**Peer-Review comments and authors responses**

"Iron Deficiency as Risk Factor for Adverse Cardiovascular Outcomes in Pediatric Heart Disease: A Systematic Review and Meta-Analysis"

We are grateful to the reviewers for their helpful reviews. We have attempted to respond to all of the reviewers' suggestions. Below is a keyed response to all comments (bold), with the corresponding changes in the revised manuscript provided *in italics* for reference.

**Reviewer 1:**

Recommendation: revisions required.

**Comments for authors:** This is a well-written and highly relevant systematic review and meta-analysis evaluating iron deficiency (ID) as a risk factor for adverse cardiovascular outcomes (ACV) in pediatric heart disease. The authors have done a fantastic job of synthesizing a limited yet critical body of literature and addressing the high risk of bias in the included studies. I appreciated the comprehensive supplementary material and the cautious tone of the conclusions. My comments below are intended to strengthen the clarity, methodological transparency, and overall impact of the manuscript.

> **Comment:** Abstract: If the word limit allows, I recommend clarifying the inclusion and exclusion criteria directly in the abstract. For instance, specifying that only peer-reviewed, English-language studies were included, and that animal studies and abstracts were excluded.

**Response:** *Thank you for your suggestion. We have added the inclusion and exclusion criteria to the abstract.*
*Line 61–62, page 2: "Studies were excluded if they reported data from adults or animals, if only abstract data were available, or if they were published in a non-English language."*

> **Comment:** It appears from the PRISMA flowchart that abstracts were excluded, but this should be explicitly stated in the exclusion criteria. Similarly, please clarify whether study protocols or grey literature were considered or excluded.

**Response:** *Thank you for your suggestion. We have added this information to the exclusion criteria. Study protocols and grey literature were excluded from this systematic review, as they do not provide data to assess the relationship between iron deficiency (ID) and adverse cardiovascular outcomes (ACV), including death, heart transplant, or mechanical circulatory support.*
*Line 178–180, page 4: "Studies including data from adults and animals, and those published in a non-English language or available only as abstracts were excluded from this analysis. Likewise, study protocols and grey literature were not included in this review."*

> **Comment:** The exact date of the search (currently noted in the results section as December 10th, 2024) should be moved to the Methods.

**Response:** *Thank you for your suggestion. We have made this change.*
*Line 189, page 5: "The search covered the databases from their inception until December 10, 2024."*

> **Comment:** Please clarify whether study selection and data extraction were done in pairs as it's noted in the risk of bias assessment, and if not, consider noting this as a limitation. Ideally, data extraction should be performed independently by two reviewers with discrepancies resolved by consensus or a third reviewer.

**Response:** *Thank you for your comment and suggestion. Study selection and data extraction were conducted by the first author (K.L.L.), as only the first author had access to the databases used for this review. We have added this as a potential limitation of the study results.*
*Line 575–581, page 18: "Study selection and data extraction were performed by the first author, which may have limited the number of included studies due to individual judgment and increased the risk of missing relevant data;*

*however, all findings were carefully discussed within the research team, and data collection followed the prespecified plan."*

> **Comment:** The inclusion of Higgins et al., 2017 outside of the formal search suggests reference mining. Please clarify how this study was identified (e.g., citation review of included articles) and reflect this in the PRISMA flowchart and methods section.

**Response:** *Thank you for your question. This is a relevant paper published by a group with high expertise in pediatric heart failure, and it was cited by one of the articles included in this review. Although our systematic review used keywords and terms that should have retrieved this paper, it was not captured in the systematic search. After discussing with the research team and reviewing all search terms, we decided to include the paper due to its relevance to the topic.*
*Reference 19: Puri K, Price JF, Spinner JA, Powers JM, Denfield SW, Cabrera AG, Tunuguntla HP, Dreyer WJ, Shah MD. Iron Deficiency Is Associated with Adverse Outcomes in Pediatric Heart Failure. J Pediatr. 2020 Jan;216:58–66.e1. doi:10.1016/j.jpeds.2019.08.060. Epub 2019 Oct 11. PMID: 31610927.*
*We clarified this in the results section:*
*Line 259-260, page 6: "This paper was identified through reference screening during the full-text review process of included articles."*

> **Comment:** I'd recommend including a list of the seven excluded full-text articles (with author names and titles) in the Supplementary Material, to enhance transparency.

