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)	Associations Between Birth Weight, Obesity, Fat Mass and Lean Mass in Korean Adolescents: The Fifth Korea National Health and Nutrition Examination Survey
AUTHORS	Kang, Myunggee; Yoo, Jung Eun; Kim, Kyuwoong; Choi, Seulggie; Park, Sang Min

VERSION 1 – REVIEW

REVIEWER	Brian Stansfield Augusta University Augusta, GA USA
REVIEW RETURNED	07-Jun-2017

GENERAL COMMENTS	I appreciate the opportunity to review this manuscript by Kang et al. The objective of this manuscript is to describe the association between birth weight and adolescent overweight/obesity and fat distribution. The authors' use a national registry to accomplish this objective. The design and statistical analysis are appropriate although alternative approaches using non-categorized continuous data would have also been appropriate and preferred by some readers/reviewers. Overall, the manuscript supports the well-recognized phenomenon that birth weight correlates directly with adolescent/adult obesity. The authors are correct in their statement that asians are under-represented in these population studies, which adds novelty to their findings as does the assessment of lean and fat mass distribution. I have a few comments and suggestions for the authors: 1. The methods section requires more detail about the KHANS. How often is the survey administered and who are the recipients of the survey. Is it entirely self-reported? It appears to be linked to health records as the authors were able to connect DXA scans with the KHANES data. If this is correct, the authors should have some information about the oversight/review of the study by a professional review board. Do the authors require consent if the study data is linked? 2. The authors approach co-morbidities in a practical way, but seem to account for some confounders that are not well-established in the literature including sleep duration and gestational age. Also, the authors do not account for current BMI in their analysis. The over-reliance on confounders may create a reversal paradox (Cole TJ Am J Epidemiol. 2005).

	This occurs when a controlled variable is in the causal pathway between the originating event and the outcome of interest, and leads to a false or exaggerated inverse relation between the two. We have accounted for this (Stansfield et al JPeds 2016) by using 2 model systems where we adjust for a group of confounders in model system 1 and then adjust for confounders + current BMI in model system 2. The change in significance is thought to be attributable to current BMI, which is closely linked to postnatal weight gain. 3. This brings up my final concern. There is much recent debate over whether birth weight or postnatal weight gain velocity are equal/disproportionate contributors to fat mass in later life. This may explain some of the J- or U-shaped relationships between birth weight and fat mass in later life. One theory, infants who are born small experience rapid catch up growth in early infancy, which sets up a conservation phenotype leading to later insulin resistance and metabolic syndrome. I would encourage the authors to consider this theory as they explain the linear relationship between birth weight and FMI and LMI as their results run counter to this hypothesis. Thank you for the opportunity to review your manuscript.

REVIEWER	Leilani Muhardi Danone Early Life Nutrition, Indonesia I am a full-time employee of Danone Early Life Nutrition, Indonesia.
REVIEW RETURNED	27-Jun-2017

GENERAL COMMENTS	The current statistical approach will not be sufficient to justify the associations between birth weight and obesity, fat/ lean mass during adolescents as suggested by the title. The language of the manuscript needs to be further improved.

REVIEWER	Rosie Cornish University of Bristol, UK
REVIEW RETURNED	12-Jul-2017

GENERAL COMMENTS	It is clearly important to establish whether previous findings relating birth weight to later obesity apply in different populations (e.g. as in this study). That said, I have some concerns about the analysis which need to be addressed. In addition, I feel that the authors should refer to a wider body of evidence when interpreting their findings. Detailed comments are given below. Main comments 1. The authors state in the discussion and in the section entitled Strengths and limitations of this study that previous studies have been small (“hardly exceed 100 people”). This is not actually correct. Doing a very quick search I found the following: (i) Monteiro POA et al. Birth size, early childhood growth, and adolescent obesity in a Brazilian birth cohort. International Journal of Obesity (2003) 27, 1274–1282. doi:10.1038/sj.ijo.0802409. (This study included over 1000 adolescents) (ii) Rossi et al. Birth weight and obesity in children and adolescents: a systematic review. Revista Brasileira de Epidemiologia (2010) 13(2) http://dx.doi.org/10.1590/S1415-790X2010000200007

