Author's response to reviews

Title: A randomized, controlled trial of dietary improvement for adults with major depression (the 'SMILES' trial)

Authors:

Felice Jacka (f.jacka@deakin.edu.au)

Adrienne O'Neil (adrienne.oneil@unimelb.edu.au)

Catherine Itsiopoulos (C.Itsiopoulos@latrobe.edu.au)

Rachelle Opie (R.Opie@latrobe.edu.au)

Sue Cotton (sue.cotton@orygen.org.au)

David Castle (David.CASTLE@svha.org.au)

Sarah Dash (s.dash@deakin.edu.au)

Cathrine Mihalopoulos (cathy.mihalopoulos@deakin.edu.au)

Mary Lou Chatterton (m.chatterton@deakin.edu.au)

Laima Brazionis (laimab@unimelb.edu.au)

Olivia Dean (olivia.dean@barwonhealth.org.au)

Allison Hodge (Allison.Hodge@cancervic.org.au)

Michael Berk (mikebe@barwonhealth.org.au)

Version: 1 **Date:** 22 Nov 2016

Author's response to reviews:

A formated response to reviewers has been included as a supplementary file

BMED-D-16-01131

A randomized, controlled trial of dietary improvement for adults with major depression (the 'SMILES' trial)

Response to Reviewer reports:

Reviewer #1:

1.1 Page 3, line 27: the comment that intervention studies are lacking is not completely true. There are a substantial number of RCTs investigating the role of nutrient supplements which could be mentioned.

Response: The possible utility of supplements as treatment options in psychiatry is of interest, but we believe a topic quite separate to an investigation of 'whole of diet' interventions. The available data repeatedly supports the contention that supplements are in no way equivalent to dietary intake. Moreover, the available (now extensive) observational and animal study evidence from the field of nutritional psychiatry points to the relationship between diet and mental health as being of far more complexity than simply an issue of nutrient deficiency, which is addressed to some extent by supplements.

However, given that many readers will likely share a similar point of view to reviewer #1, we have added a deal of extra information in the introduction, including the following sentence and recent reference (page 5, lines 81-84):

"While there are data suggesting that some nutritional supplements may be of utility as adjunctive therapies in psychiatric disorders [19], the field of research focusing on the relationships between overall dietary quality and mental disorders is new and has thus far been largely limited to animal studies and observational studies in humans."

1.2 Secondly, there are some (but not many) RCTs focusing on dietary intake in general that should be acknowledged (Forsyth A et al. Psychiatric Research 2015, Sánchez-Villegas A et al. BMC Medicine 2013). It should also be acknowledged that both trials failed to show a statistically significant impact of a dietary intervention on the prevention/treatment of depression.

Response: We have now included a discussion of these two studies in the introduction (Page 5: Lines 88-104):

"We conducted a systematic review and identified a number of interventions with a dietary change component that had examined mental health related outcomes [21]. While approximately half of these studies reported improvements in measures of depression or anxiety following the intervention, at the time of the review no studies fulfilling quality criteria had been conducted in mental health populations or had been designed to test the hypothesis that dietary improvement might result in improvements in mental health. Since then, one study has been published evaluating the possible impact of a lifestyle program, comprising both diet and exercise, on mental health symptoms in patients with depression and/or anxiety; this study failed to show any differences in symptom levels between those in the intervention and those in the attention control group [22]. On the other hand, post hoc analysis of a large-scale intervention trial provides preliminary support for dietary improvement as a strategy for the primary prevention of depression. Individuals at increased risk for cardiovascular events were randomized to a Mediterranean diet supplemented with either extra-virgin olive oil or mixed nuts, or a low-fat control diet [11]. While not statistically powered to assess the effectiveness of the intervention for preventing depression, there was evidence (albeit non-significant) of a reduced risk for incident depression for those randomized to a Mediterranean diet with nuts. This protective effect was statistically significant in those with type 2 diabetes, who comprised approximately half the sample [23].

1.3 Page 5, line 1-2: potential study participants were screened to have a poor dietary quality using the DST tool. Please add what specific score cut-point was used to identify persons with a poor dietary quality. Most importantly, also state the relevance and validity of this tool for your sample. The tool was originally developed for and validated in a much older population (persons aged 73-94 y) and was developed to detect nutritional risk in older persons.

Response: The dietary screening tool was used only to screen for eligibility for the study. It is a basic dietary tool, the contents of which were slightly modified to be relevant to the Australian diet. It asked questions about the frequency of intake of key food groups, including wholegrains, fruit, vegetables, lean meat and fish and types of fats used – all of which are relevant to the diet we were implementing. The use of this questionnaire needs to be understood in the context of the wider field of nutritional epidemiology where, at the time of the study (and still), there were extremely limited options for quickly and easily assessing diet quality. It should also be noted that we have previously conducted an extensive number of epidemiological studies using data from non-validated questionnaires, for this very reason, and have shown consistently that such

tools can be confidently used in a way that affords a ranking of participants in terms of their diet quality and yields insights into associations with mental health outcomes. The cut-point for the DST was stated in the manuscript on page 6 in the Inclusion criteria section.

1.4 Page 6, line 5-31: It remains unclear what the content of the dietary intervention was. Please describe in detail what the ModiMedDiet consists of. Was it based on nutritional advice only or was it based on a prescribed, personalized diet that met specific food group targets (and if so, what were these specific targets)?

Response: We have now added substantially to the information regarding the dietary intervention in the section focused on the dietary intervention (Page 8: Lines 160-174):

"The dietary intervention comprised personalised dietary advice and nutritional counselling support, including motivational interviewing, goal setting and mindful eating, from a clinical dietitian in order to support optimal adherence to the recommended diet. This comprised the 'ModiMedDiet', developed by RO and CI, which was based on the Australian Dietary guidelines [28] and the Dietary Guidelines for Adults in Greece [29] and is concordant with our previous dietary recommendations for the prevention of depression [30]. The primary focus was on increasing diet quality by supporting the recommended consumption of the following 12 key food groups (recommended servings in brackets): wholegrains (5-8 servings per day); vegetables (6 per day); fruit (3 per day), legumes (3-4 per week); low fat and unsweetened dairy foods (2-3 per day); raw and unsalted nuts (1 per day); fish (at least 2 per week); lean red meats (3-4 per week) [31], chicken (2-3 per week); eggs (up to 6 per week); and olive oil (3 tablespoons per day), while reducing intake of 'extras' foods, such as sweets, refined cereals, fried food, fast-food, processed meats and sugary drinks (no more than 3 per week). Red or white wine consumption beyond 2 standard drinks per day, and all other alcohol (e.g. spirits, beer), were included within the 'extras' food group. Individuals were advised to select red wine preferably and only drink with meals."

1.5 What was the adherence to the ModiMedDiet diet (mentioning the diet was 'easy to follow' (line 24) is not sufficient) and attendance rate of dietary support sessions?

Response: We note the following information already provided in the manuscript:

Page 14 reports on completer analysis / attendance rates

Page 15 (Lines 349-359) now gives further details regarding the changes in dietary intakes:

"At intervention cessation, the dietary support group had significant improvements in the consumption of the following food groups; wholegrain cereals (mean increase 1.21 (SD 1.77) servings/day); fruit (0.46 (0.71) servings/day); dairy (0.52 (0.72) servings/day); olive oil (0.42 (0.49) servings/day); pulses (1.40 (2.39) servings/week); and fish (1.12 (2.65) servings/week). With respect to the consumption of unhealthy food items, intake of extras substantially declined (mean decrease 21.76 (SD 16.01) serving/week) in the dietary support group. Conversely, there were no significant changes observed in the social support control group for any of the key food groups. These findings were confirmed by analysis of the ModiMedDiet scores: the dietary support group showed significantly greater improvement from baseline to 12 weeks on ModiMedDiet scores than controls, $t(55.6)=-4.78$, $p<0.001$; the differences remained after controlling for sex, education, physical activity and baseline BMI. Cohen's d for ModiMedDiet was 1.36 (95% CI 0.74-1.98)."

Additionally, Page 16 (Lines 369-372) and figure 3 highlight the association between adherence to the ModiMedDiet and depressive symptomatology, "Finally, dietary adherence, measured using ModiMedDiet scores, was associated with change in depression scores: there was a statistically significant difference in the change (improvement) in MADRS scores across quartiles of ModiMedDiet scores ($p=0.018$) from baseline to final assessment (Figure 3)"

1.6 Page 8, line 19: Dietary quality was assessed using the ModiMedDiet score. Again, provide detailed information how this score was created.

Response: We have added the following information to the Methods section:

Page 11; Lines 241-244:

"Dietary assessments, using 7-day food diaries, were administered at baseline and endpoint to both groups to identify dietary changes and adherence to the recommended diet; this was done by assessing change in the ModiMedDiet score, which is based on the consumption of the key food groups."

Page 8-9; Lines 177-191:

"The method for scoring the ModiMedDiet is similar to those used in PREDIMED [32] and the Framingham Offspring Cohort [33]. It is a criterion-based diet score that uses pre-defined absolute or normative goals of consumption for specific food items, independent of the individual's characteristics. It was developed based on the recommended intakes of the 12 food group components that comprise the ModiMedDiet (as above), and is out of a theoretical maximum value of 120.

1.7 Also explain on what dietary intake data the score was based (the 7-day food diary, the Cancer Council of Victoria FFQ, or the diet history obtained by the dietitian at the first session)? Also explain why 3 different tools were used to assess dietary intake at baseline and which one was used to determine the (change in) dietary quality score and why.

Response: The dietary data generated from the 7-day food diaries were utilised to generate the ModiMedDiet score (baseline, final and change). This instrument was selected as it is commonly referred to as the 'gold standard' measure of dietary intake (Hoidrup S, 2002).

We used the diet quality tool as a brief screening measure to determine eligibility for the trial. We also used the Cancer Council of Victoria Food Frequency Questionnaire to support future analyses comparing normative data from other large cohort studies in which we are involved, such as the Geelong Osteoporosis Study. This FFQ is widely used in Australia, so our SMILES CCV FFQ data will be used for this end.

As above: Page 11; Lines 241-244:

"Dietary assessments, using 7-day food diaries, were administered at baseline and endpoint to both groups to identify dietary changes and adherence to the recommended diet; this was done by assessing change in the ModiMedDiet score, which is based on the consumption of the key food groups."

1.8 Page 8, line 37: please add why only 67 persons in total were randomized while the power calculation indicated that 88 people per group were necessary. This means that the study is significantly underpowered and the observed effect could be a chance finding.

Response: We have previously (now updated) addressed the issues relating to the sample size in detail in the limitations section of the discussion. This includes a discussion of the possible reasons for our difficulties with recruitment and the impact on our analyses and results.

Page 18; Lines 423-427:

"Clearly, our results must also be considered in light of the small sample size. Failure to reach our planned sample size increases the possibility that ours was a chance finding and/or that our sample was not representative, and limited our ability to conduct subgroup analyses. It may also have inflated the effect size we observed.

And Page 19; Lines 432-434:

"A larger sample size and assessments at more than two time-points would have afforded more sophisticated statistical modelling; this should be a key focus of future replication studies".

Page 19: Lines 435-441:

"The fact that the dietary intervention group was able to make significant improvements to their diet quality suggests that dietary improvement is achievable for those with clinical depression despite the fatigue and lack of motivation that are prominent symptoms of this disorder. On the other hand, the challenges we had with recruiting this clinical population, likely due to the aforementioned symptoms and the requirement to attend the study centre on several occasions, points to the need to utilize different methods for delivering the intervention that do not require attendance with the dietitian in person, such as telephone or Skype."

1.9 Page 11, line 1-2: At baseline the ModiMedDiet score was lower for the dietary group than the control group. Provide a table with the baseline actual dietary intake (energy, nutrients and relevant food groups) of the participants in both groups as well as the actual dietary intake at 12 weeks of follow-up.

Response: As the focus of this study is depression and not the the nutritional status of participants, we respectfully decline to add a great level of detail concerning the nutrient intakes of participants in the SMILES study. However, we have added information on the lipid status of participants at baseline and endpoint (Table 2) and included the following information:

Page 13; Lines 310-312: "The dietary group had significantly lower scores on the dietary screening tool (p=0.043) and the ModiMedDiet score (p=0.031) than the social support control group at baseline, primarily due to lower intakes of fruit and higher intakes of extras. Otherwise, groups were well matched on characteristics."

Page 15: Lines 349-356: At intervention cessation, the dietary support group had significant improvements in the consumption of the following food groups; wholegrain cereals (mean increase 1.21 (SD 1.77) servings/day); fruit (0.46 (0.71) servings/day); dairy (0.52 (0.72) servings/day); olive oil (0.42 (0.49) servings/day); pulses (1.40 (2.39) servings/week); and fish (1.12 (2.65) servings/week). With respect to the consumption of unhealthy food items, intake of extras substantially declined (mean decrease 21.76 (SD 16.01) serving/week) in the dietary support group. Conversely, there were no significant changes observed in the social support control group for any of the key food groups.

1.10 Page 13, line 5: Please add the biomarker results to table 2. Was the change in biomarkers (especially LDL-cholesterol) different between the two intervention groups?

Response: The biomarker data have been detailed in Table 2. Text has also been added to Page 16; Lines 365-372 detailing the analyses.

"Changes in biomarkers are also detailed in Table 2. The only significant difference between the two groups was with respect to change in total polyunsaturated fatty acids; the social support group showed a significant drop in polyunsaturates over the 12 weeks, t(54.9)=-2.41, p=0.019."

1.11 Page 13, line 5: In 12 weeks an improvement in dietary quality will result in a reduction of LDL-cholesterol, especially in persons with a poor dietary quality at baseline. However, the change in MADRS was not correlated with any of the changes in biomarkers. This could suggest that the impact of the dietary intervention on MADRS is mediated through other variables (e.g. inflammation markers as indicated by the authors in the discussion section). Alternatively, this could indicate that the dietary intervention was not successful in changing the actual dietary behaviour and that the impact on MADRS is not due to a change in dietary intake per se but rather due to effects on e.g. structuring daily activities - meals, food preparation, food shopping etc.

Response: We agree that the benefit of dietary change to mental health could be working through a number of pathways, including those suggested. We have expanded the discussion to include these points - Page 18; Lines 409-411:

"Moreover, behavioural changes associated with food (cooking/ shopping/meal patterns) are an expected outcome of a nutrition intervention and these changes in activity may also have had a therapeutic benefit."

However, there were clear changes in dietary intakes in those in the Dietary support group and these changes were correlated with improvements in MADRS scores. This has also been discussed on Page 16: Lines 369-372 and displayed in Figure 3.

"Finally, dietary adherence, measured using ModiMedDiet scores, was associated with change in depression scores: there was a statistically significant difference in the change (improvement) in MADRS scores across quartiles of ModiMedDiet scores (p=0.018) from baseline to final assessment (Figure 3)."

1.12 Page 16, line 7: As 1) this is the first RCT showing a potential effect of dietary improvement on depression severity, 2) the results of the study cannot be fully interpreted at this point due a lack of information regarding the dietary intervention and the actual dietary intake, and 3) as the RCT was performed in a very specific subgroup of persons with depression (recruitment was very difficult as indicated in the discussion, and only those with a low dietary quality were selected) which reduces generalizability, please describe the potential implications of your study for clinical care much more careful.

Response: We have included the following on Page 19; Lines 439 – 445

"Finally, given that we recruited participants on the basis of existing 'poor' quality diet, this may limit the generalizability of our findings to the wider population of individuals with depression. However, evidence suggests that our study sample was not necessarily a special subgroup; in Australia, the recent 2014-15 Australian Health Survey tells us that only 5.6% of Australian adults had an adequate intake of vegetables and fruits. In this study, only 15 out of 166 people screened were excluded on the basis of a pre-existing 'good' diet, suggesting that – concordant with the wider population - poor diet is the norm in those with depressive illness."

And on Page 20; Lines 456-459:

"Clearly, successfully improving diet quality in patients will also benefit the physical illnesses that are so commonly comorbid with depression and which are both a cause and consequence of depression. Upskilling dietitians to best deliver this program to this patient population may also be required and members of our team are developing an educational package to this end."

1.13   Page 3, line 22: please remove the remark 'and not apparently explained by reverse causality (e.g.)'. Observational research approaches can never exclude reverse causality.

Response: We agree with the reviewer's comment and have softened the language, whilst providing extra references to studies where reverse causality has been explicitly investigated (Page 4: Lines 66-68):

"Whilst cognisant of the limitations of observational data, these associations are usually observed to be independent of socioeconomic status, education and other potentially confounding variables and not necessarily explained by reverse causality (e.g. [6-9])."

1.14 Page 14, line 37: please delete 'widespread belief' as there is also ample scientific evidence showing that a healthy diet is more costly than an unhealthy diet (e.g. Darmon, N. and A. Drewnowski, Contribution of food prices and diet cost to socioeconomic disparities in diet quality and health: a systematic review and analysis. Nutr Rev, 2015. 73(10): p. 643-60;

Kettings, C., A.J. Sinclair, and M. Voevodin, A healthy diet consistent with Australian health recommendations is too expensive for welfare-dependent families. Aust N Z J Public Health, 2009. 33(6): p. 566-72; Rao, M., et al., Do healthier foods and diet patterns cost more than less healthy options? A systematic review and meta-analysis. BMJ Open, 2013. 3(12): p. e004277.)

Response: We have now amended and added the following text (Page 17: Lines 389-395)

"While there are many data to suggest that eating a healthier diet is more expensive than an unhealthier diet [42], our detailed modelling of the costs of 20 of the SMILES participants' baseline diets compared to the costs of the diet we advocated showed that our strategy can be affordable [43]. Indeed, we estimated that participants spent an average of AU$138 per week on food and beverages for personal consumption at baseline, while the costs per person per week for the diet we recommended was AU$112 per week, with both estimations based on mid-range product costs [43]."

1.15 Table 1: The ModiMedDiet score of the control group is 44.96 in table 1, but 44.9 in table 2. This seems not correct.

Response: We noted a discrepancy in our figures and have now checked and amended these values

1.16 Table 2: Please also provide adjusted estimates.

Response: As there were no differences in adjusted versus unadjusted analyses, and cognisant of the limitations of our statistical power, we respectfully decline to add these data to the table.

1.17 Figure 2: The Y-axis should start at 0 (zero).

Response: We have made this change

1.18 The other figure 2 (?): How was adherence to the ModiMedDiet assessed? Or is the figure showing quartiles of the 12-week change in ModiMedDiet score, which is something completely different?

Response: We have now added substantial detail regarding how ModiMedDiet scores were calculated (see response 1.5), and changed the text in the manuscript to aid in clarification: (Page 16; Lines 369-372)

'Finally, dietary adherence, measured using ModiMedDiet scores, was associated with change in depression scores: there was a statistically significant difference in the change (improvement) in MADRS scores across quartiles of ModiMedDiet scores (p 0.018) from baseline to final assessment (Figure 3).''

And added to the Figure legend:

"Figure 3.   Change (improvement) in MADRS scores from baseline to endpoint across quartiles of adherence to the ModiMedDiet. Higher adherence to the diet was correlated with greater improvements in depression scores (p=0.018)"

Reviewer #2: The study represents a randomized clinical trial lasting for 3 months in which the effect of dietary counseling vs control treatment (non-dietary counseling) was tested in a sample of depressed participants. In general, the clinical trial literature in depression is plagued by very small samples, overstatement of effects, p-hacking and fishing using numerous outcomes, and questionable statistical analyses. This paper stands out in some regards because it is written in a detailed and very transparent manner, and the authors clearly mention the limitations of the present study. Overall, I recommend a major revision, and I have a larger number of suggestions for the authors.

Abstract

2.1 Methods: states "adjunctive" dietary intervention; could the authors clarify whether participants also took medication or were in psychotherapy during the trial in the abstract, if space permits? This seems quite relevant to interpret the rate of improvement.

Response: We have now done this:

"Of these, 55 were utilising some form of therapy: 21 were using psychotherapy and pharmacotherapy combined; 9 were using exclusively psychotherapy; and 25 were using only pharmacotherapy."

2.2 Results: "67 were enrolled" – how many people have no missing data?

Response: Of the 67 enrolled, 56 (83.6%) had complete data at 3 months. Please note that this information was documented in the manuscript under Completers analysis. We have also added this information to the Abstract (Lines 40-42):

"There were 31 in the diet support group and 25 in the social support control group who had complete data at 12 weeks."

2.3 Results: the authors report effect size for the MADRS main effect, but not for the remission effect (p=0.028). If possible for this chi-square test, could the authors also report the effect size here?

Response: Respectfully, The Numbers Needed to Treat (NNT) does provide an index of effect for binary outcomes and is considered to best reflect clinical significance (Kraemer, H.C., Kupfer, D.J. (2006). Size of treatment effects and their importance to clinical research and practice. Biological Psychiatry, 59, 990-996).

2.4 Conclusion: this seems quite strongly overstated. I will get to this in the remainder of the review below, but I would suggest to tone down the conclusion to be more consistent with the evidence presented in the paper.