**Response:** *Thank you for your suggestion. We have added a supplementary information 3 with excluded full texts and abstracts.*
*Line 266-267, page 6: "The titles and authors of the articles excluded after full-text review are provided in Supplementary Information 3."*

> **Comment:** Please indicate in the Methods that the REML (Restricted Maximum Likelihood) method was used to estimate between-study variance in the random-effects meta-analysis, and briefly justify this choice.

**Response:** *Thank you for your suggestion. We have added this information to the Methods section and justified the choice.*
*Line 236–239, page 5-6: "Considering the heterogeneity within the pediatric population with heart disease, a random-effects model meta-analysis was used to pool the data, with between-study variance estimated using the Restricted Maximum Likelihood (REML) method. REML was chosen because it provides less biased estimates of heterogeneity compared with alternative approaches, particularly when the number of studies is small (Piepho, 2024)."*

> **Comment:** Some minor grammatical issues should be addressed. For instance: "ID was associated ACV outcome" → "ID was associated with ACV outcome" "ID as protector factor" (including in the Figure) → "ID as a protective factor."

**Response:** *Thank you for your suggestion. We have made these corrections in both Table 2 and the figure (Forest plot).*

> **Comment:** The narrative synthesis is informative but somewhat lengthy. Please consider summarizing it more concisely.

**Response:** *Thank you for your comment. This is the new version of the narrative synthesis:*
*Lines 338-363, pages 11-12:*
*"ACV Outcomes*
*Taking into consideration the three studies with available information on ACV stratified by ID, two studies identified ID as a risk factor for ACV in pediatric patients with heart disease. (Luxford et al., 2024; Puri et al., 2020) On the other hand, one study did not show a significant effect of ID on clinical outcomes in this population.(Phillips et al., 2023)*

*Puri et al., analyzed data from 107 patients with either biventricular or single-ventricle heart disease. They found that ID significantly increased the odds of experiencing ACV events within six months, reporting an OR of 6.9 (95% CI 2.0, 24.1) for biventricular disease and OR 8.1 (95% CI 1.5, 42.3) for single-ventricle disease.ID-exposed patients also had a significantly higher cumulative length of hospital stay compared to those without ID (70% vs. 10% for biventricular and 65% vs. 2% for single-ventricle disease; p=0.001 for both).. (Puri et al., 2020)*
*Phillips et al. assessed ID as a risk factor for ACV outcomes in children with HF induced by DCM. In a group of 47 individuals, ID was not associated with ACV outcome with an HR of 1.36 (95% CI 0.61,3.04). The researchers also explored the association between ID and death, listing on the transplant registry, or VAD placement, finding no significant association.(Phillips et al., 2023)*
*Luxford et al. investigated ID as risk factor for ACV in 47 children with DCM. In this cohort, ID was associated with ACV outcome represented by the HR 3.84 (95% CI 1.27, 11.59), P=0.017 in a univariate analysis. In this analysis, the increased risk over time was nearly fourfold in the group exposed to ID. Additionally, the median cumulative length of stay at six months follow-up was higher in the group exposed to ID when compared with those IS (14.7% vs 1.1%, P=0.002). (Luxford et al., 2024)"*

> **Comment:** In addition to the limitations of the included studies, I encourage the authors to expand the Limitations section to include limitations of the review process, such as restriction to English-language studies, exclusion of grey literature, and others.

**Response:** *Thank you for this suggestion. We have already addressed these points in the limitations section.*
*Lines 579-581, page 18: "In addition, our review was restricted to English-language publications and did not include grey literature, which may have introduced language and publication bias and limited the comprehensiveness of the evidence base."*

> **Comment:** At the end of the Conclusion, consider reiterating the value of this synthesis despite limitations, particularly for guiding future prospective research and clinical screening strategies.

