Reviewer’s report

Title: Progesterone in women with arrested premature labor. A report of a randomised clinical trial and updated meta-analysis.

Version: 1 Date: 07 Mar 2017

Reviewer: Ewoud Schuit

Reviewer’s report:

I appreciate that many of my concerns have been addressed.

I do have some minor comments regarding the responses of the authors.

#7. my comment: According to Kernan et al. JCE 1999 it is advised to adjust for covariates that is stratified for in the randomization. This means regression analyses should be performed instead of the Mann-Whitney U-test, student's t-test, and Fisher's exact test.

Authors' response: To our knowledge this is not universally applied in clinical trials. Additionally, we feel that more complex analysis makes it far more difficult for the average clinician to understand and therefore trust the results. Ultimately, we performed the analysis that we decided on apriori and to change this after the fact would not be appropriate. We appreciate the advice for future trials.

My response: although it may be true that clinician may prefer simple t-tests over regression analysis, I do not think this should be an argument in refraining from regression analysis. I'm fine with leaving the t-tests in there, but it would be reassuring to know that such regression analyses would results in the same treatment effect estimates. Can the authors perform these analyses and then mention in the results section that results were similar when assessing treatment effects using regression analysis with the treatment allocation and stratification variables as covariates in the regression model?

#10. My comments: Risk of bias was assessed "using standard criteria", but items, e.g. reporting of outcomes, seem to be missing from the criteria. A better risk of bias tool may be the one developed by the Cochrane Collaboration.

Authors' response: We prefer a simplified system which has been described in the literature and we have used in a previous meta-analysis2. It is not manifestly different from the Cochrane "risk of bias" tool. It is worth noting that the risk of bias assessment is largely ignored in Cochrane reviews as to our knowledge among the obstetric reviews there are none with subgroup analysis by study quality.

My response: I'd really prefer the authors to use Cochrane's risk of bias tool as this is the current standard for assessing risk of bias of RCTs for meta-analyses. I'm not sure what the authors mean by their statement that "the risk of bias assessment is largely ignored in Cochrane reviews as to
our knowledge among the obstetric reviews there are none with subgroup analysis by study quality”. Using Cochrane’s risk of bias tool is very common in systematic reviews and meta-analyses, including those published by the Cochrane Collaboration, regardless of specialty.

#12. My comments: The authors indicate that they will use both fixed and random effects models, but from the methods it is unclear what criteria are used for the choice on the type of model, e.g. based on the level of statistical heterogeneity above a certain threshold.

Authors' response: Using fixed effects if no heterogeneity is standard and we do not think requires explanation.

My response: "no heterogeneity" is still rather vague. Although the authors state that the choice of fixed vs. random does not require explanation the authors present a fixed effects M-H analysis in figure 2, while the $i^2 = 83\%$. Based on this heterogeneity, the analysis should have been performed using a random effects model. Please re-check all analyses on whether the choice of a fixed effects model was appropriate. I still think the authors should specifically state when a fixed (e.g. $i^2 < 50\%$) and when a random (e.g. $i^2 \geq 50\%$) effects model was used.

#20. My comment: Line 140-141. Was there statistical heterogeneity between the strata of the subgroups? This information should be added to the text. Authors' response: Information is provided in the figures 2-7. We feel this would be very redundant to add to text.

My response: I disagree. Such information would be very informative as it indicates whether the treatment effect differs significantly across subgroups. Presenting a (significant) treatment effect within a subgroup without presenting the heterogeneity of the effect across subgroups may lead readers to believe that this significant effect is different from the effect in other subgroup, while it in fact may not be different. Therefore I think this p-value of heterogeneity among subgroups should be added to the text.

#23. My comment: Line 164. Remove "adequate".

Authors' response: Word is not found in line 164.

My response: please see sentence: "Several groups have found that poor adequate allocation concealment and blinding influence results, usually, by increasing positive treatment effects25-27." In this sentence, adequate is redundant.

#24. My comment: Line 173. Low power is obviously an issue of the RCT, but indeed not of the meta-analysis.

Authors' response: This is a matter of opinion. Not sure what change the reviewer is suggesting.

My response: My apologies for not being more clear. What I hinted to was that the authors could distinguish between the RCT and meta-analysis in their assessment of power. Considerin that the RCT was stopped early, power of the RCT is obviously lower than planned. THis, however is
compensated in the meta-analysis by combining data from multiple studies. The authors could explicitly mention this in their discussion.

**Are the methods appropriate and well described?**
If not, please specify what is required in your comments to the authors.

Yes

**Does the work include the necessary controls?**
If not, please specify which controls are required in your comments to the authors.

Yes

**Are the conclusions drawn adequately supported by the data shown?**
If not, please explain in your comments to the authors.

Yes

**Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?**
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I am able to assess the statistics

**Quality of written English**
Please indicate the quality of language in the manuscript:

Acceptable

**Declaration of competing interests**
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?
5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal