Author's response to reviews

Title: Meta-analysis of randomised Phase II trials to inform subsequent Phase III decisions

Authors:

Danielle L Burke (dlb735@bham.ac.uk)
Lucinda J Billingham (l.j.billingham@bham.ac.uk)
Alan J Girling (a.j.girling@bham.ac.uk)
Richard D Riley (r.d.riley@bham.ac.uk)

Version: 2  Date: 3 July 2014

Author's response to reviews: see over
Response to Referee and Editor Comments

Ref: 1581369937124657
Authors: Burke, Billingham, Girling, Riley
Title: Meta-analysis of randomised Phase II trials to inform subsequent Phase III decisions

We would like to thank Professor David Jones and Professor Lawrence Joseph for their helpful comments to improve and clarify the content in our article. We have taken these comments on board (as documented below) and also removed any unnecessary repetition.

We now respond to each of the comments in turn and detail any changes we have made to the original manuscript. The referee/editor comments are shown in bold italics, and our responses are in normal text.

N.B. The reviewers’ comments refer to page numbers in the original version.

RESPONSE TO EDITORIAL REQUESTS

1. If appropriate, please include a list of abbreviations used and their meanings, after your Conclusions.

Manuscript now includes a list of abbreviations and their meanings after the Conclusions section.

2. Please include a Competing Interests section at the end of the manuscript, before the Authors’ Contributions section. If the authors have no competing interests, please state: "The authors declare that they have no competing interests.”

Added the statement: “The authors declare that they have no competing interests” before the Authors’ Contributions section.

3. Please upload all figures as separate files via the online submission system. They should not be included within the main manuscript document.

Figures have been removed from main manuscript document and uploaded as separate files.

4. Please include a figure title and legend section in the main manuscript, after the reference list.

Manuscript now includes a figure title and legend section in the main manuscript, after the reference list.

RESPONSE TO PROFESSOR LAWRENCE JOSEPH’S COMMENTS

Major compulsory revisions:

1. It is not clear why a continuity correction would be required in a Bayesian approach to meta-analysis. I would have expected that the prior density would take care of any such need. Why is this not the case for the models used here?

This is an interesting issue. A continuity correction is avoided by the Bayesian approach where the problem studies have just one treatment arm with no events. However, computational issues remain with the Gibbs Sampling estimation using WinBUGS when some studies have zero events in both arms. We even tried using highly informative prior distributions, but the problem persisted. Thus, in this study with two zero cells, we still use the method of Sweeting et al. to add a continuity correction. This issue is now described in detail in section 2.2.3, and we have added
more clarification of why we need the continuity correction still. Furthermore, in our applied example we now also look at the influence of this study on the meta-analysis results, and find that the study (perhaps unsurprisingly) does not have much impact on the posterior estimates. Please see the start of Section 3.1.2.

2. Perhaps related to the first point, the WinBUGS program in the Appendix, there is a line to calculate variance which is similar to what one would need from a frequentist approach (essentially \(1/a + 1/b + 1/c + 1/d\)) which would seem to be unnecessary when using a Bayesian approach, where the variance would naturally arise from the posterior density, and would not need to be approximated in this way. Again, why is this calculation necessary? See also below equation (4) on page 12 of the paper. The predictive distribution can be derived without this variance formulas, by simply using existing parameters of the model to predict the "next trial" similar to those already in the meta-analysis, the total variance being the sum of within and between study variances. Thus, options 1 and 2 on pages 13 and 14 seem unnecessary as well.

Sorry for the confusion here. Let us clarify. Firstly, the only place where this variance formula is needed is when looking at the predicted distribution for the estimate and its 95% interval in a new (or 'next') trial. It is not used in the meta-analysis itself, or when obtaining a prediction interval for the true treatment effect in the new trial (see equation (3)).

We agree with the reviewer that the total variance for the 'next trial' will account for the within-study variance and the between-study variance. However, the user must actually specify the within-study variance for this new trial. In other words, there is no posterior distribution for the within-study variance for a new trial to be sampled from (as the within-study variances in the meta-analysis are assumed fixed and known). Thus, in our paper we use the frequentist approximation to the variance of a log odds ratio estimate, \(1/a+1/b+1/c+1/d\), to allow the user to specify the within-study variance for their next trial by choosing (in option (1)) particular values of c, d, and sample sizes in each group (given the posterior sample of the \(Y_{i,new}\) these choices then allow a and b to be derived in each iteration, and so \(\text{var}(\log \text{odds ratio})\) is then known. This approach mimics what would be done in real life when the next trial is actually analysed. Typically a Phase III trails is analysed using a frequentist approach, and so would use this variance formula to derive a 95% confidence interval for the treatment effect.

In this manner, equation (4) quantifies all uncertainty and variability when making predictions: the uncertainty in the meta-analysis parameter estimates, the between-study variability, and now - through the variance formula - the sampling variability of the next trial's estimate.

Further clarification of the need to actually provide the variance of the new trial's effect estimate has been added to the manuscript. Please see Section 3.1.4.