Response: We agree and have amended our discussion throughout to reflect the preliminary nature of the data. For example:

First line of Discussion:

"These results provide preliminary RCT evidence for dietary improvement as an efficacious treatment strategy for treating major depressive episodes."

And Page 20; Lines 461-470:

"In summary, this is the first RCT to explicitly seek to answer the question 'If I improve my diet, will my mental health improve?" Whilst emphasizing the preliminary nature of this study and the imperative for replication in studies with larger sample sizes, the results of our study suggest that dietary improvement guided by a clinical dietitian may provide an efficacious treatment strategy for the management of this highly prevalent mental disorder. Future work in this new field of nutritional psychiatry research should focus on replication, ensuring larger samples and more sophisticated study designs, in order to confirm effects and afford sensitivity analyses to identify predictors of treatment response."

Introduction

Overall, the introduction is very thoroughly written, detailed, and the rationale is clear and well presented.

Response: Thank you

2.5 There are hundreds of papers that try numerous treatment options for depression, from hypothermia over all sorts of food supplements to Ketamin and recently even substances aimed at growing the hippocampus. What is often missing is a hypothesis or suggested pathway on how this treatment is supposed to alleviate a common and often chronic mental illness – major depression – that is in many cases preceded by a major life event. It would be insightful for readers if the authors would try to address this point.

Response: As detailed in the introduction, this trial and its hypothesis is based on a very large and robust evidence base from the observational literature and animal studies. We have expanded the introduction and made reference to the various mechanisms of action that have been posited to explain the observed associations, along with the most relevant references (Page 4: Lines 69-80):

"Recently, a meta-analysis confirmed that adherence to a 'healthy' dietary pattern, comprising higher intakes of fruit and vegetables, fish, and whole grains, was associated with a reduced likelihood of depression in adults [1]. Similarly, another meta-analysis reported that higher adherence to a Mediterranean diet was associated with a 30% reduced risk for depression, with no evidence for publication bias [10]. The Mediterranean diet is recognized as a healthful dietary pattern and has been extensively associated with chronic disease risk reduction [11]. More recently, a systematic review confirmed relationships between unhealthy dietary patterns, characterized by higher intakes of foods with saturated fat and refined carbohydrates, and processed food products, and poorer mental health in children and adolescents [2]. Several cohort studies also reported associations between the quality of women's diets during pregnancy and the risk for emotional dysregulation in children [12-14], with new insights into potential mechanisms of action that include brain plasticity [15]; the gut microbiota [16]; and inflammatory [17] and oxidative stress [18] pathways."

2.6 The authors mention "reverse causality" before they propose their hypothesis regarding causation (food -> depression). It may be helpful to restructure this section somewhat to make that clearer (page 3 end of first paragraph).

Response: We have amended the first sentence of the introduction to make our contention regarding the nature and direction of the relationship of interest clear (Page 4: Line 60):

"There is now extensive observational evidence across countries and age groups supporting the contention that diet quality is a possible risk or protective factor for depression"

Methods

2.7 The participants were selected specifically to have poor dietary patterns. This is understandable, given the research question, but is one of the many aspects that limit the generalizability of this study (I will mention other aspects later on). It would be fair to mention this in the abstract. In a group of participants with dietary issues, dietary counseling helps 1/3 patients, moreso than no-dietary counseling. Such a conclusion seems more consistent with the evidence.

Response: As per our response to the previous reviewer on this, we have included the following on Page 19; Lines 439 – 445:

"Finally, given that we recruited participants on the basis of existing 'poor' quality diet, this may limit the generalizability of our findings to the wider population of individuals with depression. However, evidence suggests that our study sample was not necessarily a special subgroup; the recent 2014-15 Australian Health Survey tells us that only 5.6% of Australian adults had an adequate intake of vegetables and fruits. In this study, only 15 out of 166 people screened were excluded on the basis of a pre-existing 'good' diet, suggesting that – concordant with the wider population - poor diet is the norm in those with depressive illness."

2.8 "If participants were on antidepressant therapy or undergoing psychotherapy, they were required to be on the same treatment for at least two weeks prior to randomization." Do groups differ regarding psychotherapy? I only found information regarding medication later on in the MS, but I may have missed the relevant section.

Response: The number of participants in each group on both pharmaco and psychotherapy is given in Table 1 (and also now in the abstract).

2.9 The authors used "We are trialling the effect of an educational and counselling program focusing on diet that may help improve the symptoms of depression" to recruit participants. But the control counseling condition was not about dieting at all, correct? So I wonder whether people unhappy with their diet were enrolled, but this issue was only addressed in one of the two groups, which would bias the results. Similarly, control participants received financial compensation and DI participants did not – what is the rationale for this?

Response: Participants in the dietary support group were given food hampers, not as compensation, but rather to expose them to new foods and the sorts of foods and servings sizes that the dietitian was advocating as part of the dietary counselling – this has now been made clear in the manuscript – Page 9: Lines 196-7:

"In order to provide examples of serving sizes and exposure to the recommended foods, participants were also provided with a food hamper, incorporating the main components of the diet, along with recipes and meal plans."

The social support group received vouchers for movie tickets and were also offered dietary educational material at the conclusion of the study. This has now been clarified:

Page 9: Lines 211-213:

"Participants in the social support control group were provided with movie tickets as compensation for their time and participation in the study, and offered participation in a group dietary counselling session at the conclusion of the trial."

