PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Adverse events associated with pediatric massage therapy: a systematic review
AUTHORS	Karkhaneh, Mohhamad; Zorzela, Liliane ; Jou, Hsing; Funabashi, Martha; Dryden, Trish; Vohra, Sunita;

VERSION 1 – REVIEW

REVIEWER	Reviewer name: Silvano Mior Institution and Country: Canadian Memorial Chiropractic College Competing interests: None
REVIEW RETURNED	05-Nov-2019

GENERAL COMMENTS	This is an important review that will contribute to the current gap in knowledge related to AE in massage therapy (MT). The following comments are offered in an effort to add clarity to certain aspects of the paper. Title: I would recommend changing the title to a "narrative review" rather than a "systematic review", particularly after reading the methods, where inclusion criteria was broad and no quality assessment of the included articles was considered. pg 6: lines 16-23: I appreciate little has been published in this area but it would be helpful to review the attempts that have been made to describe AE in MT, in adults and/or children [e.g Posadzki et al (2013), Yin et al (2014) etc] , and identify related weaknesses. In so doing, justification for this review will be strengthened. pg 6, line 52: suggest adding how search terms were identified, and if they were modified for each of the databases searched. pg7, line 7-23: please define what is meant by "primary study"; please define the MT (you would have had to so in the search terms, thus providing guidance to readers); describe how you defined AE. Also, be specific as to what was excluded (e.g. abstracts); which conditions (inclusion suggested "not limited by medical condition"). Finally, clarify the actual selection of studies, i.e. were two reviewers involved in both the screening and determining eligibility of studies? Also, did you calculate number of discrepancies, e.g. can you calculate agreement? pg 7, lines 32-37: Tables had different headings than those listed study characteristics in the methods - please clarify. P7: Appendix A not included, thus unable to assess search terms. Pg 8: Table 1: there is an "*" following CTCAE but no legend, Pg 8, line 37: as above, if duplicate assessments by reviewers were performed, what was the agreement between reviewers? Pg 8, line 51: why was risk of bias not assessed? Please explain and also consider including in the limitation section. pg 8, line 51: please describe the "subgroup analysis" performed. Table 2: the heading "sex, grouping" is not clear - what is meant by grouping? Also, should you include "number of subjects" to reflect the numeric value in the column? the column "control" - perhaps consider "comparison" as not all appear to be control groups.

	The final column heading, "Reported conclusion" is not clear - is this referring to AE's were reported? pg 9, line 22: please specify the n refers to number of subjects per study, unlike in previous sentence where n=number of studies. pg 9, line 24: please clarify nature of "controlled studies". pg 9, line 38: the figure presented was very difficult to read. I suggest also changing format consistent with typical PRISMA reporting, as current organizations of boxes make it difficult to follow. pg 13, lines 12-14: Given the results provided in Tables 3-5, this statement is incorrect; there were no studies reported where MT was provided by a registered massage therapist. It is also unclear if a trained massage practitioner is registered in the country they practice. pg 16, line 3: please clarify what is meant by controlled clinical trial. pg 16, line 5: Please confirm that the comment "All children ..." is correct, as it appears inconsistent with values reported in Table 4. pg 21: Discussion: I would recommend that the discussion be re-written to better reflect the findings of the review. Currently, it focuses on serious AE associate with VWM. I suggest that first paragraph provide an overview of relevant findings in this review. This could be followed by discussing and contrasting the findings in context to what is known in the literature rather than focusing only on VWM. In the next paragraph you can discuss the clinical implications of the work; followed by limitations. Then close with the next steps - where should the research go in the future. pg 21, lines 7-12: I don't agree with the statement made in the last sentence of second paragraph. According to the results, there were no studies involving registered or licensed MT. There were AEs in "trained" individuals, thus the closing wording is incorrect. In consideration that no incidence or relative risk information is presented, one must take care in surmising the extent of MT being "safe" pg 23, lines 14-19: Again, disagree with the wording of the second sentence, no reported study (based on content of results and tables) included licensed or registered MT. Please revise.

