nature portfolio

Peer Review File

Birthweight Phenotype Genotype Mismatch in Cardiometabolic Programming.

Corresponding Author: Professor Craig Pennell

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 topic of birthweight phenotype, genetic predisposition and the risk of adverse metabolic health is of great interest. Using longitudinal data spanning from birth through age 27 years, on a cohort of more than 1000 individuals, this analysis used a phenotype-genotype interaction approach to study birthweight determining gene with direct measures of cardiometabolic risk. The Barker hypothesis postulates that phenotype-genotype mismatch is associated with risk of adverse cardiometabolic health. In this analysis, the researchers used offspring birthweight polygenic risk scores to support their hypothesis that the mismatch may underlie the developmental pathway to adverse cardiometabolic health. Overall, this is a very interesting analysis.

There are some points with the results that require clarification.
1. Table 1 does not display differences between the 3 groups (although line 230 states there are differences), were there any significant differences across the 3 groups for each measure?
2. Please explain why the largest phenotypic decile of the PGS effect is used as the referent? I would think the middle or normal decile should be the referent.
3. Is there a reason why the findings of DEXA fat mass and waist circumference with a significant interaction are not displayed in the main tables along with BMI and are placed in a supplementary table. These are interesting findings and reported as primary in the conclusion, thus they should be displayed in the main tables.

4. lines 302-303: what is marginal mean anthropometry, where in the results is this reported?

5. line 367, avoid the term Caucasian and use White (this terminology is from "Updated Guidance on the reporting of race and ethnicity in medical and science journals" PMID: 34402850

Reviewer #2

(Remarks to the Author)
The authors have done a commendable job of investigating the Barker hypothesis, but with a logical and novel way of defining poor fetal growth - as mismatch between birthweight genotype and phenotype. The manuscript has been written well, the findings are easy to understand and interpret.

Reviewer #3

(Remarks to the Author)
The manuscript uses a novel look at the old Barker hypothesis / Developmental Origins of Adult Health and Disease hypothesis (DoHAD). The authors use data from the Raine study in Australia, which incudes information from birth and at adult ages. The authors find that a birth weight phenotype-genotype mismatch is associated with several cardiometabolic disease risk factors in young adults.

The paper is clearly written. It uses an appropriate data resource. Further clarity is needed in a few places. Consistent language should be used when referring to the results as to whether they are from a prediction or an association analysis; it seems as though the paper uses the latter approach. In the discussion the results could be further contextualized in terms of

their effect sizes and clinical relevance.

Specific comments
P4 L4-5 The "Barker Hypothesis" has been generally replaced with the "Developmental Origins of Adult Health and Disease hypothesis". Consider rephrasing.

P4 L5 Counfounding by what? Please describe.

P5 L20 Are these analyses about prediction or association? Eventually it seems that associations are being investigated, thus this type of terminology should be used throughout the manuscript.

P6 L15 How accurate are the birth weight values in the hospital records? Has there been a validation study? If not, please describe the potential effects on the results.

P6 L8 Why were three adult ages included in the analyses? Were the outcomes chosen preferentially from one age over another?

P6 L16 The sentence about age at follow up seems out of place here. I suggest moving it.

P8 L8 How was obesity defined (e.g. which cut-off value)?

P10 L14 Given that the main analytic method was a linear mixed effects model, this suggests that the purpose of the analyses was not prediction but rather an investigation of associations. The language throughout the manuscript should reflect this.

P12 L11 It is still unclear how data from all of these adult follow ups were used in the analyses. Please specify this back in the results section.

P13 Discussion Consider adding reflections on the clinical meaning of the effect sizes observed.

P17 L29 This birth weight value is very specific; were the hospital records of sufficient quality to support this level of precision?

P19 L5-18 Although interesting, this paragraph does not relate to the results in this manuscript. Suggest shortening or deleting.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The revisions and attention to the reviewers/editors feedback have improved the readability of the manuscript. I noticed the changes to Table 3 (not table 2) where both primary and secondary outcomes are displayed. I agree that placing these findings in the main tables provides interesting data for the reader.
I agree with the authors that the Developmental origins of disease framework refers is all encompassing of low and high birth weight and risk of obesity/cardiometabolic diseases, whereas the Barker hypothesis specifically focuses on low birth and risk of metabolic disease. Thus, maintaining the Barker hypothesis terminology for this analysis is justified.

My only other comment is the use of the term adiposity- BMI is an indirect measure of adiposity, whereas waist circumference can be considered a measure of adiposity. In my read-through of this version, I do not note any issues confusing BMI with adiposity.

