PEER REVIEW HISTORY

BMJ Paediatrics Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

This paper was submitted to a another journal from BMJ but declined for publication following peer review. The authors addressed the reviewers' comments and submitted the revised paper to BMJ Paediatrics Open. The paper was subsequently accepted for publication at BMJ Paediatrics Open.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Sex Differences in Child and Adolescent Physical Morbidity: Cohort Study
AUTHORS	Sweeting, Helen; Whitley, Elise; Teyhan, Alison; Hunt, Kate

VERSION 1 - REVIEW

REVIEWER	Carsley, Sarah The Hospital for Sick Children, Canada Competing interests: None to declare
REVIEW RETURNED	18-Sep-2017

GENERAL COMMENTS	This study looks at the sex differences of 32 health states and conditions from questions asked for the ALSPAC cohort study in participants aged 4 to 13 years. A sex-age interaction was also used to determine any changing patterns as participants aged. This is a well-written study with clear objectives and results section. The main clarification would be to define the patterns of 'type 1' emerging/increasing 'excess' with more numeric precision. Emerging and increasing were separated in the figures, but addressed in the same category in the text. Introduction The introduction was clear and objectives well described. Methods The definition of 'type 1' emerging/increasing 'excess' is confusing as they are defined together and then presented in the results separately. How did the authors distinguish between what is considered emerging and what is increasing? There was no mention of including other covariates and other potential confounding factors. It would be interesting to see the description of the study population included for readers to assess the generalizability of these findings. Statistical Analysis Given the large number of analyses, some of the statistically significant associations with sex and age might just have been spurious. This should be acknowledged in the discussion of limitations. Did the authors attempt to perform more sophisticated analyses of their longitudinal data using trajectory analyses/latent growth modeling/mixed modeling? A discussion of why it wasn't used could be addressed especially since the authors acknowledge their chosen method may be crude. Why did the authors use robust standard errors to adjust for non-independence of observations instead of a generalized estimating equation or mixed model?

	Results Any other descriptors of the study population? I liked the presentation of results in the figures rather than the tables. It gave a great visual to understand better what was emerging/increasing, consistent, no difference. I still would like to have a more precise description of emerging vs. increasing since they are separated in the results. One of the objectives in the introduction was to determine "when any emerging female 'excess' occurs" however other than reporting baseline (at 57 months) and final odds ratios at 166 months, did not feel like this question was answered. Discussion The authors mention the need for future research to determine underlying mechanisms of these sex differences. I would suggest an emphasis on qualitative research here or mixed methods because they already discussed stereotypes, attitudes and expectations based on sex as possible explanations for the differences. Quantitative research may not be able to elucidate some of the next research questions if they are focused on understanding these mechanisms more fully. What are the implications for policy/practice or public health?

REVIEWER	Bagkeris, Emmanouil University College London, UK Competing interests: No competing interest
REVIEW RETURNED	05-Oct-2017

GENERAL COMMENTS	Dear authors, this is a very well written paper. I hope that my comment below will help you improve the quality of this manuscript. 1. Since the main analysis considers the data cross-sectionally, please consider providing incidence risk ratios rather than odds ratios. Please advise and reference Barros AJ, Hirakata VN. Alternatives for logistic regression in cross-sectional studies: an empirical comparison of models that directly estimate the prevalence ratio. BMC Med Res Methodol 2003; 3: 21. 2. There are nineteen symptoms and 10 infections compared between female and male children in figure 2. Have you considered correcting for multiple comparison? Will the results be the same?

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Summary
This study looks at the sex differences of 32 health states and conditions from questions asked for the ALSPAC cohort study in participants aged 4 to 13 years. A sex-age interaction was also used to determine any changing patterns as participants aged. This is a well-written study with clear objectives and results section.

REVIEWER 1, POINT 1
The main clarification would be to define the patterns of 'type 1' emerging/increasing 'excess' with more numeric precision. Emerging and increasing were separated in the figures, but addressed in the same category in the text.

AND Methods - The definition of ‘type 1’ emerging/increasing ‘excess’ is confusing as they are defined together and then presented in the results separately. How did the authors distinguish between what is considered emerging and what is increasing?

RESPONSE: We agree that we did not always explain emerging/increasing as clearly as we could have. We now more clearly distinguish and describe Types 1a (emerging female excess), 1b (increasing female excess) and 1c (disappearing male excess) in our Methods, also clearly labelling the stable patterns as 2a, 2b and 2c and noting that male excess patterns mirror those of females. We now also include the precise criteria used to define the various patterns as footnotes to the graphs. In our Results we have made sure to separate ‘emerging’ and ’increasing’ in respect of descriptions of patterns found for each measure, and our graphs and Table 2 (Summary of patterns) now all include ‘1a’, ‘1b’, etc as part of the heading for each pattern, so they can be linked to the description in the Methods.

