Obesity Facts
The European Journal of Obesity

Obes Facts 2014;7:221–232
DOI: 10.1159/000363438
Received: September 10, 2013
Accepted: November 27, 2013
Published online: June 28, 2014


**Original Article**

# Unbalanced Baseline in School-Based Interventions to Prevent Obesity: Adjustment Can Lead to Bias – a Systematic Review

Rosely Sichieri Diana Barbosa Cunha

Department of Epidemiology, Institute of Social Medicine, State University of Rio de Janeiro, Rio de Janeiro, Brazil

**Key Words**
School · Intervention · Obesity · Baseline adjustment · Randomization

## Abstract
***Background/Aims:*** Cluster designs favor unbalanced baseline measures. The aim of the present study was to determine the frequency of unbalanced baseline BMI on school-based randomized controlled trials (RCT) aimed at obesity reduction and to evaluate the analysis strategies. We hypothesized that the adjustment of unbalanced baseline measures may explain the great discrepancy among studies. ***Methods:*** The source of data was the Medline database content from January 1995 until May 2012. Our search strategy combined key words related to school-based interventions with such related to weight and was not limited by language. The participants' ages were restricted to 6-18 years. ***Results:*** We identified 146 school-based studies on obesity prevention (or overweight or excessive weight change). Of the 146 studies, 36 were retained for the analysis after excluding reviews, feasibility studies, other outcomes, and repeated publications. 13 (35%) of the reviewed studies had statistically significant ($p < 0.05$) unbalanced measures of BMI at baseline. 11 studies with BMI balanced at baseline adjusted for the baseline BMI, whereas no baseline adjustment was applied to the 5 unbalanced studies. ***Conclusion:*** Adjustment for the baseline BMI is frequently done in cluster randomized studies, and there is no standardization for this procedure. Thus, procedures that disentangle the effects of group, time and changes in time, such as mixed effects models, should be used as standard methods in school-based studies on the prevention of weight gain.


Rosely Sichieri, MD, PhD
Department of Epidemiology, Institute of Social Medicine, State University of Rio de Janeiro
Rua São Francisco Xavier, 524,7° andar, Bloco E.
Cep 20550-012, Rio de Janeiro, RJ (Brazil)
rosely.sichieri@gmail.com

*What is already known:* To date, interventions have been inconsistent in improving the BMI or body composition of children and adolescents. Uncertainty within the literature published thus far may be due to the heterogeneity of study populations and unrealistic expectations concerning the change in body mass.

*What this study adds:* Although randomized studies on average must be balanced, a greater percentage of cluster randomized studies on obesity showed statistically significant imbalance for baseline body mass. Adjusting for the baseline body mass is a wrong procedure in this scenario that increases the uncertainty. Also, few cluster randomized studies regarding obesity included around 1,000 participants – a sample size required due to the small effect on body mass that could be expected.

## Introduction

School-based interventions for obesity prevention have been conducted since the publication of two major studies on this topic in 1999: the Pathways Study conducted among American Indian school children [1] and the Planet Health Study conducted among students from Boston, MA, USA [2]. Both studies used a multicomponent intervention, had a high prevalence of obesity among the subjects, and disposed combination methods to better define obesity. Therefore, these two studies – conducted in populations with a high prevalence of obesity and using the state-of-the-art definition of obesity – were very likely to observe positive changes in obesity. However, neither study found an overall reduction in the prevalence of obesity. In the Planet Health Study, a statistically significant reduction in the prevalence of obesity was observed among girls only (decreasing from 23.6% to 20.3%), but the change was very small. Ever since these two studies were published, many more have been conducted in the USA and other countries; although many reviews have been published on the topic, the findings are still considered inconclusive, as indicated by Khambalia et al. [3]. These authors combined the findings of eight reviews, three meta-analyses, and five systematic reviews of school-based programs to prevent and control obesity and concluded that there was limited evidence to serve as a basis for recommendations on this matter. Methodological issues, such as inclusion criteria and outcome assessments, explain some of the discrepancies in these findings [4].

The heterogeneity of participants in cluster randomized trials is a potential problem in many fields of research, but it is particularly relevant in obesity studies because their outcome is almost always measured as weight or BMI change from baseline. Papers [5–8] and books [9, 10] have called attention to the controversy about whether baseline measurements should be adjusted for in this context. 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 [11].

The present study explores the imbalance of baseline groups and related methodological issues as another possible explanation for these discrepancies. This topic has not been considered in meta-analyses before, even in those accounting for the quality of the papers included. Unbalanced data at baseline in school-based studies are due to underestimated sample sizes and the cluster design, the latter because schools, not individual children, are randomized. In most studies, sample size calculations were based on a change of approximately 1 BMI unit, which is too large for most primary prevention trials. Thus, this analysis focuses on the evaluation of classes or schools for which unbalanced data in the comparison groups may represent an important source of bias.

Obes Facts 2014;7:221–232
DOI: 10.1159/000363438



The aim was to investigate the number of published school-based obesity intervention studies that used groups that were unbalanced for BMI at baseline and to study their approaches for handling the imbalance. This analysis may help researchers to better understand the uncertainty in the obesity intervention literature caused by clustering and improper analysis.

## Material and Methods

*Types of Studies Considered in This Review*

All randomized school-based intervention studies that focus on reducing excessive weight gain were included. Randomization of schools was accepted, whereas randomization of individuals was not.

*Participants*

The participants might be of either sex; and participants aged 6–18 years were included.

