PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Trial of feasibility and acceptability of routine low dose aspirin versus Early Screening Test indicated aspirin for preeclampsia prevention [TEST Study] – A multicenter randomised controlled trial
AUTHORS	Mone, Fionnuala; Mulcahy, Cecilia; McParland, Peter; Brethnach, FIONNUALA; Downey, Paul; McCormack, Dorothy; Culliton, Marie; Staunton, Alice; Cody, Fiona; Morrison, John; Daly, Sean; Higgins, John; Cotter, Amanda; Hunter, Alyson; Tully, Elizabeth; Dicker, Patrick; Alfirevic, Zarko; Malone, Fergal; McAuliffe, Fionnuala

VERSION 1 – REVIEW

REVIEWER	Prof. Jean-Christophe Gris, M.D., Ph.D. Department of Haematology, University Hospital, Nîmes and University of Montpellier, France
REVIEW RETURNED	20-Feb-2018

GENERAL COMMENTS	This is an interesting, astute and well driven study, for which the authors have to be congratulated. I have no significant reservations about the methodology and the course of the study, and about the interpretation of the data. My only comments are: 1- Routine low dose aspirin at a daily dose of 75 mg was tested. This is what is currently recommended in the United Kingdom in such a setting. However, a dose-response effect has been recently unmasked (Roberge S et al., Am J Obstet Gynecol 2017; 216(2): 110-20) which favor daily doses of aspirin greater than 100 mg, and perhaps 150 mg. To my knowledge, there is no demonstration that low-dose aspirin prescribed between 100 and 150 mg daily are associated with more safety issues than given at 75 mg daily. The chosen aspirin dose should thus be not the best therapeutic bet. An extensive comment on this point could enrich the Discussion section. 2- Willebrand factor circulating activities are highly dependent on blood groups, patients belonging to group O having significantly lower values. Low Willebrand activities may potentiate the low-dose aspirin-induced haemorrhagic risk, even if Willebrand factor increases during pregnancy. The distribution of blood groups among the 3 groups of patients may add a potentially interesting information in the table describing the patients' characteristics. 3- Aspirin adherence was also assessed via assessment of change in urinary TxB2. An exploratory analysis investigating the links between urinary TXB2 concentrations, changes in TXB2 concentrations and the secondary outcomes but also the safety outcomes would add some very interesting information to the understanding of the results.

REVIEWER	Prof Judy Simpson University of Sydney, Australia
REVIEW RETURNED	14-Mar-2018

GENERAL COMMENTS	This is a very confused and confusing paper, which violates its own study protocol. In the title it is correctly described as a pilot RCT, but the word 'pilot' does not appear at all in the text of the paper. The aims of the pilot study are clearly and properly set out in both the paper and the attached protocol, being to assess feasibility and acceptability and to obtain estimates of rates of preeclampsia and SGA. But the statistical analysis is completely inappropriate. Section 14 of the protocol clearly states, appropriately, that "Only simple descriptive statistics will be used to present the data" and that "No formal hypothesis [test] will be performed and no interim analyses are planned" and again "hypothesis testing is not a part of this pilot study and is reserved for the main study". And yet the paper is full of inappropriate hypothesis tests and an excruciating number of P-values, many testing differences in very small samples, and in Table 3 even testing both the proportion of male babies and the proportion of female babies. It is not clear whether the statistician who analysed the data is the same person who helped write the protocol, or whether s/he is one of the authors, but I strongly recommend that a statistician who can follow the protocol should be invited to be a co-author. They would also know that it is not appropriate, and violates CONSORT recommendations, to test differences at baseline – see p11 "There were no significant differences between groups at baseline [Table 1]". Furthermore, the paper is full of errors, starting in the Abstract where the number with vaginal spotting in the non-aspirin group is given as 143 instead of 28, the OR is shown as 2.6 instead of 2.1, and the final sentence of the Results gives details of hypothesis tests for preeclampsia and SGA. The 3 groups are referred to in the Abstract as (i) to (iii), but later as 1, 2, 3A and 3B. It is unclear whether the second scheduled study visit at 20-22 weeks was time to coincide with one of routine care visits, but these were at 18-20 and 25 weeks, so it appears that additional burden was imposed on the women for the study. Table 1 is useful, but the other tables are all unnecessary for the pilot study report.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Reviewer 1 comment 1; Routine low dose aspirin at a daily dose of 75 mg was tested. This is what is currently recommended in the United Kingdom in such a setting. However, a dose-response effect has been recently unmasked (Roberge S et al., Am J Obstet Gynecol 2017; 216(2): 110-20) which favor daily doses of aspirin greater than 100 mg, and perhaps 150 mg. To my knowledge, there is no demonstration that low-dose aspirin prescribed between 100 and 150 mg daily are associated with more safety issues than given at 75 mg daily. The chosen aspirin dose should thus be not the best therapeutic bet. An extensive comment on this point could enrich the Discussion section.