Importantly, participants were informed of the 50% chance of being randomised to either condition at both the screening and baseline appointments before randomisation. This was to ensure that participants understood the possibility that they would not be receiving dietary counselling before they were officially enrolled, allowing them to make an informed decision about their participation. The issue of expectation bias has been discussed in the limitations section, but now expanded upon on Page 18: Lines 414-417:

"Firstly, there is the issue of expectation bias due to the fact that we needed to be explicit in our advertising regarding the nature of the intervention and to the inability to blind the participants to their intervention group; this may have biased the results and also resulted in differential dropout rates."

2.10 I commend the authors for conducting a power analysis and writing so clearly that they did not achieve their aim. It would be even better to add to the relevant section "sample size" something like "However, due to xxx the final analytic sample included xx and xx, and only xx and xx full observations without missing data" or something similar for full transparency.

Response: We respectively disagree that this information should be presented in the sample size section as this is the description of the proposed methods, not what was actually done. We have addressed the difficulties we had with recruitment in the discussion (as above), while the section on Completers details how many participants had complete data. We have also added this information to the Abstract, so that it is immediately apparent (as above).

2.11 "All statistical analyses were conducted by an external statistician (SC), who was blind to group allocation prior to analysis." Well done!

Response: We thank the reviewer for these encouraging words.

2.12 Disclaimer: I am no expert for repeated measures mixed effects models and hope other reviewers and the editor can take a closer look at these. Especially the missing at random assumption seems highly implausible to me, because depression severity is nearly always a strong predictor for dropout, meaning follow-up data is not missing at random.

Response: The primary planned analysis method was MMRM, which assumes that the data are missing at random (MAR). We acknowledge that it may not be plausible that the data are MAR and this is why the sensitivity analyses were conducted under various assumptions. We have also discussed this clearly in the limitations section and added the following (Page 19: Lines 430-432:

"A larger sample size and assessments at more than two time-points would have afforded more sophisticated statistical modelling; this should be a key focus of future replication studies."

2.13 The authors note that "supplementary analyses" were carried out controlling for covariates, and they often present both uncontrolled and controlled analyses in the results. But not controlling for covariates opens the possibility for confounds — what is the rationale for not only reporting the properly controlled analyses in the manuscript?

Response: Because the results of the supplementary analyses did not differ from the primary main analyses, and given the issues with sample size and the limited statistical power, the primary analyses are reported.

Results

2.14 A main question for me is whether the dietary counseling sessions actually led to better diets. Could the authors answer this question somewhat more prominently in the results section? After all, they propose this as the main mechanism of symptom reduction (counseling -> change of diet -> change of depression), if I understand correctly. If power permits, the authors could even consider modeling it in a mediation model. This seems to be a quite relevant question if I understand the manuscript correctly.

Response: As per our responses to previous reviewer comments and questions, we have now added substantially to the text and the figures to illustrate the dietary change arising as a result of the intervention and its association with improvements in MADRS scores.

2.15 Do all analyses account for missing data (REML)? In any case, it is really relevant to be transparent here how many participants had full observations, in addition to "intervention, n=33; social support control, n=34".

Response: All analyses account for missing data (REML). As mentioned previously, it is stated the Results and in the Completers section how many participants in both groups had complete data.

2.16 "The dietary group had significantly lower scores on the dietary screening tool (p=0.043) and the ModiMedDiet score (p=0.031) than the social support control group at baseline." However, later on the authors state that the supplementary analysis only controlled for ModiMedDiet not dietary screening tool. If groups differ, it may be advisable to add both as covariates.

Response: The dietary screening tool is a brief measure used only to confirm eligibility. It does not give the detailed information required to fully assess diet quality; for this the ModiMedDiet score (derived from gold-standard 7-day food diaries) was used and this is now detailed in the manuscript (see previous responses).

2.17 "Those who did not complete the intervention were significantly more likely to have postsecondary education (81.8%, n=9) than those who completed (45.5%, n=25), $\chi2(1)=4.85$, p=0.028; *this relationship was particularly evident for the social support control group*, $\chi2(1)=6.92$, p=0.009." The section marked with italics implies that the second p-value is significantly smaller than the first, which would require a test.

Response: It has been made clearer on Page 14; Lines 316-319, that this effect was found for social support and not the dietary intervention group:

"Those who did not complete the intervention were significantly more likely to have post-secondary education (81.8%, n=9) than those who completed (45.5%, n=25), $\chi2(1)=4.85$, p=0.028; this relationship was observed for the social support control group, $\chi2(1)=6.92$, p=0.009 and not in the dietary support group, $\chi2(1)= 0.01$, p=0.965."

2.18 "The MMRM was rerun, adjusting for variables such as sex, education, physical activity, baseline BMI, and baseline ModiMedDiet score;" This is a good example for points (3.7) & (4.3) above.

Response: We believe that these issues have now been addressed

2.19 Browsing the results, I see 8 p-values between 5% and 1% – I very much doubt any of these will survive multiple testing. If I understand the authors correctly, they conducted a fairly large number of tests, which calls for correcting p-values. Bonferroni may be too strict due to the correlated nature of many variables, but the authors need to apply some form of correction to counter the possibility that some of the observed effects are due to false positive findings.

Response: Give that this is the first study to directly test the hypothesis that dietary improvement might lead to improvements in mental health, and should thus be considered preliminary, we always cited confidence intervals, but made no adjustments for multiple comparisons because they can result in a higher type II rate (falsely accepting the null hypothesis), reduced power, and increased likelihood of missing important findings (Rothman, K. J. (1990). "No adjustments are needed for multiple comparisons." Epidemiology 1: 43-46.).