	(This review included a number of studies of between 150 and almost 15,000 children and/or adolescents) (iii) Evensen E et al. The relation between birthweight, childhood body mass index, and overweight and obesity in late adolescence: a longitudinal cohort study from Norway, The Tromsø Study, Fit Futures BMJ Open (2017) 7:e015576. doi: 10.1136/bmjopen-2016-015576. (Just under 1000 adolescents) I would suggest the authors remove this statement. I think it would also be beneficial for them to do a more comprehensive literature search so that they can discuss their findings in the context of the current available evidence in this area. 2. As it stands, the results presented in Table 2 are quite difficult to interpret because of the small numbers (and consequently wide confidence intervals). I am not sure that dividing birth weight into quintiles is the best way to analyse the association between birth weight and overwight/obesity. A more powerful analysis would use birth weight as a continuous variable. Alternatively, to increase power a little, birth weight could be divided into 0-25th percentile, 25-75th percentile and 75-100th percentile, with the central group as the reference category. It might also be worth looking at BMI as a continuous variable and displaying the results in the same way as for FMI and LMI. This would provide a better sense of whether or not there is a trend of increasing BMI with increasing birthweight. 3. The authors adjust for smoking, drinking, physical activity, energy consumption, sleep duration, and co-morbidities potentially associated with obesity. Although these are all potentially associated with BMI/obesity, they are not causes of birth weight and therefore cannot confound the association between birth weight and BMI/obesity in adolescence. Further, it is possible that some of these variables are on the causal pathway (for example, if the hypothesis is that birth weight results in changes in appetite control then energy consumption could be on the causal pathway; in addition, birth weight has been shown to be related to glucose metabolism and blood pressure, so hypertension, diabetes, impaired fasting glucose could be on the causal pathway), in which case you definitely would not want to adjust for these factors. In summary, since it is not actually necessary to adjust for these factors (since they cannot be confounders) and it could actually be incorrect to do so, I think they should be omitted from the analysis. 4. Relating to the above, the authors give – as a strength of their study – the fact that they have adjusted for more variables than other studies. However, as mentioned above, they have included variables that should not be adjusted for but have not included key maternal factors (maternal BMI, parity, smoking during pregnancy) that could be important confounders. Although the authors mention maternal factors in the discussion, they do not explicitly acknowledge that failure to adjust for these could mean that there may be residual confounding and that this is a limitation of their study. 5. The authors have done separate tests to compare the highest and lowest birth weight quintiles to the reference group. This is not appropriate. The test for trend, together with the graphs showing the means and confidence intervals is the appropriate way of analysing these data.

	Alternatively, they could test – using a likelihood ratio test – whether birth weight overall was associated with FMI/LMI (i.e. not testing specifically for a linear trend). 6. The authors present results separately for males and females and state that there were sex differences in the analyses. However, was there actually any statistical evidence for this – i.e. was there an interaction between birth weight and sex in terms of the effect on the outcomes considered? To me, the graphs for FMI and LMI look very similar for males and females and, as stated above, the numbers for obesity and overweight are quite small and therefore these particular results are quite difficult to draw any conclusions from. I think it is appropriate to show results separately for males and females but if there is no evidence for effect modification by sex then it would be better to do the formal analysis for everyone combined to increase precision/power. Minor comments 1. The authors suggest that the fact that their data are cross-sectional means that causal inferences cannot be made and that longitudinal data are needed to clarify this. It is true that it can be difficult to make inferences about the direction of causality using cross-sectional data. However, in this particularly scenario, where the outcome (body composition / obesity etc. in adolescence) could not possibly have occurred prior to the exposure (birth weight) then the fact that the data were collected cross-sectionally makes no difference with respect to this particular issue. Obviously the fact that birth weight data were collected retrospectively (from maternal recall) means that this variable could be subject to greater error than if it had been collected prospectively. However, if you have longitudinal data it is still difficult to infer causality because you can never avoid the issue of residual confounding.

REVIEWER	Jason Oke Nuffield Department of Primary Care Health Sciences, University of Oxford
REVIEW RETURNED	13-Jul-2017

GENERAL COMMENTS	1. A complete case analysis is insufficient given the ease in which MI methods can be done using mainstream statistical software. If the reasons for the missing data are unrelated to the outcome then complete cases will be ok but I don't know if is this is a defendable assumption or not? So I would think a sensitivity analyses should be done using multiple imputation in order to include as many people as possible. This can then be used to assess the potential for bias. There is another benefit in that more powerful analyses (compared with analyses based on complete cases) can generally be done by including individuals with incomplete data. 2. Multiple comparisons The authors report a number (12 or 24) of different pairwise comparisons and report confidence intervals rather than p values. Confidence intervals that exclude an OR of 1 are being interpreted as significant. The number of comparisons means that the family wise family error rate is not 0.05 as would be believed but is inflated because of the multiple comparisons and thus the chance of a type 1 error much higher than 5%.

	This applies even when the comparisons are ad-hoc or a-priori. This needs to be addressed. Simultaneous confidence intervals that account for multiple comparisons should be used. Some minor points Use of ADA criteria for FPG – as opposed the WHO criteria should be made clear. Unsure of the validity of using term hypertension for describing the top decile (90%-100%). Suggest re-wording Wasn't clear on the term "programming of body composition" is this a technical term?

VERSION 1 – AUTHOR RESPONSE

Response to comments from the Reviewer 1:

Comment 1:
The methods section requires more detail about the KHANS. How often is the survey administered and who are the recipients of the survey. Is it entirely self-reported? It appears to be linked to health records as the authors were able to connect DXA scans with the KHANES data. If this is correct, the authors should have some information about the oversight/review of the study by a professional review board. Do the authors require consent if the study data is linked?

→ Per the reviewer's comment, we added more detail about the KNHANES in the section as following:
KNHANES is a national survey conducted every year by the Ministry of Health and Welfare. KNHANES V was performed from 2010 to 2011, and included 17,476 nationally representative non-institutionalized civilians in Korea. KNHANES V contains data on heath interviews, health behaviors, health examination, nutrition and large-scale whole-body dual energy X-ray absorptiometry (DXA). Target sampling follows a stratified multi-stage clustered probability design. More details on KNHANES have been published previously[17], and the original data are publicly available through the KNHANES website.
Also, we mentioned whether the data were self-reported as following:
Sociodemographic variables including age, sex, residential area, and household income were self-reported by the participants.
Birth weight was reported by the participants' mothers in KNHANES V.
In KNHANES V, body compositions were measured by skilled technicians, providing accurate and reliable data. All participants changed into light clothing with all jewelry removed, and then underwent DXA examinations.

