EVALUATING PROGRAM EFFECTIVENESS USING THE REGRESSION POINT DISPLACEMENT DESIGN

ARIEL LINDEN
Linden Consulting Group, Portland, OR
WILLIAM M.K. TROCHIM
Cornell University
JOHN L. ADAMS
RAND Corporation, Santa Monica, CA

Most health care initiatives are evaluated using observational study designs in lieu of randomized controlled trials (RCT) due primarily to resource limitations. However, although observational studies are less expensive to implement and evaluate, they are also more problematic in determining causality than the RCT. This trade off is most apparent in the initial planning stage of program development. An RCT is generally preferred though the cost of implementing a pilot program using the RCT might outstrip the potential benefit if the desired results are not obtained. This article describes a simple quasi-experimental model called the regression point displacement (RPD) design, which compares the prepost results of a single or multiple treatment groups to that of a control population. This design has shown great potential in evaluating health care pilot programs or demonstration projects—especially those that are community based—due to its relative ease of implementation and low cost of analysis.

Keywords: *quasi-experimental; regression point displacement; internal validity*

Most health care initiatives are evaluated using observational study designs in lieu of randomized controlled trials (RCT) due primarily to resource limitations. However, although observational studies are less expensive to implement and evaluate, they are also more problematic in determining causality than the RCT. This trade off is most apparent in the initial planning stage of program development. An RCT is generally preferred though the cost of implementing a pilot program using the RCT might outstrip the potential benefit if the desired results are not obtained. This article describes a simple quasi-experimental model called the regression point displacement (RPD) design, which compares the prepost results of a single or multiple treatment groups to that of a control population. This design has shown great potential in evaluating health care pilot programs or demonstration projects—especially those that are community based—due to its relative ease of implementation and low cost of analysis.
Keywords: quasi-experimental; regression point displacement; internal validity

Outside of clinical medicine where the randomized controlled trial (RCT) is the research method of choice, most health care initiatives are evaluated using nonexperimental or quasi-experimental study designs. The decision to use an observational design over an RCT involves several trade-offs. The RCT is generally more costly to implement on a prospective basis than many of the alternatives, requires strict adherence to protocol over an extended period of time, and is limited in generalizability of outcomes to the greater diseased population (i.e., study participants tend to be extremely narrowly focused). However, when properly conducted, RCT study outcomes are considered strong in internal validity and especially useful for causal interpretation. Nonetheless, the RCT can also be influenced by threats to validity that may lead to an incorrect determination of causality (Linden & Roberts, 2005). In contrast, quasi-experimental studies tend to be less expensive and intrusive to conduct, but their design is susceptible to threats to validity such as selection bias, statistical regression-to-the-mean and measurement error that make it more difficult to determine a causal relationship between treatment and outcome (Campbell & Stanley, 1996; Shadish, Cook, & Campbell, 2002).

Although the level of control achieved by the RCT would be welcomed by health care executives to assist in the business decision-making process, in general a substantially smaller "burden of proof" of a causal relationship would suffice. One area where the concern for accurate results does exist is in the initial planning stage of program development. Here too a trade-off exists. As shown in Figure 1, administrators must choose between widespread adoption of a yet unproven program and conducting a pilot or demonstration project to first validate its findings. If the ubiquitous implementation option

