PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Assessing the impact of adjusting for maturity in weight status classification in a cross-sectional sample of UK children
AUTHORS	Gillison, Fiona; Cumming, Sean; Standage, Martyn; Barnaby, Catharine; Katzmarzyk, Peter

VERSION 1 - REVIEW

REVIEWER	Tim Cole UCL Great Ormond Street Institute of Child Health, UK I constructed the British 1990 reference, as used in the paper.
REVIEW RETURNED	08-Feb-2017

GENERAL COMMENTS	The authors use the Khamis-Roche method to estimate maturation status in children aged 9-11, and show that rates of overweight and obesity based on BMI are reduced using biological rather than chronological age. I have some comments on the presentation and interpretation of the study findings. 1. The thesis is that parents of children reported as being overweight or obese (e.g. via the National Child Measurement Programme) are often aggrieved that the child's advanced maturation has not been taken into account, and that if it had been they would not be judged as overweight. The results confirm this to some extent, in that the overall rates of overweight and obesity fall when biological rather than chronological age is used to adjust BMI. However there is a fundamental misconception with the study. The purpose of assessing overweight/obesity in childhood is as a proxy for overweight/obesity in adulthood, where its impact on morbidity and mortality is clear and well documented. It may be that BMI appears less extreme when assessed using biological rather than chronological age, but the key question is: are more mature children at greater risk of later obesity after their maturity is taken into account? And the answer to the question is yes; there is a well-documented positive association between early puberty and later obesity - see for example this systematic review by Prentice P, Viner RM. Pubertal timing and adult obesity and cardiometabolic risk in women and men: a systematic review and meta-analysis. Int J Obes 2013;37:1036-43. Its conclusion is that "Earlier pubertal timing is predictive of higher adult BMI and greater risk of obesity." The statement on page 11 looks odd: "we are not aware of any research that has explored the effects [on BMI] of biological maturity in children and adolescents" - the literature on this is huge. And similarly, further down the page, "it is likely that early maturing

	adolescents can be a normal or healthy weight at a higher BMI than their later maturing peers". This, though not wrong, is highly misleading. 2. Some more minor comments. The paper uses the term "percentiles" throughout, which is primarily US usage, whereas "centiles" are used in the UK. It is a UK-based study, with British children and the British 1990 reference, so I respectfully suggest that centiles be used. I acknowledge my competing interest in having constructed the British 1990 reference. 3. The abstract refers to 9-11 year old children. It is overkill to also give the mean, SD and range of age. 4. Please avoid excessive numerical precision. The confidence interval for the odds ratio in the abstract extends to 19.0, which could reasonably be rounded to 19. See http://www.bmj.com/content/350/bmj.h1845. Similarly, giving F ratios (page 9) or chi-squared statistics (Table 1) to five significant digits is excessive. See http://adc.bmj.com/cgi/content/full/archdischild-2014-307149. 5. The strengths and limitations refer to adolescence. I'm not convinced that age 9-11 covers adolescence in boys. 6. Introduction line 2, 'practiced' is US spelling - please use UK spelling. 7. BMI is a ratio of weight to height, not the other way round (top of page 5). 8. Page 5 line 28, "it would is readily possible" needs attention. 9. The Methods give the median error bounds for predicted adult stature using the Khamis-Roche method. They would be worth explaining, as the mean prediction ± median error bound gives an interval containing 50% of the data. 10. The last two sentences of the first paragraph on page 8 duplicate what has gone before, and could usefully be deleted. 11. The ANOVA results (page 9) are for 3 categories of weight status and 2 sexes, giving 6 cells or 5 degrees of freedom. But they are not the appropriate d.f. to test for, as they consist of linear and quadratic weight effects, a sex effect, and interactions. The correct test would be the linear weight effect and its interaction with sex. Incidentally, what is eta squared? 12. If height is measured to the nearest 0.1 cm (Methods), why is the girl's height 166.35 cm (page 9)? 13. Figure 1 should include the raw data, and a better format would be as a bar chart. The legend needs reversing to match the two lines.

