BMJ's manuscript committee comments

1.We all found your review interesting. The effect on cancer mortality was unexpected by most of us.
Reply: As requested we have now assessed the quality of the evidence using GRADE (Grading of Recommendations, Assessment, Development and Evaluations). The findings for cancer mortality and cancer events have few events and wide confidence intervals. Both have GRADE ratings of very low quality evidence, partly related to the poorer quality of the trials for sequence generation and allocation concealment. As a consequence, we think that the cancer data, particularly cancer mortality, should be regarded with caution. However, the quality of all-cause mortality and cardiovascular events is high and cardiovascular mortality moderate.

2.Please can you draw out a little more the causes of death. You have highlighted cancer and cardiovascular disease and these appear to account for around a quarter of the 660 or so deaths. If you were able to draw up a table listing the main causes of death as reported and the numbers in each category this would make it clearer where the lower death rates were seen most or whether this is across the board.
Reply: We have provided Table 1 in the Appendix with all-cause mortality, cardiovascular mortality and cancer mortality as main causes of death from the studies. This table illustrates how frequently we were not able to obtain data on causes of death through contacting authors. Even for the largest included studies (Look AHEAD 2013, DPP 2002, TOHP II 2007) we were not able to provide data on cancer mortality. We have discussed these data in the discussion.

3.Please can you update your search for studies. The search is nearly two years old and since many of your results had wide confidence intervals, finding more studies could produce a useful improvement in precision.
Reply: We updated our search until December 2016, and thus introduce data from 6 new trials, including making contact with authors of these trials. Main findings are unchanged. The paper has been completely revised throughout.

Our statistician made the following requests for revision:
4.Authors requested unpublished outcome data, where trials were suggestive that the outcome may have been measured. The possibility that outcomes were measured but not reported/mentioned should also be considered. 3 studies were excluded due to no outcomes - were these study authors contacted for outcome data?
Reply: Yes, we contacted these three authors, who were not able to provide outcome data. We have elaborated more on the theme of missing outcome data, and the need to present these outcomes in trials in the discussion.

5.In the methods section, it is specified that a random effects meta-analysis will be used. For the primary analysis an I^2 value of 0% was found - in what situation would the authors therefore have considered a fixed effects meta-analysis? It was also unclear from the forest plot that these were the results from a random effects meta-analysis.
Reply: All the forest plot headings have now been labelled as random effects meta-analysis. Meta-analyses with small trials with few events, thus wide and greatly overlapping confidence intervals, may have I^2 that is low. We have chosen to be conservative by using random effects meta-analysis, and known heterogeneity in the trials, as demonstrated by the high heterogeneity for weight change.

6.The results section seems to come to an abrupt end with many planned analyses not discussed. Specifically

a) Rare events was a potential issue with the analysis of this data. The authors stated they would perform a Bayesian meta-analysis to address this - was this done, what were the findings? I think it is important to understand how zero events were dealt with in the analysis and some indication that the effect estimate (RR) and method of estimation was appropriate in this situation.
Reply: Analyses in Stata inserts 0.5 for zero events (in the same way as Cochrane's Review Manager). We used a Bayesian meta-analysis approach which is analogous to a multi-level model with random effects for trial and fixed effect for treatment. The results are enclosed as a supplementary file (Appendix Table 5). Estimated RRs were very similar, to the more traditional meta-analysis results reported in the main body of the manuscript. We've also included a fixed effect analysis for your information.

b) The authors also state that a meta-regression would be undertaken to investigate heterogeneity - was this done? The subgroup analyses were also not discussed in the results section.
Reply: We have now explained in the methods that we did not undertake meta-regression of disease outcomes, as a result of apparently low heterogeneity (I-squared = 0%). Subgroup analyses are now discussed in the results and discussion, but with caveats and emphasising the need for individual patient data meta-analyses to address effects in subgroups.

c) The post-hoc analyses were also not discussed in detail. The authors state "the addition of trials from Asian populations did not show significant interactions, apart from participants with cardiovascular events (I2 = 73.7%, P = 0.02)." I did not understand where this result/interpretation came from? Appendix figures 7-11 I think are missing.
Reply: Appendix Table 6 and Appendix Figures 7-11 figures have now been included and the discussion expanded in the results for post-hoc analyses.

7.Please revise your paper to respond to all of the comments by the reviewers. Their reports are available at the end of this letter, below.
Reply: Our replies are given below.