Program Development

Pilot/Demonstration Project

Randomized
Controlled Trial

Observational
Study Design

If successful

Widespread Implementation

Widespread Implementation

Randomized
Controlled Trial

Observational
Study Design

Randomized
Controlled Trial

Observational
Study Design
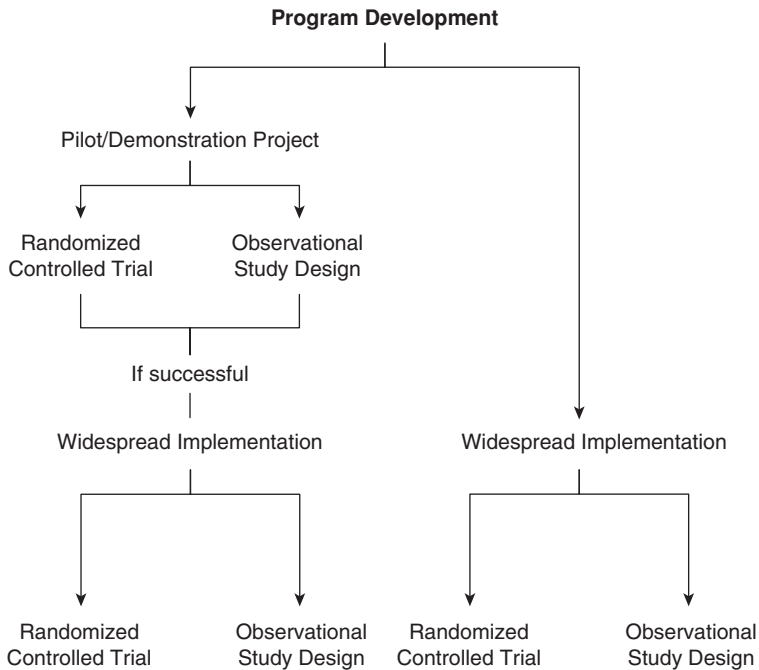
Figure 1: The Program Development and/or Implementation Process

proves unsuccessful, precious financial and human resources would have been squandered.

A pilot project provides valuable information about the viability of further program expansion at much lower resource intensity than a complete program implementation would. The choice of conducting an RCT versus implementing a quasi-experimental study design exists at the demonstration phase as well. Again, the primary trade-off is between cost and accurate interpretation of results. Even on this scale, in some cases the RCT is cost-prohibitive and thus the evaluation will naturally default to an observational study design.

One novel approach to evaluating the results of a pilot program is called the Regression Point Displacement (RPD) design. The term was coined by William Trochim and Donald Campbell (Trochim & Campbell, 1996) and represents a simple pretest-posttest quasi-experimental study design (see Figure 1) in which one experimental group (the pilot) is compared to multiple control groups on

aggregate-level data. The estimate of the effectiveness of the treatment is accomplished by comparing the vertical displacement of the treated group posttest from the pre-post regression line of the control groups (Trochim & Campbell, 1996, 1999). This article describes the RPD design and provides several examples in health care contexts to assist the reader in applying this methodology to their particular situation or setting. In addition, a comprehensive discussion of limitations of the design will be presented.

ELEMENTS OF THE REGRESSION POINT DISPLACEMENT DESIGN

The RPD design is premised on a regression model in which an outcome for the one treatment group is compared to that of a control population (see appendix). Effectiveness of the intervention is described by the beta coefficient for the dichotomous variable representing group assignment (1 = treatment, 0 = control). Although conceptually this model is straightforward and easy to compute, ensuring validity of the outcomes requires a thoughtful approach to its application and rigorous implementation of the design.

DETERMINATION OF TREATMENT GROUP ASSIGNMENT

How a particular group or unit is chosen for assignment to the intervention can affect the validity of study findings. There is broad consensus that randomization to treatment helps minimize the potential for major biases related to causal assessment. In this case, random selection of the treatment group provides protection against the investigator unconsciously (or consciously) selecting a group with unmeasured or subjective characteristics that will likely influence the treatment's effectiveness. With only one unit or group chosen for the intervention, of course, it is still possible that baseline characteristics will differ between this group and the control groups. We note that the design can be extended to incorporate covariates if baseline differences are a concern. However the RPD design does not require pretest equivalence between the intervention group and controls. The design assumes that the intervention group posttest does not differ significantly from the regression line prediction. Instead, the primary concern is whether the control groups yield an unbiased estimate of the

true population regression line and/or whether the treatment group is a member of the control group population (Trochim & Campbell, 1996, 1999).

