Appendix. Complete CASP checklist and Appraisal Summary

CASP Checklist:
For systematic reviews with meta-analysis of randomised controlled trials (RCTs)

Reviewer Name:	Bohanon Franco, C, Zackula RE, Raghuveer TS
Paper Title:	Active Treatment vs Expectant Management of Patent Ductus Arteriosus in Preterm Infants A Meta-Analysis
Author:	Santosi Buvaneswarran, MMed; Yi Ling Wong, MMed; Shen Liang, PhD; Swee Chye Quek, MD; Jiun Lee, MMed
Web Link:	
Appraisal Date:	7/15/2025

Appendix. Complete CASP checklist and Appraisal Summary

Section A: Is the basic study design valid for a systematic review?	
1. Did the systematic review address a clearly formulated research question? "Is active treatment of hemodynamically significant patent ductus arteriosus (PDA) in preterm infants better than an expectant management approach?" P: Preterm infants <33 weeks' gestation I: Active treatment of PDA C: Expectant management O: Death and/or moderate-to-severe BPD T: Treatment within the first two weeks of life	☒Yes ☐ No ☐ Can't Tell
2. Did the researchers search for appropriate study design(s) to answer the research question?	☒Yes ☐ No ☐ Can't Tell
Section B: Is the systematic review methodologically sound?	
3. Were all the relevant primary research studies likely to have been included in the systematic review? a) Searching for primary research studies A literature search was conducted by 2 independent reviewers (S.B. and Y.L.W.) using the PubMed (MEDLINE), Embase, and Cochrane Library databases, including both MeSH (Medical Subject Headings) terms and related keywords. The search was limited to studies that were published in English between January 1, 2010, and July 31, 2024, and included human participants only. Detailed search strategy is available in eTable 1. NOTE: a current Cochrane systematic review search showed more studies than reported by the authors that fit their search criteria. An independent Embase search was conducted using the PICO format, which included synonyms, produced 191 articles; whereas the flowchart in Figure 1 showed only 27 Embase articles with the Buvaneswarran et al. search terms. This may have indicated an incomplete search of articles. A more comprehensive search strategy may have been more successful.	☐Yes ☒ No ☐ Can't Tell
b) Screening primary research studies from the search	☒Yes ☐ No ☐ Can't Tell

Appendix. Complete CASP checklist and Appraisal Summary

The article clearly defined eligibility criteria, focusing on randomized clinical trials comparing active treatment of PDA with expectant management in preterm infants less than 33 weeks' gestation, with treatment initiated within the first two weeks of life. Screening was conducted independently by two reviewers, with a third author resolving discrepancies, ensuring reliability. Titles and abstracts were reviewed first, followed by full-text screening, which aligns with standard systematic review practice.	
c) Selecting primary research studies to include in the systematic review The selection of primary research studies was conducted through full-text review, with exclusions made according to the predefined eligibility criteria. While the process included a third reviewer to resolve any discrepancies in screening and data extraction, the study did not report a formal measure of inter-rater reliability, such as Cohen's kappa, which limits insight into the consistency of reviewer agreement.	☒Yes ☐ No ☐ Can't Tell
d) Summarising the search and its outputs The search and its outputs were summarized using a PRISMA flowchart (Figure 1), providing a transparent overview of the study selection process. Exclusion reasons were clearly documented and made available in a supplemental chart (eTable 2). However, the review did not include hand-searching of current or recently published meta-analyses or systematic reviews, which may have limited the comprehensiveness of the search strategy.	☒Yes ☐ No ☐ Can't Tell
4. Did the researchers assess the validity or methodological rigour of the primary research studies included in the systematic review? The researchers assessed methodological quality using the Cochrane Risk of Bias tool, noting that many studies had "some concerns." However, these evaluations were not adequately incorporated into the interpretation, and treatment arm approaches were not standardized before analysis. Importantly, the authors did not report certainty of evidence, which is a significant omission given their conclusions suggesting changes to standard of care practices. No statistical adjustments were made for the small number of included studies or small sample sizes,	☐Yes ☒ No ☐ Can't Tell

