Rebuttal letter

Manuscript ID: BMJ.2016.032914.R1

Manuscript title: FTO genotype and weight loss: a systematic review and meta-analysis of 9,563 individual participant data from eight randomized controlled trials

Thank you for the opportunity to address the comments from the Reviewers and the Committee members. The authors hope that the Reviewers and Editors will be satisfied with the further amendments which we have made to the manuscript after taking on board the feedback.

**Report from The BMJ’s manuscript committee meeting**

Detailed comments from the meeting:
First, 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.
Please also respond to these additional comments by the committee:

* Please address the points raised by our statistical consultant and included as one of the reviews.

Response: The points raised by the statistical consultant have been addressed in the responses to Reviewer #3.

* FTO is the gene with the highest association with obesity, but MC4R and TMEM18 are also important. According to the latest meta-analysis of GWAS studies, there are now 97 loci that influence body weight, but their influence is small: “The 97 loci account for ~2.7% of BMI variation.” (Nature. 2015 Feb 12;518(7538):197-206.) Please comment in the manuscript on the role of other genes on weight loss (and perhaps include a box describing these genes).

Response: The authors acknowledge that other genes, in addition to FTO, play an important role in influencing body weight. To highlight this, the following text has been added to the manuscript:

- Lines 34-38: “Specifically, 97 loci have been identified as accounting for ~2.7% of BMI variation\(^6\), of which the fat mass and obesity-associated (FTO) gene\(^9\), melanocortin 4 receptor (MC4R) gene\(^10\) and transmembrane protein 18 (TMEM18) gene\(^11\) have shown the strongest associations. Details of key genes associated with BMI are summarised in Appendix 1.”

- The following references have been added to lines 34-38:

- Lines 391-394: “An important limitation is that we evaluated the effect of FTO genotype only and, obesity risk and weight loss is influenced by multiple genes\(^50\), the effect of other obesity-related genes, such as MC4R and TMEM18, on weight loss in response to intervention remains to be determined.”
- Data supplement: A table (Appendix 1) has been added to the data supplement to highlight the key genes identified, so far, to be associated with BMI. This table is as follows:

<table>
<thead>
<tr>
<th>Known</th>
<th>Novel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fat mass and obesity associated</td>
<td>FTO (rs9939609; rs1558902)</td>
</tr>
<tr>
<td>Melanocortin 4 receptor</td>
<td>MC4R (rs6567160)</td>
</tr>
<tr>
<td>Transmembrane protein 18</td>
<td>TMEM18 (rs13021737)</td>
</tr>
<tr>
<td>Glucosamine-6-phosphate deaminase 2</td>
<td>GNPDA2 (rs10938397)</td>
</tr>
<tr>
<td>Homolog B, Endoplasmic Reticulum Export Factor</td>
<td>SEC16B (rs543874)</td>
</tr>
<tr>
<td>ATP/GTP binding protein-like 4</td>
<td>AGBL4 (rs657452)</td>
</tr>
<tr>
<td>Cell adhesion molecule 1</td>
<td>CADM1 (rs12286929)</td>
</tr>
<tr>
<td>Transcription factor 7-like 2</td>
<td>TCF7L2 (rs7903146)</td>
</tr>
<tr>
<td>Syntaxin binding protein 6</td>
<td>STXBP6 (rs10132280)</td>
</tr>
<tr>
<td>Hypoxia inducible factor 1</td>
<td>HIF1AN (rs17094222)</td>
</tr>
</tbody>
</table>

* Can you comment on the role of gene variants and weight maintenance? It may be harder for some individuals with minor FTO alleles to maintain their weight.

Response: The authors agree that there is some evidence for the role of FTO genotype in weight maintenance, but that issue was outside of the scope of the present meta-analysis. Nonetheless, the following text has been added to lines 354-369: “In a pan-European study of 742 adults, multiple SNPs were investigated for effects on weight regain over 6 months ad libitum diet following more than 8% weight loss and for whether nutrient sensitive genes modified weight regain in response to glycaemic index and high protein-based diets. No significant effects were found for the FTO SNPs, and although significant effects were found for several other SNPs, associations were no longer significant following adjustment for multiple testing. Furthermore, although a more recent study found evidence for the specific role of FTO genotype in predicting body weight maintenance, results were based on a smaller (n=128), female only sample, and so the these findings should be interpreted with caution. Given the lack of effective strategies for preventing weight regain and the evidence for an interaction between FTO genotype and physical activity on obesity risk, further research into the role of obesity-related genes and their interactions with diet and physical activity on long-term weight management are needed. Although the role of FTO, and other obesity-associated genes, in weight maintenance was outside of the scope of the present meta-analysis, it is an important consideration for maximising the long term health benefits of weight loss and should be considered in future research.”

In your response please provide, point by point, your replies to the comments made by the reviewers and the editors, explaining how you have dealt with them in the paper.

Comments from Reviewers

Reviewer: 1

Comments:
In this manuscript, the authors used 9,921 individual participant data from 9 studies for the systematic review and 8 studies of them for the meta-analysis. They found differential changes in BMI, body weight and WC in response to weight loss intervention were not significantly different between FTO genotypes. The pooled population was further stratified and analyzed in terms of BMI,
sex, ethnicity, type of intervention and duration of the intervention. Sensitivity analyses indicated that differential changes in BMI, body weight and WC by FTO genotype did not differ by intervention type, intervention length, ethnicity, sample size, sex and baseline BMI and age category.

The manuscript provided a nice review of how FTO genotype influences changes in obesity-related outcomes in randomized weight loss interventions and an implication of future clinical work. This work provides supplemental information to previous work cited in this paper. One suggestion below.

There are two studies, DPP and LookAhead, containing 2835 and 3637 individual records respectively, accounting for 65% of the total individuals of this study and contribute to the majority of the results. Are the results consistent with these two studies or not? Why? I suggest the authors to add such discussions.