Response to reviewer 1 comment 1; Agree. Addition to discussion section; A recent meta-analysis, published since completion of this study suggests that there is a dose-response effect, with higher doses of aspirin commenced prior to 16-weeks gestation, associated with a significantly greater reduction in pre-eclampsia and fetal growth restriction compared to standard lower doses.

Reference 26. Roberge S, Nicolaides K, Demers S, Hyett J, Chaillet N, Bujold E. The role of aspirin dose on the prevention of preeclampsia and fetal growth restriction: systematic review and meta-analysis. Am J Obstet Gynecol 2017;216(2):110-20

Reviewer 2 comment 2; Willebrand factor circulating activities are highly dependent on blood groups, patients belonging to group O having significantly lower values. Low Willebrand activities may potentiate the low-dose aspirin-induced haemorrhagic risk, even if Willebrand factor increases during pregnancy. The distribution of blood groups among the 3 groups of patients may add a potentially interesting information in the table describing the patients' characteristics.

Review to reviewer 1 comment 2; Agree this is very interesting, however we do not have access to this data unfortunately

Reviewer 1 Comment 3; Aspirin adherence was also assessed via assessment of change in urinary TxB2. An exploratory analysis investigating the links between urinary TXB2 concentrations, changes in TXB2 concentrations and the secondary outcomes but also the safety outcomes would add some very interesting information to the understanding of the results.

Response to reviewer 1 comment 3; Agree. TxB2 levels are associated with preeclampsia. For the purposes of this study, to assess adherence, TxB2 was only assessed in aspirin taking subjects and not those not taking aspirin, hence we have a small cohort of subjects for which results are available.

Reviewer: 2

Reviewer 2 comment 1; This is a very confused and confusing paper, which violates its own study protocol. In the title it is correctly described as a pilot RCT, but the word 'pilot' does not appear at all in the text of the paper. The aims of the pilot study are clearly and properly set out in both the paper and the attached protocol, being to assess feasibility and acceptability and to obtain estimates of rates of preeclampsia and SGA.

Response to reviewer 2 comment 1; Thankyou for reviewing the manuscript and providing us with an opportunity to improve it. The title and the abstract methods have been amended to reflect that this is a feasibility and acceptability trial.

Reviewer 2 comment 2; But the statistical analysis is completely inappropriate. Section 14 of the protocol clearly states, appropriately, that "Only simple descriptive statistics will be used to present the data" and that "No formal hypothesis [test] will be performed and no interim analyses are planned" and again "hypothesis testing is not a part of this pilot study and is reserved for the main study". And yet the paper is full of inappropriate hypothesis tests and an excruciating number of P-values, many testing differences in very small samples, and in Table 3 even testing both the proportion of male babies and the proportion of female babies.