Comment 2:
The authors approach co-morbidities in a practical way, but seem to account for some confounders that are not well-established in the literature including sleep duration and gestational age. Also, the authors do not account for current BMI in their analysis. The over-reliance on confounders may create a reversal paradox (Cole TJ Am J Epidemiol. 2005). This occurs when a controlled variable is in the causal pathway between the originating event and the outcome of interest, and leads to a false or exaggerated inverse relation between the two. We have accounted for this (Stansfield et al JPeds 2016) by using 2 model systems where we adjust for a group of confounders in model system 1 and then adjust for confounders + current BMI in model system 2. The change in significance is thought to be attributable to current BMI, which is closely linked to postnatal weight gain.

→ We agree that some confounders are not well-established in the literature and omitted those confounders from our analysis. We also agree that when examining the effect of birth weight, adjusting for current BMI could be meaningful. However, fat mass index and lean mass index, the dependent variables of our research, have strong collinearity with BMI. We thought adjusting for current BMI could result in residual confounding, and therefore could not account for it. Instead, we added in our manuscript that postnatal weight gain is another important factor in body composition and that further research is recommended. Details are appended with the next comment.

Comment 3:
This brings up my final concern. There is much recent debate over whether birth weight or postnatal weight gain velocity are equal/disproportionate contributors to fat mass in later life. This may explain some of the J- or U-shaped relationships between birth weight and fat mass in later life. One theory, infants who are born small experience rapid catch up growth in early infancy, which sets up a conservation phenotype leading to later insulin resistance and metabolic syndrome. I would encourage the authors to consider this theory as they explain the linear relationship between birth weight and FMI and LMI as their results run counter to this hypothesis.

→ We agree with Reviewer 1's comment. We are aware of the previous studies that have shown J- or U-shaped relationships between birth weight and fat mass in later life, and that the result of our research is different from them. If we had the data of postnatal weight gain, we could have made the reasons for the difference clearer, but we did not have such data. We thought that one possible explanation could be ethnic differences as we mentioned in our manuscript.
We agree that this issue is important and needs to be considered. Accordingly, we added this in the section as following:
The result of linear relationship between birth weight and later FMI contradicts some of the previous studies that have shown the J- or U-shaped relationships between birth weight and fat mass in later life. Those studies explain their results by the theory that infants who are born small experience rapid catch up growth in early infancy, which results in larger fat mass in later life[39,40]. We could not reveal clear reasons for the difference as we did not have the data of postnatal weight gain, but one possible explanation for this could be ethnic differences as mentioned above. Regarding the possible effects of postnatal weight gain on later body composition, future studies can be performed to investigate the interaction of birth weight and postnatal weight gain in terms of the effect on later body composition. Through these future studies, whether birth weight or postnatal weight gain velocity are equal or disproportionate contributors to fat mass in later life might be revealed.

Response to comments from the Reviewer 2:

Comment 1:
The current statistical approach will not be sufficient to justify the associations between birth weight and obesity, fat/ lean mass during adolescents as suggested by the title.
The language of the manuscript needs to be further improved.

→ Thank you for your comments. In the revised manuscript, we improved our statistical method in several ways: First, we re-categorized birth weight group, which was divided into five groups previously, into three groups (0-25th percentile, 25-75th percentile and 75-100th percentile), and achieved more statistically powerful results. Second, instead of doing separate tests to compare the highest and lowest birth weight quantiles to the reference group as we did previously, we only left and presented the test for trend, together with the graphs showing the means and confidence intervals (Figure 2, 3 and 4). Third, we performed a sensitivity analysis using MI methods. By this way, we tried to overcome the possibility that missing DXA data are nonrandom. After including individuals without DXA data by MI, the result had the same tendency with the complete case analysis, and we added this content to our method and discussion part.

Furthermore, we improved the quality of the English throughout the manuscript.
We hope that our efforts carried out better results, and if problems still exist with our statistical approach, please let us know.

Response to comments from the Reviewer 3:

Comment 1:
The authors state in the discussion and in the section entitled Strengths and limitations of this study that previous studies have been small ("hardly exceed 100 people"). This is not actually correct. Doing a very quick search I found the following:
(i) Monteiro POA et al. Birth size, early childhood growth, and adolescent obesity in a Brazilian birth cohort. International Journal of Obesity (2003) 27, 1274–1282. doi:10.1038/sj.ijo.0802409.
(This study included over 1000 adolescents)
(ii) Rossi et al. Birth weight and obesity in children and adolescents: a systematic review. Revista Brasileira de Epidemiologia (2010) 13(2) http://dx.doi.org/10.1590/S1415-790X2010000200007
(This review included a number of studies of between 150 and almost 15,000 children and/or adolescents)
(iii) Evensen E et al. The relation between birthweight, childhood body mass index, and overweight and obesity in late adolescence: a longitudinal cohort study from Norway, The Tromsø Study, Fit Futures BMJ Open (2017) 7:e015576. doi: 10.1136/bmjopen-2016-015576.
(Just under 1000 adolescents)

I would suggest the authors remove this statement. I think it would also be beneficial for them to do a more comprehensive literature search so that they can discuss their findings in the context of the current available evidence in this area.

