nature portfolio

Peer Review File

# Dietary intervention and BMI reduction interaction in individuals at the extremes of genetic predisposition to higher BMI: a randomized controlled trial

Corresponding Author: Dr Andrea Ganna

**This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.**

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
This single-site randomized trial (GENEROOS) tests whether genome-wide BMI polygenic score (PS) modifies weight-loss response to a 6-month, calorie-restricted dietary coaching program versus control among adults with BMI 23–36 kg/m² and no diabetes, recruited from the extremes (top/bottom 5%) of the PS distribution. The trial is thoughtfully executed, addresses a timely question at the intersection of genetics and behavioral intervention, and demonstrates strong biobank recontact feasibility (≈23% enrollment from invitations). The primary finding—no detectable PS×intervention interaction—is important for calibrating expectations about PS-guided personalization in lifestyle-based weight loss.
Major Comments
1. Overstated conclusions
The conclusions are too strong given the study's statistical power constraints. Although the authors performed power calculations and provided justification for sample size/design, the trial was powered to detect only large interaction effects (80% power ≈ ≥15% difference in intervention effect by PS group). This threshold may be too high for more nuanced gene–environment interactions. The manuscript should consistently state that the study rules out large interactions but cannot exclude moderate/smaller effects. Please temper the Abstract, Results, and Discussion accordingly.
2. Baseline characteristics across the four cells
Table 1 contrasts intervention vs control but not the four PS×arm cells (high-PS control, high-PS diet, low-PS control, low-PS diet). Given modest cell sizes, imbalances are plausible. Please provide a PS×arm baseline table and consider sensitivity analyses adjusting for any imbalanced covariates in the mixed models.
3. Rationale for BMI lower bound = 23 (vs 25) and inclusion of non-obese participants
Please justify the BMI 23 cutoff rather than the conventional 25 kg/m² threshold. Was this chosen to ensure adequate representation in the bottom 5% PS group? If so, say so explicitly and (if feasible) provide a sensitivity analysis limited to BMI ≥25 to assess robustness.
4. Intervention details & adherence metrics
The Methods describe a structured nutritional program targeting a ~500 kcal/day deficit, standardized meal plans, app-based tracking, and coach messaging. Clarify adherence: % of days with meal logging, coaching contacts per participant, attrition in app use, and distribution of achieved daily energy deficits. These will help interpret the dose of behavioral exposure. Also, please state explicitly that the intervention was purely nutritional coaching/meal planning (no formal physical activity program), aside from using activity data to estimate caloric needs. This is worth highlighting in the Discussion when interpreting generalizability to multi-component lifestyle interventions.
5. Interaction between weight change and baseline weight
Because baseline weight differs by PS and may influence achievable loss, please test Intervention × Baseline weight (continuous) and report whether individuals with higher baseline weight lost more. This helps disentangle PS vs starting weight effects (baseline weight is partially driven by PS).
6. Continuous PS and interaction modeling
Please examine continuous PS (and PS percentile) rather than top/bottom 5% cutoffs, both as a main effect and in interaction with intervention, to increase power and capture subtler gradients.
Minor Comments
• Biobank recontact feasibility: The 23.3% enrollment from invitations is excellent and strengthens the paper's implementation relevance. Consider a sentence in the discussion further highlighting the importance of biobank based

callback and what is unique about FINNGEN that enables this success, as this could be learning for other biobanks to implement, given the need for many similar studies to tests PRSs prospectively.
• Terminology
o Avoid “dietary and lifestyle intervention” if there was no structured PA component; “dietary intervention” is clearer.
o Keep phrasing consistent across the manuscript and figures.
• Power interpretation
Reiterate that failing to detect an interaction does not imply absence of effect; it indicates any effect is likely < the detectable threshold under current N and variance.
• Figure/Table polish
o Fig 2 typo: “did not enrol” → “enroll”.
o Consider moving historical/context lines (current Intro ¶ at lines ~95–105) to the Discussion if the Intro feels long.
o In Table 2, clearly label the covariance structure used (you note CS in Methods—good) and present β with 95% CIs for PS main and interaction terms prominently in text/legend.
• Exploratory signals
In Fig 3, point estimates suggest greater % loss in the bottom 5% PS than top 5% PS within the intervention (though NS). Explicitly state this as hypothesis-generating and consistent with underpowered interaction testing.
Suggestions for the Discussion
• Emphasize feasibility/engagement as a major contribution of this genetic-first, biobank callback trial. Discuss how other biobanks can leverage the same model for more similar trials.
• Clarify that the intervention was nutritional (no formal exercise program), which may limit extrapolation to comprehensive lifestyle programs.
• Note that, while PS did not modify response at detectable levels here, baseline weight and other phenotypes may still shape outcomes, hence the value of reporting the Intervention × Baseline weight interaction.

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)
I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

(Remarks on code availability)
N/A

Reviewer #3

(Remarks to the Author)
I thank the authors and the editors for the opportunity to review this interesting paper. While the study seems to be well designed and clearly explained, I do have some concerns about the statistical analysis and presentation of some results.

1) The description of the statistical analysis lacks detail, and may be incorrect. While the paper says the full protocol as available in the supplemental materials, the supplemental file available for review appears to only have additional figures.
2) Line 552 says 'highest' BIC was chosen. Usually the lowest is picked
3) LIne 554, my understanding of na.exclude is that it drops missing data points, but this is not generally considered to be 'handling' missing data. Under the MAR assumption the mixed model does let unobserved values inform the estimates of pararemeters via the covariance but this probably needs a little more explanation in the paper.
4) Importantly, there is a mismatch between the intervention and the main outcome. Line 520 states the intervention was targeted to achieve a 500 kcal.day energy intake deficit. This would, in theory, lead to a constant amount of weight loss in kg over time regardless of baseline weight. However, the main analytical outcome is % weight loss, which is, of course, depends on baseline weight. This difference could attenuate the effect the investigators are looking for.
5) Also, the main interaction being tested assumes that there is a constant difference in weight loss between groups by intervention. Figure 3, and common sense, suggests that this should not be expected to the the case. In fact Figure 3 shows that the difference in weight loss by PS groups is gradually growing within the intervention (though perhaps not significantly, and, as a minor point, it's not clear whether the error bars in figure 3 are standard errors, 95% CIs or quartiles) but not within the control group.
Table 2 shows different modeling approaches used. I would suggest that 'model 1' is not very useful as it presents a single estimate for the intervention which would average across all time points though in truth the expectation is that trajectories would diverge. Model 2 is more reasonable, allowing for intervention*time interaction. But this highlights that the proper interaction test would have been to model a three-way interaction group*time*PS and test whether that model fit better than the model without the three-way interaction.
6) The power analysis in the methods reports what was projected for equal sized groups, but should note that the actual groups did not have equal size. Overall, it's not clear that this study was adequately powered to detect the target interaction effect.
7) The exploratory analyses report results to too many decimal places that are not meaningful, and it would be more informative to provide estimates of between group differences with CIs at timepoints of interest than to report the within-group changes (also reported as 'lower levels in line 154, but I think they mean 'change given that some are <0)

8) Lines 25, 197, 216 report the threshold for inclusion of BMI >=23, even though BMI=25 is the usual threshold of overweight. A more interesting subgroup analysis may be to look at those who were actually obese at baseline. It's not clear what to make of a weight loss study that include people who don't actually need to lose weight, or at least what including those people may do to change the magnitude of associations.

(Remarks on code availability)
I somewhat cursorily reviewed the code and found that it agreed with Table 2.

Reviewer #4

(Remarks to the Author)
Recommendation: Reject

The authors have investigated whether individuals with high or low genetic predisposition to higher BMI have differential weight loss during a randomized 6-month weight-loss intervention. Overall, the results showed no effect of the genetic risk score on weight loss, which is in line with previous studies. The manuscript and topic is of clear interest and the results of potential importance as they do not support an clinically meaningful role of genotype in weight loss management. However, we have some methodological concerns that need to be addressed in order to improve the paper, including some inconsistences between the methods and trial protocol being some of the major concerns.

Major comments
- Line 143-146: Please provide treatment effect estimates, i.e., between-group estimates with corresponding CI. It is not sufficient to provide within-group changes and report p-values. This is also the case for all non-specified exploratory analyses. Although one might include observed within-group changes in the results, these are not interpreted as causal effects of the intervention and thus treatment effects should be shown and interpreted from contrasting the two arms.
- Table 2 needs to be modified as it is currently unclear. First, the number -2.97 (which we assume is the contrast between the two arms for the % change in weight with no interaction included) doesn't correspond to within-group changes of -4.70 % for the intervention group and +0.14 kg for the control group after 6 months. Please explain why, because this is somewhat confusing from the statistical analysis. Also, you have provided two different estimates (% and kg) of within-group changes in lines 138-146 for the 6-month period, although kg change was not prespecified. The results sections should therefore be checked and revised for clarity.
- Please justify providing estimates in Table 2 for baseline weight, PS group top 5%, and timepoints. These are not your intervention effects so why would these numbers be interesting to report? Also, what is meant by "The reference is the body weight change at baseline" for the superscript d – do you mean simply body weight at baseline (i.e., you look at changes within the whole group from baseline)?
- There are a couple of deviations from the ClinicalTrials protocol that have not been discussed:
o Your primary outcome in ClinicalTrials is change in BMI at month 6 and in this paper it is change in weight (although under primary outcome you only describe how BMI was assessed). No rationale for this change has been provided and needs to be clarified to the readers.
o Recruitment of individuals with a BMI between 25-35 is specified in the protocol but in the paper you describe that you recruited individuals with a BMI of 23-36.
o Recruitment of n=238 individuals in ClinicalTrials protocol but n=223 in the paper.
o No mentioning of month 2 and 4 weights for the primary outcome (or any secondary outcomes) in ClinicalTrials but these are not called "non-specified exploratory outcomes" in the paper.
o You have specified in ClinicalTrials that you will adjust for sex and age in the models but this is not specified in the methods section of the manuscript.
- There is no mentioning of any adverse events despite this being a 6-month trial – were there none?
- The protocol and SAP could unfortunately not be found in Supplementary Material.
- What was the rationale behind recruiting normal weight individuals in a study investigating the impact of genetic predisposition to obesity and effects of weight loss? It might be more interesting including only overweight or obese individuals, as a normal weight individual with BMI 23 is likely to respond differently on energy restriction diet than a person with BMI 35, irrespectively of genetic risk score. Consequently, the included individuals with normal BMI might underestimate any possible genetic effect and thereby also potentially reduce overall statistical power. The reviewer acknowledge that the author briefly mention (line 214-217) that this study chose to focus on overweight subjects, but some further discussion of the implications and interpretation of their study should be added.

Minor to moderate comments
- Abstract, line 41-42. Please report the BMI In the low and high BMI PS groups, not only the weight difference in kg.
Row 484 have an incomplete sentence "All biomarkers were analyzed in". Please revise.
- How much missing data did you have for the exploratory outcomes?
- Based on your a priori power calculation you estimated you would need 62x4=248 individuals, but you only recruited 223 – please provide justification for this and discuss its potential implications.
- The meal plans in the dietary intervention needs to be described in more detail, by e.g., providing an example of a rotating weekly meal plan. It is also unclear from the methods section whether participants received meal plans and/or actual meals. On ClinicalTrials it says "Participants will get 5 meals/day…" but in the paper you state "Participants received structured meal plans of five daily meals…". Please clarify this. As you also provided many interventions within the broader coaching program, I think it is necessary to provide more details on what was actually done. Also, the authors need to acknowledge/describe the control group as well.
- Please specify whether baseline measures of the outcomes (primary, secondary and exploratory) were adjusted for in the

analyses, as this is generally advised to do when assessing change from baseline in outcomes.
- Line 195-196 states “though diet effectiveness was similar across sexes.” – where are these results shown and were they prespecified? How was this analyzed? I am missing this information in the paper.
- The drop-out rate was very low, which is a strength that should be mentioned in the discussion together with other strengths mentioned by the authors.

- Why have you chosen to describe the population stratified on sex in table 1?
- Line 199-200 states “improvements in cardiometabolic biomarkers suggest meaningful physiological changes beyond weight loss.” – Provide justification for this statement as weight loss is known to impact meaningful physiological changes to cardiometabolic biomarkers.
- Line 211-212 beginning with “Another strength…” can be removed. This is the fundamental pillar of randomized trials and you have already stated that one strength is the randomized design.
- Line 213: In results, drop-out was 5.8% and not <5%. Please be consistent throughout the manuscript with all the numbers provided. As of now, it is difficult for the reader to navigate this manuscript due to different numbers being provided in the different sections.
- “In conclusion, this 6-month randomized dietary intervention study found no evidence that BMI polygenic scores predict differential weight loss outcomes among overweight and mildly obese non-diabetic individuals.” – please remove as prediction was not your aim.
- There are little information about intervention adherence, but it would be interesting to know if there were any differences among the low and high PS BMI groups? If so, this needs discussion and added into the interpretation.

(Remarks on code availability)

Reviewer #5

(Remarks to the Author)
I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

(Remarks on code availability)

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
Thank you for the careful and thorough responses to the reviewers and for the substantial revisions made to the manuscript. The study addresses an important question at the intersection of polygenic risk and behavioral interventions and represents an innovative application of a genetic-first recontact strategy in a randomized trial setting. The manuscript is generally well written, the analyses are appropriate, and many of the reviewers’ concerns have been satisfactorily addressed in the revised version.

I particularly appreciate the additional analyses, clarification of the statistical modeling, improved presentation of baseline characteristics across PS × intervention groups, and the expanded discussion of limitations. The study provides an important and well-executed contribution to the literature on the clinical utility of polygenic scores in lifestyle-based weight loss interventions.

That said, I still have one remaining concern regarding how the study’s limitations are communicated in the Abstract.

