Author’s response to reviews

Title: Different analysis methods of Scottish and English child physical activity data explain the majority of the difference between the national prevalence estimates

Authors:

Chloë Williamson (s1231185@sms.ed.ac.uk)
Paul Kelly (p.kelly@ed.ac.uk)
Tessa Strain (tessa.strain@mrc-epid.cam.ac.uk)

Version: 1 Date: 24 Jan 2019

Author’s response to reviews:

Response to review: Manuscript PUBH-D-18-03635 “Different analysis methods of Scottish and English child physical activity data explain the majority of the difference between the national prevalence estimates”

23rd January 2019

Dear BMC Public Health Editorial Board

We thank the editor and reviewers for the opportunity to respond to reviewer comments. Below we have provided a point-by-point response to comments provided by reviewers 1 and 2 and have detailed subsequent changes that have been made to the manuscript.

Chloë Williamson (on behalf of all authors)

Reviewer 1

Comment: The authors elegantly assessed how great an influence the difference in analysis methods between Scotland and England has for child physical activity prevalence estimates.
They concluded that the difference in the analysis method explains the majority of the difference in the child physical activity prevalence estimates between Scotland and England, thereby undermining comparability between the UK home nations. The paper addresses an important issue in the field and is well executed. I recommend publication of this article as is. The authors are nevertheless encouraged to polish the English language before publication (typos etc).

Response: Following your advice, we have conducted a thorough proof read of our revised manuscript. Corrected typos are listed at the end of this letter.

Reviewer 2

Comment: Overall Impression: This paper is simplistic in its analysis and messaging but nonetheless stresses a very important point. It is unfortunate that comparisons have been made in the past between datasets that are very clearly not comparable. Perhaps the publication of this paper will help to raise awareness of this issue.

Response: We also hope this issue receives more attention in the future. Hopefully this paper will help.

General Comments from Reviewer 2:

1. Why is the word "minimum" included in the naming of the "daily minimum method"? Does it add information beyond the simpler "daily method"?

Response: We discussed this issue between the authors and thank the reviewer for raising it. We discussed a number of alternate naming options, and all had pros and cons. After considering this in detail we believe that on balance including the word ‘minimum’ is the most appropriate option and is helpful in the method name as it concisely highlights the differences between the two interpretations of the guideline. We feel that this would be helpful in situations where our study is summarised or figures published without further context. We believe there is the potential for
these situations to arise as we know that this paper will be of interest to policy-makers and those involved in the nation health surveys’ design. Therefore, we would prefer to keep using the term “Daily Minimum Method” and abbreviation DMM.

2. From a purely mathematical perspective, estimates using the weekly average will always be higher than when using the daily method because it is simply a less strict mathematical approach. This is insinuated throughout but it may be worth explicitly stating it as such somewhere in the paper. This comment relates a little to Figure 1. In other words one should not be surprised that the values using the weekly average method are a lot higher compared to the daily method.

Response: We agree that it improves the manuscript to state this explicitly and we have added a sentence with a description of the DMM being a stricter version of the WAM (page 5, line 108).

In addition, we realise from this comment as well as Comment 5 that we need to clarify the methods better in certain areas. To address this, we have made the following amendments to the relevant paragraph in the introduction following Table 1 (beginning on page 5, line 95):

- We have made the assumption that sessions = days more explicit
- We have explained why it is necessary to make this assumption

The section used to read:

“Given that the SHeS and HSE questionnaires are so similar, yet produce such different prevalence estimates, we investigated how the data were processed. A subtle but potentially important difference between the SHeS and HSE surveys is how the frequency of activities are reported, which has implications for how the data are processed. In the SHeS, children or their parents are asked to report how many sessions of each activity they have done in the previous seven days [12]. In the HSE, activities are reported for specific days in the previous seven [14]. It is therefore not possible to ascertain whether child respondents to the SHeS achieve ≥60 minutes of MVPA on each specific day. Instead, one can calculate whether a child has reported ≥7
sessions, and whether the total weekly duration is equivalent to an average ≥60 minutes per day. For the rest of this paper, we use the term ‘Weekly Average Method’ (WAM) to describe this approach. This differs from the ‘Daily Minimum Method’ (DMM) used by the HSE where a child is judged to meet the recommendation if they report ≥60 minutes per day on each specific day of the preceding week (see Figure 1).”

