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)	Who perpetrates violence against children? A systematic analysis of age and sex specific data
AUTHORS	Knight, Louise; Devries, Karen; Abrahams, Naeemah; Bott, Sarah; Riveros, Dr Betzabe; Cappa, Claudia; Watts, Charlotte; Gannet, Katherine; Kress, Howard; Hollis, NaTasha; Peterman, Amber; Walsh, Sophie; Petzold, Max; Kishor, Sunita; Maxwell, Lauren; Chan, Ko Ling; Guedes, Alessandra; Williams, Abigail; Garcia-Moreno, Claudia

VERSION 1 - REVIEW

REVIEWER	George Nikolaidis Institute of Child Health Competing Interests: I had some consultancy relation with UNICEF before (for which some of the authors are working for); however, I do not consider that to create any kind of conflict of interest at this occasion
REVIEW RETURNED	18-Apr-2017

GENERAL COMMENTS	The manuscript presents a meta-analysis of previous research on child abuse and neglect. Given that, it represents an extremely interesting research initiative since on one hand the importance (and public health impact) of children's victimization is widely acknowledged while on the other there is shortage of population-based empirical data on the subject. However, as the manuscript itself admits, there is a prima facie issue on conceptual and methodological diversity of field research on the particular subject matter. Such conceptual and methodological discrepancies among research methods and operational definitions employed makes comparison at least challenging; as some of the questionnaires used might refer to practices submitted to children (i.e. "hitting you with bear hand") while others on broader concepts involving subjective apprehension (such as "exposure to violence"), comparison among results might end up quite misleading (resulting in differentiated incidence percentages in virtue of measurement bias and not revealing and truthful diversity); additionally, in some research protocols there is a more detailed and wide inquiry on types of violent practices in which minors might be exposed to while in others only a few practices or very broad concepts are asked the subjects to report (leading again in under- or over- estimation of overall violence exposure percentages). Such challenging issues are quite common in child maltreatment field research; and mostly can be dealt with by comparing only equivalent items of (the most commonly used) questionnaires (rather than by comparing overall

	abuse exposure percentages which might be misleading. In any case, providing some further discussion on how results could be comparable among them would merit the manuscript. On the grounds of the above (which is a wider consideration on the implementation of meta-analytic techniques on the particular field) the last paragraphs in Box 1 of the manuscript (page 5, lines 50 and beyond) as well as the paragraph in page 6 seem to introduce an additional consideration about the soundness of considering as “abuse proper” practices (i.e. of corporal punishment of harsh discipline) in societal environments that they are considered as standard ways of children’s upbringing. That line of thought, from an empirical researcher’s point of view might be totally vindicated (since it creates doubt about truthfulness and reliability of responders, differentiated ways of conceiving the same items in internationally used questionnaires etc.); however, if not clarified as such, it might leave room for misunderstanding (which I reckon that authors do not wish anyway). Therefore, maybe clarifying it would benefit the clarity of the message that the manuscript brings to the reader. In lines 51 and beyond of page 8 of the manuscript another important issue is raised, namely the problems in projecting nation-wide results from school-based field surveys. That is so, mainly because it is well established that groups of children which leave school earlier are in general more vulnerable to many kinds of victimization. Later in the text, the manuscript gives the impression that this discrepancy in expected rate of exposure to violence among children who attend school and children who have dropped out is acknowledged (page 12 etc.); still some clarification in page 8 would contribute to further clarity of the methods used and their rationale. In page 11 and elsewhere the manuscript mentions the relative lack of data from surveys in children younger than 11 years’ old. Given the fact that most commonly used questionnaires are by default not to be sued in younger children and given also the fact that such limitations mainly derive from problematic potential discrimination between “truth” and “reality” (historical truthfulness) in younger children that could be also mentioned as a main reason for this shortcoming of international research. In the second paragraph of page 14 of the manuscript (lines 14 to 25) it is reported that results of surveys on children’s victimization differ in respect to the responder, the ones by caregivers giving usually higher rates of reported victimization comparted to the ones by children themselves. Still, this is highly sensitive in respect to specific items of questionnaires (more often rates are higher when caregivers are asked to report regarding “milder”: forms of children’s victimization; in “more severe” forms percentages might be higher when children are asked directly). Of course, this comment leads us again in the initial comment on methodology and conceptual diversity of tools used in various research field studies that were introduced in the meta-analysis (which in turn makes again the case for some further clarity on the consequent limitations in virtue of that diversity). In overall, it is a rigorous piece of scientific work on an extremely important subject matter which contributes substantially to the enrichment of existing knowledge on the field and promotes building the evidence base required for sound public health and social policy design.