If a particular group is chosen for the intervention based on its outlier pretest measure, the RPD design is most comparable to the regression discontinuity (RD) design (Trochim, 1984, 1990; Linden, Adams, & Roberts, 2005) in that the outlier value corresponds to the cutoff point that determines assignment category. However, it is important for this cutoff value to represent an objectively established criterion or else the RPD design will not provide optimal control for regression to the mean effect (or potential selection bias). The RPD design corresponds to a nonequivalent group design (NEGD) if the group volunteers or is selected for participation due to other subjective reasons. This is the least desired assignment method as it is the most difficult to determine causality.

THE UNIT OF MEASURE

There is much flexibility in the RPD design methodology for determining the unit of measure. In health care, the most common units of measure for the RPD design will be physicians, clinics, hospitals (and their subunits), purchasers, counties, or other logical groupings. Aggregating outcome data from within these units (to create a rate or average) reduces variability and improves statistical power as compared to individual measurements of the unit itself (Trochim & Campbell, 1996, 1999). To clarify this point further, assume a study in which 10 medical groups are compared where 1 group received the treatment and 9 did not. The unit of measure in this study is the medical group. *Individual* group measures may include the following: type of medical group (general practice or specialty care), reimbursement method (fee schedule or capitation), location (urban or rural), and so on. With a sample size of 10 units, these individual group level measures can be widely variable and thus limit the ability to draw conclusions from the outcomes. However, measures that are *aggregated* to the group level such as; office visits per thousand patients, prescriptions per patient per year, and so on, have much less variability because the denominator for each measure is based on size of the group's population (which in most cases can be in the hundreds or thousands).

NUMBER OF TREATMENT GROUPS

In its most elementary form, the RPD design will comprise a single intervention group. However, situations may arise in which more than one group will participate in the treatment arm (i.e., when the study organizers may be contractually obligated to administer the same intervention to multiple groups). As will be discussed later, the results can then be combined to determine the joint effect estimate. As the number of experimental groups approaches that of controls, the design assumes a form that is akin to the more standard designs (depending on the method of assignment, either the RCT, the NEGD or the RD design).

Another possible variation on the RPD design relative to the number of treatment groups is the implementation of multiple interventions. For example, assume a health management program wants to test whether any of three health behavior classes affect health-related behaviors for patients. In this scenario, program administrators may designate three clinics as treatment settings, each providing one of the three classes. The RPD design can then determine whether any or all three of the classes were effective in eliciting the desired health behaviors. The limitations associated with these model enhancements will be discussed later.

CHOICE OF PREMEASURES AND POSTMEASURES

The RPD design does not require that the preprogram and postprogram measures be identical but for it to be used effectively, a substantial correlation between preprogram and postprogram measures is preferable. Therefore, in most cases, it is both practical and easier to interpret the results when the prevariables and postvariables are the same. For example, when implemented in both measurement periods, health status or quality of life instruments provide a clear and meaningful depiction of the longitudinal impact of a health management intervention at the aggregate level. Conversely, the researcher may choose to characterize the relationship between a preprogram variable such as health status and a postprogram outcome such as a hospitalization rate, as affected by an intervention.