1. Interpretation of the negative interaction result in the Abstract

Although the Discussion now appropriately acknowledges the limitations related to statistical power and the possibility that smaller interaction effects may exist, the Abstract conclusion still reads too definitively. Currently, the wording suggests that BMI polygenic scores do not substantially modify response to dietary intervention and highlights the limited utility of such scores for stratifying lifestyle interventions.

Given the study’s sample size and the detectable effect threshold described by the authors (approximately ≥15% difference in intervention effect by PS group), the trial is primarily able to rule out large interaction effects, but it cannot exclude more modest or subtle gene–environment interactions. As currently written, the Abstract may lead readers to interpret the findings as stronger evidence against any modifying effect than the data can support.

In addition, the exploratory adherence data suggest highly heterogeneous engagement with the intervention application, with meal logging showing a mean of ~64 days but a standard deviation exceeding the mean. This implies that a non-trivial fraction of participants may have had minimal behavioral exposure to the intervention, which further limits the strength of

conclusions about effect modification.

For these reasons, I recommend that the Abstract explicitly acknowledge these limitations, rather than placing them only in the Discussion or Conclusion sections.

A possible revision could be along the lines of:

“While these findings suggest that genetic susceptibility to higher BMI, as captured by current genome-wide polygenic scores, does not strongly modify short-term weight loss response to dietary intervention, modest interaction effects cannot be excluded given the study’s limited power and variability in engagement with the intervention. Larger studies with more consistent behavioral exposure will be required to clarify the role of polygenic scores in stratifying lifestyle-based weight loss strategies.”

The precise wording is of course up to the authors, but the key point is that the Abstract should clearly communicate that modest interaction effects cannot be excluded.

2. Interpretation of the intervention exposure

Related to the above point, the adherence metrics indicate substantial heterogeneity in engagement with the intervention application (e.g., mean meal logging ~64 days with SD ~67). While the exploratory analyses linking logging frequency to BMI change are helpful, the manuscript would benefit from a brief sentence acknowledging that the behavioral “dose” of the intervention was variable across participants, which may limit the ability to detect effect modification.

This does not require additional analyses but should be reflected in the interpretation of the findings.

Overall, this is a thoughtful and well-conducted study that provides valuable insight into the challenges of prospectively testing polygenic score–guided behavioral interventions. Addressing the points above—particularly clarifying the limitations in the Abstract—would strengthen the interpretation and ensure the conclusions remain appropriately calibrated to the data.

(Remarks on code availability)
not currently

Reviewer #2

(Remarks to the Author)
I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)
I thank the authors for their careful revisions and responses to prior review. This MS does present convincing evidence against the interaction tested.

I have only a minor criticism at this point.
The authors now report additional analyses related to meal and activity logging. The units and exact method are not reported which makes interpretation difficult. Given the the coefficients are small but with narrow confidence intervals, my guess is that the logging adherence was included as a number of adherent days. In this case, the size of the coefficient would not indicate the size of the effect, especially since the apparent distributions (especially for meal logging) are likely skewed. Without knowing the units the authors explanation that the associations are small is not evaluable. Given the modest sample size it is not especially plausible that an association of negligible size would be statistically significant.

If these are included please describe the methods a bit more, for example were both logging types modeled simultaneously or separately, what was the scale of the coefficient, and how did including logging change the expected weight change at a fixed level of logging, or how much of weight change do they explain, using and R^2 or something similar.

I did scan through the linked R code but did not see those analyses.

Best,
J. Richman

(Remarks on code availability)

Reviewer #4

(Remarks to the Author)
We thank the authors for their response to our comments, which to some extent clarifies several raised issues and concerns. However, for a few comments, we do not fully understand the response and would here like additional clarifications and further revisions in manuscript where needed. Please respond to our comments below to the specific responses where we ask for clarification.

1) Line 143-146: Please provide treatment effect estimates, i.e., between-group estimates with corresponding CI. It is not sufficient to provide within-group changes and report p-values. This is also the case for all non-specified exploratory analyses. Although one might include observed within-group changes in the results, these are not interpreted as causal effects of the intervention and thus treatment effects should be shown and interpreted from contrasting the two arms.

Response: We thank the reviewer for this comment. We now provide BMI change estimates for both between-group and within-group comparisons in the Results section. These changes are now reflected under the “Primary outcome” sub-section. Similarly, we have updated the “Non-specified exploratory analyses” in the Results section and added a Supplementary Table 2 to provide effect estimates with 95%CIs for all clinical lab measurements as suggested.

Results – Primary outcome: “In the diet intervention group, the adjusted mean change in BMI (kg/m2) at two months was –1.18 (95% CI, –1.38 to –0.97, n=109), at four months was –1.61 (95% CI, –1.82 to –1.40, n=108), and at six months (end of study) was –1.51 (95% CI, –1.72 to –1.31, n=106), whereas in the control group the mean change in BMI (kg/m2) at two months was –0.31 (95% CI, –0.52 to –0.11, n=107), at four months was –0.24 (95% CI, -0.45 to –0.04, n=106), and at six months (end of study) was -0.01 (95% CI, –0.22 to 0.19, n=104). Similarly, the adjusted mean change in BMI (kg/m2) in the top 5% BMI PS group at two months was -0.67 (95% CI, -0.84, -0.49, n=109), at four months was -0.85 (95% CI, -1.03, -0.67, n=108) and at six months was -0.69 (95% CI, -0.86, -0.51, n=106), whereas in the bottom 5% BMI PS group at two months was -0.82 (95% CI, -1.03, -0.61, n=109), at four months was -1.00 (95% CI, -1.21, -0.79, n=108) and at six months was -0.84 (95% CI, -1.05, -0.63, n=106).”

Comment: You have only presented the estimates for the within group changes but you have still not presented between group differences which are highly relevant to see. Thus, please present those estimates, for primary as well as the secondary outcomes. The authors should not draw conclusions from within group changes alone, which we believe is done.

Supplementary Table 2: Changes of non-exploratory clinical test values of individuals by intervention and BMI PS groups at the end of the study.

Comment: You have not presented the between group differences here, please do.

14) Line 195-196 states “though diet effectiveness was similar across sexes.” – where are these results shown and were they prespecified? How was this analyzed? I am missing this information in the paper.

Response: Very good point here, as this statement came from the initial models that we included age and sex, which were then removed from the paper. However, we now present our results adjusting for age and sex, and therefore we will keep this as is in the manuscript (Please see full model covariate estimates in Supplementary Table 1).

Comment: We do not understand; if you have adjusted for sex in the models, how can you state that there was an similar diet effect across sex? Where do you show these results, as the Full model in Suppl Table 1 do not provide data that the effect was similar. Please clarify.

17) Line 199-200 states “improvements in cardiometabolic biomarkers suggest meaningful physiological changes beyond weight loss.” – Provide justification for this statement as weight loss is known to impact meaningful physiological changes to cardiometabolic biomarkers.

Response: We thank the Reviewer for this comment. Given the major updates in our manuscript, this sentence is now modified as follows:

“Third, we did not collect detailed data on calorie intake, which limited our ability to examine the behavioral mechanisms underlying weight loss and BMI reduction. However, using data on adherence to meal and activity logging, we observed additional, but relatively small reductions in BMI beyond the effect of diet alone.”

Comment: This paragraph is unclear and needs rewriting. If you did not have data on calorie intake, how can you draw conclusions on separating specific effects on BMI beyond diet? The strongest predictor of BMI reduction in a decrease in total calorie intake. Can the authors please clarify the above reasoning and better explain what you have done.

(Remarks on code availability)

Reviewer #5

(Remarks to the Author)
I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

(Remarks on code availability)

Version 2:

Reviewer comments:

Reviewer #4

(Remarks to the Author)
The authors have responded overall to the comments raised, and have revised the manuscript accordingly. I have no further comments regarding the revised manuscript.

(Remarks on code availability)

Reviewer #5

(Remarks to the Author)
I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

(Remarks on code availability)

# REVIEWER COMMENTS

## Reviewer #1 (Remarks to the Author):

This single-site randomized trial (GENEROOS) tests whether genome-wide BMI polygenic score (PS) modifies weight-loss response to a 6-month, calorie-restricted dietary coaching program versus control among adults with BMI 23–36 kg/m² and no diabetes, recruited from the extremes (top/bottom 5%) of the PS distribution. The trial is thoughtfully executed, addresses a timely question at the intersection of genetics and behavioral intervention, and demonstrates strong biobank recontact feasibility (≈23% enrollment from invitations). The primary finding—no detectable PS×intervention interaction—is important for calibrating expectations about PS-guided personalization in lifestyle-based weight loss.

### Major Comments

1. Overstated conclusions
The conclusions are too strong given the study's statistical power constraints. Although the authors performed power calculations and provided justification for sample size/design, the trial was powered to detect only large interaction effects (80% power ≈ ≥15% difference in intervention effect by PS group). This threshold may be too high for more nuanced gene–environment interactions. The manuscript should consistently state that the study rules out large interactions but cannot exclude moderate/smaller effects. Please temper the Abstract, Results, and Discussion accordingly.

**Response:** We thank the Reviewer for the constructive feedback. We have now revised our abstract, discussion, and conclusion parts to more accurately reflect the changes we made and to better align with the statistical power limitations of our study, as shown below. To make it easier, we also kept the major changes as tracked in the manuscript.

**Abstract:** *"These findings suggest that, in this population, genetic susceptibility to higher BMI, as captured by current genome-wide polygenic scores, does not substantially modify short-term weight loss response to dietary intervention, underscoring the limited utility of these scores for stratifying lifestyle-based weight loss strategies by genetic risk."*

**Discussion:** *"However, we observed no statistically significant evidence that the BMI polygenic score modified the effect of diet intervention on BMI change. Within the limits of the study's power, these findings suggest that genetic susceptibility to higher BMI, as captured by current genome-wide polygenic scores, does not substantially alter response to dietary intervention."*

**Conclusion:** *"Modest interaction effects cannot be excluded given the sample size, and larger studies are needed to evaluate smaller effects. Our findings suggest that, in this population, current genome-wide BMI polygenic scores have limited utility for identifying individuals most likely to benefit from dietary interventions aimed at reducing body weight, underscoring the*

*importance of critically evaluating genetic tools before their integration into clinical decision-making."*

2. Baseline characteristics across the four cells
Table 1 contrasts intervention vs control but not the four PS×arm cells (high-PS control, high-PS diet, low-PS control, low-PS diet). Given modest cell sizes, imbalances are plausible. Please provide a PS×arm baseline table and consider sensitivity analyses adjusting for any imbalanced covariates in the mixed models.

**Response:** We thank the Reviewer for this suggestion to improve the descriptive Table 1 in our manuscript. We have now updated it as suggested and removed the sex-stratification to take into account suggestions from another Reviewer as well. In addition, we adjusted our models for age and sex, as can be seen from the Methods, as shown below:

**Methods – Statistical Analysis:** *"The model included a randomized intervention group (diet or control), BMI PGS group (top or bottom 5%), and their interaction, and was adjusted for age, sex, baseline BMI, and BMI measurement time."*

**Table 1:** Baseline descriptive characteristics of the GENEROOS study participants (n=223).

| Variable | Control (n=110) | | Diet intervention (n=113) | |
|---|---|---|---|---|
| | Top 5% BMI PGS (n=67) | Bottom 5% BMI PGS (n=43) | Top 5% BMI PGS (n=74) | Bottom 5% BMI PGS (n=39) |
| Age, years [mean (SD)] | 50.65 (11.21) | 53.65 (9.94) | 50.10 (10.88) | 51.43 (8.55) |
| Sex, female [n (%)] | 52 (77.6) | 30 (69.8) | 54 (73.0) | 31 (79.5) |
| Weight, kg [mean (SD)] | 89.74 (11.62) | 80.56 (11.89) | 88.58 (13.27) | 80.89 (12.11) |
| Height, cm [mean (SD)] | 168.06 (8.65) | 169.19 (8.69) | 168.56 (8.72) | 167.51 (8.54) |
| Body mass index (BMI), $kg/m^2$ [mean (SD)] | 31.72 (2.86) | 28.05 (2.79) | 31.06 (3.00) | 28.73 (2.78) |
| Total cholesterol, mean (SD), mmol/l | 4.91 (1.03) | 4.91 (0.75) | 4.69 (0.97) | 5.15 (1.0) |
| High-density lipoprotein cholesterol, mean (SD), mmol/l | 1.45 (0.40) | 1.49 (0.37) | 1.39 (0.40) | 1.61 (0.41) |
| Low-density lipoprotein cholesterol, mean (SD), mmol/l | 3.01 (0.90) | 2.94 (0.78) | 2.87 (0.85) | 3.15 (0.86) |
| Triglycerides, mean (SD), mmol/l | 1.60 (1.22) | 1.51 (0.95) | 1.48 (0.98) | 1.56 (0.79) |
| Fasting glucose, mean (SD), mmol/l | 5.64 (0.86) | 5.60 (0.74) | 5.69 (0.79) | 5.67 (0.61) |
| Hemoglobin A1c, mean (SD), mmol/l | 37.86 (5.17) | 36.95 (4.15) | 37.48 (3.84) | 37.05 (2.89) |

3. Rationale for BMI lower bound = 23 (vs 25) and inclusion of non-obese participants
Please justify the BMI 23 cutoff rather than the conventional 25 kg/m² threshold. Was this chosen to ensure adequate representation in the bottom 5% PS group? If so, say so explicitly and (if feasible) provide a sensitivity analysis limited to BMI ≥25 to assess robustness.

**Response:** This is a very good point, and we are happy to elaborate on the rationale. Our study leveraged the unique opportunity we had to invite people to enroll in a prospective study based

on their genotype. To make our process more focused on the right group, we used previously collected BMI data for these individuals during the initial invitation. However, since these data were not very recent, we allowed for some flexibility and invited individuals whose BMI was in the 23-36 kg/m$^2$ range. The idea was that once a person was interested in participating, they would provide their current BMI via the online questionnaire, which would subsequently be validated by a nurse when they visit in person. To keep our study pool engaged and the sample size as large as possible, we decided to keep all these individuals in the study until the end, even if their BMI was slightly under or over the pre-specified range of 25-35 kg/m$^2$. For transparency, the minimum and maximum BMI range in our data analysis is 23.78 - 37.45 kg/m$^2$.

