Author’s response to reviews

Title: First trimester ultrasound measurements and maternal serum biomarkers as prognostic factors in monochorionic twins: a cohort study

Authors:

Fiona Mackie (fionamackie@doctors.org.uk)
Rebecca WHITTLE (r.l.whittle@keele.ac.uk)
R. Katie Morris (r.k.morris@bham.ac.uk)
Jon Hyett (jon@fetalmedicine.org)
Richard Riley (r.riley@keele.ac.uk)
Mark Kilby (m.d.kilby@bham.ac.uk)

Version: 1 Date: 01 Mar 2019

Author’s response to reviews:

Reviewer #1: This manuscript is well written and straightforward.

Major comments:

1. The authors do not appear to adjust for overfit when presenting the c-statistics or increase in c-statistics corresponding to multivariable models resultingly, these estimates are likely overoptimistic. Could the authors update their analysis to correct for this optimism?

Thank you very much for raising this point We recognise the reviewers point, and agree that if we were aiming for the multivariable model in our paper to be used as a prediction tool, then adjustment for overfitting would be important. However, as the aim of this study was to assess individual prognostic factors, as opposed to creating a prognostic model, we have not corrected for optimism in the C statistics reported. To address the important point by the reviewer, in the Methods we now refer to ‘apparent’ C-statistics and clarify the issue of overfitting as follows:

“To gauge the potential increase in discrimination performance of a prognostic model that includes each potential prognostic factor in addition to existing factors, the change in apparent c-statistic (increase in area under the curve) for each outcome was calculated (i.e. difference in apparent c-statistic for models with standard characteristics including or excluding each factor). No adjustment for potential model overfitting was made during the calculation, as this was only for illustration of potential impact of including the factors.”
2. The authors could consider removing the univariable results as they test the adjusted associations regardless of the univariable results and that they are investigating the added prognostic value of these factors in addition to standard predictors.

Thank you, but as this is the first time these factors have been explored in this context, we would prefer to keep the univariable results in for completeness. We agree that the adjusted results are more important, and the paper focuses mainly on these.

3. A decision curve could add value here in order to compare the added prognostic benefit for the different types of factors.

Thank you for the suggestion, which we considered at length. Ultimately, as we are not saying to use our findings for clinical decision making based on predicted risks from the multivariable model, we think it may be misleading to also include a net benefit examination of the model with and without the different factors. However, for illustration we do present the predicted absolute risks from the model for hypothetical patients, and one can clearly see that there are a wide range of predicted risks across the distribution of factor values, even when holding fixed standard characteristics at their mean values. Thus, this implicitly suggests that clinical decisions (based on a threshold of risk) may indeed be changed if the factor(s) is (are) included, as the risk predictions may change importantly. But further research is needed to develop (in a larger dataset) an actual prediction model, followed by internal validation, and then potential net benefit evaluation and external validation.

Minor comments:

1. The overall column in table 1 is redundant. Perhaps the authors instead could include a column of comparative tests between no complication vs at least one complication per twin pregnancies.

Thank you for your suggestion, we have now removed this column, but haven’t included comparative tests between the 2 groups as the aim of the table was to describe the characteristics of each group, now formally compare them. If the editors think these data should be included, we are happy to include it.

2. The first line in table 2 is the only one presented without a denominator. The authors may want to add this in for consistency.

Thank you for pointing this out, we have added the denominator in.
Reviewer #2: This study investigates the individual predictive value of several first trimester markers for poor outcome in monochorionic twin pregnancies in a cohort of 177 twins. In reviewing it, I have focused on primarily on the methodology.

Major comments

It may not be entirely clear to the reader how predictive markers for adverse outcomes in twin pregnancies during the first trimester could improve clinical care. In the second paragraph, the authors mention that intensive surveillance associated with MC twin pregnancies has costs and burdens. Is the idea that low risk pregnancies could be identified where surveillance could be reduced or is it the hope that high risk pregnancies could be identified so that interventions can be taken. Maybe both are possibilities, but it would be helpful to have a little more on this in the introduction.

Thank you for your comments, we have included more detail in the introduction to further demonstrate the importance of our study. We agree that both are possibilities but the consequences of a false negative may well be fetal death, and thus we have focused on how clinical care could be improved for those at higher risk. We have also included other benefits of prognostic factor research, that are not linked to direct patient care.