→ Thank you for your comments. Although there are some studies with large population examining the relationship between birth weight and 'obesity,' not many exist that targeted birth weight and 'body composition.' That was our intention of saying that "previous studies hardly exceed 100 people." However, we agree that our expression was unclear and could be misunderstood. We omitted the part as suggested by the reviewer and corrected the manuscript. We also did a more comprehensive literature search, including the articles the reviewer suggested, to improve the quality of the research. The revised part is as the following:

The study population of this study consisted of Koreans, an ethnicity few previous studies targeted, and was relatively large (1,304 people), raising the credibility of this study.

Despite these limitations, we used a large study population (n=1,304) compared to previous studies, which increases credibility of this study.

2. As it stands, the results presented in Table 2 are quite difficult to interpret because of the small numbers (and consequently wide confidence intervals). I am not sure that dividing birth weight into quintiles is the best way to analyse the association between birth weight and overwight/obesity. A more powerful analysis would use birth weight as a continuous variable. Alternatively, to increase power a little, birth weight could be divided into 0-25th percentile, 25-75th percentile and 75-100th percentile, with the central group as the reference category. It might also be worth looking at BMI as a continuous variable and displaying the results in the same way as for FMI and LMI. This would provide a better sense of whether or not there is a trend of increasing BMI with increasing birthweight.

→ We agree with Reviewer 3's comment and reanalyzed accordingly. We divided birth weight into three groups (0-25th percentile, 25-75th percentile and 75-100th percentile), and achieved more statistically powerful results. Revised tables are shown in Table 2, Figure 2, 3 and 4.

We also looked at BMI as a continuous variable as the suggestion. There was an increasing trend of BMI with increasing birth weight (P for trend: 0.002). We added this content in the part of the manuscript as the following:
BMI of adolescents tended to increase linearly with increasing birth weight (P for trend: 0.002).

3. The authors adjust for smoking, drinking, physical activity, energy consumption, sleep duration, and co-morbidities potentially associated with obesity. Although these are all potentially associated with BMI/obesity, they are not causes of birth weight and therefore cannot confound the association between birth weight and BMI/obesity in adolescence. Further, it is possible that some of these variables are on the causal pathway (for example, if the hypothesis is that birth weight results in changes in appetite control then energy consumption could be on the causal pathway; in addition, birth weight has been shown to be related to glucose metabolism and blood pressure, so hypertension, diabetes, impaired fasting glucose could be on the causal pathway), in which case you definitely would not want to adjust for these factors. In summary, since it is not actually necessary to adjust for these factors (since they cannot be confounders) and it could actually be incorrect to do so, I think they should be omitted from the analysis.

→ We agree with Reviewer 3's comment and omitted those confounders (smoking, drinking, physical activity, energy consumption, sleep duration and comorbidities) from the analysis. The confounders we left in the revised version are sex, age, residence and household income.

4. Relating to the above, the authors give – as a strength of their study – the fact that they have adjusted for more variables than other studies. However, as mentioned above, they have included variables that should not be adjusted for but have not included key maternal factors (maternal BMI, parity, smoking during pregnancy) that could be important confounders. Although the authors mention maternal factors in the discussion, they do not explicitly acknowledge that failure to adjust for these could mean that there may be residual confounding and that this is a limitation of their study.

→ We agree with Reviewer 3's comment and added our limitations in the part as the following:
Although parental factors including maternal obesity, diabetes may influence the birth weight of their offspring, such information are not included in this survey. It is a limitation of this study that failure to adjust for these could have resulted in residual confounding.

5. The authors have done separate tests to compare the highest and lowest birth weight quintiles to the reference group. This is not appropriate. The test for trend, together with the graphs showing the means and confidence intervals is the appropriate way of analysing these data. Alternatively, they could test – using a likelihood ratio test – whether birth weight overall was associated with FMI/LMI (i.e. not testing specifically for a linear trend).

→ We agree with Reviewer 3's comment and omitted the relevant parts. Instead, we did the test for trend and presented the graphs showing the means and confidence intervals as the reviewer suggested. The revised graphs are included in Figure 2, 3 and 4.

6. The authors present results separately for males and females and state that there were sex differences in the analyses. However, was there actually any statistical evidence for this – i.e. was there an interaction between birth weight and sex in terms of the effect on the outcomes considered? To me, the graphs for FMI and LMI look very similar for males and females and, as stated above, the numbers for obesity and overweight are quite small and therefore these particular results are quite difficult to draw any conclusions from. I think it is appropriate to show results separately for males and females but if there is no evidence for effect modification by sex then it would be better to do the formal analysis for everyone combined to increase precision/power.

