The American Journal of Drug and Alcohol Abuse
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597226

A Small Sample Randomized Clinical Trial Methodology Using N-of-1 Designs and Mixed Model Analysis

Ty A. Ridenour [a]; Deanne L. Hall [b]; James E. Bost [c]
[a] Center for Education and Drug Abuse Research, University of Pittsburgh, Pittsburgh, Pennsylvania, USA [b] Department of Pharmacy, University of Pittsburgh, Pittsburgh, Pennsylvania, USA [c] Center for Research and Health Care, University of Pittsburgh, Pittsburgh, Pennsylvania, USA

Online Publication Date: 01 July 2009

To cite this Article Ridenour, Ty A., Hall, Deanne L. and Bost, James E.(2009)'A Small Sample Randomized Clinical Trial Methodology Using N-of-1 Designs and Mixed Model Analysis',The American Journal of Drug and Alcohol Abuse,35:4,260 — 266
To link to this Article: DOI: 10.1080/00952990903005916
URL: http://dx.doi.org/10.1080/00952990903005916

The American Journal of Drug and Alcohol Abuse, 35:260–266, 2009

ISSN: 0095-2990 print/ 1097-9891 online
DOI: 10.1080/00952990903005916

informa
healthcare

A Small Sample Randomized Clinical Trial Methodology Using N-of-1 Designs and Mixed Model Analysis

Ty A. Ridenour, Ph.D., M.P.E.
Center for Education and Drug Abuse Research, University of Pittsburgh, Pittsburgh, Pennsylvania, USA

Deanne L. Hall, Pharm.D.
Department of Pharmacy, University of Pittsburgh, Pittsburgh, Pennsylvania, USA

James E. Bost, Ph.D.
Center for Research and Health Care, University of Pittsburgh, Pittsburgh, Pennsylvania, USA

***Background/Objectives*: To date, research on substance abuse prevention relied extensively on large sample randomized clinical trials to evaluate intervention programs. These designs are appropriate for certain types of randomized prevention trials (e.g., efficacy or effectiveness for broad populations) but are unfeasible for other prevention science scenarios (e.g., rare pathologies, pilot studies, or replication tests at specific locales). *Methods*: An alternative randomized clinical trial is described that relies on much smaller samples, less resources than the large sample designs, randomization, N-of-1 designs for the intervention group, and mixed model analysis. *Results*: This methodology is illustrated using a small sample prevention study, which demonstrates its statistical power, flexibility, and sophistication for experimental testing of prevention-oriented research questions. *Scientific Significance*: This methodology can be applied to many existing prevention datasets to facilitate secondary analyses of existing datasets as well as novel studies. It is hoped that such efforts will include further development of the small sample design in substance abuse prevention contexts.**

Keywords Mixed modeling, N-of-1, prevention, randomized clinical trials, small samples, substance abuse

INTRODUCTION

This article proposes and demonstrates small sample randomized clinical trial (RCT) designs for substance abuse (SA) prevention. SA prevention researchers nearly exclusively utilize large samples even though alternative, clinically-oriented, RCTs such as N-of-1 designs may be more advantageous for many investigations (1, 2). Large sample RCTs (LRCT) are appropriate for efficacy or effectiveness estimates for broad populations, but are inadequate or unfeasible for many other prevention science scenarios (e.g., replication studies for specific locales or subgroups). Herein, a design is proposed that (a) is used in other clinical research, (b) utilizes small samples, and (c) retains strengths of LRCT designs.

Next, a simplistic form of this design is demonstrated that could be applied to existing SA prevention datasets. This illustration uses a RCT of a program to reduce blood glucose in diabetic patients. Diabetes can be a consequence of SA (14.6% of the sample had SA) and reducing elevated blood glucose can prevent manifold harmful medical complications of diabetes, including death. SA increases risk for and exacerbates such medical complications, making control of blood glucose particularly important in diabetic patients with SA. The power of the design is further illustrated by replicating tests among only diabetic patients with SA.

LRCT STRENGTHS AND LIMITATIONS

Strengths of traditional LRCTs include (a) randomization into control and experimental groups, (b) detection of small intervention effects, and (c) the ability to address complex research questions using a range of statistical techniques. However, one limitation of LRCTs is ambiguity regarding to whom intervention outcomes apply, typically a metaphorical person who experiences the average outcome. It would be more informative to identify intervention effects (positive, null, or negative) that are specific to subgroups, a concept that is termed person-intervention interactions (1, 4). The current reliance on LRCT and universal prevention also has led to an acceptance (perhaps even expectation) of small efficacy of interventions. Moreover, post hoc insights about person-intervention interactions often cannot be tested (e.g., due to lacking data on a particular variable) without an additional resource-intensive LRCT. The

Address correspondence to Ty A. Ridenour, Center for Education and Drug Abuse Research, 3501 Terrace St., 711 Salk Hall, University of Pittsburgh, Pittsburgh, PA 15261, USA. E-mail: tar27@pitt.edu

SRCT design proposed for SA prevention research incorporates strengths of LRCTs and N-of-1 designs.

NEEDS FOR RCTS THAT USE SMALLER SAMPLES

Small sample RCT (SRCT) designs emphasize maximizing intervention impact. In contrast to statistical notions of efficacy and effectiveness, "impact" herein refers to the degree to which individuals respond to an intervention. Efficacy and effectiveness gauge an average effect size of a prevention program; impact refers to size of individuals' responses to the program.

A SRCT (e.g., $n < 80$ per study arm) could advance SA prevention in many ways. Consider recent progress of different areas of prevention. Universal program research has been considerable, but has generated null to small efficacy with a few exceptions (5). These outcomes might be expected based on the widely-accepted multifactorial, ecological, and liability-threshold models, which each postulate that manifold risk factors each vary in salience to bias individual development toward or away from SA (3). However, rather than addressing an individual's salient risk factors, universal interventions attempt to alter one to a few factors in all persons using the same "dosage" of a manualized program. Further complicating the intervention process (and potentially mitigating universal program efficacy, effectiveness, and impact) are manifold putative intervention moderators (3–7).

By tailoring intervention to address specific risk factors in those who experience them, prevention programs may have greater impacts. To illustrate, the *Good Behavior Game* program reduces SA risk specifically in disruptive students in poorly managed classes and is based on decades of SRCTs, using specific mechanisms of behavior change (e.g., student peer pressure, teacher-delivered reinforcement) (4). Studies to elucidate such mechanisms are not amenable to LRCTs because of the smaller populations and samples with specified risk factor(s). However, one strategy for these RCTs could be to test an existing program in a well-specified population that has salient levels of the risk factor(s) that the program targets. Other critical studies not amenable to LRCTs include experiments to (a) investigate prevention of rare pathologies, (b) identify the most impactful elements of programs, (c) pilot testing, and (d) refining existing programs. Optimally, these studies would be conducted during developmental phases of prevention programs.

Person-intervention interactions are implied in all SA prevention programs because they target specific risk factors, which vary in salience among people (3–6). SRCTs could elucidate and explicate such interactions starting with secondary analyses (e.g., whether universal program efficacies result from small impacts on many persons or large impacts on few persons is largely unknown, but testable using existing datasets). Person-intervention interaction studies require resolute objectivity from scientists, who may have to disclose that an intervention is iatrogenic for certain persons. Yet, such studies can greatly advance prevention and etiology (6). Dishion et al.'s (7) report of deviancy training between recipients of a group intervention that resulted in iatrogenic outcomes was cited over 600 times as of October 25, 2008.

N-OF-1 DESIGNS

Turning to N-of-1 designs, their limitations have largely precluded their use in prevention science and possibly led to naïveté among prevention researchers about their substantial strengths (1, 8, 9). Strengths of N-of-1 designs include (a) sequencing of study phases to isolate an intervention effect, (b) focus on within-person change, (c) emphasis on impact of an intervention, (d) utility for clinical settings and not merely research, and (e) ability to increase statistical power with sample size and with number of observations per participant (2). Collectively, these strengths make N-of-1 RCTs ideal for translational research such as prevention science (1).

A classic example N-of-1 RCT is the ABAB design, consisting of four contiguous phases each of which includes multiple observations of an outcome. During the first A phase, baseline data are collected (in place of a control group). The first B phase is an intervention period. Next, the A and B phases are repeated. Thus, the second A phase outcomes are expected to return to baseline levels and outcomes during both B phases are hypothesized to be better than during the A phases. Assuming this pattern in outcomes is observed, change in outcomes is concluded to be due to the intervention. The ABAB design can be a powerful way to test intervention effects because control and intervention observations are not only equalized in theory (cf, randomization); the observations are from the same persons.

The ABAB option of N-of-1 RCTs is designed for an intervention with effects that dissipate soon after it is withdrawn. One problem with the ABAB design is that many educational or psychological interventions cannot be withdrawn per se because they involve inculcating skills that putatively are not forgotten. Alternative, single A phase, designs exist to maintain the causal implications that can be drawn from N-of-1 studies. One example is multiple baselines designs, which consist of AB phases and randomizing persons to varying start points of the B phase (e.g., four study participants could begin the B phase at their randomly-assigned 7th, 4th, 6th, and 9th observation, respectively) (8). To the extent that outcomes improve in association with beginning the B phases, stronger causal implications can be drawn about an intervention.

Unfortunately, strengths of N-of-1 designs are outweighed by their *traditional* limits, the most decisive of which are statistical. The primary limitation of traditional N-of-1 RCTs is use of single cases (9, 10), which cannot be generalized. The emphasis among N-of-1 researchers on effect size led to reliance on visual inspection of outcomes plotted on figures to determine the impact of an intervention. Evidence demonstrates that visual inspection of such plots leads to erroneous conclusions about intervention impact (e.g., true impact can be obfuscated by

variability in outcomes) (11, 12). Serial dependency (correlations between within-person observations) can invalidate traditional N-of-1 designs and may occur for several reasons such as an individual's risk factors can impinge on multiple observations, one observation may influence subsequent observations, cyclical behavior patterns (e.g., increased alcohol consumption during weekends compared to other days), and practice effects in multiple evaluations (13, 14).

Solutions to each of these N-of-1 RCT limitations arise from using multiple participants. Results can be generalized. Intervention impact can be examined using both visual and statistical tests, including confidence intervals (11–15). Serial dependency can be statistically accounted for and scientifically informative (later). Randomizing persons into intervention (N-of-1 phases) vs. control conditions can further improve: generalization, accounting for serial dependency, testing intervention effects, and impact size estimates. Moreover, the analytic techniques can include complex statistical modeling and subgroup comparisons.

PROPOSED SRCT: DRAWING FROM THE STRENGTHS OF N-OF-1 DESIGNS, LRCT, AND MIXED MODELS

SRCT analysis can consist of within-person mixed models, which resolve limitations of traditional N-of-1 analysis and facilitate testing of complex research questions (13, 14). Compared to other longitudinal within-person analytical techniques (time series, P-technique, MANOVA or MANCOVA, meta-analysis of case studies), mixed models offer (a) statistical and logistical parsimony, (b) flexibility and elegance, (c) statistics that are used increasingly in prevention and clinical research, and (d) straightforward clinical interpretation (4, 10, 13).

Mixed models are typically used in prevention to analyze multilevel clusters of persons (e.g., students, analyzed at level 1, are clustered within classrooms at level 2, which in turn are clustered within schools at level 3) (3). In contrast for SRCTs, individual differences are analyzed at level 2, and within person outcomes, are analyzed at level 1 (10, 13). Each level of mixed models consists of a regression model and can include such variables as covariates or interaction terms. In fact, certain intervention effects are tested using an interaction term (e.g., between study arm and time) in which it hypothesized that outcome slopes differ between study arms (13, 14). For the most basic SRCT (13), the level 1 or within-subjects model is:

$$Y_{ij} = b_{0i} + b_{1i} + e_{ij}, \qquad [1]$$

where Y_{ij} is an observed outcome for participant i at occasion j; b_{0i} equals the initial level of the outcome for participant i; b_{1i} equals the average change in outcome (or slope) for participant i over a specified time period (e.g., per week or month); and e_{ij} equals random error (13).

The level 2 model (differences between participants) consists of two equations:

$$b_{0i} = \beta_0 + u_{0i} \quad \text{and} \quad b_{1i} = \beta_1 + u_{1i}, \qquad [2]$$

where b_{0i}, the initial outcome of individual i, is modeled as a function of the population average initial level of the outcome (β_0) plus the individual's deviation from the population average (u_{0i}); and b_{1i}, the average change over time of individual i's outcome, is modeled as a function of the population average slope (β_1) plus the individual's deviation in slope from the population average (u_{1i}). This set of models compose the unconditional (i.e., no predictors except time) linear regression mixed model. Substituting the terms from the level 2 model into the level 1 model, the multilevel analysis can be summarized as a single model:

$$Y_{ij} = \beta_0 + u_{0i} + \beta_1 + u_{1i} + e_{ij}. \qquad [3]$$

Two augmentations to model 3 were described earlier in non-statistical terms and can be useful for SRCTs. One is the hypothesis that an intervention will improve outcomes faster than a control group (i.e., greater slope in outcomes) (14). This suggests the following augmentation:

$$Y_{ij} = \beta_0 + \beta_{01}\text{Intx}_i + u_{0i} + \beta_1 + \beta_{11}\text{Intx}_i + u_{1i}, \qquad [4]$$

where: Intx = 0 for controls, and 1 for intervention participants. Hence, β_{01} and β_{11} represent the difference between intervention and control groups in terms of average level and slope over time, respectively; these terms provide tests of intervention outcome (efficacy, effectiveness, impact).

Noortgate et al. (10) provide the second augmentation. They demonstrated how to use mixed models for an ABAB design and provided sample PROC MIXED SAS code. The model for their most detailed analysis, with estimates specific to each study phase, appears below. ABAB phases are organized into two dummy-coded phase (A vs. B) terms, which are equivalent to the Intx terms of model 4, and two dummy-coded block (1st AB vs. 2nd AB) terms:

$$Y_{ij} = \beta_0 + \beta_{0\text{B}i} + \beta_{0\text{P}i} + u_{0i} + \beta_1 + \beta_{1\text{B}i} + \beta_{1\text{P}i} + u_{1i}, \qquad [5]$$

where: $\beta_{0\text{B}i}$ and $\beta_{0\text{P}i}$ are deviations from population mean due to study block and study phase, respectively; $\beta_{1\text{B}i}$ and $\beta_{1\text{P}i}$ are deviations from population slope due to study block and study phase, respectively. Accordingly, this model provides mean and slope differences between phases and blocks as well as within and between subject variances. Again, traditional regression analyses variables can be added, as can covariance structures to account for serial dependency.

Earlier, the limitation of serial dependency was described. SAS PROC MIXED provides many covariance structures to

test and account for serial dependency, according to the structure that best fits a dataset (14,15). Evidence demonstrates that (a) results of repeated measures analysis are robust if the selected covariance structure resembles the true structure, (b) parameter estimates of fixed effects rarely change with the choice of covariance structure, but (c) erroneous selection of covariance structure can effect the precision of these estimates (14, 15). Standard tests can be used to determine the covariance structure that best fits a particular dataset (14, 15).

A final consideration for SRCT analyses is the Kenward–Roger adjusted F-test. Compared to default F-tests for parameters, the Kenward–Roger test is more precise and better controls for Type I error in analyses of small samples (15, 17).

ILLUSTRATION OF SRCT FOR PREVENTION

Existing prevention datasets rarely include phases that resemble N-of-1 designs (e.g., baselines typically consist of one wave of data). Even so, simple forms of the SRCT proposed here can be applied to existing prevention datasets (e.g., to investigate subgroups). The example dataset used here illustrates this point because it (a) is from a traditional RCT prevention study, (b) contains the fewest waves of data required for mixed models (three) thereby providing the least statistical power that can be drawn from multiple observations, (c) uses a small sample (which also constrains power), and (d) demonstrates the power that within-person mixed models can provide to test person-intervention interactions and testing of subgroups.

Background

The study tested an intervention to improve adherence to and outcomes of diabetes-related pharmacotherapy and lifestyle changes (16). Nearly 10% of U.S. citizens older than 20 years have diabetes mellitus. Although prevention of medical complications from diabetes may at first appear unrelated to SA, diabetes can be a consequence of SA, and even low levels of substance use can increase risk for and exacerbate the medical complications of diabetes such as microvascular (nephropathy, retinopathy, and neuropathy) and cardiovascular disease (17, 18). Accordingly, one form of harm reduction is prevention of medical complications among diabetic patients with SA. A subgroup of this sample had concomitant SA related to tobacco (n = 11) or alcohol (n = 1). The intervention consisted of pharmacist delivered education and monitoring to bolster adherence to medicinal regimen (19). Clinical observations led to the hypothesis that patients with uncontrolled blood glucose are more responsive to pharmacist advice about improving glucose levels. In addition to the main effect, it was hypothesized that the intervention would reduce A1C more in diabetes patients with baseline uncontrolled glucose (A1C > 8.0%) compared to other diabetes patients.

Methods

Controls received care as usual. Outcomes were percent of glycocylated hemoglobin or A1C, which measures blood glucose level. Pre-study power analyses suggested N = 150 (75 per study arm) was needed to detect the hypothesized effect of intervention using traditional statistical analyses. This sample size was considered easily attainable because the annual number diabetic of patients treated at the site is several times greater than 150. As it turned out, multiple diabetes studies occurred at the site concomitantly with this study, resulting in the availability of only 82 patients (37 intervention, 45 controls) for this study.

Analyses

Originally, 12 month outcomes were tested with ANOVA and χ^2. Fortunately, 6 month outcomes also were collected, permitting secondary analysis using the earlier SRCT analytical approach with PROC MIXED and the ML estimator. Competing models of A1C outcomes (compared to an unconditional model) consisted of either of these fixed effects: baseline A1C, study arm, interaction between baseline A1C and study arm, or a model including all of these predictors. Three covariance structures were tested to account for serial dependence: compound symmetry (correlations between observations are equivalent regardless of lag time) and the Toeplitz and autoregressive (lag 1) structures, both of which assume that within-person observations correlate more with closer temporal proximity (lag of 1 observations correlate the most, lag of 2 observations correlate less, etc.) (13–15). The autoregressive (lag 1) structure best accounted for serial dependency based on statistical fit and parsimony.