We agree with all Reviewers who raised similar concerns that the overweight/obese range is 25-35 kg/m$^2$, and we provide two sensitivity analyses to assess the robustness of our findings. The first strictly includes individuals within the 25-35 kg/m$^2$ range (n=196) and the second focuses on obese individuals only with a BMI within the 30-35 kg/m$^2$ range (n=101). As you can appreciate from the updated Table 2 below, we observed no evidence that the BMI polygenic score group is an effect modifier of the intervention on BMI change in any of the additional analyses, indicating similar intervention effects across the BMI polygenic score groups. We added this to the *"Statistical analysis"* and *"Changes compared to the registered trial protocol"* subsections in Methods, as shown below:

**Methods – Statistical Analysis:** ***"****Sensitivity analyses were conducted by refitting the primary mixed-effects model in restricted samples defined by baseline BMI. Specifically, analyses were repeated among participants with baseline BMI 25–35 kg/m² to align with the trial protocol, and among obese participants only with baseline BMI 30–35 kg/m² to assess the robustness of the primary findings to the BMI range included."*

**Methods - Changes compared to the registered trial protocol:** *"To maintain our statistical power as high as possible, we kept all individuals who enrolled in the study with a BMI ranging from 23.78 - 37.45 kg/m2, and instead of removing those outside the 25-35 kg/m2 range before enrollment, we excluded them in the analysis by performing sensitivity analyses."*

**Table 2:** Interaction between BMI polygenic score group and diet intervention on BMI change.

| **Model** | **PS Group * Diet Interaction Estimate**[a] | **95% CI** | **P Value** |
|---|---|---|---|
| Model 1[b] | -0.02 | -0.45, 0.42 | 0.94 |
| Model 2[c] | -0.006 | -0.45, 0.44 | 0.98 |
| Model 3[d] | -0.08 | -0.88, 0.72 | 0.85 |

[a]Estimates represent adjusted differences in BMI change comparing polygenic score groups between randomized diet intervention and control derived from linear mixed-effects models including fixed effects for age at baseline, sex, BMI at baseline, time, intervention group, BMI polygenic score group, time and intervention group interaction, and a random intercept for each participant. To model within-subject correlation over time, the compound symmetry was selected. [b]Model 1 includes data from all individuals ($N_{individuals}$ = 223 and $N_{observations}$ = 863) with a BMI of 23.78 - 37.45 kg/m$^2$. [c]Model 2 includes data from overweight/obese individuals ($N_{individuals}$ = 196 and $N_{observations}$ = 760) with a BMI of 25 - 35 kg/m$^2$. [d]Model 3 includes data from obese individuals ($N_{individuals}$ = 101 and $N_{observations}$ = 390) with a BMI of 30 - 35 kg/m$^2$.

4. Intervention details & adherence metrics
The Methods describe a structured nutritional program targeting a ~500 kcal/day deficit, standardized meal plans, app-based tracking, and coach messaging. Clarify adherence: % of days with meal logging, coaching contacts per participant, attrition in app use, and distribution of achieved daily energy deficits. These will help interpret the dose of behavioral exposure.
Also, please state explicitly that the intervention was purely nutritional coaching/meal planning (no formal physical activity program), aside from using activity data to estimate caloric needs. This is worth highlighting in the Discussion when interpreting generalizability to multi-component lifestyle interventions.

**Response:** We thank the Reviewer for this suggestion. Via the application data retrieval, we had access to the number of days each person logged their meals and the number of days their activity data were synced. Participants logged meals on a mean of 64.2 (SD, 66.8) days during the six-month intervention, and their activity data were recorded on a mean of 125.0 (SD, 72.8) days. To further evaluate whether interaction with the application, despite being voluntary, was associated with a differential BMI reduction, we ran linear regression models with BMI change (kg/m$^2$) as the dependent variable. We found that people with higher adherence to meal logging ($\beta_{BMI\_change}$ = -0.011 kg/m$^2$, 95%CI [-0.017, -0.007]) and activity tracking ($\beta_{BMI\_change}$ = -0.006 kg/m$^2$, 95%CI [-0.011, -0.001]) had a higher BMI reduction, however, the effect size is very small in comparison to the mean BMI change due to the intervention ($\beta_{BMI\_change}$ = -1.5 kg/m$^2$ in the same group. These exploratory findings are now reported in the Results under the *"Non-prespecified exploratory analysis"* subsection. We also updated the *"Diet intervention"* subsection in the Methods to reflect this and added a short paragraph to the Discussion.

**Results - Non-prespecified exploratory analysis***: "Participants logged meals on a mean of 64.2 (SD, 66.8) days during the six-month intervention, and their activity data were recorded on a mean of 125.0 (SD, 72.8) days. Greater meal logging adherence was significantly associated with greater BMI reduction (βBMI_change = -0.011, 95%CI [-0.017, -0.007]). Similarly, greater activity data logging adherence was significantly associated with greater BMI reduction (βBMI_change = -0.006, 95%CI [-0.011, -0.001]). However, although statistically significant, the estimated effect sizes were small relative to the overall mean BMI reduction of −1.5 kg/m² observed during the intervention, indicating that logging adherence alone explains only a limited portion of the total weight loss."*

**Methods – Diet intervention:** *"Adherence to the nutritional intervention was quantified using app-derived engagement metrics. Meal logging adherence was defined as the number of days participants logged dietary intake in the mobile application during the six-month intervention period. Activity tracking adherence was defined as the number of days during which step counts were recorded via the wearable device. Both metrics were analyzed as raw counts and as proportions relative to the total study duration."*

***Discussion:*** *"Greater adherence to meal and activity logging in the diet intervention group was associated with larger reductions in BMI but the effect sizes were relatively small compared to the effect of diet."*

For the second part of the comment about the intervention being purely nutritional with no formal physical activity coaching, we have now updated the description of the diet intervention program in the *"Diet intervention"* subsection in the Methods, as well as included a relevant sentence in the Discussion, as shown below respectively:

**Methods – Diet intervention:** *"The diet intervention program was purely focused on nutritional coaching and meal planning, which addressed eating rhythm, portion control, balanced meals, caloric needs tied to activity levels, and preparation of nutritious, simple meals. There was no formal physical activity coaching, aside from using activity data to estimate each participant's caloric needs."*

**Discussion:** *"Finally, generalizability may be limited to populations with the resources, interest, and digital literacy to engage in a structured dietary intervention focused on nutritional coaching and meal planning without a physical activity component."*

5. Interaction between weight change and baseline weight
Because baseline weight differs by PS and may influence achievable loss, please test Intervention × Baseline weight (continuous) and report whether individuals with higher baseline weight lost more. This helps disentangle PS vs starting weight effects (baseline weight is partially driven by PS).

**Response:** This is a very good point, and we agree with the Reviewer that baseline weight/BMI differs partially because of the BMI PS among participants. This was, in fact, presented in our Supplementary Figure 2. However, when we tested for the interaction between continuous baseline BMI and the intervention group (Diet vs Control), we observed no evidence that baseline BMI modified the effect of intervention on BMI change ($\beta_{BMI\_change} = -0.006$, 95% CI [-0.07, 0.08]).

6. Continuous PS and interaction modeling
Please examine continuous PS (and PS percentile) rather than top/bottom 5% cutoffs, both as a main effect and in interaction with intervention, to increase power and capture subtler gradients.

**Response:** This is a great suggestion, and it was in our initial plans when designing the study. Unfortunately, due to ethical constraints, we could not get access to their actual scores, so the grouping had to be done into the two buckets we report, the Top and Bottom 5% BMI PS, prior to recontacting by the collaborating Biobank.

## Minor Comments

7. Biobank recontact feasibility: The 23.3% enrollment from invitations is excellent and strengthens the paper's implementation relevance. Consider a sentence in the discussion further highlighting the importance of biobank-based callback and what is unique about FINNGEN that enables this success, as this could be learning for other biobanks to implement, given the need for many similar studies to tests PRSs prospectively.

**Response:** We thank the Reviewer for this suggestion. We have now highlighted this in the Discussion part, as shown below:

*"Using a genetic-first recontact strategy from a cohort of over 38,000 genotyped individuals, we recruited participants at the extremes of BMI genetic risk with a 23.3% successful enrollment rate, maximizing power to detect differential effects. Although our hypothesis was not supported, we showcase how biobank-based studies can leverage existing genetic data to recall and prospectively enroll individuals into future precision medicine trials."*

8. Terminology
o Avoid "dietary and lifestyle intervention" if there was no structured PA component; "dietary intervention" is clearer.

o Keep phrasing consistent across the manuscript and figures.

**Response:** This is now updated across the revised manuscript as suggested.

9. Power interpretation
Reiterate that failing to detect an interaction does not imply absence of effect; it indicates any effect is likely < the detectable threshold under current N and variance.

**Response:** We thank the Reviewer for this suggestion. To address this, we added the following statements below in the Discussion and Conclusion parts:

**Discussion:** *"We observed no statistically significant evidence that the BMI polygenic score modified the effect of diet intervention on BMI change. Within the limits of the study's power, these findings suggest that genetic susceptibility to higher BMI, as captured by current genome-wide polygenic scores, does not substantially alter response to dietary intervention."*

**Conclusion:** *"Modest interaction effects cannot be excluded given the sample size, and larger studies are needed to evaluate smaller effects."*

10. Figure/Table polish

o Fig 2 typo: "did not enrol" → "enroll".

**Response:** Thank you for catching this typo. Fig 2 is now corrected.

o Consider moving historical/context lines (current Intro ¶ at lines ~95–105) to the Discussion if the Intro feels long.

**Response:** We thank the Reviewer for this suggestion. As we now made some major changes in our paper, we believe this section fits better in the Introduction as compared to the Discussion part.

o In Table 2, clearly label the covariance structure used (you note CS in Methods—good) and present β with 95% CIs for PS main and interaction terms prominently in text/legend.

**Response:** This comment is addressed under comment #3 above.

11. Exploratory signals
In Fig 3, point estimates suggest greater % loss in the bottom 5% PS than top 5% PS within the intervention (though NS). Explicitly state this as hypothesis-generating and consistent with underpowered interaction testing.

**Response:** We thank the Reviewer for this suggestion, which was also picked by another Reviewer. In our initial analysis using bodyweight change (%), we derived the unadjusted estimates to plot Fig 3, which we acknowledge generated a misleading picture of our findings. Taking into consideration all the relevant comments from Reviewers, we have now updated our models to align with the registered protocol and report our findings in BMI change ($kg/m^2$). Figure 3 is now updated, reflecting these changes and with the adjusted estimates (see below). We hope the Reviewers can appreciate that there is no visual suggestion for a greater % BMI change in the bottom 5% PS compared to the top 5% within the Diet intervention group, which is supported by the proper testing of interaction in our linear mixed-effects models. The updates can be found under the "*Primary Outcome*" subsection in Results, as shown below.

**Results – Primary outcome:** *"Similarly, the adjusted mean change in BMI (kg/m2) in the top 5% BMI PS group at two months was -0.67 (95% CI, -0.84, -0.49, n=109), at four months was -0.85 (95% CI, -1.03, -0.67, n=108) and at six months was -0.69 (95% CI, -0.86, -0.51, n=106), whereas in the bottom 5% BMI PS group at two months was -0.82 (95% CI, -1.03, -0.61, n=109), at four months was -1.00 (95% CI, -1.21, -0.79, n=108) and at six months was -0.84 (95% CI, -1.05, -0.63, n=106). The trajectory of BMI change (kg/m2) between the control and diet intervention by the BMI PS groups is shown in Figure 3.*

*We did not observe a statistically significant interaction between the BMI PS group and the diet intervention when using linear mixed models, including BMI estimates from both in-person visits (baseline and end-of-study) and self-reported weight measurements at months two and four. The interaction remained null in the additional models we ran as sensitivity analyses, which included only individuals with a baseline BMI of 25-35 kg/m2 and 30-35 kg/m2, compared to the 23.78 - 37.45 kg/m2 used in the primary analysis (Table 2). The full model results from the primary analysis are shown in Supplementary Table 1. The three-way interaction between BMI measurement time, BMI PS group, and randomized intervention group was not statistically significant (P = 0.42), and inclusion of this term did not change the estimated BMI PS-by-intervention interaction."*

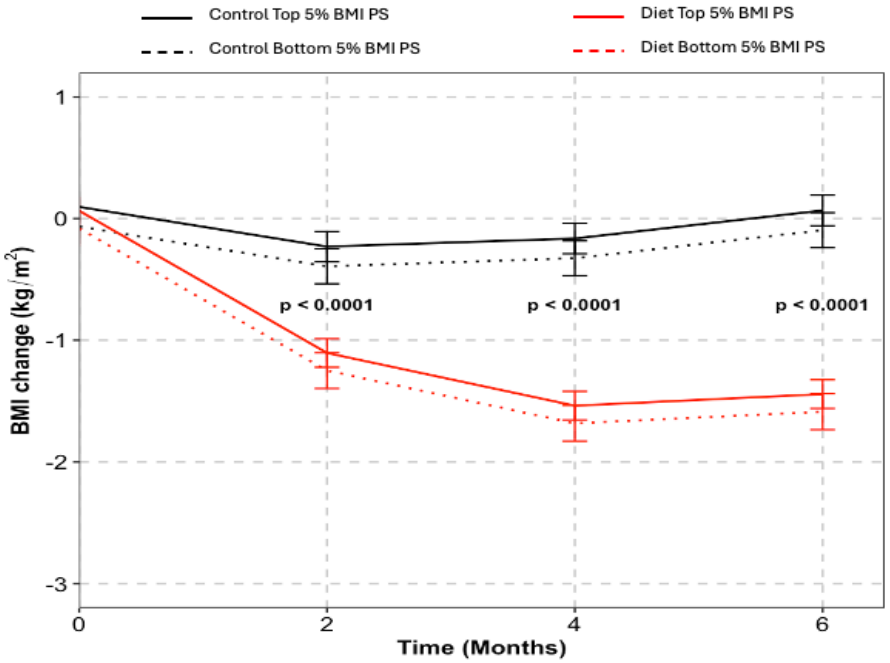



**Figure 3:** This figure shows the trajectory of BMI change (kg/m2) between the control and diet intervention by the BMI PS groups (N=223). The estimates are derived from the fully adjusted model with standard errors. Participants in the diet intervention exhibited greater BMI reduction compared with controls at two, four, and six months (P < 0.001). The P-values at each time point were calculated using the Mann-Whitney U (Wilcoxon rank-sum) non-parametric test. The effect of the diet did not differ by PGS group over time, as indicated by a non-significant treatment-by-time-by-PGS interaction (P = 0.42).