**Response:** *Thank you for your suggestion. We have revised the Conclusion to reiterate the value of our synthesis, as you recommended. We've added a new sentence that explicitly addresses how our findings can guide future research and clinical strategies.*
*Lines 629-631, page 19: "Despite these limitations, this synthesis provides valuable insights for shaping prospective research agendas and highlights the need to consider ID in clinical screening strategies for children with heart disease."*

> **Comment:** I commend the authors for their careful approach and thoughtful presentation. The paper contributes meaningfully to the literature, and my suggestions are intended to refine and strengthen what is already an important contribution.

**Reviewer 2:**

Recommendation: Accept Submission

**Comments for authors:** The paper "Iron Deficiency as a Risk Factor for Adverse Cardiovascular Outcomes in Pediatric Heart Disease: A Systematic Review and Meta-Analysis" presented a methodologically robust study about the association between iron deficiency (ID) and adverse cardiovascular outcomes (ACV) in individuals under 21 years of age diagnosed with heart disease. The study reports an important evidence gap in pediatric cardiology and it was developed according to appropriate methodological standards.

**Strengths of the study:**

> • The research question is clearly defined using the PECOS framework and supports a relevant clinical knowledge gap (the lack of pediatric-specific data in the literature).

• The review protocol was registered in PROSPERO and reported using PRISMA-P and PRISMA 2020 guidelines.

• Inclusion/exclusion criteria, definitions of ID and ACV, and explanation for the age cut-off are appropriate and supported by the literature.
• The search strategy is applied to PubMed, EMBASE, and Web of Science, with a supplementary file showing detailed terms.
• The study selection process is described, and Rayyan was used for screening. Used a CONSORT diagram to present study flow. The process is well-structured and reproducible.
• Risk of bias assessment using the ROBINS-E tool is conducted by two independent reviewers with a third for conflicting assessment.
• The choices of data synthesis using random-effects meta-analysis, and assessment of heterogeneity and publication bias are methodologically explained. Subgroup analyses were pre-specified, enhancing the strength of the method.

**Minor Areas for Improvement:**

**Comment:** Page 5 row 235: you included one study that was not identified through the primary search strategy, but the manuscript does not clarify how it was retrieved. Indicating whether it was found through hand search or reference screening would enhance methodological reproducibility.
Overall, this is a very solid review with a strong methodological process.

**Response:** *Dear Reviewer, thank you for your observation and your positive assessment of our review.*
*Page 6, line 259-260: We have now revised the "Study Selection" section to explicitly state that the study by Higgins et al. (2017) "was identified through reference screening during the full-text review process of included articles."*

**Reviewer 3:**

Recommendation: Revisions Required

**Comments for authors:** Dear Authors, thank you for the opportunity to review your manuscript entitled: "Iron Deficiency as Risk Factor for Adverse Cardiovascular Outcomes in Pediatric Heart Disease: A Systematic Review and Meta-Analysis" I greatly appreciate your effort to address an important and understudied clinical question. The relationship between iron deficiency (ID) and adverse cardiovascular outcomes (ACV) in pediatric heart disease is highly relevant, and your initiative to synthesize the available evidence is commendable. Below, I provide a series of comments organized by manuscript sections, aimed at strengthening its methodological rigor, clarity of interpretation, and clinical relevance.

**Comment:** In the abstract, you note that heterogeneity was low ($I^2$=16.25%). Given the inclusion of only three studies, did you consider clarifying that the statistical power to detect heterogeneity is limited in small samples? This would help temper the interpretation of $I^2$.

**Response:** *Thank you for this valuable comment. We agree that clarifying the limitations of interpreting $I^2$ in a small sample is important for accurate interpretation. As you suggested, we have added a sentence to the Abstract's Results section to reflect this point. Page 2, Line 70-71, we have added: "However, the power to detect heterogeneity is limited when few studies are included."*

**Comment:** Did you consider specifying in the abstract that all included studies were retrospective and single-center? This may help readers better gauge the level of evidence from the outset.

**Response:** *Thank you for this valuable suggestion. We agree that providing this information in the abstract is important for contextualizing the findings. As you recommended, we have updated the Results section of the abstract to specify that the included studies were retrospective and single-center.*
*On Page 2, Line 67, we have added: "Five retrospective single-center studies were included."*

> **Comment:** Could you further justify the decision to group congenital heart disease (CHD) and cardiomyopathies together, considering their different pathophysiologies, clinical trajectories, and responses to treatment.

**Response:** *Thank you for this valuable comment. We agree that the distinction between congenital heart disease (CHD) and cardiomyopathies is important. We have added a clarification to the Introduction to justify our approach. Our decision to group these conditions together is based on their shared clinical trajectory toward heart failure and the need for advanced interventions. This approach is also consistent with other studies in the field that analyze both etiologies together due to their shared pathophysiological pathways leading to heart failure.*
*Line 101-105, page 3: "Although CHD and cardiomyopathies differ in etiology and progression, their shared clinical trajectory toward HF and end-stage complications, such as the need for MCS or HT, provides a common framework for evaluating the role of ID.This approach is consistent with other studies in the field that analyze both etiologies together due to their shared pathophysiological pathways leading to HF."*

> **Comment:** Is it methodologically appropriate to compare the prevalence of ID in your population (hospitalized patients with severe heart disease) with the 38.6% figure from Weyand et al., which reflects healthy adolescents? A brief discussion on the comparability of these populations would be helpful.

**Response:** *Thank you for this excellent methodological point. We agree that the populations are not directly comparable and that this distinction is crucial for proper interpretation. We have added a brief discussion to the manuscript to address this.*
*Page 15, Lines 460-464: "While the 38.6% ID prevalence reported by Weyand et al. in healthy adolescents provides a useful reference point, it may not be directly comparable to the hospitalized pediatric heart disease population studied here, which likely represents a group with greater disease severity and inflammation, factors known to influence iron metabolism."*

> **Comment:** Given that the included age range spans up to 21 years, did you consider stratifying by age (e.g., children vs. adolescents), given developmental differences in iron metabolism and disease progression?

**Response:** *Thank you for this insightful suggestion. We agree that stratifying by age is an important consideration due to developmental differences. Unfortunately, this was not feasible because of the limited and inconsistent reporting of age-stratified data across the included studies, along with their small sample sizes. We have added this as a limitation to the Discussion section to enhance the transparency of our findings.*
*Lines 601-603, page 18: "Although developmental differences in iron metabolism may exist between younger children and adolescents, subgroup analysis by age was not feasible due to a lack of consistent age-stratified data across studies and small sample sizes."*

> **Comment:** Could you provide more detail on the rationale for selecting the specific cutoff points used to define ID? Some literature suggests higher ferritin thresholds in inflammatory states such as heart failure.

**Response:** *Thank you for this important comment. We agree that higher ferritin thresholds are a valid consideration in inflammatory states.*
*In our review, we intentionally prioritized studies that used a multi-marker approach to defining iron deficiency, which is considered the gold standard in pediatric and heart failure literature. This method, which combines multiple iron parameters (ferritin, serum iron, transferrin, and TSAT), helps to mitigate misclassification that can occur when relying on ferritin alone, as ferritin can be elevated by inflammation.*
*We have revised the Methods section to clarify this rationale.*
*Lines 163-168, page 4: "Therefore, to reduce misclassification due to inflammation, a combination of biomarkers was used. Specifically, ID was defined as the presence of two or more of the following: serum iron <50 μg/dL, ferritin <20 ng/mL, transferrin >300 ng/mL, and transferrin saturation (TSAT) <15%. This multi-marker definition aligns with pediatric clinical research and large-scale population surveys like NHANES. (Looker et al., 1995; Puri et al., 2020)"*

**Comment:** Given the heterogeneity in ID definitions across studies (e.g., ferritin, transferrin, TSAT), did you consider performing subgroup or sensitivity analyses based on these differing criteria?