increasing the likelihood of false positives (e.g., composite death/BPD with RR=1.10, 95% CI 1.01–1.19). Additionally, there was no assessment of publication bias, such as a funnel plot, which may have contributed to misclassification of findings. The absence of certainty ratings (e.g., GRADE/GradePro) and methodological safeguards like Sidik-Jonkman or Knapp-Hartung adjustments weakens the rigor of the review. Given these limitations, the work may have been better classified as a scoping review or narrative literature review, rather than a systematic review and meta-analysis, as it does not seem to fully meet the evidentiary standards required for the latter.	
5. Did the researchers extract, and present information from the individual primary research studies appropriately and transparently? (a) Extraction of data Data extraction was conducted by two independent reviewers with third-party adjudication, which supports transparency. Forest plots and tables were used to present study characteristics, treatment effects, and subgroup data. However, the analyses were not adjusted for the small number of included studies, and the meta-analyses were underpowered. None of the subgroups had sufficient evidence to support firm conclusions, which weakens the reliability and interpretability of the findings.	☐ Yes ☒ No ☐ Can't Tell
(b) Presentation of data The researchers did present the key characteristics of the included primary studies in tables, outlining participant numbers, demographics, interventions, comparators, outcomes, and study timeframes. Results were summarized using both tables and Forest plots, which displayed effect sizes, confidence intervals, and p-values. However, the analyses were not adjusted for the small number of included studies, limiting the reliability of findings. Additionally, no Forest plots were provided comparing critical aspects such as timing of first treatment, dosing, or other regimen-related variables, which weakens the depth and completeness of the synthesis.	☒ Yes ☐ No ☐ Can't Tell

Appendix. Complete CASP checklist and Appraisal Summary

Section C: Are the results of the systematic review trustworthy?	
6. Did the researchers analyse the pooled results of the individual primary research studies appropriately? The pooled analyses were not conducted appropriately. Several included studies had small sample sizes, raising the risk of false positives (e.g., composite death/BPD with RR=1.10, 95% CI 1.01–1.19), yet no statistical adjustments such as Sidik-Jonkman or Knapp-Hartung methods were applied. Outcomes like PVL, ROP treatment, and pulmonary hemorrhage included zero events, but no clear adjustments were reported to account for this limitation. While the authors stated that Egger and Begg tests were performed, results were not provided, and no funnel plot was included to assess potential publication bias. Additionally, no power analysis was completed. Although a random-effects model was used, the lack of adjustments for zero events, small sample sizes, and missing publication bias assessment substantially undermines the validity of the pooled results.	☐ Yes ☒ No ☐ Can't Tell
6.1 Subgroup analysis The subgroup analyses were incomplete and underpowered. Key comparisons related to treatment protocols—such as dose regimens, timing of first treatment, and duration of therapy—were missing, limiting the ability to draw meaningful conclusions across studies. Additionally, the subgroup analyses that were conducted lacked sufficient power, further weakening their reliability and reducing confidence in any observed differences.	☐ Yes ☒ No ☐ Can't Tell
6.2 Meta-regression - Did not complete a meta-regression (N/A)	☐ Yes ☐ No ☐ Can't Tell
7. Did the researchers report any limitations of the systematic review and, if so, do the limitations discussed cover all the issues you identified during critical appraisal? The researchers reported several limitations, including variability in treatment protocols, differences in BPD definitions, the absence of an IPD meta-analysis, potential type II errors in subgroup analyses, and higher rates of open-label treatment in the expectant group. They also noted that because PDA treatment was initiated	☐ Yes ☒ No ☐ Can't Tell

within the first two weeks of life (with seven trials starting within the first 72 hours), the findings may not be generalizable to infants treated later. However, despite this, the authors still presented broad composite conclusions. Importantly, while they acknowledged the small number of studies and the possibility of false negatives, they did not adjust their statistical model accordingly. This omission, particularly in subgroup analyses, means that key methodological concerns were not adequately addressed, and the reported limitations did not fully cover the issues identified in critical appraisal.	
7.1 Subgroup-analysis See above	☐ Yes ☒ No ☐ Can't Tell
7.2 Meta-regression Did not complete a meta-regression	☐ Yes ☐ No ☐ Can't Tell
8. Would the benefits of intervention outweigh any potential disadvantages, harms and/or additional demand for resources associated with acting on the results? Recommending a change to the current standard of care for aggressive PDA treatment based on this meta-analysis would be premature and potentially harmful. The conclusions were drawn from a small number of studies, many of which had methodological limitations and underpowered subgroup analyses. Acting on such limited evidence could introduce risks to patient safety and place unnecessary demands on clinical resources without sufficient justification.	☐ Yes ☒ No ☐ Can't Tell
Section D: Are the results of the systematic review relevant locally?	
9. Can the results of the systematic review be applied to your local population/in your local setting or context? Yes, the results of the systematic review have some local applicability, though with important caveats. The outcomes assessed—such as mortality and BPD—are relevant to the local population, and the baseline characteristics reported in eTable 6 appear broadly representative, even though the included studies spanned multiple countries. However, the lack of a full demographic breakdown of the maternal and infant populations limits confidence in the direct comparability to local settings. Therefore, while the findings may have some relevance,	☒ Yes ☐ No ☐ Can't Tell

