Systematic Review and Meta-Analysis of Comprehensive Behavioral Family Lifestyle Interventions Addressing Pediatric Obesity

David M. Janicke,1 PHD, Ric G. Steele,2 PHD, Laurie A. Gayes,2 MS, Crystal S. Lim,1 PHD, Lisa M. Clifford,1 PHD, Elizabeth M. Schneider,1 PHD, Julia K. Carmody,1 MS, and Sarah Westen1 MS

1 Department of Clinical and Health Psychology, University of Florida and 2 Clinical Child Psychology Program, University of Kansas

All correspondence concerning this article should be addressed to David M. Janicke, PHD, Department of Clinical and Health Psychology, University of Florida, PO Box 100165, Gainesville, FL 32610, USA.

E-mail: djanicke@phhp.ufl.edu

Received January 16, 2014; revisions received March 29, 2014; accepted April 7, 2014

Purpose To conduct a meta-analysis of randomized controlled trials examining the efficacy of comprehensive behavioral family lifestyle interventions (CBFLI) for pediatric obesity. Method Common research databases were searched for articles through April 1, 2013. 20 different studies (42 effect sizes and 1,671 participants) met inclusion criteria. Risk of bias assessment and rating of quality of the evidence were conducted. Results The overall effect size for CBFLIs as compared with passive control groups over all time points was statistically significant (Hedge’s g = 0.473, 95% confidence interval [.362, .584]) and suggestive of a small effect size. Duration of treatment, number of treatment sessions, the amount of time in treatment, child age, format of therapy (individual vs. group), form of contact, and study use of intent to treat analysis were all statistically significant moderators of effect size. Conclusion CBFLIs demonstrated efficacy for improving weight outcomes in youths who are overweight or obese.

Key words children; intervention outcome; meta-analysis; obesity.

Childhood obesity is a significant public health concern, with roughly 32% of children considered overweight or obese (Ogden, Carroll, Kit, & Flegal, 2012). The sequelae of pediatric obesity include comorbid medical complications, such as high blood pressure, abnormal lipid proteins, liver disease, sleep disordered breathing, and type 2 diabetes (Daniels, 2006). Overweight and obese children face increased risks for poor self-esteem and body image, peer victimization, weight stigmatization, depressive symptoms, and other psychological difficulties (Daniels, 2006; Puhl, Luedicke, & Heuer, 2011). Obesity in childhood also leads to increased risks of being obese in adulthood (Singh, Mulder, Twisk, van Mechelen, & Chinapaw, 2008). Pediatric interventions that lead to weight reductions have been associated with improvements in metabolic factors (Ebbeling, Leidig, Sinclair, Hangen, & Ludwig, 2003) and self-esteem (Janicke et al., 2008). There is a need for evidence-based pediatric obesity interventions to treat increased weight status in children and adolescents.

Research to elucidate factors contributing to the continual rise in rates of pediatric obesity provides evidence of environment by gene interactions contributing to child and adolescent weight status. Studies have also shown that environmental factors can alter genetic factors associated with weight (Koletzko, Brands, Poston, Godfrey, & Demmelmaier, 2012). Potential environmental factors include consuming large portions of high-calorie nutrient-poor foods, decreased engagement in physical activity, and increased time spent in sedentary behaviors (Lioret, Volatier, Lafay, Touvier, & Maire, 2009; Spear et al., 2007). One of the strongest predictors of child weight is parent weight status (Whitaker, Wright, Pepe, Seidel, & Dietz, 1997). Given that parents play a significant role in establishing patterns of eating and physical activity.
throughout childhood (Spear et al., 2007), behavioral family lifestyle interventions have been developed that focus on modifying the obesogenic family environment to address weight management in children and adolescents.

Generally, behavioral family lifestyle interventions focus on encouraging overweight and obese children and their parents to modify the family’s dietary intake, physical activity habits, or both. Dietary modifications generally target reducing the consumption of high-fat/high-calorie foods and increasing consumption of fruits and vegetables, as well as dietary monitoring through classifications systems such as the Stoplight diet (Epstein, 1993). Activity targets commonly include increasing the intensity and duration of physical activity (i.e., play, family activities, organized sports, or structured exercise program) and reducing time spent in sedentary activities (e.g., television viewing). Behavioral strategies to support the adoption of healthier lifestyle behaviors are central to these programs. Specific strategies may include the following: parent modeling; monitoring of dietary intake and physical activity; goal setting; problem solving; gradual shaping; child behavior management strategies including differential attention and contingency management; and stimulus control.

Parents are often considered a critical agent of change in behavioral lifestyle interventions, as they exercise significant control over children’s eating and physical environment and, ultimately, behaviors. As such, one or both parents are typically included in treatment. Parents are often encouraged, if not specifically targeted, to also follow healthy lifestyle treatment recommendations. The format of behavioral family lifestyle interventions varies from group based (multiple families participate at one time) and individual family based (family meets one-on-one with interventionist), to a combination of these formats. The extent to which children and adolescents are directly involved in treatment also varies across interventions. While the vast majority of interventions require children or adolescents to regularly attend intervention sessions with their parents, some recent interventions have adopted a ‘parent-only’ model, in which only the participating parent(s) attend intervention sessions during which they are taught strategies to support their child or adolescent in modifying weight-related lifestyle behaviors (Golan, Kaufman, & Shahar, 2006; Janicke et al., 2008).

The number of content components also varies across lifestyle interventions. Some interventions may include only two components. For example, an intervention may address dietary intake using behavior strategies but not address physical activity. Others may address physical activity and dietary intake but not use behavior strategies. Finally, there are those interventions that are comprehensive, in that they address all three intervention components. For the purposes of this review, interventions that included content addressing all three areas (i.e., dietary intake, physical activity, and behavior strategies) are referred to hereafter as comprehensive behavioral family lifestyle interventions (CBFLIs). Owing to the increase in outcome research examining treatments for pediatric obesity, there is a need to understand the current efficacy of these interventions, as well as to identify moderators of treatment success to inform future clinical practice and treatment outcome research.

Jelalian and Saelens (1999) conducted a review of pediatric obesity treatments in the empirically supported treatment series published in the Journal of Pediatric Psychology over a decade ago. They concluded that behavioral lifestyle interventions could be considered well-established treatments and that there was strong evidence for their short- and long-term efficacy (e.g., 5 and 10 years posttreatment) in reducing weight in school-age children. The evidence of behavioral lifestyle interventions for adolescents was only considered “promising” owing to the limited number of treatment outcome studies in adolescents at that time. However, there was no requirement that the treatments reviewed by Jelalian and Saelens (1999) be CBFLIs.