REVIEWER	Reviewer name: Nigel Hall Institution and Country: University Of Southampton UK Competing interests: None declared
REVIEW RETURNED	27-Nov-2019

GENERAL COMMENTS	A systematic review of adverse events in massage therapy in children. I have reviewed primarily the methodology, having little knowledge of the subject area. The PRISMA guidance has been followed, the search strategy is broad. I believe the authors should make more of the fact that 50% of studies made no mention of whether an AE occurred or not. It is therefore not know if it occurred and is not reported or if it never occurred. This may significantly influence the overall results and I think warrants greater discussion perhaps with a call for standardised reporting after this type of intervention. Indeed I would make a comment related to this part of the conclusion. I would be interested to see information on timing of AE in relation to timing of the MT please. What do the authors mean by ‘trained massage therapist’? The number of papers using this type of therapist is small and the conclusion can therefore be challenged. The association with volvulus without malrotation is interesting and certainly concerning.

	I would caution against advising for more research into this AE. One might argue that for such a catastrophic complication there is adequate evidence already to suggest it should not be performed. Perhaps the authors could discuss whether there is evidence to support any benefit of the intervention for this population of patients and consider updating their conclusion in light of this.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

This is an important review that will contribute to the current gap in knowledge related to AE in massage therapy (MT). The following comments are offered in an effort to add clarity to certain aspects of the paper.

Title: I would recommend changing the title to a "narrative review" rather than a "systematic review", particularly after reading the methods, where inclusion criteria was broad and no quality assessment of the included articles was considered.

RE: Thank you. We respectfully disagree with the reviewer and would like to articulate methodological details that confirm the work is best described as a systematic review (broad inclusion criteria and presence/absence of quality assessment are not what determine the difference between a systematic and narrative review).

Systematic reviews of adverse events are highly encouraged to include multiple study designs, including retrospective studies, as case reports and case series are the primary source of adverse events reporting (Loke et al., 2007). Clinical trials are well known to be poor at identifying and reporting adverse events (Ioannidis et al., 2004, Edwards and Aronson, 2000, Ioannidis and Lau, 2001, Loke and Derry, 2001, Chou et al., 2010) and a review of adverse events solely including RCTs will likely overlook significant data on adverse events (Loke et al., 2007, Ioannidis et al., 2004). In this systematic review, we included all study designs (from case reports to RCTs) to identify all adverse events reported in the literature, making it the most systematic and comprehensive report of studies on the topic.

In addition, we followed multiple steps to promote methodological rigour consistent with systematic reviews:

1- Searches: we searched 7 electronic databases. As per the Cochrane handbook (Higgins et al., 2019), a systematic review should search at least of 2 databases. Our goal was a comprehensive assessment to ensure no relevant studies were missed. In addition, we had no limitation to language.

2- Screening: first and second level screening were done by 2 independent reviewers (done independently and in duplicate), assuring the most systematic method of screening was done, as recommended by the Cochrane Handbook.

3- Data extraction: data extraction was done using standardized a priori data extraction forms and also done independently and in duplicate, as recommended by the Cochrane handbook.

4- Quality assessment: Risk of bias assessment was not performed on this review for two main reasons:

a) there is no preferred method to assess quality of reporting of case series and case reports, which represented the majority of studies (11 of 16) included in this review.

b) the ROB tool for RCTs is designed to assess risk of bias with regards to efficacy, not adverse events (the senior author is a Co-convenor with the Cochrane collaboration adverse events methods group). When the ROB tool is used, its unclear if a low risk of bias for efficacy equally represents low risk of bias for adverse events. For example, once a study is randomized and double blinded to answer the efficacy question, its unclear if the adverse event assessment received the same attention, or the adverse events were reported after the end of the trial when participant and assessor masking may no longer have been present (Ioannidis et al., 2004, Edwards and Aronson, 2000, Ioannidis and Lau, 2001, Loke and Derry, 2001, Chou et al., 2010).

