Author's response to reviews

Title: Using latent class analysis to develop a model of the relationship between socioeconomic position and ethnicity: cross-sectional analyses from a multi-ethnic birth cohort study and implications for health research

Authors:

Lesley Fairley (Lesley.fairley@bthft.nhs.uk)
Baltica Cabieses (balticabeatriz@gmail.com)
Neil A Small (N.A.Small@bradford.ac.uk)
Emily S Petherick (Emily.petherick@bthft.nhs.uk)
Debbie A Lawlor (D.A.Lawlor@bristol.ac.uk)
Kate E Pickett (kate.pickett@york.ac.uk)
John Wright (john.wright@bthft.nhs.uk)

Version: 3
Date: 30 May 2014

Author's response to reviews: see over
Dear Editor

Thank you for reviewing our manuscript, providing reviewers’ comments and inviting us to submit a revision. We have now prepared a revised manuscript that reflects the suggestions made by the reviewers and enclose this with all changes highlighted for your consideration.

Below we have provided point-by-point response to the reviewers’ comments detailing the changes we have made in response to each issue raised. This includes page and line references to where changes can be found in the revised manuscript.

Lesley Fairley (on behalf of all authors)

Reviewer: Margaret Kelaher

Reviewer’s report:
This is a well written article examining the use of latent class analysis to develop measures of socioeconomic position. The paper then explores ethnic differences in socioeconomic position using these measures. The main issue with the paper is that it does not include any health related variables which will massively influence its interest to the audience of BMC public health.

- While the analysis included in the paper does not include any health related variables we still think it will be of interest to the readers of BMC Public Health as socioeconomic position is a key exposure and confounder in most health related studies that it is important to explore different methodologies to combine several measures into one indicator. Given the multidimensional nature of SEP, the main reason for wanting combined indicators is that we think different measures of SEP might mean something slightly different, and we want to use all that information in a statistically efficient way. Thus it is important to explore different methodologies to combine several measures into one indicator. The main focus of this paper was to describe this methodology and the relationship between the resulting classes and ethnicity which we feel merits a paper in its own right. We intend to investigate how these classes are linked to different health outcomes in our multi ethnic populations in future research. We have made several changes to the manuscript strengthen the aims of this paper.
• The title has been changed to "Using latent class analysis to develop a model of the relationship between socioeconomic position and ethnicity: cross-sectional analyses from a multi-ethnic birth cohort study and implications for health research"

• The introduction and the discussion of the abstract on page 2 (lines 2-6 and lines 24-28) have been updated.

• We have also updated the introduction (page 4 lines 27-28 and page 5 lines 17-23) to be clearer about why this research is important

Reviewer: Laura Marlow

Reviewer's report:
This paper was interesting and well written throughout. The authors make a good case for what their analyses can add to our understanding of measuring SEP. The question is well defined and uses a large diverse sample of women, with multiple measures of SEP. The methods are well described and easy to follow. However, there is no discussion of the assumptions for LCA.

• The main assumption for LCA is conditional independence, that is, within each class all measures are independent as all correlation between the variables is explained through the class structure. This has been added to the methods section of the manuscript on page 9 line 27 and page 10 lines 1-2.

I am a little confused by the description of the ethnicity question used, in particular the categories Indian Caribbean and African Indian. Perhaps this is a typo or maybe the authors added these sub-categories, but they are not in the original 2001 census question. Please clarify.

• The ethnicity questions used were based on guidance from the Office for National Statistics but culturally adapted for the local Bradford population. We recognise that these descriptions are confusing and they did not play a part in our analysis so we have deleted these details from the manuscript.

I was interested to see that housing status was not included in table 6 (the LCA within Pakistani women) could the authors comment on this.

• The four classes did differ by housing status with the "Women not employed, moderate education, benefits, subjectively poor and materially deprived" group most likely to be living
in social housing, we have updated table 7 (page 26) and the discussion (page 14 line 6) to include this information.

The variation in the profile depending on which ethnic group the analysis was being run with suggests the initial categories (shown in table 3) may not be applicable to any one ethnic group. The authors should offer some interpretation of this finding and the implications it has for researchers considering the use of this method in the future.