Response: The authors thank the Reviewer for highlighting this point and we note that results from Look AHEAD and DPP have been evaluated independently in previous publications (ref 15 and 18, respectively) and have also been analysed in a pooled analysis (ref 20; Diabetes 2015 Dec; 64(12): 4312-4321). Further discussion has been added to lines 337-342: “The present meta-analysis included two studies, Look AHEAD 15 and DPP 18, which together contributed 67% (n=6472) of the total sample. Independently, neither study identified any association between FTO genotype and weight loss 15 18, which was consistent with parallel and pooled analysis of both studies 20 and with our overall findings. In our sensitivity analyses, omission of the largest study, i.e. Look AHEAD, did not change the pattern of results, thus supporting the robustness of our findings.”

Additional Questions:
Please enter your name: Bo Ji

Job Title: postdoc associate

Institution: the Jackson Laboratory

Reimbursement for attending a symposium?: No

A fee for speaking?: No

A fee for organising education?: No

Funds for research?: No

Funds for a member of staff?: No

Fees for consulting?: No

Have you in the past five years been employed by an organisation that may in any way gain or lose financially from the publication of this paper?: No

Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this paper?: No

If you have any competing interests (please see BMJ policy) please declare them here: No competing interest.
The manuscript by Livingstone et al. is very interesting individual-participant data (IPD) meta-analysis on the effect of FTO on weight loss after a dietary, exercise, and pharmacological interventions. In essence, it addresses a question that becomes more and more common: does genetic variation in the FTO predict response to weight loss interventions?

The study overcomes the many limitations of a recently published meta-analysis by Xiang et al. I read the specific meta-analysis and I agree with Livingstone et al. that the Xiang paper has some important methodological limitations. In particular, the conclusions of that paper cannot be directly inferred based on the reported analysis. As Livingstone et al mention, the absence of gene-treatment interaction is a substantial limitation for the Xiang paper.

In contrast, Livingstone et al. carefully evaluate gene-treatment interactions in their analysis. The statistical analysis is very robust and the right meta-analytical methods have been applied to address the questions of interest. In addition, the IPD nature of the analysis and the registration in PROSPERO are indicative of a well-done and credible study.

My main comments pertain to reporting issues. First, the statistical analysis could be more detailed given that there are 2 sets of coefficients as shown in Table 2. I had to read this section many times to understand exactly how the analysis was done and what was meta-analyzed and map this to the numbers in Table 2. I think that you should explain in more words in the main text what each coefficient represents.

Response: Thank you for identifying this lack of clarity.

The following text has been added to lines 153-173 to detail what each coefficient represents:

“Two sets of regression coefficients are presented in Table 2. The first set of coefficients is arm-specific, and captures within-arm change in the outcome during the course of the study per copy of the FTO minor allele. Both treatment and control arms are included in these calculations. The second set of coefficients is intervention-specific, and captures treatment vs. control differences in outcome change scores per copy of the FTO minor allele, i.e. gene X treatment interaction effects. So, for studies with K treatment arms and a single control arm, K+1 coefficients are used to capture FTO effects on within-arm change scores (set 1), and K coefficients are used to capture FTO effects on treatment vs. control differences in change scores (set 2). It is important to note that the first set of coefficients is estimated from separate groups of study participants, rendering them statistically independent and facilitating standard error calculations for the treatment vs. control contrasts given by the second set of coefficients. However, the second set of coefficients uses a common control group for each study, which introduces dependence at the study level.

This second set of regression coefficients from Table 2 was used to evaluate differences in FTO allelic effects on the outcomes of interest between intervention and control groups. Random effects models were used to estimate the pooled effect sizes and to account for both sampling error and inter-study population variation. Meta-estimates were weighted by the inverse of the variance of the effect size (that is, 1/variance), where variance took into account the two potential sources of variation (i.e. within- and between-studies variance).”

Furthermore, the following changes have been made to Table 2:
- Table title renamed: “Moderation by FTO genotype of intervention effects on change in obesity-related outcomes in randomized controlled trials”
- The second column has been renamed as “Study Arm”
- Footnote 2 has been reworded: “2. Values represent the coefficient and 95% CI for differential change in BMI, body weight, and WC between the treatment and control arms by FTO minor allele (allele-dose model was employed and coded in terms of copies of minor allele (0, 1, 2))”

Lastly, the following change has been made to the section heading in the Results in line 261: “Assessment of gene x treatment interaction effects on change in obesity-related outcomes”

Second, the numbers in the manuscript, tables, and figure 2 are slightly different due to rounding. I would recommend reporting the same number of decimal digits throughout the manuscript for consistency.

Response: Thank you for identifying this inconsistency. Results in the manuscript text and tables were presented to two decimal places whereas results in Figure 2 were presented to three decimal places. The latter have been changed to 2 decimal places.

Third, I believe that the "Study quality and publication bias" section should be expanded to include more information (see below).

Response: This section has been expanded based on the changes described below.

Some additional comments/questions. Why are both Egger's test and Begg's test used? Wouldn't one be enough?

Response: Both tests are statistical tests for assessing funnel plot asymmetry. Begg’s rank correlation test examines the association between effect size estimates and their variances. If present, it is most likely due to smaller studies reporting larger effect sizes, which is prima facie evidence of publication bias i.e. omission of small studies with small effects due to lack of statistical significance for their findings. This test is known to be underpowered, especially when the number of primary studies is small. Egger’s test has higher specificity and power than Begg’s, because it uses the actual effect sizes, rather than their ranks. In Egger’s test, the regression line between study precision (independent variable) and effect size (dependent variable) is plotted and this regression line is weighted by the inverse of variance. In the absence of publication bias, the intercept of this regression line should equal zero, thus providing a formal way of testing the null hypothesis.

In light of these arguments, we have retained Egger’s tests and removed Begg’s tests from the manuscript. The following addition was made to lines 196-199: “Small-study effects were appraised by visual inspection of funnel plots of effect size against the standard error, with asymmetry assessed formally with Egger’s test, chosen over Begg’s test for its greater specificity and power 32, where a P-value < 0.1 was considered as significant 33.”

Also, I think that it would be more appropriate to use the term small-study effects (instead of publication bias).

Response: The authors agree and have made the following changes to the manuscript:
- Line 196; Line 255 and Line 258: “small-study effects”
By the way, you mention funnel plots and Galbraith plots but I did not see these plots somewhere. Did I miss them? The Appendix that I have includes only 4 tables.

Response: Funnel and Galbraith plots have been added to the data supplement as Appendix 6 and Appendix 7 respectively.

By the way, in Appendix 3, can you provide the p-values for the meta-regressions?

Response: P values were all non-significant (as denoted in the footnote), and so the authors do not consider them relevant for inclusion.

In regards to the IPD, can you provide some details on how this was done? Currently, the “Search Strategy” and the other sections in the Methods imply that a literature-based meta-analysis was conducted. Do I understand correct that after the literature search was done, PIs from eligible studies were conducted? If so, please clarify the stages from literature search to IPD analysis.

Response: To provide further clarification, the following text was added to lines 75-81: “As detailed below, an initial systematic literature search was undertaken to identify eligible studies for inclusion in the meta-analysis. Once eligible studies were identified, the corresponding authors of eligible studies were invited to contribute individual participant data or to undertake a prescribed statistical plan (if individual participant data were not available for sharing). If authors did not reply, then the corresponding studies were excluded from the meta-analysis, but study characteristics were reported in Table 1.”

That being said, what is the publication status of the studies contributed to this IPD? Were these published and now they contributed to the IPD? Or are there some studies that were not published but were found through relevant searches and their PIs were contacted? For example, you mention that 3 studies did not contribute data -- if they are published you can report their characteristics and findings but decide not to include them in the IPD statistical synthesis.

Response: All studies that were deemed eligible had published data on weight loss and FTO genotype. As indicated above, study characteristics for those studies that did not provide individual participant data (n=3) have been added to Table 1.

Page 19: Study quality involves additional items than publication bias. How was study quality evaluated and where is the relevant information (did I not receive it from the Journal?)? You can present this in a table using the format of the Cochrane Risk of Bias tool.

Response: The Cochrane Risk of Bias tool is referenced in line 114 and line 256 and outcomes from use of this tool have been added to the data supplement in Appendix 3.

Table 1: Does the MAF value refer to the whole cohort of randomized individuals?

Response: Yes. This has been clarified in the footnote of Table 1: “MAF, minor allele frequency for all randomized individuals who provided genetic consent and whose DNA data passed quality control procedures”.

What is the relation between the studies included in this analysis compared to the Xiang et al. paper? Is there any overlap? Are the 2 sets totally different?
Response: To clarify which studies were included in Xiang et al. and the subsequent overlap with the present study, the following text has been added to lines 320-323: “Of the 10 studies included by Xiang et al., 6 overlapped with studies included in the present meta-analysis 13-17 19, with the remaining studies being either RCTs 38 that did not agree to provide individual level data for the present analysis (n=1) or non-RCTs 39-41 (n=3).”

Overall, I think that this is a well-conducted IPD meta-analysis that has considerable strengths over the Xiang paper (which, by the way, is far from an optimal meta-analysis). I include here some comments on reporting issues of the Livingstone paper and I look forward to a revised version.

Additional Questions:
Please enter your name: Orestis Panagiotou

Job Title: Researcher

Institution: National Cancer Institute

Reimbursement for attending a symposium?: No

A fee for speaking?: No

A fee for organising education?: No

Funds for research?: No

Funds for a member of staff?: No

Fees for consulting?: No

Have you in the past five years been employed by an organisation that may in any way gain or lose financially from the publication of this paper?: No

Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this paper?: No

If you have any competing interests (please see BMJ policy) please declare them here: No competing interests

Reviewer: 3

Comments:
I found this paper difficult to follow in terms of IPD and meta analyses and how these fitted together. I have many queries and points to note re the presentation and analyses as given below:

Response: The authors hope that the responses detailed below will answer these queries.

1. It is not clear how a fixed-effect, inverse-variance meta-analysis produces an overall estimate of FTO genotype where there are multiple treatment arms. Although there is a reference, a short explanation would be helpful. Additionally, a sentence on what the estimate obtained actually corresponds to in terms of the competing treatments should be given.
Response: Thank you for highlighting this lack of clarity. The following text has been added to lines 174-194:

“For studies with K>1 treatment arms and a single control arm, including all K treatment vs. control differences from Table 2 in the meta-analysis shown in Figure 2, would have invalidated the standard error calculations programmed into our meta-analytic packages, as these packages assume independent intervention effects, not ones based on comparing different treatment arms with a shared control group. For such multi-arm studies, a fixed-effect, inverse-variance meta-analysis was used to produce an overall estimate of \textit{FTO} effects across all treatment arms, in the absence of within-study heterogeneity. That overall \textit{FTO} effect across treatment arms was then compared with the \textit{FTO} effect on the placebo arm, thus satisfying statistical independence assumptions underlying standard error calculations.

A further complication arose from the fact that our meta-analytic packages require that effects be entered in terms of mean differences, their standard deviation, and sample size. As our within-study meta-analyses returned point estimates and their standard errors instead, we arbitrarily fixed our effective sample size for the combined intervention effect to the sum of the sample sizes of all K treatment groups, and then back-calculated the standard deviation required for the combined intervention effect to have the right level of precision. As the forest plot depends solely on the product of the standard deviation with the square root of the reported sample size, the arbitrariness in these calculations did not affect the validity of the forest plot itself.”

