OBJECTIVE Although the primary care setting has been recommended as an acceptable environment for pediatric overweight/obesity treatment, a quantitative analysis has not been conducted to determine the effectiveness of pediatric weight management interventions delivered in these settings. Therefore, the purpose of the current study was to conduct a meta-analysis of weight management interventions for youth in primary care settings. Method A literature search using PsycINFO and PubMed was conducted to identify articles published through October 2015. Eighteen studies (3,358 participants) met inclusion criteria; studies included a treatment and comparison group and targeted individuals or families for treatment. Study characteristics were coded, and study rigor of articles was assessed. Results The overall effect size for change in BMI in primary care weight management interventions compared to control groups was small but statistically significant (d = 0.26, 95% CI [0.14, 0.38]). The number of treatment contacts, treatment months, and visits with a pediatrician emerged as significant moderators of outcome, such that BMI reduction was positively related to greater contact. Conclusions In comparison to control conditions, weight management programs in primary care settings can be effective for BMI reduction, suggesting that primary care is a suitable setting for treatment of pediatric overweight/obesity. Additionally, treatments that were longer in duration, included more contacts (in person or phone), and included more contacts by a pediatrician had greater impact on BMI reduction. Future studies should continue to examine other aspects of acceptability and accessibility as well as demonstrate the effectiveness of interventions on improving psychosocial outcomes.

Accepted for publication in Health Psychology (APA, 2016). doi: 10.1037/hea0000381
familiar location (Barlow & Ohlemeyer, 2006; Brennan, Walkley, & Wilks, 2012; Jensen, Aylward, & Steele, 2012). For example, Barlow and Ohlemeyer (2006) questioned families regarding their reasons for dropping out of the pediatric weight management program and found that 25% of the families rated time and location as the most important barriers for treatment. Similarly, Kitscha, Brunet, Farmer, & Mager (2009) conducted a qualitative study to identify reasons for attrition from a pediatric weight management program and found that a major reason was physical barriers, including location.

One setting that may be more accessible and acceptable for families is the child’s primary care office. As defined by the Institute of Medicine (IOM; 1994), “primary care is the provision of integrated, accessible health care services by clinicians who are accountable for addressing a large majority of personal health care needs, developing a sustained partnership with patients, and practicing in the context of family and community” (p. 15). The IOM does not define primary care by the profession of the clinician, but notes that the clinician is someone who has a significant scientific knowledge-base and has direct contact with the patient. Examples provided by the IOM (1994) include physicians, physical therapists, nutritionists, and social workers. Implicit in this definition is the philosophy that primary care represents a partnership between the patient and the clinician and that it is linked to and located in the communities in which the patients live, which increases the acceptability and accessibility of the practice. Wald, Moyer, Eickhoff, and Ewing (2011) noted that, beyond accessibility and acceptability, the frequency with which children visit their primary care office makes this an ideal environment to influence the health of large numbers of children with overweight and obesity. Indeed, the Expert Committee (2007) and the USPSTF (2010) recommended primary care as a suitable setting for pediatric weight management interventions; more specifically, they recommended that primary care clinicians assess children for overweight and obesity and offer or refer them to behavioral lifestyle interventions to promote reductions in weight status.

The available empirical research to date suggests that primary care offices can be effective settings for weight management interventions. Summarizing the available literature to date, Sargent, Pilotto, and Baur (2011) conducted a systematic review of primary care interventions for pediatric obesity. Of the 18 articles reviewed, 12 reported significant effects for reducing BMI; in particular, the authors found that low-intensity interventions delivered by a physician were particularly successful. Sargent et al. (2011) also concluded that interventions were more effective when they focused on at least two targets of behavior change (e.g., reducing fat and sugar in diet, increasing physical activity). Unfortunately, however, the literature currently provides no overall quantitative estimate of the degree to which interventions based in primary care are effective. Thus, in addition to updating the qualitative review provided by Sargent et al., the current meta-analysis provides a standardized effect size to quantify the effectiveness of the weight management interventions in primary care settings compared to active, education, and passive control conditions.

A secondary aim of the current study is to examine potential moderators of treatment effects in the primary care setting. Previous studies have provided empirical support for several potential moderators. For example, Janicke and colleagues (2014) found that interventions with greater intensity and duration were associated with better weight outcomes. In addition, better weight outcomes have been shown to be related to male gender (Jelalian et al., 2008) and greater child age (Janicke et al., 2014). Thus, potential moderators of effect size that were examined include study design (type of control, study rigor, length of follow-up), treatment characteristics (number of sessions, number of treatment months, and number of visits with pediatrician), and participant characteristics (child gender and age).

METHODS

Literature Search

Relevant literature was identified in three ways. First, comprehensive literature searches were conducted using two electronic databases (i.e., PsycINFO, PUBMED). To identify articles on pediatric weight management interventions in primary care settings, the following full keyword and abbreviated search terms were used: “weight management or healthy lifestyle” and “intervention or treatment or control or RCT or comparison” and “primary care or doctor or pediatrician” and “child* or youth or family or adolescent” and “obes* or overweight or BMI or weight.” Although behavioral interventions have the most empirical support for pediatric obesity (see Janicke et al., 2014; Kitzmann et al., 2010), the term “behavioral” was not included in the search terms for the current study in order to be inclusive of all interventions offered in primary care settings, and to be able to most accurately summarize the components of the interventions in the current literature. Next, backward and forward searches were conducted to identify relevant articles that were cited within the articles included in the initial search and relevant articles that cited the articles identified in the
initial search, respectively. Lastly, relevant unpublished data were identified by searching ProQuest Dissertations and Theses and by contacting first authors of the included studies. After excluding duplicates, 2,280 articles were screened for potential inclusion in the current meta-analysis during the summer of 2014 and fall of 2015. The cutoff for inclusion in the current analysis was a publication date on or before October 13, 2015.

**Inclusion/Exclusion Criteria**

Of the articles and abstracts screened, articles were retained for the meta-analysis if they met the following criteria: (1) involved a weight management program in a primary care setting; (2) mean age of participants was between 2-18 years old at the start of the intervention; (3) mean BMI of participants was at or above the 85th percentile; (4) intervention included a treatment and comparison group; (5) primary outcome was child BMI, BMI z-score, or BMI percentile; (6) were reported in English; and (7) reported enough information about results to calculate an effect size for the outcomes. Studies from any country as well as published and unpublished studies were included in the current meta-analysis. Because primary care involves a range of clinicians (see IOM, 1994 definition), the current study did not specify the precise profession of clinician for inclusion in the current analyses; this approach allowed for an accurate summary of the current literature and the examination of related moderators that may provide evidence for what aspects of primary care settings are unique and most effective at promoting weight outcomes. Studies were excluded based on the following criteria: (1) primary outcomes only consisted of mental health concerns (e.g., health-related quality of life); (2) intervention included drug trials and/or surgery; (3) the intervention was targeted towards individuals with a chronic illnesses other than obesity (e.g., diabetes); and (4) the study was a prevention study.

Articles and abstracts were screened for potential inclusion by the first two authors, and all discrepancies were discussed. Based on the initial search, 2,280 titles and abstracts were screened using the inclusion and exclusion criteria; 88 articles were retained for further review. The second phase of screening, which included reading the entire article and applying criteria, resulted in a total of 18 articles that were eligible for the current study (see Figure 1).

**Data Extraction**

**Coding of study characteristics.** The first two authors each coded the 18 articles separately, using a coding procedure specifically established for this investigation. Discrepancies in data extraction were discussed until a decision was reached; therefore, all data was resolved to 100% agreement between the first two authors.

First, study participant characteristics were coded including child age (i.e., range, mean), percentage of the sample that was female, and the percentage of the sample that was non-Caucasian. In addition, study design characteristics (i.e., whether parents were targeted for change, who delivered the intervention, use of a manualized treatment, inclusion of a power analysis, intent-to-treat analysis, length of follow-up, country of intervention) were coded. To further examine design characteristics, study rigor was assessed using an 18-point rating scale, developed and used by Lundahl and colleagues (2010), in which higher scores were indicative of higher study quality. In accordance with the procedures described by Lundahl et al. (2010), the first two authors rated each study based on criteria such as number of participants, inclusion of a treatment fidelity measure, and use of objective measures. See online supplement for the full study rigor rating scale.

Intervention characteristics were coded to reflect the type of comparison group (active, education, or passive) and type of treatment group (e.g., diet, physical activity, screen time, disordered eating). Furthermore, the number of sessions, total treatment months, and number of meetings with a pediatrician were coded. BMI, BMI-z scores (zBMI), and/or BMI percentiles were used to calculate effect sizes.

**Coding of effect sizes.** The primary aim of the current meta-analysis was to evaluate the effectiveness of weight management interventions for overweight and obese youth in primary care settings. All of the studies identified in this search utilized designs in which participants were assigned to a treatment or control group, and participants’ BMIs were measured both before and after the intervention. Because of these design elements, effect sizes were coded according to Morris’ (2008) calculation for pretest-posttest-control designs. To take full advantage of the data provided in each study, an effect size was calculated by examining the pre-post change in weight status in the treatment group minus the pre-post change in weight status in the control group, divided by the pooled standard deviation for pre-intervention (Morris, 2008). The first two authors calculated effect sizes for all studies using this formula, regardless of whether the study reported an effect size. Positive effect sizes represent greater reductions in the outcome variables.

For studies that had two experimental groups, two separate effect sizes were calculated using the same control group. Similarly, for studies that reported
multiple outcome measures (e.g., BMI z-score, BMI percentile, BMI) or multiple time-points (e.g., pre-post, pre-follow up), individual effect sizes were first calculated. To limit interdependency of effect sizes, multiple effect sizes were then averaged to result in one effect size per study (Card, 2012).

**Statistical Analysis**

The weighted random-effects mean effect size was calculated based on the set of individual effect sizes. Random effects models assume that the variability in effect sizes is due to both sampling error and variability at the population level. In addition, the random-effects model allows inferences to be made about the effect size of a population of studies larger than only the studies included in the current analysis (Card, 2012; Cheung, 2008).

As suggested by Card (2012), the heterogeneity of effect sizes across studies (Q) was computed first. Next, the population variability was estimated in order to separate the observed heterogeneity into that due to sampling error and that due to true variability in population effect sizes. After estimating the population variability, a new, random effect weight for each study was calculated. Instead of weighting each study by the inverse of the sampling variance, the random effects model weights each study by the inverse of the sampling variance plus the variability across the population effects (Card, 2012; Cheung, 2008). The mean effect size was estimated using this new weighted value.

Next, moderator analyses were conducted using mixed-effects models. Mixed-effects models allow for the examination of fixed-effects moderators in the presence of random-effects heterogeneity (Card, 2012). A structural equation modeling approach (using Mplus) was used to estimate mixed-effects models (Cheung, 2008). To do this, a random-effects model was built by transforming the intercept and using it to predict the slope of the transformed effect size; the slope was allowed to vary randomly (Card, 2012). Each moderator was analyzed separately due to the small number of studies. Because some analyses had very small number of studies, conclusions should be interpreted with caution (Card, 2012).

Lastly, a fail-safe number was calculated to address the common concern of publication bias in meta-analyses. Publication bias is a concern because researchers are less likely to report non-significant results than significant results, and non-significant results are less likely to be published. The fail-safe number indicates the number of studies with an effect size of zero that could be added to the analysis before the overall mean effect size dropped below the minimally significant effect size (Card, 2012).

**RESULTS**

The current analysis included 18 studies, with a total of 43 effect sizes calculated due to the studies including multiple time points, experimental groups, and outcomes. As noted above, multiple individual effect sizes from each study were averaged into a single effect size to limit violations of independence. See Table 1 for a description of study characteristics.

The studies included in the current analysis were published between 2002 and 2015. Seventy-seven percent of the studies were conducted within the last five years, which indicates that research in this field is relatively new and rapidly growing. Eleven studies were conducted in the United States, two were conducted in Mexico, and two were conducted in Australia; the remaining three studies were conducted in Sweden, Italy, and the United Kingdom.

The mean number of participants included in the studies was 186.56 with sample sizes ranging from 22 to 645; the total number of participants across all 18 studies was 3,358. Three studies primarily targeted adolescents (12-17), and seven studies focused exclusively on children under the age of eight; the remaining studies included participants that crossed multiple age groups. Despite the range of ages in the studies and consistent with task force recommendations (Barlow et al., 2017, USPSTF, 2010), parents were targeted as change agents for the child’s BMI reduction in all of the studies. Seven of the studies did not report or did not have enough data to calculate effect sizes for zBMI; in these cases, BMI percentile and/or BMI were used as the primary outcome variables.

Four studies compared treatment to active control groups, seven compared treatment to education control groups, and seven compared treatment to passive control groups. Across treatment groups, all studies included at least two healthy lifestyle components (e.g., targeting diet, physical activity, sedentary activity, or disordered eating), and 16 of the 18 included three or four components. Furthermore, all included studies incorporated behavioral components (e.g., specifying behaviors to change, reinforcing positive behaviors, setting goals, changing the environment, monitoring behaviors, promoting self-management skills). The mean number of months of treatment was 8.36, with a range of 1.5-24 months. The average number of treatment contacts (in-person and phone) was 11.6, with a range of 1-51 sessions/contacts. Nine studies (50%) included a pediatrician on the treatment team; of these, three
studies included a pediatrician as the sole interventionist. Three studies (16.7%) included a psychologist (or psychology trainee) on the treatment team. Eleven studies (61.1%) included interventionists from at least two professions (e.g., pediatricians, nurses, psychologists, dieticians). Thirteen of the studies (72.2%) noted the training of specific clinical staff (e.g., a pediatrician) to deliver the intervention. Half of the studies did not include a follow-up assessment; of those that did have a follow-up, the mean length of time for the follow-up assessment was 6 months (range from 3-12 months).

Eleven studies included an a priori power analysis, and 14 studies included intent-to-treat analyses. Only two studies reported using a manualized treatment. The average total score for study rigor, according to the rating system described by Lundahl et al. (2010), was 12.4, with scores ranging from 8-15 (out of a possible 18 points).

**Overall Effect Size**

The weighted average effect size for change in BMI for overweight and obese youth in primary care weight management interventions compared to control groups over all time points was small but statistically significant (Cohen’s $d = 0.26$, 95% CI [0.14, 0.38]; Cohen 1988). As indicated in Figure 2, the range of effect sizes for the 18 studies was quite large (-0.124 to 0.7636). Only one study had a negative effect size (Martinez-Andrade et al., 2014), which indicates that this was the only study that did not result in improved outcomes compared to the control group.

**Moderator Analyses**

Reflecting the range of effect sizes obtained, a formal test of heterogeneity was significant ($Q = 57.46, p < 0.01$), and indicated a medium to large amount of heterogeneity (i.e., the ratio of between-study variability to total variability; $I^2 = 70.34%$; Huedo-Medina, Sánchez-Meca, Marín-Martínez, & Botella, 2006). In an attempt to determine the source of heterogeneity, hypothesized continuous moderators in the current analyses were examined. Child age, gender, study rigor, and length of follow-up were not significant moderators of treatment effects. However, the number of treatment contacts (by any clinician and by any means, including phone and in-person contacts; $b = 0.007, p < 0.01$), duration of treatment (in months, $b = 0.027, p < 0.01$), and number of visits with a pediatrician ($b = 0.028, p < 0.05$) were significant moderators of treatment effects. Larger effect sizes were associated with more treatment contacts, longer treatment duration, and greater number of treatment sessions with a pediatrician. The type of control group (active, education, and passive) was also examined as a categorical moderator. However, type of control group was not a significant moderator of treatment effect.

**Publication Bias**

A fail-safe number was calculated to identify the number of studies with an effect size of zero that, if included in the analyses, would reduce the overall mean effect size to below the minimally significant effect size ($d = .01$). The fail-safe analyses indicated that 29 studies with an effect size of 0 would have to be added to the analyses to reduce the overall effect size to $d = 0.1$.

**DISCUSSION**

The primary care setting has been recommended as an ideal environment for pediatric weight management interventions due to the frequency of visitation and the accessibility and acceptability of that context for families (e.g., Sargent et al., 2011; Wald et al., 2011). However, a review of the literature failed to provide a quantitative analysis of the effectiveness of weight management interventions delivered in primary care settings. The current study was designed to build upon a systematic review conducted by Sargent and colleagues (2011) by providing a quantitative analysis (i.e., meta-analysis) of effect sizes of pediatric weight management interventions in primary care settings and evaluating potential moderators of intervention outcomes. The current study is unique and advances the field because it provides a standardized effect size to quantify the effectiveness of the weight management interventions in primary care settings compared to active, education, and passive control conditions. The study is particularly important because of the recent trends toward increasing the role and presence of professional psychology in primary care settings (e.g., Palermo et al., 2014; Rozensky & Janicke, 2012).

The current results extend Sargent and colleagues’ (2011) qualitative review of the literature, and indicate that pediatric overweight/obesity interventions in primary care settings can be effective. The overall effect size on change in BMI obtained in this meta-analysis was small but significant ($d = .26$, 95% CI [.14 to .38]; Cohen, 1988). This effect size is generally consistent with previous reviews examining lifestyle interventions for overweight/obese youth. Kitzmann et al. (2010) reported an effect size of $d = .41$ (95% CI = .26 to .55) for behavioral lifestyle interventions, and Janicke et al. (2014) reported an effect size of $g = .47$ (95% CI = .36 to .58) for comprehensive behavioral lifestyle interventions. Although the confidence interval of the present study’s effect size overlaps with Kitzmann and colleagues’ (2010) results, there is
considerably less overlap with the confidence interval reported by Janicke et al. (2014). Reasons for the slightly lower effect sizes (in comparison to Janicke and colleagues’) may include the nature of the treatment conditions included: Janicke et al. (2014) included only studies with interventions that addressed dietary intake, physical activity, and behavioral strategies (i.e., comprehensive behavioral lifestyle interventions), whereas the current meta-analyses included studies that included at least two healthy lifestyle components (e.g., targeting diet, physical activity, disordered eating, and/or sedentary activity) and did not require interventions that target all three of the components.1 Though the purpose of this meta-analysis was not to examine the efficacy of specific treatment components, the difference in the effect sizes across meta-analyses may reflect positively on comprehensive behavioral lifestyle interventions.

The current study also allowed an investigation of a number of potential moderators of weight management intervention for children and adolescents delivered in primary care settings. Previous studies have found better weight outcomes related to male gender (Jelalian et al., 2008) and among older children and adolescents (Janicke et al., 2014). In contrast, child age and gender were not significant moderators of treatment effects in the current meta-analysis. Similarly, study rigor and length of follow-up did not significantly impact treatment outcomes. However, in the current study, the number of treatment contacts, number of treatment months, and number of visits with a pediatrician were particularly important, such that larger effect sizes were associated with more treatment contacts, longer duration of treatment, and greater involvement from a pediatrician. This finding is consistent with Janicke et al. (2014), who reported that larger effect sizes were associated with longer treatment (duration in months) and more treatment sessions. It is left to future studies to determine whether the impact of treatment duration and number of contacts differs depending on who is delivering the intervention or on the background and level of training of the interventionist.

Because primary care settings can be effective for pediatric weight management programs, assessment, prevention, and intervention, efforts should be incorporated universally across primary care clinicians (Barlow, 2007; USPSTF, 2010). However, even when weight management interventions are implemented, evidence from the current analyses suggests that weight management interventions in primary care settings vary greatly on a number of factors including the profession/training of the clinician, the components of the intervention, and the intensity of treatment (i.e., number of contacts). Taken together with the recommendations from the USPSTF (2010), children who visit primary care settings should be screened for overweight/obesity, and youth who are overweight or obese should receive a moderate-to-high intensity behavioral lifestyle intervention to promote healthy diet and physical activity habits using behavioral strategies (USPSTF, 2010). Additionally, the current analyses suggest that greater involvement from a pediatrician is an important component for weight management interventions in primary care settings.

The results of the present analyses suggest a number of roles for the child health psychologist in delivering effective interventions for pediatric obesity/overweight in primary care settings. As noted above, only three interventions (16.7%) had a psychologist (or graduate trainee) directly involved in the delivery of the treatment. However, beyond direct care, and given their familiarity with the application of behavioral principles to health conditions, psychologists can also facilitate the delivery of evidence-based interventions for obesity/overweight through education/training and systematic consultation to primary care clinicians, even if the psychologist her or himself is not directly involved in treatment delivery. Although the specific role of psychologists was not always specified, the majority of primary care interventions identified in the literature (72%) described at least some training for the primary care clinicians.

Despite the contributions of these analyses, there are certain limitations that must be considered. First, the fail-safe number of 29 is relatively small; this calculation suggests that only 29 articles with an effect size of zero would be needed to reduce the overall effect size of the current meta-analysis to 0.1. According to Rosenthal’s (1979) suggestion of an appropriate fail safe number (5k + 10; where k is the number of studies), a fail-safe number for the current meta-analysis that is robust to the existence of possible excluded studies is 100. Therefore, the possibility of publication bias should be considered when evaluating the results of this study. Additionally, the relatively

---

1 Beyond this difference, and despite very little overlap (n = 2) in the studies included in the two meta-analyses, the overall characteristics of the samples were quite similar. For example, the mean duration of treatment was similar (Janicke et al., 6.4 months; current study, 8.3 months), as was the methodological rigor of included studies (Janicke et al., Lundahl score of 12.3; current study mean Lundahl score of 12.4).
small number of articles included in the analysis may have made it difficult to detect moderators, such as control group type and length of follow-up, which might have been detected with the inclusion of additional studies.

Furthermore, there were a number of additional potential moderators that the literature suggests might impact the accessibility and availability of interventions; however, the current meta-analysis was not able to examine some potential moderators due to the lack of reporting consistency across studies. For example, socioeconomic status (SES) was not examined as a moderator of weight management interventions in primary care settings due to the lack of consistent reporting practices for information regarding SES. Future studies should use similar measurements of SES, such as the Hollingshead Four-Factor Index of Socioeconomic Status (Hollingshead, 1975). Similarly, articles did not uniformly provide information on participant satisfaction, acceptability of the intervention, and attendance; thus, these factors were not examined as potential moderators.

Additional potential moderators were not examined due to the nature of the studies selected for the current analyses. For example, because all of the studies in the current analyses targeted parents as change agents for the child’s BMI reduction, parental involvement was not examined as a potential moderator to help shed light on the inconsistencies in the literature regarding whether parental involvement is associated with better outcomes (Faith et al., 2012). Additional research is necessary to determine the role of parental involvement as a moderator, and more specifically, the role of parental adherence to behavioral components and which behavioral components promote better outcomes (Faith et al., 2012). The inclusion criteria for the current meta-analysis allowed studies from any country, as long as they were written in English. Health care systems vary by country, which could impact the implementation and success of weight-management interventions. Thus, it would be important to examine country as a moderator in future analyses. Additionally, a wide range of professionals (e.g., pediatricians, nurses, psychologists, dieticians) administered the interventions in the studies included within the current meta-analysis, and some studies included a team of professionals to assist with specific components of the intervention. Although examination of professional type as a moderator was not possible in the current study, the field should understand the role of the various clinicians in weight management interventions in primary care. Furthermore, it will be important for future studies to examine the role of different professionals and the level of training needed to successfully conduct pediatric overweight/obesity treatments in these settings. This information would allow recommendations for who is crucial in administering these interventions.

Another important limitation of the current meta-analysis is that it focused on solely weight change (e.g., BMI, zBMI, BMI percentile) as the primary treatment outcome. Although weight reduction is a significant area of concern, pediatric overweight/obesity also increases the risk for negative mental health outcomes (e.g., lower quality of life, impaired social functioning, increased depressive symptoms; Sawyer et al., 2011). Future reviews should also examine the impact of primary care interventions for obesity on mental health outcomes. For example, future studies may consider examining health-related quality of life (HRQOL) as an outcome measure. HRQOL is a multidimensional concept which consists of functioning across several domains including school, social, physical, and emotional (Palermo et al., 2008). Positive changes in HRQOL could potentially be more meaningful for patients and perhaps more immediate than changes in BMI.

The findings of the current meta-analysis suggest that weight management programs in primary care settings are effective for reducing BMI in comparison to active, education, and passive control conditions. Additionally, a greater number of treatment contacts, number of treatment months, and number of visits with a pediatrician appear to be positively related to reduction in BMI. The present findings are encouraging and indicate that primary care continues to be a suitable setting for pediatric overweight/obesity weight management interventions. Future research on interventions in primary care settings should continue to rigorously examine other important aspects of accessibility and acceptability, such as SES, acceptability, attendance, and satisfaction. Future research should also consider the effectiveness of pediatric weight management interventions in primary care settings on improving other psychosocial outcomes, including HRQOL.
REFERENCES


