

## *Comparative Effectiveness Research Review Disposition of Comments Report*

### **Research Review Title: Strategies To Prevent Weight Gain Among Adults**

Draft review available for public comment from May 23, 2012 thru June 28, 2012.

**Research Review Citation:** Hutfless S, Maruthur NM, Wilson RF, Gudzone KA, Brown R, Lau B, Fawole OA, Chaudhry ZW, Anderson CAM, Segal JB. Strategies to Prevent Weight Gain Among Adults Comparative Effectiveness Review No. 97. (Prepared by The Johns Hopkins University Evidence-based Practice Center under Contract No. 290-2007-10061-I.) AHRQ Publication No. 13-EHC029-EF. Rockville, MD: Agency for Healthcare Research and Quality; March 2013. [www.effectivehealthcare.ahrq.gov](http://www.effectivehealthcare.ahrq.gov).

### **Comments to Research Review**

The Effective Health Care (EHC) Program encourages the public to participate in the development of its research projects. Each comparative effectiveness research review is posted to the EHC Program Web site in draft form for public comment for a 4-week period. Comments can be submitted via the EHC Program Web site, mail or email. At the conclusion of the public comment period, authors use the commentators' submissions and comments to revise the draft comparative effectiveness research review.

Comments on draft reviews and the authors' responses to the comments are posted for public viewing on the EHC Program Web site approximately 3 months after the final research review is published. Comments are not edited for spelling, grammar, or other content errors. Each comment is listed with the name and affiliation of the commentator, if this information is provided. Commentators are not required to provide their names or affiliations in order to submit suggestions or comments.

The tables below include the responses by the authors of the review to each comment that was submitted for this draft review. The responses to comments in this disposition report are those of the authors, who are responsible for its contents, and do not necessarily represent the views of the Agency for Healthcare Research and Quality.

Commentator & Affiliation	Section	Comment	Response
<b>Peer Reviewer 1</b>	General	The report is clinically meaningful and the target population and audience are well-defined.	Thanks.
<b>Peer Reviewer 1</b>	General	Conceptually, it is difficult to understand how a weight gain prevention intervention can have self-management components (KQ1; defined as goal setting, problem solving, social support, etc. (p.19)) without having a diet or physical activity component (KQs 2 and 3; i.e., there must be a target behavior for self-management techniques). In fact, several of the self-management, diet, and activity interventions categorized as being only 1 of the 3 actually have overlapping components (e.g., diet + self-management). Thus, having separate key questions that examine the comparative effectiveness of self-management, diet, or physical activity interventions in isolation seems problematic as these truly don't exist in isolation in any well executed intervention and should, therefore, be placed in the combination category.	We re-evaluated the self-management articles. We made sure that interventions that included diet (or physical activity) + self-management were included with the combination interventions.  Added to p3 of Executive Summary and p25 of Methods under scope and key questions "Dietary and physical activity interventions inherently include some aspects of self-management. Only when self-management did not include traditional diet or physical activity components (i.e., daily weighing or regulating television viewing) was the study reported in KQ1."
<b>Peer Reviewer 1</b>	Introduction	No substantive suggestions	Thanks.

Commentator & Affiliation	Section	Comment	Response
<b>Peer Reviewer 1</b>	Methods	It would be helpful to know how the authors arrived at their definitions of clinical significance and associated cut-off scores. This information should be explicitly stated in the Review.	<p>We identified articles from the literature that reported on clinically meaningful thresholds. However, these thresholds were usually in the setting of weight loss. We identified a reference (Katan MB, Ludwig DS. Extra calories cause weight gain--but how much? JAMA 2010; 303(1):65-6.PMID:20051571) (Lovasi GS, Hutson MA, Guerra M, Neckerman KM. Built environments and obesity in disadvantaged populations. Epidemiol Rev 2009; 31:7-20.PMID:19589839) for a clinically meaningful definition for weight maintenance; we then applied this to a BMI change for an individual starting with BMI=27 (mean/median in many included studies). For waist circumference, we did not identify a threshold and thought that due to measurement error we should go no lower than 1cm.</p> <p>Following the suggestion that most adults gain 0.5kg/year, we changed our clinically meaningful threshold for the difference in weight maintenance between groups to be 0.5kg for 1 year trials (1kg for 2 yr, etc.). We created the corresponding value for BMI based on the baseline BMI of 27 (0.2 units in year 1, 0.4 units at 2 years).</p> <p>For waist circumference, we felt 1cm should be the minimum due to measurement error. We recommended the difference in groups to be at least 1 cm for year 1, 2 cm for year 2, etc.</p>
<b>Peer Reviewer 1</b>	Methods	Given that the prevention of excess weight gain is important during pregnancy, a rationale for excluding pregnant women from the review should be provided in the text, particularly given that other subpopulations at high risk for weight gain are examined.	<p>Prevention of excessive weight gain during pregnancy is beyond the scope of the review because for normal weight women, weight gain is expected during pregnancies. The inclusion criteria would have been different for this population and the clinically meaningful thresholds would have differed as well.</p> <p>We were open to including studies that examined prevention of weight gain after pregnancy as long as the women were followed at least 1 year post-partum. We did not observed any studies that met these inclusion criteria. Note that we also included Adair et al (Adair LS, Gultiano S, Suchindran C. 20-year trends in Filipino women's weight reflect substantial secular and age effects. J Nutr 2011; 141(4):667-73.PMID:21325475) which recruited women while they were pregnant.</p>
<b>Peer Reviewer 1</b>	Methods	Were only studies with objective assessments of height and weight included? If not, this should be stated. Moreover, authors may want to consider this when rating evidence as high, moderate, or low.	<p>We downgraded the quality of interventions that did not report an outcome assessor blinded to intervention group. Presumably the measurement of weight is objective but still can be subject to biased data collection.</p>

Source: <http://effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?productid=1441&pageaction=displayproduct>

Published Online: March 25, 2013

Commentator & Affiliation	Section	Comment	Response
<b>Peer Reviewer 1</b>	Methods	Search strategies are explicitly stated and logical and statistical approach seems appropriate.	Thanks.
<b>Peer Reviewer 1</b>	Results	In some places only clinical significance is reported, in other places only statistical significance is reported, whereas in others it is not clear whether there was a statistically significant difference but not a clinically significant difference (e.g., pgs 44-45). Reporting should be consistent throughout the review and increased clarity regarding statistical and clinical significance would help to interpret the findings.	We appreciate the concern for consistency throughout the report. We report now on the between group difference threshold and statistical significance throughout the report.
<b>Peer Reviewer 1</b>	Results	Throughout the results, some examples of intervention components could be added to provide the reader a better context of the interventions being presented	We have moved all of the intervention characteristics tables to the body of the report rather than the appendices.
<b>Peer Reviewer 1</b>	Results	All figures lack signs indicating direction of weight loss, etc (i.e., +/-), which makes them unclear.	This has been improved.
<b>Peer Reviewer 1</b>	Results	Figures 3 & 4: It is not clear why one would control for change in diet over time when examining the effects of dietary interventions (designed to induce change) on weight.	Thank you for the attention to detail. "Adjusted for age, race/ethnicity, baseline BMI, change in dietary and physical activity patterns over time" changed to "adjusted for age, race/ethnicity, baseline BMI, change in physical activity patterns over time."
<b>Peer Reviewer 1</b>	Results	There are indicators on the bottom of each figure that the effect either "Favors Group 1" or "Favors Group 2;" just because an intervention did not produce a clinically meaningful effect, does not necessarily mean that it favors the control group. Instead, it just means that it was not robust enough to produce a clinically meaningful difference between the two conditions.	We deleted the line for clinically meaningful difference/between group difference threshold. We kept the label to identify which intervention is favored. We agree we should delete "Favors group 2" or move to the corresponding threshold above 0. For head to head comparisons, 2 clinically meaningful thresholds would be on each figure. We would make sure that these changes do not make the figure overly busy before implementing
<b>Peer Reviewer 1</b>	Results	On a related note, in some Figures control and intervention conditions are referred to as "Group 1" or "Group 2" whereas in others they are referred to as "Control" or "Intervention" – this should be consistent throughout.	This has been changed.
<b>Peer Reviewer 1</b>	Results	I assume that the size of the effect symbol reflects the sample size for each study, however, this should be explicitly stated.	This has been changed in Figures 2-5 to have a common size of effect symbol with the N added as a column. In remaining Figures, a footnote has been added to indicate that the size of the effect symbol reflects the sample size (Renee to add footnotes).
<b>Peer Reviewer 1</b>	Results	Table 14 – Primary Aim is missing as is description of "Group 1"	Study the effect of TV watching hours on change in BMI in patients with colorectal cancer."

