PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	The accuracy of the neutrophil to lymphocyte ratio for the diagnosis of neonatal sepsis: a systematic review and meta-analysis
AUTHORS	xin, yu; Shao, Yunshuang; Mu, Wenjing; Li, Hongxu; Zhou, Yuxin; Wang, Changsong

VERSION 1 – REVIEW

REVIEWER	Martono Tri Utomo Universitas Airlangga- Dr. Soetomo Hospital, Department of Child Health
REVIEW RETURNED	07-Feb-2022

GENERAL COMMENTS	1. What is the reason for dividing the cut-off >2 and <2? is this a cut off for NLR? 2. Has there been a check on other cut-off divisions to check sensitivity, specificity, PLR, NLR and DOR? 3. How to distinguish between the abbreviations of NLR (neutrophil lymphocyte ratio) and NLR (negative likelihod ratio) in a sentence and the results?

REVIEWER	Aysegul Ozel Istanbul Universitesi
REVIEW RETURNED	07-Feb-2022

GENERAL COMMENTS	I congratulate the authors for their valuable meta-analysis.

REVIEWER	Deepak Sharma Fernandez Hospital
REVIEW RETURNED	14-Feb-2022

GENERAL COMMENTS	Thanks for giving me opportunity to review this systematic review and meta-analysis • Overall the study methodology is good. • Needs language editing. Authors need to review following issues 1. Page number 3- Neutrophil to lymphocyte ratio (NLR) is more accurate than blood culture (gold standard) in the diagnosis of neonatal sepsis????? – evidence to support this statement. 2. Page 3- This new laboratory index improves the diagnostic efficiency of neonatal sepsis, providing clinical evidence for the diagnosis of neonatal sepsis??? How laboratory index can provide

	clinical evidence???? 3. Page 4- Introduction Line no. 66-68; Due to the sensitivity of disease diagnosis methods and the timeliness and effectiveness of the whole treatment process, the mortality rate of neonatal sepsis is increasing year by year – Is Mortality rate increasing due to effectiveness of treatment process???? 4. Page no 7- line138-139: rephrase 5. Page number 8 - line no 146-150; not very clear that what authors want to convey. It is mentioned that for patient selection three references were considered high risk, please elaborate on this. 6. Your results are totally confusing. Page 9, line no. 166-170; you have mentioned pooled sensitivity and specificity, but at what ratio of neutrophil to lymphocyte (N:L) ratio these results hold true. You have not mentioned a specific value for N:L ratio. Same holds true for EOS subgroup analysis. The readers should be able to understand that if they want to apply N:L ratio in their setting to identify sepsis then what cut off for N:L ratio they should keep in their mind. 7. Line number 176 t0 183; Cutoff value>2, pooled sensitivity and specificity are, respectively 0.83(95 % CI 0.66-0.93) and 0.80(95 % CI 0.44-0.95), respectively; PLR is 4.1(95 % CI 1.0-17.2), NLR is 0.21(95 % CI 0.07-0.60), DOR is 20 (95 % CI 2-218), the area under the curve (AUC) is 0.88 (95 % CI 0.85-0.91). (4). Cutoff value <2, pooled sensitivity and specificity are, respectively 0.74(95 % CI 0.69-0.78) and 0.90(95 % CI 0.71-0.97); PLR is 7.1(95 % CI 2.3-21.8), NLR is 0.29(95 % CI 0.23-0.36), DOR is 25(95 % CI 7-88) The area under the curve (AUC) is 0.77(95 % CI 0.73-0.81). From these results what do you interpret?? Whether Cut off value to identify sepsis should be > 2 or less than 2 because from the results it seems that both cut offs can be used to identify sepsis which is not possible. <2 can be anything (0.5, 1,1.3 etc), similarly > 2 can be anything (5,10,20 etc), so is it N:L ratio of 1 (<2), or 5 (>2) which will suggest sepsis. From these results I am not clear that for identifying sepsis, what cut off value of N:L ratio should I consider to identify sepsis. I hope you are understanding what I want to convey. • Discussion – Needs more elaboration. Discuss on findings of included studies. Compare results among the included studies. Discuss diagnostic value of N:L ratio in comparison to other markers used to identify sepsis like CRP, Procalcitonin, micro ESR etc. • On ROC curve, you need to mention that at what N:L ratio you are getting best diagnostic accuracy. • You need to properly design table on characteristics of included studies as in current format it is not understandable.

REVIEWER	Ahmed Omran Suez Canal University, Pediatrics and Neonatology
REVIEW RETURNED	26-Feb-2022

GENERAL COMMENTS	Dear Editor,

	Thank you very much for giving me the opportunity to revise the manuscript “The accuracy of the neutrophil to lymphocyte ratio for the diagnosis of neonatal sepsis: a systematic review and meta-analysis” I believe that the topic of this systematic review and meta-analysis is very interesting and could add value to the written literature on neonatal sepsis diagnosis. It is a well-written paper and I recommend accepting it. Best regards, Ahmed Omran Assistant professor of Pediatrics Suez Canal University Egypt

REVIEWER	Deena Thomas Muthoot Healthcare, Neonatology
REVIEW RETURNED	12-Apr-2022

GENERAL COMMENTS	I really appreciate the meticulous work and meta-analysis done by the authors of this study. However few comments need to be addressed: 1. Study flow in Figure 1 shows 522 articles assessed for screen and 508 excluded, leaving 14 articles assessed for full text review. However the authors have mentioned 17 articles for full text review. Kindly check the numbers. 2. Pubmed screenshot search strategy shows nor (title/ abstract). Instead it should be nlr (title/ abstract). Kindly redo the search and provide the revised screenshot: it may change the number of articles screened but may or may not change the number of articles ultimately included in the study. 3. Table 2 should contain overall analysis and subgroup analysis. The subgroup analysis has been mentioned in text but no table has been provided. Kindly mention the entire detail in table and only main findings in text. The discussion should also reflect the interpretation of subgroup analysis especially NLR> 2 and NLR<2. The remaining two tables can be renamed as Table 3 and 4 respectively. 4. Kindly mention the range of cutoff of NLR used in studies included in both abstract and results. 5. The table on meta-regression shows year 'Yes' or 'No'. This doesn’t give a clear information on what year cutoff was used. Also give a brief description of the table on meta-regression. I understand that midas, reg command has been used to generate the table; however the graph so generated also shows besides non- Asian countries prospective studies are also source of heterogeneity in specificity. This is not reflected in the results. 6. Strengths and limitations need not be mentioned before the background of study. The paragraph there mentions NLR is more accurate than blood culture which is the gold standard, which needs to be corrected. 7. Please remove the numbering from data synthesis and subgroup analysis. The results need to be written in paragraph format in the main text and not point wise. 8. Kindly check the numberings of additional file. The text mentions search strategy screenshots as additional file 3, however we have received them as additional file 1. 9. The eligibility criteria mentions control as non- neonatal sepsis. Better term would have been 'no sepsis' or 'neonates without sepsis' to avoid confusion. 10. In identification of studies paragraph in result section it is better

	not to merge EONS, LONS and preterm infants in the same sentence. Kindly separate them into two sentences: one mentioning number of studies with EONS, LONS and both EONS and LONS while the other mentioning number of studies with only preterm, term or including all gestation. 11. Kindly check the version of metadisc software used. Some places mention 1.4 while some as 14. 12. Kindly check the references: reference number 10 and 13 wrongly written. Also reference number 13 Karabulut study has not been included in this meta-analysis. The study had been published in June 2021 (before the search date of Aug 2021 as per this study). Was there any particular reason for this omission? If not, inclusion of that study will change the entire results.