We added the following text to the manuscript in the data synthesis and now reads:

“Risk of bias was not assessed for two reasons: (i) the large majority of included studies (11 of 16) were case reports and case series, for which there is no universally accepted tool to assess bias; and (ii) with regards to RCTs, there is evidence that existing ROB tools are developed with efficacy as their main focus and that other factors should be considered for assessment of adverse events (Higgins et al., 2019).Until an appropriate tool is developed, we opt to not proceed with an assessment that may itself lead to bias. At present, there is no clear guidance on how to proceed with ROB assessment on studies reporting adverse events data.”

Finally, PRISMA asks authors to identify systematic reviews as such in their title(Moher et al., 2009), and we think it is important to be compliant with this reporting guideline.

pg 6: lines 16-23: I appreciate little has been published in this area but it would be helpful to review the attempts that have been made to describe AE in MT, in adults and/or children (Posadzki and Ernst, 2013) (Yin et al., 2014), and identify related weaknesses. In so doing, justification for this review will be strengthened.

RE: Thank you, we modified the text according to your suggestion. Line 16-23 and now reads:

“Most systematic reviews of pediatric massage therapy have focused on the effectiveness of MT rather than its safety (Bennett et al., 2013, Juneau et al., 2015, Shipwright and Dryden, 2012, Kulkarni et al., 2010, Beider and Moyer, 2007, Ireland and Olson, 2000, Field, 1998, Field, 1995, Moyer et al., 2011, Li et al., 2016, Field et al., 2010). One systematic review included two case reports of AEs among children (Ernst, 2003); one infant with scrotal hematoma associated with MT of a hydrocele (Ram et al., 1994) and a teenager with a popliteal artery pseudoaneurysm after MT overlying an exostosis (Kalinga et al., 1996). AEs reported for adults were more diverse, including: cerebrovascular accidents, disc herniation/spinal cord injury, thromboembolism (renal, pulmonary), soft tissue trauma, leg ulcer, and genitourinary injuries (Ernst, 2003, Posadzki and Ernst, 2013, Yin et al., 2014). These studies demonstrated that the nature and circumstances of MT AEs for adults are quite different from pediatrics.”

pg 6, line 52: suggest adding how search terms were identified, and if they were modified for each of the databases searched.

RE: We have added this info into the main document and now reads:

“We identified search terms by using key words including massage, adverse effects, and pediatrics and a combination of subject terms selected from the controlled vocabulary or thesaurus (e.g. MeSH - Medical Subject Headings, Emtree for Embase, CINAHL Headings and PsycINFO Thesaurus) with a wide range of free-text terms.”

pg7, line 7-23: please define what is meant by "primary study"; please define the MT (you would have had to so in the search terms, thus providing guidance to readers); describe how you defined AE. Also, be specific as to what was excluded (e.g. abstracts); which conditions (inclusion suggested "not limited by medical condition"). Finally, clarify the actual selection of studies, i.e. were two reviewers

involved in both the screening and determining eligibility of studies? Also, did you calculate number of discrepancies, e.g. can you calculate agreement?

RE: Thank you. We will address each point raised:

Define primary study:

A primary study is one that reports data firsthand, as opposed to a review (which summarizes information from primary studies). We have added the search strategy, so the reader can appreciate how extensive and comprehensive our search terms were and the specific massage terms used. This information has been added to the manuscript.

Define AE:

Adverse events were defined as per the Cochrane handbook as: ´adverse event is an unfavourable or harmful outcome that occurs during, or after, the use of a drug or other intervention, but is not necessarily caused by it'. This definition has also been added to the manuscript and now reads:

"´adverse event is an unfavourable or harmful outcome that occurs during, or after, the use of a drug or other intervention, but is not necessarily caused by it'"

Be specific as to what was excluded (e.g. abstracts): our study selection criteria states that we included "primary study published in a peer-reviewed journal"; hence, we have not included any abstracts, grey literature, or non peer reviewed material.

