Peer Review File

Article information: https://dx.doi.org/10.21037/tp-23-548

Reviewer A

Comment 1: Overall, the study appears to be well-conducted and provides valuable insights into the impact of ventricular dilatation on neurodevelopmental and perinatal outcomes in fetuses diagnosed with isolated fetal ventriculomegaly. The objective of the study is clearly stated in the abstract, and it's relevant to the field of prenatal medicine. The inclusion and exclusion criteria are well-defined, which helps in ensuring the relevance and quality of the studies included in the review. The methods used, including the modified Newcastle-Ottawa Quality Assessment Scale for study quality assessment and the meta-analysis, seem appropriate for this type of research. The presentation of results is clear, with specific numbers and relative risks provided. The findings suggest a significant difference in adverse outcomes between mild and severe cases of isolated fetal ventriculomegaly, which is an important contribution to the literature. The conclusions are well-supported by the results and align with the study's objectives. However, it might be beneficial to briefly discuss the clinical implications of these findings. Overall, I recommend accepting this article for publication, with the suggestion to possibly include a brief discussion of the clinical implications of the study's findings in the final manuscript. This research can contribute valuable information to the medical community regarding the management and counseling of pregnant individuals with isolated fetal ventriculomegaly.

Reply 1: Thank you for your thoughtful review of our study. We have taken your feedback into consideration and have added further information on the clinical implications of our findings in the discussion. We have expanded further on how these results may be used to aid in counselling parents of infants with isolated fetal ventriculomegaly, as well as how they may be used for risk stratification by clinicians. We believe that these additions enhance the practical relevance of the findings of this paper and provides valuable insights for healthcare providers involved in the management of patients with isolated fetal ventriculomegaly. We also agree that this information will contribute to the broader understanding and application of our research in prenatal medicine.

Changes in the text: We have added further information on the clinical implications of this research (see Page 20, line 486-497).

Comment 2: Overall, this meta-analysis is well conducted, using appropriate methods, and addresses a frequent and important problem in prenatal diagnosis: the prognosis of ventriculomegaly according to whether it is mild or severe. My main concern is that the authors classify mild and moderate ventriculomegaly in the same category. In the usual practice and literature of prenatal diagnosis, we use 3 distinct thresholds. Mild ventriculomegaly between 10 and 12mm, moderate ventriculomegaly between 12 and 15mm and severe ventriculomegaly above 15mm. Merging the mild and moderate categories is a bold move and should be justified because the ratio of isolated or associated ventriculomegaly is considerably different between these two groups, as are their outcomes. The only justification presented by the authors is “studies have reported there to be little significant difference between mild and moderate

ventriculomegaly" (l.142-143), but that's not enough. For all the studies included, the authors specify in Table 1 the thresholds used and whether they report mild and moderate ventriculomegaly in two separate categories or in the same category. It would be really interesting to present a sub-analysis comparing the outcomes of these three categories (severe vs moderate vs mild) if there is sufficient cases.
I don't know if the following suggestion is feasible or not but it would be an added value. As there may be up to 30% genetic abnormalities detected by arrayCGH and /or WES in cases of severe ventriculomegaly, could the authors detail which included studies have results for arrayCGH and WES for their cases? And differentiate isolated cases with and without arrayCGH in the analyses? As well as cases with and without WES?

Specific comments:
It would be relevant to mention in the abstract the thresholds used to define mild and severe ventriculomegaly.
The introduction is well written, with a logical structure. In terms of incidences, it would be more useful to mention agenesis of the corpus callosum than dandy-walker in the e.g., or to add it if the authors wish to retain the latter (l.64-65). Chiari malformations are often a consequence of myelomeningoceles, so I would remove them, but that's a detail.
L 81-84: There may be discrepancies between the outcomes of mild and moderate ventriculomegaly, depending on the thresholds used, but certainly not between mild to moderate and severe ventriculomegaly, as the literature is quite extensive on the poor prognosis of the latter category (e.g. DOI:10.1002/uog.19038)
The methods, inclusion criteria and research strategies are generally well described and reproducible. I would just add the exact MeSH terms that were used either in the methods or as supplementary material. It's not enough to say that keyword variants were searched for; the authors have to be transparent about the search equations used.
The presentation of the results is clear and I have no additional comments to those presented above: if possible differentiate between mild and moderate ventriculomegaly, and if possible differentiate between cases with or without abnormalities on arrayCGH or WES.
For the discussion, once again I would separate the cases with mild or moderate ventriculomegaly: either the authors can provide this additional analysis and the results will be up for discussion with the current literature, or this should be recognised in the limitations.
The same applies to prenatal genetic investigations.