---

2 References marked with an asterisk indicate studies included in the meta-analysis.


*Small, L., Bonds-McClain, D., Melnyk, B.,


**Figure 1. PRISMA Flow Diagram for Included Studies.**

- **Records identified through database searching** (n=2275)
  - Additional records identified through other sources (n=5)

  - Records screened (n=2280)
    - Records excluded: (n=2192)
      - Full-text articles assessed for eligibility (n=88)
        - Full-text articles excluded:
          - Not in primary care (n=2)
          - Prevention (n=4)
          - Incorrect age group (n=4)
          - Not enough data (n=26)
          - Not outcome of interest (n=6)
          - No comparison group (n=9)

  - Studies included in meta-analysis (n=18)
<table>
<thead>
<tr>
<th>Author et al. (Year)</th>
<th>Total Sample Size (N)</th>
<th>Study Effect size, 95% CI</th>
<th>Child Age Range (years)</th>
<th>Average Child Age (years)</th>
<th>Control Group</th>
<th>% Female</th>
<th>Country</th>
<th>Pediatrician Involved in Treatment?</th>
<th>Treatment Contacts</th>
<th>Follow up period (months)</th>
<th>Study Rigor Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Banks et al. (2012)</td>
<td>52</td>
<td>0.039 (-0.508, 0.587)</td>
<td>5-16</td>
<td>11.4</td>
<td>Active</td>
<td>N/A</td>
<td>United Kingdom</td>
<td>No</td>
<td>5</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Berkowitz et al. (2013)</td>
<td>169</td>
<td>0.007 (-0.295, 0.308)</td>
<td>12-16</td>
<td>14.6</td>
<td>Active</td>
<td>76.9%</td>
<td>United States</td>
<td>No</td>
<td>23</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Davoli et al. (2013)</td>
<td>372</td>
<td>0.454 (0.249, 0.659)</td>
<td>4-7</td>
<td>6.6</td>
<td>Education</td>
<td>61.6%</td>
<td>Italy</td>
<td>Yes</td>
<td>5</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>DeBar et al. (2012)</td>
<td>208</td>
<td>0.290 (0.018, 0.562)</td>
<td>12-17</td>
<td>14.1</td>
<td>Education</td>
<td>100.0%</td>
<td>United States</td>
<td>Yes</td>
<td>28</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Diaz et al. (2010)</td>
<td>43</td>
<td>0.521 (-0.085, 1.126)</td>
<td>9-17</td>
<td>N/A</td>
<td>Active</td>
<td>51.6%</td>
<td>Mexico</td>
<td>Yes</td>
<td>51</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Looney &amp; Raynor (2014)</td>
<td>22</td>
<td>0.087 (-0.929, 1.103)</td>
<td>4-10</td>
<td>8</td>
<td>Education</td>
<td>68.2%</td>
<td>United States</td>
<td>No</td>
<td>6</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Marild et al. (2012)</td>
<td>193</td>
<td>0.464 (0.049, 0.878)</td>
<td>8-13</td>
<td>11</td>
<td>Passive</td>
<td>57.6%</td>
<td>Sweden</td>
<td>No</td>
<td>12</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Martinez-Andrade et al. (2014)</td>
<td>306</td>
<td>-0.124 (-0.349, 0.101)</td>
<td>2-5</td>
<td>3.4</td>
<td>Passive</td>
<td>47.4%</td>
<td>Mexico</td>
<td>No</td>
<td>6</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>McCallum et al. (2007)</td>
<td>163</td>
<td>0.082 (-0.226, 0.389)</td>
<td>5-10</td>
<td>7.4</td>
<td>Passive</td>
<td>52.0%</td>
<td>Australia</td>
<td>Yes</td>
<td>4</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Authors</td>
<td>N</td>
<td>B</td>
<td>Age</td>
<td>Gender</td>
<td>Type</td>
<td>Percentage</td>
<td>Country</td>
<td>Treatment</td>
<td>No</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td>-------------------------</td>
<td>----</td>
<td>-----</td>
<td>-----</td>
<td>--------</td>
<td>--------</td>
<td>------------</td>
<td>---------</td>
<td>-----------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>Quattrin et al. (2014)</td>
<td>96</td>
<td>0.693</td>
<td>2-5</td>
<td>4.5</td>
<td>Education</td>
<td>66.7%</td>
<td>United States</td>
<td>No</td>
<td>13</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td>Resnicow et al. (2014)</td>
<td>645</td>
<td>0.764</td>
<td>2-8</td>
<td>5.1</td>
<td>Education</td>
<td>57.1%</td>
<td>United States</td>
<td>Yes</td>
<td>10</td>
<td>0</td>
<td>12</td>
</tr>
<tr>
<td>Saelens et al. (2002)</td>
<td>39</td>
<td>0.370</td>
<td>12-16</td>
<td>14.2</td>
<td>Active</td>
<td>40.9%</td>
<td>United States</td>
<td>Yes</td>
<td>12</td>
<td>3</td>
<td>14</td>
</tr>
<tr>
<td>Schwartz et al. (2007)</td>
<td>61</td>
<td>0.167</td>
<td>3-7</td>
<td>5</td>
<td>Passive</td>
<td>59.0%</td>
<td>United States</td>
<td>Yes</td>
<td>2</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Small et al. (2014)</td>
<td>60</td>
<td>0.163</td>
<td>4-8</td>
<td>5.9</td>
<td>Education</td>
<td>60.0%</td>
<td>United States</td>
<td>No</td>
<td>4</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>Taveras et al. (2011)</td>
<td>445</td>
<td>0.085</td>
<td>2-6</td>
<td>4.9</td>
<td>Passive</td>
<td>48.0%</td>
<td>United States</td>
<td>Yes</td>
<td>7</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>Tucker et al. (2013)</td>
<td>125</td>
<td>0.223</td>
<td>4-18</td>
<td>9.7</td>
<td>Education</td>
<td>49.0%</td>
<td>United States</td>
<td>No</td>
<td>6</td>
<td>6</td>
<td>11</td>
</tr>
<tr>
<td>Wake et al. (2009)</td>
<td>258</td>
<td>0.023</td>
<td>5-10</td>
<td>7.5</td>
<td>Passive</td>
<td>46.9%</td>
<td>Australia</td>
<td>Yes</td>
<td>4</td>
<td>9</td>
<td>15</td>
</tr>
<tr>
<td>Wald et al. (2011)</td>
<td>101</td>
<td>0.327</td>
<td>9-12</td>
<td>10.6</td>
<td>Passive</td>
<td>63.4%</td>
<td>United States</td>
<td>No</td>
<td>11</td>
<td>8</td>
<td>11</td>
</tr>
</tbody>
</table>