12. Suggestions for the Discussion

• Emphasize feasibility/engagement as a major contribution of this genetic-first, biobank callback trial. Discuss how other biobanks can leverage the same model for more similar trials.

**Response:** We thank the Reviewer for this suggestion. We have now further highlighted this under the "Strengths" paragraph in the Discussion, as shown below.

*“Using a genetic-first recontact strategy from a cohort of over 38,000 genotyped individuals, we recruited participants at the extremes of BMI genetic risk with a 23.3% successful enrollment rate, maximizing power to detect differential effects. Although our hypothesis was not supported, we showcase how biobank-based studies can leverage existing genetic data to recall and prospectively enroll individuals into future precision medicine trials.”*

• Clarify that the intervention was nutritional (no formal exercise program), which may limit extrapolation to comprehensive lifestyle programs.

**Response:** This was addressed above in comment 4.

• Note that, while PS did not modify response at detectable levels here, baseline weight and other phenotypes may still shape outcomes, hence the value of reporting the Intervention × Baseline weight interaction.

**Response:** This was addressed under comment #5 above.

## Reviewer #2 (Remarks to the Author):

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Reviewer #2 (Remarks on code availability):

N/A

**Response:** We thank the Reviewer for their detailed and constructive comments. We also acknowledge and appreciate the co-review conducted under the Nature Communications early career researcher initiative, which promotes training and recognition in peer review.

## Reviewer #3 (Remarks to the Author):

I thank the authors and the editors for the opportunity to review this interesting paper. While the study seems to be well designed and clearly explained, I do have some concerns about the statistical analysis and presentation of some results.

1) The description of the statistical analysis lacks detail and may be incorrect. While the paper says the full protocol as available in the supplemental materials, the supplemental file available for review appears to only have additional figures.

**Response:** We thank the Reviewer for bringing this to our attention. We have shared the approved protocol of our study during the submission process, but perhaps there was a missing link/document in the package. We now make sure that this is provided in the Supplementary Material. In addition, we made some changes in our revised manuscript's Statistical analysis section as well (lines 588-623).

2) Line 552 says 'highest' BIC was chosen. Usually the lowest is picked

**Response:** We thank the Reviewer for catching this. Indeed, although we picked the model with compound symmetry as the covariance structure based on the lowest BIC compared to the other models, we wrote the opposite here. This is now corrected in the "*Statistical analysis"* subsection in Methods as shown below:

*"The final covariance structure (i.e., CS) was selected based on convergence and the lowest Bayesian Information Criterion (BIC), indicating the optimal balance between goodness of fit and model parsimony."*

3) LIne 554, my understanding of na.exclude is that it drops missing data points, but this is not generally considered to be 'handling' missing data. Under the MAR assumption the mixed model does let unobserved values inform the estimates of pararemeters via the covariance but this probably needs a little more explanation in the paper.

**Response:** We thank the reviewer for this helpful comment. We agree that na.exclude itself does not constitute a missing data method. In our analysis, missing outcome data were addressed through likelihood-based estimation in linear mixed-effects models fitted by REML. Under the Missing At Random (MAR) assumption, this approach allows participants with incomplete follow-up to contribute all available observations to the estimation of fixed effects via the model likelihood and covariance structure, without the need for imputation. We have clarified this point in the "*Statistical analysis"* subsection in Methods and revised the text, accordingly, as shown below:

*“Missing outcome data were handled using likelihood-based estimation in linear mixed-effects models fitted by restricted maximum likelihood (REML). Under the assumption that data were missing at random (MAR), this approach allows all available observations to contribute to parameter estimation without requiring imputation, with within-subject correlation accounted for through the specified covariance structure. The na.exclude option was used to ensure appropriate alignment of fitted values and residuals while excluding missing observations from the likelihood.”*

4) Importantly, there is a mismatch between the intervention and the main outcome. Line 520 states the intervention was targeted to achieve a 500 kcal.day energy intake deficit. This would, in theory, lead to a constant amount of weight loss in kg over time regardless of baseline weight. However, the main analytical outcome is % weight loss, which is, of course, depends on baseline weight. This difference could attenuate the effect the investigators are looking for.

**Response:** Given the feedback we got from all Reviewers, it seems that presenting our findings in bodyweight change (kg and %) created confusion; therefore, we have now updated our results to match the primary outcome, which is BMI change ($kg/m^2$). With this change, it is now evident that participants in the diet intervention significantly reduced their BMI over the 6 months compared to those in the control group, irrespective of their BMI PS group. The changes are now reflected in the updated Results section and Fig 3 in the revised manuscript, as also shown below.

*“Similarly, the adjusted mean change in BMI (kg/m2) in the top 5% BMI PS group at two months was -0.67 (95% CI, -0.84, -0.49, n=109), at four months was -0.85 (95% CI, -1.03, -0.67, n=108) and at six months was -0.69 (95% CI, -0.86, -0.51, n=106), whereas in the bottom 5% BMI PS group at two months was -0.82 (95% CI, -1.03, -0.61, n=109), at four months was -1.00 (95% CI, -1.21, -0.79, n=108) and at six months was -0.84 (95% CI, -1.05, -0.63, n=106). The trajectory of BMI change (kg/m2) between the control and diet intervention by the BMI PS groups is shown in Figure 3.*

*We did not observe a statistically significant interaction between the BMI PS group and the diet intervention when using linear mixed models, including BMI estimates from both in-person visits (baseline and end-of-study) and self-reported weight measurements at months two and four. The interaction remained null in the additional models we ran as sensitivity analyses, which included only individuals with a baseline BMI of 25-35 kg/m2 and 30-35 kg/m2, compared to the 23.78 - 37.45 kg/m2 used in the primary analysis (Table 2). The full model results from the primary analysis are shown in Supplementary Table 1. The three-way interaction between BMI measurement time, BMI PS group, and randomized intervention group was not statistically significant (P = 0.42), and inclusion of this term did not change the estimated BMI PS-by-intervention interaction.”*

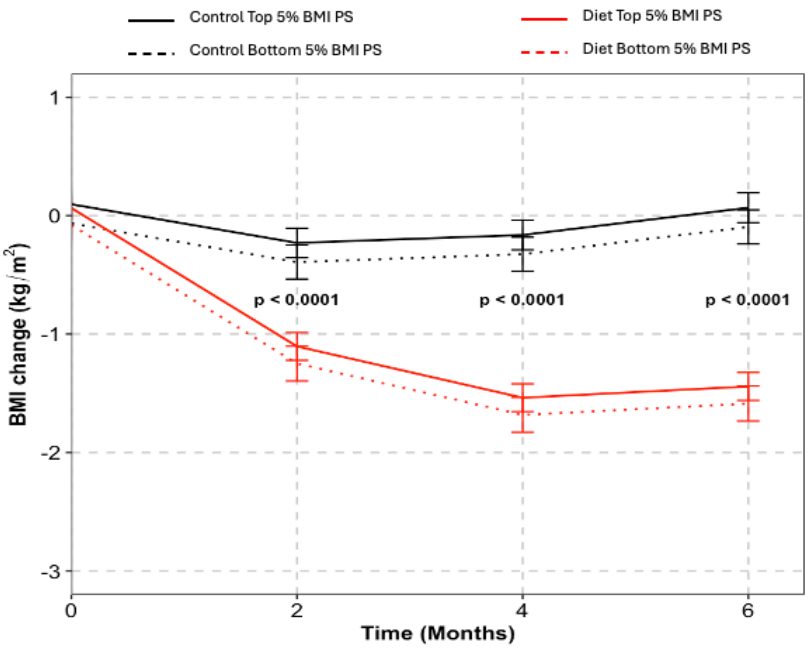



**Figure 3:** This figure shows the trajectory of BMI change (kg/m2) between the control and diet intervention by the BMI PS groups (N=223). The estimates are derived from the fully adjusted model with standard errors. Participants in the diet intervention exhibited greater BMI reduction compared with controls at two, four, and six months (P < 0.001). The P-values at each time point were calculated using the Mann-Whitney U (Wilcoxon rank-sum) non-parametric test. The effect of the diet did not differ by PGS group over time, as indicated by a non-significant treatment-by-time-by-PGS interaction (P = 0.42).

5) Also, the main interaction being tested assumes that there is a constant difference in weight loss between groups by intervention. Figure 3, and common sense, suggests that this should not be expected to the the case. In fact Figure 3 shows that the difference in weight loss by PS groups is gradually growing within the intervention (though perhaps not significantly, and, as a minor point, it's not clear whether the error bars in figure 3 are standard errors, 95% CIs or quartiles) but not within the control group.

**Response:** We thank the Reviewer for this suggestion, which was also picked by another Reviewer. In our initial analysis using bodyweight change (%), we derived the unadjusted estimates to plot Fig 3, which we acknowledge generated a misleading picture of our findings. Taking into consideration all the relevant comments from Reviewers, we have now updated our models to align with the registered protocol and report our findings in BMI change (kg/m2). Figure 3 is now updated, reflecting these changes and with the adjusted estimates **(please our response in the previous comment above).** We hope the Reviewer can now appreciate that there is no visual suggestion for a greater % BMI change in the bottom 5% PS compared to the top 5% within the Diet intervention group, which is supported by the proper testing of interaction in our linear mixed-effects models.

Table 2 shows different modeling approaches used. I would suggest that 'model 1' is not very useful as it presents a single estimate for the intervention which would average across all time points though in truth the expectation is that trajectories would diverge. Model 2 is more reasonable, allowing for intervention*time interaction. But this highlights that the proper interaction test would have been to model a three-way interaction group*time*PS and test whether that model fit better than the model without the three-way interaction.

**Response:** We thank the Reviewer for the very good point and suggestion about Table 2, which was mentioned by another Reviewer as well. Taking into consideration all the comments related to the main analysis, we now added the sensitivity analyses in Methods, as well as results from those in Table 2, as shown below. In addition, we provide the full model estimates in Supplemental Table 1.

**Methods – Statistical Analysis:** ***"****Sensitivity analyses were conducted by refitting the primary mixed-effects model in restricted samples defined by baseline BMI. Specifically, analyses were repeated among participants with baseline BMI 25–35 kg/m² to align with the trial protocol, and among obese participants only with baseline BMI 30–35 kg/m² to assess the robustness of the primary findings to the BMI range included."*

**Table 2:** Interaction between BMI polygenic score group and diet intervention on BMI change.

| **Model** | **PS Group * Diet Interaction Estimate**[a] | **95% CI** | **P Value** |
|---|---|---|---|
| Model 1[b] | -0.02 | -0.45, 0.42 | 0.94 |
| Model 2[c] | -0.006 | -0.45, 0.44 | 0.98 |
| Model 3[d] | -0.08 | -0.88, 0.72 | 0.85 |

[a]Estimates represent adjusted differences in BMI change comparing polygenic score groups between randomized diet intervention and control derived from linear mixed-effects models including fixed effects for age at baseline, sex, BMI at baseline, time, intervention group, BMI polygenic score group, time and intervention group interaction, and a random intercept for each participant. To model within-subject correlation over time, the compound symmetry was selected. [b]Model 1 includes data from all individuals ($N_{individuals}$ = 223 and $N_{observations}$ = 863) with a BMI of 23.78 - 37.45 kg/m$^2$. [c]Model 2 includes data from overweight/obese individuals ($N_{individuals}$ = 196 and $N_{observations}$ = 760) with a BMI of 25 - 35 kg/m$^2$. [d]Model 3 includes data from obese individuals ($N_{individuals}$ = 101 and $N_{observations}$ = 390) with a BMI of 30 - 35 kg/m$^2$.