REVIEWER	Chun Sen Wu University of Southern Denmark, Research Unit of Gynecology and Obstetrics, Institute of Clinical Research
REVIEW RETURNED	26-Jun-2022

GENERAL COMMENTS	This systematic review is to evaluate the accuracy of the neutrophil to lymphocyte ratio for the diagnosis of neonatal sepsis. My review will primarily focus on the statistical analysis. Major concerns: Now it is well-known that I squared is not an absolute measure of heterogeneity between the studies. Statistics used in the meta-analysis should neither be based on I squared nor based on the heterogeneity. Rather it should be based on the hypothesis whether it is a common effect, fixed effects, or random effects with some distribution. However, the authors inappropriately used the I squared to justify the selection of the statistical analysis. sROC was presented in the manuscript but the authors neither described how they did nor interpreted sROC. Furthermore, the assumption of SROC is that the primary studies are random samples of one large common study, and that differences in results are random error, which may not always be justified. It also does not account for any systematic difference between the study population. Possible solution is hierarchical SROC, which accounts for variation both within and between studies. Deeks funnel plot asymmetric analysis was used to evaluate the publication bias. As the Cochrane recommendation mentioned “review authors should remember that, because the tests typically have relatively low power, even when a test does not provide evidence of funnel plot asymmetry, bias (including publication bias) cannot be excluded”. Therefore, it was over-interpreting the test as “The results of Deeks’s funnel plot asymmetry test showed that p=0.40 and p>0.05. This result indicated that the 13 articles included had no publication bias.” Perhaps could either present a contour-enhanced funnel plot or nonparametric trim-and fill analysis of publication bias. As we all know that no diagnostic test is perfect and almost all tests will sometimes miss disease (false negative) or indicate disease in normal patients (false positive). However, False negative and false

	positive diagnoses are rarely equally important. Missing a life-threatening disease will probably be regarded by a patient (and his or her doctor) as much more important than a false positive diagnosis in a healthy patient. Neonatal sepsis is considered as a life-threating event by many if not all. The pooled false negative rate (1 – sensitivity) is on average 23% and even can be lower as 29%, which is serious. It would lose more than 20% newborn babies suffering neonatal sepsis if the diagnosis were based on the ratio. therefore, the ratio alone is a poor measure. Even pooled sensitivity from the subgroup analysis was 0.83 (95 % CI 0.68-0.91), which means the false negative would be 17% even could be lower as 32%. Based on the results, can be conclude as ""… moderate diagnostic capacity with high sensitivity … for diagnosing neonatal sepsis"? Due to sensitivity and specificity are often negatively correlated within studies, bivariate analysis has been proposed to take the correlation into the consideration.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Dr. Martono Tri Utomo, Universitas Airlangga- Dr. Soetomo Hospital

Comments to the Author:

1. What is the reason for dividing the cut-off >2 and <2? is this a cut off for NLR?

Response: Thank you for your suggestion. First, I want to explain why we performed subgroup analysis with cut-offs >2 and <2. We take into account that in the 13 studies we originally included, the cut-off value of NLR varies from 0.1- 9.4 is not equal, and its span is very large, so we intend to prove that our final results are not affected by such a large span of cut-off value by subgroup analysis with the cut-off value, which leads to the instability of the results. In the articles, we included the cut-off value. There were 8 studies with a value less than 2 and 5 studies with a value greater than 2. We consider that the number of studies with cut-offs >2 and <2 is roughly equal, so we chose 2 as the cut-off value. We admit that this approach may be difficult for readers to understand.

Considering that our search time was outdated, we researched the literature until 22.6.28. After screening, we found a newly published study [1]. With reference to your second comment, we decided to divide the 14 articles into 3 segments according to the cut-off value, namely, 0-2, 2-4, and >4. We have added a new subgroup analysis table and added it to the table. The details are as follows. From the results of the subgroup analysis with the cut-off value, it can be found that even

though the NLR cut-off value of our included studies spanned a large range, it did not affect the stability of our results.

Finally, considering that the issue of cut-off values may confuse readers, we describe the range of cut-off values for our included literature in the Abstract and Results sections. In the methodology section, we explain the purpose of the subgroup analysis in detail.

2. Has there been a check on other cut-off divisions to check sensitivity, specificity, PLR, NLR and DOR?

Response: Thank you for your suggestion. As I mentioned in your first comment, we have added a new subgroup analysis table that looks more intuitive. We used cut-off values (0-2, 2-4, >4), type of sepsis, study area, etc., were subgroup analysed.(Additional file 5) The specific subgroup analysis table is as follows:

Table: Subgroup analysis of neutrophil-to-lymphocyte ratio in the diagnosis of neonatal sepsis. (Additional file 5)

Subgroup	Study number	Sen	Spe	LR^+	LR^-	DOR	AUC
All	14 [1-14]	0.74	0.88	6.35	0.30	21.27	0.87
Neonates							
EOS	6 [4,7,9-11,14]	0.75	0.99	63.30	0.26	247	0.97
LOS	4 [2,3,12,14]	0.60	0.85	3.71	0.41	11.14	0.85
Areas							
Asian	11 [2-9,12-14]	0.75	0.83	4.40	0.30	15	0.86
Non-Asian	3 [1,10,11]	0.67	0.90	18.64	0.38	45.94	0.95
Cut off							
0-2	8 [2-4,6,8,10-12]	0.74	0.90	7.1	0.29	25	0.77
2-4	3 [5,7,13]	0.79	0.62	2.21	0.33	6.73	0.85
>4	3 [1,9,14]	0.60	0.91	9.00	0.27	31.51	0.95

Note: SEN: sensitivity; SPE: specificity; LR^-: negative likelihood ratio; LR^+: positive likelihood ratio; DOR: diagnostic odds ratio; AUC: area under the curve;

3. How to distinguish between the abbreviations of NLR (neutrophil lymphocyte ratio) and NLR (negative likelihood ratio) in a sentence and the results?

Response: We appreciate this good suggestion, and we have made it according to your ideas. We fully agree with your suggestion, considering that it will confuse readers with the neutrophil-to-lymphocyte ratio and negative likelihood ratio. To solve this problem, we decided to abbreviate the neutrophil-to-lymphocyte ratio as NLR, the negative likelihood ratio as LR^-, and the positive likelihood ratio as LR^+, and we avoided the use of abbreviations in the abstract and conclusion sections.

We sincerely hope that this revised manuscript has addressed all your comments and suggestions. We earnestly appreciate your warm work and hope that the correction will meet with approval. Once again, thank you very much for your comments and suggestions.

Reviewer: 2

Dr. Aysegul Ozel, Istanbul Universitesi

Comments to the Author:

I congratulate the authors for their valuable meta-analysis.

Response: Thank you for your recognition of our paper.

Reviewer: 3

Dr. Deepak Sharma, Fernandez Hospital

Comments to the Author:

Thanks for giving me opportunity to review this systematic review and meta-analysis

- *Overall the study methodology is good.*
- *Needs language editing.*

Response: Thank you for your recognition of our article methodology, and at the same time, we found a professional language polishing company (American Journal Experts Company.

www.aje.com) to help us improve the grammar. The verification code (**9A04-E9A1-77DB-E20C-3E8E).** We are very sorry for bothering you with this kind of mistake that we should have avoided. Please refer to the revised version.

AJE

Editing Certificate

This document certifies that the manuscript

The accuracy of the neutrophil to lymphocyte ratio for the diagnosis of neonatal sepsis: a systematic review and meta-analysis

prepared by the authors

Yu,Xin

was edited for proper English language, grammar, punctuation, spelling, and overall style by one or more of the highly qualified native English speaking editors at AJE.

This certificate was issued on **July 21, 2022** and may be verified on the AJE website using the verification code **9A04-E9A1-77DB-E20C-3E8E** .

Neither the research content nor the authors' intentions were altered in any way during the editing process. Documents receiving this certification should be English-ready for publication; however, the author has the ability to accept or reject our suggestions and changes. To verify the final AJE edited version, please visit our verification page at aje.com/certificate. If you have any questions or concerns about this edited document, please contact AJE at support@aje.com.

AJE provides a range of editing, translation, and manuscript services for researchers and publishers around the world. For more information about our company, services, and partner discounts, please visit aje.com.

Authors need to review following issues

1. Page number 3- Neutrophil to lymphocyte ratio (NLR) is more accurate than blood culture (gold standard) in the diagnosis of neonatal sepsis????? – evidence to support this statement.