Reviewer #1 (Remarks to the Author):

This topic of birthweight phenotype, genetic predisposition and the risk of adverse metabolic health is of great interest. Using longitudinal data spanning from birth through age 27 years, on a cohort of more than 1000 individuals, this analysis used a phenotype-genotype interaction approach to study birthweight determining gene with direct measures of cardiometabolic risk. The Barker hypothesis postulates that phenotype-genotype mismatch is associated with risk of adverse cardiometabolic health. In this analysis, the researchers used offspring birthweight polygenic risk scores to support their hypothesis that the mismatch may underlie the developmental pathway to adverse cardiometabolic health. Overall, this is a very interesting analysis.

There are some points with the results that require clarification.
1. Table 1 does not display differences between the 3 groups (although line 230 states there are differences), were there any significant differences across the 3 groups for each measure?

Response: We thank the reviewer for the opportunity to explain. There were significant differences as detailed in the text; however, as this table is a description of the baseline demographics in our groups (i.e. not part of an a priori hypothesis) we don't think it is appropriate to test the differences. We ran the requested test in table 2 for the primary outcomes with adjustment for confounders. Our text aims to draw attention to the demographic differences between the groups (in particular the perinatal differences) exactly to justify the confounders used in primary models.

2. Please explain why the largest phenotypic decile of the PGS effect is used as the referent? I would think the middle or normal decile should be the referent.

Response: Given the conceptual framework in figure 1A, we believe there are true FGR participants in the middle deciles, whereas the top decile is mostly without FGR. As the Barker hypothesis relates to growth restriction, we decided that the top decile was the more appropriate reference. Of course, following our analysis we also suggest that fetal growth *excess* could be a programming event because the HOMA-IR effect going from positive to negative. This could then suggest that the middle deciles were in fact a better reference after all, but we leave that debate for review articles once replication ensures the genuine nature of programming from phenotype-genotype mismatch. In the manuscript we have written "*To avoid bias from fetal growth restriction in the middle eight deciles, we used the top decile as the reference and the interaction with the lowest decile as the primary outcome*" (page 10, line 11-13)

Action: To further clarify our position, we have now added to the discussion of the HOMA-IR results "*This result questions our a priory assumption that programming is a phenomenon related exclusively to FGR and highlight fetal growth excess as a potential programming event. "* (page 16, line 14-16)

3. Is there a reason why the findings of DEXA fat mass and waist circumference with a

significant interaction are not displayed in the main tables along with BMI and are placed in a supplementary table. These are interesting findings and reported as primary in the conclusion, thus they should be displayed in the main tables.

Response: We decided to examine BMI, systolic BP, LDL-cholesterol and HOMA-IR as the primary endpoints as these were well known measures previously examined in the DOHaD literature. The risk of type 1 error was not adjusted for the multiple testing related to our secondary endpoints. Some of these were also added following our internal review of the original manuscript draft (e.g. waist circumference, non-HDL cholesterol and triglycerides). To highlight this aspect, the secondary measures were kept in the supplementary and not included in the main table as per this comment. We apologise that they were "reported as primary" as we tried to caution the reader in our discussion. For instance when writing about dyslipidemia we stated "*Our data suggest a reduction in HDL-c and an increase in triglycerides with a phenotype-genotype mismatch, but the secondary and borderline nature of the significance estimates weakens our confidence in the association*" (page 16, line 20-2). Our approach was a trade-off between methodologic ease-of-understanding with result readability. As the journal editors agree with your position we will adapt our table.

Action: We have reformatted table 2 to include results for the secondary outcomes meeting our significance threshold after a bar clearly marking the secondary nature of these outcomes. The "significant" results in the secondary outcomes have been kept without **bolding** to avoid any confusion.

4. lines 302-303: what is marginal mean anthropometry, where in the results is this reported?

Response: This was simply referring to BMI and waist circumference as one. We agree that this is unnecessarily confusing.

Action: We have changed the wording to "*as the BW-PGS increased, BMI changed from normal- to overweight*" (page 16, line 6-7)

5. line 367, avoid the term Caucasian and use White (this terminology is from "Updated Guidance on the reporting of race and ethnicity in medical and science journals" PMID: 34402850

Action: This has been corrected.

Reviewer #2 (Remarks to the Author):

The authors have done a commendable job of investigating the Barker hypothesis, but with a logical and novel way of defining poor fetal growth - as mismatch between birthweight genotype and phenotype. The manuscript has been written well, the findings are easy to understand and interpret.