NUMBER OF COVARIATES

In its basic form, the RPD design uses a simple linear regression model where the outcome variable is regressed on the one
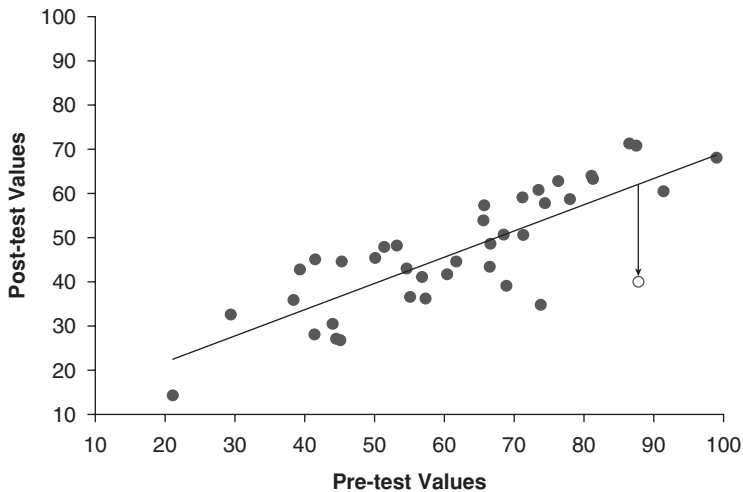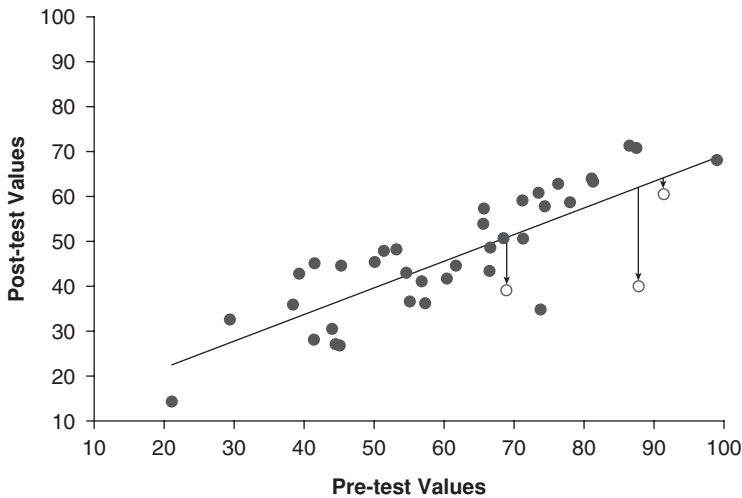
Figure 2: Graphic Display of the Single Treatment Group Regression Point Displacement (RPD) Design

NOTE: Precoordinate and postcoordinate values are plotted for each unit and/or group, and fitted with a regression trend line (diagonal line). The vertical arrow points to the treatment group's value (open circle), illustrating an RPD from the trend line.

pre-intervention variable and a program indicator. Patient- or group-level covariates may be added when deemed appropriate, but for each additional covariate an increase in the number of cases will be required to maintain the statistical power of the test (Trochim & Campbell, 1996, 1999). Naturally, covariates are more effective when multiple units are assigned to the treatment group. In the case described earlier of multiple interventions across multiple sites, dummy variables will be added to the model as covariates to represent each group and/or intervention.

VISUAL DISPLAY AND INTERPRETATION OF RPD DESIGN RESULTS

In many cases, visual inspection of the data will provide a clue as to the form that the RPD design will take. The approach is to graph the X-Y data for each group or unit using a scatterplot. Most basic computer programs with graphing capabilities can also provide a regression trend line. Figure 2 illustrates a possible result of a single treatment

Figure 3: Graphic Display of Multiple Treatment Groups Used in Regression Point Displacement (RPD) Design

NOTE: Precoordinate and postcoordinate values are plotted for each unit and/or group and fitted with a regression trend line (diagonal line). The vertical arrow points to the treatment groups' values (open circle), illustrating an RPD from the trend line.

group RPD analysis. In this hypothetical example, pre- and post-intervention score coordinates are plotted and a regression line is fitted to the data. The single intervention group value is located and visually inspected for its displacement from the regression line. A vertical line or arrow can be added to the graph in order to provide perspective of this displacement relative to the control groups' data points. In this example, the treatment shows a positive effect (a decrease in the outcome variable), as indicated by the approximate 30 point displacement from the regression line.

Figure 3 depicts a case in which two additional groups (for a total of three) were assigned to the intervention. As illustrated, the point displacement varied between all three groups, indicating a differential treatment effect.