→ Thank you for your comments. It is true that we could not give statistical evidence for the interaction between birth weight and sex in terms of the effect on the outcome. However, as the data of men only, and of women only have their own value independently, we thought that they could be analyzed and interpreted separately. Also, we found out by chi square test that the distribution of input variable (birth weight) and output variables (obesity and body composition) were very different among sex, which made us think that sex stratified analysis is necessary. And as our result was statistically significant in women but not men, possible explanation for this had to be suggested and that was why we explained sex differences.
However, we agree that interpreting them independently without statistically explaining their interaction could be controversial. In our revised manuscript, we changed our expressions and tried not to emphasize sex difference as the previous one, but emphasized the formal analysis for everyone as suggested by the reviewer. Changes can be found throughout the result and discussion section.

Minor comments

1. The authors suggest that the fact that their data are cross-sectional means that causal inferences cannot be made and that longitudinal data are needed to clarify this. It is true that it can be difficult to make inferences about the direction of causality using cross-sectional data. However, in this particularly scenario, where the outcome (body composition / obesity etc. in adolescence) could not possibly have occurred prior to the exposure (birth weight) then the fact that the data were collected cross-sectionally makes no difference with respect to this particular issue. Obviously the fact that birth weight data were collected retrospectively (from maternal recall) means that this variable could be subject to greater error than if it had been collected prospectively. However, if you have longitudinal data it is still difficult to infer causality because you can never avoid the issue of residual confounding.

→ We agree with Reviewer 3's comment and clarified the characteristics and limitations of our study in the section. The revised parts are as the following:

There are several limitations in this study. First, since cross-sectional data was used, inferences about causal effects cannot be made clearly and only assumptions about the relationship can be suggested. Although the fact that the outcome, obesity and body composition, could not possibly have occurred prior to the exposure, birth weight, helps us infer the direction of causality, further studies are needed to clarify the causality and its mechanism.

Response to comments from the Reviewer 4:

Comment 1:
A complete case analysis is insufficient given the ease in which MI methods can be done using mainstream statistical software. If the reasons for the missing data are unrelated to the outcome then complete cases will be ok but I don't know if is this is a defendable assumption or not? So I would think a sensitivity analyses should be done using multiple imputation in order to include as many people as possible. This can then be used to assess the potential for bias. There is another benefit in that more powerful analyses (compared with analyses based on complete cases) can generally be done by including individuals with incomplete data.

→ Authors appreciated the reviewer's suggestions and did an additional analysis with MI methods. As the reviewer commented, there is a possibility that missing DXA data are nonrandom and therefore sensitivity analysis is needed in this study. After including individuals without DXA data by MI, the result had the same tendency with the complete case analysis, and we added this content to our method and discussion part as the following:

Additionally, we ran a sensitivity analysis imputing the missing data with an expectation-maximization algorithm. By the use of this approach, we could avoid using biased population estimates, as there is a possibility that missing DXA data are nonrandom. The linear trend of adjusted means for BMI, fat mass and lean mass was tested after imputation.

We evaluated that even after imputing the missing data, the overall trend of positive correlation between birth weight and FMI did not change. In men, women and total population, FMI increased significantly with increasing birth weight (P for trend <0.001 for all). However, LMI increased with increasing birth weight only in men population (P for trend: 0.01).

Comment 2:
Multiple comparisons
The authors report a number (12 or 24) of different pairwise comparisons and report confidence intervals rather than p values. Confidence intervals that exclude an OR of 1 are being interpreted as significant. The number of comparisons means that the family wise family error rate is not 0.05 as would be believed but is inflated because of the multiple comparisons and thus the chance of a type 1 error much higher than 5%. This applies even when the comparisons are ad-hoc or a-priori. This needs to be addressed. Simultaneous confidence intervals that account for multiple comparisons should be used.

→ Per the reviewer's comment, we believe that our result table was presented unclearly and was misunderstood. We have not done such a lot (12 or 24) pairwise comparisons as the reviewer said. Table 2 presents 12 independent regression models we have tested, and as we reduced our exposure variable (birth weight) categories from 5 to 3 in our revised manuscript, at most two pairwise comparisons are made in each regression models. Although this may still have a problem of multiple comparisons, we believe that it would not have caused serious inflation of type 1 error. If it is our misunderstanding, please let us know.
In addition, we changed the format (order of subgroup categories) of the result table (Table 2) to improve readability.

minor points
Use of ADA criteria for FPG – as opposed the WHO criteria should be made clear.

Unsure of the validity of using term hypertension for describing the top decile (90%-100%). Suggest re-wording

Wasn't clear on the term "programming of body composition" is this a technical term?

→ We omitted comorbidities (including FPG and hypertension) from the confounders in our revised analysis according to the other reviewer's comment.
We agree that the term "programming of body composition" was unclear, and substituted this term with other expressions.

VERSION 2 – REVIEW

REVIEWER	Brian Stansfield Augusta University Augusta, GA USA
REVIEW RETURNED	04-Sep-2017

GENERAL COMMENTS	The authors have adequately answered my concerns and have revised the manuscript to meet the major concerns.