Reply 2: Thank you for your detailed feedback the classification of ventriculomegaly severity in our study. We appreciate your suggestions for sub-analysis comparing the outcomes of mild, moderate and severe ventriculomegaly. We acknowledge the conventional practice of using three distinct thresholds for ventriculomegaly severity: mild (10-12mm), moderate (12-15mm), and severe (above 15mm). Unfortunately, the studies that fit our inclusions criteria reported mild and moderate ventriculomegaly as a single entity. Thus, due to the way the studies were conducted and reported, we did not have a sufficient number of cases to conduct a meaningful sub-analysis comparing the outcomes for mild, moderate and severe ventriculomegaly; for this reason, sub-analyzing these categories with this restricted data would have limited statistical power and could lead to unreliable conclusions. While we understand the potential value of such an analysis, we believe that our study's primary aim remains valid and contributes to the existing literature in the field. For this reason and as suggested by reviewer B, we have added

this as a limitation in our discussion.

Furthermore, we thank you for your suggestion regarding the inclusion of genetic testing results such as arrayCGH and WES in our analysis. While we acknowledge the potential value of exploring the association between genetic abnormalities and isolated fetal ventriculomegaly outcomes, this aspect goes beyond the scope of our study. Our primary focus is to investigate the impact of isolated fetal ventriculomegaly on neurodevelopmental and perinatal outcomes, rather than specifically evaluating the genetic underpinnings of the condition. As such, we did not collect data on arrayCGH or WES results for the cases included in our analysis. However, we appreciate your suggestion and recognize the importance of genetic testing in the evaluation and management of fetal ventriculomegaly. Future studies specifically designed to investigate the relationship between genetic abnormalities and isolated fetal ventriculomegaly outcomes could provide valuable insights into this aspect of the condition.

Additionally, we have made various edits in response to further comments made by reviewer B. For example, in the abstract, we have included the thresholds used to define mild and severe ventriculomegaly for clarity. Moreover, we have revised the examples given in the introduction to mention agenesis of the corpus callosum and have removed chiari malformations, considering their association with myelomeningoceles. We also highly appreciate your attention to detail and the insights provided regarding the literature on the outcomes of severe ventriculomegaly and, in response to this, have made changes to this section of the text to include more accurate information on the topic.

We have also taken steps to enhance the transparency and reproducibility of our research methods, as suggested by reviewer B, by including the exact MeSH terms and key phrases used in our search strategy in the supplementary material. By providing this detailed information, we aim to facilitate a clearer understanding of our research process and ensure that our methodology is easily replicable by others in the field. We sincerely appreciate your commitment to promoting rigorous scientific practices.

Changes in the text: As requested, we have added the thresholds used to define mild and severe ventriculomegaly in the abstract (see Page 2, lines 39-40).
As suggested we have edited the introduction to include mention of agenesis of the corpus callosum as well as remove mention of chiari malformation (See Page 3, lines 63-64).
In response to the important clarification made by reviewer B regarding the extensive literature on the poor prognosis of severe ventriculomegaly, compared to mild, we have edited this section accordingly (see Page 4, lines 79-83).
As suggested, we have added the exact MeSH terms and key phrases/words used in the supplementary material (see Page 27-28, lines 665-680).
We have added additional limitations regarding the lack of sub-analysis for mild and moderate ventriculomegaly in the discussion (see Page 19, lines 468-474).

Comment 3: The manuscript titled 'Perinatal and Neurodevelopmental Outcomes of Fetal Isolated Ventriculomegaly: A systematic review and meta-analysis' by Ali et al discusses the outcomes associated with fetal isolated ventriculomegaly across 23 studies. The authors selected these studies based on specific criteria, although many of them fail to meet these

criteria. The authors aimed to include only studies published after the year 2000 due to technological advancements ('Data from studies prior to 2000 may be outdated, given the use of more advanced technology and techniques'). However, several studies within the meta-analysis, such as Graham (2001), Mercier (2001), Leitner (2004), Signorelli (2004), and Gaglioti (2005), followed populations dating back to the 1990s. Furthermore, two manuscripts, Weichert (2010) and Sethna (2011), encompassed mixed populations. Notably, certain studies included in the meta-analysis, such as those by Graham, Signorelli, and Gaglioti, did not employ fetal MRI, and the sonographic techniques utilized at that time were not comparable to modern neurosonography methods. It is very likely that non-isolated cases might have been included.

