Author's response to reviews

Title: Socioeconomic position and allostatic load: Evidence from the West of Scotland Twenty-07 cohort study

Authors:

Tony Robertson (tony.robertson@ed.ac.uk)
Frank Popham (frank.popham@glasgow.ac.uk)
Michaela Benzeval (mbenzeval@essex.ac.uk)

Version: 2 Date: 29 January 2014

Author's response to reviews: see over
Socioeconomic position across the lifecourse and allostatic load: Evidence from the West of Scotland Twenty-07 cohort study

Dear Dr Bosma,

We would like to thank you and the reviewers for your time, consideration and constructive comments on our paper submitted for publication in BMC Public Health. Please find below responses to each reviewer comment in turn. We have also submitted two versions of the updated manuscript, one with ‘track changes’ and one with all changes accepted. All references to pages/paragraphs relate to the ‘tracked changes’ version of the manuscript.

We thank you for your time and look forward to hearing from you.

Yours sincerely,

Tony Robertson
Research Fellow, SCPHRP

Phone: +44 (0) 131 651 1591
Email: tony.robertson@ed.ac.uk
EDITOR’S COMMENTS

1. Table 1 shows how the different models (accumulation, critical, and mobility) were estimated. In the text, the authors referred to references 28 and 29 for more detail. As this is crucial for the current paper, I would prefer more detail about how the different models can be truly separated. Related to this, it is not clear why the critical model should be examined by including the separate SEP measures ‘in separate models’. To truly see whether SEP1 is critical, I would like to see its influence independent of SEP2 and 3. And, suppose that upward mobility was protective (e.g. in earlier life), how sure can we be that it is an effect of upward mobility and not of SEP2 per se (as critical period)? The authors should either provide further convincing information about how to test these three models or should think about simpler models in which the three SEP measures are included in a (simple) linear regression analysis (informing us about accumulation and critical periods), and in which separate two-way interactions are separately tested between the three SEP measures (informing us on the importance of mobility).

We have added the following to Methods>Statistical analysis to improve the clarity and importance of the methods employed in the analysis:

“A structured modelling approach developed by Mishra et al was used to compare the three theoretical lifecourse models of accumulation (Figure 1a), critical periods (Figure 1b) and social mobility (Figure 1c). The basic idea of this approach is that, given three binary SEP variables, a saturated model would allow all eight possible SEP trajectories to have a different mean outcome. The saturated model is then modelled with three main effects, all two-way interactions, and the 3-way interaction, where the constant (α) is the expected mean for allostatic load for the trajectory where persons were non-manual (higher SEP) at all three time points. This modelling technique allows the direct comparison of each of the different lifecourse models (in the form of the simpler nested models) — accumulation, critical period and mobility hypotheses—to the all-inclusive (saturated) model. Using model-fit statistics, we can identify which of these simpler models has a fit as good as the saturated model. Given its simpler structure, any model found to fit the data as well as the saturated model is considered to be the most parsimonious. This structured modelling approach can provide a formal and clearer understanding of the relative merits of these alternative
hypotheses. Table 1 summarises the different lifecourse models. Two versions of the accumulation model were considered. The ‘strict’ model assumes that the longer a person spends in a lower SEP, the worse the physiological burden, irrespective of time period (i.e. having low SEP in childhood and the transition to adulthood will have an identical effect on allostatic load as having low SEP in childhood and adulthood). This model is estimated by constraining the regression coefficient between each SEP measure and allostatic load to an equal value (i.e. the mean effect of the three SEP measures). For the ‘relaxed’ model, each SEP measure is assumed to be contributing to the risk of higher allostatic load, but not necessarily in an equal manner (i.e. there is no such constraint imposed). For the critical period model, each SEP life-stage is considered to have an independent relationship with allostatic load, irrespective of SEP at other life-stages. This is estimated in the models by constraining two of the three SEP measures to equal zero. This is repeated for each of the three life-stages in turn. Finally, we have considered two mobility models. Firstly, early mobility between childhood and the transition to adulthood was considered and secondly mobility between the transition to adulthood and adulthood SEP. To estimate these effects in the models, all other SEP combinations (i.e. low SEP at both life-stages or higher SEP at both life-stages) are constrained to be zero. Only upward and downward mobility are considered, with the assumption that upward mobility will be associated with lower allostatic load and downward mobility with higher allostatic load compared to those showing stable SEP. Full model specifications are available in Supplementary Table 3.”

We have also added a Supplementary Table 3 that includes the model specifications and constraints in algebraic form.

2. The abstract’s conclusion reports that the pre-adulthood may represent a particularly sensitive period for SEP to impact on allostatic load. This seems a different message than the message in the Discussion of the paper and also different than the results in Table 2 where it is shown that in both younger cohorts the transition to adulthood period is (consistently) important for allostatic load.