Response: Thank you for your suggestion. 'Neutrophil to lymphocyte ratio (NLR) is more accurate than blood culture (gold standard) in the diagnosis of neonatal sepsis' originates from the 'Strengths and Limitations section' of this article, We admit that there are serious problems with this statement and that this statement is too absolute. The diagnosis of neonatal sepsis is often accompanied by nonspecific clinical symptoms, high false-negative rates, and delays in obtaining blood culture results. The clinical manifestations of neonatal sepsis in this study were often nonspecific. The nonspecificity of clinical manifestations increases the difficulty in the diagnosis of neonatal sepsis, and it is easily missed and misdiagnosed. Blood culture is the gold standard for diagnosing sepsis, but it is time-

consuming, with a small blood sample volume, irregular operation during blood collection, and antibacterial treatment, all of which can lead to an increase in the false-negative rate and increase the difficulty of clinical diagnosis [2]. An ideal biomarker would need to have a high degree of accuracy in the early identification of the presence or absence of neonatal sepsis to guide the initiation and duration of antibiotic therapy. The neutrophil to lymphocyte ratio is a composite marker of absolute counts of neutrophils and lymphocytes in peripheral blood. An elevated neutrophil to lymphocyte ratio means a higher inflammatory load, which means a higher neutrophil count due to active inflammation, while low lymphocyte counts are associated with defective responses to inflammatory processes [3-4]. The NLR is an easy, fast, and inexpensive laboratory indicator that can be obtained early, combined with our results (sensitivity was 0.74 (95% CI 0.61-0.83), specificity was 0.88 (95% CI 0.73-0.95), LR+ was 6.35 (95% CI 2.6-15.47), LR- was 0.30 (95% CI 0.19-0.46), the diagnostic odds ratio was 21.27 (95% CI 6.98-64.84), and the area under the curve (AUC) was 0.87 (95% CI 0.84-0.89).) It can fully demonstrate that NLR, a laboratory indicator, can improve the accuracy of diagnosing neonatal sepsis.

According to editor's comments:‘Please revise the‘Strengths and limitations of this study’ section of your manuscript (after the abstract). This section should contain up to five short bullet points, no longer than one sentence each, that relate specifically to the methods. The novelty, aims, the results or expected impact of the study should not be summarized here.’Eventually, we removed the sentence.

2. Page 3- This new laboratory index improves the diagnostic efficiency of neonatal sepsis, providing clinical evidence for the diagnosis of neonatal sepsis??? How laboratory index can provide clinical evidence????

Response: Thank you for your suggestion. As stated in your first comment response, we removed this sentence based on the editor's comments and made a comprehensive change to the Strengths and Limitations section. The specific modifications are as follows:

Before revision：

(1). As a cheap and readily available new comprehensive inflammatory indicator, Neutrophilthe neutrophil to lymphocyte ratio (NLR) is relatively stable and unaffected by in vitro blood sample processing and conventional physiological conditions.

(2). NeutrophilThe neutrophil to lymphocyte ratio (NLR) is more accurate than blood culture (gold standard) in the diagnosis of neonatal sepsis. This new laboratory index improves the diagnostic efficiency of neonatal sepsis, providing clinical evidence for the diagnosis of neonatal sepsis.

(3). Due to the limited number of articles, we cannot accurately distinguish the accuracy of the ratio of neutrophils to lymphocytes in early-onset neonatal sepsis and late-onset sepsis.

After revision：

- We conducted a comprehensive search of each literature database and formulated detailed inclusion and ranking criteria to ensure the quantity and quality of the included literature.
- Subgroup analyses were performed according to sepsis type, study area and cut-off value as described in the methodology section of this study.
- Our included articles lack more multicenter and large sample studies.
- There may be other clinical and statistical heterogeneity in the included studies.

3. Page 4- Introduction

Line no. 66-68; Due to the sensitivity of disease diagnosis methods and the timeliness and effectiveness of the whole treatment process, the mortality rate of neonatal sepsis is increasing year by year – Is Mortality rate increasing due to effectiveness of treatment process????

Response: We appreciate this good suggestion, and we have done it according to your ideas. We acknowledge that there are some issues with this statement, so we have changed it.

Before revision：

Due to the sensitivity of disease diagnosis methods and the timeliness and effectiveness of the whole treatment process, the mortality rate of neonatal sepsis is increasing year by year.

After revision：

At present, neonatal sepsis is faced with insufficient diagnostic methods, resulting in the inability to guide clinical treatment in a timely manner, thereby affecting its therapeutic effect.

4. Page no 7- line138-139: rephrase

Response: Thank you for your suggestion. Considering that this might confuse the reader, we have adjusted the paragraphs here to make it easier to read and understand.

Before revision：

After checking duplicates and reading abstracts and excluding relevant literature according to the exclusion criteria, 14 studies were finally included. The specific process is shown in Fig 1. The references were included from 2017 to 2022, with 1499 newborns, including 783 in the study group and 716 in the control group. Among them, 3 had late-onset sepsis, 5 had early-onset sepsis, and 2 were preterm infants. Eleven studies were from Asia, and 3 studies were from non-Asian countries. Basic information of the included literature is shown in Table 1.

After revision：

After checking duplicates and reading abstracts and excluding relevant literature according to the exclusion criteria, a final total of 14 studies were used for the current meta-analysis. The specific process is shown in Fig 1. Of these, 783 neonates in the sepsis group and 716 neonates in the nonsepsis group were studied and evaluated. Additional file 2 shows the major characteristics of the selected studies. The baseline information included the following parameters: the number of patients, gestational age, regions, types of sepsis, disease diagnosis methods, study design and the cut-off value of NLR.

5. Page number 8 - line no 146-150; not very clear that what authors want to convey. It is mentioned that for patient selection three references were considered high risk, please elaborate on this.

Response: We appreciate this good suggestion, and we have made it according to your ideas. We acknowledge that it was not expressed in sufficient detail here. For patient selection, three references were considered high-risk. According to the content in the QUADAS-2 checklist. Senem Alkan Ozdemir et al did not avoid unreasonable case exclusion, and RH Ruslie et al and Khadijah Rizky Sumitro et al did not avoid case–control study designs. We therefore assessed these three articles as

high risk in terms of patient selection.

6. Your results are totally confusing. Page 9, line no. 166-170; you have mentioned pooled sensitivity and specificity, but at what ratio of neutrophil to lymphocyte (N: L) ratio these results hold true. You have not mentioned a specific value for N:L ratio. Same holds true for EOS subgroup analysis. The readers should be able to understand that if they want to apply N:L ratio in their setting to identify sepsis then what cut off for N:L ratio they should keep in their mind.

Response: Thank you for your suggestion. I understand what you are trying to say. For a certain diagnostic test, multiple researchers may have conducted research, but because these studies have different degrees of random sampling error and the cut-off values of diagnostic indicators used in each of them are often different, the obtained diagnostic test accuracy indicators such as sensitivity. The degree and specificity also vary. To comprehensively analyse the results of different studies and obtain comprehensive conclusions, a systematic meta-analysis method is needed to evaluate the diagnostic tests. A systematic review/meta-analysis of diagnostic tests is a research method that comprehensively evaluates the accuracy and importance of diagnostic tests by qualitative description or quantitative analysis with synthetic receiver operating characteristic curves. Its purpose is to evaluate the accuracy of a diagnostic test for diagnosing the target disease and is the highest level of evidence in a diagnostic test. It mainly includes two aspects: 1. The technical quality evaluation of diagnostic tests, mainly from the aspects of research design, method precision, accuracy, repeatability, sensitivity, specificity, etc.; 2. The accuracy evaluation of diagnostic tests, mainly meta-analysis, was used to evaluate the sensitivity and specificity of the target disease and to report the likelihood ratio and diagnostic odds ratio.

Perhaps because the profile of subjects, the inclusion or exclusion criterion, and the criteria for threshold establishment were heterogeneous, none of the eligible studies used a unified threshold. What we meant to say was that systematic reviews/meta-analyses of diagnostic tests did not ultimately yield optimal cut-off values [5,6]. Of course, we also agree with your idea, and we have tried to find a solution. The only way is to get in touch with the authors of each included article, hoping to get their raw data, but unfortunately, this may be difficult to achieve.

We listed each cut-off value used for NLR in each study in the included study basic information table (Additional file 2), and in the results description section, we specified the range of cut-off values used

in the included studies.

7. Line number 176 t0 183; Cutoff value>2, pooled sensitivity and specificity are, respectively 0.83(95 % CI 0.66-0.93) and 0.80(95 % CI 0.44-0.95), respectively; PLR is 4.1(95 % CI 1.0-17.2), NLR is 0.21(95 % CI 0.07-0.60), DOR is 20 (95 % CI 2-218), the area under the curve (AUC) is 0.88 (95 % CI 0.85-0.91).

(4). Cutoff value <2, pooled sensitivity and specificity are, respectively 0.74(95 % CI 0.69-0.78) and 0.90(95 % CI 0.71-0.97); PLR is 7.1(95 % CI 2.3-21.8), NLR is 0.29(95 % CI 0.23-0.36), DOR is 25(95 % CI 7-88) The area under the curve (AUC) is 0.77(95 % CI 0.73-0.81). From these results what do you interpret?? Whether Cut off value to identify sepsis should be > 2 or less than 2 because from the results it seems that both cut offs can be used to identify sepsis which is not possible. <2 can be anything (0.5, 1,1.3 etc), similarly > 2 can be anything (5,10,20 etc), so is it N:L ratio of 1 (<2), or 5 (>2) which will suggest sepsis.