Not surprisingly, since the Jelalian and Saelens review, a number of systemic reviews and meta-analyses of pediatric obesity interventions have been published (Kitzmann et al., 2010; McGovern et al., 2008; Oude-Luttikhuis et al., 2009; Seo et al., 2010; Whitlock, O’Connor, Williams, Beil, & Lutz, 2010; Willfley et al., 2007). The general consensus from these published reports is that pediatric obesity interventions result in small to moderate short-term improvements in adiposity in children and adolescents. However, a thorough examination of this literature finds mixed methodology and results across meta-analyses, with great variety in the number of studies included in different meta-analyses, largely owing to differing foci and study inclusion and exclusion criteria (i.e., intent-to-treat [ITT] analysis, different intervention components, duration of treatment or follow-up analysis, randomization of treatment participants, minimum sample size requirements, adequate control condition). Moreover, a number of meta-analyses use broad inclusion criteria that allow for lifestyle interventions with diverse treatment components (two or three component programs) or interventions delivered in different settings (e.g., community, clinic, research, and school; Kitzmann et al., 2010). While informative, one disadvantage to this ‘broader approach’ is that such clinical heterogeneity precludes a meaningful answer as to the
intervention type to which an effect can be generalized (Cohen, Thombs, & Hagedoorn, 2010).

Although a number of meta-analyses include intervention studies with behavioral family components, many do not provide explicit examination of interventions with a behavioral family component or adequately define behavioral components (Kitzmann et al., 2010; McGovern et al., 2008; Seo et al., 2010; Wilfley et al., 2007). To our knowledge only three meta-analyses have explicitly examined the efficacy of pediatric lifestyle interventions that require the inclusion of a behavioral family component (Oude-Luttikhuys et al., 2009; Whitlock et al., 2010; Young, Northern, Lister, Drummond, & O’Brien, 2007). However, even these three reviews do not require the use of all three common lifestyle intervention components (i.e., dietary intake, physical activity, and behavioral strategies) for study inclusion in the meta-analytic reviews.

Whitlock and colleagues identified 11 fair- or good-quality behavioral family lifestyle intervention trials published between 2005 and 2008 addressing weight loss in overweight and obese children and adolescents ≤19 years of age (Whitlock et al., 2010). Results revealed that behavioral family lifestyle interventions of medium to high intensity (i.e., ≥26 hr of treatment contact) had moderate to large effects on weight outcomes compared with very low intensity interventions (e.g., <10 hr) in the short term, with intervention effectiveness tending to increase with more intensive interventions. The authors also concluded that there was insufficient evidence to draw conclusions on outcomes >12 months after treatment.

The Cochrane Collaborative published a review that included 54 randomized controlled trials (RCTs) published through May 2008 and focused on lifestyle interventions (i.e., dietary, physical activity, and/or behavioral oriented treatment) for children and adolescents (mean age <18 years at posttreatment) (Oude-Luttikhuys et al., 2009). However, they only identified eight behavioral family lifestyle interventions that qualified for a separate meta-analysis examining both short- and long-term outcomes in elementary-age and adolescent youth. While noting many limitations in the literature and the lack of quality data to adequately ascertain treatment efficacy, the authors concluded that behavioral family lifestyle interventions result in significant and clinically meaningful changes in weight status for children and adolescents compared with standard care and self-help, both in the short and long term.

Finally, Young and colleagues (2007) identified 44 pediatric obesity interventions across 16 studies for children aged 5–12 years. Of these 44 interventions, 31 were behavioral lifestyle interventions. Results indicated that relative to alternative treatment interventions, behavioral lifestyle interventions produce larger effects at posttreatment. Relative to Oude-Luttikhuys et al. (2009) who required intent-to-treatment analysis for study inclusion, Young and colleagues did not include intent-to-treatment analysis as an inclusion criteria. This may in part account for the larger number of interventions included and the large effect size reported by Young et al.

While these three meta-analyses provide essential information on the relative efficacy of behavioral family lifestyle interventions, the most recent review was of interventions published only through May 2008. Additionally, the target of past meta-analyses of behavioral family lifestyle interventions has been on change in weight status, and therefore, more information is needed on the impact of these interventions on key secondary outcomes such as dietary intake and physical activity, as well as potential moderators of treatment effects. Finally, none of these past reviews explicitly required that interventions were CBFLIs, meaning they did not require interventions to include all three treatment components to be included (i.e., dietary intake, physical activity, and behavioral strategies). Thus, the primary objective of this review is to evaluate the efficacy of CBFLIs in reducing or stabilizing child adiposity in overweight and obese youth. A secondary objective of this review is to evaluate the impact of CBFLIs on key secondary outcomes including child caloric intake, physical activity, sedentary behavior, and parent use of behavior management strategies. A final secondary objective of this meta-analysis is to evaluate potential moderators of treatment outcome including (1) child age; (2) child sex; (3) duration and intensity of treatment; (4) length of time from baseline to outcome assessment; (5) whether the parent(s) is targeted for health behavior or weight change; (6) type of comparison control condition; (7) methodological rigor of studies; and (8) other study characteristics (i.e., ITT analysis used, manualized intervention, exclusion of children over the 99th percentile for body mass index (BMI), form of therapy [individual vs. group vs. both], and form of contact [in-person, in-person plus internet, phone]). These moderating variables were selected based on findings from previous published studies, as well as reviews and recommendations for treating pediatric obesity (Faith et al., 2012; Jelalian & Saelens, 1999; Spear et al., 2007; Wilfley et al., 2007), and represent potential important directions for future intervention research. The proposed methods will draw heavily from the Cochrane Collaboration and represent current best practice in systematic review methodology. This review will add to the literature by examining only those studies that include behavioral lifestyle interventions that are comprehensive in nature (intervention explicitly addresses dietary intake and
Methods
Criteria for Considering Studies for This Review

Types of Studies
RCTs comparing an active CBFLI for overweight or obesity in children or adolescents with an attention/educational control, other active treatment, or wait list control were considered for this review. Studies meeting the following criteria were included:

- RCTs published in full in peer-reviewed journals, unpublished dissertations, and brief reports that include sufficient methodological details to allow (1) critical appraisal of study quality and (2) calculation of effect sizes.
- Primary aim of the trial was to evaluate the efficacy or effectiveness of an active CBFLI for overweight or obesity in children or adolescents with change in weight as the primary outcome.
- Both the treatment and the control arms of the study must have a minimum sample size of 10 at the posttreatment assessment.

Types of Participants
Children and adolescents, ≤ 19 years of age at baseline, who are classified as overweight or obese based on age and gender norms (Kuczmarski et al., 2000), were included in this review. Studies were excluded if (1) participants were from special populations, such as children or adolescents with developmental disabilities, developmental delays, or cognitive impairment, (2) child or adolescent participants were diagnosed with a chronic illness that inhibits typical growth and development (e.g., cystic fibrosis) or bulimia nervosa, (3) any of the child participants were classified as normal weight in any of the treatment arms at baseline assessment, or (4) children or adolescents with an underlying diagnosis that predisposes children to obesity or greater than normal weight gain (e.g., Prader-Willi Syndrome, Thyroid disease).

Types of Interventions
Interventions were included if (1) the active treatment included dietary, physical activity, and behavioral components that focused on change in weight and weight-related health behaviors (for explicit criteria see below), (2) the primary objectives of the intervention were to produce weight loss or prevent further weight gain in children or adolescents who were already obese or overweight at baseline, and (3) interventions were conducted in outpatient or community settings. Studies were excluded if the active intervention being evaluated (1) involved a curriculum-based school weight management program and (2) child weight was based on subjective methods of measurement (e.g., self-reported, parent-reported, interviewer estimated). Active interventions in which only parents attended treatment sessions, but child weight or adiposity was the primary outcome, were eligible for inclusion in this analysis. Studies reporting only previously reported data and studies in which medication was part of treatment were also excluded.

Interventions that were considered to include a dietary component must have included education, presentation, or discussion of healthy eating, dietary guidelines, or food preparation strategies. Interventions that were considered to address physical activity must have included one of the following: (1) education on the importance of adequate physical activity, (2) encouragement to increase physical activity, (3) guidelines for recommended levels of physical activity for children, or (4) structured physical activity time for children during intervention sessions. Finally, for an intervention to have been considered “behavioral,” the description of the active treatment must have included one of the following terms: self-monitoring, stimulus control, goal setting, positive reinforcement, differential or contingency management, behavioral parent training, or problem solving. Moreover, these terms must have described efforts to help families make changes to child dietary intake or physical activity. Only studies that included sufficient information to determine whether they met these inclusion criteria were eligible for review.

Type of Outcome Measures
Change in child weight at posttreatment was the primary outcome targeted in this meta-analysis. For inclusion, measures of child weight must have been objective, adjusted for child age and sex, and reported as one of the following: BMI, BMI z-score, BMI percentile, percent overweight, or adiposity. Data were analyzed at pretreatment and posttreatment. When available, data were also analyzed at the first follow-up visit that occurred at least 6 months postbaseline. Change in child weight must have been reported or was able to be calculated from the inclusion of pretreatment and posttreatment weight values. In addition, child weight must have been assessed at the same time points across all randomized conditions.
Search Methods for Identification of Studies

A two-phase search strategy was used to identify studies for this review. First, five electronic databases (the Cochrane Central Register of Controlled Trials [CENTRAL]; ProQuest Dissertations and Theses Full Text; PsycINFO via EBSCO; PubMed; and Web of Science) were systematically searched by one member of the research team using the following truncated key terms to identify studies for review: (child OR adolescent OR family OR parent) AND (intervention OR treatment) AND (overweight OR obesity). While the integrity of the search terms was not significantly altered, some changes in search strategies were necessary based on the requirements of the particular database used (i.e., use of MeSH terms for PubMed). Systematic searches of these databases occurred in February 2013 and again in April 2013 to allow for identification of studies published up through April 1, 2013. The search was limited to articles that were available in English.

A total of 7,734 articles were identified through the initial searches across databases, while an additional 12 articles were identified from examination of published meta-analyses and systematic reviews reporting on interventions addressing pediatric obesity (see Figure 1 for the PRISMA Flow diagram). Of these 7,746 studies, 192 were removed as they represented duplicates across one or more databases, 9 studies were removed as no full-text article of the study was available, and 7,149 were excluded based on an initial screening of the title and abstract. Study exclusion criteria for the abstract review were as follows:

1. Participants
   a. Youth >19 years of age.
   b. Less than 10 participants in any condition at posttreatment.
2. Intervention
   a. At least one arm of the intervention is not a primarily behavioral intervention.
   b. Nonrandomized controlled trial.
   c. Child weight was not a primary outcome of the intervention.
3. Comparison groups
   a. Study did not include at least one of the following control conditions: (i) no treatment control, (ii) waitlist control, (iii) education control, or (iv) treatment as usual control.

During the second phase of the search, the remaining 298 articles were assigned to one of three teams composed of two reviewers who each independently assessed the full text of the remaining articles to determine eligibility for inclusion in the analysis. When an inclusion or exclusion decision could not be reached by the two reviewers, the full team of reviewers discussed the concerns to yield a consensus. The reviewers were not blind to the names of study authors, institutions, or journals. Of the 298 full-text articles reviewed, 278 were excluded (see Figure 1 for reasons for exclusions). The two-phase process resulted in 20 studies eligible for inclusion in the current meta-analysis.

Data Extraction

A coding document was developed for data extraction procedures. This included critical study information such as references, details of participants and their demographics, aspects of the intervention or therapy, characteristics of the treatment team, the setting of the intervention, outcome measures, and statistical outcomes. Both members from each review team carried out data extraction from studies that were identified for inclusion. All intervention and outcome data were compared for consistency and resolved to 100% agreement with the assistance of the first author as needed.

Assessment of Study Rigor

Study rigor was assessed on an 18-point scale developed by Lundahl et al. (2010) and based on criteria from existing assessment instruments and approaches, such as the Cochrane system. Each study was rated by the first author based on criteria such as number of participants, attrition, quality control, inclusion of the assessment of treatment fidelity, objectivity of measurements, and reporting of follow-up data. Higher scores indicate higher study quality.

Assessment of Risk of Bias

Available information about the included trials was used to assess risk of bias in five domains: random sequence generation (failure to use a specified randomization process), allocation concealment (failure to blind investigators for future randomization assignments), blinding of outcome assessment (failure to conceal outcome assessors from participant intervention group assignment), incomplete outcome data (failure to report complete outcome data from all randomized participants), and selective reporting (failure to report results for all planned outcome data). Guidelines from the Cochrane Handbook were followed for this assessment (Higgins & Green, 2011).
Rating the Quality of Evidence and Strength of Recommendations

The Grading of Recommendations Assessment, Development and Evaluation (GRADE) system was used to evaluate the quality of evidence in this review. With the GRADE system, evidence for specific outcomes is rated across studies, rather than within individual studies. In this approach, RCTs are initially considered to provide high-quality evidence and observational studies are considered to provide low-quality evidence for estimates of intervention effects. Five factors may lead to rating down the quality of evidence (risk of bias, inconsistency, indirectness, imprecision, and publication bias) and three factors may lead to rating up (large effect size, dose–response relationship exists, and consideration of all plausible confounding variables). The quality of evidence for each outcome across studies falls into one of four categories ranging from “high” to “very low.”

Data Analysis

The primary objective of this meta-analysis was to evaluate the efficacy of CBFLIs on adiposity in children and adolescents. To control for change in weight status due to the intervention, the effect size calculated was constructed as a comparison of the change in weight status in the control condition as compared with the intervention condition. Independent effect sizes were calculated for the primary outcome variable (adiposity) in each identified study. To minimize the distortion of standard error estimates resulting from nonindependent effect size estimates (Card, 2012), multiple outcomes derived from the same sample were aggregated such that each study (or group

Figure 1. PRISMA participant flow diagram.
of studies using the same sample) contributed only one overall effect size at each assessment point for the primary outcome variable. For studies reporting multiple outcome measures assessing the same general construct (e.g., BMI, zBMI, percent overweight), an overall effect size was created by averaging the individual effect sizes. If studies reported multiple analyses using subsets of the same sample, the analysis most directly assessing change in children’s weight or adiposity was used.

Effect sizes were calculated following the aforementioned rules (e.g., one effect size per variable per study, etc.).

Hedges’ g (Hedges & Olkin, 1985) was used as the index of standard mean difference between treatment conditions in the current meta-analysis. Hedges’ g is preferred as an index of mean difference when the preponderance of studies in the sample use relatively small sample sizes, with correspondingly greater standard errors (Card, 2012). In studies where efficacy of treatment is represented via other statistics (e.g., F, t, X), these values were converted into Hedge’s g using standard conversion procedures. Cohen (1969) provided guidelines for interpreting the d effect size: small, 0.20–0.49; medium, 0.50–0.79; and large, ≥0.80. Mean effect sizes based on Hedge’s g follow these same conventions. All effect sizes were expressed in terms of 95% confidence intervals (CIs). Confidence bands that did not include zero were considered statistically significant, while those with confidence bands including zero were considered nonsignificant.

Before calculating mean effect sizes, individual effect sizes were weighted to reflect the degree to which the standard error approximates population parameters (Card, 2012). Specifically, a weighted mean effect size was calculated by computing the product of each study’s effect size by its weight, calculating the sum of each individual weighted effect size, and dividing this sum by the sum of weights.

To examine the homogeneity of effect sizes for the primary outcome variable, the Q statistic was used. A significant Q statistic indicates that within-group variability among effect sizes is greater than sampling error alone would predict. This is presumed to signal systematic differences between studies in the sample, perhaps suggesting the presence of moderator variables (Card, 2012). To preserve statistical power for the planned moderation analyses described below, a fixed effect model was used when estimating heterogeneity among the sample of studies (Card, 2012).

Continuous and categorical moderators were assessed using differing methodologies. Continuous variables (e.g., age, duration), modeled separately, were analyzed using weighted regression analyses, with studies with lower standard error contributing more weight in the regression equation. Categorical variables (e.g., gender, setting) were analyzed using analysis of variance techniques to assess differences between groups of interest. It is recommended that ~20 studies per group be included to achieve optimal power to detect group differences (Card, 2012). Therefore, lack of significant differences between groups found in these analyses may be due to insufficient power, as this study only contained 20 studies in total. For this reason, secondary analyses of interactions between moderating variables were not assessed.

To evaluate the degree to which publication bias (i.e., file-drawer problem; Rosenthal, 1979) inflated the resulting effect sizes of this meta-analysis an effect size fail-safe N was calculated (Card, 2012). An effect size fail-safe N indicates the number of unidentified studies with an average effect size that would be necessary to reduce the obtained mean effect size to the smallest meaningful effect size (ESmin). Following Rosenthal (1979), the smallest meaningful effect size was set to 0.1.

Results
Description of Studies

The literature search returned 20 studies that were eligible for inclusion based on the inclusion and exclusion criteria (see PRISMA diagram). Across these studies, 42 effect sizes were calculated to account for multiple time points and multiple outcomes. To prevent violations of independence, effect sizes from a single study were averaged to create an overall average effect size for each study, which was then used to calculate the overall effect size across all studies. Descriptions of study characteristics are presented in Table I.

The majority of the studies were published in the past 10 years, although one study was published in 1985. Of the 20 studies presented here, eight did not provide zBMI data. In these cases, BMI, % body fat, or % overweight data were used and averaged to yield a single estimate of weight change for the study. For length of study period, studies ranged from no data after measurements to follow-up 2 years after baseline. Study rigor was measured using an 18-point scale assessing methodological quality derived from Lundahl et al. (2010). This scale was used because it yielded the most comprehensive assessment of study rigor identified. The average total score on this scale was 12.3, and scores ranged from 9 to 16. The average number of participants per study was 86.28, with a range of 22–108. The total number of child participants across studies was 1,671.
Overall Weight Outcome Effect Size

The overall effect size for change in zBMI in CBFLIs as compared with that in passive control groups over all time points was statistically significant ($g = 0.473$, 95% CI [.362, .584]). This finding is consistent with Cohen’s classification of a “small effect size” (Cohen, 1969). Only one study yielded a negative effect size (Kalarchian et al., 2009), indicating that the rest of the studies yielded improved outcomes as compared with the control group. The effect size of each individual study, as well as each study’s contribution to the overall effect size, is depicted in Figure 2.

Table I. Study Characteristics

<table>
<thead>
<tr>
<th>Study authors</th>
<th>Type of study</th>
<th>Sample size</th>
<th>Child age range</th>
<th>Percent female</th>
<th>Country where study conducted</th>
<th>Race/ethnicity</th>
<th>Duration of Tx (month)</th>
<th>Parent targeted for lifestyle change</th>
<th>Outcome assessment time point(s)</th>
<th>Outcome assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boudreau, Kurowski, Gonzalez, Dimond, &amp; Oreskovic (2013)</td>
<td>2-arm RCT</td>
<td>26</td>
<td>9–12</td>
<td>51.2%</td>
<td>United States</td>
<td>L = 100%</td>
<td>6</td>
<td>No</td>
<td>6</td>
<td>BMI z-score</td>
</tr>
<tr>
<td>Bocca, Corpeleijn, Stolk, &amp; Sauer (2012)</td>
<td>2-arm RCT</td>
<td>75</td>
<td>3–5</td>
<td>72.0%</td>
<td>The Netherlands</td>
<td>N.R.</td>
<td>4</td>
<td>Yes</td>
<td>4, 12</td>
<td>BMI z-score</td>
</tr>
<tr>
<td>Dehar et al. (2012)</td>
<td>2-arm RCT</td>
<td>208</td>
<td>12–17</td>
<td>100%</td>
<td>United States</td>
<td>W = 72%</td>
<td>5</td>
<td>No</td>
<td>6, 12</td>
<td>BMI z-score</td>
</tr>
<tr>
<td>Diaz, Esparza-Romero, Moya-Camarena, Robles-Sardin, &amp; Valencia (2010)</td>
<td>2-arm RCT</td>
<td>43</td>
<td>9–17</td>
<td>51.2%</td>
<td>Mexico</td>
<td>N.R.</td>
<td>6</td>
<td>No</td>
<td>6, 12</td>
<td>BMI z-score</td>
</tr>
<tr>
<td>Golley et al. (2007)</td>
<td>2-arm RCT</td>
<td>74</td>
<td>6–9</td>
<td>63.52%</td>
<td>Australia</td>
<td>W = 98%</td>
<td>5</td>
<td>Yes</td>
<td>12</td>
<td>BMI z-score</td>
</tr>
<tr>
<td>Janicke et al. (2008)</td>
<td>3-arm RCT</td>
<td>71</td>
<td>8–14</td>
<td>60.6%</td>
<td>United States</td>
<td>W = 73%</td>
<td>4</td>
<td>Yes</td>
<td>4, 10</td>
<td>BMI z-score</td>
</tr>
<tr>
<td>Jiang et al. (2005)</td>
<td>2-arm RCT</td>
<td>68</td>
<td>12–15</td>
<td>39.7%</td>
<td>China</td>
<td>N.R.</td>
<td>24</td>
<td>Yes</td>
<td>24</td>
<td>BMI</td>
</tr>
<tr>
<td>Kalarchian et al. (2009)</td>
<td>2-arm RCT</td>
<td>192</td>
<td>8–12</td>
<td>57.0%</td>
<td>United States</td>
<td>W = 74%</td>
<td>12</td>
<td>No</td>
<td>6, 12, 18</td>
<td>% overweight</td>
</tr>
<tr>
<td>Kalavainen, Korppi, &amp; Nuutinen (2007)</td>
<td>2-arm RCT</td>
<td>70</td>
<td>7–9</td>
<td>65.7%</td>
<td>Finland</td>
<td>F = 98.3%</td>
<td>6</td>
<td>Yes</td>
<td>6, 12</td>
<td>BMI z-score</td>
</tr>
<tr>
<td>Nemet et al. (2005)</td>
<td>2-arm RCT</td>
<td>46</td>
<td>6–16</td>
<td>43.5%</td>
<td>Israel</td>
<td>N.R.</td>
<td>3</td>
<td>No</td>
<td>3, 12</td>
<td>BMI, BMI % body fat</td>
</tr>
<tr>
<td>Nemet, Barzilay-Teeni, &amp; Elaluk (2008)</td>
<td>2-arm RCT</td>
<td>22</td>
<td>8–11</td>
<td>63.6%</td>
<td>Israel</td>
<td>N.R.</td>
<td>3</td>
<td>No</td>
<td>3</td>
<td>BMI, body fat</td>
</tr>
<tr>
<td>Reinhr et al. (2010)</td>
<td>2-arm RCT</td>
<td>66</td>
<td>8–16</td>
<td>60.6%</td>
<td>Germany</td>
<td>N.R.</td>
<td>6</td>
<td>N.R.</td>
<td>6</td>
<td>BMI z-score</td>
</tr>
<tr>
<td>Sacher et al. (2008)</td>
<td>2-arm RCT</td>
<td>116</td>
<td>8–12</td>
<td>54.3%</td>
<td>United Kingdom</td>
<td>W = 50%</td>
<td>12</td>
<td>No</td>
<td>6</td>
<td>BMI z-score</td>
</tr>
<tr>
<td>Savoye et al. (2007)</td>
<td>2-arm RCT</td>
<td>174</td>
<td>8–16</td>
<td>60.3%</td>
<td>United States</td>
<td>W = 37%</td>
<td>12</td>
<td>No</td>
<td>6, 12, 24</td>
<td>Body fat</td>
</tr>
<tr>
<td>Senedlak &amp; Spence (1985)</td>
<td>3-arm RCT</td>
<td>33</td>
<td>6–13</td>
<td>N.R.</td>
<td>United States</td>
<td>N.R.</td>
<td>1</td>
<td>No</td>
<td>1</td>
<td>% overweight</td>
</tr>
<tr>
<td>Shelton et al. (2007)</td>
<td>2-arm RCT</td>
<td>43</td>
<td>3–10</td>
<td>53.3%</td>
<td>Australia</td>
<td>N.R.</td>
<td>1</td>
<td>Yes</td>
<td>3</td>
<td>BMI and WC</td>
</tr>
<tr>
<td>Vos, Huisman, Houdjik, Pijl, &amp; Wit (2012)</td>
<td>2-arm RCT</td>
<td>79</td>
<td>8–17</td>
<td>53.2%</td>
<td>The Netherlands</td>
<td>N.R.</td>
<td>3</td>
<td>No</td>
<td>3, 12</td>
<td>BMI z-score</td>
</tr>
<tr>
<td>Wafa et al. (2011)</td>
<td>2-arm RCT</td>
<td>107</td>
<td>7–11</td>
<td>53.5%</td>
<td>Malaysia</td>
<td>N.R.</td>
<td>6</td>
<td>No</td>
<td>6</td>
<td>BMI z-score</td>
</tr>
<tr>
<td>West, Sanders, Cleghorn, &amp; Davies (2010)</td>
<td>2-arm RCT</td>
<td>101</td>
<td>4–11</td>
<td>67.3%</td>
<td>Australia</td>
<td>W = 87%</td>
<td>3</td>
<td>No</td>
<td>6</td>
<td>BMI z-score</td>
</tr>
<tr>
<td>Williamson et al. (2005)</td>
<td>2-arm RCT</td>
<td>57</td>
<td>11–15</td>
<td>100%</td>
<td>United States</td>
<td>AA = 100%</td>
<td>6</td>
<td>Yes</td>
<td>6, 24</td>
<td>BMI, body fat</td>
</tr>
</tbody>
</table>

Note. AA = African American; A = Asian; F = Finnish; H = Hispanic; I = Indigenous; L = Latina; M = Mediterranean; O = other; W = White; WC = waist circumference; N.R. = not reported; RCT = randomized controlled trial.

The overall effect size for change in zBMI in CBFLIs as compared with that in passive control groups over all time points was statistically significant ($g = 0.473$, 95% CI [.362, .584]). This finding is consistent with Cohen’s classification of a “small effect size” (Cohen, 1969). Only one study yielded a negative effect size (Kalarchian et al., 2009), indicating that the rest of the studies yielded improved outcomes as compared with the control group. The effect size of each individual study, as well as each study’s contribution to the overall effect size, is depicted in Figure 2.
Overall Effect Sizes for Secondary Outcomes

Eight studies provided usable information on child caloric intake. The overall effect size for change in caloric intake in CBFLIs as compared with that in passive control groups over all time points was not statistically significant ($g = 0.086, 95\% \text{ CI } [-0.090, 0.263]$).

There was an insufficient number of studies (e.g., less than two) meeting inclusion and exclusion criteria for this review that reported on physical activity, sedentary behavior, and parent use of behavioral management strategies to allow for the calculation of an effect size based on the Cochrane review guidelines.

Posttreatment and Follow-up Assessment Weight Outcome Effect Size

Secondary analysis separately examined weight outcomes directly following completion of the intervention (posttreatment) and outcomes at later follow-up assessments. Based on 20 studies, the effect size for CBFLIs at posttreatment as compared with control groups was statistically significant ($g = 0.416, 95\% \text{ CI } [.307, .526]$). Furthermore, based on 11 studies, the effect size for CBFLIs at follow-up assessment as compared with control groups was also statistically significant ($g = 0.386, 95\% \text{ CI } [.248, .524]$). Both of these meet criteria for a “small effect size” (Cohen, 1969). The difference between the effect size for posttreatment assessments and follow-up assessments was not statistically significant.

Moderator Analyses

A test of heterogeneity was then conducted to determine the appropriateness of testing for moderating variables. The Q test of heterogeneity was statistically significant (84.446, $p < .001$), indicating that these studies are heterogeneous.
and that study variables are significantly moderating the study effect sizes. Therefore, analyses, with each moderator variable modeled separately, were conducted to assess study characteristics that may be influencing effect size outcomes. Moderating variables were included based on reviews and recommendations for treating pediatric obesity, previous research, and their potential of representing important directions for future intervention research (Faith et al., 2012; Jelalian & Saelens, 1999; Spear et al., 2007; Wilfley et al., 2007). Continuous moderator variables examined in this meta-analysis included time between baseline and posttreatment assessment, duration of treatment, number of sessions, child age, child sex (i.e., % female), methodological rigor of study, and amount of time in treatment for both child and parent. Duration of treatment, number of sessions, child age, and the amount of time that the child spent in treatment were all statistically significant moderators. Specifically, larger effect sizes were associated with longer treatment (duration in weeks), more treatment sessions, a greater total amount of time (over the course of treatment) that the child spent in treatment, and greater child age. The results of continuous moderating analyses are presented in Table II.

Categorical moderators analyzed in this meta-analysis included whether the study sample used ITT analysis, type of control group, use of a manualized treatment, BMI greater than the 99th percentile as an exclusionary criteria, if the parents were targeted for their own health behavior change, format for therapy, and form of contact with participants. Results from these analyses are presented in Table III. Use of ITT analysis, format of therapy, and form of contact were significant moderating variables, such that individual and in-person CBFLIs were associated with larger effect sizes.

### Table II. Continuous Moderator Analyses

<table>
<thead>
<tr>
<th>Moderator variables</th>
<th>$R^2$ value</th>
<th>p value</th>
<th>Beta</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time between baseline and assessment</td>
<td>.043</td>
<td>.198</td>
<td>0.208</td>
</tr>
<tr>
<td>Duration of intervention</td>
<td>.374</td>
<td>.004</td>
<td>0.612</td>
</tr>
<tr>
<td>Number of intervention sessions</td>
<td>.264</td>
<td>.020</td>
<td>0.514</td>
</tr>
<tr>
<td>Percentage of females participants in intervention</td>
<td>.068</td>
<td>.281</td>
<td>-0.234</td>
</tr>
<tr>
<td>Age of child participants</td>
<td>.305</td>
<td>.014</td>
<td>0.169</td>
</tr>
<tr>
<td>Methodological rigor of study</td>
<td>.003</td>
<td>.831</td>
<td>-0.193</td>
</tr>
<tr>
<td>Amount of minutes child in treatment</td>
<td>.795</td>
<td>.000</td>
<td>0.892</td>
</tr>
<tr>
<td>Amount of minutes parent in treatment</td>
<td>.017</td>
<td>.603</td>
<td>-0.131</td>
</tr>
</tbody>
</table>

### Effect Size Fail-Safe N Calculation

A fail-safe N was calculated to help evaluate the likelihood of our overall effect size being an overly positive representation of the true effect size. This calculation assesses the number of studies with an effect size of zero that would have to be included before the overall effect size would shrink to the smallest meaningful effect size. Our calculations indicated that ~75 studies with an effect size of zero would have to be added to this meta-analysis to bring down the overall effect size to 0.1.

### Risk of Bias

Each study was assessed on five dimensions for risk of bias. The summary risk of bias findings across all studies are presented in Figure 3. Overall the 20 studies included in this review exhibit relatively low risk of bias for random sequence generation and selective reporting. Over half of the studies examined were judged to exhibit unclear or high risk of bias for blinding of outcome assessment. In addition, over half of the studies reviewed reported insufficient methodological details and outcome data to determine if there was low or high risk of bias for incomplete outcome data and thus were classified as unclear risk of bias. Finally, over half of the studies examined exhibited high risk of bias for allocation concealment.

### GRADE Rating

A rating of “moderate” (3 on a scale of 1–4) was assigned to the quality of evidence for improvements in BMI Z-score based on the GRADE system. As all trials were RCTs, our review assumed a starting point of high-quality evidence (a rating of 4). However, the quality of evidence was inconsistent and there are some potential concerns with risk for bias (see previous section), which lowered the overall rating from high quality to moderate quality. A rating of “low quality” (2 on a scale of 1–4) was assigned to the quality of evidence for changes in caloric intake based on the GRADE system. As all trials were RCTs, our review assumed a starting point of high-quality evidence (a rating of 4). However, owing to the potential concerns for risk of bias and the fact that the measurement of dietary intake was viewed as imprecise, the GRADE rating was reduced from high quality to low quality. See Table IV for a display of the Summary of Findings Table.

### Discussion

Previous reviews examining studies published before 2008 have reported that lifestyle interventions addressing pediatric obesity are associated with improvements in weight
outcomes (Jelalian & Saelens, 1999; Kitzmann et al., 2010; Oude-Luttikhuis et al., 2009; Whitlock et al., 2010; Wilfley et al., 2007). While very informative, these reviews included heterogeneous groups of studies with varied types of lifestyle intervention programs. The current meta-analysis extends the literature by examining studies published in 2013 and being the first to examine only multi-component CBFLIs to treat obesity in children and adolescents who are overweight or obese. Specifically, our inclusion criteria required that intervention programs address three components, improvements in dietary intake, increases in physical activity, and the use of behavioral strategies to achieve improvements in these two areas. The overall effect size of 0.47 for change in BMI z-score, although small, was approaching Cohen’s requirement of 0.5 for a medium effect size. This finding is consistent with similar past reviews (Whitlock et al., 2010; Wilfley et al., 2007) and suggests that multi-component CBFLIs are efficacious in improving weight outcomes in children who are overweight or obese. Follow-up analysis found that effect sizes did not significantly differ between results at posttreatment and follow-up assessments, providing evidence that improvements in weight outcomes are maintained at follow-up. The length of follow-up assessments ranged from 10 to 24 months after baseline, with a majority (8 of 10 studies) including 12-month follow-up assessment. The analyses of effect sizes examining maintenance of outcomes at follow-up assessments were based on a relatively small number of studies. Thus, results should be interpreted with caution.

Exploratory analyses were also conducted to examine potential moderators of treatment outcome. Most notably dose of treatment, as measured by the number of intervention sessions and the number of minutes children spent in treatment sessions, was positively related to effect size. This is consistent with previous reviews (Whitlock et al., 2010; Wilfley et al., 2007). Our results, in combination with these previous reviews, provide support for the notion that greater intervention duration and intensity are associated with better weight outcomes, which is similar to findings reported in the adult obesity literature (Middleton, Patidar, & Perri, 2012). However, more research specifically examining dose–response relationships of CBFLIs to weight and behavioral outcomes is needed before definitive conclusions can be drawn.

