PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	RANDOMISED PLACEBO CONTROLLED STUDY OF THE EFFECT OF PARACETAMOL ON ASTHMA SEVERITY IN ADULTS
AUTHORS	Ioannides, Sally; Williams, Mathew; Jefferies, Sarah; Perrin, Kyle; Weatherall, Mark; Siebers, Robert; Crane, Julian; Patel, Mitesh; Travers, Justin; Shirtcliffe, Philippa; Beasley, Richard

VERSION 1 - REVIEW

REVIEWER	Seif Shaheen Queen Mary University of London, UK
REVIEW RETURNED	09-Nov-2013

GENERAL COMMENTS	The authors have carried out an RCT to determine whether regular paracetamol use makes adult asthma worse. This is an important clinical question to address given the epidemiological evidence for a link. Unfortunately the study has some limitations and was unable to reach definitive conclusions, mainly because of lack of statistical power. Main comments: 1) It might be helpful in the abstract to clarify what the difference in doubling dose for BHR means ie the direction of effect for a naïve reader. 2) I'm interested to know why the authors chose BHR as the primary outcome measure, and not ACQ or symptom severity? Was a symptom diary completed and bronchodilator use recorded? 3) Statistical methods: first sentence of second paragraph doesn't make sense to me. 4) Table 2: the difference in outcomes was 'adjusted for baseline' – do the authors mean that the difference in change (final measure minus baseline) between the two arms was calculated (this is what I would expect to have been done)? 5) The authors say that 'intention to treat' analyses were carried out but there were some drop-outs (a greater proportion in the paracetamol arm) and there is no mention of how this was addressed in the analyses. Were missing values imputed (to preserve the randomisation) or was the analysis actually restricted to those who had final outcomes measured (in which case the benefits of randomisation may have been jeopardised)? This is relevant to the different numbers cited at baseline versus 12 weeks in Table 2. 6) A limitation of the study design was the imbalance in numbers between the two treatment arms which arose following randomisation because further subjects were then found to be ineligible. I don't understand why randomisation (ideally blocked) could not have been done following assessment at the second visit? 7) One potential explanation for the low proportion of subjects in the paracetamol arm having detectable blood paracetamol levels is poor compliance. This should be acknowledged, as this was the whole point of doing these measurements. Are more sensitive assays

	available? 8) Were the participants asked to avoid taking other forms of paracetamol (eg over the counter remedies containing the drug) for the duration of the study? 9) Why do the authors think that cytokine concentrations were not measurable in a high proportion of subjects? Was there a problem with the assays? Ideally these outcomes should have been treated as continuous not binary variables. 10) The authors propose a possible future cross-over study in the discussion, although there may be uncertainties about the duration of the wash-out period. The authors are probably aware that a parallel design trial of paracetamol versus ibuprofen is underway in preschool children with asthma in the USA, although there is no placebo arm.

REVIEWER	Tricia McKeever University of Nottingham
REVIEW RETURNED	18-Nov-2013

GENERAL COMMENTS	The paper is well written and presented. THe study itself is underpowered , however this is clearly recognised as a limitation in the researched paper. The first statement of the discussion could be more clearly stated- as paracetamol would not be expected to reduce PC20 if the hypothesis is true.

REVIEWER	Dr Alemayehu Amberbir London School of Hygiene and Tropical Medicine, UK Malawi Epidemiology and Interventions Research Unit, Malawi
REVIEW RETURNED	05-Dec-2013

GENERAL COMMENTS	This is a well-performed randomized placebo controlled trial of the effect of paracetamol on asthma severity among adults. The authors have provided a balanced discussion of the findings. This is a much needed study concerning the paracetamol enigma in asthma. The findings, although negative, provide important insight on which the design of future studies can be based. However, there are a few issues that need be clarified: Major: 1. I don't fully understand the requirement to randomize subjects prior to their final eligibility screening visit which has clearly resulted in loss of power as well as unequal allocation of treatment and placebo arm. The reason given by the authors, in my view, is less compelling, and therefore implication to the findings including any source of bias should be discussed. 2. Medication compliance in the study was assessed using pill count which is prone to overestimation of compliance particularly on repeat visits as the participant aware that a pill count is going to be conducted. Moreover, serum paracetamol levels were detected in around 38% of the paracetamol group (& none in the placebo). The authors have tried to explain this in their discussions but not fully convincing. Does this partly explain some of the non-significant findings in the study? 3. Most of the studies to date (except few) have shown significant

	adverse effect of paracetamol on asthma in infants or children suggesting that paracetamol may be involved in the pathogenesis of asthma. It would have been great to see the findings of the cytokine measurements in order to understand this mechanism. Is there data on glutathione peroxidise and glutathione transferase status? 4. From the investigators point of view, what are the next questions and design options regarding the paracetamol and asthma hypothesis – similar RCTs in adults with stable asthma of larger size or well powered RCTs during pregnancy or infancy and later childhood? A balanced discussion of these in line with current findings would enlighten future studies. Minor: 1. Table 1: characteristics of participants enrolled in the study – is there any difference e.g. sex and some of the skin test results? A footnote describing this would be useful. 2. Did you record information on use of other analgesics particularly ibuprofen and/or aspirin through the 12 weeks period?

