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.

ARTICLE DETAILS

TITLE (PROVISIONAL)	The reporting of data monitoring committees and adverse events in pediatric trials: a descriptive analysis
AUTHORS	Gates, Allison; Caldwell, Patrina; Curtis, Sarah; Dans, Leonila; Fernandes, Ricardo; Hartling, Lisa; Kelly, Lauren; Vandermeer, Ben; Williams, Katrina; Woolfall, Kerry; Dyson, Michele

VERSION 1 – REVIEW

REVIEWER	Reviewer name: Peter Flom Institution and Country: Peter Flom Consulting, USA Competing interests: None
REVIEW RETURNED	29-Dec-2018

GENERAL COMMENTS	I confine my remarks to statistical aspects of this paper. These were very simple, but this was appropriate and I have no problems with the methods.

REVIEWER	Reviewer name: Shaoqing Ni Institution and Country: Children's Hospital, Zhejiang University Schollo of Medicine,Hangzhou,China Competing interests: No
REVIEW RETURNED	13-Jan-2019

GENERAL COMMENTS	1. This article evaluated the conduct and reporting of a randomly selected paediatric trials published in 2012. It is 2019 now. These data were a little old. However, it revealed some facts that need improvement in the future. 2. As DMCs is not required in all studies. Mandated reporting of DMCs is not realistic. If the author can detailed the study type that needs DMCs reporting and the required member combination of DMCs would be more reasonable and helpful.

REVIEWER	Reviewer name: Dr Florentia Kaguelidou Institution and Country: Florentia Kaguelidou MD, .Associate Professor of Clinical Pharmacology. Clinical Investigations Center, INSERM CIC 1426. Robert Debré Hospital, APHP, 48 boulevard Sérurier, 75019 Paris. University Paris 7 – Diderot, Paris, France. Competing interests: none
REVIEW RETURNED	24-Jan-2019

GENERAL COMMENTS	This article is part of a 'series' of papers describing the same 300 randomly selected, randomised controlled trials (RCTs) reporting outcomes in participants ≤ 21 years of age in 2012. After evaluating the global conduct and reporting quality of the RCTs and the consent and recruitment strategies used, the authors describe now the reporting of data monitoring committees (DMCs), safety monitoring data and adverse events in these trials.

	General Comments 1. Authors provide an overall description when there are actually three distinct aspects to describe in this paper: 1) existence of not of a DMC with definition of interim analyses and stopping rules, 2) early stopping of the RCTs and reasons for that, and 3) reporting of adverse events and harm-related outcomes. If authors agree with this remark then they should review the entire paper to reflect these 3 distinct aspects of monitoring participants' safety in RCTs. 2. Authors present a limitation to their paper that is actually a quite significant one: not all trials theoretically require a DMC (there are some clinical and methodological criteria for that, Ellenberg et al. Pediatrics, 2012). Therefore, authors should have presented the results separately for the RCTs that require DMCs and those that a priori do not require one. Describing this feature for the entire sample of trials leads to rather pessimistic conclusions. 3. Conversely, reporting of adverse events should be provided for all RCTs and reflects a different aspect of safety monitoring to be introduced, described and discussed separately. 4. Discussion is poorly written. Authors conclude with the following general and not original message: 'it is not well done and should be better done because the RCTs concern children that are vulnerable'. They do not put forward some other interesting findings such as but not limited to: drug and vaccine trials frequently report DMCs but trials evaluating medical devices and surgery or radiotherapy do not; a lot of trials report on harm-related outcomes but no one stopped early because of harm or adverse events; etc… I suggest that the authors review their findings and enrich their discussion with more interesting and original messages. Comments by section TITLE 1. What do the authors mean by 'safety monitoring data'? Does that refer to interim analyses, stopping rule and early stopping? If not, then I propose to delete this from the title. INTRODUCTION 1. First paragraph is a 'standard paragraph' about 'therapeutic orphans', not necessary. Authors should rather introduce the need for safety monitoring in RCTs, present previous findings and provide a better justification for their review. METHODS 1. Methods are reduced to a minimum because of previous reports of the same extraction and this seriously impacts the reading of the paper. 2. It is also regretful that the authors do not provide any information about the design of the study or the funding sources which would have been interesting to describe with regards to the existence or not of a DMC (do industry-sponsored trials report a DMCs more frequently than institutional trials?) 3. The authors did not retrieve any information about the independency of the DMCs' members though investigators and sponsors should not be involved in these committees. 4. Is Appendix 1 necessary? Has been provided in previous publications? 5. Appendix 2. What do the authors mean by 'did the authors plan to collect data on adverse effects/events or side effects?' and what did they expect to find as 'methods for collecting data on adverse effects?'? Are these features required to be reported in RCTs and is there a list of accepted methods or specific guidelines for that?