Regarding the suggestion to test a three-way interaction, it is an interesting proposal that asks a slightly different question, which is whether the interaction between BMI PS and intervention changes over time, as opposed to our pre-specified primary outcome that asks the question whether the intervention effect on BMI change differs by BMI PS group. To explore this suggestion, we tested a three-way interaction between BMI measurement time, BMI PS group,

and intervention group to evaluate whether the BMI PS-by-intervention interaction varied over time. This interaction was not statistically significant, and inclusion of the three-way term did not materially change the estimated PS-by-intervention effects or the study conclusions. Because the primary hypothesis concerned an overall BMI PS-by-intervention interaction and time-varying genetic modification was not prespecified, we retained the more parsimonious model for the primary analysis, but we noted in the Results the three-way interaction as a sensitivity analysis. These changes are now reflected in the revised manuscript as below: under Methods/Statistical analysis and Results/Primary outcome sections, as shown below, respectively:

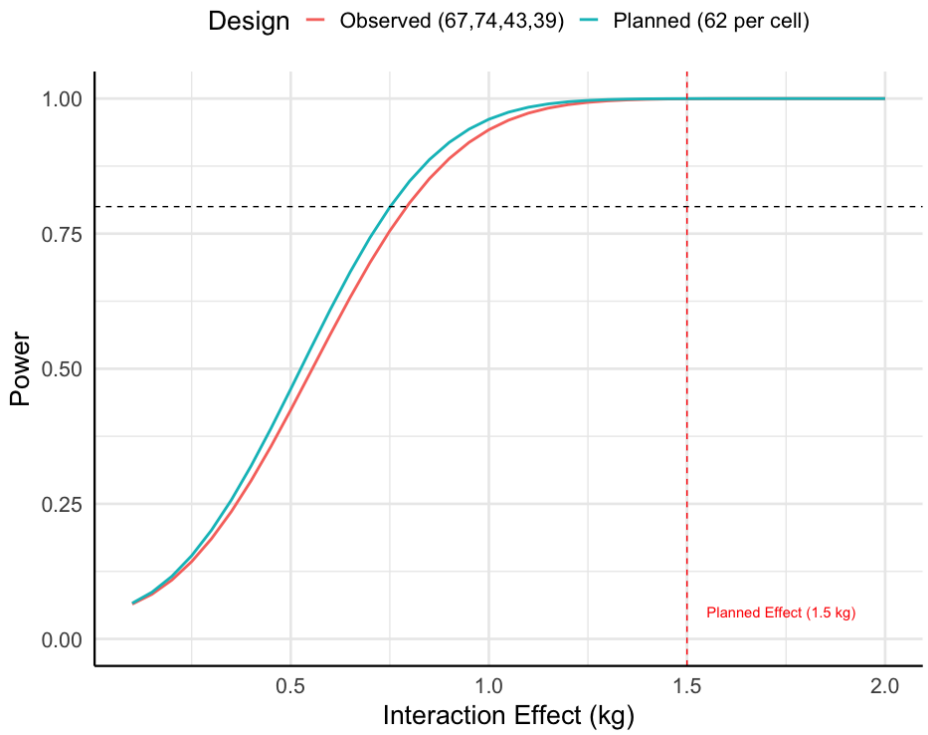
**Methods – Statistical analysis:** *"In a sensitivity analysis, we additionally tested a three-way non-prespecified exploratory interaction between BMI measurement time, BMI PS group, and randomized intervention group to assess whether the BMI PS-by-intervention interaction varied over time."*

**Results – Primary outcome:** *"The three-way interaction was not statistically significant, and inclusion of this term did not change the estimated BMI PS-by-intervention interaction."*

6) The power analysis in the methods reports what was projected for equal sized groups, but should note that the actual groups did not have equal size. Overall, it's not clear that this study was adequately powered to detect the target interaction effect.

**Response:** We agree with the Reviewer that group sizes were unequally balanced relative to the original design. To evaluate the impact on statistical sensitivity, we performed power calculations using the observed sample distribution and standard linear model assumptions ($\alpha = 0.05$). These analyses indicate ≥80% power to detect the pre-specified 15% difference in diet intervention effect by BMI PS group (our target interaction effect) despite the unequal group sizes. Although power is reduced for smaller effects compared with the planned balanced design, the study remains adequately powered to test the hypothesised effect magnitude **(see figure below)**. To enhance transparency, we have clarified this in the manuscript under the "*Changes compared to the registered trial protocol" section*.

**Methods - Changes compared to the registered trial protocol:** *"The updated power calculations reflect this revised scenario, and post hoc power calculations based on the achieved sample sizes indicate that power remained >80% for detecting the prespecified interaction effect."*




In addition, we have revised our abstract, discussion, and conclusion parts to more accurately align with the finding's interpretation based on the statistical power of our study. To make it easier, we kept the major changes as tracked in the manuscript.

**Abstract:** *"These findings suggest that, in this population, genetic susceptibility to higher BMI, as captured by current genome-wide polygenic scores, does not substantially modify short-term weight loss response to dietary intervention, underscoring the limited utility of these scores for stratifying lifestyle-based weight loss strategies by genetic risk."*

**Discussion:** *"However, we observed no statistically significant evidence that the BMI polygenic score modified the effect of diet intervention on BMI change. Within the limits of the study's power, these findings suggest that genetic susceptibility to higher BMI, as captured by current genome-wide polygenic scores, does not substantially alter response to dietary intervention."*

**Conclusion:** *"Modest interaction effects cannot be excluded given the sample size, and larger studies are needed to evaluate smaller effects. Our findings suggest that, in this population, current genome-wide BMI polygenic scores have limited utility for identifying individuals most likely to benefit from dietary interventions aimed at reducing body weight, underscoring the importance of critically evaluating genetic tools before their integration into clinical decision-making."*

7) The exploratory analyses report results to too many decimal places that are not meaningful, and it would be more informative to provide estimates of between-group differences with CIs at timepoints of interest than to report the within-group changes (also reported as 'lower levels in line 154, but I think they mean 'change given that some are <0)

**Response:** We thank the Reviewer for this suggestion. Firstly, we have now updated all our exploratory analyses results to two decimal places (lines 168-174). Secondly, the mean (SD) of all clinical markers at baseline, both by BMI PS and intervention groups, can now be seen in the updated Table 1 of the manuscript (Page 19). In addition, we now report the changes in all clinical markers (from baseline to end of the study) as suggested by BMI PS and the intervention group in Supplementary Table 2 as shown below.

**Supplementary Table 2:** Changes of non-exploratory clinical test values of individuals by intervention and BMI PS groups at the end of the study.

| Clinical test | Control – Bottom 5% BMI PS | Control – Top 5% BMI PS | Diet intervention – Bottom 5%. BMI PS | Diet intervention – Top 5% BMI PS |
|---|---|---|---|---|
| Total cholesterol, mean (95%CIs), mmol/l | -0.09 (-0.3, 0.13) | 0.05 (-0.12, 0.22) | -0.13 (-0.39, 0.13) | -0.25 (-0.38, -0.11) |
| High-density lipoprotein cholesterol, mean (95%CIs), mmol/l | -0.01 (-0.06, 0.05) | -0.01 (-0.06, 0.04) | -0.01 (-0.08, 0.06) | -0.02 (-0.06, 0.02) |
| Low-density lipoprotein cholesterol, mean (95%CIs), mmol/l | -0.13 (-0.34, 0.07) | 0 (-0.14, 0.14) | -0.19 (-0.41, 0.02) | -0.26 (-0.4, -0.12) |
| Triglycerides, mean (95%CIs), mmol/l | -0.01 (-0.28, 0.25) | -0.13 (-0.33, 0.07) | -0.25 (-0.53, 0.04) | -0.2 (-0.37, -0.04) |
| Fasting glucose, mean (95%CIs), mmol/l | 0.03 (-0.11, 0.17) | -0.02 (-0.19, 0.14) | -0.18 (-0.34, -0.01) | -0.12 (-0.28, 0.03) |
| Hemoglobin A1c, mean (95%CIs), mmol/l | 0 (-0.56, 0.56) | -0.1 (-1.13, 0.94) | -0.59 (-1.24, 0.05) | -0.64 (-1.1, -0.18) |

8) Lines 25, 197, 216 report the threshold for inclusion of BMI >=23, even though BMI=25 is the usual threshold of overweight. A more interesting subgroup analysis may be to look at those who were actually obese at baseline. It's not clear what to make of a weight loss study that include people who don't actually need to lose weight, or at least what including those people may do to change the magnitude of associations.

**Response:** We thank the Reviewer for this suggestion. We have done this as a sensitivity analysis, and we report the findings in Table 2 of our manuscript. Please refer to our response to comment 5 above.

## Reviewer #3 (Remarks on code availability):

I somewhat cursorily reviewed the code and found that it agreed with Table 2.

**Response:** We appreciate the Reviewer's time to review our code.

## Reviewer #4 (Remarks to the Author):

Recommendation: Reject

The authors have investigated whether individuals with high or low genetic predisposition to higher BMI have differential weight loss during a randomized 6-month weight-loss intervention. Overall, the results showed no effect of the genetic risk score on weight loss, which is in line with previous studies. The manuscript and topic is of clear interest and the results of potential importance as they do not support an clinically meaningful role of genotype in weight loss management. However, we have some methodological concerns that need to be addressed in order to improve the paper, including some inconsistences between the methods and trial protocol being some of the major concerns.

Major comments
1) Line 143-146: Please provide treatment effect estimates, i.e., between-group estimates with corresponding CI. It is not sufficient to provide within-group changes and report p-values. This is also the case for all non-specified exploratory analyses. Although one might include observed within-group changes in the results, these are not interpreted as causal effects of the intervention and thus treatment effects should be shown and interpreted from contrasting the two arms.

**Response:** We thank the reviewer for this comment. We now provide BMI change estimates for both between-group and within-group comparisons in the Results section. These changes are now reflected under the *"Primary outcome"* sub-section. Similarly, we have updated the *"Non-specified exploratory analyses"* in the Results section and added a Supplementary Table 2 to provide effect estimates with 95%CIs for all clinical lab measurements as suggested.

**Results – Primary outcome:** *"In the diet intervention group, the adjusted mean change in BMI (kg/m$^2$) at two months was −1.18 (95% CI, −1.38 to −0.97, n=109), at four months was −1.61 (95% CI, −1.82 to −1.40, n=108), and at six months (end of study) was −1.51 (95% CI, −1.72 to −1.31, n=106), whereas in the control group the mean change in BMI (kg/m$^2$) at two months was −0.31 (95% CI, −0.52 to −0.11, n=107), at four months was −0.24 (95% CI, -0.45 to −0.04, n=106), and at six months (end of study) was -0.01 (95% CI, −0.22 to 0.19, n=104). Similarly, the adjusted mean change in BMI (kg/m$^2$) in the top 5% BMI PS group at two months was -0.67 (95% CI, -0.84, -0.49, n=109), at four months was -0.85 (95% CI, -1.03, -0.67, n=108) and at six months was -0.69 (95% CI, -0.86, -0.51, n=106), whereas in the bottom 5% BMI PS group at two months was -0.82 (95% CI, -1.03, -0.61, n=109), at four months was -1.00 (95% CI, -1.21, -0.79, n=108) and at six months was -0.84 (95% CI, -1.05, -0.63, n=106)."*

**Results – Non-specified exploratory analyses: "***Changes in all abovementioned laboratory values by intervention and BMI PS groups are shown in Supplementary Table 2."*

**Supplementary Table 2:** Changes of non-exploratory clinical test values of individuals by intervention and BMI PS groups at the end of the study.

| Clinical test | Control – Bottom 5% BMI PS | Control – Top 5% BMI PS | Diet intervention – Bottom 5%. BMI PS | Diet intervention – Top 5% BMI PS |
|---|---|---|---|---|
| Total cholesterol, mean (95%CIs), mmol/l | -0.09 (-0.3, 0.13) | 0.05 (-0.12, 0.22) | -0.13 (-0.39, 0.13) | -0.25 (-0.38, -0.11) |
| High-density lipoprotein cholesterol, mean (95%CIs), mmol/l | -0.01 (-0.06, 0.05) | -0.01 (-0.06, 0.04) | -0.01 (-0.08, 0.06) | -0.02 (-0.06, 0.02) |
| Low-density lipoprotein cholesterol, mean (95%CIs), mmol/l | -0.13 (-0.34, 0.07) | 0 (-0.14, 0.14) | -0.19 (-0.41, 0.02) | -0.26 (-0.4, -0.12) |
| Triglycerides, mean (95%CIs), mmol/l | -0.01 (-0.28, 0.25) | -0.13 (-0.33, 0.07) | -0.25 (-0.53, 0.04) | -0.2 (-0.37, -0.04) |
| Fasting glucose, mean (95%CIs), mmol/l | 0.03 (-0.11, 0.17) | -0.02 (-0.19, 0.14) | -0.18 (-0.34, -0.01) | -0.12 (-0.28, 0.03) |
| Hemoglobin A1c, mean (95%CIs), mmol/l | 0 (-0.56, 0.56) | -0.1 (-1.13, 0.94) | -0.59 (-1.24, 0.05) | -0.64 (-1.1, -0.18) |

2) Table 2 needs to be modified as it is currently unclear. First, the number -2.97 (which we assume is the contrast between the two arms for the % change in weight with no interaction included) doesn't correspond to within-group changes of -4.70 % for the intervention group and +0.14 kg for the control group after 6 months. Please explain why, because this is somewhat confusing from the statistical analysis. Also, you have provided two different estimates (% and kg) of within-group changes in lines 138-146 for the 6-month period, although kg change was not prespecified. The results sections should therefore be checked and revised for clarity.

**Response:** We agree with the Reviewer that reporting estimates in different parts of the manuscript with different units created a lot of confusion. This has now been changed, and we have updated all relevant sections of our paper to report only the BMI change (kg/m2), which aligns with our primary outcome. For example, in Table 2, we now report the interaction estimates (primary outcome) from our models, and the full model covariate estimates are provided in the Supplementary material (Supplementary Table 1). In addition, the Results section and Figure 3 are now updated to reflect those changes.

3) Please justify providing estimates in Table 2 for baseline weight, PS group top 5%, and timepoints. These are not your intervention effects, so why would these numbers be interesting to report? Also, what is meant by "The reference is the body weight change at baseline" for the superscript d – do you mean simply body weight at baseline (i.e., you look at changes within the whole group from baseline)?

**Response:** We thank the Reviewer for this suggestion, which was also brought up by the other Reviewers. Our Table 2 is now updated to only report our pre-specified primary outcome, that is, the BMI PS x Intervention interaction (see below). All other covariate estimates are reported in the Supplementary Table 1.