Although visual inspection of the data and graphic displays provide a general idea as to the structural form of the RPD model, statistical

TABLE 1
Regression Model Statistics for a Hypothetical Simple Regression Point Displacement Design (RPD) Design

	Coefficients	SE	t Stat	p value
Intercept	8.897	5.699	1.561	.127
Pretest (X)	1.118	0.117	9.597	<.001
Group (Z)	34.171	9.742	3.508	.001
Model statistics				
Adjusted R^2	71.900%			
Standard error	9.580			
Observations	39.000			

TABLE 2
Regression Model Statistics for a Hypothetical Regression Point Displacement Design (RPD) Design in which Three Treatment Groups Were Compared to Controls

	Coefficients	SE	t Stat	p Value
Intercept	9.491	5.424	1.750	.089
Pretest (X)	1.087	0.111	9.798	<.0001
Group (Z)	22.618	5.502	4.111	.0002
Model statistics				
Adjusted R^2	74.380%			
Standard error	9.160			
Observations	39.000			

modeling must be performed to verify whether the displacement is indeed statistically significant. For readers interested in the actual statistical analysis, the appendix provides a comprehensive discussion of the modeling procedure. Table 1 presents the results of the regression analysis for the single treatment group data presented in Figure 2. There appears to be a statistically significant treatment effect as indicated by the assignment variable Z ($p = .001$), with a displacement (β^2 value) of 34.2 points from the mean. This value is only slightly higher than what was determined on visual inspection.

Table 2 presents the regression analysis results for the multiple treatment group model. Although Figure 3 shows that the intervention appeared to have a differential effect on each group's outcomes, the treatment effect is significant ($p = .0002$), with an average displacement (β^2 value) of 22.6 points from the slope. If the intervention groups are hypothesized to have a common effect, a single intervention indicator

can be used. If the multiple interventions are different in design or implementation, a separate indicator for each intervention group may be included.

APPLICATIONS OF THE RPD DESIGN IN HEALTH CARE

The RPD design is applicable to many areas of health care program evaluation. In this section, examples are provided from the various segments of the health care industry to assist the reader in generating ideas for further study.

MANAGED CARE ORGANIZATIONS

Managed care organizations (MCOs) may be in the best overall position to use the RPD design. These organizations contract with many and various types of health service providers, such as physicians, hospitals, pharmacies, ancillary providers, and so on, and have access to multiple administrative databases. These databases include claims, quality-related information such as complaints, appeals of denied services, and satisfaction surveys. Additional files include member demographics and payer profiles. Medical and financial data are already evaluated at the aggregate level to provide high level comparisons between providers of care, or between groups of members, and are typically pooled by disease or type of insurance.

In this setting, the RPD design could be used to test the utility of a pay-for-performance (P4P) initiative in which financial incentives are given to providers of health services to encourage improvement in quality of care. A clinic or medical group practice could be assigned to receive the P4P incentive whereas the rest of the clinics are reimbursed as usual. Any clinically valid indicator could be used for the pretest and posttest measure, but for illustrative purposes assume that the "percentage of congestive heart failure patients receiving Angiotensin-converting Enzyme (ACE) Inhibitors" serves as the premeasure and postmeasure. These data are captured in medical and pharmacy claims and can be readily analyzed. A clinic could be randomly assigned to the P4P initiative or be chosen due to its low ACE compliance rate. At the end of the program period a remeasurement would be performed and the appropriate statistical analysis (as described in the appendix) would be conducted. The results of the

demonstration project will assist the MCO in determining whether the P4P should be expanded to the entire network or scrapped.

HOSPITALS

Hospitals are structured organizationally in a way that is very suitable for the RPD model to be used in evaluations. For example, assume a hospital is operating at maximum capacity and determines that a systematic delay in discharging patients is creating a bottleneck for newly admitted patients awaiting a bed. Furthermore, assume that this delay is a result of physicians beginning their rounds at a late hour so that they will not be discharging some patients until later in the day. Hospital administrators could conduct a study in which one unit or floor identified as the outlier is assigned to the intervention whereas the others continue on as usual. The pretest and posttest measurement could be the % of patients discharged after 12 p.m. (with a zero percentage being ideal). The intervention could entail physicians beginning their rounds at the earliest mutually agreed-on hour, and that easier cases will be discharged first. After the study period, a remeasurement of the post-noon discharge rate will be conducted, and the unit compared statistically to the control units.