For example, the DREW study assessed the effect of three different exercise regimens vs. control: 4 kcal/kg/wk, 8 kcal/kg/wk, and 12 kcal/kg/week. All three intervention effects were reported separately in Table 2, with sample sizes of N=84, 70, 64 respectively. However, Figure 2 only included a combined \textit{FTO} effect across all N=218 intervention participants, and compared it to the \textit{FTO} effect obtained among N=60 control participants. Of note, no evidence of \textit{FTO} effect heterogeneity was obtained across exercise regimens, further justifying our decision to merge the three intervention arms. A similar approach was used for the different intervention regimens evaluated in the Food4Me study, for the two types of diet supplements evaluated in the PREDIMED study, and the lifestyle and drug interventions employed by DPP.

2. Treating covariates as continuums if they are such is generally considered preferable, since it corresponds to a feasible clinical model (which categorisation does not) and is more powerful in determining true patterns. Hence age, BMI and intervention length should be treated as continuums with appropriate modelling.

Response: Variables such as age and BMI were treated as continuums when entered in the model as covariates. As detailed in line 147-150, models were adjusted for age, sex, baseline outcome (BMI, body weight or WC), ethnicity, country/centre, socio-economic status, physical activity and smoking. Information on whether covariates were continuous or categorical has been added: “Models were adjusted for age (continuous), sex, baseline outcome (BMI, body weight or WC; continuous), ethnicity (categorical), country/centre (categorical), socio-economic status (categorical), physical activity (continuous where possible) and smoking (categorical; Appendix 4).”

However, they were discretized when used to split the sample into strata for stratum-specific analyses as specified in lines 200-204: “To explore potential sources of heterogeneity, moderation testing was conducted using intervention type (diet and diet and/or exercise), intervention length (≤6 months and >6 months), age (<50 years and ≥50 years), sex (binary), BMI (<30 kg/m2 and ≥30...)
kg/m²), and race/ethnicity (Caucasian, Black/African American and Hispanic) as putative categorical moderators.”

<table>
<thead>
<tr>
<th></th>
<th>Original (N)</th>
<th>Analysed (N)</th>
<th>Excluded² (N, %)</th>
<th>Missing outcome³ (N, %)</th>
<th>Missing covariate³ (N, %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DPP</td>
<td>2865</td>
<td>2835</td>
<td>30 (1.1)</td>
<td>28 (93.3)</td>
<td>2 (6.7)</td>
</tr>
<tr>
<td>DREW</td>
<td>464</td>
<td>278</td>
<td>186 (40.1)</td>
<td>50 (26.9)</td>
<td>136 (73.1)</td>
</tr>
<tr>
<td>Finnish DPS</td>
<td>460</td>
<td>264</td>
<td>196 (40.6)</td>
<td>70 (35.7)</td>
<td>126 (64.3)</td>
</tr>
<tr>
<td>Food4Me</td>
<td>1607</td>
<td>671</td>
<td>936 (58.3)</td>
<td>343 (73.0)</td>
<td>127 (27.0)</td>
</tr>
<tr>
<td>Look AHEAD</td>
<td>3882</td>
<td>3637</td>
<td>245 (6.3)</td>
<td>155 (63.3)</td>
<td>90 (36.7)</td>
</tr>
<tr>
<td>POUNDS LOST</td>
<td>811</td>
<td>600</td>
<td>211 (26.0)</td>
<td>166 (78.9)</td>
<td>45 (21.3)</td>
</tr>
<tr>
<td>PREDIMED</td>
<td>778</td>
<td>735</td>
<td>43 (5.5)</td>
<td>40 (93.0)</td>
<td>3 (7.0)</td>
</tr>
<tr>
<td>NUGENOB</td>
<td>764</td>
<td>543</td>
<td>221 (28.9)</td>
<td>135 (61.1)</td>
<td>86 (38.9)</td>
</tr>
<tr>
<td>Total</td>
<td>11631</td>
<td>9563</td>
<td>2068 (17.8)</td>
<td>988 (61.7)</td>
<td>615 (38.4)</td>
</tr>
</tbody>
</table>

3. Patients without complete data were excluded. How many does this apply to? What evidence is there that the included subgroup are not biased? Multiple imputation should be considered. At the very least, there should be details of how many were excluded, from which studies, and why they had incomplete data.

Response: To clarify how many participants were excluded from the analyses from each study, the following table has been added to the data supplement (Appendix 5) and referenced in line 243.

Appendix 5 Number of participants excluded from each study included in the quantitative synthesis due to missing data on outcomes and/or covariates
1. n=466 were excluded for non-missing data reasons: these participants were not overweight or obese. The percentages of individuals with missing data did not include these individuals
2. Percentage of individuals excluded is estimated as a percentage of the total original participants
3. Percentage of missing data for outcomes and covariates is estimated as a percentage of total excluded individuals

As shown in Appendix 5, a large proportion of missingness (61.7%) was due to missing outcome data, as distinct from missing data for covariates. Given that multiple imputation of outcome data under a non-ignorable missingness (NMAR) model can introduce further bias, the authors did not think that it was appropriate for use in this analysis. Missing outcome data for overweight or obese participants is a common finding in lifestyle interventions and so this has been acknowledged as a potential limitation in lines 409-411: “In addition, weight loss interventions may be biased by higher dropout rates than other RCTs,” which may have influenced our ability to identify a relationship between FTO genotype and weight loss.”

4. What is meant by "Where applicable, analytic samples were stratified at the individual level by BMI ... and/or duration of treatment"? Some of these were already adjusted for in the analyses. For the factors associated with mode of treatment (type and duration), the question is whether outcomes are different in these subgroups and so interaction models would be appropriate?

Response: Table 2 presents FTO allelic effects on treatment vs. control differences in change scores adjusted for baseline values of the outcome, age, sex, ethnicity, SES, country/centre, physical activity and smoking. In these analyses, only main effects of covariates were included in the regression models and were treated as a nuisance, rather than being of interest in themselves.