Source: <http://effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?productid=1441&pageaction=displayproduct>

Published Online: March 25, 2013

Commentator & Affiliation	Section	Comment	Response
<b>Peer Reviewer 1</b>	Results	Table 15 – based on the description it appears that there were 4 treatment arms (Food Guide Pyramid, low fat, high fruit/veg, and low fat + high fruit/veg); however, only 2 group columns are identified.	This table now has 4 columns and includes the relevant information.
<b>Peer Reviewer 1</b>	Results	Page 35, lines 27-29: should be under weight loss and not waist circumference	Changed to “There was no difference between the gym based exercise group and the control group in waist circumference among the elderly.”
<b>Peer Reviewer 1</b>	Results	Page 27, line 55 –“low risk” should be “high risk”	Changed low to high. Thank you for your attention to detail.
<b>Peer Reviewer 1</b>	Results	Page 39, lines 33-3: Indicating effects of BMI and age on intervention outcomes would be informative. Similarly, given that gender was examined separately in the Robins et al (2006) study, it would be informative to report the effects of the intervention on men (page 65, lines 15-21).	Subgroup analyses (including baseline BMI, age and sex) are reported at the end of each key question.
<b>Peer Reviewer 1</b>	Results	Page 78, lines 35-39: Lack of data on waist circumference is reported under the heading Weight Change.	This statement has been moved to the appropriate section of the report.
<b>Peer Reviewer 1</b>	Results	Page 80, lines 24-29: Subanalyses are described, but it is not clear who was in the subgroup that was analyzed.	We have clarified who the subgroups are under Subgroups for College-based interventions. We have revised the paragraph to read: “One trial evaluated the effects of the intervention on subgroups based on BMI: BMI less than or equal to 24 kg/m <sup>2</sup> and BMIs > 24 kg/m <sup>2</sup> ) at baseline. <sup>84</sup> There were no differences in 16 month BMI change between intervention and control participants who had lower BMIs at baseline. However, the higher BMI intervention group (n=11) lost 1.4 kg as compared with higher BMI controls (n=6) who gained 9.2 kg. This difference met the between group difference threshold and was statistically significant (Appendix E, Evidence Table 17).”
<b>Peer Reviewer 1</b>	Discussion	In the Executive Summary, page 14, a paragraph is lifted from the section “Research Gaps” and placed in “Conclusions;” a more thoughtful concluding paragraph would be appropriate.	Deleted “more research needed” final sentence and replaced with “Although there is no strong evidence to promote a particular weight gain prevention strategy, there is no evidence that <i>not</i> adopting a strategy to prevent weight gain is preferable.”

Commentator & Affiliation	Section	Comment	Response
<b>Peer Reviewer 1</b>	Discussion	Too much is made from the observational study that found eating potato chips and French fries is associated with weight gain. Consumption of these foods is likely just a proxy for a poor diet. Similarly, concluding that eating fewer potatoes will help to prevent weight gain overstates the effect (p. 122, line 13).	Agree with Reviewer that confounding is a concern, but the Nurses Health Study/Health Professionals Followup Study did meet our thresholds for significance and if this is the best info we have, we need to report it. We have diminished the attention paid to this finding in the ES and abstract.  (Mozaffarian D, Hao T, Rimm EB, Willett WC, Hu FB. Changes in Diet and Lifestyle and Long-Term Weight Gain in Women and Men: New England Journal of Medicine. N Engl J Med 2011; 364(25):2392-404.PMID: <a href="http://dx.doi.org/10.1056/NEJMoa1014296">http://dx.doi.org/10.1056/NEJMoa1014296</a> )
<b>Peer Reviewer 1</b>	Discussion	To explain the finding that control groups did not gain weight at rates consistent with the national average, it is suggested that weight monitoring may have helped control participants maintain their weight. Another plausible explanation is selection bias – individuals who choose to join a weight gain prevention trial may be more weight conscious than those who don't, thereby gaining less weight than the typical American over the same timeframe.	Thank you. We have included in our discussion section the following text: "Many control groups had no increase in weight over time. In the general United States population, adults gain about 0.5 kg per year. (Williamson DF. Descriptive epidemiology of body weight and weight change in U.S. adults. Ann Intern Med 1993; 119(7 Pt 2):646-9.PMID:8363190) Individuals enrolled in interventional studies may be more likely to make behavior changes regardless of the group assignment. It is possible that the knowledge that one will be evaluated on weight regularly may help people to maintain weight without an intensive intervention. This may support the use of weight surveillance interventions in a workplace or primary care setting"
<b>Peer Reviewer 1</b>	Discussion	In essence, more research is needed and I think that this comes through nicely in both the discussion and concluding comments.	Thanks!
<b>Peer Reviewer 1</b>	Clarity and Usability	This report is well organized.	Thanks
<b>Peer Reviewer 1</b>	Clarity and Usability	This report helps to identify areas for future research.	Thanks

Commentator & Affiliation	Section	Comment	Response
<b>Peer Reviewer 2</b>	General	<p>I was asked to review the Comparative Effectiveness Review entitled “Approaches to Weight Maintenance in Adults.” While clearly a lot of work went into conducting this review, I am very concerned about the basic decisions that were made in developing it and consequently the body of research reviewed. Specifically, the review lacks a basic understanding of weight regulation. Weight maintenance is a matter of balancing caloric intake and expenditure. If weight gain is occurring commonly in populations, intake is exceeding expenditure. Therefore to reduce the risk of weight gain, it is necessary to decrease caloric intake and/or increase expenditure. The major problem with this review is that studies which focus on decreasing caloric intake appear to have been excluded (because they are considered weight loss vs weight gain prevention). Thus one is left with including in this review those dietary and combination approaches that only involve changing fat intake, Mediterranean diets, reducing salt intake, slowing act of eating, etc. As suggested by this review such changes do not reduce the risk of weight gain. Those studies where caloric intake is reduced need to be included as they have the greatest chance of yielding positive results.</p>	<p>The EPC team included expertise in weight loss, weight maintenance, policy, internal medicine, diabetes, epidemiology and systematic reviews. We also obtained input from the Key Informants and a panel of technical experts when we developed the questions, scope and protocol. Expertise on the TEP included several weight loss experts. We were inclusive of studies that targeted caloric restriction for weight gain prevention we did not include studies that targeted caloric restriction for weight loss.</p> <p>“Failed” weight loss to increase weight maintenance among the overweight and obese is beyond the scope of the review. The mechanism to maintain weight among those who have lost weight or are trying to lose weight is clearly different than those who have never been overweight as evidenced by the extreme difficult individuals who have lost weight have with keeping the weight off long-term. We agree that the concept is worthy of further investigation, but the current report aims to examine primary prevention of weight gain.</p> <p>We have clarified our rationale for excluding these studies in the Introduction and Discussion. We agree that this approach is a future research need for weight gain prevention among the overweight and obese.</p>



Commentator & Affiliation	Section	Comment	Response
<b>Peer Reviewer 2</b>	General	<p>A related problem is that the review excludes studies that use weight loss as a means to prevent weight gain. This is a major concern because producing periodic weight losses, even if there is some weight regain, may well be the most effective approach to weight gain prevention. For example, the Women's Healthy Lifestyle Project is not included in this review (Kuller et al, Circulation, 2001). This trial involves 500 women who were randomized to lifestyle intervention or control and tests the hypothesis that "reducing saturated fat and cholesterol consumption and preventing weight gain by decreased caloric and fat intake and increased physical activity would prevent the rise in LDL cholesterol and weight gain during perimenopause to postmenopause". I assume this study was excluded because the intervention group, including those who were normal weight, was given a modest weight loss goal as a means to preventing weight gain. However, this study showed convincingly that this approach decreased the expected weight gain during this menopausal transition. This study, and others like it that use weight loss as a means to prevent weight gain, should be included. Omitting these studies excludes a potentially very effective approach and seriously biases this report.</p>	<p>Yes, this study was excluded because some individuals had a goal of weight loss.</p> <p>The question "What is the effectiveness of weight loss interventions for long-term weight stability?" may be a relevant question for decisionmakers but was not the subject of our review.</p> <p>Discussion, p131: "Presently, the majority of the U.S. population is overweight or obese, so weight loss interventions may be more appropriate than weight gain prevention interventions for most of the population. However, changes that prevent weight gain may be more effective in the healthy weight population than encouraging weight loss in the overweight and obese population. This, however, remains unknown. Environment level interventions can benefit people of all weights. The CDC and WHO recommend environment level interventions, despite the limited availability of long-term data, for this reason."<sup>107</sup></p>



