NUMBER ID: 02581873

PUBLICATION NAME: World Journal of Clinical Pediatrics

TITLE

Pediatric Anesthesia Emergence Delirium Scales: A diagnostic meta-analysis.

Reviewer#1:

Specific Comments to Authors:

1. The actors need to parameterize the results to see if the method is really effective. Without the defined threshold of the exam, we cannot know if it is real. So it is important to expand the discussion and the conclusion that seems to be a good technique, however more future studies are needed.

   Many thanks for the suggestion.

   Firstly, we absolutely agree with the reviewer about the parameterization of the test results as the heterogeneity of the analysis is substantial, and therefore the summary estimate might not represent the individual studies adequately. To further analyse this, we have done a leave-one-out cross validation on the data. This information has been included in the statistical analysis subsection, results and discussion sections of the revised manuscript as:

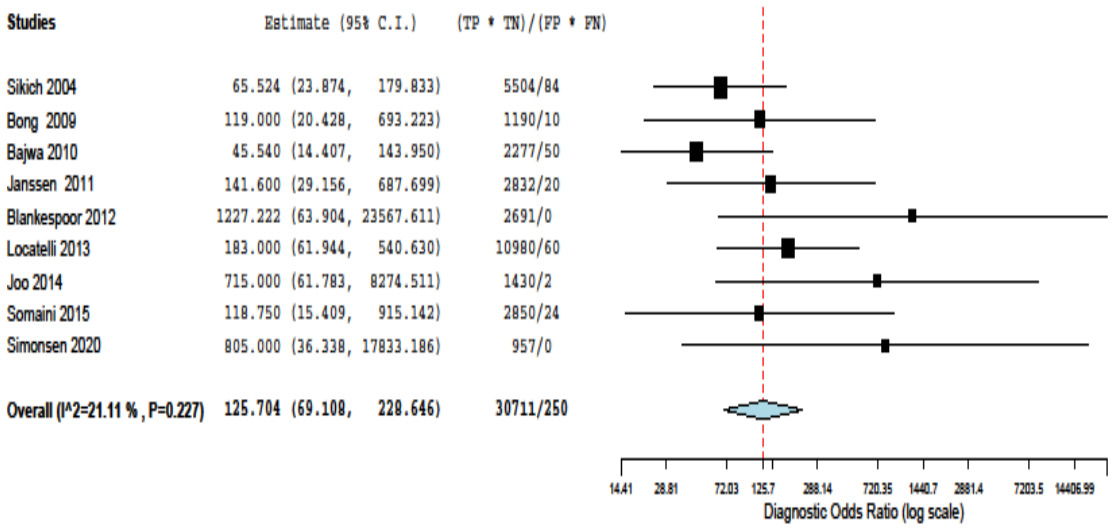
"…was done using univariate meta-regression. In addition, as the heterogeneity was substantial, it was reasoned that the summary statistics might not represent the individual studies adequately. Therefore, as a post hoc test to parametrise the summary DOR, we conducted a leave-one-out cross validation. We calculated the 95% Confidence Interval (95% CI) when indicated. The analyses were done with the METANDI module of STATA (version 16). We conducted the leave-one-out cross validation using the software Open-Meta™ meta-analysis software (Brown University, Providence RI, USA).. [11] …."

[Page: 9-10; Subsection: Statistical analysis]