- We have added an extra paragraph to the discussion on page 14 lines 9-15. We found that the two approaches we undertook showed that different components of SEP do aggregate differently within the two largest ethnic groups in this study. This suggests that there may not be one overall class descriptor that can be used across several ethnic groups and that ethnic group specific classes may need to be defined. However in this study and other studies there may be problems defining these classes based on the availability of data on different SEP measures and also the sample sizes in the different ethnic groups. In this study we were only able to define classes for the largest two ethnic groups.

Perhaps more information in table 2. I would have liked to have seen some more socio-demographic characteristics of the sample and this could also be useful for highlighting the amounts of missing data for each question.

- Table 2 (page 20-21) has been extended to include all the socioeconomic position variables considered in the latent class analysis, this also includes the rates of missing data for each item.

- The material deprivation measure is subjective and it could be that ethnic minority women are more satisfied with what they have, but there could also be some bias in reporting here. Some ethnic groups may be less comfortable reporting being unhappy with what they have. I might also question the cross-cultural appropriateness of the eleven items from the Households Below Average Income Survey [19]. These are points the authors may like to acknowledge in the discussion.

- The eleven items used to measure material deprivation were taken from a large national study, the Households Below Average Income Survey, and these items have been used previously to look at material deprivation across ethnic groups in relation to child poverty.
The questions are asked in such a way that it reduces the risk that the question is not culturally appropriate because it is linked to the respondents impression and does not imply all groups should want the same items ie they can reply “don’t want or need”.

Reviewer: Charu Mathur
Reviewer's report:

Major Compulsory Revisions

It's not clear what are the implications of the findings in terms of improving the health of these disadvantaged women, and how can these findings help future research? The authors have focused only on the methodological aspect of their research/findings and fail to answer the question "so what" or "what do we do with these findings"

- The aim of this paper was to describe how latent class analysis can be used to help us understand and combine several measures of SEP into one measure. In order to examine ethnic differences in health it is important to make appropriate adjustment for SEP, assuming that once SEP is considered any ethnic group differences are due to factors linked to ethnicity such as cultural and genetic differences. It is unclear whether gradients in health by SEP in particular ethnic groups are not always found due to poor and inappropriate measurements of SEP for such groups. The method we describe in the paper shows how several measures of SEP can be combined and we found that different components of SEP aggregate differently in two large ethnic groups in our study sample suggesting that different measures may be more useful/applicable for different ethnic groups. Further work is needed to replicate these findings in other studies and further work is needed to see how these classes are associated with different health outcomes for both the women and their offspring.

- We have made several changes to the manuscript strengthen the rationale of this paper.

- The title has been changed to "Using latent class analysis to develop a model of the relationship between socioeconomic position and ethnicity: cross-sectional analyses from a multi-ethnic birth cohort study and implications for health research"

- The introduction and the discussion of the abstract on page 2 (lines 2-6 and lines24-28) have been updated.
• We have also updated the introduction (page 4 lines 27-28 and page 5 lines 17-23) to be clearer about why this research is important

Discussions section is not well developed and needs work.
Rationale for doing analysis 1) entire sample with ethnicity as a covariate, 2) doing a subgroup analysis

• We felt it was important to do the analysis overall and stratified by ethnic group for the two largest groups in our study. This further subgroup analysis allowed us to investigate if different numbers and types of classes were found in different ethnic groups. Our stratified analysis found that the classes differed by ethnic group again highlighting that some measures of SEP may be more applicable for some ethnic groups.

• We have added an extra paragraph to the discussion on page 14 lines 9-15 discussing the implications of these finding on further research.

The authors should present a table for model selection with the fit indices (AIC,BIC etc) for models with 2-10 latent classes.

• Initially we ran models for 1 to 7 classes. We have run the extra models for the 8, 9 and 10 classes and have included model fit statistics (Log likelihood, percentage reduction in likelihood, BIC and entropy) in a table (table 3, page 22)

The authors should present the range of average posterior probabilities of membership in each class.

• For each individual we obtained the posterior probability of membership to each of the classes, however the range of these probabilities for each class is from 0 to 1. For some individuals they may be clearly more weighted into one class and for other the probabilities may be split more evenly across the classes. We used weighted multinomial regression which was weighted by the probability of belonging to each class rather than assign each individual to the class with the highest probability. Therefore we do not think adding this information would be that informative.