From these results I am not clear that for identifying sepsis, what cut off value of N:L ratio should I consider to identify sepsis. I hope you are understanding what I want to convey.

Response: Thank you for your suggestion. I understand what you want to express, and in your review comments, I can see that your main concern is the NLR cut-off value in our NLR diagnosis of neonatal sepsis. Please refer to your previous comment reply for the reply to the question related to the cut-off value.

Additionally, I would like to explain to you why we do subgroup analysis with cut-off >2 and <2, we take into account that in the 13 studies we initially included, the NLR cut-off values ranged from 0.1-9.4, Its span is very large, so we intend to prove that our final results are not affected by such a large span of cut-off value by subgroup analysis with a cut-off value, which in our included articles, the cut-off value is less than 2. 8 articles and 5 articles greater than 2. We consider that the number of articles with cut-offs >2 and <2 is roughly equal, so we chose 2 as the cut-off value. We admit that this approach may make it difficult for readers to understand. Considering our search, the time was expired, and we researched the literature until 22.6.28. After screening, we found newly published studies that could be included [2]. After our discussion, we decided to divide the cut-off value into 3 sections, which are 0-2, 2-4, and > 4. We have added a new subgroup analysis table. (Additional file 5) The details are as follows. The results of the subgroup analysis with the cut-off value. It shows no

threshold effect heterogeneity. By using metadisc1.4, we found that the Spearman correlation coefficient was -0.037 (p= 0.899) (p>0.05). Even the NLR cut-off value of the articles we included was very wide, but it did not affect the stability of our results.

Finally, considering that the issue of cut-off values may confuse readers, we describe the range of cut-off values for our included literature in the Abstract and Results sections. In the methodology section, we explain the purpose of the subgroup analysis in detail.

Table: Subgroup analysis of neutrophil-to-lymphocyte ratio in the diagnosis of neonatal sepsis. (Additional file 5)

Subgroup	Study number	Sen	Spe	LR^+	LR^-	DOR	AUC
All	14 [1-14]	0.74	0.88	6.35	0.30	21.27	0.87
Neonates							
EOS	6 [4,7,9-11,14]	0.75	0.99	63.30	0.26	247	0.97
LOS	4 [2,3,12,14]	0.60	0.85	3.71	0.41	11.14	0.85
Areas							
Asian	11 [2-9,12-14]	0.75	0.83	4.40	0.30	15	0.86
Non-Asian	3 [1,10,11]	0.67	0.90	18.64	0.38	45.94	0.95
Cut off							
0-2	8 [2-4,6,8,10-12]	0.74	0.90	7.1	0.29	25	0.77
2-4	3 [5,7,13]	0.79	0.62	2.21	0.33	6.73	0.85
>4	3 [1,9,14]	0.60	0.91	9.00	0.27	31.51	0.95

Note: SEN: sensitivity; SPE: specificity; LR^-: negative likelihood ratio; LR^+: positive likelihood ratio; DOR: diagnostic odds ratio; AUC: area under the curve;

• *Discussion – Needs more elaboration. Discuss on findings of included studies. Compare results among the included studies. Discuss diagnostic value of N:L ratio in comparison to other markers used to identify sepsis like CRP, Procalcitonin, micro ESR etc.*

Response: Thank you for your suggestion. According to your request, we have supplemented the discussion of the included articles. We discuss the diagnostic value of NLR compared with CRP,

procalcitonin, microerythrocyte sedimentation rate and other markers for differentiating sepsis. Corresponding modifications have been shown in the manuscript.

• *On ROC curve, you need to mention that at what N:L ratio you are getting best diagnostic accuracy.*

Response:

Thank you for your suggestion. As I replied in your sixth and seventh comments, we cannot obtain the best cut-off value for NLR on the ROC curve in the diagnostic meta-analysis. In the future, we hope to be able to contact the author of each article and hope to obtain comprehensive raw data to obtain the best cut-off value through statistical analysis to better guide clinical application.

• *You need to properly design table on characteristics of included studies as in current format it is not understandable.*

Response: Thank you for your suggestion. We have adjusted the table according to your request to make it easier to understand. For details, see Additional file 2.

We did our best to answer your reviewer comments and explained to you in detail the cut-off you care about in the diagnostic meta-analysis. We sincerely hope that this revised manuscript has addressed all your comments and suggestions. We earnestly appreciate your warm work and hope that the correction will meet with approval. Once again, thank you very much for your comments and suggestions. Meanwhile, we would be happy to respond to any further questions and comments that you may have.

Reviewer: 4

Dr. Ahmed Omran, Suez Canal University

Comments to the Author:

Dear Editor,

Thank you very much for giving me the opportunity to revise the manuscript "The accuracy of the neutrophil to lymphocyte ratio for the diagnosis of neonatal sepsis: a systematic review and meta-

analysis"

I believe that the topic of this systematic review and meta-analysis is very interesting and could add value to the written literature on neonatal sepsis diagnosis.

It is a well-written paper and I recommend accepting it.

Best regards,

Ahmed Omran

Assistant professor of Pediatrics

Suez Canal University

Egypt

Response: Thank you for your recognition of our paper.

Reviewer: 5

Dr. Deena Thomas, Muthoot Healthcare, Muthoot Healthcare

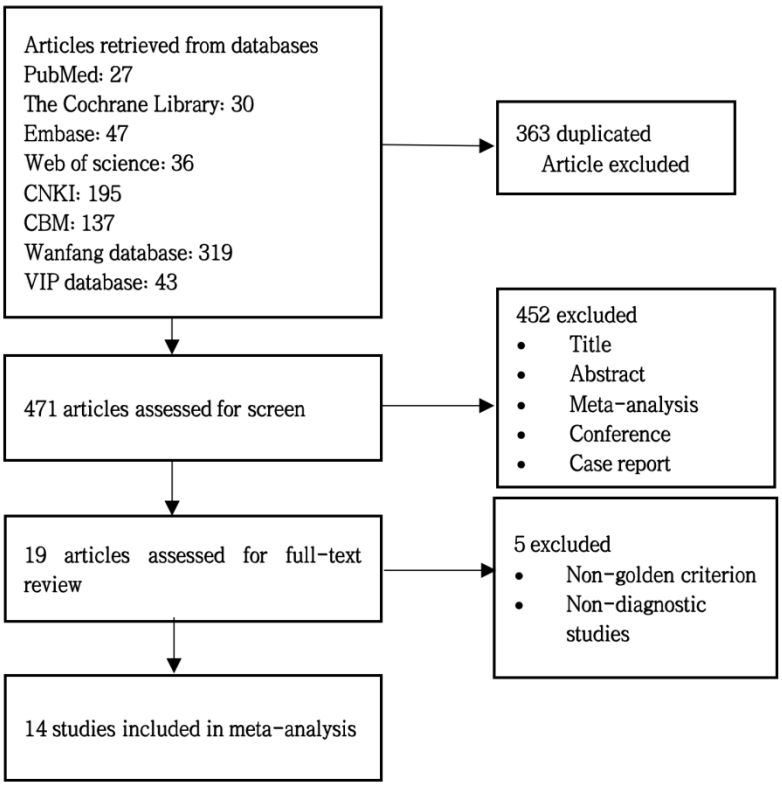
Comments to the Author:

I really appreciate the meticulous work and meta-analysis done by the authors of this study. However few comments need to be addressed:

1. Study flow in Figure 1 shows 522 articles assessed for screen and 508 excluded, leaving 14 articles assessed for full text review. However, the authors have mentioned 17 articles for full text review. Kindly check the numbers.

Response:

Thank you for your suggestion. We have carefully checked our literature screening flow chart and found this problem due to our negligence. We have researched, and the search time has been updated to 22.6.28. The search results showed that one article met the inclusion criteria [1], and we updated the flowchart.



2. PubMed screenshot search strategy shows nor (title/ abstract). Instead, it should be nlr (title/ abstract). Kindly redo the search and provide the revised screenshot: it may change the number of articles screened but may or may not change the number of articles ultimately included in the study.

