## PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

## ARTICLE DETAILS

| | |
|---|---|
| **TITLE (PROVISIONAL)** | Longitudinal study of cardio-metabolic risk from early adolescence to early adulthood in an ethnically diverse cohort |
| **AUTHORS** | Harding, Seeromanie; Silva, Maria; Molaodi, Oarabile; Enayat, Zinat; Cassidy, Aidan; Karamanos, Alexis; Read, Ursula; Cruickshank, John |

## VERSION 1 - REVIEW

| | |
|---|---|
| **REVIEWER** | Dr Yannan Jiang<br>The University of Auckland<br>New Zealand |
| **REVIEW RETURNED** | 23-Jul-2016 |

| | |
|---|---|
| **GENERAL COMMENTS** | I enjoyed reading this paper, which was well written and addressed an important research question on cardiovascular risk from early adolescence to early adulthood in different ethnic groups that could have an impact on the health service and society generally. To fully understand the study design, analysis and results, I have the following comments for the authors:<br><br>METHODS<br><br>As a longitudinal study, a total of 6643 students were recruited at the age of 11-13 years in 2002-2003. Of these, 72% participated in the first follow up at the age of 14-16 years in 2005-2006. This paper reported the results on a 10% sub-sample who took part in the pilot follow-up, with 665 participants completed the study at age 21-23 years. Although the participation rate was high (97% of invited) in an ethnically diverse cohort that was chosen to be representative by gender and socio-economic circumstances across the population, how was this sub-sample randomly chosen from the initial longitudinal cohort should be described in the study design. This information is important to evaluate the overall validity and generalisability of the results.<br><br>STATISTICAL ANALYSIS<br><br>The author stated that continuous variables were tested for normality using the Shapiro-Wilk test, was any transformation applied to those variables not normally distributed in final analysis? The sample size could be considered sufficient for central limit theorem (CLT) to apply.<br><br>Missing data are common in longitudinal studies and should be taken into account in analysis if the proportion of loss to follow up is high. The author mentioned that missing data were coded as missing and included in the analyses. How was this implemented? |

| | |
|---|---|
| | Standard statistical software normally drop all missing observations from analysis, unless a separate category is defined for Missing in categorical variables. None of the tables, however, reported this level of information.<br><br>Age is normally interpreted as a continuous measure in years. If specific age groups or bands are defined for analysis, the term needs to be used correctly. Also, BMI is not directly comparable between children and adults without standardisation. What was the rationale including adiposity measures at both 11-13y and 21-23yr in the same linear regression models for bio-markers at 21-23yr? Could change from baseline be used? Standard regression assumes independence between confounders and/or predictors, however, repeated measures over time are not. This is different from mixed-effects models with a random subject effect taking into account the correlation between repeated measures on the same subject.<br><br>RESULTS<br><br>In Table 1, sample sizes per group and overall need to be presented. Missing data should be reported (if any), and abbreviations should be explained previously (e.g. FAS). For consistency, parental diabetes at 14-16yr could report both categories.<br><br>Tables on regression models need to present all covariates fitted in the model, and consistent with those defined in Statistical Analysis. Please check the footnotes carefully. For S1, what statistics were reported in the table and what tests were used to compare which groups for those p-values?<br><br>For the results presented separately for males and females, were the analyses conducted separately for each gender? Depending on how the interaction term was added in the model, the regression estimates could be different between the sub-group analyses and the total cohort analysis with multiple covariates and their interactions. |

| | |
|---|---|
| **REVIEWER** | Linlin Li<br>Zhengzhou Univercity, China |
| **REVIEW RETURNED** | 04-Aug-2016 |

| | |
|---|---|
| **GENERAL COMMENTS** | This manuscript is a longitudinal study. Why not use COX regression? |

| | |
|---|---|
| **REVIEWER** | Markus Juonala<br>University of Turku, Finland |
| **REVIEW RETURNED** | 29-Sep-2016 |

| | |
|---|---|
| **GENERAL COMMENTS** | Specific comments:<br>1) Allostatic load should be better defined in the abstract<br>2) Is this sub-cohort representative of the total baseline population in respect of baseline risk factor and ses status?<br>3) Why is CRP threshold 0 in allostatic load?<br>4) Table 1: Abbreviations should be explained<br>5) Based on Tables 3 a and 3b this study doesn't seem to have |

| | |
|---|---|
| | enough power for ethnicity-specific analyses<br>6) Study limitations should be provided |

**VERSION 1 – AUTHOR RESPONSE**

Reviewer: 1
Reviewer Name: Dr Yannan Jiang
Institution and Country: The University of Auckland, New Zealand
Competing Interests: None declared

I enjoyed reading this paper, which was well written and addressed an important research question on cardiovascular risk from early adolescence to early adulthood in different ethnic groups that could have an impact on the health service and society generally. To fully understand the study design, analysis and results, I have the following comments for the authors:

METHODS
As a longitudinal study, a total of 6643 students were recruited at the age of 11-13 years in 2002-2003. Of these, 72% participated in the first follow up at the age of 14-16 years in 2005-2006. This paper reported the results on a 10% sub-sample who took part in the pilot follow-up, with 665 participants completed the study at age 21-23 years. Although the participation rate was high (97% of invited) in an ethnically diverse cohort that was chosen to be representative by gender and socio-economic circumstances across the population, how was this sub-sample randomly chosen from the initial longitudinal cohort should be described.
*RESPONSE: We used a stratified sampling approach. We first tried to locate the sample and, 81% (5414 of 6643) of the cohort was traced through friendship networks, social media, and community campaigns (Harding et al Lancet 2013). We then randomly selected 100 (50 per gender) in each in ethnic group, and pragmatically attempted to ensure representation across the 49 schools and across baseline socio-economic circumstances.