CLINICS

An ongoing concern among patients, and one that is typically noted in patient satisfaction surveys, is the wait time to see their physician. The wait-time reasons are many, and may be structural as well as process-driven. Assume that a clinic wants to reduce the dissatisfaction level caused by wait times by providing an entertaining health education video. To meet the RPD requirements of aggregated or pooled data, a physician would be randomly chosen as the unit of measure and all of his and/or her patients would receive the health education after signing in. A patient satisfaction survey would be used as both the pretest and posttest and would be administered at the completion of the visit. One question would relate specifically to the wait time: "On a scale of 1 – 5, were you satisfied with the wait time (1 being *very dissatisfied* and 5 being *very satisfied*)." At the end of the study period, the scores would be averaged across patients for each physician and compared statistically to determine if the video improved wait-time satisfaction.

POPULATION-BASED PROGRAMS

Disease or health management programs typically implement interventions at a population level, with the population defined as an MCO membership, employees at a company, residents of counties, states, or even at the national level (e.g., Medicare recipients). Pilot programs or demonstration projects at this stratum can be quite costly and resource intensive, thus the RPD design is uniquely suited for these endeavors. For example, assume that a State Medicaid program would like to conduct a pilot program targeting disease self-management for diabetics to determine if a large-scale implementation would be a cost-effective mechanism to improve health status and reduce costs to Medicaid. With the hypothesis being that emergency department (ED) visits could be reduced as a consequence of improved understanding of the disease process and self-management techniques, all counties within the State would be compared on baseline ED visit rates for diabetes. A county could be chosen randomly for the intervention and the program would then be implemented. At the end of the program period, the diabetes ED visit rates would be remeasured and the RPD analysis conducted. All else being equal, the regression results would indicate a positive program effect. The State could then use these figures to estimate the cost benefit of implementing the program on a State-wide basis.

VALIDITY OF THE RPD DESIGN

SELECTION BIAS

Concerns for selection bias arise when a group or unit either volunteers to participate in the program or when selection is based on subjective criteria. In this case, nearly any explanation offers a plausible rival hypothesis to a causal link between intervention and outcome (Linden, Adams, & Roberts, 2003).

As with any quasi-experimental design, the method of assignment to the intervention is the primary source of bias facing the RPD design. However, this can be partially mitigated by randomly assigning the group to the treatment. Unlike other quasi-experimental designs, baseline equivalence is not a prerequisite in the RPD, but the choice of an outlier to represent the treatment group increases the possibility of regression toward the mean. The likelihood of this

effect can be investigated by examining historical data to see if the group or unit's relative position with respect to the remainder of the sample is temporally stable.

Based on the discussion above it should be clear that the value of randomization in the RPD design is not to achieve pretest equivalence between treatment and control groups but more so as a means of ensuring that assignment to treatment is not biased (whether intentionally or not) by the researchers in favor of a particular group. Similarly, in assigning a group to treatment based on its outlier pretest value (consistent with the RD design) it is important for this cutoff value to represent an objectively established criterion. For example, the overuse of ED services by Medicaid recipients is well-documented (Cunningham, 2006). A plot of income versus ED visit rates would provide a cutoff at the Medicaid income eligibility level. Following an intervention targeting unnecessary ED utilization in this group, we would expect to see a decrease compared to controls on the regression line.

INSTRUMENTATION

Administrative databases (such as insurance claims, membership files, or electronic medical records) used for retrieving information on diagnostic measures to identify suitable participants, baseline characteristics, quality indicators, and utilization and cost values are notoriously inaccurate and may lead to spurious results (Jollis et al., 1993; Linden & Roberts, 2005).

Changes in instrumentation during repeated measurement of observations is a cause for concern. For example, data systems often become obsolete. In many cases as new systems are established, data cannot be easily moved from the retired system thereby raising the potential for missing data or a change in collection methods.