**Response:** *Thank you for this excellent suggestion. We agree that the heterogeneity in iron deficiency (ID) definitions across studies is a critical consideration.*
*We initially planned to conduct subgroup and sensitivity analyses based on these definitions, as outlined in our methods. However, due to the small number of studies and the limited availability of disaggregated data, performing a robust analysis was not feasible. We have explicitly acknowledged this limitation in the manuscript to ensure transparency.*
*Lines 610-613, page 18: "The variation in ID definitions across the studies complicates the classification of patients, limits our ability to conduct subgroup or sensitivity analyses, and may lead to misclassification bias, potentially distorting the observed relationship between ID and ACV."*

**Comment:** Did you explore whether prospective studies were excluded due to other eligibility criteria? Clarifying this would help contextualize the scope of included evidence.

**Response:** *Thank you for this thoughtful observation. We did not exclude any studies based on their prospective design alone.*
*However, after a thorough screening, we found that no prospective studies met all of our predefined eligibility criteria, specifically the requirement to report on both ID and ACV outcomes in the target population of individuals under 21 years of age with heart disease.*
*To enhance methodological clarity, we have added a sentence to the Study Selection subsection of the Results.*
*Lines 263-266, page 6: "While both prospective and retrospective studies were eligible for inclusion, no prospective studies met all criteria, particularly the requirement to report both ID and ACV outcomes in individuals under 21. As such, all included studies were retrospective in design."*

**Comment:** Applying the ROBINS-E tool is a strength. Would it be possible to include a domain-level summary table (e.g., confounding, selection bias, etc.) to enhance transparency?

**Response:** *Thank you for this valuable suggestion. We agree that a domain-level summary table would significantly enhance the transparency and clarity of our risk-of-bias assessment.*
*As you recommended, we have created a summary table that presents the ROBINS-E tool's domain-level judgments for each included study. This table will be included in our revised manuscript as a Supplementary Information 5, providing a more detailed and easily digestible overview for the reader.*

**Risk of Bias Assessment of Included Studies Using ROBINS-E**

| Study | Confounding | Selection of participants | Measurement of exposure | Post-exposure interventions | Missing data | Measurement of outcome | Selection of reported result | Overall risk of bias |
|---|---|---|---|---|---|---|---|---|
| Higgins, 2017 | Some concerns | Low | High | High | Very high | Low | Very high | Very high |
| Puri, 2020 | High | Low | High | High | Very high | Low | Very high | Very high |
| Phillips, 2023 | High | Low | High | High | Very high | Low | Very high | Very high |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Luxford, 2024 | High | Low | High | High | Very high | Low | Very high | Very high |
| Newland, 2024 | High | High | High | High | Very high | High | Very high | Very high |

**Comment:** Since some studies reported hazard ratios (HR) while others used odds ratios (OR) derived from raw frequencies, did you consider whether this heterogeneity in effect measures could affect the validity of the pooled analysis? Might it be appropriate to either exclude HR-based studies or clearly caution readers about the differences in interpretation?

**Response:** *Thank you for this valuable and relevant comment. We agree that combining different effect measures such as odds ratios (ORs) and hazard ratios (HRs) is a critical methodological consideration.*
*As you noted, HRs and ORs represent distinct measures of effect and their pooling can introduce interpretive limitations. To address this, we have added a statement to the Limitations section to explicitly caution readers about this.*
*Lines 617-621, page 19: "Furthermore, while all effect estimates in the meta-analysis were expressed as ORs, two of the original studies reported HRs, which capture time-to-event relationships. To maintain consistency, we calculated ORs using raw frequencies. However, this approach may oversimplify the temporal dimension of outcomes and limit comparability. Future meta-analyses with access to individual participant time-to-event data would permit more robust pooled HR estimates."*

**Comment:** You mention planned subgroup analyses (e.g., by sex, heart disease type), but none were presented. Could you explain why they were not performed and whether future work might explore them?

