## 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)** | Socioeconomic inequalities in functional somatic symptoms by social and material conditions at four life course periods in Sweden: a decomposition analysis |
| **AUTHORS** | San Sebastian, Miguel; Hammarström, Anne; Gustafsson, Per |

## VERSION 1 - REVIEW

| | |
|---|---|
| **REVIEWER** | Mohammad Hassan Emamian<br>Shahroud University of Medical Sciences, Shahroud, Iran |
| **REVIEW RETURNED** | 26-Sep-2014 |

| | |
|---|---|
| **GENERAL COMMENTS** | The reviewer also provided a marked copy with detailed comments. Please contact the publisher for full information about it. |

| | |
|---|---|
| **REVIEWER** | Bjorn Hallerod<br>Department of Sociology and Work Science<br>University of Gothenburg<br>Sweden |
| **REVIEW RETURNED** | 07-Jan-2015 |

| | |
|---|---|
| **GENERAL COMMENTS** | The paper tries to explain socioeconomic inequalities in health using a novel decomposition strategy and excellent longitudinal data. The central question is not if there exist a SES health gradient but why. Even though data are longitudinal they are treated as repeated cross-sections. By doing so the authors want to analyze if the same factors explains the SES gradient at different life phases. I am very sympathetic to the basic outline and I fully agree with the authors about the importance of this type of studies. However, I do think that the analysis suffers from a number of significant shortcomings.<br><br>1. The operationalization of SES is problematic. SES is dichotomized into low and high SES. The former includes manual workers and the latter non-manual workers and self-employed. For a start this is problematic. SES related differences are mainly driven by high skilled non-manual workers whereas the differences between low skilled non-manuals and skilled manuals often are negligible. A large group of self-employed are in fact doing manual work (self employed carpenters etc). A bigger problem is that SES at age 16 is based on parents' occupation, while SES at age 21, 30, and 42 is based on respondents' occupation. Thus, in the first cross-section the authors are analyzing the impact of class (SES) of origin on health, in the latter cross-sections the impact of individual SES but without acknowledge the fundamental difference between these two measures. The easy way out is of course to drop the age 16 |

| | |
|---|---|
| | analysis. To make it even more problematic SES at age 21 and 30 is based on a mix of occupation and education. To summarize; the aim is to understand SES related health differences but that cannot be achieved without a distinct and consistent measure of SES. This paper lacks such a measure.<br><br>2. Because different 'adversities' are used for the different cross-sections it is very hard, not to say impossible, to address 'to what degree the same determinants operates across the life course…..' (page 4). This has a number of serious consequences. The first is obvious. Why comparing four different cross-sections if they are not comparable? Apart from this there is a lack of theoretical considerations, i.e., what kinds of social mechanisms mediating the impact of SES are the empirical indicators indicating? Because there is no such reasoning the discussion and conclusion becomes a repetition of the results but without tapping into a more theoretical and explanatory phase.<br><br>3. Finally I think that the paper lacks references to longitudinal research in general and research on accumulation of advantages and disadvantages in particular.<br><br>For these reasons I cannot recommend the paper for publication. |

**VERSION 1 – AUTHOR RESPONSE**

Reviewer: 1
Reviewer Name Mohammad Hassan Emamian, MD, MPH, PhD
Institution and Country Shahroud University of Medical Sciences, Shahroud, Iran
Please state any competing interests or state 'None declared': None

This is an interesting study that seeks to assess socioeconomic inequalities in functional somatic symptoms. Overall, the topic is important; however, a number of issues may limit the value of the article, I hope below comments improve the quality of the paper.

Title

1. I find the title in its current state to be confusing. Socioeconomic inequalities in health, is a wide concept, while the authors try to decompose socioeconomic inequalities in functional somatic symptoms. I would consider editing for clarity's sake.

We agree with this comment. The title has been changed to: "Socioeconomic inequalities in functional somatic symptoms by social and material conditions at four life course periods in Sweden: a decomposition analysis"

Abstract

2. It does not adequately set up the rationale for the study.

Re-reading the abstract, we agree, and have therefore modified the abstract to better present the rationale:

" … but current research is lacking on key points, including determinants of socioeconomic differences in health, and not the least variations of these determinants over the life course."
3. The results section needs more explanation. What is the gap? You must explained the results

according to method of analysis (Oaxaca decomposition)

To clarify the results section in the abstract, we have revised it to more specific formulations, including adding explanation to the gap.