REVIEWER	Emma Eyre Coventry University, UK
REVIEW RETURNED	10-Mar-2017

GENERAL COMMENTS	This is a well written paper which addresses an important topic around classifications of weight status in children and the additional consideration of maturing in improving this. The paper makes a novel contribution to the field. The work has been conducted on a large sample, is current and conclusions have important messages for practice. The methods are described in sufficient detail. The authors acknowledge some of the limitations of their work.Further comments i have: It would be useful to understand why Kamis-Roche Method was the chosen method for determining maturation over other available methods (such as Mirwald). Additionally, could self reported parental height have impacted on your findings? Given the work around increased fat mass around the trunk during time leading to peak height velocity, i wondered if the authors had considered or obtained measures of waist circumference? Is the high participation rate of white children (98%) symbolic of the general childhood population in the southwest?

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Tim Cole
UCL Great Ormond Street Institute of Child Health, UK Please state any competing interests or state ‘None declared’:I constructed the British 1990 reference, as used in the paper.
---

1. The thesis is that parents of children reported as being overweight or obese (e.g. via the National Child Measurement Programme) are often aggrieved that the child's advanced maturation has not been taken into account, and that if it had been they would not be judged as overweight. The results confirm this to some extent, in that the overall rates of overweight and obesity fall when biological rather than chronological age is used to adjust BMI.

However there is a fundamental misconception with the study. The purpose of assessing overweight/obesity in childhood is as a proxy for overweight/obesity in adulthood, where its impact on morbidity and mortality is clear and well documented. It may be that BMI appears less extreme when assessed using biological rather than chronological age, but the key question is: are more mature children at greater risk of later obesity after their maturity is taken into account? And the answer to the question is yes; there is a well-documented positive association between early puberty and later obesity - see for example this systematic review by Prentice P, Viner RM. Pubertal timing and adult obesity and cardiometabolic risk in women and men: a systematic review and meta-analysis. Int J Obes 2013;37:1036-43. Its conclusion is that "Earlier pubertal timing is predictive of higher adult BMI and greater risk of obesity."

- Professor Cole raises an important point that we can certainly address in more detail in the paper.

Further, and while we acknowledge that some of our assumptions were not made explicit, we would argue that the study is not misconceived. To clarify, we fully agree with Professor Cole that most research suggests that obesity in childhood tracks to adulthood, which is when many negative implications for health occur. Similarly, we agree that early maturing children are more likely to be overweight in adulthood – and as such our finding in relation to pubertal timing is certainly not unexpected. Yet, we are not attempting to contest either of these positions nor can we via our cross-sectional design. Rather, we are addressing a different question of whether by using the current methods of weight classification we are correctly distinguishing between children who are biologically more advanced, and children who are clinically overweight.

- We have attempted to make this more apparent in our Introduction section; we have added a new first sentence confirming the link between childhood obesity and adult health risk, and a second sentence in paragraph 1 of page 2 where maturity is discussed making explicit the recognition that early maturing children are at greater risk of lifelong obesity.

The statement on page 11 looks odd: "we are not aware of any research that has explored the effects [on BMI] of biological maturity in children and adolescents" - the literature on this is huge.

We agree with the Reviewer, that this statement did not clearly articulate our intended purpose, and thus could be misleading. Accordingly, we have now amended this paragraph as follows:

“Past work has explored how moderating factors such as sex, race and ethnicity may influence the accuracy of BMI in predicting health risk,[33,34] however whereas the impact of puberty on BMI at a given chronological age is well established, few studies have attempted to quantify the impact of biological maturity has on the accuracy of weight classifications.[35] A sensitivity and specificity analysis of BMI in classifying obesity (as measured by body fat mass established through DXA scans, establishing puberty through tanner scales) in adolescents of all ages in New Zealand reported 6-12% of misclassification [35]. Nonetheless, the present study is the first to demonstrate how weight classification may account for children’s maturity status in addition to age and sex when benchmarking BMI against growth reference charts, and to report on the likely effects (in terms of changes to weight classifications) of doing so.”