Results

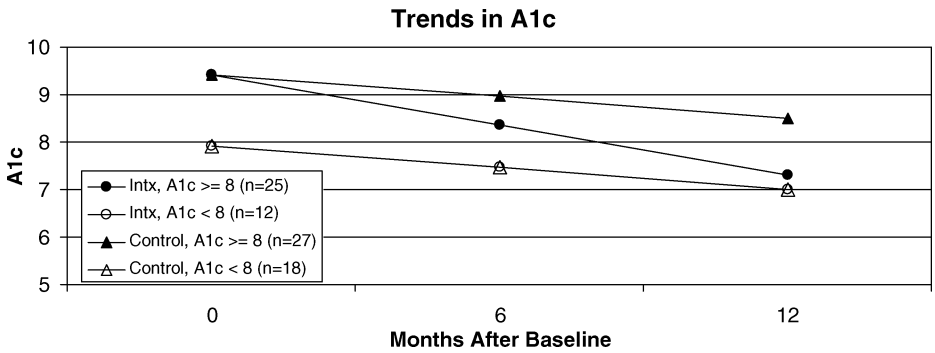
Twelve month outcomes were consistent with hypotheses. Mean A1C decreases of 1.1% vs. .6% occurred in intervention and control participants, respectively. An ANOVA testing the interaction resulted in F = 1.18 (p = .33). The clinical goal of A1C <= 7.0% occurred in 41% of intervention participants and 25% of controls. However, none of these results generated p-values < .05 because of the small sample.

Compared to an unconditional model, adding baseline A1C level as a predictor improved the model fit (Table 1). Of the other predictors tested, only the intervention × baseline A1C interaction further improved model fit with statistical robustness. Notably, statistical detection of the interaction occurred using four small subgroups (n of 25, 12, 27, and 18) and generated p-values that exceeded .05. In statistical terms, the best fitting model was:

$$Y_{ij} = 7.93 - .46(\text{wave}) + 1.50(\text{baseline A1C level}) - .61(\text{intervention} \times \text{baseline A1C level; a slope}).$$

The Table 1 figure illustrates within-person change of the four subgroups using the REML estimator (13–15). Pharmacist

TABLE 1
Tests of whether pharmacist intervention improves blood glucose in diabetics



	Unconditional Model	Baseline A1C	Intervention × Baseline A1C
LR χ^2	839.5	814.8	810.0
LR χ^2 difference, df	N/A	24.7, 1*	4.8, 1df*
AIC	847.5	824.8	822.4
BIC	857.1	836.8	836.4

Note: N = 82. LR χ^2 is the likelihood-ratio chi-square; differences between nested models are testable using the chi-square distribution. LR χ^2 difference here presents results of a χ^2 test compared to the preceding model (*indicates $p < .05$). AIC is the Akaike's Information Criterion and BIC is the Bayesian Information Criterion; better fit to the data is indicated by smaller values for both criteria.

intervention impacted only diabetic patients with baseline A1C > 8.0%; however, for those patients, the impact was considerable. On average, their A1C dropped to nearly the goal of A1C <= 7.0% in contrast to control participants whose A1C dropped on average to about 8.5%. The clinical implication is clear; diabetic patients with A1C > 8.0% could benefit from the intervention to better manage blood glucose.

To further demonstrate how SRCT can test person-intervention interactions using subsamples of existing datasets, analyses were repeated with the 12 participants with SA. Half were serendipitously randomized into each study arm. Compared to the unconditional model, the best fitting model from Table 1 improved fit to the data with a likelihood-ratio $\chi^2 = 4.9$, $p < .10$ (a trend). Inspection of subgroup A1C slopes revealed a potentially greater intervention effect for those with SA. Slopes increased slightly in all SA participants but those with baseline A1C > 8.0% who received the intervention - their average outcome was A1C = 7.5%.

DISCUSSION

This manuscript presented needs, rationale, and a methodology for SRCTs. Certain strengths of this design were demonstrated, including its statistical power. Although the example did not feature important aspects of SRCTs (e.g., A and B phases), it did illustrate how the SRCT analysis could be applied to many existing SA prevention datasets. It is hoped that more frequent mixed model analysis of N-of-1 designs could clarify impacts of existing prevention programs for specific types of persons and lead to further development of N-of-1 designs for SA prevention. Other methodologists have progressed this technique, such as meta-analysis of small sample and N = 1 studies, SAS code for multiple phase within-subject studies, and detailed analysis of individual study participants (10, 11, 13, 14). However, little such progress has occurred within SA prevention. The example used herein advanced SA SRCTs by using a behavioral prevention among persons with SA, albeit to prevent secondary outcomes.