I don't agree with the choice of the wording "clinical utility" as I think this should be reserved when looking at the impact on health and cost outcomes when a predictor is implemented in practice. What the authors refer to as "clinical utility" is just dichotomization of the prognostic factor to help simplify the results for the reader.

Thank you for your comment, we did debate this phrase at the time of writing the initial manuscript as it didn’t sit entirely perfectly with us. We have changed it to ‘Translating the prognostic effects into absolute risk’ as this is what we have actually done.

The exclusion of twins without aneuploidy screening may make the results less generalizable if a large percentage of women don't participate (e.g. younger women). Please reflect on this in the discussion.
Thank you for comment, we agree this is a limitation and have included it in the discussion.

The authors did not adjust for gestational age at sampling because it was perfectly correlated with a factor under investigation, crown rump length. I don't think that gestational age should have been excluded for this reason if it thought to be an important confounder.

The reason that we did not adjust for gestation age was that CRL was a prognostic factor under investigation. We decided to calculate gestational age based on CRL as this is best practice because using the last menstrual period is known to be less accurate due to relying on women remembering, and many women having an irregular cycle. We agree this is a limitation, but as there was no significant difference between the complicated group and non-complicated group with regards to CRL, we believe we can do no more.
In terms of the maternal complications and composite outcomes, I wonder why you included gestational diabetes (half of all the complications) and hypertensive disorders. Do they occur more often in twin pregnancies? If not, you may just have the same findings as any study trying to predict these disorders (e.g. same findings as a study looking at biomarkers to predict GDM in all pregnancies). I think the more important outcomes to focus on are the ones that are specifically related to twin pregnancies.

Yes they do occur more often in twin pregnancies, but this is why we have focused on the fetal complications and included the maternal composite outcome in the additional files and not main manuscript.

Minor comments

Ln 171. Do you mean that you used WHO growth curves? It's not entirely clear from this sentence.

Yes we do, our apologies, we have changed the wording to reflect this, thank you for pointing this out.

Ln 178-179. The neonatal and maternal composite outcomes should be specified in the text, not in an additional file. It's also not clear if the neonatal outcome in this sentence is defined the same as the primary outcome, which is also a neonatal outcome.

Thank you for your comment. As there is a limited word count for the manuscript, and it was not in our primary outcomes, or a focus of this manuscript, we felt it was better to move these definitions to an additional file. The primary outcome and neonatal outcomes are very different. The only postnatal, and thus neonatal measure used in the primary outcome is birthweight as there is no 100% accurate antenatal assessment of fetal weight. The other conditions included in the primary outcome all occur antenatally, not in the neonatal period.

Ln 196-199. It is not entirely clear from this sentence that you looked at the "added" value of the prognostic markers on top of the "existing factors". Please reword.

Our apologies. We have reworded this section and hope it is clearer ‘To gauge the potential increase in discrimination performance of a prognostic model that includes each potential prognostic factor in addition to existing factors, the change in apparent c-statistic (increase in area under the curve) for each outcome was calculated (i.e. difference in apparent c-statistic for models with standard characteristics including or excluding each factor).’

Ln 211-214. While I can understand that dichotomizing the prognostic factors at commonly used thresholds makes the results clinically interpretable, I think I would have preferred to have (also) seen a decision curve analysis.
Thank you for the suggestion. See our response to Reviewer 1. We are not saying to use our findings for clinical decision making based on predicted risks from our model, so think it may be misleading to also include a net benefit examination of the model with and without the different factors. However, for illustration we do present the predicted absolute risks from the model for hypothetical patients, and one can clearly see that there are a wide range of predicted risks across the distribution of factor values, even when holding fixed standard characteristics at their mean values. Thus, this implicitly suggests that clinical decisions (based on a threshold of risk) may indeed be changed if the factor(s) is (are) included, as the risk predictions may change importantly.

Ln 220-222. It's not clear why the patient characteristics were dichotomized when looking at clinical utility.

We are sorry but we do not understand the reviewer’s comment. The patient characteristics were not dichotomise i.e. we did not divide the data into groups, we used mean values from the study cohort to create a hypothetical patient.

Table 2. Include total N. It should also be obvious what the composite outcomes are in the table. Some outcomes could have some more definition (e.g. spontaneous ptb), for example in footnotes.