**Response:** *Thank you for this relevant observation. We agree that subgroup analyses are crucial for exploring potential modifiers and were a pre-specified component of our review protocol.*
*We were unable to perform the planned subgroup analyses (e.g., by sex, heart disease type, or study quality) due to significant methodological limitations. The primary challenges were the small number of eligible studies (n=3 for ACV outcomes) and, critically, the lack of consistently reported, disaggregated data in the original publications. Without stratified data, any subgroup analysis would have been statistically underpowered and could have led to misleading conclusions.*
*We have now explicitly acknowledged this limitation in the Limitations section of the manuscript.*
*Lines 610-611, page 18: "The small number of studies included also limited the subgroup analysis planned in the methodology plan."*

**Comment:** Two studies (Higgins and Newland) did not report the primary outcome (ACV). Could you clarify whether their inclusion was solely for prevalence estimation, and whether separating the analyses might avoid methodological confusion?

**Reponse:** *Thank you for this important clarification request. You are correct in your assessment.*
*The studies by Higgins et al. (2017) and Newland et al. (2024) were included exclusively for our secondary objective of estimating the prevalence of iron deficiency. They did not report on the primary outcome of ACV, and their data were therefore not included in the meta-analysis for ACV outcomes.*
*We have clarified this distinction in the Results section of the manuscript.*
*Lines 385-388, page 12-13: "Of the five studies included, three reported data on adverse cardiovascular outcomes (Phillips, Luxford, and Puri). These were included in the meta-analysis. The remaining two studies (Higgins and Newland) did not report ACV outcomes and were therefore considered only in the prevalence analysis."*

**Comment:** The Luxford study yielded an OR with a wide confidence interval (1.03–77.37). Did you consider conducting a sensitivity analysis excluding this study to evaluate its impact on the pooled effect?

**Response:** *Thank you for this valuable suggestion. We recognize that the wide confidence interval of the Luxford study's OR could influence the pooled effect, and a sensitivity analysis is a standard approach to evaluate this. However, we did not perform a sensitivity analysis for this reason due to the very limited number of included studies (n=3 for the meta-analysis). Excluding the Luxford study would have reduced the total number to only two studies, which would have compromised the statistical power and robustness of our findings.*
*We have acknowledged this limitation and express our intent to update the meta-analysis in the future as more studies become available, which will allow for more robust analyses, including sensitivity analyses.*
*Lines 610-611, page 18: "The small number of studies included also limited the subgroup analysis planned in the methodology plan."*

> **Comment:** Did you consider reporting additional heterogeneity indicators (e.g., $\tau^2$) or including a prediction interval to better inform readers, especially given the small number of studies?

**Response:** *Thank you for this thoughtful suggestion. We agree that a small number of studies can limit the interpretability of heterogeneity statistics.*
*In our meta-analysis of three studies, we chose to apply a random-effects model to account for potential heterogeneity, even though the $I^2$ value was low. This approach minimizes the risk of overinterpreting the findings.*
*We did not report additional indicators like $\tau^2$ or prediction intervals because these measures are statistically unstable and unreliable with so few studies.*
*We agree that these measures would be valuable for informing readers, and we plan to incorporate them into future updates of this meta-analysis as more studies become available.*

> **Comment:** In Phillips' study, no association was found between ID and ACV. Might this be explained by less confounding due to comparable disease severity between groups? A brief reflection on this would strengthen the narrative synthesis.

**Response:** *Thank you for this insightful comment. We agree that the differences in findings between studies, including the lack of association in the Phillips study, may be influenced by confounding factors such as disease severity.*
*We have already addressed this possibility in the Discussion section, under the heading "Discrepancies in Findings Across Individual Studies." As you noted, the differences in disease severity, and other factors, could indeed explain the variance in results.*
*On Page 16, line 513-515: "The discrepancies among these findings likely stem from multiple factors, including differences in how ID was defined and variations in disease severity, which can lead to confounding by severity and exposure bias."*

> **Comment:** The pooled OR (4.44) is substantial. Do you believe this effect size might be inflated due to the lack of adjustment for key confounders (e.g., ventricular function, natriuretic peptides, nutritional status, socioeconomic status)?

