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