Clarify which conditions were excluded: We did not exclude based on condition; all studies were eligible if the participant received massage therapy. This has been clarified in the manuscript.

Clarify the selection of studies:

We have added that study selection and data extraction were done by two independent reviewers and now reads:

"Two reviewers performed study selection, independent and in duplicate, and discrepancies were resolved by consensus or a third reviewer."

Did you calculate number of discrepancies?

As recommended by the Cochrane collaboration, discrepancies were discussed, and consensus was reached between the two reviewers. This has been added to the manuscript. It is not recommended to report agreement (kappa), hence this was not done (Moher et al., 2009).

pg 7, lines 32-37: Tables had different headings than those listed study characteristics in the methods - please clarify.

RE: We have changed the sentence and now reads:

"Specific information included: indication for treatment, AE severity, N of event, age, (age/weight and age at AE for newborn babies), gender, type of MT, MT practitioner, outcome with severity grade, and notes from the authors."

P7: Appendix A not included, thus unable to assess search terms.

RE: Search strategy is added as: Appendix: Medline search strategy for pediatric massage therapy adverse events

Pg 8: Table 1: there is an "*" following CTCAE but no legend,

RE: Thank you. Table 1 has been re-done.

Pg 8, line 37: as above, if duplicate assessments by reviewers were performed, what was the agreement between reviewers?

RE: Differences were discussed, and consensus was reached between the two reviewers. There were no residual discrepancies. This information has been added to the manuscript.

Pg 8, line 51: why was risk of bias not assessed? Please explain and also consider including in the limitation section.

RE: Risk of bias assessment was not performed on this review for two main reasons:

a) there is no preferred method to assess quality of reporting of case series and case reports, which represented the majority of studies (11 of 16) included in this review.

b) the ROB tool for RCTs is designed to assess risk of bias with regards to efficacy, not adverse events (the senior author is a Co-convenor with the Cochrane collaboration adverse events methods group). When the ROB tool is used, its unclear if a low risk of bias for efficacy equally represents low risk of bias for adverse events. For example, once a study is randomized and double blinded to answer the efficacy question, its unclear if the adverse event assessment received the same attention, or the adverse events were reported after the end of the trial when participant and assessor masking may no longer have been present (Ioannidis et al., 2004, Edwards and Aronson, 2000, Ioannidis and Lau, 2001, Loke and Derry, 2001, Chou et al., 2010).

We added the following text to the manuscript and now reads:

“Risk of bias was not assessed for two reasons: (i) the large majority of included studies (11 of 16) were case reports and case series, for which there is no universally accepted tool to assess bias; and (ii) with regards to RCTs, there is evidence that the ROB tools are mainly developed with efficacy in focus and for assessment of adverse events other factors should be considered that largely influence adverse events reporting . Until an appropriate tool to assess adverse events data in studies included in a systematic review of adverse events is developed, we opt to not proceed with such assessment which may lead to bias itself, since there is no clear guidance on how to proceed with ROB assessment on studies reporting adverse events data”(Higgins et al., 2019, Chou et al., 2010, Ioannidis et al., 2004, Edwards and Aronson, 2000, Ioannidis and Lau, 2001, Loke and Derry, 2001).

pg 8, line 51: please describe the "subgroup analysis" performed.

RE: we added the following sentence to the manuscript and now reads:

“Subgroup analysis was done separating study designs by severity of adverse events to assess if any specific type of massage would lead to more severe events.”

Table 2: the heading "sex, grouping" is not clear - what is meant by grouping?

RE: We deleted the original Table 2 based on the Editor’s suggestion and the current table 2 does not have this statement.

Also, should you include "number of subjects" to reflect the numeric value in the column?

