi M, et al. Clinical and functional outcomes after 2 years in the early detection and intervention for the prevention of psychosis multisite effectiveness trial. Schizophr Bull 2014;sbu108.
[44] Miller G, Pinto D, Vera-Hernández M. Risk protection, service use, and health outcomes under Colombia's health insurance program for the poor. Am Econ J Appl Econ 2013;5(4):61−91.
[45] Nishi A, McWilliams JM, Noguchi H, Hashimoto H, Tamiya N, Kawachi I. Health benefits of reduced patient cost sharing in Japan. Bull World Health Organ 2012;90(6):426−435a.
[46] Pierce L, Dahl MS, Nielsen J. In sickness and in wealth: psychological and sexual costs of income comparison in marriage. Pers Soc Psychol Bull 2013;39(3):359−74.
[47] Sloan FA, Hanrahan BW. The effects of technological advances on outcomes for elderly persons with exudative age-related macular degeneration. JAMA Ophthalmol 2014;132(4):456−63.
[48] 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 2014;cmaj-140900.
[49] 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.
[50] Weaver, Cavaleri M, Olin S, Wisdom J. Implementation of CBT for youth affected by the World Trade Center disaster: matching need to treatment intensity and reducing trauma symptoms. J Trauma Stress 2010;23(6):699−707.
[51] Yörük Ertan C, Yörük BK. The impact of drinking on psychological well-being: evidence from minimum drinking age laws in the United States. Soc Sci Med 2012;75(10):1844−54.
[52] Bonell CP, Hargreaves J, Cousens S, Ross D, Hayes R, Petticrew M, Kirkwood BR. Alternatives to randomization in the evaluation of public health interventions: design challenges and solutions. J Epidemiol Community Health 2011;65:582−7.
[53] Cook TD, Wong VC. Empirical tests of the validity of the regression discontinuity design. Ann Econ Stat 2008;91/92:127−50.
[54] Buddelmeyer H, Skoufias E. An evaluation of the performance of regression discontinuity design on PROGRESA World Bank Policy Research Working Paper 3386, The World Bank, September 2004.
[55] Finkelstein MO, Levin B, Robbins H. Clinical and prophylactic trials with assured new treatment for those at greater risk: I. A design proposal. Am J Public Health 1996;86:691−5.
[56] Finkelstein MO, Levin B, Robbins H. Clinical and prophylactic trials with assured new Clinical and prophylactic trials with assured new treatment for those at greater risk: II. Examples treatment for those at greater risk: II. Examples. Am J Public Health 1996;86:696−705.
[57] Mosteller F. The promise of risk-based allocation trials in assessing new treatments. Am J Public Health 1996;86:622−3.