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)	Ultrasound Imaging versus Palpation method for Diagnostic
	Lumbar Puncture in Neonates and Infants: A Systematic Review
	and Meta-analysis
AUTHORS	Olowoyeye, Abiola; Fadahunsi, Opeyemi; Okudo, Jerome;
	Opaneye, Oluwakare; Okwundu, Charles

VERSION 1 – REVIEW

REVIEWER	Reviewer name: Peter Flom
	Institution and Country: Peter Flom Consulting USA
	Competing interests: None
REVIEW RETURNED	15-Jan-2019

GENERAL COMMENTS	I confine my remarks to statistical aspects of this paper.
	The general approach is fine; I have some issues that need to be resolved before I can recommend publication.
	p. 3. line 45: What was the conflict among the studies? Please list effect sizes found in earlier work.
	Table 2: Please clarify what the first two columns in teh "Anticipated absolute effects" are. Are these rates of failure and traumatic tap?
	Please show the results for the secondary outcomes in section 2 b and c on p. 11

REVIEWER	Reviewer name: Saleh Almenawer
	Institution and Country: McMaster University, Canada
	Competing interests: None
REVIEW RETURNED	07-Feb-2019

GENERAL COMMENTS	In this study, the authors conduct a systematic review and meta- analysis comparing ultrasonography-guided lumbar puncture to traditional palpation-based lumbar puncture in the pediatric population. The topic of this study is of interest, especially to ER physicians, pediatricians, and those working to perform lumbar punctures in low-resource settings. We have the following concerns with this manuscript:
	 The authors cite a failure rate of lumbar puncture of 35%-65% in the pediatric population. This is a very large range. It would be more useful for the authors to present one concrete value that is either an average of what is available in the literature, or a weighted pooled estimate from their own analyses. The authors investigate the need for altering needle direction
	and re-insertion of the needle as one combined outcome.

Re-insertion of the needle carries different physiological risks than simple re-direction of the bevel. A new puncture creates a second opportunity for CSF leak, increases the opportunity for traumatic tap, and is thoroughly separate from the first puncture. It would be useful for authors to separate this topic into two outcomes: success on first attempt (need to re-insert needle), and need to re-orient needle. This would be more informative for clinicians.

- 3) The authors' review includes only 4 trials, all of which are relatively recent. The search strategy seems to exclude some key databases, such as EMBASE, and the Web of Science. Moreover, the authors do not provide clarification on whether they searched the grey literature. Including additional databases, as well as hand-searching and searching the grey literature may yield more eligible studies, which could further increase the statistical power of the review and provide clinicians a more comprehensive picture of the topic.
- 4) The authors state that punctures were performed by a variety of different health care practitioners, including physicians, medical students, nurse practitioners, and housestaff. All of these individuals have different levels of training and comfort in performing lumbar puncture. This is a significant confounder when measuring success of puncture and should be discussed further in the discussion section.
- 5) If possible, the authors should calculate pooled estimates (weighted averages) of each of their outcomes in order to provide clinicians insight into the true rates of outcomes with and without ultrasound guidance.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

I confine my remarks to statistical aspects of this paper.

The general approach is fine; I have some issues that need to be resolved before I can recommend publication.

1. p. 3. line 45: What was the conflict among the studies? Please list effect sizes found in earlier work.

Thanks for your comment; we modified the sentence to say: The few studies in this population have conflicting results on benefits of ultrasound imaging for pediatric lumbar puncture that this review aims to address.

The effect size or difference between the studies is what this entire meta-analysis investigates. We showed that in figure 3 based on the guided writing format for this review

Table 2: Please clarify what the first two columns in the "Anticipated absolute effects" are. Are these rates of failure and traumatic tap?

Thanks for the comment: we noticed the table headings were missing and fixed them. They now read as "risk with palpation" and "risk with ultrasound imaging"

Please show the results for the secondary outcomes in section 2 b and c on p. 11

Thank you for the comment, the results were adjusted thus:

- a. Needle insertion attempts: The studies23-26 looked at the number of needle attempts and they found that there was no difference in both groups in terms of the number of attempts required to obtain CSF successfully during a diagnostic lumbar puncture. Gorn et al23 reported that the number of attempts between the ultrasound and palpation groups were not significantly different. Neal et al24 reported that the median number of attempts was 1 in the ultrasound group and 2 in the palpation group but this was not statistically significant. Similarly, Kessler et al25 reported 2 attempts in both the ultrasound and palpation groups. Lam et al26 reported a median of 1.5 attempts compared to 2 attempts in the ultrasound group as against the palpation group respectively.
- b. Needle redirections: No study addressed the number of redirections.
- c. Length of Procedure: Two studies25,26 looked at differences based on length of procedure and both studies could not find a statistically significant difference in the median duration of the procedures in both the intervention and control groups. Kessler et al reported a median of 1.6 minutes (IQR=0.8 to 13.4) versus 4.2 minutes (IQR=0.8 to 5.2) in the ultrasound vs. the palpation group without any statistical significance especially when median ultrasound duration of 4.6 minutes(IQR=3 to 6.8) is added. Similarly, Lam et al26 had median duration of 197 seconds vs. 146 seconds in the ultrasound vs. the palpation group respectively without any statistical significance.