And similarly, further down the page, "it is likely that early maturing adolescents can be a normal or healthy weight at a higher BMI than their later maturing peers". This, though not wrong, is highly misleading.

- We would be happy to further resolve any potential for our statements being misleading, as we have been careful throughout the manuscript to ensure that we do not claim more than is reasonable from our data. Drawing from other authors who have raised concern that failing to account for biological maturity very likely results in the misclassification of overweight in adolescents for the same reasons as we have stated in the sentence that Professor Cole highlights as problematic, so we are not entirely sure in what way it is misleading. We had already included the caveat following on from the sentence to make it clear that we are not trying to definitively claim that recategorisation/ adjustment necessarily links to health risk – i.e., that this still needs to be tested. We are happy to look at this again if the problem could be clarified. We hope the text added to the Introduction and Discussion sections in response to other comments to better articulate our points in the context of past work is helpful in doing this. For example;

Page 5: “While there is reliable evidence that earlier puberty is associated with
a greater risk of obesity, and thus that the two may be somewhat conflated,[13] researchers have also raised the question of whether it is appropriate to judge weight status based on BMI during puberty

when some increase in body fat is normal and healthy.[14,15]”

Page 11 - exert set out in response to the previous comment, above.

Bini V, Celi F, Berioli MG, Bacosi ML, Stella P, Giglio P, Tosti L & Falorni A (2000): Body mass index in children and adolescents according to age and pubertal stage. Eur. J. Clin. Nutr. 54, 214.

Daniels SR, Khoury PR & Morrison JA (1997): The utility of the body mass index as a measure of fatness in children and adolescents: differences by race and gender. Pediatrics 99, 804.

O'Dea J & Abraham S (1995): Should body mass index be used in adolescents? Lancet 345, 657.

Taylor, R. W., Falorni, A., Jones, I. E., & Goulding, A. (2003). Identifying adolescents with high percentage body fat: a comparison of BMI cutoffs using age and stage of pubertal development compared with BMI cutoffs using age alone. European journal of clinical nutrition, 57(6), 764-769.

2. Some more minor comments. The paper uses the term "percentiles" throughout, which is primarily US usage, whereas "centiles" are used in the UK. It is a UK-based study, with British children and the British 1990 reference, so I respectfully suggest that centiles be used. I acknowledge my competing interest in having constructed the British 1990 reference.

- Thank you for raising this point. The term “percentile” has been changed to “centile” throughout.

3. The abstract refers to 9-11 year old children. It is overkill to also give the mean, SD and range of age.

- Based on this comment, we have removed the additional terms in the Abstract.

4. Please avoid excessive numerical precision. The confidence interval for the odds ratio in the abstract extends to 19.0, which could reasonably be rounded to 19. See http://www.bmj.com/content/350/bmj.h1845.

Similarly, giving F ratios (page 9) or chi-squared statistics (Table 1) to five significant digits is excessive. See http://adc.bmj.com/cgi/content/full/archdischild-2014-307149.

- We have amended the reporting of data from the convention of 2 decimal places, to match the references provided and the ‘rule of 4’.

5. The strengths and limitations refer to adolescence. I'm not convinced that age 9-11 covers adolescence in boys.

We have amended the Strengths and Limitations section to ensure we now refer to children and/or adolescents.

6. Introduction line 2, 'practiced' is US spelling - please use UK spelling.

We have now amended the spelling of this word.

7. BMI is a ratio of weight to height, not the other way round (top of page 5).

We have now corrected this in the text on page 5.

8. Page 5 line 28, "it would is readily possible" needs attention.

- We have corrected this error on page 5. The sentence now reads: “it would be possible to…”

9. The Methods give the median error bounds for predicted adult stature using the Khamis-Roche method. They would be worth explaining, as the mean prediction ± median error bound gives an interval containing 50% of the data.

- We have added a definition for median error to the document (page 7). This now reads; “(i.e. the confidence interval within which 50% of the cases for true height will fall)”

10. The last two sentences of the first paragraph on page 8 duplicate what has gone before, and could usefully be deleted.