We have now emphasised the preliminary nature of our study and softened our conclusions to reflect this.

Discussion

2.20 "However, our original power calculations were based on a very small effect size; arguably this would not have been clinically significant […] The fact that a large effect size was observed with a far smaller sample size than planned supports the efficacy and clinical utility of our approach."

The contemporary literature on power and effect size would strongly disagree with this statement. A recent example is the hypothermia study that found an incredibly large effect size – which makes the effect exactly that: incredible. In the present manuscript, the authors find an effect size 5 times as large as antidepressants that have been used and developed for over half a century – in an underpowered study with plenty of limitations. I don't think this "supports the efficacy and clinical utility of our approach". This is also the case due to the considerably large CI of the effect size "(95% CI -1.73, -0.59)", implying a lack of power to determine the effect size clearly.

Response: We appreciate the reviewer's viewpoint and have added the following to what is already a fairly lengthy and comprehensive discussion of the limitations of our study and sample size, as well as describing our results as "preliminary" (as above)

Page 18: lines 421-425:

"Clearly, our results must also be considered in light of the small sample size. Failure to reach our planned sample size increases the possibility that ours was a chance finding and/or that our sample was not representative, and limited our ability to conduct subgroup analyses. It may also have inflated the effect size we observed."

2.21 Could the authors compare the response and remission rates to prior studies? If I understand correctly, about 1/3 patients in the main intervention group improved, which seems somewhat similar to results from the placebo group of clinical trials in depression if I am not mistaken.

Response: Indeed, rates of remission of one third are in line with remission rates seen with agents of known efficacy in depression, such as antidepressant medications. As such, our result is concordant with what one would expect with known efficacious agents, supporting the utility of our strategy for the therapy of depression

Reviewer #3: This study is a randomized controlled study and reports that when subjects with depression are given an improved diet, they experienced an improvement in their symptoms. However, the study has several weaknesses. In particular, the sample size was rather small while the duration was fairly short (only 31 persons in the diet group completed the study; the duration was 12 weeks). The study is therefore preliminary in nature. However, (based on the literature review reported in the paper) this is the first randomized controlled study to investigate the impact of improved diet on mental disorders. For that reason the findings are likely to be of much interest and may spark more research in the area.

Several corrections need to be made to the paper.

3.1. The Abstract should state the number of subjects in each group that completed the study. This is essential so that readers can judge the reliability of the findings based only on the abstract.

Response: This has now been done:

"There were 31 in the diet support group and 25 in the social support control group who had complete data at 12 weeks."

3.2 P 6, line 8, I find it strange that of all the many dietary recommendations available, they would base their diet in part on Greek dietary recommendations published in 1999 (ref 18).

Response: The ModiMedDiet used in the trial was based, in part, on the Mediterranean diet, which has a very extensive evidence base for its healthful properties and demonstrated associations with depression. The introduction has been augmented to provide further information regarding the Mediterranean diet, including references to key studies and a meta-analysis. Page 4: lines 71-74:

"Similarly, another meta-analysis reported that higher adherence to a Mediterranean diet was associated with a 30% reduced risk for depression, with no evidence for publication bias [10]. The Mediterranean diet is recognized as a healthful dietary pattern and has been extensively associated with chronic disease risk reduction [11]."

And on Page 5: Lines 96-104

"On the other hand, post hoc analysis of a large-scale intervention trial provides preliminary support for dietary improvement as a strategy for the primary prevention of depression. Individuals at increased risk for cardiovascular events were randomized to a Mediterranean diet supplemented with either extra-virgin olive oil or mixed nuts, or a low-fat control diet [11]. While not statistically powered to assess the effectiveness of the intervention for preventing depression, there was evidence (albeit non-significant) of a reduced risk for incident depression for those randomized to a Mediterranean diet with nuts. This protective effect was statistically significant in those with type 2 diabetes, who comprised approximately half the sample [23]."

3.3 Same page, line 24, does an alcohol intake of "2% of E" mean a maximum intake of 2%. What if subjects do not want to drink alcohol?

Response: Yes, ideally a maximum alcohol intake of 2% of energy intake was recommended. Individuals who did not consume alcohol were not encouraged to commence drinking. We have added to the manuscript to clarify this point (Page 8: Lines 171-174):

"Red or white wine consumption beyond 2 standard drinks per day, and all other alcohol (e.g. spirits, beer), were included within the 'extras' food group. Individuals were advised to select red wine preferably and only drink with meals."

3.4 Same sentence. What is meant by "fibre/other 3% of E"?

Response: Energy intake is comprised primarily of the macronutrients; proteins, fats, and carbohydrates / sugars. Fibre and other nutrients in our food supply contribute the remainder.

3.5 P 8, line 21, Here the authors state that the "key food groups" are vegetables, fruits, olive oil, and legumes. This makes no sense.

Response: Concordant with the traditional Mediterranean diet, the intervention diet (ModiMed) is rich in vegetables, fruit and whole grains, with an emphasis on increased consumption of oily fish, extra virgin olive oil, legumes and raw unsalted nuts. Hence, these items are considered "key food groups" that make up the dietary pattern. However, we have now rewritten that section and this sentence is removed.