Minor essential revisions:

1. How were the number of iterations and burn-in used in the MCMC results chosen? 100,000 seems like overkill as burn-in, and I wonder why such a high number was chosen. Is there a particular problem with convergence in this model?

We can reassure the reviewer that there was not a problem with convergence in this model. The large burn in and iteration length were selected for two reasons. Firstly, the trials each have small sample sizes, therefore, we were cautious that the posterior distributions would be wide, which requires a larger number of iterations so that the posterior density is smooth. Secondly, the model only takes minutes to run so it was easy to run for longer to be more cautious than perhaps
necessary. Different starting values were used to check convergence and different burn in and iteration lengths were tried and the results were the same.

RESPONSE TO PROFESSOR DAVID JONES’ COMMENTS

Major compulsory revisions

1. Must perform and summarise results of a (critical) literature review of methods of meta-analysis available for and used in the context of synthesizing Phase II study results. This literature is not particularly extensive, but if necessary the emphasis could be on Bayesian random effects models. One example likely to be included is the study by Assifi et al in Surgery 2011:150;466.

Thanks for this valid and important point. Since we only detailed one example in the paper for illustration, we agree it is important for the reader to know whether this is an isolated example, or actually typical of other meta-analyses of Phase II trials. To address this, we searched for other recent meta-analyses of Phase II trials that aimed to evaluate a treatment effect. Although the example suggested in your comment by Assifi et al. is a meta-analysis of phase II trials, we have not included it here as it includes non-randomised trials. Of the other examples, we found that many used meta-analysis methods with similar statistical issues to those in the Eikelboom example. In particular, $I^2$ is often used to make decisions about fixed or random effects models; inverse-variance methods are preferred to exact binomial likelihoods; and prediction intervals are rarely considered. We have summarised these issues in a new section of the Discussion, entitled 'Relevance of our work to recent meta-analyses of Phase II trials' (please see page 22). We hope this reassures the reviewer that our findings in the Eikelboom example are not an isolated example, and that the same methodological issues arise in other papers also.

Discretionary revisions:

2. Some tendency to repeat material about the sequence of issues to be addressed in 2.2 in following results and Discussion sections could be reduced.

Some of the repetition about continuity corrections, heterogeneity, how to interpret fixed and random-effects meta-analyses, modelling correct binomial distribution have been removed.

3. Model 2 (p8, section 2.2.1): Since this section is mostly presented at a tutorial level, the provenance and interpretation of the suggested priors, and the reason for the use of $I(0,)$ should perhaps be set out explicitly.

Some further interpretation and explanation of the prior distributions have been added to this section as suggested. Please see Section 2.2.1

4. Sections 2.2.3 and (especially) 2.2.4 are long, if reasonably clear. The paper would read better if the more technical details of implementation were moved to appendices/supplementary materials, leaving relatively brief outlines of approaches in the main text.
As mentioned, the article has been greatly reduced, especially Section 2.2.3 and 2.2.4. Statistical details of how to estimate the variance have been moved to the Supporting Information.

5. **Section 3.2.2:** The possibility indicated here is of course worth mentioning, but more specific indications of how to proceed are needed to make it practically useful, and potential dangers of use of strong priors not based on evidence of plausibility should perhaps be noted. Fig 6 could go in supplementary materials.

Figure 6 has been moved to the Supporting Information (this is now Figure 7). We have added some more guidance about how to derive these sceptical prior distributions based on external evidence or clinical guidance. We have also been more explicit about the guidance that is given by Spiegelhalter et al. for readers to use. Please see section 3.2.2 in the revision.

6. **Conclusions:** Although these currently reflect the comparison with Eikelboon’s approach/results, this is rather limited. Could brief conclusions/recommendations for good practice in meta-analysis of Phase II trials not be added?

A table of recommendations for good practice has been added to the conclusions.

7. **Abstract – Results l2:** ‘.. encompass the subsequent Phase III trial results’?

Abstract - Results line 2-amended sentence.

8. **Abstract – Results l4:** 0.82 or 0.83 (as in 3.1.3)?

Where the text originally reported 0.827 this has been amended to 0.824.

9. **p9 l-4:** I dislike the ‘…will likely be..’ expression used here and elsewhere. [However, I concede that it may be regarded as quite acceptable by many other readers!]

Where the phrase “will likely be” was used, this has been changed to “will be” throughout the text.

10. **p12 l10:** ‘…typical number (<10) of studies..’?

The wording has been changed to say that identifying causes of heterogeneity is difficult if there are few studies (for example <10).

11. **p18 l-11:** ‘.. actually entirely plausible..’ This seems a little over-emphatic, even if technically correct.

This wording has been changed to say “plausible” rather than “actually entirely plausible”.

12. **P20 l-1:** ‘..such as in the example by..’

Changed wording in manuscript to suit.

Minor essential revisions:
13. P21 l12: ‘including’, not ‘included’

Changed “included” to “including”.

We thank the two reviewers again for their constructive comments.