Author’s response to reviews

Title: Many continuous variables should be analyzed using the relative scale: a case study of β2-agonists for preventing exercise-induced bronchoconstriction

Authors:
Harri Hemilä (harri.hemila@helsinki.fi)
Jan Friedrich (friedrichj@smh.ca)

Version: 3 Date: 07 Aug 2019

Author’s response to reviews:

Many continuous variables should be analyzed using the relative scale:
 a case study of β2-agonists for preventing exercise-induced bronchoconstriction

Responses to Reviewer #3
2019-8-7
Harri Hemilä and Jan Friedrich (HH+JF)

Reviewer reports:
Reviewer #3: I read the revised paper by Hemilä and Friedrich and the authors' response to my comments, and I have some remaining comments.

1. page 5, Statistical Methods, lines 19-29: Why don't the authors write their three models as model equations in the usual way? It would be much easier to understand. With $X = FEV_1$ and $Y = FEV_2$ the model with intercept and slope (for example) would look like

$$Y_i = \alpha + \beta X_i + \epsilon_i.$$

HH+JF: Very good suggestion. When we had been working on the manuscript, the models that we compared seemed so simple that we did not see a need for the explicit formulas. We understand that stating the explicit formulas makes the text easier for the readers. We added the formulas in the revised version of the manuscript.

2. page 12, line 12 (and probably also other places): Please adhere to the correct notation in R. For example, metacont() is a function, not a procedure.
HH+JF: Corrected.
3. page 12 and Table 5: I now see that the authors calculate the standard errors and confidence intervals on the relative scale in the same way as on the absolute scale such that they obtain symmetric CIs also on the relative scale. The question, however, is whether this makes sense. A relative scale is chosen just in cases where effects are thought to impact values in a multiplicative way. In my view it does not make sense then to calculate standard errors like on an absolute scale and thus mixing a multiplicative model with additive procedures. The authors should reflect this objection and give a justification, instead of giving lengthy instructions for use. See also points 4 and 7.

HH+JF: We do not quite agree with this comment.

One of us (JF) has written about the analysis of continuous outcomes on the relative scale based on the SE on the log scale, calculated with the Delta approach [ref 5 in our paper], which leads to asymmetric CIs on the absolute scale (but symmetric on the log scale).

However, we do not feel there is justification to require that the relative effect (ratio) on continuous outcomes should always be analysed with the above approach. There are many approaches to calculate CIs for ratios.

In linear regression, the slope gives the relative effect, i.e. the ratio between change in Y from change in X, and the 95% CI for the slope is calculated from the SE of the slope. In our linear regression analyses in Figures 2 and 4, we calculate the relative effect, and the CIs with these calculations are symmetric since the CI is based on the SE on the absolute scale. In multivariate linear regression, several continuous X variables give slopes that show the relative effect of various X variables on the outcome Y, conditionally that the other X variables are constant. We do not agree with the statement “In my view it does not make sense then to calculate standard errors like on an absolute scale”.

Furthermore, let us consider another illustrative example. Let the duration of hospital stay be 8 days in the control group (SD=2 and N=10) and 6 days in the intervention group (SD=2 and N=10). On the absolute scale, this leads to 2 days shorter hospital stay in the intervention group with 95% CI from 0.12 days to 3.88 days (t-test based CI). This 2-day effect corresponds to 25% shorter hospital stay with the 95% CI from 1.5% to 48.5%, based on the CI calculated on the absolute scale (i.e. 1.5% = 0.12/8). We do not agree that this kind of CI calculated from the absolute scale is inappropriate in such an example.

Linear transformations do not influence this kind of analysis. We can use meters or inches, but the physics remains the same. Similarly, we can measure hospital stay in hours, which would lead to the multiplication of estimate and its 95% CI by 24. Division of the findings by the control group duration is mathematically fully analogous to using a different unit of time, in that case, 1 unit of time is the duration in the control group, but the relations between SDs and the hospital stays remain identical, and the transformation does not influence eg the t-test and the validity of the analysis.

If the SD is proportional to the size of the outcome, and the sizes of the outcomes in two groups differ dramatically, then the SDs also differ dramatically. Let us consider that the duration of
hospital stay is 8 days in the control group and 2 days in the intervention group (and SD proportional to the mean). In such a case, the calculation of relative effect seems much more appropriate to be carried out by using the Delta approach [ref 5 in our paper], which would lead to asymmetric CIs (though symmetric on the log scale).

There are many ways to calculate CIs for ratios. In a previous paper [ref 9 in the manuscript], HH compared the relative effect confidence intervals for particular IPD disease duration data; calculated on the basis of the absolute scale, on log transformed IPD, on the Delta approach of group means and SE, on Fieller's approach (https://en.wikipedia.org/wiki/Fieller%27s_theorem), and on Bootstrapping.