Response:

Thank you for your suggestion. We have carefully checked the retrieval formula of PubMed. There is indeed the problem you mentioned. We reretrieved 'NLR', and the retrieval time was updated to 22.6.28. The retrieval results added a new study that met the inclusion criteria. [1], and in the editor's opinion: '- Please include, as a supplementary file, the precise, full search strategy (or strategies) for all databases, registers, and websites, including any filters and limits used. Please include these are lists, rather than screenshots.' We have updated Additional file 1 (Search strategy), and we have replaced the previous screenshots by making a table at the request of the editor. The table includes the details of the literature search (database name, website, search results, detailed retrieval formula, etc.), we reextracted the data and updated the screening flow chart, quality evaluation, statistical analysis and other steps. The newly included articles did not have an obvious impact on the original results, and we revised them accordingly. Both have been presented in the manuscript.

3. Table 2 should contain overall analysis and subgroup analysis. The subgroup analysis has been

mentioned in text, but no table has been provided. Kindly mention the entire detail in table and only main findings in text. The discussion should also reflect the interpretation of subgroup analysis especially NLR> 2 and NLR<2. The remaining two tables can be renamed as Table 3 and 4 respectively.

Response:

Thank you for your suggestion. We have added a new subgroup analysis table at your request, but we have named it Table 4 (Additional File 5) considering the order in which it appears in the text. Considering that we have added a new included article, we performed a subgroup analysis with the cut-off value divided into three segments, 0-2, 2-4, and >4. We explained this in the Discussion, and the subgroup analysis table is as follows:

Table: Subgroup analysis of neutrophil-to-lymphocyte ratio in the diagnosis of neonatal sepsis.(Additional file 5)

Subgroup	Study number	Sen	Spe	LR^+	LR^-	DOR	AUC
All	14 [1-14]	0.74	0.88	6.35	0.30	21.27	0.87
Neonates							
EOS	6 [4,7,9-11,14]	0.75	0.99	63.30	0.26	247	0.97
LOS	4 [2,3,12,14]	0.60	0.85	3.71	0.41	11.14	0.85
Areas							
Asian	11 [2-9,12-14]	0.75	0.83	4.40	0.30	15	0.86
Non-Asian	3 [1,10,11]	0.67	0.90	18.64	0.38	45.94	0.95
Cut off							
0-2	8 [2-4,6,8,10-12]	0.74	0.90	7.1	0.29	25	0.77
2-4	3 [5,7,13]	0.79	0.62	2.21	0.33	6.73	0.85
>4	3 [1,9,14]	0.60	0.91	9.00	0.27	31.51	0.95

Note: SEN: sensitivity; SPE: specificity; LR^-: negative likelihood ratio; LR^+: positive likelihood ratio; DOR: diagnostic odds ratio; AUC: area under the curve;

4. Kindly mention the range of cutoff of NLR used in studies included in both abstract and results.

Response:

Thank you for your suggestion. Our NLR cut-off value range for the included studies was 0.1-9.4. At your request, we have supplemented the range of cut-off values used in the included studies in the Abstract and Results sections.

5. The table on meta-regression shows year 'Yes' or 'No'. This doesn't give a clear information on what year cutoff was used. Also give a brief description of the table on meta-regression. I understand that midas, reg command has been used to generate the table; however, the graph so generated also shows besides non Asian countries prospective studies are also source of heterogeneity in specificity. This is not reflected in the results.

Response:

Thank you for your suggestion. We are in the meta-regression analysis (2019 is the cut-off year, '≥2019' is yes, and '＜2019' is no), and we marked the cut-off year in the year section of the meta-regression analysis table. Our meta-regression results were changed because we reperformed our literature search and included a newly published study,[1] which indicated prospective studies as a source of specific heterogeneity (p=0.01). Based on your suggestion, we have briefly described the meta-regression table in the Results section.

Sensitivity and Specificity

Parameter	category	nstudies	Sensitivity	p1	Specificity	p2
Asia	Yes	11	0.75 [0.63 - 0.88]	0.92	0.84 [0.70 - 0.98]	0.28
	No	3	0.67 [0.40 - 0.95]	.	0.98 [0.92 - 1.00]	.
Year	Yes	10	0.69 [0.55 - 0.83]	0.08	0.87 [0.74 - 1.00]	0.87
	No	4	0.83 [0.68 - 0.98]	.	0.91 [0.76 - 1.00]	.
Preterm	Yes	2	0.71 [0.40 - 1.00]	0.73	0.81 [0.44 - 1.00]	0.91
	No	12	0.74 [0.62 - 0.86]	.	0.89 [0.79 - 1.00]	.
Prospectivestudy	Yes	3	0.84 [0.68 - 1.00]	0.62	0.98 [0.92 - 1.00]	0.01
	No	11	0.70 [0.57 - 0.83]	.	0.83 [0.70 - 0.97]	.

Joint Model

Parameter	category	LRTChi2	Pvalue	I2	I2lo	I2hi
Asia	Yes	2.74	0.25	27	0	100
	No	.	.	.	.	.
Year	Yes	1.82	0.40	0	0	100
	No	.	.	.	.	.
Preterm	Yes	0.31	0.86	0	0	100
	No	.	.	.	.	.
Prospectivestudy	Yes	5.28	0.07	62	15	100
	No	.	.	.	.	.

6. Strengths and limitations need not be mentioned before the background of study. The paragraph there mentions NLR is more accurate than blood culture, which is the gold standard, which needs to be corrected.

Response:

Thank you for your suggestion. 'Neutrophil to lymphocyte ratio (NLR) is more accurate than blood culture (gold standard) in the diagnosis of neonatal sepsis' originates from the 'Strengths and Limitations section' of this article, We admit that there are serious problems with this statement and that this statement is too absolute.

The diagnosis of neonatal sepsis is often accompanied by nonspecific clinical symptoms, high false negative rates, and delays in obtaining blood culture results. The clinical manifestations of neonatal sepsis in this study were often nonspecific. The nonspecificity of clinical manifestations increases the difficulty in the diagnosis of neonatal sepsis, and it is easily missed and misdiagnosed. Blood culture is the gold standard for diagnosing sepsis, but it is time-consuming, with a small blood sample volume, irregular operation during blood collection, and antibacterial treatment, all of which can lead

to an increase in the false-negative rate and increase the difficulty of clinical diagnosis [2]. An ideal biomarker would need to have a high degree of accuracy in the early identification of the presence or absence of neonatal sepsis to guide the initiation and duration of antibiotic therapy. The neutrophil to lymphocyte ratio is a composite marker of absolute counts of neutrophils and lymphocytes in peripheral blood. An elevated neutrophil to lymphocyte ratio means a higher inflammatory load, which means a higher neutrophil count due to active inflammation, while low lymphocyte counts are associated with defective responses to inflammatory processes [3-4]. The NLR is an easy, fast, and inexpensive laboratory indicator that can be obtained early, combined with our results (sensitivity was 0.74 (95% CI 0.61-0.83), specificity was 0.88 (95% CI 0.73-0.95), positive likelihood ratio was 6.35 (95% CI 2.6-15.47), negative likelihood ratio was 0.30 (95% CI 0.19-0.46), diagnostic odds ratio was 21.27 (95% CI 6.98-64.84), and area under the curve (AUC) was 0.87 (95% CI 0.84-0.89)). It can fully demonstrate that NLR, a laboratory indicator, can improve the accuracy of diagnosing neonatal sepsis.

According to editor's comments:‘Please revise the ‘Strengths and limitations of this study’ section of your manuscript (after the abstract). This section should contain up to five short bullet points, no longer than one sentence each, that relate specifically to the methods. The novelty, aims, the results or expected impact of the study should not be summarized here.’Eventually, we deleted the sentence.

7. Please remove the numbering from data synthesis and subgroup analysis. The results need to be written in paragraph format in the main text and not point wise.

Response:

Thank you for your suggestion. We removed the numbering from the data synthesis and subgroup analysis as you requested and expressed our results in the main text in paragraph form.

8. Kindly check the numberings of additional file. The text mentions search strategy screenshots as additional file 3,however we have received them as additional file 1.

Response:

Thank you for your suggestion. We have changed the search strategy's Additional file 3 to Additional file 1.

9. The eligibility criteria mention control as non- neonatal sepsis. Better term would have been 'no sepsis' or 'neonates without sepsis' to avoid confusion.

Response:

Thank you for your suggestion. To avoid confusion, we changed ‘nonneonatal sepsis’ to ‘neonates without sepsis’ according to your suggestion.