STATISTICAL ANALYSIS
The author stated that continuous variables were tested for normality using the Shapiro-Wilk test, was any transformation applied to those variables not normally distributed in final analysis? The sample size could be considered sufficient for central limit theorem (CLT) to apply.
*RESPONSE: All continuous various variables used were normally distributed. These were WHtr, sBP, dBP, Total cholesterol, HDL, HbA1c.
Missing data are common in longitudinal studies and should be taken into account in analysis if the proportion of loss to follow up is high. The author mentioned that missing data were coded as missing and included in the analyses. How was this implemented? None of the tables, however, reported this level of information.
*RESPONSE: Missing observations were coded as a separate category for covariates that were not continuous. We have amended Table 1 to show % missing in these variables. Missing data were excluded for continuous variables or outcomes. We have amended Table 2 to include the effect of the not stated categories. The % missing for anthropometry, sBP and dBP was<5% and for blood bio-markers 24%. We have added this information to the methods.

Age is normally interpreted as a continuous measure in years. If specific age groups or bands are defined for analysis, the term needs to be used correctly. Also, BMI is not directly comparable between children and adults without standardisation. What was the rationale including adiposity measures at both 11-13y and 21-23yr in the same linear regression models for bio-markers at 21-23yr? Could change from baseline be used? Standard regression assumes independence between confounders and/or predictors, however, repeated measures over time are not. This is different from mixed-effects models with a random subject effect taking into account the correlation between

repeated measures on the same subject.
*RESPONSE:
Re age: the age bands are tight in DASH. At baseline they were 11y-13y and at the follow-up they were 21y-23y. Hence we used age as a continuous measure in years. We have amended the text.
Re BMI: We have omitted the BMI analyses.
Re independence of repeated measures in the linear regression models: We reran the models as suggested. We were particularly interested in examining whether adiposity at 11-13y, a critical time in adolescent development, had an impact on outcomes at 21-23y. We have presented both the influence of adiposity measures at 11-13y (Tables 3a and 3b), and change from baseline (S3 and S4). As expected from the previous analyses, change in adiposity measures was significantly associated with HDL and Total Cholesterol.

RESULTS
In Table 1, sample sizes per group and overall need to be presented. Missing data should be reported (if any), and abbreviations should be explained previously (e.g. FAS). For consistency, parental diabetes at 14-16yr could report both categories.

*RESPONSE:
Re sample size: Added
Re missing data: The table is now large so we omitted parental diabetes which was not used in the analyses. Family affluence and education were also not used (as were not associated with the outcomes in the univariate analyses) but we've left these in the table to provide an overall SEC profile of the sample.
Re abbreviations: Amended.

Tables on regression models need to present all covariates fitted in the model, and consistent with those defined in Statistical Analysis. Please check the footnotes carefully. For S1, what statistics were reported in the table and what tests were used to compare which groups for those p-values?

*RESPONSE: Covariates added to Table 2, the only table with missing covariates. Footnotes checked. P-values were derived from simple linear/logistic regression models with ethnicity as an independent variable. This was added to the footnote of S1.

For the results presented separately for males and females, were the analyses conducted separately for each gender? Depending on how the interaction term was added in the model, the regression estimates could be different between the sub-group analyses and the total cohort analysis with multiple covariates and their interactions.

*RESPONSE: The gender specific results were extracted from gender stratified models. We have amended the methods. Due to the small sample size we kept interactions to a minimum, i.e. ethnicity*age and ethnicity*adiposity measure. They were not significant.

Reviewer: 2
Reviewer Name: Linlin Li
Institution and Country: Zhengzhou Univercity, China
Competing Interests: No conflict of interest

This manuscript is a longitudinal study. Why not use COX regression?
*RESPONSE: Cox regression is usually used to investigate the effect of several variables upon the time a specified event takes to happen, such as death. Here all exposures and outcomes were measured at the same time for all participants.

Reviewer: 3
Reviewer Name: Markus Juonala
Institution and Country: University of Turku, Finland
Competing Interests: None declared

Specific comments:

1) Allostatic load should be better defined in the abstract
*RESPONSE: Amended

2) Is this sub-cohort representative of the total baseline population in respect of baseline risk factor and ses status?
*RESPONSE: Please see response to reviewer 1 above. We sampled to ensure representation across by ethnicity and SEC at baseline, the key social risk factors that DASH was set up to investigate in relation to health.

3) Why is CRP threshold 0 in allostatic load?
*RESPONSE: We apologise for this error – CRP was excluded from the allostatic score as more than half of the participants had a level of 0.

4) Table 1: Abbreviations should be explained
*RESPONSE: Amended – in full at first mention.

5) Based on Tables 3 a and 3b this study doesn't seem to have enough power for ethnicity-specific analyses
*RESPONSE: This is possible as we did not set out to conduct a pilot follow-up study powered to investigate ethnic differences in these outcomes. The following has been added to the text. 'The small sample size of the ethnic groups prohibited robust testing of ethnic specific effects. The primary aims of the pilot follow-up were to locate the diverse groups in the cohort, investigate whether they would take participate in a subsequent follow-up study and agree for their parents to be invited to join DASH, and whether they would consent to the different measures, notably the bio-markers.Despite these small numbers, however, these findings in a small age range provide a robust platform for planning future studies. '
6) Study limitations should be provided
*RESPONSE: See response above. We expanded this section in the discussion and added to bullet points in 'Strengths and limitations' after abstract.

**VERSION 2 – REVIEW**

| | |
|---|---|
| **REVIEWER** | Markus Juonala<br>University of Turku, Finland |
| **REVIEW RETURNED** | 09-Nov-2016 |

| | |
|---|---|
| **GENERAL COMMENTS** | The reviewer completed the checklist but made no further comments. |