

This is an Open Access article licensed under the terms of the Creative Commons Attribution-NonCommercial 3.0 Unported license (CC BY-NC) (www.karger.com/OA-license), applicable to the online version of the article only. Distribution permitted for non-commercial purposes only.

Letter to the Editor

Concerning Sichieri R, Cunha DB: Obes Facts 2014;7:221–232. The Assertion that Controlling for Baseline (Pre-Randomization) Covariates in Randomized Controlled Trials Leads to Bias Is False

Peng Li^a Andrew W. Brown^a John A. Dawson^a Kathryn A. Kaiser^a
Michelle M. Bohan Brown^b Scott W. Keith^c J. Michael Oakes^d
David B. Allison^a

^aOffice of Energetics and Nutrition Obesity Research Center, University of Alabama at Birmingham, Birmingham, AL, USA; ^bDepartment of Food, Nutrition, and Packaging Sciences, Clemson University, Clemson, SC, USA; ^cDivision of Biostatistics, Department of Pharmacology and Experimental Therapeutics, Sidney Kimmel Medical College, Thomas Jefferson University, Philadelphia, PA, USA; ^dDivision of Epidemiology and Minnesota Population Center, University of Minnesota, Minneapolis, MN, USA

We read with interest the article ‘Unbalanced Baseline in School-Based Interventions to Prevent Obesity: Adjustment Can Lead to Bias – a Systematic Review’ [1], hereafter ‘*the article*’. We agree with the authors that more rigor is needed in research on obesity treatment and prevention, and in the design, analysis, and reporting of cluster randomized controlled trials (cRCTs) [2], also called group randomized trials.

Unfortunately, rather than offering clarifying information, *the article* is based on incorrect statistical reasoning and inaccurate statements about what past publications have shown. The fundamental conclusion as stated in its title and elsewhere in *the article* is incorrect. For example, the statement ‘Although adjusting for the baseline values of parameters (*sic, variables – Li et al.*) that are highly influenced by baseline values is a standard procedure, this approach can bias the results ...’ is simply untrue. Such erroneous conclusions could lead researchers to avoid legitimate power-enhancing analytic methods, and should be retracted.

Adjusting for pre-randomization covariates in randomized trials does not introduce bias nor invalidate significance tests. This is known from statistical principles and requires neither simulation nor meta-analyses. By definition and design, in randomized experiments pre-randomization covariates are independent of treatment assignment, with the exception of chance deviations which are accommodated in the calculation of frequentist significance tests and their associated p values. If the outcome variable (Y) is measured pre-randomization (Y₀) and at the end of the study (Y₁), then using either Y₁ or Y₀ – Y₁ as outcomes and either

David B. Allison, Ph.D.
Office of Energetics and Nutrition Obesity Research Center
University of Alabama at Birmingham, Ryals Building, Room 140J
1665 University Boulevard, Birmingham, Alabama 35294, USA
dallison@uab.edu

controlling or not controlling for Y_0 as a covariate are legitimate analyses [3], among other analytical choices. The inclusion of Y_0 as a covariate, though, will often increase statistical power for a given effect size when the null hypothesis is false if appropriate assumptions are met [4, 5]. This is true for both cRCTs and ordinary RCTs [6].

The article's erroneous conclusion is based on two fallacious lines of reasoning. First, *the article* miscites past literature. *The article* states, 'A computer simulation study which compared the biases in the estimated treatment effect, with and without adjusting for measurement error at baseline and for different levels of baseline imbalance, concluded that adjusting for baseline leads to bias, especially when sample sizes are small.' However, the cited paper [7] found no evidence of bias in estimation of treatment effects when controlling for covariates with ordinary least squares methods (again, as knowable from statistical principles). Rather, that paper found bias due to a particular measurement error correction – a different matter entirely. *The article* also states that papers and books 'have called attention to the controversy about whether baseline measurements should be adjusted for in this context' and cites six references in support. The cited paper [7] is one of those, but we note that it states 'Controversy exists in the literature about whether baseline measurement error should be adjusted for in this context'. Furthermore, the cited references are definitely not addressing only 'this context' of cRCTs. One of the books they cite makes a statement contrary to *the article's* claim: 'In general, the analysis of longitudinal data from a randomized trial is the only setting where we recommend adjustment for baseline through analysis of covariance' [8]. These six cited references address the context of controlling for covariates in observational (nonrandomized) studies, whereas *the article* is explicitly about cRCTs.