We have deleted these as suggested.

11. The ANOVA results (page 9) are for 3 categories of weight status and 2 sexes, giving 6 cells or 5 degrees of freedom. But they are not the appropriate d.f. to test for, as they consist of linear and quadratic weight effects, a sex effect, and interactions. The correct test would be the linear weight effect and its interaction with sex. Incidentally, what is eta squared?

- In response to the comments above about the appropriateness of the analysis, we have taken further statistical advice. We agree that we could anticipate a non-linear increase in the dependent variable with weight category (there being likelihood that the point of inflection occurs between the healthy weight and overweight categories, and that overweight and obese categories are relatively similar), that there may be a quadratic relationship represented by the interaction term, but that the relationship between sex and the dependent variable is linear: we hope this is a correct interpretation of the Reviewer’s points.

However, having discussed our options for dealing with this, we still believe that an ANOVA is appropriate for this analysis given that we are primarily interested in the differences between groups (rather than the linear relationships or trends underpinning group allocation). The options for reflecting the potential presence of quadratic data in ANOVA include transforming the variables, or performing a contrast analysis (i.e., whereby both overweight and obese categories are contrasted with the normal weight category). In either case, we generate the same F statistic (i.e., the same level of between-group differences) and the same degrees of freedom (i.e., a test of the main effect). Thus, we have retained our results as they currently stand after this scrutiny. However, we would be happy to consider alternatives if greater guidance as to what alternative test is sought could be provided.

- In response to the final comment, we have now included information on our selection and interpretations of the effect size indicators we have used in the Methods section (page 8). (i.e., $\eta 2$ indicates the effect size of F statistics; values ≥ 0.022 are considered a small but meaningful effect, ≥ 0.059 a moderate effect and 0.14 and upwards a large effect [41].)

12. If height is measured to the nearest 0.1 cm (Methods), why is the girl's height 166.35 cm (page 9)?

- This figure contains the additional decimal point as it represents a mean of values, rather than an

individual child's measurement (we had been adhering to conventions regarding the use of 2 decimal places). In line with comment 4, we have now reduced the number of decimal points reported here and throughout the paper.

13. Figure 1 should include the raw data, and a better format would be as a bar chart. The legend needs reversing to match the two lines.

- We have now presented the information in Figure 1 as a bar chart as requested.

Reviewer: 2
Emma Eyre
Coventry University, UK
Please state any competing interests or state 'None declared': None declared
---

This is a well written paper which addresses an important topic around classifications of weight status in children and the additional consideration of maturing in improving this. The paper makes a novel contribution to the field.

- We are grateful for the Reviewers' positive comments regarding the contribution that this paper could make to the field.

The work has been conducted on a large sample, is current and conclusions have important messages for practice. The methods are described in sufficient detail. The authors acknowledge some of the limitations of their work. Further comments i have:

It would be useful to understand why Kamis-Roche Method was the chosen method for determining maturation over other available methods (such as Mirwald). Additionally, could self reported parental height have impacted on your findings?

We selected the Kamis-Roche method as this has been shown to be a reliable and acceptable means of assessing maturity status in many countries. The Mirwald method has been shown to be quite sensitive to systematic error associated with both age and maturity status - as such it would be a less reliable and valid indicator of maturity in this instance. For example in boys, it under predicts age at peak height velocity in younger boys and early maturing boys and over predict age at peak height velocity in older boys and late maturing boys. These errors can be up be up to 2 to 3 years in some instances (see Malina & Koziel, 2014).
Malina, R. M., & Kozieł, S. M. (2014). Validation of maturity offset in a longitudinal sample of Polish girls. Journal of sports sciences, 32(14), 1374-1382.
We adjusted for anticipated over-estimates in parental self-reported height in line with published guidelines. While not perfect, this has been found to be the most appropriate means of acknowledging and controlling for self-reported data. This approach is stated in page 7, middle of paragraph 1 (now highlighted).