STATISTICAL POWER

Although intuitively it may seem that the low sample sizes typically used in the RPD design would reduce statistical power accordingly, it must be remembered that each point will often represent aggregated group data and not individual values. Such group means, totals or rates, are typically more stable and precise than within-group data (Trochim & Campbell, 1996, 1999) and they correspondingly increase

the strength of the prepost measure correlation coefficient (which is also a factor in the available power). Given that only one group is assigned to the intervention, power can be increased as a result of either a rise in the number of controls or demonstrating a large program effect. The estimated sample or effect sizes needed to achieve the desired statistical power can be determined relatively easily (Linden, Adams, & Roberts, 2004b).

MODEL SPECIFICATION ISSUES

The use of a linear regression model when the data are nonlinear (e.g., is a polynomial, log, or other such function) may lead to dramatically incorrect estimates of effect. This is akin to the specification problem in RD design (Trochim, 1990). Identifying and controlling for nonlinearity is explained in the appendix.

There are some subtle model specification issues that depend on the homogeneity of the treatment effect. This design is more sensible if the treatment effect is not heterogeneous and a function of group characteristics. For example, if the treatment success is dependent on group characteristics (e.g., the treatment is more successful with strong support from the provider) a single group that has a low treatment effect can obscure what would be a larger average treatment effect in a study with more treatment groups. One example of heterogeneity with respect to group characteristics would be a 50% rate of failure of implementation. Then we would see either zero or twice the average treatment effect, depending on which type of group was selected. This suggests avoiding this design unless group characteristics and structure are thought to have little effect on outcome.

EXTERNAL VALIDITY

Similar to the RCT or the RD design, the RPD design does not ensure external validity. As most programs are implemented in a particular setting with a given set of individuals under a specific set of circumstances, the results may be limited in their generalizability across persons, settings, treatments, and outcomes. Therefore, to increase the generalizability potential of the study outcomes, the intervention must be implemented in a group that is most similar to that of the control population (Linden, Adams, & Roberts, 2004a).

CONCLUSION

This article has introduced the RPD design as being a practical study design in health care. It is a simple yet effective model for comparing the prepost results of a single or multiple treatment groups to that of a control population. As a result, this design shows great potential in evaluating health care pilot programs or demonstration projects due to its relative ease of implementation and low cost of analysis. The RPD design has the most power to detect program effectiveness when the effect size effect of the intervention is substantial, when the prepost outcome measure correlation is high, and/or when the control population is sufficiently large.

APPENDIX
REGRESSION-POINT DISPLACEMENT MODEL SPECIFICATION

BASIC MODEL

The basic regression-point displacement (RPD) model is as follows (Trochim, 2001):

$$Y = \beta_0 + \beta_1 X_i + \beta_2 Z_i + e_i \qquad (1)$$

Where Y is the dependent or outcome variable, β_0 is the intercept, Z_i is the dichotomous assignment variable (1 = treatment, 0 = control), β_1 is the pretest coefficient, β_2 is the estimated treatment effect, and e_i is the random error term. The main treatment effect is easily identified by a statistically significant p value in the β_2 coefficient. Because the units of analysis are groups it is important to weight the regression to reflect the differing sample sizes.

ANALYZING MULTIPLE INTERVENTIONS

In situations where there is more than one intervention, Equations 1 can be expanded to include multiple dichotomous assignment variables Z_i to represent the various intervention arms. Using an example of three mutually exclusive interventions implemented with 3 groups, the additional independent variables will be as follows:

Z_1: 1 = intervention A, 0 = control
Z_2: 1 = intervention B, 0 = control
Z_3: 1 = intervention C, 0 = control

Accordingly, it can then be determined whether a treatment effect is evident between (a) each intervention level compared to the control group and (b) each level of intervention compared to each other. In this case there are several possible models that may be employed. If the interventions are similar in design and implementation the most powerful analysis will be the inclusion of a single indicator for all of the intervention groups. A more conservative approach would be to include one indicator per intervention group and jointly test for an intervention effect with an F-test. If the interventions are different in either design or implementation individual t tests on the intervention group indicators will allow each to stand on its own.