*Types of Intervention*

Studies dealing with intervention in terms of dietary advice for students intended to reduce weight were included in this review. Furthermore, studies that compare the effects of dietary advice versus no dietary advice or dietary advice versus physical activity advice were included.

*Types of Outcome Measures*

As the main outcome, we chose changes in weight over time or related measures, e.g., weight gain, overweight, obesity, BMI, or BMI z-score.

Secondary outcomes were changes in food consumption and physical activity.

*Search Strategy for Identification of Studies*

Medline was searched to identify relevant literature. There were no language restrictions for search terms or trial inclusion. The search strategy combined 'intervention at school' or 'school-based', and 'randomized' or 'clustered' with key words related to weight ('obesity', 'weight', 'body mass index', 'weight gain' or 'overweight'). All articles published between January 1995 and May 2012 were regarded as eligible. In 1995, the first trial on the prevention of cardiovascular disease among children using a school-based design was published, and obesity-related, school-based studies have been appearing since then. The search started on May 25, 2012, and updates were included through June 8, 2012.

*Review Methods*

The papers were reviewed by the two authors (RS and DBC) independently. Relevant studies were determined by the initial search of electronic databases and subsequent screening by the lead reviewer (RS) and a double-check by the co-reviewer (DBC). During this initial screening, articles could be rejected if the reviewer inferred from the title and/or the abstract that it did not meet the inclusion criteria.

Both reviewers independently collected data from each study using a data extraction form. This form included authors, country and year of publication, number of schools randomized, sample characteristics (size and age, baseline data, and the main findings with and without adjustment. The balance of baseline measures was also investigated, and studies with unbalanced baselines were defined by a statistically significant degree of imbalance ($p \leq 0.05$) for baseline BMI or related measures, such as bioimpedance or the prevalence of overweight and obesity.

## Results

Of 257 papers taken into account, 146 were related to the review subject according to their titles. These 146 works included 17 reviews, 37 papers that reported only the design of the study or pilot results, and 3 opinion or position papers. In 12 papers, randomization was not conducted at the school level, 2 articles reported no baseline data, 1 presented results of

Obes Facts 2014;7:221–232
DOI: 10.1159/000363438




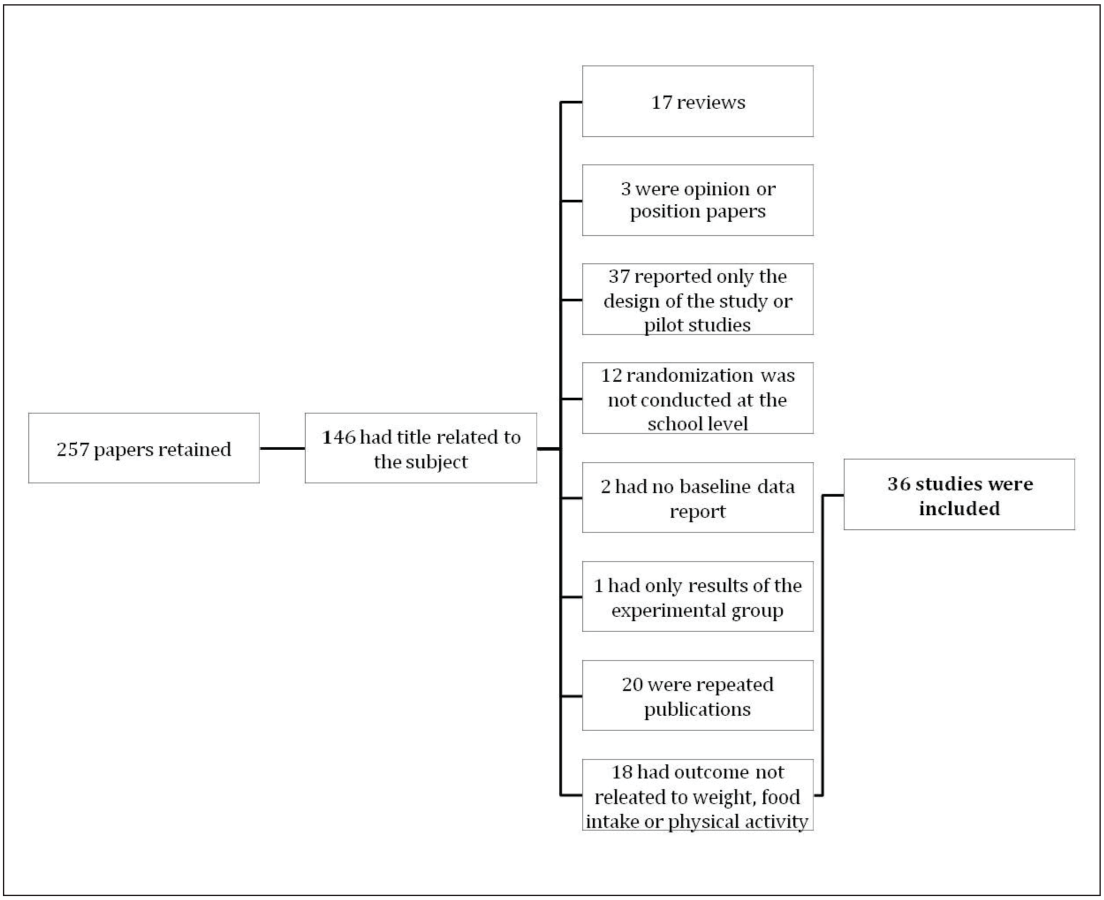


**Fig. 1.** Flowchart of the selection of the studies.

the experimental group only, 20 were repeated publications (differing in analysis or years of follow-up), and 18 only provided outcome data for diabetes behaviors that were not related to food intake or physical activity or outcome data for food sales (fig. 1). A total of 36 studies were included in the analysis (table 1).