Inequalities in symptoms between blue- and white-collar socioeconomic groups increased along the life course, from 3.46 units at age 16 to 4.92 units at age 42. In the decomposition analysis, a high proportion of the health gap between socioeconomic groups could be explained by social and material living conditions…

4. The conclusion section must be according to results.

Yes, we agree with this comment. The following text has been added to re-direct the conclusion closer to the results, rather than looking forward.

"Socioeconomic inequalities in functional somatic symptoms increased along the life course in this Swedish cohort. A considerable portion of the social gaps in health was explained by concurrent social and material conditions, and the importance of specific adversities was dependent on the life course stage."

Introduction

5. Lines 48-55, there are many other studies regarding socioeconomic decomposition such as:

http://www.ncbi.nlm.nih.gov/pubmed/25068048

http://www.ncbi.nlm.nih.gov/pubmed/23636851

http://www.ncbi.nlm.nih.gov/pubmed/24979299

http://www.ncbi.nlm.nih.gov/pubmed/24171933

Although the references the reviewer provided concern other health and SES measures as well as other types of decomposition factors, we agree that they all are relevant for the present study and have therefore added them to the introduction.

There are also many other studies, decomposing socioeconomic inequalities using decomposition of Concentration Index.

We agree with the reviewer that there are other studies which decompose the socioeconomic inequality of health measured by the concentration index, but this is another conceptual and analytical approach to the one used in our research, which is not directly comparable to the one of the present study. Not to stray too far from the target, we therefore prefer to not go into that, admittedly substantial, literature, for this specific manuscript.

Methods

6. Page 6-7 – Why you dichotomized the adversities variables? You may use continuous or ordinal variables in analysis, leading to more precise results.

Adversity variables were either based on binary questionnaire items, or dichotomized partly in order to keep them a constant range of all adversities used in the analysis, partly to keep consistency because this is the way they have been using in other studies from our group. See for instance these references:
Gustafsson PE, Hammarström A. Socioeconomic disadvantage in adolescent women and metabolic syndrome in mid-adulthood: an examination of pathways of embodiment in the Northern Swedish Cohort. Soc Sci Med. 2012 May;74(10):1630-8
Gustafsson PE, Hammarström A, San Sebastian M. Cumulative contextual and individual disadvantages over the life course and adult functional somatic symptoms in Sweden. Eur J Public Health. 2014
Gustafsson, P.E., Janlert, U., Theorell, T., Westerlund, H., & Hammarstrom, A. Social and material adversity from adolescence to adulthood and allostatic load in middle-aged women and men: Results from the Northern Swedish Cohort. Annals of Behavioral Medicine 2012; 43, 117-128
7. Page 8 – The way of dividing population in two low and high SES is an important factor in decomposition analysis. Dividing the population to SES according to their occupation and level of education is not a correct method. Is there any study regarding this type of classification in Sweden?

Not sure if we understand this point. As mentioned in the article, we classified socioeconomic status by own occupation according to the classification scheme of Statistics Sweden, and dichotomized into manual workers (blue-collar) vs. non-manual employees and self-employed (white-collar) which is a standard procedure both in Sweden and internationally.
See for instance:
Rydwik E, Welmer AK, Angleman S, Fratiglioni L, Wang HX. Is midlife occupational physical activity related to disability in old age? The SNAC-Kungsholmen study. PLoS One 2013;8(7):e70471.
Morikawa Y, Martikainen P, Head J, Marmot M, Ishizaki M, Nakagawa H. A comparison of socio-economic differences in long-term sickness absence in a Japanese cohort and a British cohort of employed men. Eur J Public Health 2004;14(4):413-6.
We have also used this type of classification in our previous studies (see for instance those mentioned in the previous comment).

8. Page 9, line 19 – I appreciate the authors description of the Oaxaca decomposition model. As this technique might not be familiar to many readers, a bit more detail on why it was used and what information it provides would be useful. The following sentence needs more details.

“The latter, unexplained, component captures all potential effects of differences in unobserved variables”

The authors also do not present a true interpretation about negative and positive estimates in Oaxaca models.