Response to reviewer 2 comment2; For feasibility trials CONSORT advise 'Results of any other analyses performed that could be used to inform the future definitive trial' http://www.consort-statement.org/extensions/overview/pilotandfeasibility. If one was to proceed with a larger appropriately powered study to determine the clinical efficacy and safety of routine low dose aspirin

use versus screening indicated aspirin, knowledge of the proportion of low-risk women that developed pre-eclampsia and associated confidence intervals in the absence of an accepted statistic in the literature may serve useful for a power calculation. The trial protocol also points toward assessing differences in secondary outcomes between groups. The secondary outcomes table has now been removed from the primary manuscript and added as a supplementary table. Gender of babies delivered has been removed from the table.

Reviewer 2 comment 3; It is not clear whether the statistician who analysed the data is the same person who helped write the protocol, or whether s/he is one of the authors, but I strongly recommend that a statistician who can follow the protocol should be invited to be a co-author. They would also know that it is not appropriate, and violates CONSORT recommendations, to test differences at baseline – see p11 “There were no significant differences between groups at baseline [Table 1]”.

The trial statistician is Dr Patrick Dicker who is a co-author on the paper and both co-wrote the study protocol and this manuscript. The term ‘there were no significant differences at baseline’ has now been removed thankyou for highlighting this.

Reviewer 2 comment 4; Furthermore, the paper is full of errors, starting in the Abstract where the number with vaginal spotting in the non-aspirin group is given as 143 instead of 28, the OR is shown as 2.6 instead of 2.1, and the final sentence of the Results gives details of hypothesis tests for preeclampsia and SGA. The 3 groups are referred to in the Abstract as (i) to (iii), but later as 1, 2, 3A and 3B.

Response to reviewer 2 comment 4; Thankyou, errors have been corrected and preeclampsia and SGA data removed from the abstract. Three group naming/labels clarified in abstract.

Reviewer 2 comment 5; It is unclear whether the second scheduled study visit at 20-22 weeks was time to coincide with one of routine care visits, but these were at 18-20 and 25 weeks, so it appears that additional burden was imposed on the women for the study.

Response to reviewer 2 comment 5; Women attended at 20-22 weeks [the time in Ireland when they attend for their structural fetal anatomy scan] and underwent this scan by the study sonographer, which was documented separately in the clinical notes, also relieving some of the clinical burden from the sonography department. Addition to Methods section to clarify this; Participants underwent two scheduled study visits, at study recruitment and at 20-22 weeks (to coincide with their fetal anatomy scan which was performed at the same time)

Reviewer 2 comment 6; Table 1 is useful, but the other tables are all unnecessary for the pilot study report.

Response to reviewer 2 comment 6; Table 3 is now a supplementary table only

D. FORMATTING AMENDMENTS (if any)

Comment D1.Figure Resolution'
- Please provide another copy of your figures with better qualities and please ensure that Figures are of better quality or not pix-elated when zoom in. NOTE: They can be in TIFF or JPG format and make sure that they have a resolution of at least 300 dpi. Figures in PDF, DOCUMENT, EXCEL and POWER POINT format are not acceptable.

Response D1; Figure resolution amended to 300dpi. Note updated version of Figure 1 content

2.Supplementary File Format
- Please re-upload your supplementary files in PDF format.

Response D2; Changed to PDF and uploaded

VERSION 2 – REVIEW

REVIEWER	Jean-Christophe Gris, M.D., Ph.D. Department of Haematology, University Hospital, Nîmes, and Universty of Montpellier, France
REVIEW RETURNED	03-Apr-2018

GENERAL COMMENTS	Congratulations for this original study, the results of which, within the limits of the chosen therapeutic option, open interesting perspectives of a dedicated RCT. In this future trial, all conventional determinants of individual bleeding risk should be carefully integrated as exploratory covariates of the observes bleedings. Similarly, the biological response to low-dose aspirin should be analyzed in its ability to predict clinical events.