10. In identification of studies paragraph in result section it is better not to merge EONS, LONS and preterm infants in the same sentence. Kindly separate them into two sentences: one mentioning number of studies with EONS, LONS and both EONS and LONS while the other mentioning number of studies with only preterm, term or including all gestation.

Response:

Thank you for your suggestion. Considering that this might confuse the reader, we have adjusted the paragraphs here to make it easier to read and understand.

Before revision：

After checking duplicates and reading abstracts and excluding relevant literature according to the exclusion criteria, 14 studies were finally included. The specific process is shown in Fig 1. The references were included from 2017 to 2022, with 1499 newborns, including 783 in the study group and 716 in the control group. Among them, 3 had late-onset sepsis, 5 had early-onset sepsis, and 2 were preterm infants. Eleven studies were from Asia, and 3 studies were from non-Asian countries. Basic information of the included literature is shown in Table 1.

After revision：

After checking duplicates and reading abstracts and excluding relevant literature according to the exclusion criteria, a final total of 14 studies were used for the current meta-analysis. The specific process is shown in Fig 1. Of these, 783 neonates in the sepsis group and 716 neonates in the nonsepsis group were studied and evaluated. Additional file 2 shows the major characteristics of the selected studies. The baseline information included the following parameters: the number of patients, gestational age, regions, types of sepsis, disease diagnosis methods, study design and the cut-off

value of NLR.

11.Kindly check the version of metadisc software used. Some places mention 1.4 while some as 14.

Response:

Thank you for your suggestion. We checked the metadisc software version in the manuscript and found that one place is metadisc 14, which has been changed to metadisc 1.4.

12. Kindly check the references: reference number 10 and 13 wrongly written. Also reference number 13 Karabulut study has not been included in this meta-analysis. The study had been published in June 2021 (before the search date of Aug 2021 as per this study). Was there any particular reason for this omission? If not, inclusion of that study will change the entire results.

Response:

Thank you for your suggestion. We have checked our references carefully. Due to our negligence, there may have been citation errors in the endnote export process. After our careful check, we found that the reference number 13 Karabulut study did not meet the inclusion criteria of our article. Because the diagnosis of neonatal sepsis in this article was based on clinical manifestations rather than blood cultures (the gold standard), we did not include this article to ensure the quality of the article. Furthermore, we have checked all references and completed revisions.

Before revision：

reference number [10]: Sumitro K R, Utomo M T, Widodo A D W. Neutrophil-to-Lymphocyte Ratio as an Alternative Marker of Neonatal Sepsis in Developing Countries[J]. *Oman Medical Journal*, 2021, 36(1): e214-e214.

reference number [13]: Karabulut B, Alatas S O. Diagnostic Value of Neutrophil to Lymphocyte Ratio and Mean Platelet Volume on Early Onset Neonatal Sepsis on Term Neonate[J]. *Journal of Pediatric Intensive Care*, 2021, 10(02): 143-147.

After revision：

reference number [10]: Alkan Ozdemir S, Arun Ozer E, Ilhan O, et al. Can neutrophil to lymphocyte ratio predict late-onset sepsis in preterm infants? [J]. *Journal of Clinical Laboratory Analysis*, 2017: e22338.

reference number[13]: Sumitro K R, Utomo M T, Widodo A. Neutrophil-to-Lymphocyte Ratio as an Alternative Marker of Neonatal Sepsis in Developing Countries[J]. *Oman Medical Journal*, 2021, 36(1): e214-e214.

We sincerely hope that this revised manuscript has addressed all your comments and suggestions. We earnestly appreciate your warm work and hope that the correction will meet with approval. Once again, thank you very much for your comments and suggestions.

Reviewer: 6

Dr. Chun Sen Wu, University of Southern Denmark, Odense University Hospital

Comments to the Author:

This systematic review is to evaluate the accuracy of the neutrophil to lymphocyte ratio for the diagnosis of neonatal sepsis.

My review will primarily focus on the statistical analysis.

Major concerns:

Now it is well-known that I squared is not an absolute measure of heterogeneity between the studies. Statistics used in the meta-analysis should neither be based on I squared nor based on the heterogeneity. Rather it should be based on the hypothesis whether it is a common effect, fixed effects, or random effects with some distribution. However, the authors inappropriately used the I squared to justify the selection of the statistical analysis.

Response: Thank you for your valuable comment. We admit that we have not considered this aspect

carefully. We have made corresponding revisions in the statistical methods section of the manuscript. The specific revisions are as follows:

Before revision：

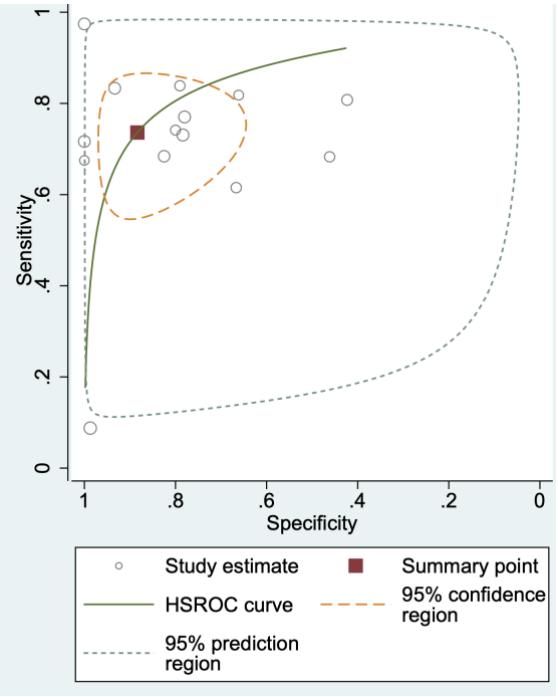
The I^2 test evaluated study heterogeneity. I^2>50% indicated that the heterogeneity generated in the study would have a specific impact. Meta Disc1.4 software was used to analyse the threshold effect heterogeneity. If the effect sizes of the studies were homogeneous, a fixed-effects model was used; if they were heterogeneous, a random-effects model was used; if there was heterogeneity between the studies, the source of the heterogeneity was further explored, and threshold effect and nonthreshold effect analyses were carried out.

After revision：

The heterogeneity of the included studies was evaluated by the Cochrane Q test and I^2 statistic. I^2 could be calculated from the formula $I^2=100\%\times(Q–df)/Q$. If I^2 is<50% or the *p* value is>0.1, a fixed effects model is used for pooling the data, whereas if I^2 is>50% or the *p* value is<0.1, then there is more heterogeneity among studies, and a bivariate random effects model is used for pooling the data; if I^2 is<50% or the *p* value is<0.1, a fixed effects model could be used; if I^2 is>50% or the *p* values>0.1, a bivariate random effects model could be used. If there was heterogeneity between the studies, the source of the heterogeneity was further explored, threshold effect and nonthreshold effect analyses were carried out Meta Disc1.4 software was used to analyse the threshold effect heterogeneity.

sROC was presented in the manuscript but the authors neither described how they did nor interpreted sROC. Furthermore, the assumption of SROC is that the primary studies are random samples of one large common study, and that differences in results are random error, which may not always be justified. It also does not account for any systematic difference between the study population. Possible solution is hierarchical SROC, which accounts for variation both within and between studies.

Response: Thank you for your suggestion. At your request, we have explained SROC in detail in the manuscript. We also performed HSROC based on your suggestion. For details, please refer to the image below.