caution is warranted in applying the conclusions without more detailed demographic alignment.	
Section E: Will the implementation of the results represent greater value for your service users or population?	
10. If actioned, would the findings from the systematic review represent greater or additional value for the individuals or populations for whom you are responsible? If implemented, the findings suggesting a more conservative approach could, in theory, provide added value by improving outcomes and reducing unnecessary interventions and costs. However, the validity of these results is undermined by significant methodological weaknesses, including underpowered analyses, lack of certainty assessments, and inadequate adjustments for small sample sizes. While the outcomes studied are applicable to the local population, it is likely that the findings of the systematic review will not confer greater or additional benefit or value on the individuals and populations for whom you are responsible, given the limitations of the evidence base.	☐ Yes ☒ No ☐ Can't Tell

What is your conclusion about the systematic review – can it be used to support evidence-based decision-making? *Based on the vast methodological concerns, the lack of a certainty of evidence assessment, potential for publication bias, and missing statistical adjustments, the results and conclusions presented by the authors should not be used to support proper evidence-based decision-making.*	☐ Yes ☒ No ☐ Can't Tell

Appendix. Complete CASP checklist and Appraisal Summary

CASP General SR Checklist: Collation of critical appraisal responses

Yes	Checklist question	Can't tell	No
	A. Is the basic study design valid for a systematic review?		
X	1. Did the systematic review address a clearly formulated research question?		
X	2. Did the researchers search for appropriate study designs to answer the research question?		
	B. Is the systematic review methodologically sound?		
	3. Were all relevant primary research studies likely to have been included in the systematic review?		X
	4. Did the researchers assess the validity or methodological rigour of the primary research studies included in the systematic review?		X
	5. Did the researchers extract, and present information on the individual primary research studies appropriately and transparently?		X
	C. Are the results of the systematic review trustworthy?		
	6. Did the researchers analyse the pooled results of the individual primary research studies appropriately?		X
	7. Did the researchers report any limitations of the systematic review and, if so, do the limitations discussed cover all the issues in your critical appraisal?		X
	8. Would the benefits of intervention outweigh any potential disadvantages, harms and/or additional demand for resources associated with acting on the results?		X
	D. Are the results of the systematic review relevant locally?		
	9. Can the results of the systematic review be applied to your local population/in your local setting or context?	X	
	E. Will the implementation of the results represent greater value for your service users or population?		
	10. If actioned, would the findings from the systematic review represent greater or additional value for the individuals or populations for whom you are responsible?		X

APPRAISAL SUMMARY: *List key points from your critical appraisal that need to be considered when assessing the validity of the results and their usefulness in decision-making.*		
Positive/Methodologically sound	**Negative/Relatively poor methodology**	**Unknowns**
• The review addressed a clearly formulated research question with an appropriate PICO framework. • Eligibility criteria were clearly defined (RCTs, preterm infants <33	• Literature search was incomplete; independent searches yielded more eligible articles than reported, suggesting missed studies. • No inter-rater reliability measure (e.g., Cohen's kappa) was reported.	• Generalizability to populations beyond the first two weeks of treatment remains unclear. • Lack of full demographic breakdown of maternal and infant

weeks, PDA treatment vs expectant management, treatment initiated within two weeks). • Screening and selection were performed by two reviewers with a third to resolve discrepancies. • PRISMA flow diagram was provided, and reasons for exclusion were transparent. • Study characteristics, participant data, and intervention details were summarized in tables and Forest plots, including effect sizes, confidence intervals, and p-values. • Outcomes assessed (mortality, BPD, etc.) are relevant and broadly reflective of the local population.	• No hand-searching of references or recent reviews was performed. • Cochrane Risk of Bias assessments were conducted but not meaningfully integrated into interpretation of results. • No certainty of evidence assessment (e.g., GRADE/GradePro) was reported. • Statistical limitations were not addressed: no adjustment for small study numbers, zero-event outcomes, or underpowered subgroups. • Egger and Begg tests were claimed but not reported; no funnel plot was provided to assess publication bias. • Subgroup analyses were incomplete (e.g., dose, timing, regimen) and underpowered; no meta-regression performed. • Limitations acknowledged by the authors did not adequately cover major issues, such as failure to adjust models for small samples or risk of false positives. • Conclusions suggesting changes to standard of care were drawn despite insufficient evidence, making the recommendations potentially harmful. • Overall methodological weaknesses undermine validity and suggest the study may be more appropriate as a scoping or narrative review rather than a systematic review with meta-analysis.	participants limits certainty about applicability across diverse settings. • The true presence or absence of publication bias cannot be assessed due to missing funnel plots and unreported test results.