Commentator & Affiliation	Section	Comment	Response
<b>Peer Reviewer 2</b>	General	Similarly, the Pound of Prevention was excluded because the reviewers noted that some individuals in the study were trying to lose weight. However, this study did NOT target weight loss. This is one of the best weight gain prevention studies and should definitely be included. The fact that the individuals are seeking to lose weight is not different from almost many other studies reviewed here. Consider especially the observational studies, where weight loss is a goal of many participants at different times during the period of follow-up. While the randomized component of Pound of Prevention did not yield positive results, in subsequent analyses of this trial, which are more similar to observational studies reported here, those participants who increased self-weighing (Linde, et al, Annals Behav Med, 2005), increased physical activity and decreased fat intake (Sherwood, IJO, 2000) had the best maintenance of weight. Finally in a subsequent report these authors Jeffery et al, Int J of Obesity, 2002) showed that while the study did not target weight loss, 9% of the participants lost 5% or more of their weight at Year 1; these participants were the ONLY group that was below baseline at Year 3, again supporting the contention that producing some weight loss may be the most effective approach to weight gain prevention.	We also struggled with excluding Pound of Prevention given its name and goal. However, individuals were allowed to choose a weight loss goal and the study was excluded for this reason. We address the exclusion of Pound of Prevention in the discussion. The reviewer brings up an important point regarding observational studies. The investigators of observational studies are likely uncertain regarding the intent of the participants' weight goals. Because we included interventional studies with unclear intent, we also included observational studies. We have pointed out in the discussion that at least some of the individuals in the studies may have had weight loss goals and the results may reflect the effects of behaviors that lead to weight loss and weight gain prevention.
<b>Peer Reviewer 2</b>	General	Further evidence that initial weight loss may be important for weight gain prevention comes from the Nurses Health Study (Field, A. IJO, 2001); although the Nurses Health Study is included, I do not see mention of this specific analysis. In this paper, it was shown that women who lost weight initially subsequently experienced greater weight regain than those who remained stable—yet, despite this regain, these women ended up with less overall weight gain than those who remained weight stable.	Mozaffarian 2011 pooled data from the Nurses Health Study, Nurses Health Study II and Health Professionals Followup Study. The 2011 study updates the 2001 study. (Field AE, Wing RR, Manson JE, Spiegelman DL, Willett WC. Relationship of a large weight loss to long-term weight change among young and middle-aged US women. International Journal of Obesity 2001; 25(8):1113-21.PMID: http://search.ebscohost.com/login.aspx?direct=true&db=psych&AN=2001-05766-001&site=ehost-live Alison.Field@channing.harvard.edu)

Commentator & Affiliation	Section	Comment	Response
<b>Peer Reviewer 2</b>	General	Another major concern is the use of observational studies, along with the clinical trials in this evidence based report. The reviewers note that observational studies provide a lower level of scientific evidence than trials and try to deal with this by referring to these as “approaches”. However, this wording does not satisfactorily reduce the fact that findings from these observational studies are given a fair amount of attention in the report, for example in the Abstract. There is no evidence that avoiding potato chips or French fries will prevent weight gain. I would recommend that this report be limited to actual trials.	Consistent with the AHRQ Methods Manual for Systematic Reviews, we included observational studies if they could strengthen the body of evidence. Given that we had few trials on a number of the interventions, we chose to include observational studies. We were highly selective in the observational studies we included with strict criteria for confounding control. Additionally, the majority of RCTs do not reflect real world situations that need to occur for long-term prevention of weight gain.
<b>Peer Reviewer 2</b>	General	The other major concern with this review is with the selection of topics. I feel strongly that neither studies of Adults with Cancer nor Studies of Adults with Mental Illness should be included in this report. These individuals are gaining weight due to treatments received; thus approaches to preventing this weight gain may be specific to the pathways by which the medication affects weight regulation. I would strongly recommend removing these sections as I feel they detract from, rather than improve the report.	The subgroups were discussed with the Key Informants and Technical Expert Panel. They agreed with the inclusion of these subgroups in the report.
<b>Peer Reviewer 2</b>	General	While I feel that weight changes in response to external influences, such as medications, seem outside this review, if adults with cancer and adults with mental illness are included, this reviewer wonders why the weight gain with smoking cessation is not included. Smoking cessation, like these medications, disrupts the regulation of body weight and produces weight gain. Trials of ways to prevent this weight gain have been reported and if sections on cancer and mental illness are seen as appropriate, then smoking cessation should be included as well.	There is a 2012 Cochrane review on this topic <a href="http://summaries.cochrane.org/CD006219/interventions-for-preventing-weight-gain-after-smoking-cessation">http://summaries.cochrane.org/CD006219/interventions-for-preventing-weight-gain-after-smoking-cessation</a> <a href="http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD006219.pub3/abstract;jsessionid=C68B5F692D4588D4C549821EB556CC25.d03t04">http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD006219.pub3/abstract;jsessionid=C68B5F692D4588D4C549821EB556CC25.d03t04</a>  We chose not to duplicate the Cochrane group's efforts.
<b>Peer Reviewer 2</b>	Abstract	The title is an inaccurate reflection of the topic covered. The document should be re-titled: Comparative effectiveness of approaches for prevention of weight gain. This is the topic addressed in each of the questions posed and far better conveys the domain covered by this report.	We agree and this has been changed.

Source: <http://effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?productid=1441&pageaction=displayproduct>

Published Online: March 25, 2013

Commentator & Affiliation	Section	Comment	Response
<b>Peer Reviewer 2</b>	Abstract	The abstract suggests that 2 approaches were effective at preventing weight gain, with moderate strength of evidence; one of these, the college course, produced a statistically significant effect of 1.3 kg at 24 months – the authors state (p. 71) that this change is not considered clinically significant, but then this effect is called significant in the Abstract. Moreover, as college freshman may still be growing, BMI would be the preferable measure to us in this study– there were no differences between groups for BMI. The second study compared aerobic, resistance and no exercise on weight gain in 101 women beginning exercise at home for patients receiving cancer chemotherapy and a steroid. The results of the 2 exercise interventions which reduced weight gain from 5.9 kg over 1 year in the control group, to a weight loss of .4 kg in resistance and 2.5 kg in aerobic, are hard to explain. Aerobic exercise for 20 minutes per day, 4 days per week for a year is about 400 – 600 kcal expenditure per week and would be expected to produce about a 2.5 to 3 kg weight loss over a year. The results reported suggest that the aerobic exercise intervention reduced weight loss by 8.4 kg! I was unable to find this study in my library (it is in a very low impact journal) but I have concerns about it and especially about highlighting it in the abstract.	We changed the criteria for a meaningful between group difference threshold. The studies that met the criteria for a meaningful difference have changed after annualizing the threshold. Also note that all studies are associated with low strength of evidence indicating that future research may change the inference.
<b>Peer Reviewer 2</b>	Abstract	The abstract then notes several Effective (word used in the abstract!) approaches with “low strength of evidence” –There really is no evidence to support that these changes are “effective” Rather, these behaviors have been associated with less weight regain, but no trials have shown that making these specific changes are indeed effective. It is critical that the word effective be removed from this sentence and that the sentence states that these findings were from observational trials.	We reported our findings consistent with GRADE. “Low” strength of evidence indicates low confidence that the evidence reflects the true effect and further research is likely to change our confidence in the estimate of the effect and is likely to change the estimate

Commentator & Affiliation	Section	Comment	Response
<b>Peer Reviewer 2</b>	Abstract	The issue of blinding of assessors is important, but this review emphasizes it too much. In a study of weight – where weights are objectively measured on scales (most of which are digital) it is actually hard to see how lack of blinding could influence results. The report notes that the risk of bias was downgraded for “not blinding the assessor to the intervention.” It is unclear what is meant here—I think that in most studies, the assessor knows what interventions are being studied, but they do not know which group the particular participant is in.	We downgraded evidence unless the study specifically stated that the assessor was unaware of which group the participant was assigned to. It is possible that we are downgrading evidence due to poor reporting. As the CONSORT checklist has been around since 1996 and the majority of included trials were published after this time, we feel that the authors had the opportunity to practice good reporting and downgrading is acceptable.
<b>2</b>	General	Finally a smaller point, one of the key questions covered is Orlistat. Orlistat is not used alone; it is by protocol used in combination with a low fat diet; Orlistat is recommended for weight loss; weight maintenance; and prevention of weight regain – prevention of weight gain is not listed as a use	We agree with the reviewer. Because orlistat is available over the counter, we included it in our review in case there was evidence that it was being used for weight maintenance in the absence of prior weight loss. Peer reviewed reports of orlistat being used in this manner were not identified.
<b>Peer Reviewer 2</b>	General	In conclusion, I would recommend MAJOR SUBSTANTIVE CHANGES to this review. Most notably, as stated, I feel that studies using calorie restriction or weight loss as a means to prevent weight gain are very important to include as they likely represent the most effective strategy. I feel that the observational studies should be excluded for the review as they are potentially misleading. Finally, I do not feel that studies of cancer patients and psychiatric patients belong in this report. I hope these important concerns are seriously considered.	Thank you for your comments. In addition to the specific responses above, we summarize our responses: <ol style="list-style-type: none"> <li>1. We were open to including studies of caloric restriction to prevent weight gain. We did not include studies of weight loss. Prevention of weight gain was the nominated topic for this review. Based on what our team and experts felt were recent trends in research, we expanded the population of the review to include prevention of weight gain among overweight and obese persons.</li> <li>2. We selected observational studies with careful control for confounding and selection bias. Including these studies was endorsed by the Technical Expert Panel. Observational studies have the benefit of measuring real world behaviors for a longer period of time than most trials.</li> <li>3. Our technical experts endorsed inclusion of the subgroups. All subgroups were chosen because they are at greater risk of weight gain or are a high priority population for prevention of weight gain based on the improved health outcomes that are thought to be associated with weight gain prevention.</li> </ol>
<b>Peer Reviewer 4</b>	General	The topic is important and timely, the key questions are appropriate, and several aspects of the design are strong.	Thanks