This was an error and should have read ‘the transition to adulthood may represent a particularly sensitive period for SEP...’ This has been corrected.
3. Dichotomising the biomarkers for each separate cohort and sex results in problems comparing between cohorts and sexes. It is clear that the measures are summed afterwards to create the allostatic load measure, but much of the effect of cohort/age and sex is artificially ‘deleted’ by this procedure. It is then also a bit odd to see that sex is still adjusted for. Table 1 showing no differences between sexes cannot be interpreted easily, given the above.

In analysis of this study sample, we believe it is important to separate the three cohorts in order to prevent age effects distorting the associations identified if analysed as one sample. If we had not standardised by age cohort, one would find high values of allostatic load in the oldest cohort and lower values in the younger cohorts, reducing the variation that could be tested for in the analysis. Although we are not directly testing for differences between cohorts, this standardisation still allows visual comparisons between the cohorts and cohort-specific SEP-allostatic load associations to be identified. A similar strategy was employed for sex, although we did not separate the sexes in the analysis as this would have further reduced the power and added an additional question (i.e. are there sex differences in the association?). While standardising by sex did remove sex-specific variation in each biomarker, it might not fully account for differences in the overall allostatic load score e.g. perhaps a greater proportion of women than men had just one biomarker in the highest quartile of risk rather than several.

4. The Discussion reports (on page 10) “Although the accumulation model was identified as the best-fit model, it was apparent that a simple ‘summing of the risk’ approach might not allow us to fully understand the changing nature of the association between allostatic load and SEP over the lifecourse?”. This is unclear. Does it refer to the 1970s cohort where the ‘relaxed accumulation model’ shows the best fit? Or does it refer to the absence of an effect in the 1930s cohort and the difference with the younger cohorts? This needs further explanation. See also Conclusion where this is reported again.

We have edited the aforementioned sentence in the discussion to improve clarity: “Although the accumulation model was identified as the best-fit model in the 1970s and 1950s cohorts (with a longer time spent with higher SEP associated with lower allostatic...
load), it was apparent, given the relaxed accumulation model was the best-fit model in the 1970s cohort and there were significant life-stage differences in the 1950s cohort, that a simple ‘summing of the risk’ approach (assuming all life-stages pose identical risks) might not allow us to fully understand the changing nature of the association between allostatic load and SEP over the lifecourse.”

5. When discussing the cohort effects (page 11), it is not clear how higher life expectancies and more communicable diseases in younger cohorts might have resulted in stronger SEP-allostatic load associations.

We have edited the aforementioned sentence in the discussion to improve clarity:

“For example, the meaning of SEP has changed for the different cohorts (e.g. the growing importance of education in people’s lives with younger birth cohorts); life expectancy has increased with younger cohorts (i.e. they may be physiological ‘younger’ at older ages than previous cohorts); and the pattern of diseases has also altered across cohorts (e.g. shift from higher prevalence of communicable to non-communicable disease, with these different disease-types potentially impacting differently on physiological burden across the body).”

**REVIEWER ONE’S COMMENTS**

1. An inclusion of sociodemographic characteristics is necessary to include. A better characterization of the study sample in general is needed. The details are described elsewhere, as the authors mention, but readers wholly unfamiliar with the study sample will still want to know more about the sample than is currently provided in the methods section.

We have included more details about the study design and demographics in ‘Methods>Study Sample’ section, including an additional supplementary table detailing the demographics across the five waves of the study:

“The Study has two subsamples: the regional sample, a two-stage stratified random sample of people living in the Central Clydeside Conurbation, West of Scotland (a socially mixed and mainly urban area) and the localities sample of people from two areas of the city of Glasgow. The target sample at W1 for each cohort was 1,500; the overall achieved sample
was 4,510 (1970s n=1515; 1950s n=1444; 1930s n=1551). Baseline respondents have been shown to be representative of the general population of the sampled area.[19] The study design is described in more detail elsewhere.[20] The Tayside Committee on Medical Research Ethics approved the study."

“Approximately 53% of the respondents were female across the cohorts and this was stable over time (Supplementary Table 1).[21] Analysis of baseline data for those who participated at each wave showed that men, people from manual classes (lower SEP) and those with poor starting health were less likely to remain in the study, and in each case this was particularly true of the 1930s cohort. The latter was mainly due to mortality, with nearly 37% of this cohort having died by W5. Among those in the 1950s and 1930s cohorts the proportions reporting poor health increased over time as they aged, but for the 1970s cohort it was relatively stable until the most recent wave when there was a drop in those reporting poor health.”

2. The methods section should include information about the sample selection and recruitment. How did these participants get into the study? What is their representativeness?

Please see above response to Comment 1.

3. I was curious as to why the authors constructed a dichotomous allostatic load indicator. There are, as the authors mentioned, three different systems operating simultaneously. Why not separate by these different systems? Given the SEM and the broad range of biomarkers listed, it seems the authors would want to examine allostatic load from several different perspectives.