From these 36 studies, 13 displayed unbalanced outcome measures at baseline (table 2), and most of them were based on mean BMI or BMI classification. Adjustments varied across the analyzed material. Some of the studies balanced at baseline were adjusted for baseline values of BMI and other variables. Thus, in 11 studies with balanced anthropometric measures at baseline, an analysis was conducted after adjusting for BMI [2, 12–21]. Conversely, no baseline adjustment was applied in the 5 studies that were unbalanced at baseline [22–26].

## Discussion

Our analysis showed that 35% of the reviewed studies used unbalanced baseline BMI. This indicates that the school clustering design poses a methodological challenge for analyzing the results because anthropometric measures at baseline are one of the most important factors to explain changes over time in BMI or related measures.

Obes Facts 2014;7:221–232
DOI: 10.1159/000363438



**Table 1.** School-randomized studies for obesity prevention: comparison of intervention(s) and control groups – unbalanced data refer mainly to measures of body composition

| Author, year (study acronym) | Country | Number of schools intervention/control number of participants | Age or grade | Baseline characteristics | Outcomes and adjustments | Results |
|---|---|---|---|---|---|---|
| Lubans et al., 2012 [30] (NEAT Girls) | Australia | 6/6 matched regions and school socioeconomic status (SES) 179/178 | 12 – 14 years | balanced BMI mean and classification | BMI, % fat (with and without adjustment for baseline) | 1 year follow-up; no association with and without adjustment |
| Story et al., 2012 [26] (Bright Start) | USA | 7/7 267/187 | kindergarten & 1st graders; indigenous reservation | unbalanced; BMI z-scores and classification | food intake and PA in school and family; adjusted for age, gender, and SES | 1 year follow-up; mean BMI and BMI z-score increased; reduction of overweight |
| Williamson et al., 2012 [31] (LA Health) | USA | 5/6/6 primary prevention / primary & secondary (school) / control 713/760/587 | 4th–6th graders; 10.5 ± 1.2 years; rural area | unbalanced BMI classification | food intake and PA; adjusted for baseline value | 28 months follow-up; small changes in body fat; combining intervention group |
| Puder et al., 2011 [12] (BALLABEINA) | Switzerland | 20/20 342/310 | preschool classes | balanced | PA and BMI; adjusted for baseline values, age, sex, SES, and linguistic region | 1 year follow-up; no effects on BMI; aerobic fitness increased |
| Rush et al., 2012 [15] (ENERGIZE) | New Zealand | 62/62 matched urban/rural and SES 692/660 | 5 – 7 years and 10 – 12 years | balanced | food intake and PA; adjusted for baseline | 2 year follow-up; no changes |
| Bjelland et al., 2011 [32] (HEIA) | Norway | 12/25 510/910 | 6th graders | unbalanced | sugar-sweetened beverages and screen time; adjusted for baseline | 8 months follow-up; behavior changed in girls only |
| Jansen et al., 2011 [13] | Netherlands | 10/10 matched proportion of migrants and neighborhood 1,240/1,382 | 6 – 12 years; 3rd–8th graders | balanced BMI classification; unbalanced age | food intake and PA; adjusted for baseline, SES, gender, grade, and ethnic background | no effects on BMI; intervention effect on the prevalence of overweight in grades 3 – 5 |
| Thivel et al., 2011 [22] | France | 14/5 229/228 | 6 – 10 years | unbalanced; % obese: 27%/20% | diet and PA; no adjustments | no effects on BMI or BMI classification; fitness improved |
| Aburto et al., 2011 [16] | Mexico | 8/8/11 PA: control vs. 50 min/week vs. 100 min/week 259/260/332 | primary schools | balanced mean BMI and % normal BMI; unbalanced age | adjusted for baseline | improvement according to intervention in PA; no BMI results |