In Table 3 these analyses were further stratified by putative categorical moderators of FTO effects. Cross-stratum comparisons tested a moderation hypothesis, whereas within-stratum significance tests help elucidate the nature of the interaction, if present. Variables employed as stratification factors in Table 3 were removed as covariates from the corresponding regression model.
The following text has been added to lines 200-210: “To explore potential sources of heterogeneity, moderation testing was conducted using intervention type (diet and diet and/or exercise), intervention length (≤6 months and >6 months), age (<50 years and ≥50 years), sex (binary), BMI (<30 kg/m² and ≥30 kg/m²), and race/ethnicity (Caucasian, Black/African American and Hispanic) as putative categorical moderators. Between-stratum comparisons in Table 3 were used to test the hypothesis that the relationship between FTO genotype and intervention effects on change in obesity-related outcomes differed across levels of the moderator, whereas within-stratum tests helped elucidate the nature of the interaction, if present. Variables employed as stratification factors were removed as covariates from the corresponding regression model. In addition, intervention length and sample size were investigated as continuous variables in moderation analyses.”

In addition, to improve the clarify of the manuscript, the results from the moderation testing were combined with the sensitivity analyses section in line 276.

5. For the secondary analysis, AA/AT vs TT, please explain how this relates to the primary and number of alleles. (Maybe this is obvious to all clinical readers? But nonetheless would be helpful to clarify.)

Response: The following text has been added to lines 211-213: “As a secondary analysis, FTO genotypes were recoded using a dominant model, where individuals with one (AT) and two (AA) copies of the minor allele were grouped together and compared with individuals with no copies of the minor allele (TT).”

6. It seems a bit arbitrary to perform a sensitivity analysis that just removes the study that contributed to most to analysis of BMI. How was this decision taken? If doing such a removal (which I don't think is usual for these types of analysis), why not also for the other outcomes?

Response: The authors agree that this is not clear. Look AHEAD was identified as the largest study and contributed the most participants to the analyses and so, as a sensitivity analysis, this study was removed from the analysis of BMI, BW and WC (lines 215-218). As detailed in lines 290-293, results after the removal of the largest study are presented for BMI, BW and WC: “Following removal of the largest study, i.e. Look AHEAD, overall effect sizes for BMI, body weight and WC in Figure 2 changed to -0.09 ([95% CI: -0.22 to 0.04] kg/m²; P=0.17), -0.21 ([-.56 to 0.14] kg; P=0.25) and -0.01 ([-.45 to 0.42] cm; P=0.96) respectively, but failed to attain statistical significance.”

7. There is a typo in the first paragraph of the results. There were 12 (not 13) studies excluded as they did not report results for the FTO genotype reported in figure 1, and this also corresponds to the numbers removed to give 9 remaining.

Response: Thank you for highlighting this error. This has been amended accordingly.

8. The inclusion criteria are studies designed to induce weight loss. Why then are 9 studies identified to be included in the review with 1 then excluded from the meta-analysis since it gave results for weight maintenance rather than weight loss? The role of the DiGenes study is very confusing to me. The results state that IPD from all 9 studies were analysed, but figure 1 states that 9 were in the qualitative synthesis only. It seems to me that the DiGenes study plays a minimal role in the paper, being excluded from table 2 and beyond, and yet the numbers for that study are included even in the title total. I wonder whether this paper should be included at all.
Response: The authors agree that the inclusion of Diogenes in the present paper is not consistent with the inclusion criteria. To avoid confusion, Diogenes has been excluded and the title, text and tables have been updated accordingly.

Furthermore, to show that the focus of the manuscript is on weight loss, “reduction” has been specified in line 9.

9. Quality assessment breakdown and the funnel plot should be presented.

Response: The Cochrane Risk of Bias tool is referenced in line 114 and line 256 and has been added to the data supplement in Appendix 3. Funnel plots have been added to Appendix 6.

10. Main outcome is change in BMI, body weight and WC from baseline but the studies included have follow up times ranging from 10 weeks to 3 years. How are these combined? Are the allele effects and treatment vs control differences in table 2 comparable since over such different time frames? How can they be interpreted? I appreciate that there is a secondary analysis that considers duration alongside a host of other potential covariates and is found non-significant (Appendix 3) but this is a very cursory look, effectively comparing DREW, Food4Me and NUGENOB with Look AHEAD, POUNDS LOST and PREDIMED (6 month dichotomy). Intervention length should be properly investigated as a continuum. I do not think that an analysis combining outcomes at 10 weeks with those after 3 years really gives me interpretable measures.

Response: The authors acknowledge that intervention length is an important consideration in the evaluation of weight loss interventions. To improve the evaluation of intervention length on the relationship between FTO genotype and weight loss, intervention length was treated as a continuum (estimated by number of weeks). The results from this moderation analysis was non-significant.

The following additions were made to the manuscript to reflect this additional analysis:
- Lines 209-210: “In addition, intervention length was investigated as continuous variable in moderation analyses.”
- Lines 283-286: “In addition, investigation of intervention length as continuous variable did not significantly affect the relationship between FTO genotype and treatment-control differences in change in obesity-related outcomes (BMI, body weight and WC) following weight loss interventions.”

11. Table 2: (i) The Food4Me groupings are not clear. Why is T3 the only phenotype and control ‘genotype’? (ii) MD should be defined (PREDIMED)

Response: To clarify these points, the following text has been added to the Table 2 footnote: “PN, Personalised nutrition based on information on diet (T1), diet and phenotype (T2) or diet, phenotype and genotype (T3); MD, Mediterranean diet”. In addition, the study arms names for Food4Me have been updated in Table 2.

12. It seems remarkable (figure 2) that estimated heterogeneity (I2 coefficient) is zero when combining the effects of such different interventions (drugs vs exercise vs different dietary combinations), although the confidence intervals are very wide. Does this require comment?