Response: Thank you.

Reviewer #3 (Remarks to the Author):

The manuscript uses a novel look at the old Barker hypothesis / Developmental Origins of Adult Health and Disease hypothesis (DoHAD). The authors use data from the Raine study in Australia, which incudes information from birth and at adult ages. The authors find that a birth weight phenotype-genotype mismatch is associated with several cardiometabolic disease risk factors in young adults.

The paper is clearly written. It uses an appropriate data resource. Further clarity is needed in a few places. Consistent language should be used when referring to the results as to whether they are from a prediction or an association analysis; it seems as though the paper uses the latter approach.

Action: This use of prediction has been corrected to association for models of adult cardiometabolic measures. For ease of understanding we have kept the term "genetically predicted birthweight" when discussing the BW-PGS in plain terms and hope this is acceptable.

In the discussion the results could be further contextualized in terms of their effect sizes and clinical relevance.

Response: We appreciate this comment; however, as we have discussed there are several necessary steps before our findings can be thought of as clinically relevant.

1) Replicating the interaction in independent datasets.
2) Determining whether gestational age adjusted birthweight is a better phenotypic trait as compared to crude birthweight
3) What the optimal cut-point is or whether a non-categorical model can be derived to gauge programming effects
4) Optimising genetic predictors of birthweight with larger GWAS
5) Making sure that our results are not subject to "winners curse" (i.e. bloated effect sizes)

In our opinion, the text is stretched to the limit of scientific soundness in the following sentence regarding potential clinical application: "*Identification of individuals with a birthweight phenotype-genotype mismatch could allow early identification of children at future risk.*" (Page 3, line 1-2).

Action: To clarify our position we have added to the limitations: "*For these reasons, we also believe it is premature to speculate on the clinical relevance of birthweight phenotype-genotype mismatches.*" (page 19, line 4-5)

Specific comments
P4 L4-5 The "Barker Hypothesis" has been generally replaced with the "Developmental Origins of Adult Health and Disease hypothesis". Consider rephrasing.

Response: We reference the work by Barker and Hales (ref 3) which specifically outlines the central idea that poor fetal growth leads to cardiometabolic programming – i.e. the Barker hypothesis. To our understanding the term "Developmental Origins of Adult Health and Disease hypothesis" is much broader, referring to a myriad of early programming events, not examined in

our paper. For this reason we prefer to keep the wording "Barker hypothesis".

P4 L5 Counfounding by what? Please describe.

Response: For the association between birthweight and cardiometabolic risk, we could list smoking, poor diet, poverty, genetics, race, etc. but the problem remains that there is a *potential* for unmeasured confounding. The problem is addressed in Mendelian Randomisation which is effective in handling even unidentified confounders (if the assumptions hold) and in our phenotype-genotype interaction as explained in figure 1. We state "*Potential confounding of this association has hindered causal inference*". (Page 4, Line 5)

P5 L20 Are these analyses about prediction or association? Eventually it seems that associations are being investigated, thus this type of terminology should be used throughout the manuscript.

Action: This use of "prediction" has been corrected to "association" for models of adult cardiometabolic measures, throughout. For ease of understanding we have kept the term "genetically predicted birthweight" when discussing the BW-PGS in plain terms and trust this is acceptable. We note, that the word "association" risks undermining the importance of the finding. Given the conceptual framework laid out in the directed acyclical graphs of Fig1 the associations are likely causal; therefore, we have added to the "strengths" section of the discussion: "*..and, if replicated, provide causal insight*" (Page 18, Line 5-6)

P6 L15 How accurate are the birth weight values in the hospital records? Has there been a validation study? If not, please describe the potential effects on the results.

Response: While there is no validation study for the accuracy of birth weights in the hospital record, they are measured by trained and experienced nurses on scales that are regularly checked and calibrated. In addition, any error would be random (sometimes high and sometimes low) and the effect of this would be to reduce power without inducing bias.

P6 L8 Why were three adult ages included in the analyses? Were the outcomes chosen preferentially from one age over another?

Response: In order to maximise the sample size due to different participants attending follow-up at different ages, we included the three follow-ups available at the time of analysis. This is stated in the sentence "*Every eligible participant was included to maximize the study size.*" (page 6, line 8)

The linear mixed model with clustering at participant ID level allows (in simple terms) for estimating the average effect of the BW-PGS on the cardiometabolic outcomes across the 3 ages while accounting for effects related to the age of the follow up (for instance increasing BMI with increasing age), and so we used all available information at all follow ups.