*Table 1 continued on next page*

Obes Facts 2014;7:221–232
DOI: 10.1159/000363438



**Table 1.** Continued

| Author, year (study acronym) | Country | Number of schools intervention/control number of participants | Age or grade | Baseline characteristics | Outcomes and adjustments | Results |
|---|---|---|---|---|---|---|
| Llargues et al., 2011 [23] (AVall) | Spain | 8/8<br>272/236 | 5 – 6 years | unbalanced mean BMI; 16.9/16.4 ($p = 0.02$) | diet and PA; adjusted for school | difference after 2 years: –0.85/1.74 kg/m$^2$ |
| Hoffman et al., 2011 [33] | USA | 2/2<br>149/148 | 5 – 6 years | unbalanced z-score mean BMI; 0.80/0.93 | fruits and vegetables; adjusted for baseline; sex, race | 3.5 years follow-up; no change in BMI; change in fruits not sustained |
| Nemet et al., 2011 [34] | Israel | 15/15<br>417/795 | 5 – 6 years | balanced mean BMI | diet and PA; no adjustments | BMI not changed; fitness improved |
| Greening et al., 2011 [14] (TEAM) | USA | 1/1<br>204/246 | 6 – 10 years | balanced; % BMI > 95th percentile | diet and PA; adjusted for baseline values | % body fat reduced in intervention vs. control ($p = 0.02$); no change in prevalence or mean BMI |
| Neumark-Sztainer et al., 2010 [17] (New moves – girls) | USA | 3/3<br>182/174 | 15.8 ± 1.2 years; only girls | balanced % BMI; classification and % body fat | diet and PA; adjusted for baseline, age, and race | BMI and % body fat not changed; fitness and sedentary behavior improved |
| Toruner and Savaser, 2010 [35] | Turkey | 1/1<br>overweight and obese > 90th percentile of Turkish children<br>41/40 | 4th graders | balanced mean BMI | diet and PA; no adjustment | 1 year follow-up; BMI means reduced in intervention; knowledge scores improved |
| Foster et al., 2010 [36] (HEALTY) | USA | 21/21<br>2307/2296 | 6th graders; 11.3 ± 0.6 years | balanced; % BMI > 85th percentile | Diet and PA; no adjustments | 2 years follow-up; no change in BMI > 85th percentile; mean BMI, z-score, waist circumference reduced ($p = 0.04$) |
| Krimler et al., 2010 [37] (KISS) | Switzerland | 16/12<br>297/205 | 1st and 5th graders | unbalanced; overweight (Swiss centiles) | PA; adjustment for grade, sex, baseline values | 9 months follow-up; significant differences for mean BMI and sum of skinfolds ($p < 0.01$) |
| Francis et al., 2010 [38] | Trinidad and Tobago | 5/6<br>299/280 | 6th graders; 9 – 11 years | unbalanced; % BMI > 95th percentile (23.6%/12.9%) | diet and PA; adjusted for SES, gender, age, BMI baseline | 3 months follow-up; intervention had favorable change in diet, without difference in PA and eating attitude |

*Table 1 continued on next page*

Obes Facts 2014;7:221–232
DOI: 10.1159/000363438



**Table 1.** Continued

| Author, year (study acronym) | Country | Number of schools intervention/control number of participants | Age or grade | Baseline characteristics | Outcomes and adjustments | Results |
|---|---|---|---|---|---|---|
| Singhal et al., 2010 [39] | India | 1/1<br>matched SES<br>102/108 | 15 – 17 years | balanced BMI, unbalanced waist circumference (WC) and waist-hip ratio (W-HR) (p = 0.02 for both) | diet; no adjustments | 6 months follow-up; no change in the primary outcome BMI; decrease in mean values of WC (0.02) and W-HR (0.02) |
| Donnelly et al., 2009 [40] (PAAC) | USA | 14/10<br>814/713 | 2nd and 3rd graders | balanced | change in BMI; adjusted for grade, age, and gender | 3 years follow-up; no difference in BMI |
| Singh et al., 2009 [41] (Dutch Obesity Intervention in Teenagers; DOiT) | Netherlands | 10/8<br>632/476 | 12 – 14 years | balanced obesity; unbalanced overweight boys; 11.9%/18.9% | diet; adjusted for baseline values | reduction of skinfold thickness in the short and long terms; no changes in BMI |
| Muckelbauer et al., 2009 [42] | German | 17/15<br>1,641/1,309 | 2nd and 3rd graders | balanced BMI and BMI classification | beverage consumption; no adjustments | reduction in prevalence of overweight & obesity; 0.69 (0.48 – 0.98); reduction in juice consumption; no reduction in soft drinks; increased water intake |
| Marcus et al., 2009 [24] (STOPP) | Sweden | 5/5<br>1,670/1,465 | 1st–4th graders | unbalanced; overweight/obesity = 20%/16% | diet and PA changes due to changes in school environment; unadjusted | 4 years follow-up; decrease by 3.2% (20.3 to 17.1%) in intervention; increase of 2.8% (16.1 to 18.9%) in control. |
| Graf et al., 2008 [43] (CHILT) | German | 12/5 | primary schools | unbalanced | physical performance (PP); adjusted for age, sex, baseline | 4 years follow-up; PP improved; prevalence and incidence of obesity not affected |
| Gutin et al., 2008 [44] (Georgia FitKid) | USA | 9/9<br>603/584 | 3rd graders; 8.5 ± 0.6 years | balanced; % body fat and BMI z-score classification | PA adjusted for sex, race, age, and economic disadvantage status | 3 years follow-up; no change in BMI or waist; positive changes in fitness vanish during summer periods |
| Kipping et al., 2008 [45] | England | 10/9<br>331/348 | 9 – 10 years | balanced BMI and BMI classification | healthy eating, PA and TV viewing; adjusted for age, sex, and baseline characteristic | 5 months follow-up; positive changes in PA; no changes in BMI or screen time |
| Sichieri et al., 2008 [29] | Brazil | 23/24<br>526/608 | 9 – 12 years | balanced | sodas; age-adjusted | 1 year follow-up; no overall effect; girls overweight at baseline had a reduction in BMI |