Response: The authors were also surprised by this findings and double checked results by running the analyses in Revman, STATA, and the R package “meta”. Although the packages differed in how they calculated the 95% confidence intervals for the I2 coefficient, all point and interval estimates were consistent in identifying low levels of heterogeneity in the data (I2=0).
13. Appendix 4, Meta regression: BMI seems to be counted as both categoric and continuous. See notes above re continuums. Many of the confidence intervals will be very wide and some discussion should be given to the power of the study to identify differences of clinical importance.

Response: As detailed in the response to point 2 and in lines 200-210, BMI and related anthropometric variables were treated as a continuum when entered in the model as covariates, but as categorical when used as stratifying variables. For the purposes of the moderation analyses, BMI, BW and WC were assessed according to BMI category (<30 kg/m\(^2\); ≥30 kg/m\(^2\)).

The authors would also like to highlight that these post hoc calculations are exploratory analyses and that the focus of the manuscript is on the main outcomes. We do not wish to over-emphasise the outcomes from these post hoc investigations.

14. Table 3: If age and BMI were treated as the continuums suggested then these analyses would not need to exclude NUGENOB and PREDIMED. Clarify what the 'P' columns are. In particular, are the 2nd set of P values for each comparison tests of interaction? So that there may be a significant difference for the over 50s but this is not significantly different to the change seen in the younger age group?

Response: As detailed above, age and BMI were not treated as a continuum for the purposes of the moderation analyses. The reason for not using continuum was that it was not possible, given that each study was analysed independently (according to the same statistical procedure) and that two studies (Look AHEAD and DPP) conducted the statistical analyses themselves and provided the summary data. NUGENOB and PREDIMED were thus excluded because, as denoted in the footnote to Table 3, the age distributions of the study participants were not being suitable for the chosen cut offs. I.e. the cut offs were <50 years and ≥50 years and the age ranges in NUGENOB and PREDIMED were 50-80 years and 20-50 years, respectively. Similarly for BMI, the cut offs were <30 kg/m\(^2\) and ≥30 kg/m\(^2\) and the BMI range in NUGENOB was 26.4-66.1 kg/m\(^2\).

Both P values are produced in Revman upon generation of a stratified meta-analysis forest plot. The first P value represents the test for significance of the effect size within a given stratum compared with zero (e.g. whether the change in BMI is significant in individuals over the age of 50 years only). The second P value represents the outcome of the test for significance of the effect size between strata (e.g. whether the change in BMI differs significantly between individuals under and over the age of 50 years).

To clarify interpretation of Table 3, the following changes have been made:

- Table 3 title: “Stratified analyses to identify potential moderators of the relationship between FTO genotype and intervention effects on change in obesity-related outcomes”
- The footnotes of Table 3: “1, Values for “mean difference” represent linear regression coefficients for differential change between treatment (combined) and control arms in BMI, body weight or waist circumference by FTO genotype (allele-dose model) and 95% CI. All results have been stratified by levels of the putative moderator. 2, P values indicate whether within-stratum “mean differences” are significantly different to zero. 3, P values indicate whether “mean differences” differ significantly between strata.”
- Lines 277-286: “The relationship between FTO genotype and treatment-control differences in change in obesity-related outcomes (BMI, body weight and WC) following weight loss interventions was not influenced by study type, study length, sex, race/ethnicity, or BMI category (Table 3). Change in BMI per copy of the FTO minor allele was significant in individuals ≥50 years (beta: -0.23 [95% CI: -0.44 to -0.22] kg/m2; I20 (0 to 75); P=0.03), but
not in individuals <50 years (beta: -0.04 [95% CI: -0.26 to 0.18] kg/m²; I² = 0 to 75); however, the corresponding moderation test failed to attain significance (P = 0.75). In addition, intervention length modelled as a continuous variable did not significantly affect the relationship between \( FTO \) genotype and treatment-control differences in change in obesity-related outcomes (BMI, body weight and WC) following weight loss interventions.”

15. Appendix 4 and associated text (second last sentence of results): The numbers do not seem to correspond. For example, how does an average increase of 1.35 (1.17, 1.56) in BMI per copy of the \( FTO \) minor allele relate to the values in the first column of appendix 4 table (0.42, 0.37 etc.)? If these are coefficients from the baseline values regression, then editing the footnote to state that these are 'differences' rather than 'change' would be clearer for the reader.

Response: The authors acknowledge that this section was not clear. The last sentence of the paragraph referred to a sensitivity analysis pertaining to change in obesity-related outcomes following the intervention, and so has been moved to earlier in the section. Text corresponding to the appendix (now Appendix 8) pertained to an investigation into associations between \( FTO \) genotype and baseline BMI, BM and WC and so has been moved to a new paragraph under the heading “Baseline Association Testing”. In addition, further text has been added to this paragraph and to the Methods section to clarify that the data in the table represent the individual regression coefficients for differences in baseline BMI, BW and WC by \( FTO \) genotype and that the data in the text represent the fixed effect inverse variance meta-analysis effect size for all studies:

- Lines 222-224: “These analyses were adjusted for age, sex, ethnicity, country/centre, socio-economic status, physical activity and smoking and aggregated using a fixed-effects, inverse variance meta-analysis.”
- Lines 295-301:
  "Baseline Association Testing
  To facilitate better interpretation of any effect of \( FTO \) genotype on weight loss following an intervention, we investigated the association between \( FTO \) genotype and obesity measures at baseline. \( FTO \) genotype differences in BMI, body weight and waist circumference at baseline for each study are presented in Appendix 8. Collectively, for each copy of the \( FTO \) minor allele, BMI, body weight and WC increased by 0.31 ([95% CI: 0.14 to 0.47] kg/m²; P < 0.001), 0.89 ([0.45 to 1.32] kg; P < 0.001) and 0.63 ([0.28 to 0.98] cm; P < 0.001) respectively.”