False negatives: **fn** True negatives: **tn**

Refining starting values:

Iteration 0: log likelihood = **-98.339951**
Iteration 1: log likelihood = **-98.335698**
Iteration 2: log likelihood = **-98.335631**
Iteration 3: log likelihood = **-98.335637**

Performing gradient-based optimization:

Iteration 0: log likelihood = **-98.335637**
Iteration 1: log likelihood = **-98.335438**
Iteration 2: log likelihood = **-98.335438**

Meta-analysis of diagnostic accuracy

Log likelihood = **-98.335438** Number of studies = **1**

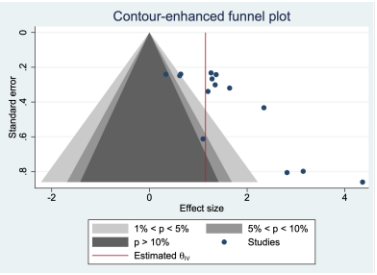
	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Bivariate						
E(logitSe)	**1.0262**	**.3019493**			**.4343898**	**1.61800**
E(logitSp)	**2.028404**	**.5136037**			**1.021759**	**3.03504**
Var(logitSe)	**1.141125**	**.5122329**			**.4734155**	**2.75057**
Var(logitSp)	**3.021848**	**1.594691**			**1.074178**	**8.50098**
Corr(logits)	**-.1289933**	**.3228453**			**-.6487794**	**.472866**
HSROC						
Lambda	**2.899164**	**.521423**			**1.877193**	**3.92113**
Theta	**-.1405004**	**.3559702**			**-.8381891**	**.557188**
beta	**.4869269**	**.344331**	**1.41**	**0.157**	**-.1879494**	**1.16180**
s2alpha	**3.234852**	**1.643139**			**1.19533**	**8.75429**
s2theta	**1.048249**	**.4732491**			**.4326884**	**2.5395**
Summary pt.						
Se	**.7361784**	**.0586445**			**.6069214**	**.834520**
Sp	**.8837472**	**.0527667**			**.7353151**	**.954132**
DOR	**21.21277**	**12.01602**			**6.989283**	**64.3816**
LR+	**6.332565**	**2.863118**			**2.610533**	**15.3613**
LR-	**.2985261**	**.0667921**			**.1925462**	**.462838**
1/LR-	**3.349791**	**.7494805**			**2.16058**	**5.1935**

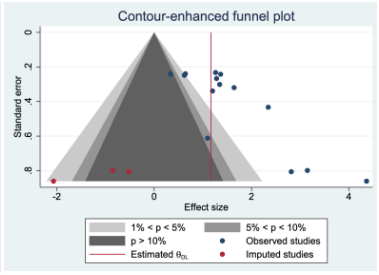
Covariance between estimates of E(logitSe) & E(logitSp) **-.0170471**

DeeksDeeks' funnel plot asymmetric analysis was used to evaluate the publication bias. As the Cochrane recommendation mentioned, "review authors should remember that, because the tests typically have relatively low power, even when a test does not provide evidence of funnel plot asymmetry, bias (including publication bias) cannot be excluded". Therefore, it was over-interpretingoverinterpreting the test as follows: "The results of Deeks'sDeeks' funnel plot asymmetry test showed that p=0.40 and p>0.05. This result indicated that the 13 articles included had no publication bias." Perhaps could either present a contour-enhanced funnel plot or nonparametric trim-and fill analysis of publication bias.

Response: Thank you for your suggestion. Considering that the search time is outdated, we researched the literature. The search time was updated to 2022.06.28. The search results added new literature. [1]

Taking into account your suggestion, we carried out the contour-enhanced funnel plot, and indeed we found bias in the results, so we performed a clip-and-fill method, and the results showed that publication bias did not have a large impact on the stability of our results. Corresponding modifications are presented in the manuscript. For details, please refer to the following figure:






```
Effect-size label: Effect size
      Effect size: d
        Std. err.: sed

Nonparametric trim-and-fill analysis of publication bias
Run estimator, imputing on the left

Iteration                      Number of studies =    17
  Model: Random-effects          observed =           14
Method: DerSimonian-Laird        imputed =             3

Pooling
  Model: Random-effects
Method: DerSimonian-Laird
```

Studies	Effect size	[95% conf.	interval]
Observed	1.402	1.032	1.772
Observed + Imputed	1.167	0.766	1.567

As we all know that no diagnostic test is perfect, and almost all tests will sometimes miss disease (false negative) or indicate disease in normal patients (false positive). However, False negative and false positive diagnoses are rarely equally important. Missing a life-threatening disease will probably be regarded by a patient (and his or her doctor) as much more important than a false positive diagnosis in a healthy patient.

Neonatal sepsis is considered as a life-threating event by many if not all. The pooled false negative rate (1 – sensitivity) is on average 23% and even can be lower as 29%, which is serious. It would lose more than 20% newborn babies suffering neonatal sepsis if the diagnosis were based on the ratio. therefore, the ratio alone is a poor measure. Even pooled sensitivity from the subgroup analysis was 0.83 (95 % CI 0.68-0.91), which means the false negative would be 17% even could be lower as 32%. Based on the results, can be conclude as ""… moderate diagnostic capacity with high sensitivity … for diagnosing neonatal sepsis"?

Response: Thank you for your suggestion. We acknowledge that we did not consider the risk of underdiagnosis rates for neonates in our Conclusions. In our results, the sensitivity and specificity of NLR in diagnosing neonatal sepsis were 0.74 and 0.88, respectively. The sensitivity and specificity of IL-8 in the diagnosis of neonatal sepsis were found to be 0.78 and 0.84 in the study by Min Zhou et al. [8]. The sensitivity and specificity of neutrophil CD64 in the diagnosis of neonatal sepsis were found in the study by Ji, Dai et al. to be 0.80 and 0.83, respectively. [9] Additionally, in the study of Lv, Bokun et al., it was found that the sensitivity and specificity of TNF-α in the diagnosis of early-onset sepsis and late-onset sepsis were sensitivity = 0.66, specificity = 0.76 and sensitivity = 0.68, specificity = 0.89, respectively. [10] Compared with IL-8, CD-64 and TNF-α, NLR has more diagnostic advantages, but compared with other indicators, such as PCT [11], presepsin[12] and other indicators, the diagnostic advantages are limited. As you said, no diagnostic indicator is perfect, and we are also aware of the harm caused by missed clinics for neonatal sepsis. Most of the diagnostic indicators

should be combined with specific clinical manifestations and other laboratory tests to play a role in their diagnosis. Therefore, we have revised our conclusions and hope that our reply can be recognized by you.

Before revision：

The neutrophil to lymphocyte ratio has a moderate diagnostic capacity with high sensitivity and specificity for diagnosing neonatal sepsis. It can provide a reference value for the early diagnosis of neonatal sepsis.

After revision：

Our findings suggest that the neutrophil-to-lymphocyte ratio is a useful indicator for the diagnosis of early neonatal sepsis, but it still needs to be combined with other laboratory tests and specific clinical manifestations.

sensitivity and specificity are often negatively correlated within studies, bivariate analysis has been proposed to take the correlation into the consideration.

Response: Thank you for your suggestion. We performed Spearman's correlation analysis on sensitivity and specificity according to your request. Spearman's rho= -0.0508 indicates that sensitivity and specificity are negatively correlated. Prob > |t| = 0.8631 showed that the difference was not statistically significant.

```
 Number of obs =      14
Spearman's rho =      -0.0508

Test of H0: Sens and Spec are independent
    Prob > |t| =       0.8631
```

I understand that you are an expert in statistics, thank you for your professional statistical opinion for our study. We have revised our statistics section according to your request, but about the sroc, we

have not included in the text Use hsroc instead, with regards to the model usage for diagnostic meta-analysis, I wonder if you meant to let us change to bivariate random-effect model. [13] Excluding my lack of knowledge in statistics, we will also be happy to respond to any further questions and comments you may have. We sincerely hope that this revised manuscript has addressed all your comments and suggestions. We earnestly appreciate your warm work and hope that the correction will meet with approval. Once again, thank you very much for your comments and suggestions.

VERSION 2 – REVIEW

REVIEWER	Deena Thomas Muthoot Healthcare, Neonatology
REVIEW RETURNED	08-Aug-2022

GENERAL COMMENTS	All the comments have been adequately addressed in the revised manuscript.