REVIEWER	Leilani Muhardi Danone Nutricia Early Life Nutrition Indonesia Full time employee of Danone Nutricia Early Life Nutrition Indonesia.
REVIEW RETURNED	17-Sep-2017

GENERAL COMMENTS	The findings of the study are of value to the current body of literature. It has large number of study participants, long follow-up time and add more insights on Asians. However, the coherence of the narrative doesn't flow well for the readers. Thus a major rework to link ideas and findings; as well as to improve clarity of the message need to be undertaken. Further clarifications are needed for example for the following statements in the abstract: Page 2- line 27 the reason to use of Multivariable logistic regression and Least Square means if the statistical method is to be incorporated in the abstract. Ln 37- clarity on the comparison group for the statement " women in the highest 25th percentile of birth weight is more likely to be obese". Ln 40 - with p trend of 0.03, could we state that the FMI significantly increased with increasing birth weight? Table 2: Please explain why adjusted OR is made with adjustment to sex while the data is already presented separately between men and women. In the conclusion (page 13), could we state that in women, high birth weight is associated with general obesity based on available data? As it is an observational cohort and no post-natal information was obtained, perhaps we can only relate the high birth weight among women with higher risk of general obesity?

REVIEWER	Rosie Cornish University of Bristol, UK
REVIEW RETURNED	14-Sep-2017

GENERAL COMMENTS	The authors have addressed most of my previous concerns. I have some further comments relating to two of my previous points, plus four additional comments. 1. I previously said that the fact that the data are cross-sectional is not really the issue in this particular study. I do not think the new wording is really clear in this regard. I would suggest removing the sentence that has been added and simply replacing the term “cross-sectional” with “observational”. I don’t think you need to talk about the direction of causality since this will be obvious to any reader. 2. I previously suggested displaying the results for BMI in the same way as for FMI and LMI. I think it is quite important to include a figure for BMI (perhaps separately for males and females) as per the figures given for FMI and LMI. This can be included in addition to Table 2 and will make it much clearer for the reader in terms of assessing whether there is an association between birthweight and BMI. As it stands, it is not easy to judge this from 12 different odds ratios and their confidence intervals. 3. The crude odds ratios in Table 2 are not those obtained from the numbers given in the table. For example, the proportion of females overweight in the 75-100th percentile group for birthweight is higher than the proportion overweight in the 25-75th percentile group, yet the odds ratio is less than one. This occurs in several places in the table. Are these errors or is it because sampling weights have been applied? If it is the latter then this needs to be explained in the statistical methods section of the text. Also if sampling weights have been used to calculate the odds ratios, then the weights should also be applied when calculating the proportions overweight and obese, otherwise these results look very strange. 4. On a related note, please could the authors clarify that when calculating odds ratios for overweight, obese individuals have been omitted from the analysis (i.e. they are not included in the denominator)? 5. The authors state in the abstract and the results that an association between birthweight and fat mass index was seen in females but not males. However, the figures suggest the reverse. Have the figures been labelled incorrectly? 6. It is not sufficient to simply report the fact that multiple imputation was used and that the results were similar to those obtained from the complete case analysis. At the very least, I would expect there to be a table showing characteristics of those with and without missing data to show whether those with missing (i) birthweight and (ii) DXA data were different in terms of key characteristics compared to those who contributed to the analysis. In particular, it would be useful to know if those with missing DXA data were different in terms of their BMI. Secondly, further details should be given regarding the imputation models – what variables were included, how many datasets were imputed, and so on. The authors may want to refer to the following paper (or others) to get an idea of what they should be reporting: Rezvan PH et al. The rise of multiple imputation: a review of the reporting and implementation of the method in medical research. BMC Medical Research Methodology 2015; 15:30. https://doi.org/10.1186/s12874-015-0022-1.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Brian Stansfield
Institution and Country: Augusta University, Augusta, GA, USA
Please state any competing interests: None declared

Please leave your comments for the authors below
The authors have adequately answered my concerns and have revised the manuscript to meet the major concerns.

Reviewer: 2
Reviewer Name: Leilani Muhardi
Institution and Country: Danone Nutricia Early Life Nutrition Indonesia
Please state any competing interests: Full time employee of Danone Nutricia Early Life Nutrition Indonesia

Please leave your comments for the authors below
The findings of the study are of value to the current body of literature. It has large number of study participants, long follow-up time and add more insights on Asians. However, the coherence of the narrative doesn't flow well for the readers. Thus a major rework to link ideas and findings; as well as to improve clarity of the message need to be undertaken.

 Our study was not a follow-up study, but instead we used a cross-sectional data for the analyses. (Birth Weight was self-reported in the survey.)
We agree with the reviewer's comment that the improvement in clarity was needed. We thoroughly reviewed our manuscript and improved the coherence and the quality of the English. Changes can be found throughout the manuscript in the Marked Copy version.

Further clarifications are needed for example for the following statements in the abstract:

Page 2- line 27
the reason to use of Multivariable logistic regression and Least Square means if the statistical method is to be incorporated in the abstract.

 We revised the abstract by describing what were calculated by those statistical methods. The changes are as following:
Adjusted odds ratio (aORs) and 95% confidence intervals (CIs) were calculated by multivariable logistic regression analyses to determine the association between birth weight and being overweight or obese. Furthermore, adjusted mean values for body mass index (BMI), fat mass index (FMI), and lean mass index (LMI) according to birth weight were calculated by multiple linear regression analyses.

Ln 37- clarity on the comparison group for the statement " women in the highest 25th percentile of birth weight is more likely to be obese".