There are significant differences in baseline variables which may be clinically important. Why has this happened within RCTs? What effect might this have on the interpretation of the results? Those with the \( FTO \) alleles are significantly larger at baseline, so might they be expected to respond more to any intervention?

Response: As reported in line 218, the authors investigated differences in baseline BMI, BW and WC by \( FTO \) genotype, not by treatment arm. In additional models examining effects of \( FTO \) genotype on change in BMI, BW and WC, baseline values of the outcome were included as covariates in all of our regression models. This is specified in line 222. As these models were calculated separately by study arm, they allow for participants with higher initial baseline weight to benefit more from active treatment, possibly counteracting any adverse \( FTO \) effects at baseline.

16. Authors of eligible studies were approached to either contribute IPD for analysis or to perform analyses themselves according to a prescribed statistical plan. How many followed either route? I think it may be useful to give this information, even if only for others considering a similar analysis. Presumably, it is accepted that all analyses were ultimately undertaken following the statistical plan as envisaged? Is there any evidence of this? What checks were undertaken? For example, if authors
performed their own analyses but did not submit the data was this verified against their published results? If IPD was submitted was this also verified against published results?

Response: An overview of the IPD methodology has been added to lines 75-81: “As detailed below, an initial systematic literature search was undertaken to identify eligible studies for inclusion in the meta-analysis. Once eligible studies were identified, corresponding authors of eligible studies were invited to contribute individual participant data or to undertake a prescribed statistical plan (if individual participant data were not available for sharing). If authors did not reply then they were excluded from the meta-analysis, but study characteristics were reported in Table 1.”

Furthermore, information on the number of studies that provided IPD and summary statistics has been added to lines 234-237: “A total of 8 studies were included in the present meta-analysis. Individual participant data were analysed from all 8 studies: 5 studies provided individual participant data and 3 studies performed our prescribed statistical analysis and provided summary level data.”

Regarding the analysis, log files returned by the three studies that performed the statistical analysis according to the statistical plan were checked for consistency with the prescribed plan. All three studies performed the analysis according to the plan and provided us with the data. Details of this step has been added to lines 139-142: “Log files returned by studies that provided summary level data were checked for consistency with the prescribed statistical plan. Results from all individual participant data analyses were compared with published studies to ensure consistency in the pattern of results.”

Three studies refused to participate. The only information given from these is the total number of participants, but it would be helpful, if possible, to have more information as to why they refused and any specific differences in their included patients.

Response: These studies were not included in the present meta-analysis because although authors were contacted (by email) but we did not receive any response. As a result, we were unable to provide information on their reasons for not participating. Information on these studies and the participants have been added to the qualitative synthesis in Table 1 and in lines 241-254.

17. Discussion: What is meant by 'carefully controlled RCTs'?

Response: The words “carefully controlled” have been removed.

18. Implications: The authors state that the predisposition associated with FTO can be counteracted through interventions, but if those with the allele start off more obese (point 15) and lose no more (outcomes from study) are the effects actually counteracted? (Also, last sentence abstract conclusions.)

Response: All our regression models for change in obesity-related outcomes were arm-specific, and included baseline values of the outcome as covariates. Hence, they allowed for treatment effects to differ by baseline values of the outcome, allowing more obese subjects to benefit more from the intervention, and partially counteracting adverse FTO effects on baseline weight. Whether adverse FTO effects were fully counteracted in practice would require fitting mediational models at the study level that combined direct FTO effects on outcome change scores with indirect FTO effects via higher baseline values of the outcome. Fitting and meta-analyzing such study-level effects is indeed possible, but lies beyond the scope of the current paper.
We have, therefore, removed the statement that genetic predisposition to obesity associated with the minor FTO allele can be fully counteracted through dietary, physical activity or drug-based weight-loss interventions:

- Line 26: “These findings show that individuals carrying the minor allele respond equally well to dietary, physical activity or drug-based weight loss interventions and thus genetic predisposition to obesity associated with FTO minor allele can be at least partly counteracted through such interventions.”
- Lines 418-421: “Importantly, our findings show that the genetic predisposition to obesity associated with the FTO minor allele can be at least partly counteracted through dietary, exercise or drug-based weight loss interventions and that those carrying the minor allele respond equally well to such interventions.”

The final sentence of the implications does not relate to the results from this study.

Response: The final sentence of the implications is designed to highlight the fact that we did not observe a deleterious effect of FTO genotype on a wide range of weight loss interventions. This is important because it provides evidence that lifestyle behaviours, notably diet and exercise, should remain the focus of public health strategies for obesity management. This sentence has been revised as follows: “Future public health strategies for the management of obesity should aim to induce long-term improvements in lifestyle behaviours, such as eating patterns and in physical activity, since these will be effective in achieving sustained weight loss irrespective of FTO genotype.”

Additional Questions:
Please enter your name: Angela Wade
Job Title: Professor of Medical Statistics
Institution: UCL Institute of Child Health
Reimbursement for attending a symposium?: No
A fee for speaking?: No
A fee for organising education?: No
Funds for research?: No
Funds for a member of staff?: No
Fees for consulting?: No
Have you in the past five years been employed by an organisation that may in any way gain or lose financially from the publication of this paper?: No
Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this paper?: No
If you have any competing interests (please see BMJ policy) please declare them here: None
**Information for submitting a revision**

Deadline: Your revised manuscript should be returned within one month.

How to submit your revised article: Log into http://mc.manuscriptcentral.com/bmj and enter your Author Center, where you will find your manuscript title listed under "Manuscripts with Decisions." Under "Actions," click on "Create a Revision." Your manuscript number has been appended to denote a revision.