Thus, the relevant section of our methods now reads:
“… patterns of sex-by-age differences, which a previous systematic review conceptualised in terms of four ‘types’,[ref] and are defined here as:
• ‘Type 1’: an emerging/increasing female ‘excess’, or disappearing male excess with age, occurring because female rates increase more than those of males or decrease less than those of males, resulting in a marked sex-by-age interaction. Type 1 patterns therefore include: (a) a male ‘excess’ reversing to a female ‘excess’ or no sex difference at younger ages, but a female ‘excess’ at older ages (emerging female ‘excess’); (b) a female ‘excess’ at younger ages, increasing with age (increasing female ‘excess’); or (c) a male ‘excess’ at younger ages, but no sex difference at older ages (disappearing male ‘excess’). For ‘Type 1’ patterns, the odds of morbidity among females compared to males start below, at, or above unity and increase with age.
• ‘Type 2’: (a) stable female ‘excess’; (b) stable lack of a sex difference; or (c) stable male ‘excess’. For ‘Type 2’ patterns, the odds of female versus male morbidity are consistently either above, below or at unity
• ‘Type 3’: variations on an emerging/increasing male ‘excess’, or disappearing female ‘excess’ (the reverse of ‘Type 1’). For ‘Type 3’ patterns, the odds of female versus male morbidity start above, at, or below unity and decrease with age.
• ‘Type 4’: mixed/unclassifiable patterns.
(Note precise definitions are included as footnotes to the results graphs.)”

Examples of the precise criteria used to define the various patterns now included as footnotes to the praphs include:
“1a Emerging female excess defined as: odds ratios (ORs) at youngest age(s) show a marked male excess or no sex difference; ORs at oldest age(s) show a marked female excess; sex-by-age interaction is marked.
2a Consistent female excess defined as: majority of ORs show a marked female excess; sex-by-age interaction is not marked.”

REVIEWER 1, POINT 2
There was no mention of including other covariates and other potential confounding factors. It would be interesting to see the description of the study population included for readers to assess the generalizability of these findings.

RESPONSE: We took the decision not to adjust for confounders as we were interested in how morbidity rates varied with age and sex irrespective of socio-economic status or other potential confounders. We acknowledge it could be interesting to include these in further analyses, as a separate paper, to explore whether they might help account for any of the changing rates observed

here (if, for example, socio-economic status was associated with a morbidity measure in different ways for younger/older males/females).

We agree that it would be useful for readers to have access to a basic description of the study population, although word limits mean this must be brief. We now therefore include an additional supplementary table (new Supplementary table 2) which describes the characteristics of those included/not included in the analytical samples. In respect of generalisability, our comment on this table within our Methods reads: “Supplementary Table 2 shows the characteristics of those included/not included in the samples and demonstrates that due to sample attrition, those included were more likely to be first-born children of mothers in a first marriage, of higher socio-economic status, and who had never smoked”. We also include a sentence on this within our study limitations as follows: “Further, although there was little evidence of differential attrition according to child sex, the sample was more advantaged than the general population, thus potentially limiting generalisability of the findings”.

REVIEWER 1, POINT 3
Given the large number of analyses, some of the statistically significant associations with sex and age might just have been spurious. This should be acknowledged in the discussion of limitations.

RESPONSE: We now acknowledge this (but also note our focus on patterns, and the predominance of one type) in our limitations section, as follows:
“… and the fact that we conducted analyses on 32 morbidity measures introduces the possibility of spurious (chance) significance for some sex-by-age interactions due to multiple testing. However, our focus was on consistent patterns of odds ratios with age and these are unlikely to have arisen purely by chance. In addition, there is no reason to think that spurious interactions would occur more often for measures showing a ‘Type 1’ pattern. ”

REVIEWER 1, POINT 4
Did the authors attempt to perform more sophisticated analyses of their longitudinal data using trajectory analyses/latent growth modeling/mixed modeling? A discussion of why it wasn’t used could be addressed especially since the authors acknowledge their chosen method may be crude.

RESPONSE: We believe that one of the main strengths of these analyses is their simplicity and the fact that we were able to present our results in simple graphical form, which this reviewer kindly described as a “great visual to understand better what was emerging/increasing, consistent, no difference”. We think the benefits of this in making our results accessible to a general readership outweigh any which might ensue from more complex analyses which would be unlikely to change our conclusions. In retrospect, our phrase “Our categorisation of patterns of sex-by-age differences … could be regarded as crude” was not the best way to describe our own analyses, and we have replaced this with “could be regarded as simple”.