<table>
<thead>
<tr>
<th>Moderator variables</th>
<th>Categories</th>
<th>N</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intent to treata</td>
<td>Yes (1)</td>
<td>7</td>
<td>0.658 (0.491, 0.824)</td>
</tr>
<tr>
<td></td>
<td>No (2)</td>
<td>13</td>
<td>0.290 (0.141, 0.440)</td>
</tr>
<tr>
<td>Type of control group</td>
<td>Standard care</td>
<td>10</td>
<td>0.430 (0.280, 0.572)</td>
</tr>
<tr>
<td></td>
<td>Waitlist</td>
<td>9</td>
<td>0.470 (0.290, 0.650)</td>
</tr>
<tr>
<td>Use of manualized treatment</td>
<td>Yes</td>
<td>5</td>
<td>0.546 (0.314, 0.747)</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>7</td>
<td>0.360 (0.115, 0.605)</td>
</tr>
<tr>
<td></td>
<td>Not reported</td>
<td>8</td>
<td>0.437 (0.279, 0.596)</td>
</tr>
<tr>
<td>BMI &gt; 99th as exclusion</td>
<td>Yes (1)</td>
<td>4</td>
<td>0.352 (0.279, 0.596)</td>
</tr>
<tr>
<td></td>
<td>No (2)</td>
<td>11</td>
<td>0.392 (0.239, 0.546)</td>
</tr>
<tr>
<td></td>
<td>Not reported</td>
<td>5</td>
<td>0.812 (0.546, 1.075)</td>
</tr>
<tr>
<td>Parent targeted for change</td>
<td>Yes (1)</td>
<td>6</td>
<td>0.268 (0.049, 0.486)</td>
</tr>
<tr>
<td></td>
<td>No (2)</td>
<td>12</td>
<td>0.502 (0.366, 0.638)</td>
</tr>
<tr>
<td></td>
<td>Not reported</td>
<td>2</td>
<td>0.689 (0.265, 1.113)</td>
</tr>
<tr>
<td>Form of therapya</td>
<td>Individual</td>
<td>2</td>
<td>1.291 (0.889, 1.693)</td>
</tr>
<tr>
<td></td>
<td>Group</td>
<td>13</td>
<td>0.372 (0.240, 0.504)</td>
</tr>
<tr>
<td></td>
<td>Both</td>
<td>5</td>
<td>0.426 (0.240, 0.665)</td>
</tr>
<tr>
<td>Form of contacta</td>
<td>In person</td>
<td>14</td>
<td>0.642 (0.498, 0.786)</td>
</tr>
<tr>
<td></td>
<td>Phone only</td>
<td>1</td>
<td>0.219 (–0.363, 0.801)</td>
</tr>
<tr>
<td></td>
<td>In person+Other</td>
<td>5</td>
<td>0.176 (–0.007, 0.359)</td>
</tr>
<tr>
<td>Timing of measurement</td>
<td>Post-Treatment</td>
<td>20</td>
<td>0.416 (0.307, 0.526)</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td>11</td>
<td>0.386 (0.248, 0.524)</td>
</tr>
</tbody>
</table>

Note. aIndicates statistically significant moderating variable.

bDue to multiple measurements taken within the same study, this number does not represent studies, instead representing measurements taken within a study.

Figure 3. Risk of bias.
Table IV. Summary of Findings for Multi-Component Behavioral Family Interventions to Treat Obesity in Children and Adolescents

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Patient or population: Overweight and obese children and adolescents</th>
<th>Illustrative comparative risks (95% CI)</th>
<th>Relative effect (95% confidence interval)</th>
<th>No of participants (studies)</th>
<th>Quality of the evidence (GRADE)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMI z-score (overall)</td>
<td>Investigators used objective measurement of anthropometric variables.</td>
<td>Assumed risk: Control, Behavior Family Therapy</td>
<td>The overall mean decrease in BMI z-score in the intervention groups was 0.47 standard deviations more than the control groups (0.36–0.58 lower)</td>
<td>-</td>
<td>1671 (20 studies)</td>
<td>moderate</td>
</tr>
<tr>
<td>Caloric Intake</td>
<td>Investigators used objective measurement of caloric intake (i.e., 24-hr recalls, food frequency questionnaires).</td>
<td></td>
<td>The overall mean decrease in caloric intake in the intervention groups was 0.09 standard deviations more than the control groups (&lt;0.09–0.26 lower)</td>
<td>-</td>
<td>542 (8 studies)</td>
<td>low quality</td>
</tr>
</tbody>
</table>

Note: GRADE Working Group grades of evidence.

High quality (+++): Further research is very unlikely to change our confidence in the estimate of effect.

Moderate quality (+++): Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low quality (+++): Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

Very low quality (+++): We are very uncertain about the estimate.

The level of parent involvement relative to weight outcomes in pediatric lifestyle interventions has received growing attention in the literature in recent years. Neither the amount of time parents spent in treatment nor whether the parents were targeted for their own lifestyle behavior change was related to child weight outcomes in this analysis. This is consistent with a review by Faith and colleagues (2012) who found limited support for the notion that greater parental involvement in treatment leads to better child weight outcomes. However, Faith and colleagues did find that greater parent adherence to core behavioral strategies was related to greater child weight status outcomes. It may be that the actual application of behavioral strategies to address lifestyle issues within the home leads to better child weight outcomes, which was not assessed in our analysis. As noted by Faith et al. (2012), only a few studies have reported on this important relationship (Epstein, 1993; Wrotniak, Epstein, Paluch, & Roemmich, 2005). Future intervention trials should include data collection to examine if parent use of behavioral strategies mediates the relationship between treatment participation and key outcomes such as changes in dietary intake, physical activity, and weight status.

A number of other potential moderators of effect size outcomes were examined in this review. Age was a moderator, such that larger effect sizes for weight outcomes were positively associated with the mean age of child participants. This is consistent with a review by Oude-Luttikhuis (2009) who found larger effects for studies with children ≥12 years of age compared with studies with children <12 years of age. Weight outcomes were moderated by form of therapy (individual vs. group) and form of contact (in person vs. phone only vs. in-person plus other), such that therapy with individual families and therapy delivered via in-person contacts related to larger effect sizes. However, the number of studies included in this moderator analyses were heavily skewed toward group and in-person contacts that these findings must be viewed as preliminary. The current review also found that studies that used ITT analysis (k = 7) were associated with better weight outcomes than studies that did not use ITT analysis. While one could speculate that studies including ITT analysis are methodologically stronger than those studies that do not include ITT analysis, our moderator analysis examining methodological rigor of included studies was not significant and thus did not support this conclusion.

Data from eight studies found that CBFLIs do not have significant effects on change in caloric intake. This is surprising given the effects that CBFLIs were found to have on weight change. It is likely that the imprecise nature of assessing caloric intake at least partially impacted these results (Schoeller, 1995), as it seems unlikely that the weight changes observed were entirely due to increases in energy expenditure. These findings highlight the need for innovative future research designs to examine the specific...
mechanisms by which the interventions exert their influence. One such example may be the use of sequential, multiple assignment, randomized trials (SMART Designs) that provide data to enable assessment of optimal component sequencing that can facilitate the development of high-quality adaptive interventions (Almirall, Compton, Gunlicks-Stoessel, Duan, & Murphy, 2012).