P6 L16 The sentence about age at follow up seems out of place here. I suggest moving it.

Action: This sentence has been removed.

P8 L8 How was obesity defined (e.g. which cut-off value)?

Action: The word obesity (signifying a category BMI>30) has been changed to the non-categorical "adiposity" throughout.

P10 L14 Given that the main analytic method was a linear mixed effects model, this suggests that the purpose of the analyses was not prediction but rather an investigation of associations. The language throughout the manuscript should reflect this.

Action: This use of "prediction" has been corrected to "association" for models of adult cardiometabolic measures, throughout. For ease of understanding we have kept the term "genetically predicted birthweight" when discussing the BW-PGS in plain terms and hope this is acceptable.

P12 L11 It is still unclear how data from all of these adult follow ups were used in the analyses. Please specify this back in the results section.

Response: In essence, a linear mixed model fits a line to all 3 outcomes of interest over time and models this as a function of the variable of interest, in this case the difference between genetically determined weight and actual birth weight. Where an outcome is missing, it essentially imputes it. We agree that it is perhaps easier for the reader if we repeat the essentials of the method before outlining our results.

Action: Prior to the results description we have now included the following: "*At ages 20, 22, and 27, there were 1097, 893, and 856 participants respectively, with 118 in both the largest and smallest phenotypic groupings. These participants were pooled and included in adjusted mixed effect models of repeated cardiometabolic measures from ages 20 – 27 years.*" (Page 12, line 11-13)

P13 Discussion Consider adding reflections on the clinical meaning of the effect sizes observed.

Response: We appreciate this comment; however, as we have discussed there are several necessary steps before our findings can be thought of as clinically relevant.

1) Replicating the interaction in independent datasets.
2) Determining whether gestational age adjusted birthweight is a better phenotypic trait as compared to crude birthweight
3) What the optimal cut-point is or whether a non-categorical model can be derived to gauge programming effects
4) Optimising genetic predictors of birthweight with larger GWAS
5) Making sure that our results are not subject to "winners curse" (i.e. bloated effect sizes)

In our opinion, the text is stretched to the limit of scientific soundness in the following sentence regarding potential clinical application: "*Identification of individuals with a birthweight phenotype-genotype mismatch could allow early identification of children at future risk.*" (Page 3, line 1-2).

Action: We have added to the limitations: "*For these reasons, we also believe it is premature to speculate on the clinical relevance of birthweight phenotype-genotype mismatches.*" (page 19, line 4-5)

P17 L29 This birth weight value is very specific; were the hospital records of sufficient quality to support this level of precision?

Response: The reviewer is correct and we have modified our description. Regarding the precision of the birthweight, please see our comment above.

Action: We have now written "*Bias may be introduced in these models due to co-linearity between BW-PGS and crude BW, and the cut-off at 2710 g was arbitrarily chosen based on our sample distribution of BW.; nevertheless, this analysis suggested that BW-PGS could have a programming role in neonates with a BW of 2710 or less*" (Page 17, line 17-20)

P19 L5-18 Although interesting, this paragraph does not relate to the results in this manuscript. Suggest shortening or deleting.

Response: We agree with the reviewer. This was originally a two piece article submitted simultaneously – one detailing the theoretical framework of the phenotype-genotype interaction, the other with a practical application (i.e. the results outlined in the article). We were encouraged by Nature Communications to submit just one paper to Communications Medicine, and we have done our best to synthesise the two, and make the description of the broader application less voluminous.

Reviewer comments and response

Format:

We have kept our comments in red font, with division of our **response** and **action** when relevant. Citations are kept italic and in "*brackets*" with reference to the clean manuscript page and line number.

Reviewer comments

Reviewer #1 (Remarks to the Author):

The revisions and attention to the reviewers/editors feedback have improved the readability of the manuscript. I noticed the changes to Table 3 (not table 2) where both primary and secondary outcomes are displayed. I agree that placing these findings in the main tables provides interesting data for the reader.
I agree with the authors that the Developmental origins of disease framework refers is all encompassing of low and high birth weight and risk of obesity/cardiometabolic diseases, whereas the Barker hypothesis specifically focuses on low birth and risk of metabolic disease. Thus, maintaining the Barker hypothesis terminology for this analysis is justified.

My only other comment is the use of the term adiposity- BMI is an indirect measure of adiposity, whereas waist circumference can be considered a measure of adiposity. In my read-through of this version, I do not note any issues confusing BMI with adiposity.

Response: To the best of our reading the reviewer seems satisfied with our changes in all respects, and so we leave the manuscript unaltered.