

Reviewer #1: Primary Reviewer

As the authors indicate, the US has experienced rapid increases in childhood obesity during the past two decades and a decrease in the median age of menarche. While both of these "trends" provide independent cause for concern to physical and mental/psychological health, few longitudinal studies of differences in growth by timing of menarche have been conducted- particularly on girls younger than five years of age. This study was designed to address this question and included the following specific aims: At what ages were growth differences evident in early, mid- and late onset menarche groups?; Were the associations between growth and menarcheal timing similar for white and African-American girls?

In addition to the potential significance, the study has several strengths. The review of literature includes and addresses seminal and recent relevant articles to support the significance and identify gaps in the current state of the science. Important to emphasize is that key longitudinal studies that include data on both age of menarche and height and/or weight measures in girls \leq five years of age were based on small and fairly homogenous samples. These studies did not allow for examination of the key questions posed in this study; results did not provide information regarding the timing of menarche on growth patterns in height and BMI from early childhood through late adolescence.

Another strength is the source, diversity, size of the sample and number of eligible participants (90%) who actually participated in the study.

The detailed data analysis section (pages 8-10) clearly explains the approach(es) considered and justifies the approach selected. Strengths and limitations (as well as information yield) with both fixed and mixed methods model approaches are detailed.

The discussion section does not go beyond the data/ results presented; however, a few points warrant clarification. It is indicated (page 14) that results suggest that the "pathway through which growth patterns and maturation differences must begin in childhood". A bit more elaboration on this observation would strengthen the manuscript. It is noted (page 15) that four differences are important to emphasize when comparing results across race. These important points merit elaboration. The new finding that "the majority of African-American females may be experiencing childhood growth patterns that have been linked to significant chronic health problems" also merits more explanation/ emphasis.

Response: We have revised the discussion section extensively in an effort to go beyond summarizing the results. We have organized the discussion around the two research questions.

The study has important limitations that are outlined by the authors including self or mother-reported age of menarche. While staging of menarche (early, mid- and late onset) has no accepted standard definition, the methods used in this study attempt to follow those used in similar key longitudinal studies. Some discussion of the limitations of method used in this study in relation to results observed should be included.

Response: This is an important point. The paragraph on limitations now discusses the limitations of using menarche and self or mother reports of menarche as a proxy for pubertal timing. Since we are stratifying on reports, which likely contain measurement error, there is noise in our classification of the three staging groups. In addition menarche is only a proxy for the timing of a complex nexus of pubertal events, including adrenarche, thelarche and pubarche, which could be measured more precisely with data from physical exams conducted more frequently than every two years. The misclassification introduced by noisy measure of pubertal timing acts like classical measurement error

on an explanatory variable- namely it will bias estimates of size effects towards zero. We outline these points in a revised limitations paragraph that explains how the limitations will bias our results towards the null finding of no differences between the staging groups.

More specific suggestions for future research- based on results obtained and the state-of the science should also be included.

Response: We have included a discussion of implications in a new and expanded section entitled, "Conclusions and Implications". We specifically identify research questions and future directions for research that are a direct result of the study reported here.

Finally, given the selected publication outlet, Nursing Research, the potential implications for healthcare warrant elaboration.

Response: We appreciate the opportunity to elaborate upon the implications of our findings for clinical practice and have done so in the last section of the paper.

Reviewer #2:

General comments: This paper presents interesting ideas regarding the stated aim of the research, to examine the association between timing of menarche on growth patterns in height and BMI from early childhood through late adolescence in a racially diverse, longitudinal sample. The longitudinal design and large, diverse sample represents strengths of this study.

The title and abstract communicate an adequate representation of the intent and findings of the manuscript.

Introduction and Background: The significance of the manuscript is introduced with the discussion of childhood obesity. The authors provide a plausible argument that menarche onset may be associated with subsequent BMI and height.

Literature review: A careful review addresses the limitations of published studies, providing the background for the gap in knowledge that would be addressed through the findings of this manuscript. The authors posit that obesity contributes to early menarche and subsequent growth patterns, in contrast with the prevailing idea that early menarche predisposes to obesity. The authors highlight the importance of tracking girls "over the complete age range" as an indication of the novel nature of the study, underscoring the significance of establishing the critical period for intervention to prevent obesity.

Style of presentation: The manuscript is written in a clear, concise manner overall.

Research design: The authors utilized an appropriate design to address the stated research questions: (1) At what ages were growth differences evident in early, mid and late onset menarche groups? (2) Were the associations between growth and menarcheal timing similar for the white and African-American girls? The sample was derived from the CNLSY79 via biennial interviews including self-report and objective measures. There was ambiguity regarding the age of entry of the girls in the sample, a significant concern in light of the author's statements of the novel nature of this manuscript.

Response: This is an extremely important point. In the introduction we report that the median age at which African-American girls were first interviewed was 30 months (2.5 years of age) and the median

age at which white girls were first interviewed was 22 months (~=1.8 years of age). We restate the median age at which girls first enter the study in the data section.

The data was limited to the mean number of interviews in each group.

Response: We believe that the mean numbers of interviews becomes more meaningful when taken in context of the median age of entry into the sample.

Also unclear is the statement on page 9, line 18, referring to 'estimates' to predict height and weight.

Response: We recognize that both Reviewer 2 and Reviewer 3 found ambiguity in the source of the “estimated” or “predicted values” of BMIs and heights. We regret that in an attempt to minimize the number of tables, we caused unnecessary confusion about the source of estimates used in our graphs. We now report the coefficient estimates of the key variables in the mixed models of BMI and height as a function of chronological age in Tables 2 and 3. We also include tables of the predicted values (and associated 95% confidence intervals) at selected ages- 2,3,4,5,6,7,8 and 20 years of age in Tables 4 and

5. The predicted values were used in the graphs. We have stated that a complete set of parameter estimates, tables of predicted values, and graphs are available on request.