Source: <http://effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?productid=1441&pageaction=displayproduct>

Published Online: March 25, 2013

Commentator & Affiliation	Section	Comment	Response
<b>Peer Reviewer 4</b>	General	The major problems, that each require major revision, are in the selection of the clinically meaningful outcome	The threshold for the clinically meaningful outcome was revised and is not referred to as a meaningful between group difference.
<b>Peer Reviewer 4</b>	General	The major problems, that each require major revision, in the grading of the evidence	Evidence was graded consistent with the EPC Methods Guide using GRADE. In our methods section and appendices we are transparent in showing how we arrived at our strength of evidence grades
<b>Peer Reviewer 4</b>	General	The major problems, that each require major revision, in the omission of observational studies from the pooled estimates (meta-analysis Figures).	In general observational study estimates are pooled separately from trials. We can work to get both study designs in the same figure stratified by type, when relevant.
<b>Peer Reviewer 4</b>	General	The search for KQ6 also does not appear adequate, based on specified inclusion/exclusion criteria.	We also expected to identify more evidence in this area given the interest and publicity surrounding the topic. We found primarily cross-sectional studies, not serial cross-sectional studies (which we had intended to include). We had searched the grey literature during the initial review process to identify missed studies and found none. We searched the gray literature again for the final report.
<b>Peer Reviewer 4</b>	General	Several other key but relatively minor revisions are also needed, detailed below. With appropriate changes, this report will be very important and useful.	Thanks
<b>Peer Reviewer 4</b>	Introduction	Citation for average annual weight gain in the US is from 1993 - this should be updated using more recent data.	We have added an additional citation. (Katan MB, Ludwig DS. Extra calories cause weight gain--but how much? JAMA 2010; 303(1):65-6.PMID:20051571) (Lovasi GS, Hutson MA, Guerra M, Neckerman KM. Built environments and obesity in disadvantaged populations. Epidemiol Rev 2009; 31:7-20.PMID:19589839)
<b>Peer Reviewer 4</b>		Also, authors state that average US adult weight gain is between 0.5 to 1 kg per year – this upper range seems far too high. (The average American gains 44 pounds between age 20 and 40? Implausible). Average US adult weight gain is probably closer to 0.4 to 0.6 kg/year. Also, the distribution (e.g., 10, 25, 50, 75, and 90th percentiles) of annual weight gain in US adults should be detailed.	We have changed our clinically meaningful threshold to 0.5 kg/year.
<b>Peer Reviewer 4</b>	Introduction	Additional major morbidity caused by adiposity should be added, e.g. gallstones, back pain, sleep apnea, depression.	We did not intend to have an exhaustive list.

Source: <http://effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?productid=1441&pageaction=displayproduct>

Published Online: March 25, 2013

Commentator & Affiliation	Section	Comment	Response
<b>Peer Reviewer 4</b>	Introduction	Given the focus on weight gain, and that adiposity is roughly continuous risk factor for most diseases, the relevant topic in the Introduction should not be obesity per se (i.e., a somewhat artificial, dichotomous endpoint of BMI>30), but adiposity. For instance, gaining adiposity to go from a BMI of 22 to 24, or from 26 to 28, is associated with poor health outcomes. So, the term "obesity" should be replaced with "adiposity" throughout the Introduction, with discussion of this latter concept.	Team discussed terminology to be applied were relevant throughout report. Our team, technical experts and stakeholders thought that obesity was appropriate. No change made.
<b>Peer Reviewer 4</b>	Introduction	Page 18, the statement that overweight individuals have similar or better health outcomes than normal weight individuals is flawed and needs revision. Careful analyses in several cohorts have demonstrated that such findings are entirely due to confounding by smoking and reverse causation. Thus, the public health goal should not be keeping adult weights between 18.5 and 29.9 (as stated by the authors, and in stark disagreement with every national, international, or health advocacy organization guideline), but maintenance of normal weight and prevention of further weight gain among overweight or obese individuals.	Agree that statement is surrounded by controversy. Deleted statement as not essential to report.
<b>Peer Reviewer 4</b>	Introduction	The Introduction should include a detailed section (2-3 pages) on methodologic issues for studying long-term weight gain. Specifically, this should include the challenges in studying an endpoint of relatively small magnitude (~0.5 kg/year), a value that is often within the range of year-to-year individual variability (let alone month-to-month or even day-to-day variability), and so requires both (a) considerable numbers of individuals and (b) many years of follow-up to detect valid and reasonably precise effects. Given these specific issues, this section should also discuss the strengths and limitations of various study designs to evaluate long-term weight gain. For example, almost explicitly, a perfectly or nearly perfectly performed, unbiased, long-term RCT of weight gain is impossible. First, for most interventions, participant blinding would be impossible. Second, for many interventions, randomization would be unethical. Third, given the duration of follow-up needed to confirm true long-term effects (3-5 years +), cross-over and noncompliance would seriously limit validity of any trial. Fourth, the numbers of subjects	Thank you for these helpful comments that we have incorporated into the Discussion.  One of our suggested future research needs is identifying the intended weight goal in observational studies.

Source: <http://effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?productid=1441&pageaction=displayproduct>

Published Online: March 25, 2013

Commentator & Affiliation	Section	Comment	Response
		needed to detect small but clinically relevant effects (e.g., changes of 0.1-0.2 kg/year, which would be quite relevant for an outcome in which the average is 0.5 kg/year) would be prohibitively large. Thus, for nearly all relevant interventions, a true long-term (3-5 years +), double-blind, placebo-controlled, sufficiently powered RCT of long-term weight gain with little cross-over, noncompliance, or loss to follow-up is impossible to design or implement. Long-term observational studies also have inherent limitations, but actually may be a better design than most RCTs for determining the best unbiased estimates of true effects on long-term weight gain for many relevant interventions, provided that exposures, outcomes, and relevant covariates are appropriately measured, loss to follow-up is evaluated and considered, and appropriate analyses and careful adjustment for confounding are performed, including relevant sensitivity analyses. Consequently, for this specific research question, the usual criteria for ideal evidence must be carefully considered and revised to recognize the inherent methodologic challenges of studying long-term weight gain (see comments in Methods, below). All these key issues should be described in the detail in the Introduction.	
<b>Peer Reviewer 4</b>	Methods	Many strengths are evident. These include appropriate scope and key questions; appropriate inclusions and exclusions, in particular the evaluation of both trials and observational studies; appropriately broad extracted outcomes; and reasonable approaches to timing and setting.	Thanks



Commentator & Affiliation	Section	Comment	Response
<b>Peer Reviewer 4</b>	Methods	<p>The major revisions that are needed, which will substantially improve the validity and utility of this report, are in the selection of the clinically meaningful outcome and in the grading of the evidence: In addition to considering statistical significance, the authors define a clinically meaningful outcome as at least 0.8 BMI units, 2.5 kg weight, or 2 cm waist circumference. Such thresholds may make sense for intensive studies of substantial weight loss among obese individuals, but have no applicability to the outcome of interest: long-term weight gain. First, these outcomes must be annualized, i.e., in evaluating the clinical relevance of any study's finding on long-term weight gain, the duration of follow-up is crucial. For example, for the outcome of long-term weight gain, a 1 kg difference over 2 years is quite different than a 1 kg difference over 10 years. Thus, these thresholds must be annualized, i.e., as the smallest clinically relevant difference in BMI, weight, or waist circumference per year of follow-up. Second, the magnitude of the clinically relevant difference must be selected in light of the average outcome that is actually observed and that we are trying to prevent. Given that average US adult weight gain is about 0.5 kg/year, any intervention that would prevent a reasonable proportion of this weight gain would be quite meaningful. The choice of proportion is somewhat subjective, but 20-25% seems reasonable. Thus, any intervention that leads to a 0.1 to 0.125 kg/year difference in weight gain would be quite clinically meaningful, on both an average individual level and a population level. Weight gain is a complex, multifactorial process – if we could identify a single intervention that would wipe out 20% of average weight gain (or even 10%!), that would be of tremendous clinical relevance.</p>	<p>Reviewer brings up a great point regarding small differences effects on population level. We changed the threshold and annualized it based on this recommendation.</p>