Awareness of the potential limitations of this SRCT may improve these types of studies by accounting for them. First, although one advantage of SRCTs is that use of well-specified samples provides clarity about to whom outcomes apply, replication is essential. Whereas with LRCTs it can be unclear to whom an efficacy estimate applies because the sample is so diverse, with SRCTs it can be unclear how well an intervention effect generalizes to other populations. It also is possible that intervention impact may be partly due to site-specific factors such as a talented deliverer of a prevention program. Generalization between subgroups and intervention deliverers are only two of the factors that ought to be tested prior to conducting LRCTs. Fortunately, SRCT replication is a reasonable and efficient approach compared to LRCTs.

A second limitation is that burden on participants in a SRCT is greater than a LRCT; greater number of data collection waves may impede retention of study participants. In studies that address needs (e.g., risk factors) of patients, clients, or children,

study compliance is more likely. Also, in many contexts it is possible to collect multiple waves of data and multiple phases without increasing participant burden. For example, the earlier illustrative study of diabetes patients is being evolved into an ABA design, using patient medical records of A1C before and after the intervention (the two A phases) as well as to add data points during the B phase. Using data from study participant records for A phases may be feasible in a range of other settings from education to criminal justice.

OPPORTUNITIES AFFORDED BY SRCTS

Certain opportunities are uniquely availed with SRCT methods. Of these, testing for person-intervention interactions may provide the greatest short-term benefits. Many SA prevention RCTs have been conducted, permitting re-analyses to test program impact in subgroups. Based on the complexity of SA etiology, it is reasonable to hypothesize that these types of studies could lead to important advances, especially for selective/indicated prevention.

The potential importance of the Kenward–Roger adjusted test (15, 17) in SRCT analyses can be illustrated with the present results. Using the default SAS statistical test, the best fitting model in Table 1 generated the same parameter estimates but a $p < .001$ (compared to the $p < .05$ reported earlier). Avoiding Type I errors is important for identifying intervention strategies that offer nil or minimal impact overall as well as in subgroups. Power analyses also are needed for these types of statistical techniques.

An important aspect of SRCTs is their adaptability to clinical settings. Blanco et al. (22) demonstrated that large proportions of the SA population would be excluded from SA treatment RCTs because of study exclusion criteria. Such discontinuity between SA prevention RCT samples and the targeted population also may occur for reasons such as exclusion criteria, family attrition in family-oriented prevention studies, or differences between school personnel who are vs. are not willing to participate. SRCTs offer the techniques for persons in applied settings to conduct replication tests of prevention impact.

Importantly, N-of-1 designs historically have been short-term studies. With some exceptions, prevention science lacks studies of short-term impact of prevention programs. Perhaps participant attitudes, behavior, skills, or knowledge improve very soon after certain prevention sessions, but then return to baseline levels. Such an intervention effect would not be detected using standard prevention methodologies but could be investigated using SRCTs.

To reiterate, while the most commonly utilized LRCTs provide SA prevention scientists with important tools, these designs are inadequate for many critical prevention science research questions. Variant forms of the SRCT could provide the apparatus to advance important areas of prevention that to date have been slow to evolve. Strengths of SRCTs were demonstrated: flexibility for increasing statistical power, testing complex intervention research questions, direct clinical implications, clinically replicable RCTs, ability to detect person-intervention interactions, and estimation of intervention impact. All of these aspects and additional strengths of the SRCT meet important ongoing scientific and pragmatic needs in SA prevention science.

ACKNOWLEDGMENT

This study was funded by NIDA (P50-05605).

Declaration of Interest

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the article.