Please refer to: O’Donnell O, van Doorslaer E, Wagstaff A et al (2008) Analyzing Health Equity Using Household Survey Data A Guide to Techniques and Their Implementation. The World Bank, Washington, DC

The concept of the unexplained component has bee now expanded. It reads:

“The latter, unexplained component, captures the part of the gap that remains unexplained after the health determinants are taken into account; it is seen as an indication of unequal treatment (discrimination) of the SES groups, and/or also differences in omitted determinants of health”

Not sure why the reviewer mentions that we do not present a true interpretation about the negative and positive estimates. In the text, we state:

"Decomposition estimates can be negative or positive; negative estimates indicate that the variable in question contributes to the inequality in the direction which runs counter to the overall inequality. Thus, a positive sign indicates that the gap in the variables favor the most advantaged group (high SES in our case)" which is backed by a reference (Stewart Willams, 2009).

9. page 9, line 36 – Before running the decomposition command (Oaxaca in STATA) the author must confirm the existence of an SES inequality in a multivariate regression model with SES as an independent variable. If SES does not have any significant effect in this model (this is probably true at age 16 in this study) then decomposition will not be necessary. (Please refer to above mentioned reference).

While we are aware that this is the strategy proposed by O`Donnell et al, their theoretical or conceptual arguments to follow these steps are unclear. The key point to conduct an Oaxaca decomposition analysis is that there is a health gap between a dichotomized socioeconomic variable and not the statistical association that there might between these two variables.

As the reviewer points there is not a significant association in the regression model between SES and FSS at age 16 but there is however a difference in health of 0.19 units. Should not be we interested in exploring which factors explain that difference because the regression model was insignificant?

Another related point is the justification to include variables in the Oaxaca analysis based on the regression results. This can be misleading. In our case at age 16, sex is significant but not parents separation in the regression model; in the Oaxaca analysis both variables explained the same amount of the difference (6.25% thought in opposite directions). Based on the regression analysis we would have missed the contribution of the variable separation.

10. Page 10, line 3 – When you find any interaction between variables, then repeating the analysis for subgroups is recommended. The author does not report any interaction between gender and other variables (this is possible in Oaxaca command), therefore they must capture the role of gender simply by adding this variable to decomposition modes.

Again, this line of argument represents a particular strategy of data analysis. We prefer to use stratification based on conceptual arguments more than statistical ones, i.e. a theory-driven rather than a data-driven approach. Based on our previous research, we know that there are strong differences in socioeconomic inequalities in health by gender. This clarification has been added to the text.

"Given that our previous research has shown significant socioeconomic differences in health between men and women 37,38, analyses were also performed stratified by gender."

As far as we know the Oaxaca command from Stata can calculate interaction terms measuring the simultaneous effect of differences in endowments and coefficients but it does not provide an indication among which variables the interaction exists.

Results

11. Please add 95% CI in decomposition tables.

We can do this though we would prefer not since the tables are already so dense in terms of the amount of numbers (it is unusual to present decomposition analysis along the life course). We have made efforts to present the results in a way as concise as possible given the amount of point estimates which are completely necessary to report for our research questions. Confidence intervals would not add anything qualitatively new to the inferences and the interpretation of the results. But if the reviewer or editor considers appropriate to include them even in the light of the density issue, we can of course easily do it.

12. Surprisingly the authors omitted the independents variables in unexplained parts and report only constant and total share. Please report the results completely and explain about data in results and discussion section. Also in explained part many variables have missing values!

The information required by the reviewer is provided by the command Oaxaca in Stata and therefore this could be done. At the same time, not all information provided by Stata commands are always necessary. In this case, we consider that i) the interpretation of the specific contribution of each variable to the unexplained part is limited (is discrimination or differences in the effects (betas) of the variables?); ii) we do not think that information adds any relevance to the key research question of this study which is to measure the part of the gap in mean health status due to differences in specific variables; iii) a less important reason can be related to the clarity of the table. Given the amount of explanatory variables included in this study, it would double the size of the tables, and as noted above, we have attempted to select only the most important estimates in order to make the results section more accessible for a broader audience.

Nevertheless, if the reviewer or editor would insist on reporting the unexplained part in addition to the explained ones, we are happy to comply.

13. Prospective design is considered as strength of this study, but the author analyzed data as four cross-sectional studies. They must benefit from more appropriated analysis for repeated measures data such as GEE models. This is especially important for causal inference.

GEE models or repeated measurements analyses are not very helpful for the research question of this study: to decompose differences in health according to socioeconomic status. This would completely change the aims of the manuscript. We have already used this principal approach (more specifically, by longitudinal multilevel models) in a previous publication but answering a complete different research question. See for instance the reference below.