RE: We now provide this information under the heading “n event/N population”.

the column "control" - perhaps consider "comparison" as not all appear to be control groups.

RE: Modified as per suggestion.

The final column heading, "Reported conclusion" is not clear - is this referring to AE's were reported?

RE: The purpose of this column was to summarize was the overall conclusion about the

intervention or the case report of AE. We modified it to:

“Notes from study authors”

pg 9, line 22: please specify the n refers to number of subjects per study, unlike in previous sentence where n=number of studies. pg 9, line 24: please clarify nature of "controlled studies".

RE: Done.

pg 9, line 38: the figure presented was very difficult to read. I suggest also changing format consistent with typical PRISMA reporting, as current organizations of boxes make it difficult to follow.

RE: Done.

pg 13, lines 12-14: Given the results provided in Tables 3-5, this statement is incorrect; there were no studies reported where MT was provided by a registered massage therapist. It is also unclear if a trained massage practitioner is registered in the country they practice.

RE: We removed that sentence from the paragraph.

pg 16, line 3: please clarify what is meant by controlled clinical trial.

RE: We meant a non-randomized controlled trial (NRCT). We modified that and now reads:

“ There was one non-randomized controlled trial (NRCT)”

pg 16, line 5: Please confirm that the comment "All children ..." is correct, as it appears inconsistent with values reported in Table 4.

RE: Thank you, this has been changed to:

“Those children …”

pg 21: Discussion: I would recommend that the discussion be re-written to better reflect the findings of the review. Currently, it focuses on serious AE associate with VWM. I suggest that first paragraph provide an overview of relevant findings in this review. This could be followed by discussing and contrasting the findings in context to what is known in the literature rather than focusing only on VWM. In the next paragraph you can discuss the clinical implications of the work; followed by limitations. Then close with the next steps - where should the research go in

RE: We have modified the discussion as per your suggestion. We kept one paragraph on VWM since we believe the finding of VWM in preterm infants after abdominal massage was a significant finding for this review.

Now the discussion reads:

“To the best of our knowledge, this is the first systematic review to primarily assess AEs associated with pediatric MT.

After extensive search in multiple databases and including non-English studies we identified 64 adverse events associated with pediatric massage. These events were further classified in mild, moderate and severe by our team. Moderate events were the most frequently identified, closely

followed by mild and severe adverse events. Of serious AE identified, abdominal massage to stimulate peristalsis in preterm infants was identified as been associated with multiple reports of volvulus without malrotation.

Volvulus without malrotation (VWM) is considered a rare event, often associated with preterm birth (<30weeks GA) and low birth weight (<1000 g) (Maas et al., 2014, Zweifel et al., 2013, Billiemaz et al., 2001, Yarkin et al., 2019, Mark et al., 2013, Raherison et al., 2012, Drewett and Burge, 2009). We identified 17 events of volvulus without malrotation in pre term infants on this review, mostly very low birth weight (VLBW). In five VWM, the authors did not considered MT as a risk factor for the volvulus while for the remaining 12 events, MT was considered a risk factor by authors (Maas et al., 2014). The identified cases of VWM led to a routine change in their centers to avoid abdominal massage in preterm infants due the risk of volvulus. All the authors report that changing their practice to avoid abdominal massage in preterm neonates generated no further events of VWM in their centers (Maas et al., 2014).The large majority of the events reported in this review occurred in infants born at <30 weeks GA and weighted <1000g, which may be a confounder, since VLBW is consider a risk factor for VWM. Others have reported findings of VWM in pre term infants. Horsch has reviewed cases of volvulus in surgical referral center over 8 years in Germany and found 3 preterm infants with VWM, of these, 1 received properistaltic abdominal massage (Horsch et al., 2016). Yarkin after reviewing 2 years of data from an epidemiological survey identified that 12 of 15 cases of low birth weight (<1500g) pre term infants with VWM had abdominal massages prior to the event, although due the uncontrolled study design they were unable to determine abdominal massage as a direct risk factor (Yarkin et al., 2018).. Horsch and Yarkin were not included this review because they did not look into MT specifically. Several authors (Maas et al., 2014, Horsch et al., 2016, Berseth, 1989Kargl, 2015 #481) have speculated on other risk factors for VWM, including immaturity of the intestine resulting in prolonged transit times and stool retention, use of continuous positive airway pressure (CPAP) or other intensive respiratory support, and female gender. Several of these risk factors were also identified in the included reports and it makes difficult to assure causality of MT in the development of VWM.