REFERENCES

1. Guyatt GH, Haynes RB, Jaeschke RZ, Cook DJ, Green L, Naylor CD, Wilson MC, Richardson WS. Users' Guides to the Medical Literature: XXV. Evidence-Based Medicine: Principles for Applying the Users' Guides to Patient Care. JAMA 2000; (284):1290–1296.
2. Venter A, Maxwell SE. Maximizing power in randomized designs when N is small. In Hoyle RH. (ed.) Statistical Strategies for Small Sample Research. Thousand Oaks, CA: Sage, 1999; pp. 31–58.
3. Ridenour TA, Stormshak EA. Introduction and rationale for individualized substance abuse prevention from an ontogenetic prevention. Am J Drug Alc Abuse 2009;(35):1–3.
4. Kellam SG, Brown CH, Poduska J, Ialongo N, Wang W, Toyinbo P, Petras H, Ford C, Windham A, Wilcox HC. Effects of a universal classroom behavior management program in first and second grades on young adult behavioral, psychiatric and social outcomes. Drug Alc Depend 2008; (95S):S5–S28.
5. Derzon JH, Sale E, Spring JF, Brounstein P. Estimating intervention effectiveness: Synthetic projection of field evaluation results. J Primary Prevent 2005; (26):321–343.
6. Lilienfeld SO. Psychological treatments that cause harm. Perspect Psychol Sci 2007; (2):53–70.
7. Dishion TJ, McCord J, Poulin F. When interventions harm: Peer groups and problem behaviors. Am Psychol 1999;(54):755–764.
8. Ferron J, Sentovich C. Statistical power of randomization tests used with multiple-baseline designs. J Exper Ed 2002; (70):165–178.
9. Onghena P, Edgington ES. Customization of pain treatments: Single-case design and analysis. Clin J Pain 2005; (21):56–68.
10. Noortgate WVD, Onghena P. Combining single-case experimental data using hierarchical linear models. School Psychol Quart 2003; (18):325–346.
11. Blouin DC, Riopelle AJ. On confidence intervals for within-subjects designs. Psychol Meth 2005; (10):397–412.
12. Franklin RD, Gorman BS, Beasley TM, Allison DB. Graphical display and visual analysis. In Franklin RD, Allison DB, Gorman BS, (eds.) Design and Analysis of Single-Case Research. Mahwah, NJ: Erlbaum, 1997.
13. Hedeker D, Gibbons RD. Longitudinal Data Analysis. Hoboken, NJ: Wiley & Sons, 2006.
14. Singer JD, Willett JB. Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence. New York: Oxford University Press, 2003.
15. Littell RC, Milliken GA, Stroup WW, Wolfinger RD, Schabenberger O. SAS for mixed models (2nd ed.). Cary, NC: SAS Press, 2006.
16. Hall DL, Dang Q, Fischer G, Kapoor WN. Pharmaceutical intensive treatment of Type 2 Diabetes (PITT-DM Pilot). In review.

17. Sikkink J, Fleming M. Adverse health effects and medical complications of alcohol, nicotine and drug use. In Fleming MF, Barry KL, (eds.). Addictive Disorders. St. Louis: Mosby Year Book, 1992; pp. 145–168.
18. Spencer EA, Pirie KL, Sevens RJ, Beral V, Brown A, Liu B, Green J, Reeves GK. Diabetes and modifiable risk factors for cardiovascular disease: The prospective Million Women Study. Eur J Epidemiol 2008;(23):793–799.
19. Hall DL, Kapoor WN, Fischer G, Fevrier LL, Simak DM. Pharmaceutical Intensive Treatment of Type 2 Diabetes Mellitus (PITT-DM Pilot). Pharmacotherapy 2006; (26):e112.
20. Blanco C, Olfson M, Okuda M, Nunes EV, L SM, Hasin DS. Generalizability of clinical trials for alcohol dependence to community samples. Drug Alc Depend 2008; (98):123–128.