There are several examples of how to operationalize allostatic load, as highlighted by Juster, 2010*. The method we used involved dichotomising each biomarker measure into low and high based on being in the highest risk quartile (‘1’) versus not (‘0’). The final allostatic load score was then based on summing each of these binary scores into a continuous outcome ranging from 0 to 9. This method has been widely used in the literature, is relatively easy to understand (compared to using z-scores, for example) and captures the idea of allostatic
load better, in our opinion, than using clinical cut-offs or a higher cut-off (that will capture greater levels of clinical disease, rather than physiological dysregulation). While it is possible to break down the allostatic load score into its constituent parts, we feel this would move the focus away from the true concept of allostatic load, i.e. being about the body as a whole and the effect of multiple physiological changes working in a synergistic fashion to affect health.


4. The authors indicate that using SEP as a binary variable could be a limitation. But was this the only way that the data were collected or was making SEP a binary variable a decision on the authorship team? Again, there could be a lot of variance lost by restricting to a binary variable.

The decision to dichotomise was driven by the modelling structure we based our analysis on. Here it was necessary to limit the SEP measures to binary measures in order to keep the number of potential SEP trajectories manageable (i.e. 8 different trajectories, given three time points and a binary measure). Greater numbers of categories would have resulted in very complex results and reduced the sample size in many of the trajectories. Furthermore, dichotomising the measures allowed comparability between the social class and education measures. Dichotomising social class in this way (manual vs non-manual) is common and well-validated in the UK. The education split was based on proportions in each of the two groups and previous knowledge of the meaning of education in the UK during this time-period. We have discussed these issues in Discussion>Strengths and limitations section:

“We have followed a modelling approach which uses binary SEP measures at different stages of the lifecourse. Reducing these indicators to binary variables does simplify the information measured and may not allow us to identify non-linear patterns of association with allostatic load. Using SEP measures with multiple categories would increase the complexity of the models, as well as increasing less common SEP trajectories.”

5. A deeper understanding and explanation of the model fit statistics for SEM are needed. The authors mentioned p-values but there are other indicators that are typically used in
SEM (RMSEA, CRI). If there is a reason the authors are not using these fit statistics, they should say so in the analysis section. Additionally, I would have appreciated a visual representation of the model the authors endeavor to test since they are using SEM.

Our aim in the analysis was to replicate the methods employed by Mishra et al (2009)* who developed this specific analysis strategy for comparing the different lifecourse models. We did consider trying other model fit statistics but felt it was important to replicate the previous methods for consistency. The p-values are linked to the use of the F-test of model fit as used by Mishra et al. In addition there are limitations using Stata with clustered data and weighting that limits the available outputs, including other fit statistics such as AIC.


To visually represent the different theoretical models we have created a figure, which can be found in Methods>Statistical analysis (Figure 1).

**REVIEWER TWO’S COMMENTS**

1. Although the authors refer to the ‘cohort profile’, I would still like to see some more detailed information on loss to follow up and attrition in the description of the study sample. For example, how many have died, what about the SEP of people lost to follow up, why were data of 540 people incomplete? The authors might also refer to the supplementary table here.

   Please see above response to Reviewer 1, Comment 1.

2. In the ‘statistical analyses’ paragraph, the authors mention that “all analyses were weighted using inverse probability weights to correct for bias due to attrition”. Please explain (related to my first comment)!

   We have edited this statement to improve its clarity:
“All analyses were weighted to the living baseline sample at the time of the wave 5 interviews using inverse probability weights. [31] Weighting the analysis sample in this way inflates the weight for subjects who are underrepresented due to missing data in order to reduce bias introduced by changes in the sample characteristics over time (e.g. those with lower SEP being more likely to drop out of the study).”

3. Is it possible that ‘multiple comparisons’ affected the significance of findings?

Although there are multiple analyses, each analysis is of a simple model compared to a more complex model in which it is already nested. This non-independence between all the models would negate the multiple comparisons problem that arises in analyses where the comparisons are made between a number of independent predictors and an outcome.

4. I thought it was interesting that upward mobility was associated with higher allostatic load (in contrast to the expected direction). Could the authors suggest a reason for this finding?

Given that the mobility models did not fit the data well (as shown by the tests carried out in the analysis) we did not want to elaborate on this finding in the manuscript. We feel it is important to focus only on those models identified has having the best fit as this is the main focus of the paper.

5. “This study was the first to formally compare the different SEP lifecourse models for their association with allostatic load. Accumulated SEP was the strongest predictor of allostatic load, although a simple ‘summing of the risks’ may not encapsulate the entire nature of the association between allostatic load and SEP across the lifecourse.” What do these findings imply for future research and/or policy?

We have added the following to our conclusions at the end of the manuscript:

“This finding highlights that when considering the links between socioeconomic circumstances and allostatic load (and health more generally), we must consider the association across the lifecourse and not assume short-term interventions will have significant effects.”