Commentator & Affiliation	Section	Comment	Response
<b>Peer Reviewer 4</b>	Methods	<p>In grading the strength of evidence, the authors use traditional comparative-effectiveness criteria. These criteria have often been applied to medical treatments such as drugs, devices, or (less frequently) procedures, in which long-term, randomized, double-blind, placebo-controlled, adequately powered trials can often be performed. As discussed above (see Introduction), the outcome (long-term weight gain) and relevant interventions here do not allow, by nature, such studies. Thus, while it seems enticing to fall back on “established” evidence grading criteria, this topic requires careful consideration of the true likely quality of different types of evidence for this specific outcome and these specific interventions; and corresponding careful modification of the evidence grading system to account for the relevant methodologic issues. The same domains can be used, but with revision of the criteria for each. Also, given the outcome, a fifth domain should be added: duration of follow-up. (For studying this outcome, studies of 1-2 years duration are simply robust than studies of 3-4+ years (or ideally, 5-10+ years) duration.) In considering the criteria for each domain, the frank inability to perform a true long-term, double-blind, placebo-controlled, sufficiently powered RCT of long-term weight gain with little cross-over, noncompliance, or loss to follow-up must be recognized. Consequently, for most interventions, the best possible evidence available from trials will be from well-done shorter-term (1-2 year) RCTs, that are not so long-term as to be hopefully flawed by cross-over and noncompliance, but that also would only provide moderate evidence due to their relatively short duration and (for most interventions) inability to be double-blinded. Similarly, given the relative strengths of well-done prospective observational studies for this outcome and most relevant interventions, a well-done, long-term, prospective observational study would also provide moderate evidence.</p>	<p>We agree that a future research need is identifying the best tool to assess quality for non-drug/device interventions.</p> <p>We did modify the Downs and Black criteria when assessing study quality because we agree that a double blind trial standard is unreasonable in this setting (Appendix F).</p> <p>Overall, we felt the domains of Risk of Bias, Consistency, Directness and Precision remain appropriate for the SOE. We have graded outcomes consistent with the EPC Methods Guide, which is based upon the GRADE system.</p> <p>We did include observational studies and we recognize, as does the reviewer, the value of observational studies for this question.</p>

Commentator & Affiliation	Section	Comment	Response
<b>Peer Reviewer 4</b>	Results	The results should be updated throughout to reflect a more relevant annualized outcome (e.g., differences of about 0.1 to 0.2 kg/year or more, or equivalent BMI or WC changes per year), and also the revised evidence grading.	We updated the way that we identified meaningful differences. Our thresholds are annualized.
<b>Peer Reviewer 4</b>	Results	The figures should include pooled analyses from all the studies, i.e., including prospective studies. The main inclusion criteria appropriately include observational studies – these should be pooled by relevant intervention as well.	We did not pool the results of any studies. Because many observational studies provided categories of exposure (quintiles, for example), the observational study results do not contribute to the figures.
<b>Peer Reviewer 4</b>	Results	KQ6: Many, many more observational studies have been done (hundreds) of environmental level exposures and differences in BMI among adults. Were these excluded because they did not assess weight change per se, i.e., because most were cross-sectional? This should be clarified in the Methods, i.e., specifying that cross-sectional studies were excluded a priori. To be precise, such studies are still trying to make inference about what BMI would have been in the counterfactual instance of the alternative environmental setting, which in a sense is no different than weight change. Of course, the temporal direction of the association cannot be ascertained in a cross-sectional study; if this latter point is the reason for exclusion of their studies (which would be fine), then this should be specified in the Methods.	Thank you for helping us to phrase our intent. This was nicely stated. Studies that did not report on weight change over at least one year among the same group of people (or people residing in the same region) were excluded. The majority of cross-sectional studies were excluded for this reason.
<b>Peer Reviewer 4</b>	Discussion	The discussion will need substantial revisions once the definition of a clinically meaningful threshold is corrected (see above), as well the evidence grading criteria.	Agreed. We made these changes after modifying the threshold.
<b>Peer Reviewer 4</b>	Discussion	For environmental approaches, the exclusion of the vast majority of the literature due to its cross-sectional nature should be specified, along with the major limitations of this approach. Specifically, although biased results are minimized, there is a very high likelihood of type II error due to omission of most of the data.	We considered adding type 2 error comment to limitations of our approach. However, our stakeholders and technical experts felt that requiring at least one year of followup was appropriate.

Source: <http://effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?productid=1441&pageaction=displayproduct>

Published Online: March 25, 2013

Commentator & Affiliation	Section	Comment	Response
<b>Peer Reviewer 4</b>	Discussion	In considering the utility and applicability of this report, the bar for “high” evidence should be discussed and possibly re-visited. Given that, for the great majority of relevant interventions, this bar is impossible to meet, might not the bar for “high” be simply the best feasibly-achievable evidence? In other words, should two best types of achievable evidence (well-done shorter-term RCTs, and well-done long-term observational studies) actually provide “high” rather than “moderate” evidence? The downside is that this would not be consistent with other reports on other outcomes; the upside is that this would provide us with a more valid assessment of the “best” potential evidence here.	Given our modification of assessing Risk of Bias and study quality of trials to not require double blind trials, we feel that our revisions make high SOE achievable. Given other reviewers concerns regarding unmeasured confounding in observational studies and unknown intent in these studies, trials of weight gain prevention are appropriate. Understanding when these trials are efficacy compared to effectiveness trials will be important.
<b>Peer Reviewer 4</b>	Discussion	Future research should also include the vital need to create appropriate and relevant evidence grading criteria for complex, multifactorial, long-term exposures and outcomes such as those here, the influences of habitual diet, activity, and environmental exposures and interventions on long-term weight gain.	Agreed.
<b>Peer Reviewer 4</b>	Clarity and Usability	If the main methodologic issues (clinically relevant threshold; evidence grading) are corrected, this will be quite useful. As it stands, this could be more misleading than helpful.	Substantive changes made to threshold. Evidence grading changed based on the new threshold. Thank you for suggesting these changes. We agree the report is more useful based on these suggestions.
<b>TEP 1</b>	General	The report is very well done. It is disappointing that there are not more relevant research studies in this area. The major impact of the report will be to say that there is just not enough data available to determine effectiveness of current strategies for preventing weight gain. Perhaps the major impact of the report will be to facilitate more research in this area. Could it, for example, be used by NIH or USDA to fund more targeted research in this area?	Thanks! PCORI has requested that applicants identify topic areas identified as future research needs in AHRQ reports. We hope that NIH and USDA will also find the report and Future Research Needs document useful.
<b>TEP 1</b>	Introduction	The introduction is appropriate.	Thanks
<b>TEP 1</b>	Methods	I felt that the methods were appropriate.	Thanks
<b>TEP 1</b>	Results	It is a very long report for essentially the conclusion that there are insufficient data to answer the questions asked. However, it may be useful for researchers in this area.	We have tried to make as compact as possible using tables and figures instead of text to ease reading.

Commentator & Affiliation	Section	Comment	Response
<b>TEP 1</b>	Discussion	I do believe the future research section could be strengthened. Since the conclusion is that there is not enough research, the solutions to this could be explored more. Rather than just saying we need more research, can you speculate on how that might happen.	This information will be covered in the future research needs report in more detail. The future research needs document will be based on the low and insufficient evidence from this report and a panel of informants will work to rank the future high priority topics for future research.
<b>Peer Reviewer 5</b>	General	This report attempts to address a clinically meaningful question but due to lack of evidence or lack of strength of the current evidence few recommendations can be made at this time. However, the report identified key areas for future research, which will have implications for clinical practice.	Thanks
<b>Peer Reviewer 5</b>	General	The target population and audience are mentioned in the preface. It would be useful if these were also stated more explicitly in the body of the report.	No change made
<b>Peer Reviewer 5</b>	General	The key questions are appropriate and clearly stated.	Thanks
<b>Peer Reviewer 5</b>	Introduction	In general the introduction is clearly written and provides useful background information. However, the following specific points should be considered.	Thanks
<b>Peer Reviewer 5</b>	Introduction	Page 18- line 20. Studies such as the Flegal study cited here, which suggest that overweight individuals may have better morbidity and mortality outcomes than normal weight individuals have been criticized for not taking issues such as reverse causation into account that can result in biased estimates. The authors should be more cautious in the interpretation of such data.	Deleted sentence suggesting overweight may have better outcomes.
<b>Peer Reviewer 5</b>	Introduction	Strategies to prevent weight gain in normal weight individuals i.e. transitioning from normal weight to overweight are also important from a public health standpoint, particularly for certain ethnic groups such as South Asians, who experience increases in metabolic risk at lower BMI values. This should also be reflected in the background.	We agree with the Reviewer that the meaningful BMI thresholds for overweight and obesity may differ by ethnic group. We consider multiple measures of weight for this reason. We did not describe the controversies surrounding the thresholds for weight status in the report as we felt the discussion may distract rather than add to the results.
<b>Peer Reviewer 5</b>	Introduction	Page 22-line 33-39 (in Table 1, weight gain prevention outcomes): "Maintenance of weight within same BMI category as the baseline measure". Individuals closer to the border of a different BMI category may have greater chance of changing BMI categories despite having less weight gain. How might this be addressed?	Very few studies reported on changes in BMI categories. We also considered BMI measured linearly.