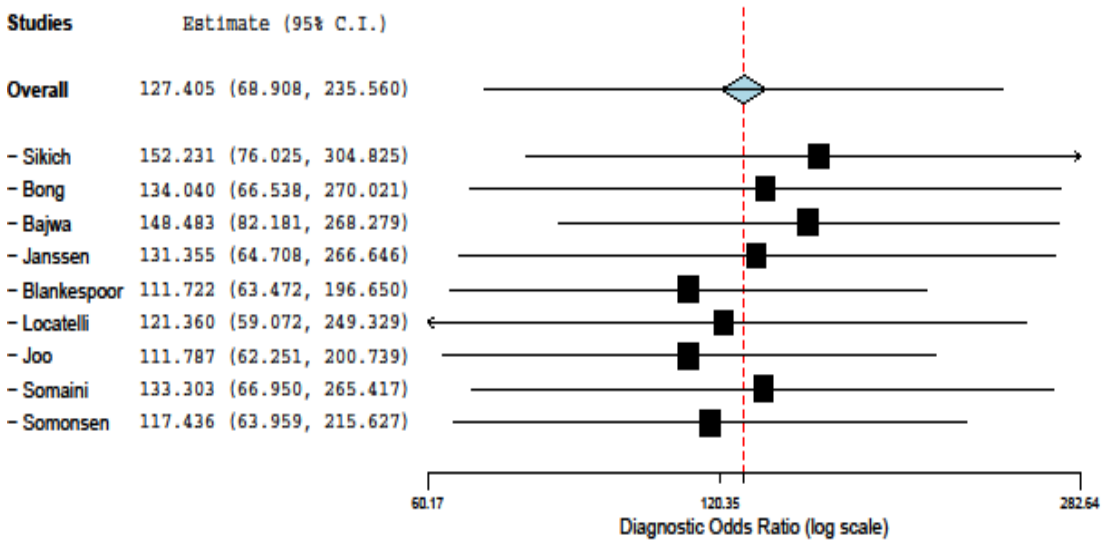
Secondly, the leave-one-out cross validation demonstrated that the summary Diagnostic odds ratio (DOR) increased or decreased when certain individual studies where left-out, presenting the influence of each study on the summary DOR.

The summary DOR and leave-one-out cross validation DOR (with OpenMeta) are given below:

<u>Summary DOR</u>



Leave-one-out cross validation DOR



This information has been added as text in the revised manuscript as:

"…The summary DOR for all PAEDS cut-off scores together was 148.33 (95% CI=48.32, 455.32). With the leave-one-out cross validation, the individual

studies significantly contributed to the summary DOR in a descending order from the study by Sikich et al [8] at the top [DOR=152.23 (95%CI= 76.23, 304.82), followed by Bajwan et al [13] [DOR=148.48 (95%CI= 82.18, 268.27), Bong et al [12][DOR=134.04 (95%CI= 66.53, 270.02), Somaini et al [17][DOR=133.30 (95%CI= 66.95, 265.41), Janssen et al [14][DOR=131.35 (95%CI= 64.70, 266.64), Locatelli et al [15][DOR=121.36 (95%CI= 59.72, 249.32), Simonson et al [18][DOR=117 (95%CI= 76.23, 304.82), Joo et al [16][DOR=111.78 (95%CI= 62.25, 200.73) and finally Blankespoor et al [9][DOR=111.72 (95%CI= 63.47, 196.65)."

[Page: 11; Section: Results]

Thirdly, we have included the strength of the cross validation in the discussion section as:

"….There was substantial heterogeneity in the diagnostic accuracy parameters of the PAEDS, which was partly explained by the setting of the occurrence of EmD and the reference standard used. The role of each individual study in the summary DOR was further explored with a range of 111 to 152, adding strength to the method of this meta-analysis. The…"

[Page: 12; Section: Discussion]

Finally, the reference for the leave-one-out cross validation DOR done with the OpenMeta is added in the revised manuscript as:

"…11. Wallace BC, Schmid CH, Lau J, Trikalinos TA. Meta-Analyst: software for meta-analysis of binary, continuous and diagnostic data. BMC Med Res Methodol. 2009 Dec 4;9:80. [doi: 10.1186/1471-2288-9-80. PMID: 19961608]."….

[Page: 18; Section: Reference]

Authors should present summary statements about the characteristics, quantity, quality, consistency of findings, and applicability of the studies included in the review.

Thank you for the suggestion. We have revised the manuscript and summary statements have been included characteristics, quantity, quality, consistency of findings, and applicability as:

Table 1: The data on the methodology and epidemiology of included studies. [7, 9,11-17]

(Page: 20; Table 1).