**Table 2:** Interaction between BMI polygenic score group and diet intervention on BMI change.

| Model | PS Group * Diet Interaction Estimate[a] | 95% CI | P Value |
|---|---|---|---|
| Model 1[b] | -0.02 | -0.45, 0.42 | 0.94 |
| Model 2[c] | -0.006 | -0.45, 0.44 | 0.98 |
| Model 3[d] | -0.08 | -0.88, 0.72 | 0.85 |

[a]Estimates represent adjusted differences in BMI change comparing polygenic score groups between randomized diet intervention and control derived from linear mixed-effects models including fixed effects for age at baseline, sex, BMI at baseline, time, intervention group, BMI polygenic score group, time and intervention group interaction, and a random intercept for each participant. To model within-subject correlation over time, the compound symmetry was selected. [b]Model 1 includes data from all individuals ($N_{individuals}$ = 223 and $N_{observations}$ = 863) with a BMI of 23.78 - 37.45 kg/m$^2$. [c]Model 2 includes data from overweight/obese individuals ($N_{individuals}$ = 196 and $N_{observations}$ = 760) with a BMI of 25 - 35 kg/m$^2$. [d]Model 3 includes data from obese individuals ($N_{individuals}$ = 101 and $N_{observations}$ = 390) with a BMI of 30 - 35 kg/m$^2$.

4) There are a couple of deviations from the ClinicalTrials protocol that have not been discussed:

o Your primary outcome in ClinicalTrials is change in BMI at month 6 and in this paper it is change in weight (although under primary outcome you only describe how BMI was assessed). No rationale for this change has been provided and needs to be clarified to the readers.

**Response:** We thank the Reviewer for pointing this out. Our initial idea of reporting body weight change instead of BMI was to make it easier for people to understand and reflect the results of the study. However, it created more confusion as this was a common comment from all the Reviewers. As explained in previous comments above, we are now reporting all our results as BMI change.

o Recruitment of individuals with a BMI between 25-35 is specified in the protocol but in the paper you describe that you recruited individuals with a BMI of 23-36.

**Response:** We thank the Reviewer for this comment. As this was brought up by Reviewer 1, we recommend referring to our response to Reviewer 1's comment 3 above.

o Recruitment of n=238 individuals in the ClinicalTrials protocol but n=223 in the paper.

**Response:** We thank the Reviewer for pointing out this discrepancy. Indeed, there is a mismatch in the reported number of actual enrolled individuals in the ClinicalTrials protocol (n=238) compared to the paper (n=223). The correct number is 223, as also described in the

Results section in the paper. We noticed that the discrepancy came from the fact that in ClinicalTrials, we intended to choose the option “Anticipated” (instead of “Actual”) for the number of 238 participants section under Enrolment. This is now updated in the ClinicalTrials portal to reflect the actual enrolment and align with the number we report in the paper.

In addition, we provide information to the Reviewer to assess that there were no significant differences between PGS and intervention groups of these 15 people who consented to the study, completed the baseline questionnaire, booked an appointment for a lab visit, but never showed up. In the Table below, we used Pearson's Chi-squared test with Yates' correction and found no statistically significant difference ($p = 0.13$) in the proportion of participants not included in the analysis, even though it was slightly higher in the Bottom 5% BMI PGS group compared with the Top 5% BMI PGS group. We repeated the same test with the intervention groups as well (Control vs Diet) with similar results ($p = 0.30$). The proportion of participants not included in the analysis was slightly higher in the Control group compared with the Diet group, but this difference was not statistically significant.

| | Included in analysis (No) | Included in analysis (Yes) |
|---|---|---|
| Bottom 5% BMI PGS | 9 | 82 |
| Top 5% BMI PGS | 6 | 141 |

| | Included in analysis (No) | Included in analysis (Yes) |
|---|---|---|
| Control | 10 | 110 |
| Diet | 5 | 113 |

o No mentioning of month 2 and 4 weights for the primary outcome (or any secondary outcomes) in ClinicalTrials, but these are not called “non-specified exploratory outcomes” in the paper.

**Response:** We thank the reviewer for raising this important issue regarding trial registration and outcome specification. As registered on ClinicalTrials.gov, the prespecified primary outcome of the trial was to test the interaction between the top and bottom 5th percentile of the BMI polygenic score distribution and BMI change between baseline and end-of-study (at six months), and this endpoint remains unchanged. As also mentioned in the protocol, we aimed to measure BMI at months two and four, although, as correctly pointed out, these were not specified as primary/secondary outcomes. The use of those additional data in the manuscript reflects the use of a prespecified repeated-measures data structure, rather than the introduction of additional primary or secondary outcomes.

Specifically, the reasoning for using all the repeated measures we had about BMI was purely based on the fact that linear mixed-effects models incorporate all available repeated measurements to estimate the trajectory of BMI change over time while accounting for within-participant correlation. This modelling approach does not redefine the primary outcome, nor does it alter the interpretation of the baseline-to–six-month effect, which remains consistent when analysed independently.

We acknowledge that the intermediate time points were not explicitly listed as trial outcomes in the registration (despite being mentioned in the protocol). In the revised manuscript, we therefore clarify that analyses incorporating the two- and four-month measurements represent exploratory longitudinal analyses intended to characterise temporal patterns and improve estimation efficiency, rather than to introduce additional confirmatory endpoints. Importantly, inclusion of these measurements does not change the direction or significance of the prespecified primary outcome.

We have revised the Methods and Results to explicitly distinguish prespecified outcomes from exploratory longitudinal modelling and to ensure full transparency with respect to the trial registration.

o You have specified in ClinicalTrials that you will adjust for sex and age in the models but this is not specified in the methods section of the manuscript.

**Response:** This is a very good point, which we have now addressed by adding those variables in our models and updating the *"Statistical analysis"* sub-section in Methods as shown below:

*"The model included a randomized intervention group (diet or control), BMI PGS group (top or bottom 5%), and their interaction, and was adjusted for age, sex, baseline BMI, and BMI measurement time."*

5) There is no mentioning of any adverse events despite this being a 6-month trial – were there none?

**Response:** We thank the reviewer for this suggestion. We have now added an explicit statement regarding adverse events in the Results under the *"Primary outcome"* sub-section, stating that no intervention-related adverse events were reported during the study period, and clarified in the Methods under the "*Diet intervention"* sub-section that adverse events were monitored throughout the trial.

6) The protocol and SAP could unfortunately not be found in the Supplementary Material.

**Response:** We are sorry to see that this was not made available to the Reviewers, as expected. We include this again in the package for review.

7) What was the rationale behind recruiting normal weight individuals in a study investigating the impact of genetic predisposition to obesity and effects of weight loss? It might be more

interesting including only overweight or obese individuals, as a normal weight individual with BMI 23 is likely to respond differently on energy restriction diet than a person with BMI 35, irrespectively of genetic risk score. Consequently, the included individuals with normal BMI might underestimate any possible genetic effect and thereby also potentially reduce overall statistical power. The reviewer acknowledge that the author briefly mention (line 214-217) that this study chose to focus on overweight subjects, but some further discussion of the implications and interpretation of their study should be added.

**Response:** This is a very good point, and we are happy to elaborate on the rationale. Our study leveraged the unique opportunity we had to invite people to enroll in a prospective study based on their genotype. To make our process more focused on the right group, we used previously collected BMI data for these individuals during the initial invitation. However, since these data were not very recent, we allowed for some flexibility and invited individuals whose BMI was in the 23-36 kg/m$^2$ range. The idea was that once a person was interested in participating, they would provide their current BMI via the online questionnaire, which would subsequently be validated by a nurse when they visit in person. To keep our study pool engaged and the sample size as large as possible, we decided to keep all these individuals in the study until the end, even if their BMI was slightly under or over the pre-specified range of 25-35 kg/m$^2$. For transparency, the minimum and maximum BMI range in our data analysis is 23.78 - 37.45 kg/m$^2$.

We agree with all Reviewers who raised similar concerns that strictly speaking, the overweight/obese range is 25-35 kg/m$^2$, but in our primary analysis we included all individuals we had data for with BMI range of 23.78 - 37.45 kg/m$^2$. To assess the robustness of our findings, we provide two sensitivity analyses to assess the robustness of our findings. The first strictly includes individuals within the 25-35 kg/m$^2$ range (n=196) and the second focuses on obese individuals only with a BMI within the 30-35 kg/m$^2$ range (n=101). As you can appreciate from the updated Table 2 below, we observed no evidence that the BMI polygenic score group is an effect modifier of the intervention on BMI change in any of the additional analyses, indicating similar intervention effects across the BMI polygenic score groups. In other words, within our cohort we did not observe meaningful BMI PS-dependent differences in response to dietary intervention. We added this to the *"Statistical analysis"* and *"Changes compared to the registered trial protocol"* subsections in Methods, as well as a section in Discussion, as shown below:

**Methods – Statistical Analysis:** ***"****Sensitivity analyses were conducted by refitting the primary mixed-effects model in restricted samples defined by baseline BMI. Specifically, analyses were repeated among participants with baseline BMI 25–35 kg/m² to align with the trial protocol, and among obese participants only with baseline BMI 30–35 kg/m² to assess the robustness of the primary findings to the BMI range included."*

**Methods - Changes compared to the registered trial protocol:** *"To maintain our statistical power as high as possible, we kept all individuals who enrolled in the study with a BMI ranging*

*from 23.78 - 37.45 kg/m2, and instead of removing those outside the 25-35 kg/m2 range before enrollment, we excluded them in the analysis by performing sensitivity analyses."*

**Table 2:** Interaction between BMI polygenic score group and diet intervention on BMI change.

| **Model** | **PS Group * Diet Interaction Estimate**[a] | **95% CI** | **P Value** |
|---|---|---|---|
| Model 1[b] | -0.02 | -0.45, 0.42 | 0.94 |
| Model 2[c] | -0.006 | -0.45, 0.44 | 0.98 |
| Model 3[d] | -0.08 | -0.88, 0.72 | 0.85 |

[a]Estimates represent adjusted differences in BMI change comparing polygenic score groups between randomized diet intervention and control derived from linear mixed-effects models including fixed effects for age at baseline, sex, BMI at baseline, time, intervention group, BMI polygenic score group, time and intervention group interaction, and a random intercept for each participant. To model within-subject correlation over time, the compound symmetry was selected. **[b]**Model 1 includes data from all individuals ($N_{individuals}$ = 223 and $N_{observations}$ = 863) with a BMI of 23.78 - 37.45 kg/m$^2$. **[c]**Model 2 includes data from overweight/obese individuals ($N_{individuals}$ = 196 and $N_{observations}$ = 760) with a BMI of 25 - 35 kg/m$^2$. **[d]**Model 3 includes data from obese individuals ($N_{individuals}$ = 101 and $N_{observations}$ = 390) with a BMI of 30 - 35 kg/m$^2$.

**Discussion:** *"In contrast, recent bariatric surgery studies have reported that BMI polygenic scores are associated with postsurgical weight loss trajectories, with lower genetic risk linked to greater long-term weight reduction. These studies typically involve individuals with severe obesity undergoing profound anatomical and metabolic alterations. By design, our study excluded individuals with extreme obesity, prior bariatric surgery, or diabetes and evaluated a non-surgical dietary intervention producing more modest weight changes. It is therefore plausible that genetic influences on weight loss are more detectable in the context of the large physiological perturbations induced by bariatric surgery than in lifestyle-based interventions, where behavioral and environmental factors may play a comparatively greater role. Notably, the mean BMI reduction achieved through the dietary intervention (−1.5 kg/m²) was half of the baseline difference in BMI between BMI polygenic score extremes (3 kg/m²). While the intervention was similarly effective across genetic risk groups, the magnitude of BMI reduction was modest relative to the baseline differences in BMI associated with polygenic score groups. This observation may partly explain why no differential diet intervention response by genetic risk was detected."*

## Minor to moderate comments

8) Abstract, line 41-42. Please report the BMI In the low and high BMI PS groups, not only the weight difference in kg.

**Response:** The abstract is now updated to reflect the major changes in the manuscript, including this suggestion.

9) Row 484 have an incomplete sentence “All biomarkers were analyzed in”. Please revise.

**Response:** We thank the Reviewer for catching this. This was a sentence that should have been completely removed during the previous editing versions. We have now removed it.

10) How much missing data did you have for the exploratory outcomes?

**Response:** We thank the reviewer for the suggestion to provide details on missing data for the exploratory outcomes. Overall, the missingness for all the labs included in the exploratory analysis was very low, with the baseline at 2/223 and follow-up at 13/223 individuals with missing measures. Both baseline and follow-up missingness for all six labs were similarly distributed across the intervention and BMI PS groups. When tested in a multivariable logistic regression including intervention group, PS BMI group, age, sex, and baseline BMI, no covariate was significantly associated with missingness (all $p > 0.2$), indicating that missing data were unlikely to be systematically related to observed participant characteristics. Related to this comment, we now added the relevant information in the Methods “*Statistical analysis*“, and Results *“Non-prespecified exploratory analysis”* sub-sections, as shown below.