*Table 1 continued on next page*

DOI: 10.1159/000363438



**Table 1.** Continued

| Author, year (study acronym) | Country | Number of schools intervention/control number of participants | Age or grade | Baseline characteristics | Outcomes and adjustments | Results |
|---|---|---|---|---|---|---|
| Plachta-Danielzik et al., 2007 [19] (KOPS) | Germany | 14/32<br>780/4217 | 6 and 10 years | BMI percentiles; balanced | diet and PA; adjusted for BMI at baseline, sex, and SES | 4 years follow-up; positive changes of BMI only in the high SES |
| Jiang et al., 2007 [46] | China | 2/3<br>1,029/1,369 | 8.3 ± 1.5 years | unbalanced BMI classification | adjusted for baseline and family SES | 3 years follow-up; prevalence of overweight and obesity reduced |
| Spiegel and Foulk, 2006 [47] | USA | 16<br>534/479 | | balanced classes; randomized for each school | diet and PA; unadjusted | 1 year follow-up; positive shifts in BMI, fruits and vegetables, and PA |
| Simon et al., 2004 [20] (ICAPS) | France | 4/4<br>475/479 | 6th grades | balanced | PA; adjusted for baseline, age, and overweight | 6 months follow-up; improvement of activity patterns |
| Lohman et al., 2003 [48] (Pathways) | USA | 21/20<br>705/663 | 3rd–5th grades; Indian children | balanced % body fat and BMI | diet and PA; unadjusted | 3 years follow-up; no change in % body fat and BMI |
| Pate et al., 2005 [21] (LEAP) | USA | 12/12<br>1,523/1,221 | 8th grades; girls only | balanced BMI classification; schools paired by SES | PA; adjusted for baseline and race/ethnicity | 1 year follow-up; increase of vigorous activity; no changes in BMI |
| James et al., 2004 [49] (CHOPPS) | England | 15/14<br>325/319 | 7 – 11 years | balanced BMI classification | beverages; unadjusted | 1 year follow-up; decrease in consumption of sodas and BMI reduction |
| Sahota et al., 2001 [50] | England | 5/5<br>314/322 | 8.4 ± 0.6 years | unbalanced; paired schools by SES; z-score 0.12/0.04 | diet and PA; unadjusted | 1 year follow-up; no change in BMI score or classification |
| Gortmaker et al., 1999 [2] (Planet Health) | USA | 5/5<br>641/654 | 11.7 ± 0.7 years | balanced; paired schools by SES | diet, PA and TV viewing habits; adjusted for age, race, and baseline | 2 years follow-up; prevalence of obesity was reduced only among girls |

PA = Physical activity.

Obes Facts 2014;7:221–232
DOI: 10.1159/000363438



**Table 2.** Information about response, balancing, adjustment, and sample size of the selected school-randomized studies for obesity prevention

| Author, year | Response | Balancing | Adjustment | Sample size |
|---|---|---|---|---|
| Lubans et al., 2012 [30] | not significant | balanced | not adjusted | 357 |
| Story et al., 2012 [26] | significant | unbalanced | not adjusted | 454 |
| Williamson et al., 2012 [31] | significant | unbalanced | adjusted | 2,060 |
| Puder et al., 2011 [12] | not significant | balanced | adjusted | 652 |
| Rush et al., 2012 [15] | not significant | balanced | adjusted | 1,352 |
| Bjelland et al., 2011 [32] | not significant | unbalanced | adjusted | 1,420 |
| Jansen et al., 2011 [13] | not significant | balanced | adjusted | 2,622 |
| Thivel et al., 2011 [22] | not significant | unbalanced | not adjusted | 457 |
| Aburto et al., 2011 [16] | not significant | balanced | adjusted | 851 |
| Llargues et al., 2011 [23] | significant | unbalanced | not adjusted | 508 |
| Hoffman et al., 2011 [33] | not significant | unbalanced | adjusted | 297 |
| Nemet et al., 2011 [34] | not significant | balanced | not adjusted | 1,212 |
| Greening et al., 2011 [14] | not significant | balanced | adjusted | 450 |
| Neumark-Sztainer et al., 2010 [17] | not significant | balanced | adjusted | 356 |
| Toruner and Savaser, 2010 [35] | significant | balanced | not adjusted | 81 |
| Foster et al., 2010 [36] | not significant | balanced | not adjusted | 4,603 |
| Krimler et al., 2010 [37] | significant | unbalanced | adjusted | 502 |
| Francis et al., 2010 [38] | not significant | unbalanced | adjusted | 579 |
| Singhal et al., 2010 [39] | not significant | balanced | not adjusted | 210 |
| Donnelly et al., 2009 [40] | not significant | balanced | not adjusted | 1,527 |
| Singh et al., 2009 [41] | not significant | unbalanced | adjusted | 1,108 |
| Muckelbauer et al., 2009 [42] | significant | balanced | not adjusted | 2,950 |
| Marcus et al., 2009 [24] | significant | unbalanced | not adjusted | 3,135 |
| Graf et al., 2008 [43] | not significant | unbalanced | adjusted | 615 |
| Gutin et al., 2008 [44] | not significant | balanced | not adjusted | 1,187 |
| Kipping et al., 2008 [45] | not significant | balanced | adjusted | 679 |
| Sichieri et al., 2008 [29] | not significant | balanced | not adjusted | 1,134 |
| Plachta-Danielzik et al., 2007 [19] | significant | balanced | adjusted | 4,997 |
| Jiang et al., 2007 [46] | significant | unbalanced | adjusted | 2,398 |
| Spiegel and Foulk, 2006 [47] | significant | balanced | not adjusted | 1,013 |
| Simon et al., 2004 [20] | not significant | balanced | adjusted | 954 |
| Lohman et al., 2003 [48] | not significant | balanced | not adjusted | 1,368 |
| Pate et al., 2005 [21] | not significant | balanced | adjusted | 2,744 |
| James et al., 2004 [49] | significant | balanced | not adjusted | 644 |
| Sahota et al., 2001 [50] | not significant | unbalanced | not adjusted | 636 |
| Gortmaker et al., 1999 [2] | not significant | balanced | adjusted | 1,295 |