	The CONSORT guidelines recommend to report 'All important harms or unintended effects in each group' in the results section. RESULTS 1. Reporting of interim analyses is described for the 55 RCTs that reported a DMC but stopping rules are described for the entire cohort of RCTs (n=300). Why? Both interim analyses and stopping rules should be described for the 55 RCTs with DMCs. Do the authors suggest that all RCTs should define 'stopping rules'? 2. Early stopping and reasons for that is a separate question that is not always related to adverse events as also suggested by the findings of this review. DISCUSSION 1. As commented previously, authors should review entirely their discussion to put forward some more pertinent and interesting messages. ABSTRACT - Line 17: '…describe the findings descriptively…' could also be written as '…describe the findings…' - Line 30: 'The reporting of adverse events and harm-related endpoints varied by the reported presence of a DMC (P < 0.001) and the nature of the intervention (P = 0.002)'. Authors should provide precisions on the direction of the effect (e.g. reporting was more frequent with RCTs that reported a DMC) together with the p-value.

VERSION 1 – AUTHOR RESPONSE

Response to the Comments from Reviewer 1

1. I confine my remarks to statistical aspects of this paper. These were very simple, but this was appropriate and I have no problems with the methods.

Response: Thank you for lending your expertise to the review of our statistical methods.

Response to the Comments from Reviewer 2

1. This article evaluated the conduct and reporting of a randomly selected paediatric trials published in 2012. It is 2019 now. These data were a little old. However, it revealed some facts that need improvement in the future.

Response: You are correct, as the data are from 2012 the findings may not be reflective of present-day conduct and reporting in pediatric trials. To acknowledge this, we have added the following to the Strengths and Limitations section: "We evaluated trials published in 2012, providing a baseline for ongoing evaluation of safety monitoring procedures in pediatric trials; however, the findings may not be reflective of present-day conduct and reporting."

2. As DMCs is not required in all studies. Mandated reporting of DMCs is not realistic. If the author can detailed the study type that needs DMCs reporting and the required member combination of DMCs would be more reasonable and helpful.

Response: You are correct, not every trial will require a DMC. In response to your comment and to those of Reviewer 3, we have revised the introduction to include a description of the roles of DMCs in trials, their recommended composition (membership), the types of trials that absolutely require a DMC, and those where a DMC is strongly recommended.

To make room for this new information in the introduction, we have deleted the first ("standard") introductory paragraph that was present in the previous draft.

Within our discussion, we have deleted "the mandated reporting of DMCs and harms" and replaced this with "…the vigilant review of the reporting of DMCs and adverse events data by reviewers and editors of academic journals…".

Response to the Comments from Reviewer 3

1. Authors provide an overall description when there are actually three distinct aspects to describe in this paper: 1) existence of not of a DMC with definition of interim analyses and stopping rules, 2) early stopping of the RCTs and reasons for that, and 3) reporting of adverse events and harm-related outcomes. If authors agree with this remark then they should review the entire paper to reflect these 3 distinct aspects of monitoring participants' safety in RCTs.

Response: You are correct, we aimed to investigate the reporting of (a) a data monitoring committee, its members and their responsibilities; (b) interim analyses, stopping rules, and early stopping (with reasons); and (c) adverse events and harm related endpoints. We acknowledge that these are distinct, yet related concepts, and have revised our introduction to better reflect our aim. We have deleted the first ("standard") introductory paragraph and replaced this with a description of the role of data monitoring committees in trials, including the establishment of stopping guidelines, interim analyses of the data, and recommendations for modification, continuation, or early stopping. We have followed that with a paragraph describing trials that require a data monitoring committee, and ones where a data monitoring committee should be strongly considered. We have commented briefly on the typical composition (membership) and roles of data monitoring committees (also addressing the comments from Reviewer 2). Finally, we having included a paragraph specifically discussing the role of data monitoring committees and the collection and reporting of harms data in pediatric trials.