Gustafsson PE, San Sebastian M. When does hardship matter for health? Neighborhood and individual disadvantages and functional somatic symptoms from adolescence to mid-life in The Northern Swedish Cohort. PLoS One 2014;9(6):e99558.

The prospective design is nevertheless an important methodological strength even when performing repeated cross-sectional analyses, since it addresses the wealth of potential confounders which would arise e.g. if the analysis would have been on an age-heterogeneous cross-sectional sample stratified by age. By using prospective data in the present manuscript, we remove many confounders simply owing to the age groups being different individuals differing in many other respects than age. This methodological strength has been clarified in the methodological section of the manuscript:

Compared to a cross-sectional study, the prospective design means that many confounders are addressed by design, specifically those that would produce biased results in different age groups solely because of the groups comprising different individuals who differ in many other respects than being of different ages.

Discussion and conclusion

14. These section must completely rewrite after repeating the analysis.

Since we have argued that a reanalysis of the data is not necessary, the discussion and conclusion have not been changed.

Reviewer: 2
Reviewer Name Bjorn Hallerod
Institution and Country Department of Sociology and Work Science
University of Gothenburg
Sweden
Please state any competing interests or state 'None declared': Non declared

The paper tries to explain socioeconomic inequalities in health using a novel decomposition strategy and excellent longitudinal data. The central question is not if there exist a SES health gradient but why. Even though data are longitudinal they are treated as repeated cross-sections. By doing so the authors want to analyze if the same factors explains the SES gradient at different life phases. I am very sympathetic to the basic outline and I fully agree with the authors about the importance of this type of studies. However, I do think that the analysis suffers from a number of significant shortcomings.

1. The operationalization of SES is problematic. SES is dichotomized into low and high SES. The former includes manual workers and the latter non-manual workers and self-employed. For a start this is problematic. SES related differences are mainly driven by high skilled non-manual workers whereas the differences between low skilled non-manuals and skilled manuals often are negligible. A large group of self-employed are in fact doing manual work (self employed carpenters etc). A bigger problem is that SES at age 16 is based on parents' occupation, while SES at age 21, 30, and 42 is based on respondents' occupation. Thus, in the first cross-section the authors are analyzing the impact of class (SES) of origin on health, in the latter cross-sections the impact of individual SES but without acknowledge the fundamental difference between these two measures. The easy way out is of course to drop the age 16 analysis. To make it even more problematic SES at age 21 and 30 is based on a mix of occupation and education. To summarize; the aim is to understand SES related health differences but that cannot be achieved without a distinct and consistent measure of SES. This paper lacks such a measure.

As the reviewer knows, the manual/non-manual SES distinction is well established in social epidemiological research, both in Sweden and internationally. The decomposition analysis requires a dichotomization of the exposure variable. In this case, we took the decision to dichotomize between manual vs non-manual(+self-employed) to follow some of the epidemiological tradition and to be coherent with our previous research. See for instance: Gustafsson PE, Hammarström A. Socioeconomic disadvantage in adolescent women and metabolic syndrome in mid-adulthood: an examination of pathways of embodiment in the Northern Swedish Cohort. Soc Sci Med 2012;74(10):1630-8.
Other kind of dichotomizations, as the reviewer suggests, could have been done. However we would fall in the same problem. For instance, self-employed is a very heterogeneous group, and to include them in the non-manuals is also questionable since they could be earning much more than manual workers and therefore be very different groups. To have only the low skilled non-manual as the exposed group would not be possible due to the low sample size.

The reviewer sees a problem in the use of SES in different ways at different age periods. But this is an established integral approach of the entire epidemiological subfield labeled life course epidemiology (see for instance Galobardes et al. Measuring socioeconomic position in health research. Br Med Bull. 2007;81-82:21-37). While we agree that the measures are indeed very different since one is representing your own SES and the other your parents', we consider the measure differences are inherent to the life course period under study – there is no directly corresponding measure of your own occupational class in adolescence (although some school variable, such as gymnasium program would be possible and is tried out e.g. by some Swedish researchers such as Kurt Hagqvist). In this manuscript, we do want to tie in to the existing social and life course epidemiological literature, and here parental SES is by a far margin the most common measure of childhood/adolescent SES. We have added a sentence in the methodological considerations.

While we are aware that the SEP of the parents might not really represent that of the children, it is common in social epidemiology to use individual occupational measures to characterize the SEP of others connected to them 57.