Massage therapy has been studied in children, but with the main focus of assessing efficacy, very little has been reported on adverse events(Bennett et al., 2013, Juneau et al., 2015, Shipwright and Dryden, 2012, Kulkarni et al., 2010, Beider and Moyer, 2007, Ireland and Olson, 2000, Field, 1998, Field, 1995, Moyer et al., 2011, Li et al., 2016, Field et al., 2010). In contrast, adults report a wide range of adverse events associated with massage, from mild events to severe, including: cerebrovascular accidents, thromboembolism (renal, pulmonary), skin injury, leg ulcer, genitourinary injuries disc herniation, soft tissue trauma, neurologic compromise, spinal cord injury, dissection of the vertebral arteries" (Yin et al., 2014). (Ernst, 2003) (Posadzki and Ernst, 2013).

Similar to other studies (Anderson and Choonara, 2010) this review reveals a lack of adverse event reporting. The majority of included studies did not report if the intervention was associated with any adverse event or not, leading to significant reporting bias from the primary studies and carried into the review level, confirming the true magnitude of adverse events related to massage in pediatrics remains unknown.

Strengths and limitations

This review was comprehensive in searching peer-reviewed publications in seven databases from inception onwards, not limiting by language. We minimized subjective errors by independent duplicate screening, study selection, and data extraction. This SR meticulously collected data from included studies and classified severities of AEs according to well-established criteria. We have incorporated to this review all relevant studies, including case reports, case series, and clinical trials in an effort to be comprehensive. In addition, we reported studies that reported if AEs occurred or not, whether they mentioned if AEs were assessed or not, except for case reports which were only included if an event was reported.

As any study, this review has limitations. The majority of the AEs, including 4 out of 5 severe AE, were described in case series/reports. Although case studies are well-recognized sources of reports on severe, unusual, and/or rare AEs, conclusions drawn from such sources are limited and subject to confounding. Additionally, given the limited information and heterogeneity of the data reported in the included studies, it was not possible to identify any conclusive association between MT and specific AEs, except for ba sha massage which caused a characteristic bruising.

Clinical relevance of review findings:

The most significant review finding warranting clinical surveillance is the association between abdominal massage and VWM in preterm infants. There are several confounders impeding this review to define a causal relationship between neonatal abdominal massage and VWM, but the existing reports should be taken into careful consideration and abdominal massage should be considered with caution on this population."

pg 21, lines 7-12: I don't agree with the statement made in the last sentence of second paragraph. According to the results, there were no studies involving registered or licensed MT. There were AEs in "trained" individuals, thus the closing wording is incorrect. In consideration that no incidence or relative risk information is presented, one must take care in surmising the extent of MT being "safe"

RE: We have removed this sentence.

pg 23, lines 14-19: Again, disagree with the wording of the second sentence, no reported study (based on content of results and tables) included licensed or registered MT. Please revise.

RE: We have removed this sentence.

Reviewer: 2

Comments to the Author

A systematic review of adverse events in massage therapy in children. I have reviewed primarily the methodology, having little knowledge of the subject area. The PRISMA guidance has been followed, the search strategy is broad.

I believe the authors should make more of the fact that 50% of studies made no mention of whether an AE occurred or not. It is therefore not know if it occurred and is not reported or if it never occurred. This may significantly influence the overall results and I think warrants greater discussion perhaps with a call for standardised reporting after this type of intervention. Indeed I would make a comment related to this part of the conclusion.