Adjusting for baseline measures was frequently utilized in the observed studies, even among those with balanced BMI at baseline. However, adjusting for baseline BMI may bias the results – an issue that has been well discussed [5–11]. For example, a book by Fitzmaurice et al. [9] discussed adjustment for baseline response using measures on the weight gain of infants aged 12 to 24 months. In this specific case, data were unbalanced at baseline because boys are heavier than girls. Both sexes gained the same amount of weight within 12 months, and it was concluded that there was no gender effect on body weight change. However, if the analysis includes adjustment for baseline values, boys gain more weight than girls. Findings such as this one, known as Lord's paradox, have generated a heated debate among analysts. In 1967, Lord described this paradox within a linear model framework: $E(XA) = E(YA)$ and $E(XB) = E(YB)$, but $E(YB|XB = x) - E(YA|XA = x) > 0$, uniformly in x (i.e., $E(YB - XB|XB = x) - E(YA - XA|XA = x) > 0$). According to Lord, the marginal group means seem to indicate no group effect (i.e., $E(YB - XB) = E(YA - XA)$), yet the comparison of conditional expectations appears to contradict the lack of a group differential effect [51].

For experimental designs, the same principles apply with the additional challenge that data are expected to be balanced at baseline. The reviewed data show that cluster designs favor imbalance, unless the number of clusters is high. We have also shown that adjustment for BMI at baseline is frequently performed.

Although adjusting for the baseline values of parameters that are highly influenced by baseline values is a standard procedure, this approach can bias the results towards an alternative hypothesis – if the control group has a greater BMI – or towards the null – if the experimental group has a greater BMI. Therefore, by forcing a baseline balance in experimental studies, a spurious relationship between treatment and outcome can be observed. Using procedures such as those from mixed-effects models represents a better way to attain results. This modeling allows testing the time effect, the treatment effect per se, and time × treatment effect (the variable that indicates change). In this type of analysis, change over time can be tested clearly without needing to adjust for baseline. The authors of the present study observed this bias in an unbalanced cluster school-randomized study, where the adjustment for baseline changed the result from a lack of association to a statistically significant association [52].

Other possible sources of discrepancy were related to the use or interpretation of the outcome measures. Table 1 shows that the reviewed studies used BMI change, BMI z-score change, prevalence of overweight and obesity, or a combination of these. In relation to the use of BMI z-score change, two studies demonstrated that changes in BMI were less subject to error compared to BMI z-scores [27, 28]. The differences were due to abrupt changes in BMI z-scores and changes in the variance of BMI z-scores with growth.

Another source of discrepancy in the studies using prevalence as an outcome was the analysis of overweight and obesity as independent outcomes. For example, in the prevention trial in American Indian children [26], the intervention was not associated with statistically significant changes in variables measured on a continuous scale: BMI, BMI z-scores, skinfolds, and percentage of body fat. However, analysis of BMI as a categorical variable showed a significant decrease only in the prevalence of overweight, and the authors concluded a need for primary prevention because overweight but not obesity was reduced. However, the overall prevalence of excessive weight (overweight plus obesity at the end of the study) decreased from 42.88% to 41.13%, similar to the results from the analysis of BMI as a continuous variable. In addition, studies have shown that interventions have a greater effect on those who are overweight or obese at the beginning of the study [3, 29], indicating that primary prevention has not been achieved.

In conclusion, unbalanced BMI values at baseline, the inadequacy of z-scores as an outcome, and a misleading definition of primary prevention of obesity may explain the controversial results of school-based obesity interventions. A pooled analysis of these studies, using mixed-effects models without adjustment for baseline, may help to better summarize these results.

## Acknowledgements

This review was funded by the Brazilian National Research Council (CNPq senior fellowship) through a sabbatical at Harvard University to RS and a fellowship to DBC from the Brazilian Federal Agency for the Improvement of Higher Education (CAPES).

## Disclosure Statement

The authors have read and approved this version of the manuscript. None of the authors have any conflicts of interest.