REVIEWER	Prof Judy Simpson University of Sydney, Australia
REVIEW RETURNED	18-Apr-2018

GENERAL COMMENTS	The statistical analysis of this revised paper remains inappropriate. In their response to my previous comment (labelled 2 by them) the authors have quoted selectively from the CONSORT extension to randomised pilot and feasibility trials. In this 2016 BMJ paper, Eldridge et al state in Box 1 “Formal hypothesis testing for effectiveness (or efficacy) is not recommended. The aim of a pilot trial is not to assess effectiveness (or efficacy) and it will usually be underpowered to do this.” Despite explicitly stating in Section 15 of the revised protocol, appropriately, that “hypothesis testing is not a part of this pilot study and is reserved for the main study”, the authors of this revised paper have made no change to the statistical analysis except to move one table to the supplementary material. So the paper remains full of underpowered hypothesis tests and littered with inappropriate P-values. And they have not removed from the Conclusion the statement about not detecting a difference in rates of preeclampsia. The authors have argued that the CONSORT feasibility trials extension advises that “Results of any other analyses performed that could be used to inform the future definitive trial” may be reported, but in the CONSORT paper the example given for Item 18 is that of a sensitivity analysis which examined odds ratio estimates in subgroups and concluded that a certain subgroup should be included in the full RCT. Thus Item 18 is shown not to be advocating against the major recommendation not to do formal hypothesis testing.

VERSION 2 – AUTHOR RESPONSE

Editorial Comment 1: Editorial comment 1 response: Like reviewer 2, we do not feel that the use of p-values is appropriate, given that the study was not powered to detect any differences. Your response regarding the CONSORT statement “Results of any other analyses performed that could be used to inform the future definitive trial” does not mean that you should provide p-values, as p-values do not inform future trials – this statement implies that effect estimates and their confidence intervals should be reported, in order to inform sample size estimation for future trials. Therefore, you would need to state that the analysis focuses on confidence interval estimation rather than hypothesis testing because the pilot study has not been powered to detect significant differences between groups. You should also remove all p-values. If possible we recommend that you consult a statistician before submitting the next revision.

Editorial Comment 1 Response: Thankyou for your coment and statistical review. Our study statistician (Dr Patrick Dicker) has reviewed the manuscript and all p-values have been removed and replaced by confidence intervals where appropriate.

Reviewer 1 Comment 1: Congratulations for this original study, the results of which, within the limits of the chosen therapeutic option, open interesting perspectives of a dedicated RCT. In this future trial, all conventional determinants of individual bleeding risk should be carefully integrated as exploratory covariates of the observes bleedings. Similarly, the biological response to low-dose aspirin should be analyzed in its ability to predict clinical events.

Reviewer 1 Comment 1 Response: Thankyou for your comment, we will aspire to incorporate these suggestions in such a future study

Reviewer 2 Comment 1: The statistical analysis of this revised paper remains inappropriate. In their response to my previous comment (labelled 2 by them) the authors have quoted selectively from the CONSORT extension to randomised pilot and feasibility trials. In this 2016 BMJ paper, Eldridge et al state in Box 1 “Formal hypothesis testing for effectiveness (or efficacy) is not recommended. The aim of a pilot trial is not to assess effectiveness (or efficacy) and it will usually be underpowered to do this.” Despite explicitly stating in Section 15 of the revised protocol, appropriately, that “hypothesis testing is not a part of this pilot study and is reserved for the main study”, the authors of this revised paper have made no change to the statistical analysis except to move one table to the supplementary material. So the paper remains full of underpowered hypothesis tests and littered with inappropriate P-values. And they have not removed from the Conclusion the statement about not detecting a difference in rates of preeclampsia.

Reviewer 2 comment 1 response: Many thanks, we agree entirely and have removed all p-values and replaced them with confidence intervals where appropriate. This statement has been removed from the conclusion section

Reviewer 2 comment 2: The authors have argued that the CONSORT feasibility trials extension advises that “Results of any other analyses performed that could be used to inform the future definitive trial” may be reported, but in the CONSORT paper the example given for Item 18 is that of a sensitivity analysis which examined odds ratio estimates in subgroups and concluded that a certain subgroup should be included in the full RCT. Thus Item 18 is shown not to be advocating against the major recommendation not to do formal hypothesis testing.

Reviewer 2 comment 2 response: Agree, this has clearly been misinterpreted by ourselves. As stated above formal hypothesis testing has been omitted.