Comments from Reviewers

Reviewer: 1

This is a carefully conducted metaanalysis which contributes to knowledge by providing some further evidence of low (saturated) fat weight reducing diets in overweight or obese individuals. The inability to demonstrate any effect at all with regard to cardiovascular events or mortality and a trend towards a reduction of cancer risk are perhaps surprising. The authors explain this in terms of the fewer trials. However from what we have learned from the metaanalyses of the dietary trials of CVD risk reduction and from what we know of the nutritional causes of cancer, one would be inclined to expect that with relatively short term intervention trials, one would be more likely to see a reduction in cardiovascular events than any reduction in cancer.
I think it would be essential in the discussion to expand on this aspect and not simply to dismiss the unexpected findings as being due to insufficient power.
Reply: We have expanded the discussion to cover these points. The findings for cancer mortality and cancer events have few events and wide confidence intervals. Both have GRADE ratings of very low quality evidence, partly related to the poorer quality of the trials for sequence generation and allocation concealment. As a consequence, we think that the cancer data, particularly cancer mortality, should be regarded with caution.

The findings for all-cause mortality and cardiovascular events are rated by GRADE as high quality, and cardiovascular mortality as moderate quality. Thus, more reliance should be placed on these than the cancer data.

I note the various subgroup analyses. Was there some reason for not also examining the outcomes by duration of followup and extent of weight loss? I appreciate that such analyses may be limited by the number of trials and the dominance of one study but any trends may at least provide some clues. As the data stand at present and given the rather dismissive tone of the discussion, the sceptic may be equally dismissive of the finding of reduction of total mortality presumably largely explained by a reduction in cancer deaths with no change in CVD events or mortality in short-term trials.
Reply: We would disagree that the reduction in all-cause mortality is largely explained by a reduction in cancer deaths for reasons explained above. We don't have adequate data to tease out the reasons for the reduction in all-cause mortality, as shown in the Appendix Table 1 with details on causes of death that we were able to obtain.

We undertook nine subgroup/sensitivity analyses (age, gender, BMI, type 2 diabetes, ethnicity, physical activity, allocation concealment, follow-up, including Asian groups with lower BMIs) for our six outcomes. Examination of the effects of lowering BMI to include more trials with Asian populations was the only *post hoc* analysis. Although interesting to examine effects by degree of weight loss or length of follow-up, we think that these topics would be better explored in an individual patient data meta-analysis, as the ability to find meaningful results from a subgroup analysis for these variables is limited.

Reviewer: 2

This is a well-conducted systematic review and meta-analysis of the effects of weight-loss interventions on mortality (all-cause and cause-specific) in adults. There are a number of strengths: the use of a Bayesian approach in the presence of sparse data; consideration of different BMI cutpoints by ethnicity; the careful documentation of adjudication of events; PROSPERO registration; and a careful analysis of sources of bias. I have but a few suggestions for clarity.

MAJOR
1. Would the authors consider a summary GRADE assessment to express the certainty of the estimates? See http://www.gradeworkinggroup.org/ for guidelines on how to implement. I think this approach would help give readers an overview of the strengths and weaknesses of the body of evidence on which this review is based.
Reply: we have now provided a summary of findings table and GRADE assessments for mortality, cardiovascular and cancer event data.

2. Would the authors please justify the decision to only perform sensitivity analyses on 2 domains of the Cochrane ROB tool (allocation concealment and follow-up)? Why were the other domains not included in the sensitivity analyses?
Reply: we were well aware that we already have many sensitivity/subgroup analyses, and likely risked spurious findings as a consequence. The introduction of GRADE has allowed us to provide greater clarity around the overall influence of study quality on our results.

MINOR
1. Was inter-rater agreement on full-text review calculated (i.e. kappa)? Could it be?
Reply: This was not done, and it is not possible to do this reliably in retrospect.

2. Under results, it appears that the number of trials meeting each characteristic are reported sometimes as numbers (i.e. "four trials with Asian population") and other times as numbers+percent of trials ("Only three trials [6.3%]"). Would the authors consider reporting consistently both the n of trials and the % of the total, please?
Reply: Yes, this has been done.

3. Would the authors also please consider the number of trials and events on which each of the RR reported under the "Meta-analyses" section is based-- e.g. "n=4 trials, 84 events; RR 0.81, 95% CI: 0.70 to 0.94; I2=0%).
Reply: Yes, this has been done as requested in the main text. Word limits preclude giving all of this information in the abstract.

4. Under "Discussion", how was the "absolute relative risk reduction of approximately 1%" figure derived?
Reply: We have removed this 1% figure and provided detail in the abstract and discussion (including confidence intervals) linked directly to the GRADE assessment in terms of deaths per 1000 participants, including 95% confidence intervals.