The definition of overweight by BMI should be included.

Response: We have added the definition of overweight when it is first introduced in the results section.

Data analysis: The authors provided adequate support for model selection, imputed data as well as the use of self-report. Figures are explained clearly. The tables provide a succinct summary of findings for significant differences in height and BMI.

Organization: The authors presented this manuscript in a clear, logical and organized manner.

Reviewer #3:

The introduction is well-written and highlights an important topic that has implications for nursing. The analytical method used also may be applied to other areas that are of interest to readers and researchers. While the aim of the paper is to describe the results, rather than be a methods paper, I think it would be helpful to provide a bit more background information. It is not clear how you got from curves to ages in Tables 1 and 2. I went back to Biro and Demerath articles and it looks like they actually used statistical techniques in interpreting differences at specific points in their curves?

Response: We too went back to the Biro and Demerath articles and noted that they reported coefficient estimates for the models used to generate the data on which their graphs are based. Both Reviewer 2 and 3 pointed out the need to provide the parameter estimates that were the source of the data points used in the graphs. We therefore include tables of the coefficient estimates of the key variables in the mixed models of BMI and height as a function of chronological age in Tables 2 and 3. We also include tables of the predicted values (and associated 95% confidence intervals) at selected ages- 2,3,4,5,6,7,8 and 20 years of age in Tables 4 and 5. The predicted values were used in the graphs. We have stated that a complete set of parameter estimates, tables of predicted values, and graphs are available on request.

You say that you used Wald tests (page 10) - why are these results not shown?

Response: We did not report the actual test statistics to conserve space. However, we report that Wald tests for race differences by menarcheal staging group rejected the null hypothesis of no race differences for both BMI and height. We provide the actual value of the test statistics in the table below.

Wald-test statistics- Race differences by menarcheal staging group

	Early	Middle	Late
Race differences in BMI by staging group Wald test- chi2(4 degrees of freedom)	40.61	61.82	74.35
Race difference in height by staging group Wald test- chi2(3 degrees of freedom)	65.03	56.80	61.02

We also report that Wald tests for staging group differences within each race rejected the null hypothesis of no staging group differences for both BMI and height. We provide the actual values of the test statistics in the table below.

Wald-test statistics- Menarcheal staging group differences by race

	African-Americans	Whites
Staging group differences in BMI Wald test- chi2(8degrees of freedom)	71.22	138.07
Staging group differences height Wald test- chi2(6 degrees of freedom)	228.40	352.72

We are happy to include any additional tables of coefficients, predicted values and test statistics if the referees and editors believe that there to be scientific merit in the additional information. We have tried to make clear that complete results are available upon request.

I am not familiar with this longitudinal analysis, so can not judge the adequacy of your approach, but found the presentation of the results difficult to follow.

Response: Upon rereading the results section we agree that it was difficult to follow. We have substantially reorganized the section, focus on the two main research questions of this paper. We hope that the referees find the revisions to be more accessible.

Although the discussion section is well-written and recognizes the limitations, the clinical, NURSING implications could be further highlighted.

Response: Both referee 1 and 3 suggesting expanding the discussion to include the implications of the study for clinical nursing practice. We have reorganized the discussion by focusing separately on the practice implications of the study and the research implications of the study. We agree that this was a weakness of our previous draft and hope that our discussion of the healthcare practice implications will be useful to practitioners.

Page 4 - lines 3-4: you are referring to actual z-scores here?

Response: We believe we have addressed this point by changing the language from “BMIz scores” to “BMI z-scores.”

Results - I think a table comparing the African-American and white girls would be helpful. You have a lot of information in the text (page 7) and I think this makes it difficult to get a sense of the comparisons, whereas if it was delineated in a table, it would be clearer.

Response: We appreciate all of the comments helping us to clarify the different kinds of comparisons that we are making. We included a table, now Table 1, which concisely reports the age ranges for the three race-specific staging groups.

Reviewer #4:

This manuscript used a well-known data (NLSY79), conducted a longitudinal analysis to explore the relationship between BMI and height differences by age of menarche between African-American and White girls. In terms of the overall quality of the paper, I think it is very good. The data cleaning, description of the procedure, data analysis, figures and tables were all very nicely done. The authors did not mention anything about Asian girls though, I know there were Asian people in the sample, maybe the number was too sample. Is it possible to add another group of Asian girls to the research, I think that will further enhance the quality of the paper. However, I doubt the importance of the topic. For me, this kind of research are better published elsewhere rather than Nursing Research, I might be wrong though. On the other hand, I think we need to encourage more nurses conducting research using secondary data like this one.

Response: There are only 18 mothers whose primary ancestry is Asian. The distribution of maternal ancestry is as follows: Chinese=2; Filipino=13; Hawaiian/Pacific Islander=2; Japanese=0; Korean=1. There are 24 girls born to these mothers. The distribution of daughter's ancestry is as follows: Chinese=2; Filipino=18; Hawaiian/Pacific Islander=3; Japanese=0; Korean=1. As stated in the description of the sample we focus on non-Hispanic whites, excluding those of Asian descent, and the African-Americans. We recognize the importance of conducting this kind of research on large, representative samples of Asian-Americans, as well as Hispanics. In our revised discussion of the research implications of our work, we explicitly recognize the importance of studying both of these important groups.