Source: <http://effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?productid=1441&pageaction=displayproduct>

Published Online: March 25, 2013

Commentator & Affiliation	Section	Comment	Response
<b>Peer Reviewer 5</b>	Introduction	Similarly, "Among those with BMI from 18.5 to 30.0 at the first measure, maintenance of weight as non obese."	The proposed topic was weight maintenance among normal weight individuals. We hoped to examine predictors of weight gain prevention among those who were not obese. BMI is a standard measure to categorize weight.
<b>Peer Reviewer 5</b>	Introduction	Individuals with lower BMI at first measure may be more likely to maintain a non-obese status compared to an individual with a greater BMI at first measure, despite having more weight gain. How might this be addressed?	Very few studies reported on changes in BMI categories. We also considered BMI measured linearly.
<b>Peer Reviewer 5</b>	Methods	Inclusion/exclusion: Justifiable	Thanks
<b>Peer Reviewer 5</b>	Methods	Search strategies: Logical and clear	Thanks
<b>Peer Reviewer 5</b>	Methods	Definitions for outcome measures: Page 22 line 28, adverse outcomes for diet should also include physical discomfort i.e. distension etc. It seems that weight maintenance is the primary outcome of interest while other measures such as adverse events, intermediate outcomes and clinical outcomes are secondary outcomes. It may be useful to categorize the outcomes in this manner	We considered reframing as secondary outcomes in methods. We looked for any adverse event with an emphasis on those specified. Very few trials reported on adverse events. No observational study reported on adverse events.
<b>Peer Reviewer 5</b>	Methods	Rating the body of evidence: Page 27 line 20-22: What is meant by non-randomized studies here? Are these observational studies? If so it may not make sense to automatically assign these studies a moderate risk of bias (as opposed to low risk of bias) because a well conducted prospective cohort study could provide more robust data for these questions compared to trials, particularly for dietary measures. Since only studies with a duration of at least 1 year are included, adherence could be a major problem for trials, in which case, observational studies might be better suited.	Changed non-randomized to observational.  We assigned risk of bias consistent with the EPC Methods Guide.
<b>Peer Reviewer 5</b>	Methods	Rating the body of evidence: Please provide rational for why these domains (page 27 line 10) were chosen to rate the evidence. i.e. why was directness chosen? It is possible that a given study not designed to address a particular question can provide just as strong or stronger evidence than a study specifically designed to answer the question, depending on study design and execution.	We used GRADE as stated in the EPC Methods Guide. We agree that a study not specifically designed to prevent weight gain can provide strong evidence, we included these studies for this reason.

Source: <http://effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?productid=1441&pageaction=displayproduct>

Published Online: March 25, 2013

Commentator & Affiliation	Section	Comment	Response
<b>Peer Reviewer 5</b>	Methods	Rating the body of evidence: Although the reference is provided, it would be useful for the Downs and Hill criteria (page 27, line 12) or internal validity and reporting bias to be explained briefly.	The citation provided includes the text used in the original form.
<b>Peer Reviewer 5</b>	Methods	Rating the body of evidence: Page 27, line 17: Please describe what sort of internal validity concerns were considered and how poor reporting was defined	We used GRADE as stated in the EPC Methods Guide. The major internal validity concerns with confounding and selection bias. Poor reporting was defined using the assessment tool of Downs and Black.
<b>Peer Reviewer 5</b>	Methods	Rating the body of evidence: Page 27 line 38: When grading precision for the outcome of adherence, why was a total number of fewer than 400 participants in all studies for a given comparison and outcome considered to be imprecise?	Because no measures of variability were available for adherence, we needed to assess precision in a different manner than the other outcomes. We chose 400 to represent enough participants to say that the adherence estimates reported would be generalizable to other populations. This number would be equivalent to 4 trials of 100 individuals receiving intervention or fewer larger trials.
<b>Peer Reviewer 5</b>	Results	The results section is nicely organized and provides sufficient detail in text, figures and tables and appendices. As far as I can tell, the authors did not omit any study without justification and included studies meet the inclusion/exclusion criteria.	Thanks
<b>Peer Reviewer 5</b>	Results	Page 35 line 47: why is mortality considered an adverse event? should this be a clinical outcome?	Agreed that mortality could be an adverse event or a clinical outcome. The stakeholders and TEP did not comment on our placement of mortality under adverse events. No change made.
<b>Peer Reviewer 5</b>	Results	Page 37 line 10-13: The study by Mozaffarian (ref 56) was intended to identify dietary and lifestyle predictors of long-term weight gain so it is not clear why the authors have stated otherwise.	The study may have been designed to do so, but the source studies (NHS, NHS 2, HPFS) were not designed to study long-term weight gain.
<b>Peer Reviewer 5</b>	Results	Applicability regarding observational studies for self-management (page 37) and diet (page 40): It does not make sense to comment on the fact that the study did not report on adherence or include clinical outcomes when they are observational (i.e. adherence is not a consideration) or not intended to look at other clinical outcomes.	Agreed. We would have reported adherence if this was reported in intervention studies. For KQ 1, no self-management intervention trials and three observational trials were included.
<b>Peer Reviewer 5</b>	Results	Table 3: Duration of Mozaffarian study was reported in the paper as changes in lifestyle behaviors and weight changes within 4-year periods over a period of 20 years in the NHS, 12 years in the NHS II, and 20 years in the HPFS. Please confirm whether duration is reported in other studies where it is stated that duration was not reported.	The maximum duration of followup was reported. The total number of person years across all studies was also reported. The mean duration of followup was not. Because we did not calculate the average followup for other studies by dividing the total person years by the total number of participants, we did not perform the calculation for this study either. No change made.

Source: <http://effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?productid=1441&pageaction=displayproduct>

Published Online: March 25, 2013



Commentator & Affiliation	Section	Comment	Response
<b>Peer Reviewer 5</b>	Results	Page 76 line 41: It is my understanding that only studies with a duration of at least one 1 year were included in this review. If studies did not report duration how can this be verified?	We have corrected this.
<b>Peer Reviewer 5</b>	Results	Page 81 line 26- The PREDIMED study was designed for cardiovascular disease prevention not weight maintenance so it is not surprising that the interventions did not prevent weight gain.	We included studies not designed to prevent weight gain but that tested our interventions of interest because they can also provide evidence.
<b>Peer Reviewer 5</b>	Discussion	The discussion does a good job stating the implications of the major findings clearly and identifying future areas of research. The limitations are described adequately, however, the following should be considered:	Thanks
<b>Peer Reviewer 5</b>	Discussion	Page 114- line 30: Why weren't sugar- sweetened beverages also mentioned from the study by Mozaffarian. According to that study, sugary beverages were the third greatest predictor of weight gain following potatoes and potato chips (consider for results section).	We include beverages in our review. Only the Nurses Health Study 1 and 2 and Health Professionals Followup Study had information. The sugar sweetened beverages did not meet the meaningful between group difference threshold as increased consumption resulted in only 1 pound of weight gain per 4 years.  (Mozaffarian D, Hao T, Rimm EB, Willett WC, Hu FB. Changes in Diet and Lifestyle and Long-Term Weight Gain in Women and Men: New England Journal of Medicine. N Engl J Med 2011; 364(25):2392-404.PMID: <a href="http://dx.doi.org/10.1056/NEJMoa1014296">http://dx.doi.org/10.1056/NEJMoa1014296</a> )
<b>Peer Reviewer 5</b>	Discussion	Page 115 line 10: The lack of findings regarding low-fat intervention is not surprising given what we know about low fat diets i.e. that the fat is usually replaced by refined carbohydrate, which has been associated with weight gain. Low-fat diets also tend to be less satiating.	Agreed
<b>Peer Reviewer 5</b>	Discussion	Page 120 line 27: This is an important point. Observational studies may actually be better suited to answer the questions at hand, since adherence tends to wane as duration of study increases, which is likely problematic when looking at trials of at least one year duration. This was not captured in the rubric used to evaluate study strength and should be mentioned as a potential limitation. Moreover it seems that observational studies were given a lower initial level of strength (moderate bias) compared to randomized trials (low bias)- if they were included among non-randomized studies.	Observational studies are given a lower starting SOE according to GRADE.  Potential methodologic future research need could be focused on understanding adherence to lifestyles in observational studies.