- Permission to publish reviewer 2's review was not received.

REVIEWER	Guoqing Hu Xiangya School of Public Health, Central South University Competing Interests: None
REVIEW RETURNED	08-May-2017

GENERAL COMMENTS	This study involves a lot of work and focuses on an important health topic. The results are important for understanding global child abuse and developing specific solutions. But the following key issues need to be clarified. 1. I have been thinking whether use prevalence or incidence in this study for several hours. Considering that most operational definitions for violence against children that are included in the appendix, are based on retrospective survey or self-report and the violence events are often discontinuous, probably “incidence” is better than “prevalence” for this study. 2. Because available data sources cover a long time period, whether there are some intervention effects for the countries with data for many years. If it is true, it may not suitable to put them together to generate Meta estimates. How to treat such effects should be mentioned in the appended methods and approaches if they exist. 3. It is great that the authors used meta-regression to correct the effects of various definitions and methodologies (annex 4). I think the main results of model fit statistics should be listed since the results highly depend on the reliability and validity of model estimates. Specific comments Title: Although 13,830 separate age and sex specific prevalence estimates from 600 population or school-based representative datasets and 43 publications obtained of 171 countries, I am afraid the estimates may be not world representative, especially for age- and sex-specific estimates because data sources do evenly distribute across countries. Probably, a title “Who perpetrates violence against children? A systematic analysis of available age and sex specific data”. Abstract: 1. Please add a subsection for outcome measure on page 3. 2. Results (lines 24-26 on page 3): Importance numbers or percentage should be inserted, such as prevalence in the past 12 months and percentage of major perpetrators. What this study adds: Lines 13-22 on page 4: important prevalence rates should be added here. Introduction Lines 2-13 on page 6: It looks that a clear operation definition should be given for child violence. Methods Lines 46-49 on page 7: Was reproducibility coefficient assessed for literature screening and data extraction for two researchers? Lines 22-24, 41-44 on page 8: Please justify the choice of two kind statistical models and add a reference to them.

	Lines 49-50 on page 8: The statistical processing for studies where the recall period was below one year does not look reasonable. Empirically, such processing may underestimate the rates especially when the recall period is very short. According to annex 3, I notice that the classification of being threatened is regarded as physical violence in some studies, but as emotional violence in other studies. The inconsistency is hard to understand. How do the authors treat it? In addition, Does the analysis by region in annex 6 make sense in terms of high scarcity of data in some regions? Results Lines 25-29 on page 11: Please provide the full titles of abbreviations under the table. Figures 2-4: It would better to read if a title is given to x axis (including the unit of prevalence rate). In addition, it may be valuable to present the percentages of child violence made by all other perpetrators, in addition to the major perpetrator. Lines 15-45 on page 12, lines 24-35 on page 17, page 21: Figure 2 and the description are hard to read. What does “recent physical and emotional violence perpetrated by household members” mean? Probably, a short note under the figure would facilitate the reading. Figures 5 and 6: Probably, it is helpful to add the survey time to the two figures. Discussions The discussion is well written.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1:

Comments:
The manuscript presents a meta-analysis of previous research on child abuse and neglect. Given that, it represents an extremely interesting research initiative since on one hand the importance (and public health impact) of children’s victimization is widely acknowledged while on the other there is shortage of population-based empirical data on the subject.
However, as the manuscript itself admits, there is a prima facie issue on conceptual and methodological diversity of field research on the particular subject matter. Such conceptual and methodological discrepancies among research methods and operational definitions employed makes comparison at least challenging; as some of the questionnaires used might refer to practices submitted to children (i.e. “hitting you with bear hand”) while others on broader concepts involving subjective apprehension (such as “exposure to violence”), comparison among results might end up quite misleading (resulting in differentiated incidence percentages in virtue of measurement bias and not revealing and truthful diversity); additionally, in some research protocols there is a more detailed and wide inquiry on types of violent practices in which minors might be exposed to while in others only a few practices or very broad concepts are asked the subjects to report (leading again in under- or over-estimation of overall violence exposure percentages). Such challenging issues are quite

common in child maltreatment field research; and mostly can be dealt with by comparing only equivalent items of (the most commonly used) questionnaires (rather than by comparing overall abuse exposure percentages which might be misleading. In any case, providing some further discussion on how results could be comparable among them would merit the manuscript.

We fully agree with the reviewer that comparing different studies with divergent methodologies and conceptualisations of violence is a difficult task. As the reviewer notes, this is the main task in any evidence synthesis effort. Rather than doing a straightforward meta-analysis where estimates from studies with divergent methods are simply averaged together, our methodology involves estimation to overcome some of the differences between studies. The reviewer mentions several measurement criteria which can cause differences in prevalence measures—whether measures ask about specific acts of violence, or more general subjective measures of exposure, and the number of questions included in a survey to measure violence exposure.

We have attempted to explain our stance on what counts as 'violence' for the purposes of this analysis in Box 1, where we state we include all measures of violent acts whether or not participants defined these as violence. For all included studies, we recorded whether survey questions ask about specific acts, versus more subjective questions about 'exposure to violence'. Annex 3 lists the questions asked in each study. We coded each definition and added these as covariates in meta-regression models, and have used these models to adjust estimates up or down based on these differences as far as possible within each type of violence (for example, for emotional violence from household members, all definitions of this form of violence were coded and added to a meta-regression model to produce estimates for that form of violence.) The separate models for each form of violence are described in Annex 4. We believe that this is a useful first step in drawing together literature on the key research questions we ask here, however we also agree that there are limitations to this approach. In particular, some of the school based surveys also have asked fewer questions about exposure to different acts of violence relative to more specialist violence studies, and thus may be more prone to misclassification of violence exposure relative to estimates of household and intimate partner violence.

We have added further text on limitations in the discussion, which now reads: "Although every effort was made to adjust for differences in measurement of violence across studies there may be residual confounding related to both definitions of violence (including whether studies asked about experience of specific acts of violence and how many questions they asked) and other study quality variables. These differences may in part explain age, sex and regional differences in prevalence estimates. Further, the school-based studies tended to include fewer questions about experience of different specific acts of violence, thus the school-based estimates may be more prone to misclassification of violence exposure relative to estimates of household and intimate partner violence. For most countries, data were only available from one or two survey years—pooling data from different years may obscure trends in the prevalence of violence over time."

On the grounds of the above (which is a wider consideration on the implementation of meta-analytic techniques on the particular field) the last paragraphs in Box 1 of the manuscript (page 5, lines 50 and beyond) as well as the paragraph in page 6 seem to introduce an additional consideration about the soundness of considering as "abuse proper" practices (i.e. of corporal punishment of harsh discipline) in societal environments that they are considered as standard ways of children's upbringing. That line of thought, from an empirical researcher's point of view might be totally vindicated (since it creates doubt about truthfulness and reliability of responders, differentiated ways of conceiving the same items in internationally used questionnaires etc.); however, if not clarified as such, it might leave room for misunderstanding (which I reckon that authors do not wish anyway). Therefore, maybe clarifying it would benefit the clarity of the message that the manuscript brings to the reader.

Thank you for pointing this out, and we certainly want to avoid misunderstanding. We have shortened and clarified the text in box 1.

In lines 51 and beyond of page 8 of the manuscript another important issue is raised, namely the problems in projecting nation-wide results from school-based field surveys. That is so, mainly because it is well established that groups of children which leave school earlier are in general more vulnerable to many kinds of victimization. Later in the text, the manuscript gives the impression that this discrepancy in expected rate of exposure to violence among children who attend school and children who have dropped out is acknowledged (page 12 etc.); still some clarification in page 8 would contribute to further clarity of the methods used and their rationale.

We again agree that there are certainly differences between school-based survey estimates and national survey estimates. In our analysis, the only surveys which tend to ask about exposure to violence from students are school-based surveys (few population based surveys include these measures). We have in fact adjusted estimates on *school violence* from school based surveys downward to reflect the fact that not all children in a given setting actually attend school, and therefore are not exposed to school violence. This is to make the denominators across household, school and intimate partner violence syntheses the same so relative prevalence at the population level can be discussed. We have clarified in text in the methods section, which now reads: "Our goal is to understand who the most common perpetrators of violence by age and sex in the whole population are, and therefore to compare prevalence across groups with different denominators. This required us to adjust estimates on school violence from school-based surveys, and intimate partner violence from ever-partnered young people, as not all young people attend school or are in intimate relationships (and therefore by definition are not exposed to these forms of violence). Estimates provided with students as the denominator were adjusted by the WHO regional estimation proportion of students attending primary and secondary schools[15].Estimates provided with the ever-partnered proportion of the survey population as the denominator were adjusted by the proportion of country populations which had ever had sex by age 20 years (using DHS data[16]), to make them reflect the prevalence of different forms of partner violence in the entire population (rather than only the ever-partnered population)."

Levels of violence among children not attending school and not likely to be in population based surveys are a separate matter—we acknowledge in our limitations section that the data pooled in our analysis do not include these children, and that levels of violence in these populations are likely to be higher.

In page 11 and elsewhere the manuscript mentions the relative lack of data from surveys in children younger than 11 years' old. Given the fact that most commonly used questionnaires are by default not to be sued in younger children and given also the fact that such limitations mainly derive from problematic potential discrimination between "truth" and "reality" (historical truthfulness) in younger children that could be also mentioned as a main reason for this shortcoming of international research.

We have added a line to the discussion to make this clear. The start of the relevant paragraph now reads: "There is a clear need for more data on the experiences of younger children, particularly around family and sexual violence. This may stem partly from investigators' concerns about the validity of younger children's survey responses."

In the second paragraph of page 14 of the manuscript (lines 14 to 25) it is reported that results of surveys on children's victimization differ in respect to the responder, the ones by caregivers giving usually higher rates of reported victimization comparted to the ones by children themselves. Still, this is highly sensitive in respect to specific items of questionnaires (more often rates are higher when caregivers are asked to report regarding "milder": forms of children's victimization; in "more severe" forms percentages might be higher when children are asked directly). Of course, this comment leads us again in the initial comment on

methodology and conceptual diversity of tools used in various research field studies that were introduced in the meta-analysis (which in turn makes again the case for some further clarity on the consequent limitations in virtue of that diversity).

We have added a line to the relevant paragraph, which now reads: "As seen with estimates of violence from caregivers/families, prevalence varies widely depending on whether caregivers are reporting or children are reporting. Perhaps contrary to expectation, caregiver reports, mainly from the MICS, yield higher prevalence estimates versus children's own reports (which tend to be from other surveys). The MICS data on emotional violence, as analysed here, include items measuring shouting, screaming and calling a child stupid or lazy (Annex 3), which may occur frequently but may not be viewed as particularly traumatic, which likely will increase prevalence estimates. Caregivers may be more likely to report these less severe acts of violence, relative to more severe forms. Children may also be more likely to recall incidents which were severe or traumatic for them, thus biasing self-reported estimates downwards.**"**

In overall, it is a rigorous piece of scientific work on an extremely important subject matter which contributes substantially to the enrichment of existing knowledge on the field and promotes building the evidence base required for sound public health and social policy design.

Thank you.

Reviewer 3:

Abstract:
1. Please add a subsection for outcome measure on page 3.
2. Results (lines 24-26 on page 3): Importance numbers or percentage should be inserted, such as prevalence in the past 12 months and percentage of major perpetrators.
What this study adds:
Lines 13-22 on page 4: important prevalence rates should be added here.

