Author’s response to reviews

Title: Between-centre differences and treatment effects in randomized controlled trials: A case study in traumatic brain injury

Authors:

Hester F Lingsma (h.lingsma@erasusmc.nl)
Bob Roozenbeek (b.roozenbeek@erasusmc.nl)
Pablo Perel (Pablo.Perel@lshtm.ac.uk)
Ian Roberts (Ian.Roberts@lshtm.ac.uk)
Andrew I Maas (Andrew.maas@uza.be)
Ewout W Steyerberg (e.steyerberg@erasusmc.nl)

Version: 3 Date: 1 July 2011

Author’s response to reviews:

Dear Editor,

We would like to thank you very much for considering our manuscript ‘Between-centre differences and treatment effects in randomized controlled trials: A case study in traumatic brain injury’ for publication in Trials and sending it out for review. We also would like to thank the reviewers for their thoughtful and valuable comments. We addressed all comments below, and are pleased to hereby submit the revised version of the manuscript. We hope you will decide to accept it for publication in Trials.

Yours sincerely, on behalf of the co-authors,

Hester Lingsma

Reviewer 1

1. It is not clear whether the analysis of variation in outcomes is for those in the control group only or those in both control and intervention groups. It might be more natural to consider variations in those without treatment, but I don’t think this is what the authors did.

Indeed we did consider the variation in outcome in the whole population, not separate in the two treatment groups. We do not think that separating the treatment groups would make a difference since we expect the randomization to be balanced within centres and countries. In addition, the variation in outcome remains present with treatment effect in the model.

2. To test the null hypothesis of no centre effects, by comparing model fits with and without inclusion of (random) centre effects, one of the most natural tests would be a goodness of fit test. It might be useful to include such a test although it would not influence over all conclusions.

We agree, indeed we did compare the different models based on -2 log
likelihood. The result of this test was not reported, to avoid confusion with the p-value from the test for tau2 being different from 0, which leads to the same conclusion.

3. A random effects (random intercept) model allows for variation between centres in the control group (call model A). The inferences of the outcome data (before treatment) reveal that there are considerable variations between centres in outcomes and so that the random intercept model is warranted. The random slope model allows for variation between centres in treatment effect. This model fit compared with the model with A provides a test of goodness of fit for comparison of the null hypothesis of no variation between centres in treatment effects. Goodness of fit tests would therefore allow direct evaluation of this hypothesis.

We agree, although we evaluated model A in the total group of patients. We tested the null hypotheses of no presence of differences in outcome and treatment effect with tau2 and by comparing the deviances, but reported only the former. See also response to comment 1, 2 and 4.

4. An alternative approach (as opposed to comparisons of deviances) would be to simply see if the tau-squared parameters were significantly different from zero, as the authors do. In these model fits the tau-squared parameters (for the random slope) is significantly different from zero, and so the treatment effect does vary between centres.

This is indeed the approach we have taken. The p values for tau2 being different from 0 are reported in the results section.

5. However, the authors compare the average (pooled) treatment effect between models, and finding that they are similar, conclude that the treatment effect does not vary by centre.

Besides the p value for tau2, we have reported the estimated treatment effects since these indicate clinical relevance of the differences between the models, regardless of the significance.

6. On face value these findings seem inconsistent – a tau-squared estimate which is significant yet the conclusion that there is no sign random variation in treatment effects. This finding needs justification. Statistically I don’t believe that it is correct that there is no variation between centres in treatment effects. It might be that this difference is so small that it is not clinically meaningful, but if so this must be clearly stated.

We agree with the reviewer on this issue. Our intention is to show that despite large variation in outcome, the variation in treatment effect is small (clinically irrelevant), and the point estimate is not affected. But indeed the conclusion that there is no variation in treatment effect is statistically incorrect. We have formulated our findings and conclusion more cautious throughout the paper.

8. At equation 8 have the authors used the correct formula for the 95% range? 95% CIs usually contain terms for the standard error of the parameters, whereas
in 8 it is the variance which is used.
The formula is in fact correct; the range of ORs is an expression of the width of the distribution of centre effects. It is not a confidence interval but a 95% range.

9. It might be that a better representation of variation between centres is provided by the 95% prediction interval, which would be something like
\[ \text{betahat} \pm (1.96 \times \text{sqrt(tausquared+var(betahat)))} \]
This is an interesting thought, we agree that including both the within and between centre variance gives a good indication of the between-centre differences. The prediction interval would be however a value per centre, which also requires a summary measure again. We therefore decided to stick with our current parameter of between centre variance tau2.

10. Page 7 PLATO study, with only a handful of counties it could be argued that there will be insufficient groups from which to estimate variation.
We agree, and this is added to the discussion.

11. Abstract results – it is not clear which models all of the quoted ORs pertain to.
We tried to clarify this.

12. Generally the clarity of the write up could be improved throughout.
We tried to clarify the text throughout the manuscript.

13. The models listed in the appendix don’t seem to be consistent with those described in the methods – the authors describe random intercept and slope model yet in the appendix it is only the random intercept or random slope, and no model is listed in which both are included.
You are correct, this model is now included in the appendix

Reviewer 2

1. Page 3 "filed" should be "field"
Thank you, this is changed

2. While imbalances do occur, i.e. by center, the goal of randomization is eliminate selection and prognostic biases. The authors should address whether there were imbalances in randomization. Did analysis of the center effect eliminate imbalances?
Thank you for this interesting but complicated point. When there are imbalances within centres, this might affect the treatment effect in that particular centre. With overall treatment effect estimation imbalances are leveled out, but with large between centre differences in outcome, this approach might cause bias. Analysis of the centre effects (i.e. stratification by centre) does not eliminate imbalances, but shows the differences in treatment effect which might be caused by imbalance. The estimation of random slopes offers an intermediate approach,
with estimation of one overall treatment effect with between centre variance, but
does not eliminate imbalances.

3. For the discussion, would the authors consider that analysis by center a
standard or optional?
We would consider the analysis by centre optional. Only when centre differences
in treatment are expected, or when a centre of country specific estimates are
required (e.g. for approval) estimation of the treatment effect using a model with
random slopes could be considered. This is added to the discussion.

4. Would the effect be different for procedural trials or for those with substantial
treatment effect?
Yes, we believe that between-centre differences could have more impact when
the intervention is more complex, for example in a procedural trial. Differences in
outcome might be caused by differences in structure and processes of care, i.e.
quality. These differences in quality might also affect the execution and thus the
effect of a more complex study intervention. This is also mentioned in the
discussion. We have no reason to assume that the effect of centre-differences is
dependent on the magnitude of the treatment effect.

6. Did center effect modify the covariate effect on outcome or did the covariate
modify the center effect?
This relates to your point on imbalances. When there is imbalance in a centre,
the covariate effect (treatment effect) might modify the centre effect. The
mechanism we assume however, and what trialists fear, is that the centre effect
modifies the covariate effect.

7. Do the authors know the reasons for the center variability on outcome, patient
or center level variables?
The differences in outcome between centres are not explained by patient
characteristics, since we adjusted for these. Our finding that part of the
between-centre differences were in fact between country differences indicate that
outcome differences are explained by centre level characteristics, but we could
not further study this in our data. This remark is added to the discussion.