However, the most important issue with this meta-analysis is that the authors did not consider that severe ventriculomegaly can be secondary to hydrocephalus and therefore cannot be considered isolated. The prognosis can vary significantly, and these cases don't meet the definition of isolated ventriculomegaly. For instance, Litwinska's study included cases of progressive ventriculomegaly that underwent in utero shunting. This study investigates a different condition where not only the prognosis differs but also the fetal procedures applied alter the neurodevelopmental outcome. Similarly, Kennely's study, among the 19 isolated fetuses with ventriculomegaly, included two cases with cortical malformations (resulting in termination of pregnancy) and six cases that underwent cephalocentesis (indicating hydrocephalus, hence not isolated VM), leading to perinatal death. Moreover, only five of these cases underwent fetal MRI. The eight surviving infants all had additional malformations, none meeting the criteria for truly isolated ventriculomegaly. The same scenario is observed in the Ge et al manuscript.

Additionally, the authors included studies that sourced data from national registries, featuring only postnatal follow-ups. As previously mentioned, this provides minimal information for fetal medicine specialists engaged in fetal consultations (e.g., Hannon and Sethna articles). The Breeze study, focusing solely on neonatal outcomes, should be excluded from the analysis due to its short-term prognosis.

In summary, most studies within this meta-analysis are outdated, failing to meet current standards in neurosonography and fetal MRI. Importantly, many fetuses included in the analysis were not genuinely isolated, and some underwent risky intrauterine procedures.

Regarding lines 76-77: "In non-IVM, fetal outcomes are often related to the associated findings, which makes counseling parents somewhat easier." This sentence can be perceived as offensive to both parents and fetal specialists engaged in counseling. Fetal specialists should always acknowledge that fetal consultations must navigate the realm of uncertainty, requiring specific expertise and empathy.

Tab 3 Chu is misspelled as chun

Reply 3: Thank you for your insightful observations and comments. We acknowledge the discrepancy between our inclusion criteria and the time periods over which particular studies extracted data. While our aim was to focus on studies published after the year 2000 to account for technological advancements, we understand that certain studies, such as those by Graham, Mercier, Leitner, Signorelli, and Gaglioti, included populations dating back to the 1990s. We have addressed this limitation in our manuscript's discussion, highlighting the potential impact of this shortfall on the robustness of our findings. Despite these limitations, we made efforts to ensure consistency in our data extraction and analysis processes to minimize bias and enhance

the reliability of our results.

Furthermore, your guidance regarding the inclusion criteria for severe ventriculomegaly in our meta-analysis is highly appreciated. We recognise the importance of accurately categorizing cases to reflect their prognosis treatment accurately. In response to the concerns raised we have decided to refine our inclusion criteria to explicitly encompass cases of severe ventriculomegaly that progressed to hydrocephalus – these cases would fall under 'apparently' isolated severe ventriculomegaly. Likewise, it is worth noting that the two cases of severe ventriculomegaly with associated cortical malformations in Kennely's study were removed from our original analysis; hence why the number of cases from this particular study is 17, instead of 19, and we have added in table 4 to further clarify this detail. This adjustment in our criteria will enable us to more precisely analyse and report the outcomes and prognoses associated with these conditions, thereby acknowledging the spectrum of severity and potential progression observed in ventriculomegaly cases. By incorporating this clarification, we aim to ensure that our study accurately reflects the diverse clinical scenarios encountered in the management of fetal ventriculomegaly. We appreciate your thoughtful input and believe that this refinement will enhance the validity and clinical relevance of our findings.
In response to your concern, we acknowledge the inclusion of studies relying on national registries with only postnatal follow-up data, as noted in the Hannon and Sethna articles, as well as the Breeze study's focus on short-term neonatal outcomes. Our intention is to capture a broad spectrum of neurodevelopmental outcomes in our analysis, irrespective of the follow-up period. This approach allows us to provide a comprehensive overview of the available data on ventriculomegaly outcomes. We recognise the potential limitations that this approach may pose, particularly for fetal medicine specialists seeking more detailed prognostic information during fetal consultations. To address this, we will explicitly highlight this aspect in the limitations section of our study. By acknowledging these limitations, we aim to ensure that readers are informed about the scope and potential implications of our inclusion criteria on the applicability of our findings to fetal medicine practice. Through this acknowledgment, we strive to provide transparency regarding the limitations of our study while also emphasizing the need for cautious interpretation of our results in clinical practice.