I would be interested to see information on timing of AE in relation to timing of the MT please.

What do the authors mean by 'trained massage therapist'? The number of papers using this type of therapist is small and the conclusion can therefore be challenged.

The association with volvulus without malrotation is interesting and certainly concerning. I would caution against advising for more research into this AE. One might argue that for such a catastrophic complication there is adequate evidence already to suggest it should not be performed. Perhaps the authors could discuss whether there is evidence to support any benefit of the intervention for this population of patients and consider updating their conclusion in light of this.

Thank you for your comments.

RE: We have modified the text according to your suggestions, please see below:

1. I believe the authors should make more of the fact that 50% of studies made no mention of whether an AE occurred or not

RE: Thank you for pointing out this relevant point. We added a paragraph in the discussion highlighting this issue and modified our conclusion as well:

´Similar to other studies (Anderson and Choonara, 2010) this review reveals a lack of adverse event reporting. The majority of included studies did not report if the intervention was associated with any adverse event or not, leading to significant reporting bias from the primary studies and carried into the review level, confirming the true magnitude of adverse events related to massage in pediatrics remains unknown.´

Conclusion: We were able to identify a range of AEs associated with MT use, from mild to severe. Unfortunately, the majority of included studies did not report if an adverse event occurred or not, leading to publication bias. This review reports an association between abdominal massage with volvulus without malrotation in preterm infants, it is still to be defined if this is casual or not, but the our findings warrant caution in the use of abdominal massage in pre term infants until further investigation is done in the subject.

2- I would be interested to see information on timing of AE in relation to timing of the MT please.

RE: Studies reporting the timing of massage in relation to AE is entered into the tables, including when this information was not reported.

3- What do the authors mean by 'trained massage therapist'? The number of papers using this type of therapist is small and the conclusion can therefore be challenged.

RE: We have removed this statement form the conclusion.

4- The association with volvulus without malrotation is interesting and certainly concerning. I would caution against advising for more research into this AE. One might argue that for such a catastrophic complication there is adequate evidence already to suggest it should not be performed. Perhaps the authors could discuss whether there is evidence to support any benefit of the intervention for this population of patients and consider updating their conclusion in light of this.

RE: We agree the finding of VWM in newborns receiving abdominal massage is concerning, but another reviewer asked us to reduce the discussion around VWM, so we have kept it to one paragraph to accommodate both of your suggestions. We have modified our conclusion to reflect a more cautionary warning for this population. We also have listed the proposed benefits of abdominal massage. Please see the discussion and conclusion that now reads:

"Volvulus without malrotation (VWM) is considered a rare event, often associated with preterm birth (<30weeks GA) and low birth weight (<1000 g) (Maas et al., 2014, Zweifel et al., 2013, Billiemaz et al., 2001, Yarkin et al., 2019, Mark et al., 2013, Raherison et al., 2012, Drewett and Burge, 2009). We identified 17 events of volvulus without malrotation in preterm infants on this review, mostly very low birth weight (VLBW). In five VWM, the authors did not considered MT as a risk factor for the volvulus while for the remaining 12 events, MT was considered a risk factor by authors (Maas et al., 2014). The identified cases of VWM led to a routine change in their centers to avoid abdominal massage in preterm infants due the risk of volvulus. All the authors report that changing their practice to avoid abdominal massage in preterm neonates generated no further events of VWM in their centers (Maas et al., 2014).