The second erroneous line of reasoning *the article* offers involves their meta-analysis. *The article* compared articles which did and which did not adjust for baseline values of the outcome and found that, on average, results regarding treatment effects differed. Even if we take this meta-analytic finding as correct at face value, it has no bearing on bias of treatment effect estimates. *The article's* meta-analysis is of cRCTs, but its evaluation of an association between analytic procedure and results across cRCTs constitutes an observational analysis. The meta-analyzed cRCTs that did or did not employ covariate adjustment might differ in many ways which could account for the observed difference in results. If the difference in treatment point estimates (not their variances) *within* each cRCT varied significantly as a function of whether a pre-randomization covariate was used, that would be a curious thing indeed and hard to explain, but that is not what is presented in *the article*.

Another serious error beyond those concerning baseline adjustment concerns sample sizes. On page 222 *the article* states that a sample size of 1,000 is needed to detect meaningful changes in BMI. For a simple pretest-posttest cRCT with m persons nested within g groups across c experimental conditions, there will be $c \times g \times m$ total persons, or sample size N . While the impact of clustering, measured by the intraclass correlation coefficient, on the design effect can be profound, the key issue for statistical power in a cRCT is the degrees of freedom (df) for the between-group variance, sometimes written τ . This is the number of conditions multiplied by the number of groups minus one, or $df = c \times (g - 1)$ [9]. This is why cRCTs need a large number of groups, not necessarily a large number of persons. For example, *the article* cites a paper [10] which was a cRCT with 454 children nested in 7 schools in each of 2 experimental conditions. Inference for the primary analysis was based not on $df \approx 454$, but $df = 12$, as demanded by statistical theory. The presentation of study sample sizes listed in table 2 of the article is therefore misleading. Note that a cRCT with one group per condition and 500 persons per group does in fact have 1,000 subjects. But it has 0 df for assessing the treatment effect. *The article* does not adequately recognize the hierarchical nature of cRCTs.

In sum, *the article's* [1] conclusions are at odds with established statistical principles and based on erroneous reasoning and interpretations of prior literature. *The article's* conclu-

Li et al.: Concerning Sichieri R, Cunha DB: Obes Facts 2014;7:221–232. The Assertion that Controlling for Baseline (Pre-Randomization) Covariates in Randomized Controlled Trials Leads to Bias Is False

sions are not only wrong, but may inappropriately dissuade readers from powerful analytic choices and lead other readers to incorrectly conclude that published cRCTs which have used covariate adjustment are invalid or that large numbers of persons, not groups, are essential. Given the above, we believe *the article* should be retracted.

Disclosure Statement

Supported in part by NIH grants R25HL124208, P30DK056336, and T32HL072757. The opinions expressed are those of the authors and do not necessarily represent those of the NIH or any other organization.

References

- 1 Sichieri R, Cunha DB: Unbalanced baseline in school-based interventions to prevent obesity: adjustment can lead to bias – a systematic review. *Obes Facts* 2014;7:221–232.
- 2 Crespi CM, Maxwell AE, Wu S: Cluster randomized trials of cancer screening interventions: are appropriate statistical methods being used? *Contemp Clin Trials* 2011;32:477–484.
- 3 Huck SW, McLean RA: Using a repeated measures ANOVA to analyze the data from a pretest-posttest design: a potentially confusing task. *Psychol Bull* 1975;82(4):511–518.
- 4 Allison DB: When is it worth measuring a covariate in a randomized clinical trial? *J Consult Clin Psychol* 1995; 63:339–343.
- 5 Lazar AA, Zerbe GO: Solutions for determining the significance region using the Johnson-Neyman type procedure in generalized linear (mixed) models. *J Educ Behav Stat* 2011;36:699–719.
- 6 Moerbeek M: Power and money in cluster randomized trials: when is it worth measuring a covariate? *Stat Med* 2006;25:2607–2617.
- 7 Chan SF, Macaskill P, Irwig L, Walter SD: Adjustment for baseline measurement error in randomized controlled trials induces bias. *Control Clin Trials* 2004;25:408–416.
- 8 Fitzmaurice GM, Laird NM, Ware JH: *Applied Longitudinal Analysis*. Hoboken, Wiley, 2011.
- 9 Hannan PJ: Experimental social epidemiology: controlled community trials; in Oakes JM, Kaufman JS (eds): *Methods in Social Epidemiology*. San Francisco, Jossey-Bass / Wiley, 2006, pp 335–364.
- 10 Story M, Hannan PJ, Fulkerson JA, Rock BH, Smyth M, Arcan C, Himes JH: Bright Start: description and main outcomes from a group-randomized obesity prevention trial in American Indian children. *Obesity* (Silver Spring) 2012;20:2241–2249.