And now reads:

“Given that the SHeS and HSE questionnaires are so similar, yet produce such different prevalence estimates, we investigated how the data were processed. A subtle but potentially important difference between the SHeS and HSE surveys is how the frequency of activities are reported, which has implications for how the data are processed. In the SHeS, children or their parents are asked to report how many sessions of each activity they have done in the previous seven days [12], making it impossible to ascertain whether child respondents to the SHeS achieve ≥60 minutes of MVPA on each specific day. In order to estimate the frequency element of the recommendation, the SHeS makes the assumption that the first 7 sessions are carried out on different days. Recommendation compliance is then estimated by the proportion of children that report ≥7 sessions, and whose total weekly duration is ≥420 minutes (7*60). For the rest of this paper, we use the term ‘Weekly Average Method’ (WAM) to describe this approach. In the HSE, activities are reported for specific days in the previous seven [14], and a child is judged to meet the recommendation if they report ≥60 minutes per day on each specific day of the preceding week (the ‘Daily Minimum Method’ (DMM); see Figure 1). The DMM is therefore a mathematically stricter version of the WAM, additionally requiring the child to have been active on every single day and to have undertaken ≥60 minutes on each day. This stricter threshold will always lead to lower prevalence estimates.”

We have also edited Figure 1 (page 5, line 112) so that it no longer has a common summary box. Rather, the total weekly time is split into the two separate methods straight away. We hope that this helps to highlight the difference of the methods using number of days versus number of sessions.

3. The introduction suggests that, in the past, others have made comparisons between groups of different age ranges. It is well established that PA decreases with age from childhood to adolescence so of course varying the age range in any comparative work will have an impact.
The authors' conclusions in this paper stress harmonization in analytical approach but perhaps this recommendation should be broader to include simple aspects like age range as well.

Response: We agree this is an important factor to consider if prevalence estimates are to be comparable. Indeed, this issue is covered in more depth in a forthcoming paper (accepting to BJSM pending revisions) that covers all aspects of comparability between UK surveillance methods.

We have added the following pieces of text to the the relevant discussion section to explain our reasoning.

Beginning on page 11, line 259:

“Through the research questions posed and the study design chosen, we were able to isolate the issue of the recommendation interpretation from other persisting concerns around the validity and reliability properties of the measurement instruments and the different age ranges used in the national prevalence estimates.”

And on page 12, line 274:

“We also chose to focus entirely on the issue of analytical approach and so have not quantified the effect of other differing factors in the published national prevalence figures such as differing age ranges.”

4. The paragraph in the introduction talking about the debate about measuring compliance to PA guidelines using accelerometers when they are derived using mostly self-reported PA data doesn't really fit in this paper. If the authors wish to mention it in passing, the discussion may be a better place for it. It just seems to veer off track in the introduction which does a good job of building the rationale for this particular paper.
Response: We agree that this paragraph does disrupt the flow and have thus removed it from the paper.

5. Just a point of clarification - it is noted that the SHeS asks parents whether 7 or more PA sessions have occurred in previous week and then the weekly average is calculated. When the authors were running the analysis for this paper, did they require that children had an average \( \geq 60 \) minutes per day and 7+ sessions in past week or just an average \( \geq 60 \) minutes per day? The 7+ sessions in week does not really fit with the guideline and is not an interpretation that others have used (to my knowledge). Did the authors look at how the inclusion of this additional requirement affected the estimates?

Response: Thank you for raising this as an unclear point. As discussed in the response to Comment 2, we have reordered the relevant paragraph (page 5, line 95) and edited the text to clarify that the WAM originates from the need for the SHeS to approximate sessions as days. We have also made it clearer in that text and Figure 1 that reporting 7+ sessions is a requirement of the WAM (see response to Comment 2 for further detail on this).