We have also carefully reviewed the following sentence in our manuscript: "In non-IVM, fetal outcomes are often related to the associated findings, which makes counseling parents somewhat easier." We have taken your concerns into consideration and have removed part of this sentence to ensure that it does not convey any unintended offence to parents or fetal specialists engaged in counseling. We appreciate your insight and emphasize the importance of acknowledging the complexities and uncertainties involved in fetal consultations with empathy and expertise.

Thank you for bringing your concerns to our attention and we appreciate your thorough review of our work.

Changes in the text: We have added additional limitations in the discussion, based on concerns raised by reviewer C (see Page 18, lines 438-443).
We have altered the inclusion criteria to include severe ventriculomegaly that progressed to hydrocephalus (see Page 6, lines 143-145)

We have inputted additional limitations, with reference to the follow-up time in the assessment of neurodevelopmental outcomes (see Page 19, lines 451-457)
Deleted part of the sentence, to avoid unintended offence to parents or fetal specialists engaged in counseling (see Page 3, line 75)

Comment 4: A meta-analysis on this topic is important for prognosis and parent counseling.
A few questions and suggestions:
1) Why were studies where gestational age at diagnosis was not reported excluded? Was this because it was unclear if diagnoses was pre- or post- natal? If pre-natal, since GA at diagnosis wasn't a variable in the analysis, it shouldn't mater if specific GA at diagnosis was reported or not as long as it was pre-natally diagnosed.
2) Delay in milestones on formal testing and presence of neurological disorders seems like a very broad definition of neurodevelopmental delay. Especially since hearing impairment, seizure, and cerebral palsy is included ... how many studies would be excluded if you only looked at those with neurodevelopmental testing?
3) Please list in the table which tests were used by each study and whether neurologic disorders was included in definition of neurodevelopmental delay.
4) Meta-analysis was only performed for studies that included both mild and severe cases to compare differences (odds ratio) in outcome between mild and severe cases. You can report a weighed average across studies described in the "qualitative synthesis" sections.
5) It is concerning that the proportion of patients with neurodevelopmental delay differed so greatly between patients included in the qualitative and quantitative analyses (for mild, 11% qualitative vs 8.43% quantitative; for severe, 58% in qualitative vs 45% quantitative). Why do you think this was the case? What differed between the studies included in the qualitative vs quantitative analyses? Inclusion criteria? Age at assessment? Definition of "neurodevelopmental delay"?

Reply 4: Thank you for your intuitive comment regarding the exclusion of studies where gestational age at diagnosis was not reported. We agree with your input that the gestational age should not necessarily matter as long as the diagnosis was made prenatally. However, the issue we encountered was the lack of clarity regarding the exact time-point of diagnosis in many studies. To address this, we had to assume that studies not explicitly stating the gestational age at diagnosis and not specifying that the diagnosis was made prenatally had to be excluded. The absence of this information made it challenging to ensure consistency in our analysis and interpretation of results. We will expand on this aspect in the manuscript to provide further clarification regarding the rationale behind the exclusion of studies where gestational age at diagnosis was not reported. We appreciate your attention to detail and your contribution to improving the clarity of our methodology.

We appreciate your observation regarding the broad definition of neurodevelopmental delay used in our study. We agree that the definition encompasses a wide range of outcomes, including neurological disorders such as hearing impairment, seizures and cerebral palsy. The reason for this broad definition stems from the diversity of methods used to assess neurodevelopmental delay across the included studies. While a minority of studies utilised specific screening tests or scales, such as the Bayley scale, the majority relied on alternative methods. These methods included the diagnosis of specific neurodevelopmental conditions

based on ICD codes, telephone interviews with parents, or questionnaires regarding the achievement of neurodevelopmental milestones. As a result, employing a narrower definition based solely on neurodevelopmental testing would lead to the exclusion of a significant number of studies from our analysis. We acknowledge the potential limitations of this broad definition and will ensure to provide transparency regarding the various assessment methods used; we have provided a table, as suggested, to portray the various methods of neurological assessment employed in the included studies. Thank you for raising this important point.

We appreciate your suggestion regarding the analysis of outcomes between mild and severe cases. Our meta-analysis aimed to delineate differences in neurodevelopmental outcomes by comparing these two classifications. However, your comment raises an important point about presenting a weighted average across studies, particularly in the 'qualitative synthesis' section. Could you please specify the outcomes or variables for which you recommend calculating a weighted average? This clarification will help us refine our analysis and ensure that our findings are as informative and relevant as possible for our audience.