EXPANDED MODEL TO HANDLE NONLINEARITY

Equation 1 specifies only the existence of a main treatment effect. However, there may be a concern that the data are nonlinear and thus estimating a linear model may bias the results. Nonlinearity can occur if the variables have are not normally distributed, there are chance outliers, or there exists a floor and/or ceiling effect (Shadish, Cook, & Campbell, 2002; Trochim, 1984). When nonlinearity is suspected, higher order terms should be added to the equation. As with the regression discontinuity design, Trochim (1984, 1990) suggests overfitting the model by one order higher than what the model indicates on visual inspection. Therefore, if the data appear to be linear, squaring the term should suffice. Although this may result in overfitting the model, any nonsignificant terms can be removed from the analysis later resulting in a correctly specified and efficient model. A model with the additional terms (Shadish, Cook, & Campbell, 2002; Trochim 1984, 1990) is as follows:

$$Y = \beta_0 + \beta_1 Z_i + \beta_2 (X_i) + \beta_3 (X_i)^2 + \beta_4 (X_i)^3 + e_i \qquad (2)$$

where the terms are as defined in Equation 1, with the addition of higher order terms $(X_i)^2$ and $(X_i)^3$.

REFERENCES

Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.

Cunningham, P. J. (2006). Medicaid/SCHIP cuts and hospital emergency department use. *Health Affairs (Millwood)*, *25*, 237-247.

Jollis, J. G., Ancukiewicz, M., Delong, E. R., Pryor, D. B., Muhlbaier, L. H., & Mark, D. B. (1993). Discordance of databases designed for claims payment versus clinical information systems: Implications for outcomes research. *Annals of Internal Medicine*, *119*, 844-850.

Linden, A., Adams, J., & Roberts, N. (2003). An assessment of the total population approach for evaluating disease management program effectiveness. *Disease Management*, *6*, 93-102.

Linden, A., Adams, J., & Roberts, N. (2004a). Using an empirical method for establishing clinical outcome targets in disease management programs. *Disease Management*, *7*, 93-101.

Linden, A., Adams, J., & Roberts, N. (2004b). The generalizability of disease management program results: Getting from here to there. *Managed Care Interface*, *17*, 38-45.

Linden, A., & Roberts, N. (2005). A user's guide to the disease management literature: Recommendations for reporting and assessing program outcomes. *American Journal of Managed Care*, *11*, 81-90.

Linden, A., Adams, J., & Roberts, N. (2005). Evaluating disease management program effectiveness: An introduction to the regression-discontinuity design. *Journal of Evaluation in Clinical Practice*, *11*, XX-XXX.

Shadish, S. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.

Trochim, W. (1984). *Research design for program evaluation: The regression-discontinuity approach*. Beverly Hill, CA: Sage.

Trochim, W.M.K. (1990). The regression discontinuity design. In L. B. Sechrest, E. Perrin, & J. Bunker (Eds.), *Research methodology: Strengthening causal interpretations of nonexperimental data AHCPR Conference Proceedings, PHS 90-3545* (pp. 119-139). Rockville, MD: Agency for Health Care Policy and Research.

Trochim, W.M.K., & Campbell, D. T. (1996). *The regression point displacement design for evaluating community-based pilot programs and demonstration projects*. Unpublished manuscript. Retrieved October 20, 2005, from http://www.socialresearchmethods.net/research/RPD/RPD.pdf

Trochim, W., & Campbell, D. (1999). Design for community-based demonstration projects. In D. T. Campbell & M. J. Russo (Eds.), *Social experimentation* (pp. 313-337). Thousand Oaks, CA: Sage.

Trochim, W.M.K. (2001). *The research methods knowledge base* (2nd ed). Cincinnati, OH: Atomic Dog.