REVIEWER	Chun Sen Wu University of Southern Denmark, Research Unit of Gynecology and Obstetrics, Institute of Clinical Research
REVIEW RETURNED	22-Aug-2022

GENERAL COMMENTS	I am very happy to see that the methodologies have been improved a lot! However, there is still a very important issue should be handled together with a few minor things. An important issue As I mentioned before “Now it is well-known that I squared is not an absolute measure of heterogeneity between the studies. Statistics used in the meta-analysis should neither be based on I squared nor based on the heterogeneity.” Please see the reference: Borenstein M, Higgins JP, Hedges LV, Rothstein HR.Res Basics of meta-analysis: I2 is not an absolute measure of heterogeneity. Synth Methods. 2017 Mar;8(1):5-18. doi: 10.1002/jrsm.1230. Epub 2017 Jan 6.PMID: 28058794 Borenstein M.J Research Note: In a meta-analysis, the I2 index does not tell us how much the effect size varies across studies. Physiother. 2020 Apr;66(2):135-139. doi: 10.1016/j.jphys.2020.02.011. Epub 2020 Apr 16.PMID: 32307309 If you have access to YouTube, you can also watch video about the topic: I-squared Webinar - I2 Does Not Tell Us How Much the Effect Size Varies (https://www.youtube.com/watch?v=38wRNJIcqe0&list=PLX-KmXTn6p4jfu3ZOaqsCVD6afq0ilYUn&index=7) Therefore, the paragraph interpreting I squared as heterogeneity must be further revised. The authors of the above publications suggest using prediction interval (for example in the forest plot) to

	show heterogeneity. A few minor things sROC is from Summary Receiver Operating Characteristic Curve, therefore the line 136-137 (a combined receiver operating characteristic curve (SROC) fitting analysis) is better revised as for example “summary receiver operating characteristic (SROC) curve analysis” The sROC curve is a plot of the true positive rate (sensitivity) as a function of the false positive rate (1-specificity). Therefore, the x axis of the figure should be revised as “1-specificity” instead of “specificity”

VERSION 2 – AUTHOR RESPONSE

Reviewer: 6

Dr. Chun Sen Wu, University of Southern Denmark, Odense University Hospital

Comments to the Author:

I am very happy to see that the methodologies have been improved a lot! However, there is still a very important issue should be handled together with a few minor things.

An important issue

As I mentioned before “Now it is well-known that I squared is not an absolute measure of heterogeneity between the studies. Statistics used in the meta-analysis should neither be based on I squared nor based on the heterogeneity.” Please see the reference: Borenstein M, Higgins JP, Hedges LV, Rothstein HR.Res Basics of meta-analysis: I2 is not an absolute measure of heterogeneity. Synth Methods. 2017 Mar;8(1):5-18. doi: 10.1002/jrsm.1230. Epub 2017 Jan 6. PMID: 28058794

Borenstein M.J Research Note: In a meta-analysis, the I2 index does not tell us how much the effect size varies across studies. Physiother. 2020 Apr;66(2):135-139. doi: 10.1016/j.jphys.2020.02.011. Epub 2020 Apr 16. PMID: 32307309

If you have access to YouTube, you can also watch video about the topic: I-squared Webinar - I^2 Does Not Tell Us How Much the Effect Size Varies (https://www.youtube.com/watch?v=38wRNJIcqe0&list=PLX-KmXTn6p4jfu3ZOaqsCVD6afq0ilYUn&index=7)

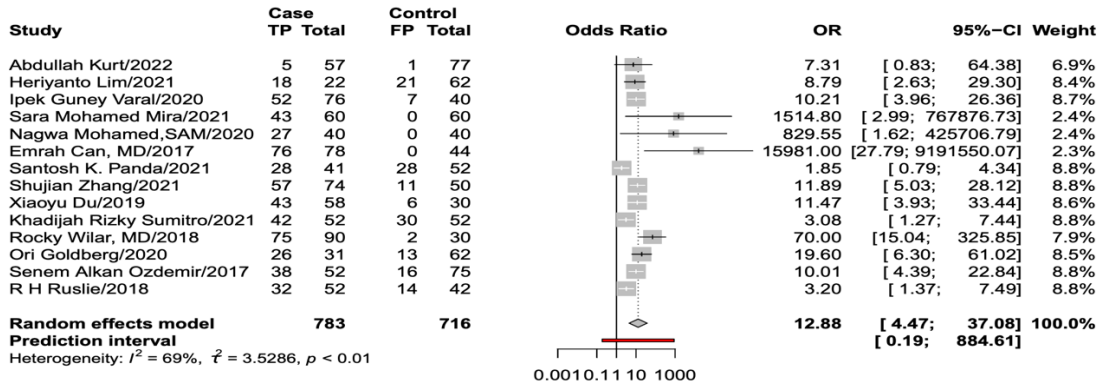
Therefore, the paragraph interpreting I squared as heterogeneity must be further revised. The authors of the above publications suggest using prediction interval (for example in the forest plot) to show

heterogeneity.

Response： Thank you for your suggestion. I am very sorry that I did not understand your meaning in the first revision and brought you unnecessary trouble. We have studied two references as you suggested and watched the video link you sent. According to your suggestion, We searched for relevant information and completed our statistical work through R (version 3.6.0). we modified the heterogeneity methodology section and the corresponding results section of the article, and we used the prediction interval in the forest plot as the standard to measure the heterogeneity. I hope our revision meets your requirements. If there are any shortcomings, we would be happy to respond to any further questions and comments that you may have.

The details are as follows:

Statistical heterogeneity was assessed using forest plots with 95% prediction interval, the tau-squared (τ^2) value and I^2 statistic. The 95% prediction interval was applied to estimate the effect size range in further studies [1]. If there was heterogeneity between the studies, the source of the heterogeneity was further explored, and threshold effect and nonthreshold effect analyses were carried out. **(Page 6，line104-106)**



The results show that, τ^2=3.5286, I^2=69% and 95% prediction interval is (0.19, 884.61).

Reference

1. Borenstein M. Research Note: In a meta-analysis, the I^2 index does not tell us how much the effect size varies across studies. J Physiother. 2020 Apr;66(2):135-139.

A few minor things

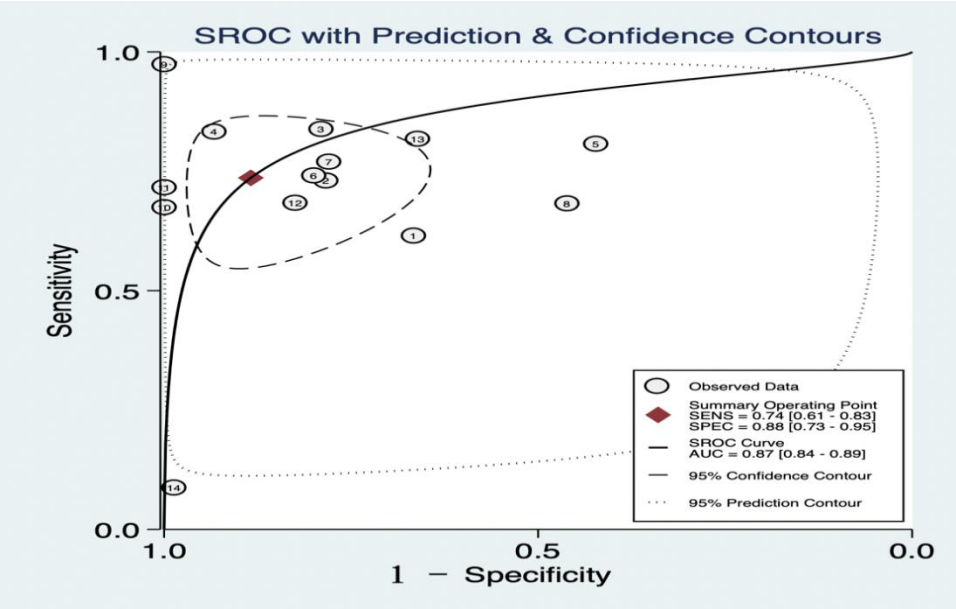
sROC is from Summary Receiver Operating Characteristic Curve, therefore the line 136-137 (a

combined receiver operating characteristic curve (SROC) fitting analysis) is better revised as for example "summary receiver operating characteristic (SROC) curve analysis"

Response : Thank you for your advice, According to your requirements, we have changed line 136-137 (a combined receiver operating characteristic curve (SROC) fitting analysis) to "summary receiver operating characteristic (SROC) curve analysis"

The sROC curve is a plot of the true positive rate (sensitivity) as a function of the false positive rate (1-specificity). Therefore, the x axis of the figure should be revised as "1-specificity" instead of "specificity"

Response：Thank you for your suggestion. According to your requirements, we have changed the x axis of the figure.



We sincerely hope that this revised manuscript has addressed all your comments and suggestions. Thank you very much again for your patient and detailed explanation of the problems in our article, which makes the methodology part of our article more perfect. We can assure you that you are a very responsible reviewer,We earnestly appreciate your warm work again. And we look forward to hearing from you regarding our submission. We would be happy to respond to any further questions and comments that you may have.

VERSION 3 – REVIEW

REVIEWER	Chun Sen Wu University of Southern Denmark, Research Unit of Gynecology and Obstetrics, Institute of Clinical Research
REVIEW RETURNED	06-Nov-2022

GENERAL COMMENTS	Thanks for the efforts! I have gone through the revised manuscript and I do not have further comments!