By removing the age 16 analyses, as the reviewer seem to suggest, we would limit the scope of the manuscript substantially by only considering the adult life course. To us, the fact that we do indeed have a measure point during adolescence is one of the key strengths of this study.

The use of concurrent highest education as a proxy for SES for some cases missing information on occupation at age 21 and 30, we agree, is not an optimal situation, but was weighed against the option of excluding the cases with missing data, with the resulting lesser statistical power, and even more importantly, an imminent risk for introduction of selection bias. Nevertheless, this has added as a specific methodological issue in the revised methodological discussion.

There also seem to have arisen some misunderstandings about the purpose of decomposition analysis. It is not that we are analyzing "the impact of class (SES) of origin on health" but what factors might be explaining a health difference that is observed between two socioeconomic groups at four different periods of life, as stated in the manuscript.

In the revised manuscript, we have tried to provide a more accessible description of the decomposition analysis in the statistics section.

2. Because different 'adversities' are used for the different cross-sections it is very hard, not to say impossible, to address 'to what degree the same determinants operates across the life course…..' (page 4). This has a number of serious consequences. The first is obvious. Why comparing four different cross-sections if they are not comparable? Apart from this there is a lack of theoretical considerations, i.e., what kinds of social mechanisms mediating the impact of SES are the empirical indicators indicating? Because there is no such reasoning the discussion and conclusion becomes a repetition of the results but without tapping into a more theoretical and explanatory phase.

We agree that the word "same" should not be used and has been substituted for "similar or different". However, we disagree with the main point of the reviewer. Neither measures of socioeconomic status nor adversities can look the same across the life course because life is different, and also the practical challenges of preserving decades-long cohorts; this is standard issue in the field of life course epidemiology. And this issue goes deeper than that; even events or conditions that operationally are identical do carry different meanings at different periods of the life course. Consider the death of parents during childhood vs. in mid-adulthood; one would be considered a traumatic event, the other a normative life event in the sense that most people do experience it, and adapt to it successfully.

Nevertheless, these conceptual, methodological and practical complexities are not, in our opinion, a reason to avoid empirical examinations which include some kind of comparison between different life course periods. While we agree with the reviewer that the age groups are not directly comparable, several issues are indeed kept fairly constant, such as the health outcome measure and not the least the sample by virtue of the prospective design. With this as a background, we see no lack in stringency with reporting to what degree observed adversities explain the SE gap in health at different life course periods.

This study falls under the theoretical frameworks of the social determinants of health inequalities and the life course perspective as presented in the introduction. In order to keep the paper's 'readability', we decided not to focus in-depth in theoretical frameworks specially when there are many and complex results that needs to be at last commented upon. In this case, it would be however difficult to discuss in the line of the reviewer (social mechanisms mediating the impact of SES) because of the restriction imposed by the decomposition analysis itself. Other researchers in our group are doing for instance structural equation modeling to disentangle the role of mediators.

To our knowledge, the present study is the first of its kind directly examining the explanatory factors for SE inequalities in health in a prospective cohort spanning over decades, and therefore we want to keep a tempered discussion about the results. We also appreciate that writing styles may vary across disciplines, and that within epidemiology a concise discussion fairly close to the results is common practice.

3. Finally I think that the paper lacks references to longitudinal research in general and research on accumulation of advantages and disadvantages in particular.

As we understand it, the present study is neither framed under a “strict” life course approach which usually requires longitudinal analysis, nor focuses on different explanatory models (such as the accumulation mentioned by the reviewer) over the life course. In the revised manuscript we have now added a few references related to the life course approach.

Whereas a considerable body of research focusing on social determinants of health has adopted a life course approach 18-19,…

For the empirical application and literature on such kind of approaches, we also refer to our previous publications:

Gustafsson PE, Hammarström A, San Sebastian M. Cumulative contextual and individual disadvantages over the life course and adult functional somatic symptoms in Sweden. Eur J Public Health 2014; pii: cku213.

Gustafsson PE, San Sebastian M. When does hardship matter for health? Neighborhood and individual disadvantages and functional somatic symptoms from adolescence to mid-life in The Northern Swedish Cohort. PLoS One 2014;9(6):e99558

Gustafsson PE, San Sebastian M, Janlert U, Theorell T, Westerlund H, Hammarström A. Life-course accumulation of neighborhood disadvantage and allostatic load: empirical integration of three social determinants of health frameworks. Am J Public Health 2014;104(5):904-10.