**Methods - Statistical analysis***: “The number and proportion of missing laboratory values were tabulated by intervention and PS BMI group. Associations between missingness and baseline covariates were assessed using logistic regression.”*

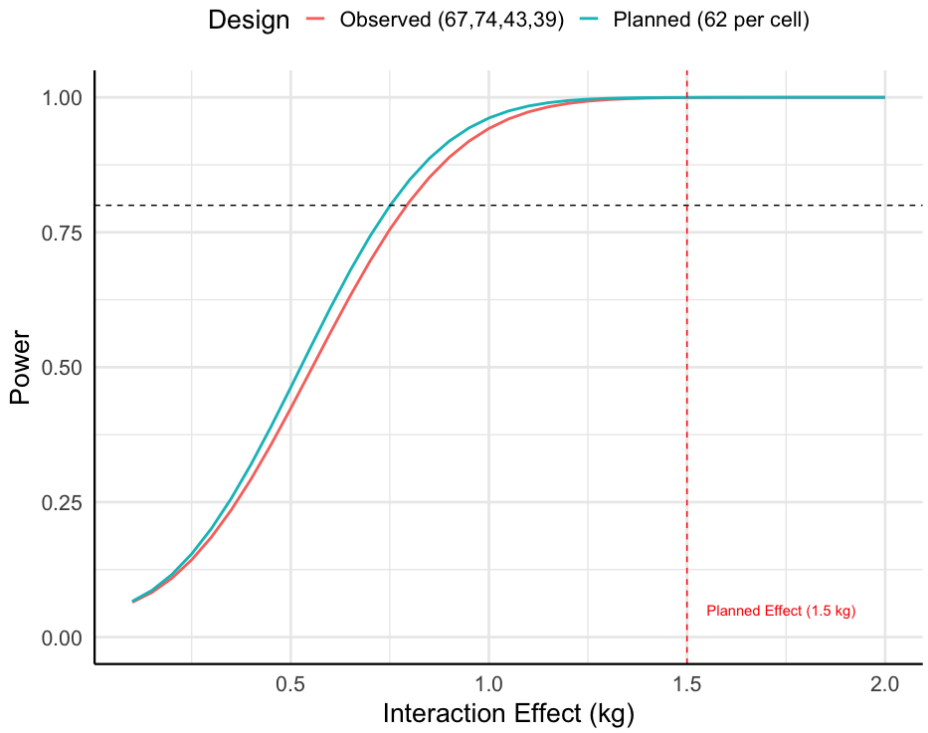
**Results - Non-prespecified exploratory analysis:** *“Missing baseline and follow-up laboratory values were low (5.8%) and similarly distributed across intervention and PS BMI groups. Logistic regression including intervention, PS BMI group, age, sex, and baseline BMI showed no significant associations with missingness (all P > 0.2), supporting the assumption that missing data were not systematically related to observed characteristics.”*

11) Based on your a priori power calculation you estimated you would need 62x4=248 individuals, but you only recruited 223 – please provide justification for this and discuss its potential implications.

**Response:** We agree with the Reviewer that group sizes were unequally balanced relative to the original design. To evaluate the impact on statistical sensitivity, we performed power calculations using the observed sample distribution and standard linear model assumptions (α =

0.05). These analyses indicate ≥80% power to detect the pre-specified 15% difference in diet intervention effect by BMI PS group (our target interaction effect) despite the unequal group sizes. Although power is reduced for smaller effects compared with the planned balanced design, the study remains adequately powered to test the hypothesised effect magnitude **(see figure below)**. To enhance transparency, we have clarified this in the manuscript under the "*Changes compared to the registered trial protocol" section*.

**Methods - Changes compared to the registered trial protocol:** *"The updated power calculations reflect this revised scenario, and post hoc power calculations based on the achieved sample sizes indicate that power remained >80% for detecting the prespecified interaction effect."*




In addition, we have revised our abstract, discussion, and conclusion parts to more accurately align with the finding's interpretation based on the statistical power of our study. To make it easier, we kept the major changes as tracked in the manuscript.

**Abstract:** *"These findings suggest that, in this population, genetic susceptibility to higher BMI, as captured by current genome-wide polygenic scores, does not substantially modify short-term weight loss response to dietary intervention, underscoring the limited utility of these scores for stratifying lifestyle-based weight loss strategies by genetic risk."*

**Discussion:** *"However, we observed no statistically significant evidence that the BMI polygenic score modified the effect of diet intervention on BMI change. Within the limits of the study's power, these findings suggest that genetic susceptibility to higher BMI, as captured by current genome-wide polygenic scores, does not substantially alter response to dietary intervention."*

**Conclusion:** *"Modest interaction effects cannot be excluded given the sample size, and larger studies are needed to evaluate smaller effects. Our findings suggest that, in this population, current genome-wide BMI polygenic scores have limited utility for identifying individuals most likely to benefit from dietary interventions aimed at reducing body weight, underscoring the importance of critically evaluating genetic tools before their integration into clinical decision-making."*

12) The meal plans in the dietary intervention needs to be described in more detail, by e.g., providing an example of a rotating weekly meal plan. It is also unclear from the methods section whether participants received meal plans and/or actual meals. On ClinicalTrials it says "Participants will get 5 meals/day…" but in the paper you state "Participants received structured meal plans of five daily meals…". Please clarify this. As you also provided many interventions within the broader coaching program, I think it is necessary to provide more details on what was actually done. Also, the authors need to acknowledge/describe the control group as well.

**Response:** We thank the Reviewer for this comment. As explained in the revised manuscript, participants assigned to the diet intervention group received through the application personalized structured meal plans (and not ready-made meals) that included five meals per day. In the same way, we refer to a "recommended food plan" in the ClinicalTrials portal, as shown below:

"*Participants will get 5 meals/day based on their calorie needs minus 500 kcal/day less than their daily calorie expenditure. All meals will be tracked via a mobile application, including any meals outside the recommended food plan".*

However, participants had the option to substitute a meal for another of equal energy profile from a long list of options. For example, if someone did not like the suggested breakfast option of 100g Greek yogurt with one apple for Monday morning, they could substitute it with 2 slices of bread and 2 slices of cheese. We believe the sentence in the Diet intervention section *"Participants could follow or adapt the plan to meet their energy target by voluntarily logging meals in the ViaEsca mobile app to monitor intake."* adequately explains that participants were not provided a ready-made meal, but they were coached on how to prepare the suggested meal plan (or then modify per their own preferences and tastes from the provided options within the application). In the Methods "*Diet intervention"* sub-section, we provide an example of how a rotating suggested weekly meal plan looked, as suggested (Supplementary Material File 4).

Last, we now added a sentence explaining how individuals in the control group were handled in the study (Recontacting/Methods), as shown below:

*"Participants who were assigned to the control group were informed by the research nurse that no nutritional coaching or similar advise will be given to them during the study and that the team would be reaching out to them to report their body weight two more times, at months two and four, and to book their end-of-study in-person visit at the clinic for blood collection and body weight measurement six months after their enrollment that will sucessfully conclude their participation."*

13) Please specify whether baseline measures of the outcomes (primary, secondary and exploratory) were adjusted for in the analyses, as this is generally advised to do when assessing change from baseline in outcomes.

**Response:** We thank the Reviewer for the opportunity to clarify this important point. We do agree that adjusting for baseline BMI when the change in BMI is your outcome is recommended. In fact, in all the models related to the main outcome (Table 2 and Fig 3) we adjusted for several covariates, including baseline BMI. This is now clarified in Statistical analysis section in methods as shown below:

*"The model included a randomized intervention group (diet or control), BMI PGS group (top or bottom 5%), and their interaction, and was adjusted for age, sex, baseline BMI, and BMI measurement time."*

14) Line 195-196 states "though diet effectiveness was similar across sexes." – where are these results shown and were they prespecified? How was this analyzed? I am missing this information in the paper.

**Response:** Very good point here, as this statement came from the initial models that we included age and sex, which were then removed from the paper. However, we now present our results adjusting for age and sex, and therefore we will keep this as is in the manuscript (Please see full model covariate estimates in Supplementary Table 1).

15) The drop-out rate was very low, which is a strength that should be mentioned in the discussion together with other strengths mentioned by the authors.

**Response:** We thank the Reviewer for this suggestion. This is now highlighted in the Discussion part, where we outline the strengths of our study, as shown below:

*"Despite the potential for higher dropout in control groups, the dropout rate was 5.8% and similar across the intervention groups, highlighting strong participation retention and supporting the robustness of our findings."*

16) Why have you chosen to describe the population stratified on sex in table 1?

**Response:** As part of the major revisions we made in our manuscript, Table 1 is now shown without any stratifications.

17) Line 199-200 states “improvements in cardiometabolic biomarkers suggest meaningful physiological changes beyond weight loss.” – Provide justification for this statement as weight loss is known to impact meaningful physiological changes to cardiometabolic biomarkers.

**Response:** We thank the Reviewer for this comment. Given the major updates in our manuscript, this sentence is now modified as follows:

*“Third, we did not collect detailed data on calorie intake, which limited our ability to examine the behavioral mechanisms underlying weight loss and BMI reduction. However, using data on adherence to meal and activity logging, we observed additional, but relatively small reductions in BMI beyond the effect of diet alone.”*

18) Line 211-212 beginning with “Another strength…” can be removed. This is the fundamental pillar of randomized trials and you have already stated that one strength is the randomized design.

**Response:** This is now removed, as suggested.

19) Line 213: In results, drop-out was 5.8% and not <5%. Please be consistent throughout the manuscript with all the numbers provided. As of now, it is difficult for the reader to navigate this manuscript due to different numbers being provided in the different sections.

**Response:** We thank the Reviewer for catching this. It is now updated to the correct number, which is 5.8%.

20) “In conclusion, this 6-month randomized dietary intervention study found no evidence that BMI polygenic scores predict differential weight loss outcomes among overweight and mildly obese non-diabetic individuals.” – please remove as prediction was not your aim.

**Response:** We thank the Reviewer for the valid comment. We have now updated our conclusion part to reflect the major revisions in our manuscript, as shown below:

*“In conclusion, this prospective randomized dietary intervention study found no statistically significant evidence that BMI polygenic scores modified the intervention effect on BMI change among overweight and mildly obese non-diabetic individuals. Modest interaction effects cannot be excluded given the sample size, and larger studies are needed to evaluate smaller effects. Our findings suggest that, in this population, current genome-wide BMI polygenic scores have limited utility for identifying individuals most likely to benefit from dietary interventions aimed at reducing body weight, underscoring the importance of critically evaluating genetic tools before their integration into clinical decision-making.”*

21) There are little information about intervention adherence, but it would be interesting to know if there were any differences among the low and high PS BMI groups? If so, this needs discussion and added into the interpretation.

**Response:** We thank the Reviewer for this suggestion. Via the application data retrieval, we had access to the number of days each person logged their meals and the number of days their activity data were synced. Participants logged meals on a mean of 64.2 (SD, 66.8) days during the six-month intervention, and their activity data were recorded on a mean of 125.0 (SD, 72.8) days. To further evaluate whether interaction with the application, despite being voluntary, was associated with a differential BMI reduction, we ran linear regression models with BMI change ($kg/m^2$) as the dependent variable. We found that people with higher adherence to meal logging ($\beta_{BMI\_change}$ = -0.011 $kg/m^2$, 95%CI [-0.017, -0.007]) and activity tracking ($\beta_{BMI\_change}$ = -0.006 $kg/m^2$, 95%CI [-0.011, -0.001]) had a higher BMI reduction, however, the effect size is very small in comparison to the mean BMI change due to the intervention ($\beta_{BMI\_change}$ = -1.5 $kg/m^2$ in the same group. These exploratory findings are now reported in the Results under the *"Non-prespecified exploratory analysis"* subsection. We also updated the *"Diet intervention"* subsection in the Methods to reflect this and added a short paragraph to the Discussion.

**Results - Non-prespecified exploratory analysis***: "Participants logged meals on a mean of 64.2 (SD, 66.8) days during the six-month intervention, and their activity data were recorded on a mean of 125.0 (SD, 72.8) days. Greater meal logging adherence was significantly associated with greater BMI reduction (βBMI_change = -0.011, 95%CI [-0.017, -0.007]). Similarly, greater activity data logging adherence was significantly associated with greater BMI reduction (βBMI_change = -0.006, 95%CI [-0.011, -0.001]). However, although statistically significant, the estimated effect sizes were small relative to the overall mean BMI reduction of −1.5 kg/m² observed during the intervention, indicating that logging adherence alone explains only a limited portion of the total weight loss."*

**Methods – Diet intervention:** *"Adherence to the nutritional intervention was quantified using app-derived engagement metrics. Meal logging adherence was defined as the number of days participants logged dietary intake in the mobile application during the six-month intervention period. Activity tracking adherence was defined as the number of days during which step counts were recorded via the wearable device. Both metrics were analyzed as raw counts and as proportions relative to the total study duration."*

***Discussion:*** *"Greater adherence to meal and activity logging in the diet intervention group was associated with larger reductions in BMI but the effect sizes were relatively small compared to the effect of diet."*

## Reviewer #5 (Remarks to the Author):

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

**Response:** We thank the Reviewer for their detailed and constructive comments. We also acknowledge and appreciate the co-review conducted under the Nature Communications early career researcher initiative, which promotes training and recognition in peer review.

**REVIEWER COMMENTS**

**Reviewer #1 (Remarks to the Author):**

Thank you for the careful and thorough responses to the reviewers and for the substantial revisions made to the manuscript. The study addresses an important question at the intersection of polygenic risk and behavioral interventions and represents an innovative application of a genetic-first recontact strategy in a randomized trial setting. The manuscript is generally well written, the analyses are appropriate, and many of the reviewers' concerns have been satisfactorily addressed in the revised version.

I particularly appreciate the additional analyses, clarification of the statistical modeling, improved presentation of baseline characteristics across PS × intervention groups, and the expanded discussion of limitations. The study provides an important and well-executed contribution to the literature on the clinical utility of polygenic scores in lifestyle-based weight loss interventions.

That said, I still have one remaining concern regarding how the study's limitations are communicated in the Abstract.

1. Interpretation of the negative interaction result in the Abstract

Although the Discussion now appropriately acknowledges the limitations related to statistical power and the possibility that smaller interaction effects may exist, the Abstract conclusion still reads too definitively. Currently, the wording suggests that BMI polygenic scores do not substantially modify response to dietary intervention and highlights the limited utility of such scores for stratifying lifestyle interventions.

Given the study's sample size and the detectable effect threshold described by the authors (approximately ≥15% difference in intervention effect by PS group), the trial is primarily able to rule out large interaction effects, but it cannot exclude more modest or subtle gene–environment interactions. As currently written, the Abstract may lead readers to interpret the findings as stronger evidence against any modifying effect than the data can support.

In addition, the exploratory adherence data suggest highly heterogeneous engagement with the intervention application, with meal logging showing a mean of ~64 days but a standard deviation exceeding the mean. This implies that a non-trivial fraction of participants may have had minimal behavioral exposure to the intervention, which further limits the strength of conclusions about effect modification.

For these reasons, I recommend that the Abstract explicitly acknowledge these limitations, rather than placing them only in the Discussion or Conclusion sections.