We did not look at the proportion of children who achieved 420 mins/week but reported <7 sessions because no UK survey uses this method. This will however be an interesting factor to consider in the future as there is the potential for the number of days requirement to be dropped from the UK guidelines.

6. The fourth paragraph of the discussion mentions the need to create surveillance recommendations when developing guidelines. There is an example of this in Canada. See: Tremblay, Carson et al. Canadian 24-hour movement guidelines for children and youth: An integration of physical activity, sedentary behaviour, and sleep. Applied Physiology Nutrition Metabolism 2016; 41: S311-S327.

Response: Thank you for bringing our attention to this article. We have added a sentence and cited this article in the discussion section (page 11, line 240).
“In light of this, we call on developers of PA recommendations to also include guidance on how to collect, process and analyse data. This should be considered a necessary step in the implementation of any recommendations.”

And now reads:

“In light of this, we call on developers of PA recommendations to also include guidance on how to collect, process and analyse data. An example of this comes from Canada where specific surveillance guidance was provided to accompany the new 24-hour PA guidelines for children and youth [20]. This should be considered a necessary step in the implementation of any recommendations.”

7. At the end of the 3rd last paragraph of the discussion, it is mentioned that others have shown the same gap between estimates of adherence to PA guidelines due to differences in analytical approaches. This issue is also discussed in detail in the Canadian context: see Colley, Carson et al. Physical activity of Canadian children and youth, 2007 to 2015. Health Reports 2017; 28(10): 8-16.

Response: Thanks for highlighting this reference. We have cited this study in the relevant section in our discussion (beginning on page 11 line 248).

The text used to read:

“These results are comparable to a recent study that analysed accelerometry data from a Scottish child cohort study through both the WAM and DMM approaches [23]. The DMM produced an estimate of 11% whilst the WAM produced an estimate of 68%. Taken together, they show that it is important not to conflate the issues of whether device-based surveillance of child PA should be introduced in Scotland with the issues of the high prevalence estimates. Whilst there may be many other relevant considerations for changing national surveillance of child PA to accelerometry, generating comparable prevalence figures to the other national surveys within the UK could be achieved through changes to the questionnaire and/or analysis method.”
“These results are comparable to Scottish and Canadian studies that analysed accelerometry data using two analytical approaches: (i) the DMM as defined in this study and (ii) a weekly total ≥420 minutes [23]. The DMM produced estimates of 11% in Scotland and 2.9-6.2% in Canada across different survey years. The weekly total approach produced estimates of 68% and 28.8-41.0%, respectively. Taken together, they show that it is important not to conflate the issues of whether device-based surveillance of child PA should be introduced in Scotland with the issues of the high prevalence estimates. Whilst there may be many other relevant considerations for changing national surveillance of child PA to accelerometry, generating comparable prevalence figures to the other national surveys within the UK could be achieved through changes to the questionnaire and/or analysis method.”

8. The authors reported that the magnitude of difference differed by age and sex in the results however, not mention of this was made in the discussion. It would be nice to include some discussion of this finding and offer some insights into why this was observed.

Response: As we did not investigate explanatory factors in our study, we are cautious of overinterpreting this result. However, we certainly agree the finding that some differences were significant across age group when we compared the WAM and DMM using HSE data should be discussed, as this indicates a difference between the two methods. We have now included a paragraph on this in our discussion (beginning on page 10, line 212).

The new paragraph reads:

“The largest difference between the DMM and WAM estimates was in 11-12 year olds and the lowest in 13-15 year olds. By definition, the differences occur when children undertake a total ≥420 minutes/week but not in a way that is evenly distributed throughout the week. Therefore, a potential reason for the high difference amongst 11-12 year olds is that this is a transitional age where children move from primary to secondary school. This might mean a shift from daily informal PA, such as active play at lunchtime, to longer but less frequent bouts of PA, such as gym sessions or afterschool sports. The smallest difference amongst 13-15 year olds may be explained by low total weekly durations, meaning neither the DMM nor WAM definition of “active” was met. Future research could investigate patterns in which PA is accumulated in
different age groups to establish why some are more affected by a difference in analysis methods than others.”