## VERSION 2 – REVIEW

| | |
|---|---|
| **REVIEWER** | Michael Gamborg<br>Institute of Preventive Medicine/Institut for Sygdomsforebyggelse |

| | |
|---|---|
| | Frederiksberg Hospital<br>Hovedvejen<br>Nordre Fasanvej<br>Frederiksberg |
| **REVIEW RETURNED** | 22-May-2015 |

| | |
|---|---|
| **GENERAL COMMENTS** | I find the paper well writen, but it needs little attention to the statistical details.<br><br>The reviewer also provided a marked copy with detailed comments. Please contact the publisher for full information about it. |

**VERSION 2 – AUTHOR RESPONSE**

Reviewer: 1
Reviewer Name: Michael Gamborg
Institution: Institute of Preventive Medicine/Institut for Sygdomsforebyggelse

Abstract
You mention that the inequalities increase by from 3.46 to 4.92, in which unit?

Thanks for pointing this issue, it was a mistake from our side. The numbers provided (3.46 to 4.92) are the means of the FSS outcome, not of the gap. We have now deleted the numbers from the sentence.

The FSS measure is a summative score of ten Likert scales, as described in the methods. As such, it is not expressed in a meaningful unit of measurement but only as a score.

And to claim that the difference increase, but there are no analysis done to support this claim.

In this sentence we refer that the gap in FSS between low and high SES increased in the sample over time. This can be observed in Table 2 where it goes from 0.19 at age 16 to 0.94 at age 43, and we do have specifically tested whether the gap is significant at each age. The reviewer is entirely correct in that the particular claim of increased gaps is not specifically tested, but based on the numerical increase in the mean differences in combination with the test whether the gaps where significant. The point with this sentence was not to make inferential claims, but to describe the gaps seen in the data – the gaps which we then decompose with inferential statistics.

In order to make this clearer, we have added “numerically” or “in the sample” in the text when we refer to these patterns in the sample, to distinguish it from stochastically founded and inferential claims. We have also reworded the discussion regarding this in softer ways to reflect that this particular claim is not inferentially supported.
You talk about mental health in the abstract, but your outcome is FSS, does FSS measure mental health?

We see different authors emphasizing and describing the (causal and conceptual) relation between FSS, somatization and mental health differently. FSS has been used in the literature as a subjective proxy measure of mental health. FSS in adolescence has been shown to be prognostic of various mental disorders in adulthood (Campo, 2012; Bhoman et al, 2012) and are closely interrelated with anxiety and depressive symptoms (Janssens et al, 2010)

However, we agree with the reviewer that it is indeed an undue simplification or even misconstruction to simply categorize them as mental health problems. While we feel that a conceptual/terminological discussion is outside the scope of this empirically oriented ms, in order to address the reviewer's concern and to avoid confusions, we have eliminated the term mental from the abstract.

References

Campo JV (2012) Annual research review: functional somatic symptoms and associated anxiety and depression—developmental psychopathology in pediatric practice. J Child Psychol Psychiatry 53: 575–592.
Bohman H, Jonsson U, Paaren A, von Knorring L, Olsson G, et al. (2012) Prognostic significance of functional somatic symptoms in adolescence: a 15-year community-based follow-up study of adolescents with depression compared with healthy peers. BMC Psychiatry 12: 90.
Janssens KA, Rosmalen JG, Ormel J, van Oort FV, Oldehinkel AJ (2010) Anxiety and depression are risk factors rather than consequences of functionalsomatic symptoms in a general population of adolescents: the TRAILS study. J Child Psychol Psychiatry 51: 304–312.

Analysis

You use chi-sq test in the tables that needs to be mentioned in the methods.

Thanks for the observation. This has now been mentioned in the Methods section (page 9). The sentence reads now:

Descriptive statistics of the health outcome and the different variables by SES were carried out. The t test for the mean FSS and the chi square test for proportions to assess statistical differences were also applied.

You do not mention how you compare the mean FSS between SES groups, but I assume it is t-test, it needs to be described, and the actual p – values are also interesting, or even better the confidence intervals for the difference.

See above. Modifications in Table 1 have also been done.

If I am not mistaken then you do assume that FSS is normally distributed, even though it takes the values 0, 1, 2,.., 26. That seems not to be a reasonable assumption. I suspect that the distribution does have a lot of 0's and is skewed. I could be wrong, but as I comment on later the tables does not help me much.

The FSS measure does indeed display a positively skewed distribution, similarly at each age. However, it does have a peak >0 (i.e. the floor effect does not seem to include the population mean/median), with a gradual slope without additional peaks, a fairly large range with a underlying continuous distribution, and most importantly, no outliers. As such, while the measure indeed looks like most "problem scales" based on Likert items and as such is not normally distributed, it is certainly not a strong deviation. With the exception of extreme outliers, double peaks or such severe deviations from the normal distribution, this kind of scales are used routinely with parametric methods, since they are considered rather robust to deviations of assumptions. Moreover, to our knowledge, there is no suitable non-parametric alternative to Oaxaca-decomposition.

However, we agree with the principal point of the reviewer, and have therefore added the following in the limitations section of the ms (page 13):

The health outcome (FSS) measures were slightly skewed, which would indicate a deviation from a parametric assumption. Therefore, the precise estimates should be interpreted with caution.

In the presentation of your results there is too little focus on the precision of your results, tables 2-4 needs confidence intervals, so that the reader can get an idea about the accuracy of your results.

Confidence intervals have been added to Tables 2-4.

I do have a problem with doing a decomposition analysis of a difference that might not even be there. The difference in FSS among the 16 year olds is not significant, so there might actually not be any difference, and then it does not make sense to decompose it. It also means that the denominators in the relative measures are small making the results unstable. But that will probably be visible looking at the CIs.

We see the point of the reviewer which we acknowledge as very relevant within classical regression frameworks. The key point to conduct an Oaxaca decomposition analysis is however that there is a health gap between a dichotomized socioeconomic variable and not the statistical association that there might between these two variables.

As the reviewer points there is not a significant association in the regression model between SES and FSS at age 16, indicating that we cannot be confident that the sample difference reflect a true population difference. There is however a sample difference in health of 0.19 units, and by performing the decomposition analysis on this sample difference we simply ask which factors might explain that observed difference. Not performing the decomposition of the non-significant gap would run the risk of type 2 error, and could be problematic e.g. in the case of future meta-analyses where even non-significant estimates are of relevance. Therefore, we rather keep the results as they are. However, we have revised the results and discussion to clearly reflect that the difference was not significant, and that the decomposition analysis therefore should be interpreted very cautiously.

In order to explain these health gaps, decomposition analyses were run for each of the four ages; for completeness sake also at age 16 despite this health gap being small and non-significant and therefore should be interpreted cautiously (page 10)

While the results should be interpreted cautiously in light of the non-significant health gap, our findings at age 16… (page 13)

Another related point is the justification to include variables in the Oaxaca analysis based on the regression results. This can be misleading. In our case at age 16, sex is significant but not parents separation in the regression model; in the Oaxaca analysis both variables explained the same amount of the difference (6.25% though in opposite directions). Based on the regression analysis we would have missed the contribution of the variable separation.

This dataset could be viewed as one dataset with repeated measures of outcome and exposures, or as 4 separate datasets, could you please reflect a little over the consequences of ignoring the repeated measures aspect of the data structure in your analysis.

Though as the reviewer mentions, the original database has a repeated structure because it is a cohort, in our study the data were analysed in a cross sectional way. To our knowledge, it is not possible to use the Oaxaca decomposition technique in a repeated measure design. As we see it, the main strength of performing these repeated c/s analyses on a prospective cohort is that many potential inter-individual sources of confounding and selection bias, which would be very manifest if we were to perform the same analyses on one cross-sectional sample stratified by age, are controlled by design, by virtue of the cohort data.

Discussion

You repeat the claim that the difference in FSS increases, but to do so you need to make a statistical analysis of this claimed increase, so that you make sure that it is statistical significant (it is not possible to ignore the repeated structure of the data when doing that).

This has been addressed now in the analysis and results section.

Tables:

Units need to be mentioned, I assume that table 1 is using percent.

Table 1 is presenting the units at the end of the heading (%). Mean and standard deviations have also been added for the health outcome.

The only representation of your outcome is the mean by gender and exposure, not even a standard deviation or range, which need to be improved.

Standard deviations have now been added to Table 1, and the ranges to the methods section on page 6.

An most importantly, tables 2-4 needs confidence intervals, to support your claim “that a considerable portion of the social gaps in health are explained…” that might not be true if the CIs gets close to zero.

Confidence intervals have been added to Tables 2-4.