Source: <http://effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?productid=1441&pageaction=displayproduct>

Published Online: March 25, 2013

Commentator & Affiliation	Section	Comment	Response
<b>Peer Reviewer 5</b>	Discussion	Page 121 line 28: The finding that controls were gaining less weight than the general population possibly because of weight surveillance may also be a reason why there was little benefit of the interventions observed.	Agreed. This is the point that we are trying to make.
<b>Peer Reviewer 5</b>	Clarity and usability	The report is well structured and clearly written. The conclusions are useful and point to areas of further research as well as a few potential points of intervention for policy.	Thanks
<b>Peer Reviewer 5</b>	Comment	The report is an important contribution in that it identifies gaps in the literature and potential areas for future research that could have important implications for obesity policy and clinical practice. It is not surprising that the strength of the evidence for the majority of comparisons were low, based on the method used to evaluate study strength. Well-conducted observational studies that adjust for potential confounders may be better suited to these questions since they don't have the problems of adherence and high drop out rates (due to participant burden) that many trials experience. This was not captured in the rubric used to evaluate study strength and could be a potential limitation. In general, applying these types of scoring systems can result in missing important aspects of individual studies that are not otherwise captured in the scoring system. Another potential limitation is lack of discussion about specific interventions in a given study in relation to results. For example the intervention arms in the PREDIMED trial were not designed to prevent weight gain, so it is not surprising that there was no difference between intervention and control for BMI and weight. Also regarding the Women's Health Initiative, we know that a major problem with low-fat diets is substitution of refined carbohydrate, which is associated with weight gain. These sorts of details are important when interpreting the data.	We modified existing quality scores to accommodate these suggestions.

Commentator & Affiliation	Section	Comment	Response
<b>Public Comment (Esai Inc.)</b>	See attached	JHU Summary of the Public comments from Esai: Esai Inc. begins their comments by describing their company. They follow by commending the broad approach taken by the EPC. They support our decision to use clinically meaningful difference and understand our lack of findings about orlistat but caution that the lack of findings about orlistat should not be generalized. Esai closes the letter by thanking AHRQ for the opportunity to review this article.	The EPC appreciates the comments from Esai Inc., but believe that none of the comments require a direct response.
<b>Public Comment (National Heart Lung Blood Institute)</b>	Comment	While we agree that there is a need for a systematic review of the literature on prevention of weight gain in adults, we wonder if this report will be useful to the public, including providers, health plans, and government programs. Most of the approaches identified are not stand-alone interventions to prevent weight gain, but are factors associated with less weight gain; these kinds of findings are not clearly distinguished in the evidence statements and may lead to confusion in interpreting the recommendations. Weight gain prevention needs to take into account a decrease in calories and/or an increase in physical activity. We are concerned that studies that focused on decreased calorie intake appear to have been excluded from the review. We suggest that there be substantial revision to this report before it is released.	After changing our threshold of a meaningful between group difference, more trials are now considered potentially effective.  We would have included studies on decreasing caloric intake for weight gain prevention. However, the studies we identified on decreased caloric intake included participants with a goal of weight loss and were excluded.
<b>Public Comment (NHLBI)</b>	Executive Summary	The title of the report is misleading and should be changed. This report is about weight gain prevention as stated throughout the document, especially in the conclusions (e.g., p. v). Weight maintenance is usually used to refer to maintenance of lost weight. However, it appears that prevention of weight gain is the focus of this report and thus the title should be changed to reflect this.	We agree and this has been changed
<b>Public Comment (NHLBI)</b>	Executive Summary	It would be useful to have a definition of comparative effectiveness as it applies to the interventions reviewed in this report.	We defined clinically meaningful thresholds which could be considered definitions of comparative effectiveness for the interventions.

Source: <http://effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?productid=1441&pageaction=displayproduct>

Published Online: March 25, 2013

Commentator & Affiliation	Section	Comment	Response
<b>Public Comment (NHLBI)</b>	Executive Summary	The Applicability section seems to contradict basic tenets of clinical trial methodology. "Because adherence was poor in many interventions, the results may have been more useful if they had been reported by adherence status." Conclusions from RCTs are based on intention-to-treat principles. If the intervention wasn't strong enough to induce behavior changes, basing conclusions on just those who were adherent violates those principles. Adherence is critical but should influence the quality and strength of the evidence.	Added "...in addition to the intent-to treat analyses." to the sentence.
<b>Public Comment (NHLBI)</b>	Methods	What is the basis and evidence for defining the "clinically meaningful difference" in BMI?	<p>We identified articles from the literature that reported on clinically meaningful thresholds. However, these thresholds were usually in the setting of weight loss. We identified a reference for a clinically meaningful definition for weight maintenance; (Katan MB, Ludwig DS. Extra calories cause weight gain--but how much? JAMA 2010; 303(1):65-6.PMID:20051571))Lovasi GS, Hutson MA, Guerra M, Neckerman KM. Built environments and obesity in disadvantaged populations. Epidemiol Rev 2009; 31:7-20.PMID:19589839) we then applied this to a BMI change for an individual starting with BMI=27 (mean/median in many included studies). For waist circumference, we did not identify a threshold and thought that due to measurement error we should go no lower than 1cm.</p> <p>Following the suggestion that most adults gain 0.5kg/year, we changed our clinically meaningful threshold for the difference in weight maintenance between groups to be 0.5kg for 1 year trials (1 kg for 2 yr, etc.). We created the corresponding value for BMI based on the baseline BMI of 27 (0.2 units in year 1, 0.4 units at 2 years).</p> <p>For waist circumference, we felt 1cm should be the minimum due to measurement error. We recommended the difference in groups to be at least 1 cm for year 1, 2 cm for year 2, etc.</p>

Source: <http://effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?productid=1441&pageaction=displayproduct>

Published Online: March 25, 2013

Commentator & Affiliation	Section	Comment	Response
<b>Public Comment (NHLBI)</b>	Methods	The definition of weight gain prevention is too narrowly defined. Long-term weight gain prevention can include some weight loss and it appears that studies in which some participants lost weight were excluded, even if it wasn't a stated goal of the study. This appears to be the reason why the Pound of Prevention Study (Jeffery RW et al, Am J Public Health 1999), one of the largest, well-conducted adult prevention trials, was excluded.	<p>We identified articles from the literature that reported on clinically meaningful thresholds. However, these thresholds were usually in the setting of weight loss. We identified a reference for a clinically meaningful definition for weight maintenance; we then applied this to a BMI change for an individual starting with BMI=27 (mean/median in many included studies). For waist circumference, we did not identify a threshold and thought that due to measurement error we should go no lower than 1cm.</p> <p>Following the suggestion that most adults gain 0.5 kg/year, we changed our clinically meaningful threshold for the difference in weight maintenance between groups to be 0.5kg for 1 year trials (1kg for 2 yr, etc.). We created the corresponding value for BMI based on the baseline BMI of 27 (0.2 units in year 1, 0.4 units at 2 years)..</p> <p>For waist circumference, we felt 1cm should be the minimum due to measurement error. We recommended the difference in groups to be at least 1 cm for year 1, 2 cm for year 2, etc.</p> <p>"Failed" weight loss to increase weight maintenance among the overweight and obese is beyond the scope of the review. The mechanism to maintain weight among those who have lost weight or are trying to lose weight is clearly different than those who have never been overweight as evidenced by the extreme difficult individuals who have lost weight have with keeping the weight off long-term. We agree that the concept is worthy of further investigation, but the current report aims to examine primary prevention of weight gain.</p> <p>We have clarified our rationale for excluding these studies in the Introduction and Discussion. We agree that this approach is a future research need for weight gain prevention among the overweight and obese.</p>
<b>Public Comment (NHLBI)</b>	Methods	The majority of trials reviewed were not designed to prevent weight gain making it difficult to assess comparative effectiveness of interventions from them.	Agreed. We included this design issue in our assessment of Directness for this reason.