□ We specified the reference group in the abstract as following:
Individuals within the highest 25th percentile in birth weight were more likely to be overweight (aOR=1.64, 95% CI=1.05-2.57) compared to adolescents within the 25th and 75th percentile in birth weight. Female adolescents who were in the highest 25th percentile in birth weight were more likely to be obese (aOR=2.15, 95% CI=1.06-4.37) compared to those within the 25th and 75th percentile in birth weight.

Ln 40 - with p trend of 0.03, could we state that the FMI significantly increased with increasing birth weight?

□ We considered P value under 0.05 as statistically significant. As this had not been clear in the previous manuscript, we added it in the 'Methods' section as following:
A P value of less than 0.05 was considered statistically significant.

Table 2:
Please explain why adjusted OR is made with adjustment to sex while the data is already presented separately between men and women.

□ We performed analysis with the total population to see the generalisability of the study, and adjusted for sex, which is an important factor regarding obesity. And then we additionally performed sex-stratified analysis to see the relationships in each sex.

In the conclusion (page 13), could we state that in women, high birth weight is associated with general obesity based on available data? As it is an observational cohort and no post-natal information was obtained, perhaps we can only relate the high birth weight among women with higher risk of general obesity?

□ Thank you for your comments. We agree that inferences about causal effects cannot be made clearly in this study, and that our expressions could be misleading. We corrected the explanation as your suggestion as following:
Particularly in women, high birth weight is associated with higher risk of general obesity.

Reviewer: 3
Reviewer Name: Rosie Cornish
Institution and Country: University of Bristol, UK
Please state any competing interests: None declared

Please leave your comments for the authors below
The authors have addressed most of my previous concerns. I have some further comments relating to two of my previous points, plus four additional comments.

1. I previously said that the fact that the data are cross-sectional is not really the issue in this particular study. I do not think the new wording is really clear in this regard. I would suggest removing the sentence that has been added and simply replacing the term "cross-sectional" with "observational". I don't think you need to talk about the direction of causality since this will be obvious to any reader.

□ We agree with the reviewer's comment. We removed the sentence about the direction of causality, and replaced the term "cross-sectional" with "observational" in the 'Discussion' section.

2. I previously suggested displaying the results for BMI in the same way as for FMI and LMI. I think it is quite important to include a figure for BMI (perhaps separately for males and females) as per the figures given for FMI and LMI. This can be included in addition to Table 2 and will make it much clearer for the reader in terms of assessing whether there is an association between birthweight and BMI. As it stands, it is not easy to judge this from 12 different odds ratios and their confidence intervals.

□ Per the reviewer's comment, we added figures that present relationship between birth weight and BMI in total population and also males and females separately. They are shown in figure 2 and in the 'Results' section as following:
BMI of adolescents tended to increase linearly with increasing birth weight in total participants, men and women (P for trend: <0.01, 0.02, and 0.05 respectively) as presented in Figure 2

3. The crude odds ratios in Table 2 are not those obtained from the numbers given in the table. For example, the proportion of females overweight in the 75-100th percentile group for birthweight is higher than the proportion overweight in the 25-75th percentile group, yet the odds ratio is less than one. This occurs in several places in the table. Are these errors or is it because sampling weights have been applied? If it is the latter then this needs to be explained in the statistical methods section of the text. Also if sampling weights have been used to calculate the odds ratios, then the weights should also be applied when calculating the proportions overweight and obese, otherwise these results look very strange.

□ Thank you for the comment. As the reviewer said, the problem was because sampling weights have been applied when analysing the odds ratio. We agree that the mismatching between the proportion and the odds ratio could look strange, so we added the proportions of overweight and obese after applying sampling weights in table 3.

4. On a related note, please could the authors clarify that when calculating odds ratios for overweight, obese individuals have been omitted from the analysis (i.e. they are not included in the denominator)?

□ Obese individuals are not omitted from the analysis of overweight, and are included in the denominator in our study. We intended to see the relation of birth weight and overweight in total population, despite it includes those who have higher BMI then the criteria of being overweight.

5. The authors state in the abstract and the results that an association between birthweight and fat mass index was seen in females but not males. However, the figures suggest the reverse. Have the figures been labelled incorrectly?

□ Thank you for the comments. There were some problems in the uploading process, and as you said the ordering and labelling of the figures were twisted. Sorry for the confusion, and we have double-checked the matching of tables and figures with our manuscripts in the revised version.

6. It is not sufficient to simply report the fact that multiple imputation was used and that the results were similar to those obtained from the complete case analysis. At the very least, I would expect there to be a table showing characteristics of those with and without missing data to show whether those with missing (i) birthweight and (ii) DXA data were different in terms of key characteristics compared to those who contributed to the analysis. In particular, it would be useful to know if those with missing DXA data were different in terms of their BMI. Secondly, further details should be given regarding the imputation models – what variables were included, how many datasets were imputed, and so on. The authors may want to refer to the following paper (or others) to get an idea of what they should be reporting: Rezvan PH et al.

The rise of multiple imputation: a review of the reporting and implementation of the method in medical research. BMC Medical Research Methodology 2015; 15:30. https://doi.org/10.1186/s12874-015-0022-1.