Thank you for your comment. We have amended the table to make it clearer.

Line 361. I wouldn't describe TTTS as a confounder of the association between biomarkers and IUFD, but rather an intermediate. Explain that if TTTS is detected and treatable, then this would weaken the association observed between biomarkers and IUFD. I also didn't find this a logical following sentence. "This is an interesting finding as consequently PlGF could be viewed as a marker of severity of TTTS."

Thank you for your comment, we have removed the word ‘confounders’ from the whole paper. With regards your second point, we have reworded this sentence so it is hopefully clearer.

Reviewer #3: This is an important study trying to identify markers of MCDA twin complications.

Although the study is robust, I do have some comments.

Line 80: "Such obstetric surveillance requires ultrasonographic expertise, health economic resources, is time consuming, and targets all MC twins as a 'high risk population'. ". Currently, all MC twins are considered high risk and a proposed screening approach may be beneficial but currently not available.
Yes we agree and are glad you think it is beneficial. This is one reason why we are looking for prognostic factors which ultimately may allow us to develop models that predict which patients are high risk and which are low risk.

Line 122: Pragnancies…guidelines: Please define guidelines or reference as a minor flaw of the study later on

We have added in the references, thank you. Although we agree this could be seen as a minor flaw, the benefit is that this is representative of the care these women receive in real life, and a screening test would need to be able to be applicable to twin pregnancies at all units.

Line 150: This is the major problem of the study as you did not convert the values into Multiple of Median (MoM). As the values change through out pregnancy you need to convert all biomarkers into MoM. The NT values need to be expressed as percentiles as well.

We did not convert to MoM because we chose to analyse biomarkers on their original scale, rather than in terms of MoMs, as all centres used the same method of measurement. Further, we prefer to include adjustment factors (which are sometimes embedded in the MoM calculation) directly in the multivariable model. Indeed, in our experience, the MoMs are not consistent at different gestational ages and it does not resolve the problem. See https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1682205/

“We show that a constant threshold-MoMs cutoff for MSAFP values actually refers to different percentiles of MSAFP levels at different gestational ages and that the combining of MoMs values between centers and gestational ages, such as suggested by Wald et al. for deriving a patient-specific risk index, is highly questionable. The results presented in this paper are quite general and will apply to all situations where MoMs are used”

176: define Iatrogenic PTB

We do not believe that iatrogenic PTB needs defining, particularly given that it follows a sentence regarding spontaneous PTB.

Line 215 - same as above: use perzentiles

We chose to use the percentage discordance cut-offs used in existing literature, to enable comparison, as opposed to centiles.

Please provide a chart showing number of complications (for example TTTS) and consequences: for example Laser or Amniotic Fluid reduction etc... Also provide the GA of the performed procedure. The same goes for TAPS: most of TAPS are a complication of laser therapy and not a problem per se.

Thank you for your comment, but we intentionally have not included these data as the aim is the investigate which prognostic factors will predict which pregnancies will develop TTTS, not what the outcome of the TTTS will be, that is irrelevant in this context as there are other variables
which need to be considered and were beyond the scope of this paper, and ultimately would not affect whether a pregnancy develops TTTS or not. We have reviewed our cohort and neither of the 2 TAPS cases were associated with laser therapy.

Line 256 Please define dysfunction

We have stated in the Additional File 1, that by dysfunction we mean failure, which in our cohort was associated with acute fatty liver, acute kidney injury from a massive obstetric haemorrhage, and renal failure secondary to HELLP. We have not included this detailed information as the maternal outcomes were not the primary outcomes for the study, but we can include this information if the editors think it would be helpful.

343 - Please specify continues measurements as I do not see them. Also when and why where those measured. They are not performed routinely.

We are not sure what the reviewer meant by this comment. By continuous measurements, we meant that the data were not dichotomised/divided into groups. We kept the data as continuous data as dichotomising measurements loses data as we have said on lines 459-462. Therefore, our prognostic effects relate to a 1-unit change in a factor on its continuous scale.

Together, the understanding of the presented numbers strongly depend on the GA and performed procedures. Therefore, I do recommend to present MoMs and to specify the collective more closely.

The authors thank all 3 reviewers for their expert and helpful comments. We hope that we have responded to them adequately.