The large majority of the events reported in this review occurred in infants born at <30 weeks GA and weighted <1000g, which may be a confounder, since VLBW is consider a risk factor for VWM. Others have reported findings of VWM in pre term infants. Horsch has reviewed cases of volvulus in surgical referral center over 8 years in Germany and found 3 preterm infants with VWM, of these, 1 received properistaltic abdominal massage (Horsch et al., 2016). Yarkin after reviewing 2 years of data from an epidemiological survey identified that 12 of 15 cases of low birth weight (<1500g) pre term infants with VWM had abdominal massages prior to the event, although due the uncontrolled study design they were unable to determine abdominal massage as a direct risk factor (Yarkin et al., 2018).. Horsch and Yarkin were not included this review because they did not look into MT specifically. Several authors (Maas et al., 2014, Horsch et al., 2016, Berseth, 1989Kargl, 2015 #481) have speculated on other risk factors for VWM, including immaturity of the intestine resulting in prolonged transit times and stool retention, use of continuous positive airway pressure (CPAP) or other intensive respiratory support, and female gender. Several of these risk factors were also identified in the included reports and it makes difficult to assure causality of MT in the development of VWM.

Massage therapy has been studied in children, but with the main focus of assessing efficacy, very little has been reported on adverse events(Bennett et al., 2013, Juneau et al., 2015, Shipwright and Dryden, 2012, Kulkarni et al., 2010, Beider and Moyer, 2007, Ireland and Olson, 2000, Field, 1998, Field, 1995, Moyer et al., 2011, Li et al., 2016, Field et al., 2010). In contrast, adults report a wide range of adverse events associated with massage, from mild events to severe, including: cerebrovascular accidents, thromboembolism (renal, pulmonary), skin injury, leg ulcer, genitourinary injuries disc herniation, soft tissue trauma, neurologic compromise, spinal cord injury, dissection of the vertebral arteries" (Yin et al., 2014). (Ernst, 2003) (Posadzki and Ernst, 2013).

Similar to other studies (Anderson and Choonara, 2010) this review reveals a lack of adverse event reporting. The majority of included studies did not report if the intervention was associated with any adverse event or not, leading to significant reporting bias from the primary studies and carried into the review level, confirming the true magnitude of adverse events related to massage in pediatrics remains unknown.

Strengths and limitations

This review was comprehensive in searching peer-reviewed publications in seven databases from inception onwards, not limiting by language. We minimized subjective errors by independent duplicate screening, study selection, and data extraction. This SR meticulously collected data from included studies and classified severities of AEs according to well-established criteria. We have incorporated to this review all relevant studies, including case reports, case series, and clinical trials in an effort to be comprehensive. In addition, we reported studies that reported if AEs occurred or not, whether they mentioned if AEs were assessed or not, except for case reports which were only included if an event was reported.

As any study, this review has limitations. The majority of the AEs, including 4 out of 5 severe AE, were described in case series/reports. Although case studies are well-recognized sources of reports on severe, unusual, and/or rare AEs, conclusions drawn from such sources are limited and subject to confounding. Additionally, given the limited information and heterogeneity of the data reported in the included studies, it was not possible to identify any conclusive association between MT and specific AEs, except for ba sha massage which caused a characteristic bruising.

Clinical relevance of review findings:

The most significant review finding warranting clinical surveillance is the association between abdominal massage and VWM in preterm infants.

There are several confounders impeding this review to define a causal relationship between neonatal abdominal massage and VWM, but the existing reports should be taken into careful consideration and abdominal massage should be considered with caution on this population."

VERSION 2 – REVIEW

REVIEWER	Reviewer name: Nigel Hall Institution and Country: University of Southampton, UK Competing interests: None declared
REVIEW RETURNED	10-Mar-2020

GENERAL COMMENTS	The authors have made significant changes to this manuscript which have improved it form a methodological perspective. However these change are so extensive it now appears to be a paper written largely by the reviewers rather than the authors themselves. I remain unconvinced of the evidence of benefit of massage therapy.

VERSION 2 – AUTHOR RESPONSE

Thank you very much for this opportunity and for the valuable suggestions to improve our manuscript. This is a systematic review of adverse events of massage therapy in children, as an attempt to shine a light on the side effects of the interventions. The efficacy of massage therapy in children was not the focus of our manuscript and therefore not answered here.

All the suggested modifications were made into the manuscript.