 We agree with the reviewer's comment. We have thoroughly reviewed the paper the reviewer referred to and added more information on the multiple imputation. We added the table showing the comparison between those with and without missing data in table 2, and added more details on the imputation models as following:

'Methods'
Additionally, we ran a sensitivity analysis imputing the missing data as there is a possibility that missing data are nonrandom. Five imputations for missing data were performed with the use of the Markov chain Monte Carlo multiple imputation technique. Variables adjusted for the primary analysis were included for imputation. The pooled effect on the primary outcome was analysed after multiple imputation. A P value of less than 0.05 was considered statistically significant.

'Results'
Table 2 describes characteristics of three groups: those with complete data (n=884), those with missing values on birth weight or BMI (n=206), and those with missing values on DXA (n=420). There were no differences in the distribution of characteristics, including birth weight, BMI, FMI and LMI among the three groups. However, those without birth weight data had higher percentage of those living in the capital area, and being in the lowest tertile of household income compared to those with complete data. Furthermore, both of the distribution of area of residence and household income differed significantly from the complete case (P value <0.01 for both area of residence and household income).

VERSION 3 – REVIEW

REVIEWER	Rosie Cornish University of Bristol, UK
REVIEW RETURNED	18-Oct-2017

GENERAL COMMENTS	The authors have addressed most of my previous concerns. I have another query about the analysis used to examine the association between being overweight and obese and a couple of further comments about the multiple imputation. Logistic regression of overweight/obese 1. I don't really understand the rationale for including obese individuals in the denominator when looking at the risk of being overweight. From a public health point of view you are not interested in whether someone is at greater risk of being overweight than being either normal weight or obese; you are interested in whether they are at greater risk of being either overweight or obese compared to normal weight. It follows from this that the appropriate way to analyse the data would be to use multinomial logistic regression (with normal weight, overweight, and obese as three categories within one single outcome variable).

	Multiple imputation 1. The authors state that, “Variables adjusted for the primary analysis were included for imputation”. They also need to explicitly state that birthweight, BMI and DXA data were included in all imputation models. 2. In their initial analysis the authors included several variables (energy consumption, physical activity, and so on) that were potentially predictive of obesity, fat mass and lean mass. Did they consider including these as auxiliary variables in the multiple imputation models? If not, I would recommend their inclusion as this could potentially impact on the results. 3. Although initial texts on multiple imputation suggested that a very small number of imputed datasets would be sufficient, subsequent research has recommended carrying out more imputations than this. I suggest that the authors look at such recommendations and increase the number of imputed datasets accordingly.

VERSION 3 – AUTHOR RESPONSE

Reviewer: 3
Reviewer Name: Rosie Cornish
Institution and Country: University of Bristol, UK
Please state any competing interests: None declared

Please leave your comments for the authors below

The authors have addressed most of my previous concerns. I have another query about the analysis used to examine the association between being overweight and obese and a couple of further comments about the multiple imputation.

Logistic regression of overweight/obese

1. I don’t really understand the rationale for including obese individuals in the denominator when looking at the risk of being overweight. From a public health point of view you are not interested in whether someone is at greater risk of being overweight than being either normal weight or obese; you are interested in whether they are at greater risk of being either overweight or obese compared to normal weight. It follows from this that the appropriate way to analyse the data would be to use multinomial logistic regression (with normal weight, overweight, and obese as three categories within one single outcome variable).

 Thank you for your comment. We agree that a “noncase” in overweight analyses should be only those with normal weight, not including those with obesity. After considering reviewer’s comment and reviewing previous studies with same outcomes (overweight, obesity) as our study, we decided to compare those with normal weight and overweight, and then compare those with normal weight and obese, respectively. The overall trend of results did not change, and the revised parts are as following:

'Methods'
Participants being neither overweight nor obese were used as a “noncase” in both overweight and obesity analyses.

Multiple imputation
1. The authors state that, “Variables adjusted for the primary analysis were included for imputation”. They also need to explicitly state that birthweight, BMI and DXA data were included in all imputation models.

2. In their initial analysis the authors included several variables (energy consumption, physical activity, and so on) that were potentially predictive of obesity, fat mass and lean mass. Did they consider including these as auxiliary variables in the multiple imputation models? If not, I would recommend their inclusion as this could potentially impact on the results.

3. Although initial texts on multiple imputation suggested that a very small number of imputed datasets would be sufficient, subsequent research has recommended carrying out more imputations than this. I suggest that the authors look at such recommendations and increase the number of imputed datasets accordingly.

 Thank you for your comments. Per the reviewer’s comment, we explicitly stated that birthweight, BMI and DXA data were included in imputations. Also, we added health behaviors information and health status data to auxiliary variables, and also explicitly stated them. For the number of imputations, after reviewing recent studies on multiple imputation, we increased the number of datasets to 40. Changes made in the manuscript are as follows:

'Methods'
Markov chain Monte Carlo multiple imputation technique was used to create forty datasets, following conservative recommendations[19]. In addition to variables included in the primary analyses (e.g. sociodemographic factors, birthweight, BMI, and DXA data), health behaviours information (e.g. smoking, drinking, physical activity, and energy consumption) and health status related to obesity (e.g. hypertension, diabetes, asthma, and impaired fasting glucose) were included as auxiliary variables.

VERSION 4 – REVIEW

REVIEWER	Rosie Cornish University of Bristol, UK
REVIEW RETURNED	04-Dec-2017

GENERAL COMMENTS	My previous concerns have now been addressed.