Reviewer: 3 (Lay review)

As a patient who has struggled with obesity for many years, the issues and questions addressed by this paper are relevant to me, and I believe they would be of interest to others like me and their carers.

The advice to investigators of weight reducing diets to always report clinically important outcomes and adverse effects should be highlighted. I cannot identify any areas that are missing from the paper.

The outcomes analysed in the paper match those that are important to patients like me. I am not aware of other outcomes that could have been considered in this analysis.

The methods used here are beyond the comprehension of most patients, so beyond stating in the conclusion that the study provides further evidence to support existing public health measures I cannot suggest anything further the authors could do to assist doctors in discussions with patients.

There was no patient involvement in this research, which is often appropriate for the type of analysis undertaken. I do not see any way that patients could have assisted in its design or the analysis. Patients could have been involved though in the presentation of the conclusion, to help ensure it could be understood by a lay audience. This might help if the study is picked up by the mainstream press or media.
Reply: we thank the lay reviewer very much indeed for their comments. Should this work be published, we will work with our university's Public Engagement with Research Unit and Press Office to help ensure that the findings can be understood by lay people.

Reviewer: 4

This review has many strengths. These include:
-pre-registration of the protocol and clarity about the analyses that were decided post-hoc -the comprehensive search, based partly on a unique database maintained in Aberdeen -contact with authors to seek further information -clear attention to bias and the role it could have had on the

results As a result, the review has produced valuable new information and I know of no comparable review. This review is very timely, because it speaks to contemporary concerns about the value of weight loss programmes in improving health and the value of advising low fat diets. This is likely to be a highly valued review and attract a good deal of media coverage and thus inform guidelines and public debate. The paper is thoughtfully discussed and the conclusions drawn seem appropriate to me.

I can think of no major issues for the authors to address. I suppose a limitation of the review is that the authors sought information on deaths only from authors where this appeared to have been collected. It's quite likely that some trials may have had these data but not hinted as such and therefore the data were missed. However, I can think of no reason to imagine that missing data, if it were collected, is likely to be different from that where the authors were able to obtain data and so I see no risk of bias.
Reply: We endeavoured to be as inclusive as possible in our contact with authors, and particularly sought data where participants were older, or trials larger, or if some data on morbidity had been collected, e.g. hospital admissions. We agree that ideally it would have been better to collect data from many more trials, but struggled to get data from even the obvious candidate trials.

Secondly, trial authors have various ways of presenting weight loss data (ie imputing/not imputing for missing data). It's not clear whether the authors made any allowances for this or they simply took the authors data as presented. This is not a big issue, because the data are there only to give context to the review. However, it might be nice to be clear.
Reply: Yes, our data on weight change were presented mainly to illustrate that weight loss occurred, so that participants were changing their lifestyles, rather than tie the degree of weight loss to outcomes. Where available we used intent to treat data for weight change, with the most conservative assumptions if also available, e.g. baseline observation carried forward data provided by the investigators for those who dropped out.

Thirdly, all trials suffer loss to follow-up and this could also affect loss of death data. What was the denominator used in the mortality data? Was it all randomised, which assumes people who were lost did not die, or all followed up? There's no perfect way to do this, but it would be nice to spell this out. This only really matters if loss to follow-up differs by trial arm. In general, there's a tendency for people who do not lose weight to be less willing to be followed up and this could lead to higher loss from the control groups, which, if this affects death information, falsely reduce the death rate in the control group, so it is unlikely to explain the results. Was there evidence that loss to follow-up varied by arm?
Reply: The denominators used for all-cause mortality, cardiovascular mortality, cancer mortality, participants with new cardiovascular event or new cancer are all persons randomised. This has been added to the methods. Across all the included trials 84% of participants in intervention groups completed, compared with the same figure (84%) in control groups.

Fourthly, low fat diets have become controversial, rightly or wrongly. This makes the low fat diet issue more prominent than probably it should be. It sounds like the authors made a few suppositions about the type of diet and I wonder if it is worth being more explicit about how they made these inferences?
Reply: Extensive efforts were made to accurately characterise the diets, even if the trial investigators gave little information in the publications. In many cases investigators referred back to previous trials or published materials, which did give sufficient information to confirm that dietary advice was for diets low in fat, presuming that their references to those publications were accurate, e.g. trials referred back to the US Diabetes Prevention Program, particular versions of

Dietary Guidelines for Americans, or the Dietary Approaches to Stop Hypertension diet. These references to older, primary sources were followed up by the authors of this paper.