3.6 P 10, line 53, do not start a sentence with a number unless written as words.

Response: this has been amended

3.7 Same line, eligible is misspelled.

Response: we could not find this mistake

3.8 P 12, line 2, NNT should be rounded to 4.1. Make the same change in the abstract.

Response: this has been amended

3.9 Table 2, physical activity, remove the final decimal place.

Response: Respectfully, this will render these values discordant with the rest of the figures in the table. We will ask the editor to make a decision regarding this point.

Reviewer #4: This is a manuscript describing the results of a randomized trial to evaluate a dietary intervention for depressions symptoms.

There are several aspects of the study design that appear to allow for possible threats to the internal validity of the trial.

4.1 First, the method of allocation concealment is inadequately described. It's indicated on page 8, line 54, that "the randomization allocation was managed by the trial dietitians or 'befrienders' in order to maintain blinding." However it's unclear what this means, the method of allocation concealment should be clearly described; that is, what methods were implemented to prevent foreknowledge or guessing of upcoming treatment assignments by either prospective participants or others involved in interacting with those prospective participants prior to randomization? If no such methods were used, the manuscript should clearly say so.

Response: We have addressed the randomisation and blinding aspects of the trial extensively in the two sections dedicated to these issues. We have now added extra information to the section on Blinding (Pages 11 and 12: Lines 271-274):

"Participants were clearly instructed only to contact the dietician/befriender personally, to avoid contact with the research assistant, and voice messages were checked daily by the dietician/befriender to avoid unintended contact or information on participants' allocation."

4.2 Furthermore, it's indicated that a 2x2 randomized block design was used, but in an unblinded (or partially blinded) trial a block size of 4 is insufficient to support allocation concealment - this should be noted as a limitation of the trial.

Response: We have done so. Page 18; Line 417-19:

"Moreover, in regards to our randomisation process, a block size of four, whilst recommended for small sample sizes to avoid imbalances in allocation, may have been insufficient to support allocation concealment."

4.3 Finally, the clear and substantial differences in follow-up causes major validity concerns, and neither the complete-case nor the BOCF analyses address these concerns; since MADRS tended to decrease, on average, in both groups, these analyses would tend to bias the results away from the null. A more convincing sensitivity analysis would be a best/worst case analysis in which all intervention participants who dropped out would be assumed to have no change (i.e., worst case) but the control dropouts would be assumed to have substantial positive change.

Response: Given the substantial power limitations imposed by the small sample size and simple study design, we believe these sorts of extra sensitivity analyses to be unviable given the statistical power issue. We have discussed the limitations of the simple study design and limited sample size extensively, and noted that our findings should be regarded as 'preliminary' (first line of Discussion), but have also emphasised this further in the manuscript:

Page 18: Line 421-425:

"Clearly, our results should also be considered in light of the small sample size. Failure to reach our planned sample size increases the possibility that ours was a chance finding and/or that our sample was not representative, and limited our ability to conduct subgroup analyses. It may also have inflated the effect size we observed."

Page 19: Line 430-432

"A larger sample size and assessments at more than two time-points would have afforded more sophisticated statistical modelling; this should be a key focus of future replication studies."

Page 20: Lines 461-62

"Whilst emphasising the preliminary nature of this study and the imperative for replication in studies with larger sample sizes….

And Page 20: Lines 464-66:

"Future work in this new field of nutritional psychiatry research should focus on replication, ensuring larger samples and more sophisticated study designs, in order to confirm effects and afford sensitivity analyses to identify predictors of treatment response.

4.4 A confidence interval should really be included for the estimated NNT, particularly since this value is based on a very limited number of events (10 and 2 in the intervention and control groups, respectively). The point estimate carries no useful information by itself.

Response: The confidence intervals have now been reported in the Abstract (Page 2: Line 46) and in the manuscript on Page 14: Line 333

4.5 Page 9 lines 50-52 and Table 1 - it is generally inappropriate to conduct statistical tests for balance of baseline characteristics in a randomized trial (see, e.g., Senn, Stat. in Med 1994, pp 1715-1726, or the CONSORT Explanation and Elaboration document item 15). Contrary to the statement in the manuscript on page 9, line 47, these comparisons are not in accordance with ICH E9. These p-values should be removed.

Response: The p values have been taken out of Table 1. Reference to the analysis of baseline differences has been taken out in the Data Analysis and Results sections.

4.6 It's correctly noted (line 14, page 10) that mixed models allow for inclusion of all available participant data. However, it's unclear from Table 2 and Figure 1 that this was actually done. Were available data for participants who dropped out included in the model or not?

Response: All available data is included in Table 2. This has been clarified in the footnote of Table 2.

4.7 It's unclear what the statement that "non-parametric statistics were used when assumptions for parametric methods were violated" (line 26, page 10) refers to. When were the parametric models used and when were non-parametric method used, and what criteria were used to decide?

Response: This sentence has now been removed.

4.8 In Table 1, again, please exclude the p-values.

Response: We have done so.

4.9 In Table 2, the within-group, within-time standard errors should be removed. Standard errors are inferential statistics and within-group inferential statistics are not supported by randomization.

Response: We believe that the standard errors should remain as they provide an indication of the extent of within-group and within-time variability.

4.10 Figure 2: a bar chart is a poor means for summarizing the results, plus the results are redundant with Table 2, suggest removing. Furthermore, it's unclear what the error bars in this figure represent but, as noted above, it's generally unnecessary to present within-group inferential statistics in a randomized trial.

Response: We believed that the figure gave a useful visual to non-technical readers to illustrate the differences between the groups on their depression scores; however, we are happy to remove it at the editor's discretion.