**Response:** *Thank you for this insightful comment. We agree that the substantial pooled odds ratio may be inflated due to the lack of adjustment for key confounders.*
*As you noted, a significant limitation of the current literature is the absence of consistently reported, adjusted analyses that account for critical factors such as ventricular function, natriuretic peptides, nutritional status, and socioeconomic status.}*
*As we discussed, the larger effect size observed in our meta-analysis (OR 4.44) is likely related to this lack of adjustment, especially when compared to large, well-adjusted studies in adults with heart failure that report smaller effect sizes (e.g., adjusted HRs of 1.58 and 1.42).*
*This discussion in our manuscript acknowledges this limitation and cautions readers about the potential for confounding, providing a more nuanced interpretation of our findings.*
*Lines 485-488, page 16: "The larger effect size in this review compared to previous studies in adults with HF may be related to the lack of adjustment for confounders, such as sex, race, socioeconomic status, and markers of disease severity like natriuretic peptides, ventricular dysfunction, and functional class."*

**Comment:** When comparing to adult studies, could you clarify that those studies adjusted for multiple covariates, unlike the pediatric studies included here, which might partially explain the larger observed effect?

**Response:** *Thank you for this insightful observation. We agree that the differences in covariate adjustment are a critical factor in explaining the larger effect size observed in our meta-analysis compared to adult studies.*
*As you noted, this point has already been addressed in the Discussion section.*
*Lines 485-488, page 16: "The larger effect size in this review compared to previous studies in adults with HF may be related to the lack of adjustment for confounders, such as sex, race, socioeconomic status, and markers of disease severity like natriuretic peptides, ventricular dysfunction, and functional class."*

**Comment:** While your clinical recommendation for ID screening in this population is understandable, would it be more cautious to frame it as a hypothesis for future prospective research, given the observational design and high risk of bias?

**Reponse:** *Thank you for this valuable observation. We agree that, given the observational design and high risk of bias in the included studies, it is crucial to frame clinical recommendations with appropriate caution.*
*This point has already been explicitly addressed in the Conclusion section of our manuscript.*
*Lines 629-634, page 19: "However, the very high risk of bias across the included studies limits the generalizability of these findings. ... Future research should focus on patient-centered outcomes, employ rigorous methodology, and broad inclusion criteria to enhance the quality and generalizability of findings."*

**Comment:** You briefly mention that oral iron supplementation is often ineffective. Would it be appropriate to also emphasize that, to date, there is no solid evidence that treating ID in pediatric patients with heart disease reduces the risk of adverse outcomes?

**Response:** *Thank you for this valuable point. We agree that it's crucial to emphasize the current lack of evidence on the effectiveness of iron deficiency (ID) treatment in reducing adverse outcomes in this patient group.*
*As you noted, we have already addressed this in the Discussion section. On Page 17, lines 554-562, we highlight the lack of robust evidence for this specific patient population by stating:*
*"Iron supplementation is an established intervention for adults with HF, with intravenous iron shown to improve exercise capacity, quality of life, and potentially reduce hospitalizations for HF. (Graham et al., 2023) In pediatric patients, oral iron replacement is the first-line treatment for those presenting with ID, offering a relatively inexpensive approach to address this nutritional deficiency. However, a retrospective cohort study of 20 pediatric patients, who had iron levels measured before and after 3 months of iron supplementation, found that 55% of patients remained with ID despite of oral iron therapy. (Puri et al., 2022) Although this systematic review supports ID as a risk factor for ACV outcome in pediatric patients with heart disease, oral iron replacement may not necessarily reduce the risk of ACV events in this group of patients."*

I thank you again for your important contribution to pediatric cardiology and for raising awareness about the possible implications of iron deficiency. I hope that these observations will be helpful in strengthening your manuscript. I look forward to your responses or clarifications.