Within the results section, the findings related to each aim are presented separately: first the findings related to data monitoring committees, followed by the findings related to interim analyses, stopping rules, and early stopping, and finally the findings related to adverse events and harm related endpoints. We have separated Table 1 into two tables, one including the results related to data monitoring committees and the other including the results related to harms. Moreover, we have edited our discussion to include a paragraph specific to each of these.

2. Authors present a limitation to their paper that is actually a quite significant one: not all trials theoretically require a DMC (there are some clinical and methodological criteria for that, Ellenberg et al. Pediatrics, 2012). Therefore, authors should have presented the results separately for the RCTs that require DMCs and those that a priori do not require one. Describing this feature for the entire sample of trials leads to rather pessimistic conclusions.

Response: We agree, not all trials require a data monitoring committee, although most pediatric trials should strongly consider establishing one. Nevertheless, our investigation into a "snapshot" of a large sample of all pediatric trials published in one year remains useful and informative; it allows us to compare the reporting of pediatric trials (in general) over time with previous studies that have used similar approaches, including our own investigation of trials published in 2007 (Hamm et al., BMC Pediatr 2010;10:96).

We appreciate your suggestion to report the results separately for trials that require data monitoring committees and those that do not; however, dichotomizing the sample of trials as suggested would not be a straightforward task. First, there exists variation in the criteria for pediatric trials that require a data monitoring committee.

Although some trials may clearly require a data monitoring committee (e.g., those that investigate new interventions, or those that recruit from high-risk populations), for others the decision whether to establish one is less clear-cut, and involves consideration of various factors. Some of these factors, such as the feasibility and timeframe for the intervention, are independent of the nature of the intervention and not available in published reports.

Furthermore, guidance from regulatory bodies, sponsors, funders, and academic institutions on the establishment of data monitoring committees leave room for interpretation. This includes the recommendations in the cited publication on data monitoring committees in pediatric trials (Ellenberg et al., Pediatrics 2012;129(Suppl 3):S132-7). Although the criteria for requiring a data monitoring committee are clear, these are not fully operationalized, leaving room for subjectivity. For example, Ellenberg et al. suggest that trials with "a large sample size" warrant a data monitoring committee, without explicitly defining at what point a sample size would be considered "large".

Given the various factors required for deciding on the establishment of a data monitoring committee, and on the ambiguity regarding which trials require one, we have not dichotomized our sample as suggested. We maintain that an analysis of a large sample of trials published over one year remains highly informative and facilitates comparisons to previous work. Nevertheless, to address your concerns we have added the following statement to the Strengths and Limitations section of the manuscript: "We evaluated trials published in 2012, providing a baseline for ongoing evaluation of safety monitoring procedures in pediatric trials; however, the findings may not be reflective of present-day conduct and reporting. Moreover, because we investigated a random sample of trials, not all of the trials would have required a DMC. Nevertheless, ongoing evaluation of the state of the research is needed to evaluate changes over time and identify the areas in most need of attention. The random nature of our sample facilitates comparisons to previous studies, including a similar descriptive analysis of pediatric trials published in 2007."

3. Conversely, reporting of adverse events should be provided for all RCTs and reflects a different aspect of safety monitoring to be introduced, described and discussed separately.

Response: We agree, the reporting of adverse events should be provided for all trials, and be reported separately. Indeed, we have provided data on the reporting of adverse events and harm-related endpoints for all 300 trials in the sample, not only those that reported establishing a data monitoring committee. As previously noted (see response to Reviewer 2 and to your first comment), we have rewritten the introduction to include pertinent background information both about data monitoring committees and the reporting of harms in pediatric trials. Within the Results we have included separate headings each of the three priority areas (i.e., the reporting of data monitoring committees; interim analyses, stopping rules, and early stopping; and harms and harm related endpoints) that we sought to investigate. We separated Table 1 into two tables, one addressing our findings related to data monitoring committees and the other addressing our findings related to the reporting of harms. We have also revised our discussion to include a paragraph on each of the three separate, but related issues.