Editorial Comments:

Results - first line - is the word "weighted" appropriate here? I think you mean just sample sizes. Same applies for column headings in Table 2. The concern is that "weighted" may infer the number of people in the population that your sample represents, rather than the number in the sample yourself which is what I think you are presenting.

Response: Thanks for raising this point. By using the word “weighted” we mean that weighting variables have been applied. The Scottish Health Survey and the Health Survey for England require the use of weights to ensure that weighted sample has a similar sex and age distribution to their target populations. They also account for bias introduced by non-response. The surveys provide these weights for use.

Weighting the data does not change the overall sample size by very much. In our analyses, the numbers are:

- SHeS: weighted 965, unweighted 946
- HSE: weighted 3840, unweighted 4155

This is not a reflection of the size of the population that the analysis represents (i.e. the children of Scotland and England). Instead it is a slightly redistributed sample of those interviewed to better approximate representativeness of the wider population.

For clarity, and because survey sampling methodology is not the key focus of the paper, we chose to present the weighted numbers only. To ensure we were not misleading readers, we included the word weighted alongside any sample size presented. We feel this is an appropriate balance of detail and clarity but would be happy to take any of the following actions if preferred:

- We could remove all references to weighted, but include a more explicit statement in the methods to indicate we only report weighted values
- We could additionally report the unweighted sample sizes alongside the weighted values
List of corrected typos and minor changes

- Page 2, line 27: “(PA)” inserted to make room for additions whilst remaining within word count
- Page 2, line 29: “guidelines” changed to “recommendation” for consistency
- Page 2, line 30: “monitor” corrected to “monitors”
- Page 2, line 31: “in England” added for clarity
- Page 2, line 31: “physical activity” changed to “PA”
- Page 2, line 32: “monitor” corrected to “monitors”, “undertake” corrected to “undertakes”
- Page 2, line 32: “daily minimum method” corrected to “Daily Minimum Method”
- Page 2, line 33: “physical activity” changed to “PA”
- Page 2, line 33: “over 60” changed to “≥60” for accuracy
- Page 2, line 33: “in Scotland” added for clarity
- Page 2, line 34: “weekly average method” corrected to “Weekly Average Method”
- Page 2, line 34: “how this” corrected to “how much this”
- Page 2, line 35: “influences” inserted
- Page 2, line 36: “Physical activity” changed to “PA”
- Page 2, line 37: “965” corrected to “947”
- Page 2, line 40: “the way the questions are asked” changed to “questionnaire design”
- Page 2, line 48: “physical activity” changed to “PA”
- Page 2, line 49: “(leaving approximately 12 percentage points representing true differences or related to questionnaire differences)” inserted
Page 2, line 50: “to determine” corrected to “determine”

Page 3, line 64: “HBSC” inserted for clarity

Page 3, line 69: a space added after “points at most [8].”

Page 4, line 124: “would consider to be of at least moderate intensity” corrected to “would be considered of at least moderate intensity”

Page 4, line 88: Table 1 – “no School Travel estimate” in headings corrected to “no school travel estimate” for consistency

Page 5, line 114: “We hypothesise that it is this different analysis method” corrected to “We hypothesise that it is this difference in analysis method”

Page 5, line 115: “in the SHeS and HSE” added for clarity

Page 6, line 133: “previously” corrected to “previous”

Page 7, line 147: added “are” to “they are split”

Page 8, line 183: Closed bracket added to end of “(28.5 (95% CI: 25.5-31.6)).”

Page 9, line 189: Table 3, furthest right column: “DMM” corrected to “WAM”

Page 10, line 226: “county Y” corrected to “country Y”

Page 11, line 236: “activity completed” correct to “activity is completed”

Page 11, line 236: “7 day” corrected to “7-day”

Page 11, line 243: “difference” corrected to “differences”

Page 11, line 257: duplicate “directly” removed

Page 12, line 265: “during 2018” changed to “for publication in 2019”

Page 12, line 282: “to determine” corrected to “determine”