REVIEWER 1, POINT 5
Why did the authors use robust standard errors to adjust for non-independence of observations instead of a generalized estimating equation or mixed model?

RESPONSE: We would make the same points here, around the benefits of keeping results simple and accessible. In addition, we explored the impact of non-independence of observations and found this to be minimal.

REVIEWER 1, POINT 6
Any other descriptors of the study population?

RESPONSE: As above (point 2), we now include a table describing those included/excluded and refer to this in both our Methods and Discussion.

REVIEWER 1, POINT 7
I liked the presentation of results in the figures rather than the tables. It gave a great visual to understand better what was emerging/increasing, consistent, no difference. I still would like to have a more precise description of emerging vs. increasing since they are separated in the results.

RESPONSE: Thank you! And as above (point 1) we now clearly separate emerging and increasing in the results in respect of each measure (although still using 'emerging/increasing' as a shorthand in the initial summary descriptions).

REVIEWER 1, POINT 8
One of the objectives in the introduction was to determine "when any emerging female 'excess' occurs" however other than reporting baseline (at 57 months) and final odds ratios at 166 months, did not feel like this question was answered.

RESPONSE: We now note the age by which each 'emerging excess' is markedly above/below unity (due to the scale of the graphs this is more clearly available via the supplementary tables). The relevant sentences within our Results now read as follows:
"… a 'Type 1a' female 'excess' emerging by 128 months for parent-reported poor general health in their child over the last month … and the last year …
… a female 'excess' emerged in respect of rates of parent-reported high temperature by 128 months and, more markedly, rash by 81 months …
… no sex difference in respect of pain in arms/legs at younger ages … but a small male 'excess' had emerged by 140 months …
… but a female 'excess' emerged for eye infections by 103 months and ear infections by 81 months …"

REVIEWER 1, POINT 9
The authors mention the need for future research to determine underlying mechanisms of these sex differences. I would suggest an emphasis on qualitative research here or mixed methods because they already discussed stereotypes, attitudes and expectations based on sex as possible explanations for the differences. Quantitative research may not be able to elucidate some of the next research questions if they are focused on understanding these mechanisms more fully.

RESPONSE: We have removed "quantitative" from the relevant sentence in our final paragraph, and this now reads "There is also a need for further qualitative or experimental studies to examine the social and/or biological mechanisms underlying these findings[refs]".

REVIEWER 1, POINT 10
What are the implications for policy/practice or public health?

RESPONSE: We have now added the following to our final paragraph: "This pattern of 'excess'

female morbidity by puberty highlights important inequities, with public health implications, not least in future health service usage". We have also expanded on our second 'What this study adds' point so this now includes the need for further studies "and understand longer term implications for sex differences in adult health".

Reviewer: 2

Dear authors, this is a very well written paper. I hope that my comment below will help you improve the quality of this manuscript.

REVIEWER 2, POINT 1
Since the main analysis considers the data cross-sectionally, please consider providing incidence risk ratios rather than odds ratios. Please advise and reference Barros AJ, Hirakata VN. Alternatives for logistic regression in cross-sectional studies: an empirical comparison of models that directly estimate the prevalence ratio. BMC Med Res Methodol 2003; 3: 21.

RESPONSE: We would like to thank the reviewer for bringing this paper to our attention. The paper suggests a number of alternatives to logistic regression, and highlights "the most common mistake associated with odds ratios in our experience: the authors "forget" what their measure of association is and make interpretations such as "the exposed group has a risk of illness four times greater than the non-exposed group"." We have been careful in our Results not to use the term 'risk', and report only exact odds ratios. Given this, we believe that logistic regression is not inappropriate and that a general audience would be most comfortable with seeing odds ratios in the context of a binary outcome variable, as we have here. However, given the points made in the Barros paper, we have now repeated our analyses (cross-sectional sample) using Poisson regression, and new Supplementary tables 6-8 show these results (risk ratios and p for sex by age interaction) in comparison with those from the logistic regressions (odds ratios with p for sex by age interaction) and can confirm that our conclusions are unchanged. We refer to these additional analyses in our Methods section as follows:
"Finally, given that some have suggested Poisson as an alternative to logistic regression for analysis of cross-sectional studies with binary outcomes,[ref] we also conducted Poisson regression analyses on the cross-sectional samples. Results (expressed as risk ratios, rather than odds ratios) were very similar to those obtained via logistic regression (see Supplementary Tables 6-8)."

REVIEWER 2, POINT 2
There are nineteen symptoms and 10 infections compared between female and male children in figure 2. Have you considered correcting for multiple comparison? Will the results be the same?

RESPONSE: This was also picked up by Reviewer 1 (point 3) and, as noted above is now acknowledged in our limitations section