**Reviewer 4:**

Recommendation: Revisions Required

**Comments for authors:** Dear Authors, Congratulations on conducting a well-structured systematic review addressing an important and understudied topic. The manuscript is clearly written and presents interesting findings. However, minor revisions are necessary.

**Comment:** In the introduction, it's important to emphasize why pediatric-specific evidence is needed.
**Response:** *Thank you for this valuable comment. We agree that emphasizing the need for pediatric-specific evidence is a crucial part of the introduction.*
*As you suggested, we have revised the Introduction to highlight the distinct differences between pediatric and adult*

*populations. We have added a new paragraph that explains why adult-specific findings may not be directly applicable to children and adolescents.*
*Lines 116-123, page 3: "Although robust evidence exists in adult populations with HF, where iron deficiency is a well-established risk factor for adverse outcomes, it remains unclear whether these findings can be extrapolated to pediatric patients. Children and adolescents with heart disease differ from adults in terms of cardiac physiology, iron metabolism, growth and developmental stages, and responses to treatment. For pediatric patients, no previous systematic review has examined the association between ID and ACV outcomes, and international guidelines only consider treating ID if it is linked with anemia. (Kirk et al., 2014) This underscores the need for pediatric-specific evidence to guide clinical decision-making."*

> **Comment:** Address more explicitly how the variability in ID definition might have influenced the pooled OR (clarify if adjustments for confounding factors were made when calculating the OR, as this might overestimate the association between ID and adverse cardiovascular outcomes ).

**Response:** *Thank you for raising this important point. The observed pooled OR may be an overestimation due to the unadjusted or minimally adjusted nature of the included studies and the variability in ID definitions.*
*As you suggested, we have explicitly addressed these methodological limitations in the Discussion and Limitations sections of our manuscript to ensure a more cautious interpretation of the findings.*
*Lack of Adjustment for Confounders*
*We have compared our pooled OR to the smaller, adjusted hazard ratios (HRs) reported in large adult studies. As we've noted in the manuscript, we believe the difference in effect size is directly related to the lack of adjustment for key confounders in the pediatric studies, such as socioeconomic status, natriuretic peptides, and ventricular dysfunction.*
*On Page 16, Lines 485-488, we state: "The larger effect size in this review compared to previous studies in adults with HF may be related to the lack of adjustment for confounders, such as sex, race, socioeconomic status, and markers of disease severity like natriuretic peptides, ventricular dysfunction, and functional class."*
*Variability in ID Definitions*
*In the Limitations section, we highlight how the diverse definitions of ID across studies complicate patient classification and may introduce misclassification bias, which could distort the observed relationship between ID and ACV. This variability also limited our ability to perform meaningful subgroup or sensitivity analyses.*
*On Page 18, Lines 612-617, we state: "The variation in ID definitions across the studies complicates the classification of patients, limits our ability to conduct subgroup or sensitivity analyses, and may lead to misclassification bias potentially distorting the observed relationship between ID and ACV. Establishing a standardized definition of ID in pediatric patients is essential for improving evidence, supporting universal screening, and guiding future studies on iron replacement therapy in children with heart disease."*

> **Comment:** In the result section, mentioning that two of the studies could not be included in the meta-analysis is important, as this might introduce selection bias.

**Response:** *We thank the reviewer for this helpful suggestion. We've revised the Results section to clearly explain why two of the included studies were not part of the meta-analysis, and we've added a note about the potential for selection bias.*
*On page 12-13, lines 385-388:*
"*Of the five studies included, three reported data on adverse cardiovascular outcomes (Phillips, Luxford, and Puri). These were included in the meta-analysis. The remaining two studies (Higgins and Newland) did not report ACV outcomes and were therefore excluded from the meta-analysis; they were considered only in the prevalence analysis.*"