* Included an *a priori* power analysis; ^b^ Included intent-to-treat analyses; ^c^ Indicated use of manualized treatment; ^d^ Study rigor scores could range from 0-18, with higher scores indicating higher study quality.
**Figure 2. Forest plot for included studies to depict how the individual studies contributed to the overall weighted average effect size**

Banks, Sharp, Hunt, & Shield, 2011  
0.039 (-0.508, 0.587)

Berkowitz et al., 2013  
0.007 (-0.295, 0.308)

Davoli et al., 2013  
0.454 (0.249, 0.659)

DeBar et al., 2012  
0.290 (0.018, 0.562)

Diaz et al., 2010  
0.521 (-0.085, 1.126)

Looney & Raynor, 2014  
0.087 (-0.929, 1.103)

Marild et al., 2012  
0.464 (0.049, 0.878)

Martinez-Andrade et al., 2014  
-0.124 (-0.349, 0.101)

McCallum et al., 2007  
0.082 (-0.226, 0.389)

Quattrin et al., 2014  
0.693 (0.281, 1.105)

Resnicow et al., 2015  
0.764 (0.565, 0.962)

Saelens et al., 2001  
0.370 (-0.263, 1.003)

Schwartz et al., 2007  
0.167 (-0.482, 0.816)

Small et al., 2013  
0.163 (-0.354, 0.680)

Taveras et al., 2011  
0.085 (-0.103, 0.272)

Tucker et al., 2013  
0.223 (-0.129, 0.575)

Wake et al., 2009  
0.023 (-0.221, 0.268)

Wald, Moyer, Eickhoff, & Ewing, 2011  
0.327 (-0.140, 0.794)

*Note.* The diamonds represent the effect size of the individual study, and the bars represent the 95% confidence interval. The central vertical line represents no treatment effect, or the null hypothesis.