The difference in the proportion of patients with neurodevelopmental delay between the qualitative and quantitative analyses raises important considerations regarding potential discrepancies in study characteristics and methodologies. Several factors could contribute to this variation. Firstly, differences in inclusion criteria may have influenced the composition of studies included in each analysis. Studies included in the qualitative analysis may have varied in terms of patient populations, diagnostic criteria, and study designs, potentially leading to differences in observed outcomes. Additionally, variations in the age at assessment and the definition of "neurodevelopmental delay" across studies could have impacted the results. Studies may have assessed neurodevelopmental outcomes at different time points during infancy or childhood, leading to variability in the identification and reporting of delays. Moreover, the methods used to define and assess neurodevelopmental delay may have differed between studies included in the qualitative and quantitative analyses. Some studies may have employed more rigorous diagnostic criteria or utilized standardized assessment tools, while others may have relied on less specific or subjective measures. Overall, the discrepancies observed highlight the importance of carefully considering study characteristics and methodologies when interpreting research findings. Further exploration of these factors may provide valuable insights into the observed differences in neurodevelopmental outcomes between studies included in qualitative and quantitative analyses.

Changes in the text: We have clarified the exclusion criteria, with reference to excluding studies where there was a lack of clarity on the gestational age at diagnosis (see Page 6, Lines 136-137)
We have added an additional table, clarifying the neurological assessments employed by each included study in the supplementary material (see supplementary materials, Page 29)

Comment 5: This is a valuable work that sheds a bit more light on a prenatal finding that can be very difficult to explain to the parents. Clinical practice needs more papers like this. Good job.

Reply 5: Thank you very much for your positive feedback and encouraging words. We are

delighted to hear that you found our work valuable and that it contributes to shedding more light on the complexities of prenatal findings, particularly in difficult situations for parents. Your support motivates us to continue our efforts in producing research that is beneficial to clinical practice. We appreciate your recognition of our work and are grateful for your kind words.

Reviewer B

1. For the fourth affiliation, if “Paediatric Neurosciences” is a department, please provide it as “Department of Paediatric Neurosciences”.
Reply: The affiliation "Department of Paediatric Neurosciences" has been correctly updated.

2. Please provide a department for the 1st and 3rd affiliation if there is any.
Reply: Departments for the 1st and 3rd affiliations have been added as requested.

3. The corresponding author’s affiliation should also be listed in the affiliation list of all authors.
Reply: The corresponding author’s affiliation is now listed in the affiliation list of all authors.

4. The structure of an original article’s abstract should be “Background, Methods, Results, Conclusions”.
Reply: The abstract has been restructured to follow the "Background, Methods, Results, Conclusions" format.

5. Based on the journal guideline, please replace the subtitle, “Conclusions and Implications” with “Conclusions”.
Reply: The subtitle "Conclusions and Implications" has been replaced with "Conclusions".

6. Please define NICU upon first use in the Main Text.
Reply: The abbreviation NICU has been defined upon its first use in the Main Text.

7. “Pagani et al. reported a prevalence of 7.9% which is lower than our findings.”
A reference is needed in the above sentence.
Reply: A reference has been added to support the statement by Pagani et al.

8. Supplementary Material is not cited in the Main Text.
Reply: Clarified that there is no supplementary material, only an appendix.

9. Figure 2-4: Kumar 2018 should be Kumar 2020. Please confirm.
Reply: The citation for Kumar has been corrected to 2020.

10. Figure 2: The study of Doğan Durdağ 2019 is Ref 22 instead of Ref 23.
Reply: The reference for Doğan Durdağ 2019 has been correctly listed as Ref 22.

11. All the abbreviations in the figure(s) and table(s) should be defined in the explanatory legend.
Reply: All abbreviations in figures and tables have been defined in the legends.

12. Tables should be provided in editable format and the reference numbers of the studies are suggested to be added.
Reply: Tables have been provided in an editable format, and references have been added as suggested.

13. Table 2: There are only 19 papers listed. Please confirm.
Reply: Confirmed that Table 2 correctly lists 19 papers.

14. Table 5: There are only 13 studies listed, but 14 studies are reported in the legend.
Reply: revised.

15. Table 1/3/5: Please check the correctness of the published year for Ge et al. and be consistent with the bibliography.
Reply: We have checked the published year for Ge et al. and updated the tables appropriately.

16. Ref 37 was cited right after Ref 35 without Ref 36 cited in between in the Main Text. Please renumber the references to meet the consecutive standard.
Reply: We have edited references 36-39, both within the main text and bibliography, to meet the consecutive standard.