Consistent with the other systematic reviews in this special issue, the GRADE system (Guyatt et al., 2011) was used to assess the strength and quality of the evidence of our findings. Evidence for changes in BMI z-score from the reviewed studies was assigned a “moderate” GRADE rating, indicating that “further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.” However, evidence for our finding on caloric intake was assigned a “low quality” rating, indicating that further research is likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate. Characteristics of the studies included in this meta-analysis that strengthened the GRADE rating for weight outcomes included the use of RCTs, evidence of a dose–response relationship between weight and treatment (e.g., better weight outcomes were associated with greater treatment duration and intensity), and an overall small effect size approaching a medium effect. There were, however, also some weaknesses in this literature, which precluded a stronger GRADE rating for each outcome. For example, 75% of studies were rated as having unclear or high risk of bias of incomplete reporting of outcome data due to large amounts of missing data, insufficient details regarding reasons for missing data, or the use of inadequate methods for managing missing data, which diminishes confidence in estimation of treatment effects. Moreover, half the studies included in this analysis were rated high on risk of bias for allocation concealment (i.e., investigators were not blind to randomized assignment of participants). In addition, effect sizes across studies for weight outcomes were somewhat inconsistent. Finally, as noted previously, the imprecise nature of assessing caloric intake negatively impacted our Grade rating for changes in caloric intake.

There are a variety of steps that can be taken by future research to address these methodological limitations and improve the scientific rigor in this area. First, many studies were excluded from the meta-analysis owing to failure to provide sufficient information to allow for calculation of effect size, or properly classify the study intervention or methodology. Similar concerns have been noted by Klesges and colleagues, who called for researchers to improve the reporting of contextual and generalizability elements central to translational research (Klesges, Dzewaltowski, & Glasgow, 2008). We echo this call and also encourage researchers to more clearly operationally define and describe both the active intervention and control condition on which the intervention is tested against. Second, the overall GRADE rating for this analysis was negatively impacted by the high risk of bias in a number of areas. To reduce this risk of potential bias and increase our confidence in the efficacy of CBFLIs, it is also important for future studies to ensure proper concealment of randomization and treatment condition assignment to investigators and assessment staff. With regards to data analysis, future studies should conduct ITT analyses when examining primary and secondary outcomes, and should clearly document why data are missing and what strategies and assumptions used for handling missing data.

There are some important additional limitations of the current analysis that have implications for future research. The clinical significance of weight outcomes associated with CBFLIs is an ongoing question. Few studies in this review reported on metabolic parameters associated with weight changes. While these data are admittedly difficult to gather with children, it will be critical to document improvements not only in weight outcomes but also metabolic parameters. Similarly, only eight studies in the literature contained sufficient information to examine change in dietary intake. Moreover, there were not enough articles included to calculate effect sizes for both physical activity and sedentary behavior. Future studies of lifestyle interventions for obesity should include sufficient objective information to allow effect sizes to be calculated for dietary intake, physical activity, and sedentary behavior, as these effect sizes likely moderate overall effect sizes for lifestyle interventions. Admittedly, reliable and valid assessment of dietary intake is an ongoing challenge. Repeated 24-hr recalls offer the best combination of feasibility and strong psychometric properties, but are expensive, time-consuming, and still subject to self-reporting bias. Strategies to reduce reporting bias and further improve the reliability and validity of assessing dietary intake are sorely needed. In addition, compared with other meta-analyses, only a relatively small number of studies and participants were included in this meta-analysis. As such, findings from our analysis of moderators should be viewed with caution. Finally, there is the risk of publication bias in any meta-analysis; the discovery of additional studies of CBFLIs might have affected the results obtained here. However, the fail-safe calculation indicated that ~75 studies with an effect size of 0 would need to be added to this meta-analysis to result in an effect size of 0.1.

The potential for adverse events is important to consider when evaluating interventions to address childhood
obesity, not only in terms of the potential for delayed linear growth but also for negative outcomes such as the development of disordered eating behaviors and reductions in self-esteem (Satter, 2005). Eight of the 20 studies included in this analysis reported on adverse events. Only four of these reported on potential changes in problematic eating attitudes and behaviors or self-esteem (DeBar et al., 2012; Janicke et al., 2008; Jiang, Xia, Greiner, Lian, & Rosenqvist, 2005; Williamson et al., 2005). One study specifically assessed for slowed linear growth (Golley, Magarey, Baur, Steinbeck, & Daniels, 2007), while three other studies reported on “generic adverse events” (Nemet et al., 2005; Reinhler et al., 2010; Sacher et al., 2010). Across these eight studies there was no data suggestive of an increase in negative outcomes or adverse events subsequent to participating in CBFLIs. This is consistent with the findings from previous reviews (Carter & Bulik, 2008; Oude-Luttikhuis et al., 2009; Whitlock et al., 2010). While testing the efficacy or effectiveness of CBFLIs on reducing or slowing weight gain in youth has important public health implications, it is equally important that families participating in these interventions are not exposed to additional causes of harm. Therefore, researchers should place a high priority on reporting information on the number and types of adverse events associated with tested interventions.

Given the lack of interventions addressing obesity in young children, as well as recent findings showing that children who are overweight at 5 years of age are four times more likely to be obese between 5 and 14 years of age compared with children who are normal weight (Cunningham, Kramer, & Narayan, 2014), there is great need for more prevention and intervention research focusing on the impact of behavioral family interventions in families of young children. Another important question is the issue of generalizability, as well as translation and dissemination to real world settings. There is growing support for the notion that longer and more intense treatment results in better weight outcomes. However, the question arise as to how can higher intensity interventions that seem to be related to better outcomes be delivered in real-world settings by individuals with less expertise than many of the obesity experts that served as interventionists in these outcome studies? Moreover, how can we help families sustain participation for the longer durations necessary for better weight outcomes, given the documented barriers that many families face in real world settings (Kottyan, Kottyan, Edwards, & Unaka, 2014; Lim & Janicke, 2013)? Future research is needed to develop and examine innovative strategies for implementing CBFLIs that can be delivered with fidelity, reach large numbers of families in community settings, and that are feasible for both families and providers. Telehealth, technology-aided, and internet-based social media interventions are areas that hold promise and are deservedly garnering more attention (Davis, Sampilo, Gallagher, Landrum, & Malone, 2013). Finally, developing broader multi-layered intervention efforts that include CBFLIs within larger community change efforts, such as by Rogers and colleagues (Rogers et al., 2013), can build awareness, as well as environmental and social supports that have the potential to facilitate long-term behavior change that can reach larger numbers of children and families.

Conclusion

The current study provides evidence that CBFLIs addressing pediatric obesity lead to improvements in child weight outcomes. While the overall effect size was small, it approached the medium range. There was no difference in effect sizes for weight outcomes at posttreatment relative to long-term follow-up. Greater duration and intensity of treatment, as well as greater child age, were all related to better weight outcomes. Future research is needed to better document changes in caloric intake, physical activity, and metabolic parameters associated with participation in CBFLIs. Future research is also needed to develop innovative strategies for feasibly implementing CBFLIs in real-world settings while preserving the core components of treatment with established efficacy.

Conflicts of interest: None declared.

References