Obes Facts 2014;7:221–232
DOI: 10.1159/000363438



## References

1 Davis SM, Going SB, Helitzer DL, Teufel NI, Snyder P, Gittelsohn J, Metcalfe L, Arviso V, Evans M, Smyth M, Brice R, Altaha J: Pathways: a culturally appropriate obesity-prevention program for American Indian schoolchildren. Am J Clin Nutr 1999;69(4 suppl):796S–802S.
2 Gortmaker SL, Peterson K, Wiecha J, Sobol AM, Dixit S, Fox MK, Laird N: Reducing obesity via a school-based interdisciplinary intervention among youth: Planet Health. Arch Pediatr Adolesc Med 1999;153:409–418.
3 Khambalia AZ, Dickinson S, Hardy LL, Gill T, Baur LA: A synthesis of existing systematic reviews and meta-analyses of school-based behavioural interventions for controlling and preventing obesity. Obes Rev 2012;13:214–233.
4 Doak C, Heitmann BL, Summerbell C, Lissner L: Prevention of childhood obesity – what type of evidence should we consider relevant? Obes Rev 2009;10:350–356.
5 Cologne JB: Re: 'when is baseline adjustment useful in analyses of change? An example with education and cognitive change'. Am J Epidemiol 2006;164:1138–1139;author reply 1139–1140.
6 Marcon A, Accordini S, de Marco R: Adjustment for baseline value in the analysis of change in FEV1 over time. J Allergy Clin Immunol 2009;124:1120.
7 McArdle PF, Whitcomb BW: Improper adjustment for baseline in genetic association studies of change in phenotype. Hum Hered 2009;67:176–182.
8 Pencina MJ, D'Agostino RB, Beiser AS, Cobain MR, Vasan RS: Estimating lifetime risk of developing high serum total cholesterol: adjustment for baseline prevalence and single-occasion measurements. Am J Epidemiol 2007;165:464–472.
9 Fitzmaurice GM, Laird NM, Ware JH: Applied Longitudinal Analysis. Hoboken, NJ, Wiley, 2011.
10 Tu Y-K, Gilthorpe MS: Statistical Thinking in Epidemiology. Boca Raton, FL, CRC Press, 2012.
11 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.
12 Puder JJ, Marques-Vidal P, Schindler C, Zahner L, Niederer I, Bürgi F, Ebenegger V, Nydegger A, Kriemler S: Effect of multidimensional lifestyle intervention on fitness and adiposity in predominantly migrant preschool children (Ballabeina): cluster randomised controlled trial. BMJ 2011;343:d6195.
13 Jansen W, Borsboom G, Meima A, Zwanenburg EJ, Mackenbach JP, Raat H, Brug J: Effectiveness of a primary school-based intervention to reduce overweight. Int J Pediatr Obes 2011;6:e70–77.
14 Greening L, Harrell KT, Low AK, Fielder CE: Efficacy of a school-based childhood obesity intervention program in a rural southern community: TEAM Mississippi Project. Obesity (Silver Spring) 2011;19:1213–1219.
15 Rush E, Reed P, McLennan S, Coppinger T, Simmons D, Graham D: A school-based obesity control programme: Project Energize. Two-year outcomes. Br J Nutr 2012;107:581–587.
16 Aburto NJ, Fulton JE, Safdie M, Duque T, Bonvecchio A, Rivera JA: Effect of a school-based intervention on physical activity: cluster-randomized trial. Med Sci Sports Exerc 2011;43:1898–1906.
17 Neumark-Sztainer DR, Friend SE, Flattum CF, Hannan PJ, Story MT, Bauer KW, Feldman SB, Petrich CA: New moves – preventing weight-related problems in adolescent girls: a group-randomized study. Am J Prev Med 2010;39:421–432.
18 Singh AS, Chin A Paw MJ, Brug J, van Mechelen W: Dutch obesity intervention in teenagers: effectiveness of a school-based program on body composition and behavior. Arch Pediatr Adolesc Med 2009;163:309–317.
19 Plachta-Danielzik S, Pust S, Asbeck I, Czerwinski-Mast M, Langnäse K, Fischer C, Bosy-Westphal A, Kriwy P, Müller MJ: Four-year follow-up of school-based intervention on overweight children: the KOPS study. Obesity (Silver Spring) 2007;15:3159–3169.
20 Simon C, Wagner A, DiVita C, Rauscher E, Klein-Platat C, Arveiler D, Schweitzer B, Triby E: Intervention centred on adolescents' physical activity and sedentary behaviour (ICAPS): concept and 6-month results. Int J Obes Relat Metab Disord 2004;28(suppl 3):S96–S103.
21 Pate RR, Ward DS, Saunders RP, Felton G, Dishman RK, Dowda M: Promotion of physical activity among high-school girls: Aa randomized controlled trial. Am J Public Health 2005;95:1582–1587.
22 Thivel D, Isacco L, Lazaar N, Aucouturier J, Ratel S, Doré E, Meyer M, Duché P: Effect of a 6-month school-based physical activity program on body composition and physical fitness in lean and obese schoolchildren. Eur J Pediatr 2011;170:1435–1443.
23 Llargues E, Franco R, Recasens A, Nadal A, Vila M, Pérez MJ, Manresa JM, Recasens I, Salvador G, Serra J, Roure E, Castells C: Assessment of a school-based intervention in eating habits and physical activity in school children: The AVall study. J Epidemiol Community Health 2011;65:896–901.
24 Marcus C, Nyberg G, Nordenfelt A, Karpmyr M, Kowalski J, Ekelund U: A 4-year, cluster-randomized, controlled childhood obesity prevention study: STOPP. Int J Obes (Lond) 2009;33:408–417.
25 Sahota P, Rudolf MC, Dixey R, Hill AJ, Barth JH, Cade J: Evaluation of implementation and effect of primary school based intervention to reduce risk factors for obesity. BMJ 2001;323:1027–1029.
26 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.
27 Berkey CS, Colditz GA: Adiposity in adolescents: change in actual BMI works better than change in BMI z score for longitudinal studies. Ann Epidemiol 2007;17:44–50.

28 Cole TJ, Faith MS, Pietrobelli A, Heo M: What is the best measure of adiposity change in growing children: BMI, BMI %, BMI z-score or BMI centile? Eur J Clin Nutr 2005;59:419–425.
29 Sichieri R, Paula Trotte A, de Souza RA, Veiga GV: School randomised trial on prevention of excessive weight gain by discouraging students from drinking sodas. Public Health Nutr 2009;12:197–202.
30 Lubans DR, Morgan PJ, Okely AD, Dewar D, Collins CE, Batterham M, Callister R, Plotnikoff RC: Preventing obesity among adolescent girls: one-year outcomes of the nutrition and enjoyable activity for teen girls (NEAT Girls) cluster randomized controlled trial. Arch Pediatr Adolesc Med 2012;166:821–827.
31 Williamson DA, Champagne CM, Harsha DW, Han H, Martin CK, Newton RL Jr, Sothern MS, Stewart TM, Webber LS, Ryan DH: Effect of an environmental school-based obesity prevention program on changes in body fat and body weight: a randomized trial. Obesity (Silver Spring) 2012;20:1653–1661.
32 Bjelland M, Bergh IH, Grydeland M, Klepp KI, Andersen LF, Anderssen SA, Ommundsen Y, Lien N: Changes in adolescents' intake of sugar-sweetened beverages and sedentary behaviour: results at 8 month mid-way assessment of the HEIA study – a comprehensive, multi-component school-based randomized trial. Int J Behav Nutr Phys Act 2011;8:63.
33 Hoffman JA, Thompson DR, Franko DL, Power TJ, Leff SS, Stallings VA: Decaying behavioral effects in a randomized, multi-year fruit and vegetable intake intervention. Prev Med 2011;52:370–375.
34 Nemet D, Geva D, Eliakim A: Health promotion intervention in low socioeconomic kindergarten children. J Pediatr 2011;158:796–801.e1.
35 Toruner EK, Savaser S: A controlled evaluation of a school-based obesity prevention in Turkish school children. J Sch Nurs 2010;26:473–482.
36 HEALTHY Study Group, Foster GD, Linder B, Baranowski T, Cooper DM, Goldberg L, Harrell JS, Kaufman F, Marcus MD, Treviño RP, Hirst K: A school-based intervention for diabetes risk reduction. N Engl J Med 2010; 363:443–453.
37 Kriemler S, Zahner L, Schindler C, Meyer U, Hartmann T, Hebestreit H, Brunner-La Rocca HP, van Mechelen W, Puder JJ: Effect of school based physical activity programme (KISS) on fitness and adiposity in primary schoolchildren: cluster randomised controlled trial. BMJ 2010;340:c785.
38 Francis M, Nichols SS, Dalrymple N: The effects of a school-based intervention programme on dietary intakes and physical activity among primary-school children in Trinidad and Tobago. Public Health Nutr 2010;13: 738–747.
39 Singhal N, Misra A, Shah P, Gulati S: Effects of controlled school-based multi-component model of nutrition and lifestyle interventions on behavior modification, anthropometry and metabolic risk profile of urban Asian Indian adolescents in North India. Eur J Clin Nutr 2010;64:364–373.
40 Donnelly JE, Greene JL, Gibson CA, Smith BK, Washburn RA, Sullivan DK, DuBose K, Mayo MS, Schmelzle KH, Ryan JJ, Jacobsen DJ, Williams SL: Physical activity across the curriculum (PAAC): a randomized controlled trial to promote physical activity and diminish overweight and obesity in elementary school children. Prev Med 2009;49:336–341.
41 Singh AS, Chinapaw MJ, Brug J, van Mechelen W: Process evaluation of a school-based weight gain prevention program: the Dutch Obesity Intervention in Teenagers (DOiT). Health Educ Res 2009;24:772–777.
42 Muckelbauer R, Libuda L, Clausen K, Toschke AM, Reinehr T, Kersting M: Promotion and provision of drinking water in schools for overweight prevention: randomized, controlled cluster trial. Pediatrics 2009;123:e661–667.
43 Graf C, Koch B, Falkowski G, Jouck S, Christ H, Staudenmaier K, Tokarski W, Gerber A, Predel HG, Dordel S: School-based prevention: effects on obesity and physical performance after 4 years. J Sports Sci 2008;26: 987–994.
44 Gutin B, Yin Z, Johnson M, Barbeau P: Preliminary findings of the effect of a 3-year after-school physical activity intervention on fitness and body fat: the Medical College of Georgia Fitkid Project. Int J Pediatr Obes 2008; 3(suppl 1):3–9.
45 Kipping RR, Payne C, Lawlor DA: Randomised controlled trial adapting US school obesity prevention to England. Arch Dis Child 2008;93:469–473.
46 Jiang J, Xia X, Greiner T, Wu G, Lian G, Rosenqvist U: The effects of a 3-year obesity intervention in schoolchildren in Beijing. Child Care Health Dev 2007;33:641–646.
47 Spiegel SA, Foulk D: Reducing overweight through a multidisciplinary school-based intervention. Obesity (Silver Spring) 2006;14:88–96.
48 Lohman T, Thompson J, Going S, Himes JH, Caballero B, Norman J, Cano S, Ring K: Indices of changes in adiposity in American Indian children. Prev Med 2003;37:S91–96.
49 James J, Thomas P, Cavan D, Kerr D: Preventing childhood obesity by reducing consumption of carbonated drinks: cluster randomised controlled trial. BMJ 2004;328:1237.
50 Sahota P, Rudolf MC, Dixey R, Hill AJ, Barth JH, Cade J: Randomised controlled trial of primary school based intervention to reduce risk factors for obesity. BMJ 2001;323:1029–1032.
51 Chen A, Bengtsson T, Ho TK: A Regression Paradox for Linear Models: Sufficient Conditions and Relation to Simpson's Paradox. Am Stat 2009; 63:218–225.
52 Cunha DB, de Souza BN, Pereira RA, Sichieri R: Effectiveness of a randomized school-based intervention involving families and teachers to prevent excessive weight gain in Brazil. PLoS One 2013;8:e57498.