A possible revision could be along the lines of:

"While these findings suggest that genetic susceptibility to higher BMI, as captured by current genome-wide polygenic scores, does not strongly modify short-term weight loss response to dietary intervention, modest interaction effects cannot be excluded given the study's limited power and variability in engagement with the intervention. Larger studies with more consistent behavioral exposure will be required to clarify the role of polygenic scores in stratifying lifestyle-based weight loss strategies."

The precise wording is of course up to the authors, but the key point is that the Abstract should clearly communicate that modest interaction effects cannot be excluded.

**Response:** We thank the reviewer for this thoughtful comment. We agree that the abstract should more explicitly acknowledge that modest interaction effects cannot be excluded given the study's sample size, and we have revised it accordingly, as you can see below. However, we respectfully disagree with linking the variability in logging adherence to the abstract-level conclusions. Our primary outcome (BMI) was directly and systematically measured in all participants and is independent of voluntary logging behavior. The adherence data were non-prespecified and exploratory, and we believe prominently featuring their limitations in the abstract would misrepresent the reliability of our primary findings. Nonetheless, as you correctly pointed out, we have ensured these nuances are clearly addressed in the Discussion.

**Abstract:** *"These findings suggest that, in this population, genetic susceptibility to higher BMI, as captured by current genome-wide polygenic scores, does not substantially modify short-term weight loss response to dietary intervention. However, given the study's sample size, modest interaction effects cannot be excluded, and larger studies are needed to clarify the role of polygenic scores in stratifying lifestyle-based weight loss interventions."*

2. Interpretation of the intervention exposure

Related to the above point, the adherence metrics indicate substantial heterogeneity in engagement with the intervention application (e.g., mean meal logging ~64 days with SD ~67). While the exploratory analyses linking logging frequency to BMI change are helpful, the manuscript would benefit from a brief sentence acknowledging that the behavioral "dose" of the intervention was variable across participants, which may limit the ability to detect effect modification.

This does not require additional analyses but should be reflected in the interpretation of the findings.

**Response:** We thank the reviewer for this suggestion. We agree that engagement with the intervention application was variable across participants, as reflected by the heterogeneity in meal logging frequency (mean 64.2, SD 66.8 days). We have added a brief acknowledgment of this variability in the manuscript. However, we would like to clarify that our primary outcome, BMI, was directly and systematically measured in all participants, independent of logging adherence. The variability in app engagement, therefore, does not affect the reliability of our BMI-based analyses, including the genetic interaction analysis, which is based on objectively measured phenotypic data rather than self-reported behavioral engagement.

**Discussion:** *"Third, we did not collect detailed data on calorie intake, which limited our ability to examine the behavioral mechanisms underlying weight loss and BMI reduction. Using data on adherence to meal and activity logging, we observed that greater adherence was associated with larger reductions in BMI, though effect sizes were relatively small. Engagement with the intervention application was variable across participants (mean 64.2, SD 66.8 days logged), which should be considered when interpreting these exploratory analyses. Importantly, BMI was directly and systematically measured in all participants, independent of logging behavior, and the*

*variability in app engagement does not affect the reliability of our primary BMI-based findings."*

Overall, this is a thoughtful and well-conducted study that provides valuable insight into the challenges of prospectively testing polygenic score–guided behavioral interventions. Addressing the points above—particularly clarifying the limitations in the Abstract—would strengthen the interpretation and ensure the conclusions remain appropriately calibrated to the data.

Reviewer #1 (Remarks on code availability):

not currently

**Reviewer #2 (Remarks to the Author):**

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

**Response:** We thank the Reviewer for their detailed and constructive comments. We also acknowledge and appreciate the co-review conducted under the Nature Communications early-career researcher initiative, which promotes training and recognition in peer review.

**Reviewer #3 (Remarks to the Author):**

I thank the authors for their careful revisions and responses to prior review. This MS does present convincing evidence against the interaction tested.

I have only a minor criticism at this point.
The authors now report additional analyses related to meal and activity logging. The units and exact method are not reported, which makes interpretation difficult. Given that the coefficients are small but with narrow confidence intervals, my guess is that the logging adherence was included as a number of adherent days. In this case, the size of the coefficient would not indicate the size of the effect, especially since the apparent distributions (especially for meal logging) are likely skewed. Without knowing the units the authors explanation that the associations are small is not evaluable. Given the modest sample size it is not especially plausible that an association of negligible size would be statistically significant.

If these are included please describe the methods a bit more, for example were both logging types modeled simultaneously or separately, what was the scale of the coefficient, and how did including logging change the expected weight change at a fixed level of logging, or how much of weight change do they explain, using and R^2 or something similar.

I did scan through the linked R code but did not see those analyses.

**Response:** We thank the reviewer for this comment. To clarify the methods, meal logging and activity logging were each modeled separately in univariate regression models against BMI change, with the predictor unit being the number of days logged during the six-month intervention. However, we would like to be cautious about over-interpreting these findings for several reasons. First, both analyses were exploratory and not prespecified as primary outcomes. Second, the logging data were incomplete across participants as reflected by the large standard deviations, indicating substantial variability and a likely skewed distribution. Third, all participants had the option to contact diet coaches by chat or phone during the program, which may have supported adherence independently of app logging, information we unfortunately did not capture.

Given these limitations, scaling the coefficient to an SD increase would imply a level of precision that the data do not support. Just for your reference, meal logging adherence explains 17% ($R^2$=0.17) and activity logging 4% ($R^2$=0.04) of the variance in six-month BMI change. We note that the relatively higher $R^2$ for meal logging compared to its small per-day beta coefficient is partly a mathematical consequence of the large variance in logging days (SD 66.8), in other words, a large spread in the predictor can inflate $R^2$ even when the per-unit effect is small. This further supports our caution against over-interpreting these associations, as the apparent explanatory power is partly driven by the very heterogeneity in logging behavior that limits the reliability of these data. While statistically significant, these associations explain only a modest proportion of overall BMI reduction and should be interpreted cautiously given the exploratory and incomplete nature of the logging data. Below, you can see the relevant modifications we made to the paper in response to this comment.

**Methods - Statistical analysis:** *"Meal logging and activity logging were each modeled separately in univariate regression models against BMI change, with the predictor unit being the number of days logged during the six-month intervention."*

**Discussion:** *“Third, we did not collect detailed data on calorie intake, which limited our ability to examine the behavioral mechanisms underlying weight loss and BMI reduction. Using data on adherence to meal and activity logging, we observed that greater adherence was associated with larger reductions in BMI, though effect sizes were relatively small. Engagement with the intervention application was variable across participants (mean 64.2, SD 66.8 days logged), which should be considered when interpreting these exploratory analyses. Importantly, BMI was directly and systematically measured in all participants, independent of logging behavior, and the variability in app engagement does not affect the reliability of our primary BMI-based findings.”*

**Reviewer #4 (Remarks to the Author):**

We thank the authors for their response to our comments, which to some extent clarifies several raised issues and concerns. However, for a few comments, we do not fully understand the response and would here like additional clarifications and further revisions in manuscript where needed. Please respond to our comments below to the specific responses where we ask for clarification.

1) Line 143-146: Please provide treatment effect estimates, i.e., between-group estimates with corresponding CI. It is not sufficient to provide within-group changes and report p-values. This is also the case for all non-specified exploratory analyses. Although one might include observed within-group changes in the results, these are not interpreted as causal effects of the intervention and thus treatment effects should be shown and interpreted from contrasting the two arms.

**Reviewer's comment:** You have only presented the estimates for the within group changes but you have still not presented between group differences which are highly relevant to see. Thus, please present those estimates, for primary as well as the secondary outcomes. The authors should not draw conclusions from within group changes alone, which we believe is done.

**Response:** We thank the reviewer for this comment, and we apologize for not being able to fully understand your point previously. We hope this now addresses your request.

**Results - Primary outcome:** *"In the diet intervention group, the adjusted mean change in BMI (kg/m$^2$) at two months was −1.18 (95% CI, −1.38 to −0.97, n=109), at four months was −1.61 (95% CI, −1.82 to −1.40, n=108), and at six months (end of study) was −1.51 (95% CI, −1.72 to −1.31, n=106), whereas in the control group the mean change in BMI (kg/m$^2$) at two months was −0.31 (95% CI, −0.52 to −0.11, n=107), at four months was −0.24 (95% CI, -0.45 to −0.04, n=106), and at six months (end of study) was -0.01 (95% CI, −0.22 to 0.19, n=104).* ***The difference in BMI change between the diet intervention and control groups was -0.87 (95% CI, -1.15 to -0.58) at two months, -1.37 (95% CI, -1.66 to -1.08) at four months, and -1.50 (95% CI, -1.79 to -1.21) at six months.***

*The adjusted mean change in BMI (kg/m$^2$) in the top 5% BMI PS group at two months was -0.67 (95% CI, -0.84, -0.49, n=109), at four months was -0.85 (95% CI, -1.03, -0.67, n=108) and at six months was -0.69 (95% CI, -0.86, -0.51, n=106), whereas in the bottom 5% BMI PS group at two months was -0.82 (95% CI, -1.03, -0.61, n=109), at four months was -1.00 (95% CI, -1.21, -0.79, n=108) and at six months was -0.84 (95% CI, -1.05, -0.63, n=106).* ***The difference in BMI change between the Top 5% and Bottom 5% BMI PS groups was 0.15 (95% CI, -0.12 to 0.42) at two months, 0.15 (95% CI, -0.13 to 0.43) at four months, and 0.15 (95% CI, -0.12 to 0.42) at six months.*** *The trajectory of BMI change (kg/m$^2$) between the control and diet intervention by the BMI PS groups is shown in Figure 3."*

2) Supplementary Table 2: Changes of non-exploratory clinical test values of individuals by intervention and BMI PS groups at the end of the study.

**Reviewer's comment:** You have not presented the between group differences here, please do.

**Response:** We thank the reviewer for this comment, and we apologize for not being able to fully understand your point previously. We hope this now addresses your request.

**Results - Non-prespecified exploratory analysis:** "*The between groups difference in total cholesterol was −0.20 mmol/l (95% CI: −0.38 to −0.02), LDL cholesterol −0.18 mmol/l (95% CI: −0.34 to −0.02), HDL cholesterol −0.01 mmol/l (95% CI: −0.06 to 0.04), triglycerides −0.14 mmol/l (95% CI: −0.35 to 0.07), HbA1c −0.56 mmol/l (95% CI: −1.29 to 0.17), and glucose −0.14 mmol/l (95% CI: −0.30 to 0.02).*"

14) Line 195-196 states "though diet effectiveness was similar across sexes." – where are these results shown and were they prespecified? How was this analyzed? I am missing this information in the paper.

**Reviewer's comment:** We do not understand; if you have adjusted for sex in the models, how can you state that there was an similar diet effect across sex? Where do you show these results, as the Full model in Suppl Table 1 do not provide data that the effect was similar. Please clarify.

**Response:** We thank the reviewer for this important clarification. You are correct that adjusting for sex in the model does not allow us to conclude that diet effectiveness was similar across sexes. To properly test for sex differences in intervention effectiveness, a sex × time interaction term would be required, which was not part of our pre-specified analysis plan. What we wanted to report is that sex was included as a covariate in the adjusted model, and it was not a significant predictor of BMI change. We have removed the inference about similar effectiveness across sexes and corrected the language used in the manuscript, as shown below:

**Discussion:** *"Our study has several limitations. First, there was a higher proportion of women (74.9%) than men, likely due to known sex differences in volunteer participation, though sex was not a statistically significant predictor of BMI change in the adjusted model."*

17) Line 199-200 states "improvements in cardiometabolic biomarkers suggest meaningful physiological changes beyond weight loss." – Provide justification for this statement, as weight loss is known to impact meaningful physiological changes to cardiometabolic biomarkers.

**Response:** We thank the Reviewer for this comment. We have revised this statement to more accurately reflect the findings of our study as follows:

**Discussion:** *"Improvements in key cardiometabolic biomarkers, including total and LDL cholesterol, suggest that the weight loss achieved was sufficient in magnitude and duration to produce detectable and meaningful changes in cardiometabolic risk factors across genetic risk groups."*

“Third, we did not collect detailed data on calorie intake, which limited our ability to examine the behavioral mechanisms underlying weight loss and BMI reduction. However, using data on adherence to meal and activity logging, we observed additional, but relatively small reductions in BMI beyond the effect of diet alone.”

**Reviewer's comment:** This paragraph is unclear and needs rewriting. If you did not have data on calorie intake, how can you draw conclusions on separating specific effects on BMI beyond diet? The strongest predictor of BMI reduction in a decrease in total calorie intake. Can the authors please clarify the above reasoning and better explain what you have done.

**Response:** We thank the reviewer for this important clarification. The original sentence was misleading, as it implied we could separate the effect of logging adherence from the effect of diet without having calorie intake data. We have revised the paragraph to remove this claim entirely. The revised text now accurately describes logging adherence as an exploratory measure of engagement with the intervention application, associated with BMI reduction but not interpreted as a mechanism independent of dietary behavior. We have also added clarification that BMI was directly and systematically measured in all participants, independent of logging behavior, ensuring that the variability in app engagement does not affect the reliability of our primary findings.

**Discussion:** *“Third, we did not collect detailed data on calorie intake, which limited our ability to examine the behavioral mechanisms underlying weight loss and BMI reduction. Using data on adherence to meal and activity logging, we observed that greater adherence was associated with larger reductions in BMI, though effect sizes were relatively small. Engagement with the intervention application was variable across participants (mean 64.2, SD 66.8 days logged), which should be considered when interpreting these exploratory analyses. Importantly, BMI was directly and systematically measured in all participants, independent of logging behavior, and the variability in app engagement does not affect the reliability of our primary BMI-based findings.”*

**Reviewer #5 (Remarks to the Author):**

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

**Response:** We thank the Reviewer for their detailed and constructive comments. We also acknowledge and appreciate the co-review conducted under the Nature Communications early-career researcher initiative, which promotes training and recognition in peer review.