VERSION 1 – AUTHOR RESPONSE

Reviewer - Seif Shaheen

The authors have carried out an RCT to determine whether regular paracetamol use makes adult asthma worse. This is an important clinical question to address given the epidemiological evidence for a link. Unfortunately the study has some limitations and was unable to reach definitive conclusions, mainly because of lack of statistical power.
Main comments:
1) It might be helpful in the abstract to clarify what the difference in doubling dose for BHR means ie the direction of effect for a naïve reader.

Thank you, we have now provided this information within the abstract.

2) I'm interested to know why the authors chose BHR as the primary outcome measure, and not ACQ or symptom severity?

Thank you, we have now provided greater justification in the choice of BHR as an objective well standardised physiological measure of asthma severity, recommended for monitoring the effects of therapy which may modify asthma severity, with appropriate references.

Was a symptom diary completed and bronchodilator use recorded?

To avoid an excessive number of secondary outcome variables and to limit the requirement for the participations during the 12 week study period, diaries were limited to daily medication use (to encourage compliance) and morning and evening peak flow in the final intervention week. Symptoms were assessed by measurement of ACQ5 at 6 weeks and 12 weeks.

3) Statistical methods: first sentence of second paragraph doesn't make sense to me.
This paragraph has been modified as follows:
"The proportion of participants with at least one asthma exacerbation was compared as an absolute risk difference, with an appropriate confidence interval, because in the event there were no asthma exacerbations in one of the randomised groups so that a relative risk could not be calculated. Simple

t-tests were used to compare mean values for ALT by randomised group. FeNO, eosinophil count and IgE were logarithm transformed because of skewed distributions, and the difference in logarithms was compared by a t-test. For those three variables with a logarithm transformation, the exponent of the difference in logarithms is interpreted as the ratio of mean values."

4) Table 2: the difference in outcomes was 'adjusted for baseline' – do the authors mean that the difference in change (final measure minus baseline) between the two arms was calculated (this is what I would expect to have been done)?

ANCOVA was used with the baseline value as a co-variate. This method of analysis will always produce an estimate as least as precise as the reviewer's suggestion of analysing the change from baseline, but will usually give estimates that are more precise than that.

5) The authors say that 'intention to treat' analyses were carried out but there were some drop-outs (a greater proportion in the paracetamol arm) and there is no mention of how this was addressed in the analyses. Were missing values imputed (to preserve the randomisation) or was the analysis actually restricted to those who had final outcomes measured (in which case the benefits of randomisation may have been jeopardised)? This is relevant to the different numbers cited at baseline versus 12 weeks in Table 2.

Please see response to editorial comments.

Those participants who were dispensed randomised medication were analysed. There was insufficient association between observed and unobserved variables to carry out multiple imputation in this study.

6) A limitation of the study design was the imbalance in numbers between the two treatment arms which arose following randomisation because further subjects were then found to be ineligible. I don't understand why randomisation (ideally blocked) could not have been done following assessment at the second visit?

Likewise, we wanted to be able to randomise only participants whose eligibility was confirmed following assessment at the second visit, however the Wellington Hospital Pharmacy was unable to prepare randomised medication to deliver to the MRINZ research offices in a timely manner. In retrospect we should have sent the randomised treatments by courier to the participants the day after Visit 2, however at the time we thought it important to dispense the randomised medication face-to-face at Visit 2.

7) One potential explanation for the low proportion of subjects in the paracetamol arm having detectable blood paracetamol levels is poor compliance. This should be acknowledged, as this was the whole point of doing these measurements. Are more sensitive assays available?

The potential for paracetamol levels that were below the detection cut-off to reflect low compliance is now acknowledged in the discussion.

8) Were the participants asked to avoid taking other forms of paracetamol (eg over the counter remedies containing the drug) for the duration of the study?

Thank you for the opportunity to provide further details regarding this issue. All participants were instructed to avoid taking other forms of paracetamol (including over-the-counter remedies containing paracetamol) or NSAIDs for the duration of the study. All participants were provided with a prescription for codeine to use as an analgesic during the study. This is now stated in the methods.

9) Why do the authors think that cytokine concentrations were not measurable in a high proportion of subjects? Was there a problem with the assays? Ideally these outcomes should have been treated as continuous not binary variables.

According to the package inserts, all serum samples of TNF-α, IL-4, IL-5 and IL-13 from healthy volunteers measure below the lowest standard. As a result our findings are not unexpected as our subjects had mild asthma. The proportion with undetectable levels was so high that rank based tests didn't produce meaningful results.

10) The authors propose a possible future cross-over study in the discussion, although there may be uncertainties about the duration of the wash-out period. The authors are probably aware that a parallel design trial of paracetamol versus ibuprofen is underway in preschool children with asthma in the USA, although there is no placebo arm.

Thank you, this issue is now raised in the discussion in which it is now stated "Important issues with the design of such a study are the duration of both the treatment periods and the crossover period". We have also included the following comment regarding the important requirement for a placebo arm: "It would be important if possible to include a placebo rather than ibuprofen arm, as NSAIDs may have the potential to both cause NSAID-induced bronchospasm, as well as reducing asthma severity with long term use.[29]"

Reviewer - Tricia McKeever

The paper is well written and presented. THe study itself is underpowered , however this is clearly recognised as a limitation in the researched paper. The first statement of the discussion could be more clearly stated- as paracetamol would not be expected to reduce PC20 if the hypothesis is true.

The hypothesis was paracetamol would increase BHR (i.e. reduce PC20). To avoid confusion we have now referred to BHR, not PC20.

Reviewer - Dr Alemayehu Amberbir

This is a well-performed randomized placebo controlled trial of the effect of paracetamol on asthma severity among adults. The authors have provided a balanced discussion of the findings. This is a much needed study concerning the paracetamol enigma in asthma. The findings, although negative, provide important insight on which the design of future studies can be based. However, there are a few issues that need be clarified:

Major:
1. I don't fully understand the requirement to randomize subjects prior to their final eligibility screening visit which has clearly resulted in loss of power as well as unequal allocation of treatment and placebo arm. The reason given by the authors, in my view, is less compelling, and therefore implication to the findings including any source of bias should be discussed.

Please see response to editor and Reviewer 1, point 5.

2. Medication compliance in the study was assessed using pill count which is prone to overestimation of compliance particularly on repeat visits as the participant aware that a pill count is going to be conducted. Moreover, serum paracetamol levels were detected in around 38% of the paracetamol

group (& none in the placebo). The authors have tried to explain this in their discussions but not fully convincing. Does this partly explain some of the non-significant findings in the study?

Please see response to Reviewer 1, point 7.

3. Most of the studies to date (except few) have shown significant adverse effect of paracetamol on asthma in infants or children suggesting that paracetamol may be involved in the pathogenesis of asthma. It would have been great to see the findings of the cytokine measurements in order to understand this mechanism. Is there data on glutathione peroxidise and glutathione transferase status?

There is no data on glutathione peroxidase and glutathione transferase status.

4. From the investigators point of view, what are the next questions and design options regarding the paracetamol and asthma hypothesis – similar RCTs in adults with stable asthma of larger size or well powered RCTs during pregnancy or infancy and later childhood? A balanced discussion of these in line with current findings would enlighten future studies.

Thank you, we have now extended the discussion in the second to last paragraph of the discussion to propose randomised placebo-controlled trials of paracetamol use in pregnancy, infancy and later childhood and the risk of developing asthma.

Minor:
1. Table 1: characteristics of participants enrolled in the study – is there any difference e.g. sex and some of the skin test results? A footnote describing this would be useful.

We do not believe it is useful to do formal statistical tests of the distribution of baseline variables, as it inflates the Type I error rate for the study as a whole and it is uncertain for each variable what a scientifically meaningful difference is (and hence lack of statistical power to detect this). It is also unclear what hypothesis the variable by variable tests are actually testing (for example randomization), and finally if there are variables of interest a better technique is to pre-specify the clinically important ones in the analysis. We feel that this is unlikely to be useful this far into the submission process now.

2. Did you record information on use of other analgesics particularly ibuprofen and/or aspirin through the 12 weeks period?

Please see response to Reviewer 1, point 8.

VERSION 2 – REVIEW

REVIEWER	Dr Alemayehu Amberbir London School of Hygiene and Tropical Medicine, UK Malawi Epidemiology and Interventions Research Unit, Malawi
REVIEW RETURNED	22-Dec-2013

GENERAL COMMENTS	I have read the revised version with interest, and have no further comments to add. Thank you for inviting me to review this manuscript.