Reviewer: 2

In this study, the authors conduct a systematic review and meta-analysis comparing ultrasonography-guided lumbar puncture to traditional palpation-based lumbar puncture in the pediatric population. The topic of this study is of interest, especially to ER physicians, pediatricians, and those working to perform lumbar punctures in low-resource settings. We have the following concerns with this manuscript:

1) The authors cite a failure rate of lumbar puncture of 35%-65% in the pediatric population. This is a very large range. It would be more useful for the authors to present one concrete value that is either an average of what is available in the literature, or a weighted pooled estimate from their own analyses

Thank you for the comment. We adjusted the text to say up to 65%

2) The authors investigate the need for altering needle direction and re-insertion of the needle as one combined outcome. Re-insertion of the needle carries different physiological risks than simple re-direction of the bevel. A new puncture creates a second opportunity for CSF leak, increases the opportunity for traumatic tap, and is thoroughly separate from the first puncture. It would be useful for authors to separate this topic into two outcomes: success on first attempt (need to re-insert needle), and need to re-orient needle. This would be more informative for clinicians

Thanks for the comments: The initial write up was done for brevity and need to keep a certain word count. We however adjusted the outcomes into 2 separate outcomes as initially proposed in out outcome definition and the text now reads.

a. Needle insertion attempts: The studies23-26 looked at the number of needle attempts and they found that there was no difference in both groups in terms of the number of attempts required to obtain CSF successfully during a diagnostic lumbar puncture. Gorn et al23 reported that the number of attempts between the ultrasound and palpation groups were not significantly different. Neal et al24 reported that the median number of attempts was 1 in the ultrasound group and 2 in the palpation group but this was not statistically significant. Similarly, Kessler et al25 reported 2 attempts in both the ultrasound and palpation groups. Lam et al26 reported a median of 1.5 attempts compared to 2 attempts in the ultrasound group as against the palpation group respectively.

- b. Needle redirections: No study addressed the number of redirections.
- 3)The authors' review includes only 4 trials, all of which are relatively recent. The search strategy seems to exclude some key databases, such as EMBASE, and the Web of Science. Moreover, the authors do not provide clarification on whether they searched the grey literature. Including additional databases, as well as hand-searching and searching the grey literature may yield more eligible studies, which could further increase the statistical power of the review and provide clinicians a more comprehensive picture of the topic.

Thanks for the comment. The low number of studies reflects the limited amount of pediatric research. Our reported strategy however was incomplete as the librarian sent over the search strategy and from her report, she searched multiple databases. We thus modified the initial description to include all the searched databases. PRISMA-P guidelines were used in the preparation of the systematic review and meta-analysis.24 We searched, the National Library of Medicine through PubMed (from 1990 to March 2018), the Cochrane Central Register of Controlled Trials (to the first quarter of 2018) and ClinicalTrials.gov (to the first quarter of 2018). We also searched EMBASE, Scopus and Web of science (to the first quarter of 2018). The search was done without any restrictions by language or publication status. The search strategy was conducted with the aid of the institutional librarian. The following search terms were used: "ultrasound" and "lumbar puncture". The search strategy also included multiple synonyms, abbreviations, and related keywords for each of these terms. We also examined the reference lists of retrieved original and review articles for any additional studies.

4)The authors state that punctures were performed by a variety of different health care practitioners, including physicians, medical students, nurse practitioners, and housestaff. All of these individuals have different levels of training and comfort in performing lumbar puncture. This is a significant confounder when measuring success of puncture and should be discussed further in the discussion section

Thanks for the comment: It is true that the results of the individual studies may be confounded by the performers experience however, it can also be a strength as the individual studies aim to adopt a generalizable sample population.

We thus added to the text The inclusion of front level providers like medical students, residents, fellows, NPs and Attendings in the included studies is a strength of this review as these are the usual performers of lumbar punctures in medical care.

5) If possible, the authors should calculate pooled estimates (weighted averages) of each of their outcomes in order to provide clinicians insight into the true rates of outcomes with and without ultrasound guidance.

Thanks for the comments, were possible the weighted averages are provided. For other outcomes, we have expanded the results section and added: Needle insertion attempts: The studies23-26 looked at the number of needle attempts and they found that there was no difference in both groups in terms of the number of attempts required to obtain CSF successfully during a diagnostic lumbar puncture. Gorn et al23 reported that the number of attempts between the ultrasound and palpation groups were not significantly different. Neal et al24 reported that the median number of attempts was 1 in the ultrasound group and 2 in the palpation group but this was not statistically significant. Similarly, Kessler et al25 reported 2 attempts in both the ultrasound and palpation groups. Lam et al26 reported a median of 1.5 attempts compared to 2 attempts in the ultrasound group as against the palpation group respectively.

c. Needle redirections: No study addressed the number of redirections.

Length of Procedure: Two studies25,26 looked at differences based on length of procedure and both studies could not find a statistically significant difference in the median duration of the procedures in both the intervention and control groups. Kessler et al reported a median of 1.6 minutes (IQR=0.8 to 13.4) versus 4.2 minutes (IQR=0.8 to 5.2) in the ultrasound vs. the palpation group without any statistical significance especially when median ultrasound duration of 4.6 minutes(IQR=3 to 6.8) is added. Similarly, Lam et al26 had median duration of 197 seconds vs. 146 seconds in the ultrasound vs. the palpation group respectively without any statistical significance