4. Discussion is poorly written. Authors conclude with the following general and not original message: 'it is not well done and should be better done because the RCTs concern children that are vulnerable'. They do not put forward some other interesting findings such as but not limited to: drug and vaccine trials frequently report DMCs but trials evaluating medical devices and surgery or radiotherapy do not; a lot of trials report on harm-related outcomes but no one stopped early because of harm or adverse events; etc… I suggest that the authors review their findings and enrich their discussion with more interesting and original messages.

Response: Thank you for your feedback on our discussion. As mentioned in our responses to your previous comments (#1 and #3), we have revised our discussion to include a paragraph for each of the issues that we investigated, as well as to include the specific concerns that you have identified. We hope that you find the revised discussion to be more interesting.

5. TITLE. What do the authors mean by 'safety monitoring data'? Does that refer to interim analyses, stopping rule and early stopping? If not, then I propose to delete this from the title.

Response: Thank you for the suggestion. We have removed "safety monitoring data" from the title. The revised title, though more succinct, adequately reflects the aims of the study.

6. INTRODUCTION. First paragraph is a 'standard paragraph' about 'therapeutic orphans', not necessary. Authors should rather introduce the need for safety monitoring in RCTs, present previous findings and provide a better justification for their review.

Response: Thank you for the constructive feedback on our introduction. As suggested, we have deleted the first ("standard") paragraph from the introduction. As mentioned in our response to Reviewer 2 and to your previous comment (#1), we have revised the introduction to include paragraphs specific to the aims of our study. These include information about the need for safety monitoring in trials, types of trials that do and do not require a data monitoring committee, and the data monitoring committee's membership and roles. Moreover, we have included a short paragraph to provide context and to justify our aims. This includes previous evidence of the inadequate reporting of data monitoring committees and of harms in pediatric trials, and the publication of evidence-based standards for conduct and reporting.

7. METHODS. Methods are reduced to a minimum because of previous reports of the same extraction and this seriously impacts the reading of the paper.

Response: We appreciate your insight and have expanded the Methods (within reason, given that these were published previously) to address your concerns.

8. METHODS. It is also regretful that the authors do not provide any information about the design of the study or the funding sources which would have been interesting to describe with regards to the existence or not of a DMC (do industry-sponsored trials report a DMCs more frequently than institutional trials?)

Response: Thank you for highlighting this important consideration. We have added information about the funding sources of the trials within the first paragraph of the results, where we describe the characteristics of the included trials. Because many trials were sponsored by more than one type of funding source, we dichotomized the trials into industry-sponsored and non industry-sponsored trials to investigate the potential relationship between funding source and the reported presence of a data monitoring committee. Our results can be found in Table 2. We have also added these to the abstract.

We have provided information about the designs of the trials in a previous publication, and have also noted in the first paragraph of the results section that the majority of the trials were parallel RCTs, and nearly all were efficacy or superiority trials. Because there was minimal variation in the designs of the included trials, we did not believe that an investigation of the presence of a data monitoring committee stratified by design would be meaningful for this sample.

9. METHODS. The authors did not retrieve any information about the independency of the DMCs' members though investigators and sponsors should not be involved in these committees.

Within the revised introduction, we emphasize that to provide credible and unbiased monitoring of ongoing trials, members of the data monitoring committee should be independent of the trial sponsor; however, we did not aim to appraise the independency of members of a data monitoring committee from the sponsors or members of the investigative team. Moreover, it is not feasible from published reports to assess independency with any validity.

Response: We nevertheless acknowledge the critical importance of independency of members of the data monitoring committee from the sponsors and investigators. In response to your comment, we have added the following to the Discussion, under the heading "Implications for research and practice": "As it was not feasible in this study to validly appraise the independency of members of the DMCs from trial sponsors or investigators (which is necessary to ensure unbiased monitoring), we cannot draw any conclusions regarding DMC conduct. Future studies may consider addressing this knowledge gap."