Figure 2: The Quality appraisal using the revised Diagnostic Accuracy Studies (QUADAS-2) for individual studies (2A) and average quality across studies (2B).

Consistency measures have been included as I^2 index, and substantial heterogeneity has been analysed further with and has been included in the revised manuscript as :

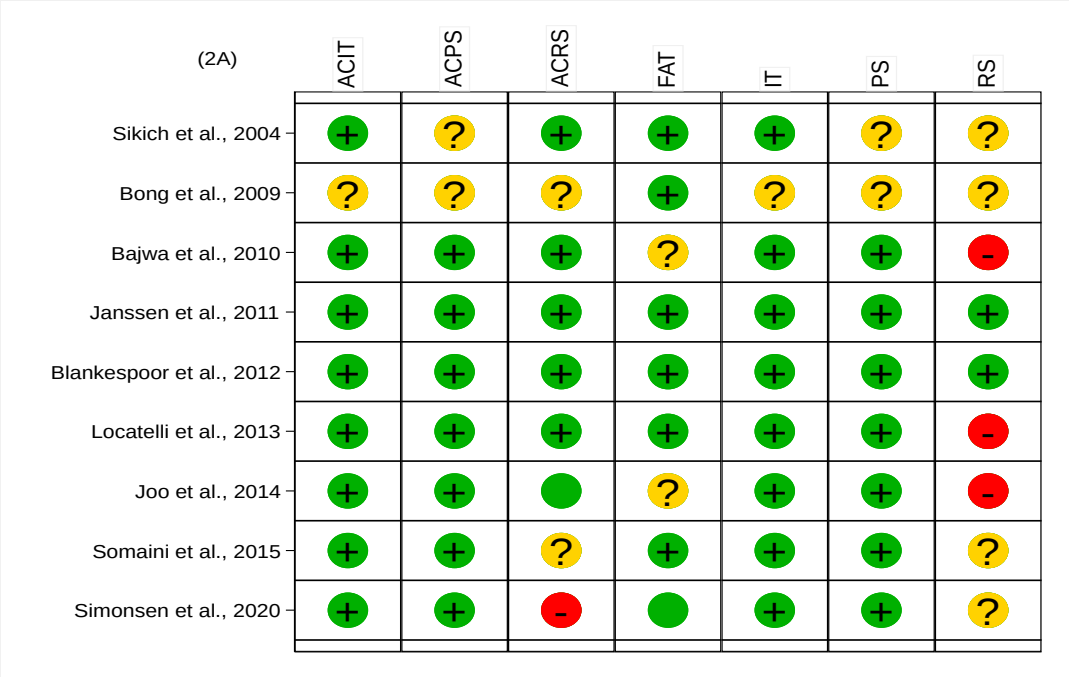
"...Expecting heterogeneity to start with, the use of random effects models, exploring the heterogeneity with meta-regression, subgroup analysis and the leave-one-out cross validation has strengthened the meta-analysis."…