Thank you for these comments. We have structured the abstract according to BMJ Pediatrics guidelines. We have multiple outcome measures, and have described the overall categories under the objectives section. We would ideally also add percentage as suggested, however because we have produced age and sex specific rates, there are no overall specific percentages that we can add which are appropriate for an abstract.

Introduction
Lines 2-13 on page 6: It looks that a clear operation definition should be given for child violence.

We have included Box 1 to explain how we conceptualized violence. The forms of violence modelled are described under each Figure, and the detailed definitions in each study are described in Annex 3.

Methods
Lines 46-49 on page 7: Was reproducibility coefficient assessed for literature screening and data extraction for two researchers?

We have clarified our procedure in the methods section, based on comments from another reviewer. The relevant text now reads: “Screening of abstracts and full text articles was performed by KM, LM, and AW. KM performed initial screening to remove irrelevant titles. Due to the volume of results, double screening of abstracts was not employed. Instead, KM, LM and AW screened a subset of 150 articles together using standardised inclusion criteria, discussing application of the criteria until consistency was reached. Remaining abstract screening was done by one reviewer. Data on study characteristics and quality were extracted by KM or LM, into a customised Google form database created by LM. KM and LM discussed any questions on a weekly basis.”

Lines 22-24, 41-44 on page 8: Please justify the choice of two kind statistical models and add a reference to them.

We have conducted standard meta-regressions and meta-analyses, and have added references in the text for these [15] and [16].

Lines 49-50 on page 8: The statistical processing for studies where the recall period was below one year does not look reasonable. Empirically, such processing may underestimate the rates especially when the recall period is very short.

We agree that we have made a conservative assumption, but we feel that this is appropriate and more defensible that attempting to adjust up these estimates. In practice, no household or IPV studies had recall periods under one year, but 53% of school based studies did. Since violence exposures can be chronic and episodic, it is likely that a proportion of children who have experienced violence in the past 3 months have also experienced it between 3 months and one year ago. Judging how many may have experienced violence only between 3 months and one year ago is difficult, and there is a lack of empirical data on which to base such an adjustment assumption. Hence, we decide to conservatively estimate prevalence as the shorter period prevalence.

According to annex 3, I notice that the classification of being threatened is regarded as physical violence in some studies, but as emotional violence in other studies. The inconsistency is hard to understand. How do the authors treat it?

You have highlighted an issue in the field of violence research in general, and you are correct that how threats of physical violence are classified does differ between studies. Re-classifying all of these to be consistent was beyond the scope of the current analysis, and we have used author definitions of each form of violence.

In addition, Does the analysis by region in annex 6 make sense in terms of high scarcity of data in some regions?

As the aim of this paper was to produce overall estimates, rather than estimates by region, we have decided to remove Annex 6.

Results

Lines 25-29 on page 11: Please provide the full titles of abbreviations under the table.

These are defined in the text of the article, and we prefer to leave as abbreviations under the table to save space.

Figures 2-4: It would better to read if a title is given to x axis (including the unit of prevalence rate). In addition, it may be valuable to present the percentages of child violence made by all other perpetrators, in addition to the major perpetrator.

Lines 15-45 on page 12, lines 24-35 on page 17, page 21: Figure 2 and the description are hard to read. What does "recent physical and emotional violence perpetrated by household members" mean? Probably, a short note under the figure would facilitate the reading.

Thank you for noting this—our figure legends defining and describing how to read the figures seem to have been disconnected from the figures during the submission process (they are included at the end of the article text rather than next to the figures). We also have decided not to present the forms of violence from different perpetrators on the same graph since we would not be able to include confidence intervals for the estimates while maintaining readability.

Figures 5 and 6: Probably, it is helpful to add the survey time to the two figures.
Thank you for this suggestion, we have added the survey times underneath the figure in a footnote (since there were relatively few surveys, this saves space in the graph).

Discussions
The discussion is well written.

Thank you.