You will be unable to make your revisions on the originally submitted version of the manuscript. Instead, revise your manuscript using a word processing program and save it on your computer. Once the revised manuscript is prepared, you can upload it and submit it through your Author Center. When submitting your revised manuscript, you will be able to respond to the comments made by the reviewer(s) and Committee in the space provided. You can use this space to document any changes you make to the original manuscript and to explain your responses. In order to expedite the processing of the revised manuscript, please be as specific as possible in your response to the reviewer(s). As well as submitting your revised manuscript, we also require a copy of the manuscript with changes highlighted. Please upload this as a supplemental file with file designation ‘Revised Manuscript Marked copy’. Your original files are available to you when you upload your revised manuscript. Please delete any redundant files before completing the submission.

When you revise and return your manuscript, please take note of all the following points about revising your article. Even if an item, such as a competing interests statement, was present and correct in the original draft of your paper, please check that it has not slipped out during revision. Please include these items in the revised manuscript to comply with BMJ style (see: http://www.bmj.com/about-bmj/resources-authors/article-submission/article-requirements and http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists).

Items to include with your revision (see http://www.bmj.com/about-bmj/resources-authors/article-types/research):

1. What this paper adds/what is already known box (as described at http://resources.bmj.com/bmj/authors/types-of-article/research)

2. Name of the ethics committee or IRB, ID# of the approval, and a statement that participants gave informed consent before taking part. If ethics committee approval was not required, please state so clearly and explain the reasons why (see http://resources.bmj.com/bmj/authors/editorial-policies/guidelines.)

3. Patient confidentiality forms when appropriate (see http://resources.bmj.com/bmj/authors/editorial-policies/copy_of_patient-confidentiality).

4. Competing interests statement (see http://resources.bmj.com/bmj/authors/editorial-policies/competing-interests)

5. Contributorship statement+ guarantor (see http://resources.bmj.com/bmj/authors/article-submission/authorship-contribution)

6. Transparency statement: (see http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists/transparency-policy)
7. Copyright statement/licence for publication (see http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists/copyright-open-access-and-permission-reuse)

8. Data sharing statement (see http://www.bmj.com/about-bmj/resources-authors/article-types/research)

9. Funding statement and statement of the independence of researchers from funders (see http://resources.bmj.com/bmj/authors/article-submission/article-requirements).

10. Patient involvement statement (see http://www.bmj.com/about-bmj/resources-authors/article-types/research).

11. Please ensure the paper complies with The BMJ’s style, as detailed below:

a. Title: this should include the study design eg "systematic review and meta-analysis."

b. Abstract: Please include a structured abstract with key summary statistics, as explained below (also see http://resources.bmj.com/bmj/authors/types-of-article/research). For every clinical trial - and for any other registered study- the last line of the abstract must list the study registration number and the name of the register.

c. Introduction: This should cover no more than three paragraphs, focusing on the research question and your reasons for asking it now.

d. Methods: For an intervention study the manuscript should include enough information about the intervention(s) and comparator(s) (even if this was usual care) for reviewers and readers to understand fully what happened in the study. To enable readers to replicate your work or implement the interventions in their own practice please also provide (uploaded as one or more supplemental files, including video and audio files where appropriate) any relevant detailed descriptions and materials. Alternatively, please provide in the manuscript urls to openly accessible websites where these materials can be found.

e. Results: Please report statistical aspects of the study in line with the Statistical Analyses and Methods in the Published Literature (SAMPL) guidelines http://www.equator-network.org/reporting-guidelines/sampl/. Please include in the results section of your structured abstract (and, of course, in the article’s results section) the following terms, as appropriate:

i. For a clinical trial: Absolute event rates among experimental and control groups; RRR (relative risk reduction); NNT or NNH (number needed to treat or harm) and its 95% confidence interval (or, if the trial is of a public health intervention, number helped per 1000 or 100,000.)
ii. For a cohort study: Absolute event rates over time (eg 10 years) among exposed and non-exposed groups; RRR (relative risk reduction.)
iii. For a case control study: OR (odds ratio) for strength of association between exposure and outcome.
iv. For a study of a diagnostic test: Sensitivity and specificity; PPV and NPV (positive and negative predictive values.)
v. For a systematic review and/or meta-analysis: Point estimates and confidence intervals for the main results; one or more references for the statistical package(s) used to analyse the data, eg RevMan for a systematic review. There is no need to provide a formal reference for a very widely used package that will be very familiar to general readers eg STATA, but please say in the text which
version you used. For articles that include explicit statements of the quality of evidence and strength of recommendations, we prefer reporting using the GRADE system.

f. Discussion: To minimise the risk of careful explanation giving way to polemic, please write the discussion section of your paper in a structured way. Please follow this structure: i) statement of principal findings of the study; ii) strengths and weaknesses of the study; iii) strengths and weaknesses in relation to other studies, discussing important differences in results; iv) what your study adds (whenever possible please discuss your study in the light of relevant systematic reviews and meta-analyses); vi) meaning of the study, including possible explanations and implications for clinicians and policymakers and other researchers; vi) how your study could promote better decisions; vi) unanswered questions and future research

g. Footnotes and statements

Online and print publication: All original research in The BMJ is published with open access. Our open access policy is detailed here: http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists/copyright-open-access-and-permission-reuse. The full text online version of your article, if accepted after revision, will be the indexed citable version (full details are at http://resources.bmj.com/bmj/about-bmj/the-bmjs-publishing-model). The print and iPad BMJ will carry an abridged version of your article. This abridged version of the article is essentially an evidence abstract called BMJ pico, which we would like you to write using the template downloadable at http://resources.bmj.com/bmj/authors/bmj-pico. Publication of research on bmj.com is definitive and is not simply interim "epublication ahead of print", so if you do not wish to abridge your article using BMJ pico, you will be able to opt for online only publication. Please let us know if you would prefer this option. If your article is accepted we will invite you to submit a video abstract, lasting no longer than 4 minutes, and based on the information in your paper’s BMJ pico evidence abstract. The content and focus of the video must relate directly to the study that has been accepted for publication by The BMJ, and should not stray beyond the data.

Response: The above listed items have been addressed in the revised manuscript.

END

References