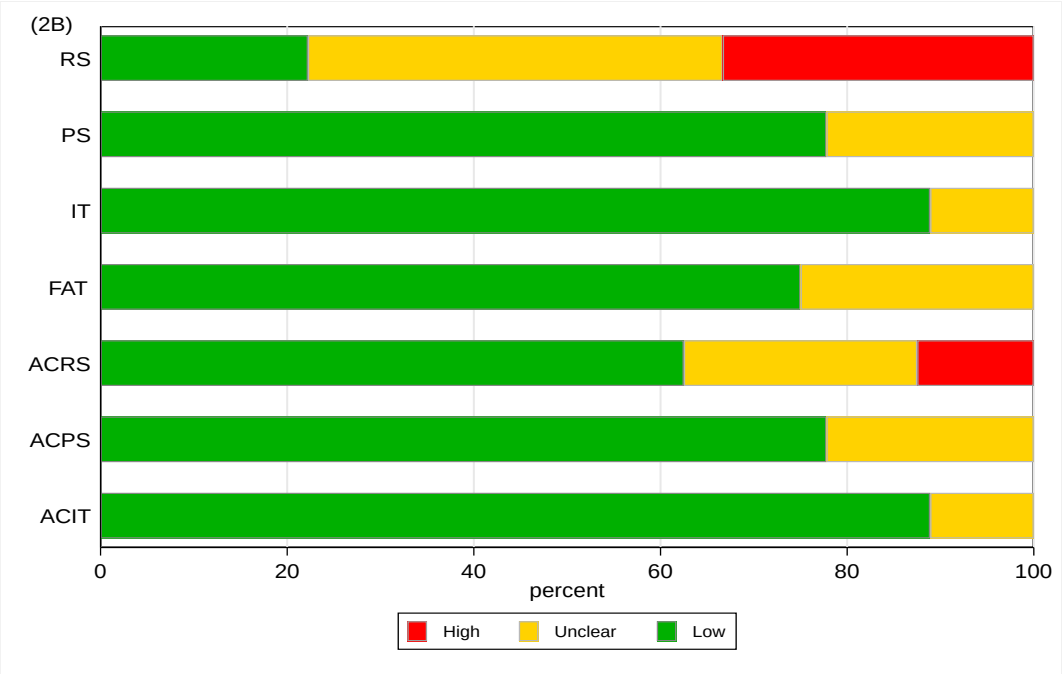
(Page 14: Section: Discussion)

Table 1: The data on the methodology and epidemiology of included studies.[7, 9,11-17]

Study	Sample size	Prevalence of PED	Sn(%)	Sp(%)	Setting	Age (years)	PEDS Cut-off	Reference standard
Sikich et al., 2004	100	11%	64	86	OP	1.6-2	≥10	Dimenhydrinate treatment
Bong et al., 2009	136	8.6%	85	96	OP	2-12	≥10	Clinical observation
Bajwa et al., 2010	117	32%	100	95	IP	1-18	≥12	Clinical observation
Janssen et al., 2011	154	16.9%	91	98	IP	1-17	≥8	DSM-IV interview for delirium
Blankespoor et al., 2012	144	16%	100	97	IP	1-18	≥8	Clinical observation
Locatelli et al., 2013	260	25%	93	94	IP	1-3	≥9	Clinical observation
Joo et al., 2014	90	25.5%	94	97	IP	2-5	≥16	Clinical observation
Somaini et al., 2015	150	21%	96	80	IP	1-7	≥9	Clinical observation
Simonsen et al., 2020	100	13.2%	86	100	IP	0.25-16	≥10	Clinical observation

Figure 2: The Quality appraisal using the revised Diagnostic Accuracy Studies (QUADAS-2) for individual studies (2A) and average quality across studies (2B).





It is important to highlight the strengths of the evidence as well as its potential limitations. When the review contains many large studies with very similar results, this may be mentioned as reinforcing the strength of the evidence.

Thank you for the suggestion. We had a review of the primary data and we find that none of the studies had duplicated data sets, same study sample/population, and similar selection process of participants or same group of authors with similar interpretation of results. This information has been included in the reviewed manuscript in the discussion section as:

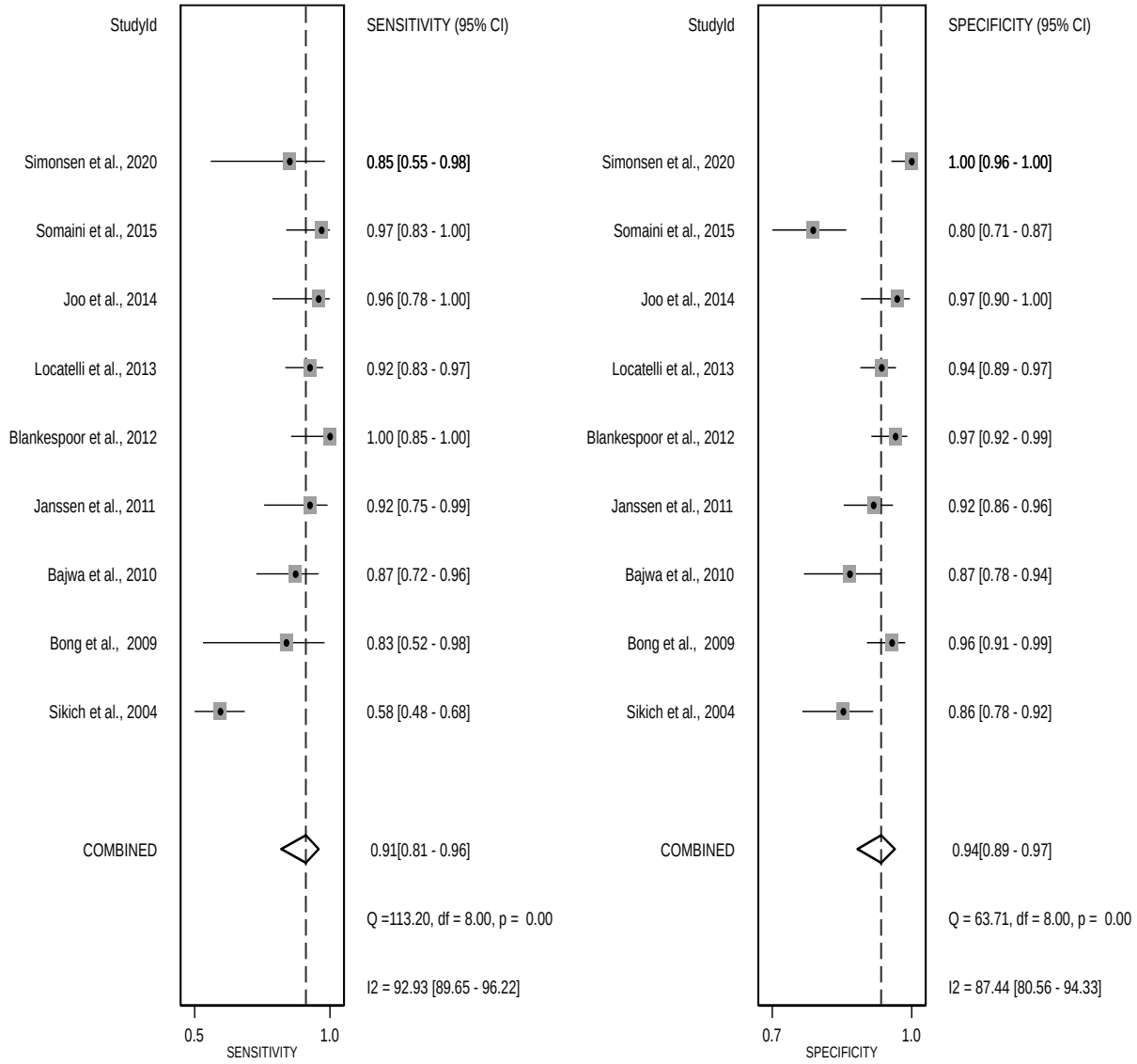
"… There was no publication bias. The quality appraisal showed that ***the most*** *common bias across studies was documenting the reference standards and applicability of the reference standards. Overall, the studies were of moderate quality. The absence of very large studies, duplicated data sets, same study sample/population, and similar selection process of participants or same group of authors with similar interpretation of results has minimized the skewing of our summary findings.*

The AUC-SROC for PAEDS in diagnosing…"

[Page: 13; Section: Discussion]

For comparative questions, the evidence will be stronger if all results were obtained in fully paired (within-study) or randomised comparative accuracy studies, and if the superiority of one test over another is consistent across included studies.

Thank you for the suggestions. We have compared the paired sensitivity and specificity within each study. The analysis is shown below.



We have summarised the findings as text in the revised manuscript as:

"… When we analysed the sensitivity-specificity pair within studies, most of the studies had a higher specificity than sensitivity [8, 12, 15, 16, 18) . However, two studies each had higher sensitivity than specificity [9,17] or equal sensitivity and specificity [13, 14]."….