Given the work around increased fat mass around the trunk during time leading to peak height velocity, I wondered if the authors had considered or obtained measures of waist circumference?

Dr Eyre raises an important point about the limitations of BMI in estimating overweight and obesity, and the changes in body composition during puberty. We did measure waist circumference of the

children in our sample, however for the purpose of this article chose to stick to BMI as this is the measurement that is used in most national measurement initiatives. Past discussion also concludes that BMI is a more useful indictor (e.g., Katzmarzyk et al, 2007), as a result of the lack of national reference data for children's waist circumference, and lack of mapping between children's waist circumference and clinical health risk.

Katzmarzyk, P. T., Janssen, I., Morrison, K. M., & Tremblay, M. S. (2007). Classification of overweight and obesity in children and adolescents. CMAJ, 176(8S), 27-32.

Is the high participation rate of white children (98%) symbolic of the general childhood population in the southwest?

We acknowledge that the proportion of white children in our analytical sample is not representative of the country as a whole, but in answer to Dr Eyre's question, it is more (if not totally) representative of the local population. We have added a note to this effect on page 6.

VERSION 2 – REVIEW

REVIEWER	Tim Cole UCL Great Ormond Street Institute of Child Health UK I constructed the British 1990 growth reference.
REVIEW RETURNED	21-Apr-2017

GENERAL COMMENTS	The authors have tried hard to respond to my comments, but in terms of my main point not entirely successfully. 1. The problem I think is that we see the research question differently. What they show is that high BMI in puberty is associated with advanced developmental age, and that if developmental age is adjusted for the BMI appears to be less high. My response is that the risk for later adverse outcome associated with the high BMI has simply been partitioned into two components, the advanced developmental age and the correspondingly lower BMI. But the adjustment does nothing about the child's future risk, it just rebadges it, so I'm unclear about its value. However I think we have to agree to differ, so I won't pursue it further. Most of my other points have been addressed. 9. Note it is Khamis-Roche not Kamis-Roche (page 7). 11. My previous point about the ANOVA remains. To state that "children were significantly more likely to be classified as overweight or obese if they were biologically more mature" indicates a significant linear trend, which should be 1 not 5 d.f. As I said before it should be the overall linear weight trend, and separately the interaction of this trend with sex (which is very obviously present in Figure 1). 12. The fact that height was obtained as the mean of two measurements is not mentioned in the Methods.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 1

(remaining points to be resolved)

9. Note it is Khamis-Roche not Kamis-Roche (page 7).

- We are grateful to the Reviewer for identifying this error, and have now corrected it in the text.

11. My previous point about the ANOVA remains. To state that "children were significantly more likely to be classified as overweight or obese if they were biologically more mature" indicates a significant linear trend, which should be 1 not 5 d.f. As I said before it should be the overall linear weight trend, and separately the interaction of this trend with sex (which is very obviously present in Figure 1).

- Thank you for this clarification, and we can see how our wording in describing the analysis caused this issue. We have amended the text to clarify this analysis, in line with the written APA guidelines for reporting ANOVA. As such we have presented the F statistic for each variable as it is discussed, as we believe is what you are suggesting. We note, that the df for weight category is 2, as there are 3 groups (rather than 1, as you suggest), but the df for gender is of course 1. We hope this now addresses your point fully. The text now reads (page 9, line 8 onwards):

“The results of a 2-way (gender and weight status) ANOVA indicated that there was a significant difference in biological maturity across weight categories ($F(2,401)=38$, $p<0.001$; $\eta 2=0.16$), gender ($F(1,401)=422$, $p<0.001$; $\eta 2=0.51$), and a significant interaction term ($F(2,401)=5.5$, $p=0.005$; $\eta 2=0.03$). The data show a trend for girls to be more biologically mature than boys at this age, for biological maturity to be more advanced in higher weight categories, and for the difference in biological maturity between weight categories to be more pronounced in girls (Figure 1).”

12. The fact that height was obtained as the mean of two measurements is not mentioned in the Methods.

- We have added a statement to clarify the measurement procedure in the Methods (page 6).