10. METHODS. Is Appendix 1 necessary? Has been provided in previous publications?

Response: You are correct, we have included the details of the search in previous publications, and it was included here for transparency. In response to your suggestion, we have revised the manuscript to instead include references to the previously published works so that readers will know where seek the details.

11. METHODS. Appendix 2. What do the authors mean by 'did the authors plan to collect data on adverse effects/events or side effects?' and what did they expect to find as 'methods for collecting data on adverse effects?'? Are these features required to be reported in RCTs and is there a list of accepted methods or specific guidelines for that? The CONSORT guidelines recommend to report 'All important harms or unintended effects in each group' in the results section.

Although these items may not be part of CONSORT, they are integral to interpreting the findings of the trial. For example, some trials may report adverse events without any indication as to a plan to collect these data, nor any indication of how they were collected. Inadequate transparency in the methods makes it very difficult to interpret the validity of the findings. As stated in the revised methods, our data extraction guide was adapted from a previous study, with input from clinical and methodological experts.

For the question, "Did the authors plan to collect data on adverse effects/events or side effects?", we were looking for the reporting of a plan to collect adverse events data within the methods of the publication (or in the trial registry or protocol). A couple of examples include: "Infants were closely monitored for adverse events […]. Adverse events were reported to the study Data and Safety Monitoring Board." and "Secondary outcomes were […] the frequency of adverse side effects in each medication group."

For the question, "Was a method for collecting data on adverse effects stated?", we were looking for the reporting of a method to collect adverse events data. A couple of examples include: "All complications were recorded until the patients were discharged" and "Child health workers were asked to report any problems, including neonatal seizures, local skin burns, […]".

Response: We have added these examples to the Appendix to help readers understand the questions.

12. RESULTS. Reporting of interim analyses is described for the 55 RCTs that reported a DMC but stopping rules are described for the entire cohort of RCTs (n=300). Why? Both interim analyses and stopping rules should be described for the 55 RCTs with DMCs. Do the authors suggest that all RCTs should define 'stopping rules'?

Response: Thank you for taking note of this inconsistency. You are correct, we should have reported the reporting of interim analyses and stopping rules for the trials that reported having a DMC. We have edited the table such that the denominator for this variable is 55 (i.e., the trials that reported the presence of a DMC).

13. RESULTS. Early stopping and reasons for that is a separate question that is not always related to adverse events as also suggested by the findings of this review.

Response: You are correct, early stopping could be related to adverse events, or could be due to other reasons (e.g., funding limitations); however, we aimed to investigate the reporting of early stopping as recommending that a trial be stopped early is one task typically undertaken by the data monitoring committee. Indeed, of the trials that reported early stopping in our sample, 38% were stopped due to futility and 15% due to clear benefit of the treatment being investigated. For this reason, we consider this outcome to be relevant in the context of our study. We have clearly indicated in the description of our findings that none of the trials in our sample reported stopping early due to evidence of harm. We have also highlighted this fact within our revised discussion.

14. DISCUSSION. As commented previously, authors should review entirely their discussion to put forward some more pertinent and interesting messages.

Response: Thank you for the feedback on our discussion. As mentioned in our responses to your previous comments (#1, #3, and #4), we have revised our discussion substantially to address your concerns.

15. ABSTRACT, Line 17: '…describe the findings descriptively…' could also be written as '…describe the findings…'

Response: You are correct. We have revised the sentence to read: "We report the findings descriptively…".

16. ABSTRACT, Line 30: 'The reporting of adverse events and harm-related endpoints varied by the reported presence of a DMC ($P < 0.001$) and the nature of the intervention ($P = 0.002$)'. Authors should provide precisions on the direction of the effect (e.g. reporting was more frequent with RCTs that reported a DMC) together with the p-value.

Response: We agree that this would be more informative. We have edited our abstract to include the following sentence: "Trials that reported a DMC compared to those that did not were more likely to report adverse events (n = 43/55, 78% vs. 100/245, 41%, $P < 0.001$) and harm-related endpoints (n = 52/55, 95% vs. 163/245, 67%, $P < 0.001$)." We made other minor edits to the abstract to maintain the 300 word count limit.