[Page: 11; Section: Results]

Authors should be aware that in this section they are expected to discuss the strengths and weaknesses of the review with regards to estimation of accuracy, and not the strengths and weaknesses of the evidence with regards to policy making decisions which would rely on other properties of the test, including its impact on patient outcomes and cost.

Thank you for the suggestion. Policymaking might depend on many non-research factors.

In the revised manuscript, the strengths and weaknesses have been rewritten as:

"…*However, some of the above findings should be interpreted in the context of the study limitations and strengths. There was substantial heterogeneity in the diagnostic accuracy parameters of the PAEDS, which was partly explained by the setting of the occurrence of EmD and the reference standard used. The role of*

each individual study in the summary DOR was further explored with a range of 111-152, adding strength to the method of this meta-analysis. The PAEDS threshold effect has to be further studied with larger meta-analysis. Expecting heterogeneity to start with, the use of random effects models, exploring the heterogeneity with meta-regression, subgroup analysis and the leave-one-out cross validation has strengthened the meta-analysis. Furthermore, not to compromise the diagnostic accuracy of PEADS in diagnosing EmD from other post-anaesthetic emergent problems like pain and agitation we excluded those studies with such conditions in this meta-analysis."....

[Page: 1; Section: Discussion]

---

(1) *Science editor:*

5 Issues raised: (1) The title is too long, and it should be no more than 18 words:

Many thanks for the suggestion.

We have reduced the length to 8 words as:

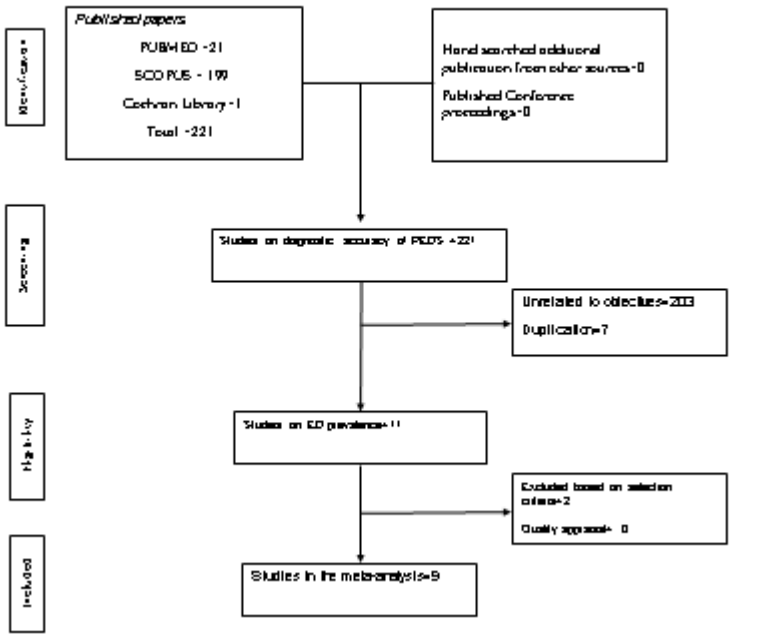
Pediatric Anesthesia Emergence Delirium Scale: A diagnostic meta-analysis.

---

(2) The authors did not provide original pictures. Please provide the original figure documents. Please prepare and arrange the figures using PowerPoint to ensure that all graphs or arrows or text portions can be reprocessed by the editor.

Thanks for your suggestion. All seven pictures have been prepared as suggested and uploaded.

Figure1: The PRISMA flowchart of studies included in the diagnostic meta-analysis for Pediatric Anesthesia Emergence Delirium Scale.



(3) ***Company editor-in-chief:*** I have reviewed the Peer-Review Report, full text of the manuscript, and the relevant ethics documents, all of which have met the basic publishing requirements of the World Journal of Clinical Pediatrics, and the manuscript is conditionally accepted.

Many thanks for the conditional acceptance.