Commentator & Affiliation	Section	Comment	Response
<b>Public Comment (NHLBI)</b>	Methods	To our knowledge, there are very few randomized trials that test weight gain prevention approaches. It appears that this may have influenced AHRQ's decision to include observational studies in a review addressing comparative effectiveness of interventions (n=12 of 44 overall were observational and 11 of 22 were observational that provided evidence for the general population). Evidence from observational studies can inform about associations, but it is unclear how some of these studies address comparative effectiveness of interventions. At the very least, evidence from observational studies should be summarized separately and distinguished from that of RCT.	In general, GRADE recommends grading the strength of evidence of the body of literature, not segregated by study design.  Well designed observational studies are often used for testing the comparative effectiveness of interventions. They require very careful control of confounders, particularly confounding by indication.  Throughout the results, observational studies are summarized after the trials.
<b>Public Comment (NHLBI)</b>	Methods	Studies with any sample size were accepted, suggesting that some studies may have been underpowered.	This limitation is captured in Precision.
<b>Public Comment (NHLBI)</b>	Results	The approaches found to be "effective" (e.g., watching less television, eating meals prepared at home, avoiding potato chips/French Fries) are not "interventions" or stand-alone weight control strategies, and were not tested in isolation of each other or other strategies for preventing weight gain. Importantly, most of these "approaches" were not tested in randomized controlled trials and thus the term 'effective' should not be used in the evidence statements (Results, p. v).	Consistent with the AHRQ Methods Manual for Systematic Reviews, we included observational studies if they could strengthen the body of evidence. Given that we had few trials on a number of the interventions, we chose to include observational studies. We were highly selective in the observational studies we included with strict criteria for confounding control. Additionally, the majority of RCTs do not reflect real world situations that need to occur for long-term prevention of weight gain.
<b>Public Comment (NHLBI)</b>	Results	Only two small trials provided evidence on college-based approaches, which seems inadequate for a 'moderate' strength of evidence.	We grade the strength of evidence based on four domains: risk of bias, consistency, directness and precision. The number of studies in a body of evidence do not directly contribute to these four domains.
<b>Public Comment (NHLBI)</b>	Discussion	Use of the word 'effective' is inappropriate	We define "effective" in the Methods section and use those criteria to interpret if a finding is effective or not based on the evidence identified as part of the comparative effectiveness review. We then tie that finding with the strength of evidence which allows us to make a judgment if the current finding may change with additional information from future studies.

Commentator & Affiliation	Section	Comment	Response
<b>Peer reviewer 6</b>	General	The topic of the report is both significant and meaningful. The report itself is also clinically meaningful yet not in the way the authors originally hoped. Instead of being able to identify promising evidence based practices for the primary prevention of weight gain, it confirms the findings of a prior study by revealing the paucity of quality research addressing this important topic.	Thank you
<b>Peer reviewer 6</b>	General	While the authors cite their rationale for their choice in stratifying their analyses by the different categories, it would have perhaps been more useful (if the available data allow) to look at the effectiveness of the strategies with reference to weight status, gender, race/ethnicity and/or socio-economic variables. Also, the categories used (e.g. general pop. vs. work-based vs. college-based, etc.) are not mutually exclusive and very likely overlapped. Addressing the amount of overlap for each of the sub-populations would be helpful.	The source studies did not provide sufficient information for us to comment on overlap.
<b>Peer reviewer 6</b>	General	the authors don't initially identify a specific population or populations in their research aims, e.g. do they want to prevent healthy weight individuals from becoming overweight or are they more concerned with preventing overweight individuals from becoming obese? As the authors indicate, the obese population has seen significant growth in the last three decades whereas the percentage of the population that is overweight has stayed relatively static. This is mentioned as a potential future area of research.	The nominated topic focused on weight maintenance among normal weight adults. We expanded the population to include overweight and obese individuals based on input from externals informants and experts.
<b>Peer reviewer 6</b>	General	The authors (via the Key Questions) attempt to stratify and examine the self-management, dietary and physical activity approaches individually yet these three are often highly intertwined.	The source studies did not provide sufficient information for us to comment on overlap.
<b>Peer reviewer 6</b>	Introduction	Given that these topics were discussed in later sections of the report, it would be helpful to include mention of the differential challenges/successes of weight maintenance among socio-economic groups as well as the specific challenges among those taking medication that induces weight gain.	Added to introduction.



Commentator & Affiliation	Section	Comment	Response
<b>Peer reviewer 6</b>	Methods	Inclusion/exclusion criteria are clearly stated with the exception of the major decision to exclude studies that mention weight loss as an aim of any study participant. This is of particular concern because it causes the exclusion of the best known weight gain prevention trial, "Pound of Prevention." Little detail is provided regarding the decision to exclude this specific study and the potential impact on the systematic review conclusions. More broadly, although the argument is made that motivations may be different for participating in a weight maintenance vs. a weight loss or weight loss maintenance study, many of the strategies are likely to be quite similar and prior systematic reviews included these studies.	The original nomination and scope of this review focused on interventions to maintain weight. Input from key informants and the technical expert panel affirmed the importance of a review focusing on this particular area. We were aware of the Pound of Prevention trial. Though it was ultimately excluded from this review, we have mentioned it in our discussion section
<b>Peer reviewer 6</b>	Methods	Given the importance of the transition from overweight to obesity in terms of health outcomes, a key question specifically addressing this point would have been helpful. Similarly, a key question addressing known transitional times when weight gain is a problem could have provided more insight for providers. Both of these areas are addressed somewhat in the presentation of results and discussion but less explicitly than if guided by a key question.	The source studies did not provide sufficient information for us to report on this area.
<b>Peer reviewer 6</b>	Methods	It is presumed, but not explicitly stated, that all authors participated in the literature search and review. Were the same two authors responsible for performing the review to identify the relevant publications? (page 4, lines 15-16)	About 10 reviewers participated in the review process.
<b>Peer reviewer 6</b>	Methods	The authors offer no justification for the selection of the clinically meaningful differences criteria (e.g. 0.8 units BMI, 2.5 kg weight or 2 cm waist circumference) (page 4, lines 52-55)	We updated our thresholds based on the peer review comments. (Katan MB, Ludwig DS. Extra calories cause weight gain--but how much? JAMA 2010; 303(1):65-6.PMID:20051571)(Lovasi GS, Hutson MA, Guerra M, Neckerman KM. Built environments and obesity in disadvantaged populations. Epidemiol Rev 2009; 31:7-20.PMID:19589839)
<b>Peer reviewer 6</b>	Methods	The authors discuss the importance of adherence, but do not describe what percent of studies reported on adherence. Was the type of study (e.g. the weight maintenance method utilized) associated with reporting on adherence?	We reported on each study that mentioned adherence.

Source: <http://effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?productid=1441&pageaction=displayproduct>

Published Online: March 25, 2013

Commentator & Affiliation	Section	Comment	Response
<b>Peer reviewer 6</b>	Results	Reporting on the demographics (BMI, race/ethnicity, etc.) within each of the stratifications (e.g. among “adults from a general population” and “work-based approaches”) would have been helpful. Some sections do while others don’t – this may simply be due to data availability.	We reported when the source studies reported.
<b>Peer reviewer 6</b>	Discussion	Would have liked more discussion by the authors concerning their important comments on page 7 of the executive summary stating that “The majority of interventional studies were randomized trials that were not designed to prevent weight gain.” (lines 4-5) and subsequently “Only one of the observational studies came from a cohort that was explicitly designed to measure weight change over time.” (lines 7-8). Especially in light that the overall aim of the paper (as stated by the authors, page 1, line 38) is to “...review studies of strategies for weight maintenance in adults” which arguably implies that the majority of studies under review will share this aim (e.g. at least one of their top research aims will be focused on the primary prevention of weight gain).	<p>We chose to be inclusive and allow studies that specifically mentioned that they aimed to prevent weight gain as well as studies that provided evidence without this design in mind.</p> <p>Discussion, p 145: Very few studies had a stated goal of weight maintenance or weight gain prevention. Goals to change dietary intake and physical activity commonly reported on weight without a weight-related goal. We excluded studies that explicitly mentioned that at least some of the patients had a goal of weight loss. The best known weight gain prevention trial was excluded for this reason, the Pound of Prevention trial.<sup>107</sup> Only one of the included observational studies was nested within a cohort whose original design had a weight related outcome of interest.<sup>67</sup></p> <p>The goals of the studies are now provided in study characteristics tables and text. For example, Results p40: “The stated goal of the intervention in three trials was to prevent weight gain,<sup>57,58,60</sup> prevent increase in percent body and abdominal fat in one trial,<sup>53</sup> change diet or physical activity patterns in five trials,<sup>49,51,54,56,59</sup> improve a cognition score among elderly patients in one trial,<sup>55</sup> and reduce the risk of cardiovascular disease in the community wide trial.<sup>50n</sup>”</p>
<b>Peer reviewer 6</b>	Discussion	Perhaps add more discussion on the differences between primary weight maintenance and weight loss maintenance related to behaviors, motivation, etc. Notably, the authors exclude studies that focus on weight loss and weight loss maintenance yet don’t seem to provide justification as to why they are outside the scope of this review. For those not familiar with the subject matter, their rationale may not be clear.	There is a vast literature on weight loss and weight maintenance after weight loss. Because there are many factors that may drive differences in preventing weight gain and weight loss, we focused on study on preventing weight gain.

Commentator & Affiliation	Section	Comment	Response
<b>Peer reviewer 6</b>	Clarity and usability	Overall, the report is clear and usable. Additional discussion of some of the points mentioned above would improve one's ability to apply the findings to practice or policy. One small and (perhaps trivial critique), is that the semantics of some sentences are sometimes "bulky" and could perhaps flow better if revised. As an example, page 7, lines 36-37, "The strength of evidence is low that workplace based approaches do not prevent weight gain" might be better stated as "Workplace based approaches do not appear to prevent weight gain. However, the strength of the findings supporting this conclusion is low."	Thank you.