See the comparison:

Obviously, the various approaches led to different 95% CIs as they are based on different reasoning. The Delta approach uses the SE of the log(outcome) and leads to symmetry of CI on the log scale, but the Fieller approach does not, while bootstrapping can lead to many kinds of forms of CI depending on the experimental data as the observed data do not need to follow the normal distribution on either the absolute or the log scales. There are papers that have compared eg the Delta method and the Fieller method for the analysis of ratio data (eg https://doi.org/10.1198/tast.2010.08130), but such issues are very far from the primary focus of our paper.

The focus of reviewer's above comment is on our Figs 5B and 5C.

In Fig. 5C, we use the SEs that we get from the linear regression models and, as described above, calculating the 95% CI for slope from the SE of the slope is the standard approach in linear regression.

In Fig 5B, we use the SE that is based on the absolute scale, as described in current Table S3. That SE is based on paired data, and therefore the Delta method cannot be used in the same way as was described by JF [ref 5 in our manuscript]. Although we want to present Fig 5 as an illustration that the relative scale analyses can be carried out with the combination of a spreadsheet program and a standard meta-analysis program, Fig. 5 is somewhat distant from our main issue (i.e. absolute scale vs relative scale). Therefore, we would not like to make the Fig 5 description and analysis more complex. The main purpose of Figure 5 is to demonstrate that relative scale effects can be calculated even if the particular statistical program might not have such an option readily available.

Reviewer writes "The authors should reflect this objection and give a justification, instead of giving lengthy instructions for use".

We do not agree that our instructions are lengthy. We assume that this comment refers to "a simple approach to pool results of study-level data on the relative scale when this option is not available in a statistical program is to normalize the results of the studies by dividing the absolute mean effects and their SD values by the placebo group mean outcome value (Table S3). Such a transformation can easily be done with a spreadsheet program and the transformed data can be
entered in a standard statistical program for meta-analysis." This text is just 75 words in a manuscript of 5700 words.

After the above text, we added a new sentence "In a spreadsheet program, the study results can also be combined on the logarithmic scale so that the pooled results are back transformed to the original scale in the meta-analysis program [5]." but we do not think it is relevant in our manuscript to discuss such issues in more depth. The interested readers can look eg ref 5.

In summary, in linear regression, the 95% CIs for relative effects (slope) are symmetric and we disagree with the comment "it does not make sense then to calculate standard errors like on an absolute scale". There are many different ways to calculate CIs for relative effects (ratio) but that field is far from the main focus of our paper. We do not consider that our way of presenting Fig 5B is inappropriate in the context, and as noted above Fig 5C is based on the standard approach in linear regression.

4. page 18/19, Legends to Figures. These legends are much too long. A figure legend is thought to precisely describe what is seen in the figure, not more. Particularly, it is not the place for adding extensive reflections to the text, to describe multiple single individual points on the figure, or to comment, interpret or discuss the results. Particularly, the legend to Figure 2 should be tightened.

HH+JF: We substantially shortened the Fig 2 legend. Although to a large extent figures should be intelligible alone, the Fig 2 legend had become too long. We deleted issues that are discussed in the text section and we refer to Fig 1 for help in the interpretation of Fig. 2. We re-read the other figure legends, but we did not consider that the other legends are too long, and reviewer does not specify any other figure that is too long.

5. page 18/19, legends to Figures 2 and 4: continuous line -&gt; solid line (four times).

HH+JF: Done

6. page 19, legend to Figure 4: "Each circle indicates one study and the area of the circle is proportional to the square root of the number of participants and represents the weighting of the study": This doesn't make sense to me. The usual weighting in a meta-analysis is by inverse variance, which for a continuous endpoint is proportional to the sample size, n. To have the area of the circles proportional to the weights, they must be proportional to n, that is, the *radius* should be proportional to sqrt(n) (not the area!).

HH+JF: There were some confusions in our calculations and presentation.
In Figure 4 plotting, the diameter was correctly related to the square root of N, which means that the area was related to “N”. Thus, there was an error in the Fig 4 legend. We corrected this to “area of the circle is proportional to the number of participants”. Thus, the areas were correct, but the legend was not.

However, when checking this issue further, we found that in the calculation of the linear regression we had used the square root of N as the weight, which was an error in that case. We recalculated the model with the correct weight (i.e. “N”) and revised the text and tables with the corrected numbers. The changes were small.

7. page 22, Table 5 seems unnecessary. My problem with the symmetric CIs was not that I was not able to calculate them. Rather, I doubt that they make sense. See point 3 above.

HH+JF: We moved Table 5 to Additional file 1 as Table S3. It seems to us that some readers might not see how the calculations for Fig 5 were done. As to the symmetric and asymmetric CIs, see our comments above.

(A minor point. You wrote in your response: "We do not add the reference suggested by the reviewer since we should renumber all the later references and there would be a high risk of some errors remaining in such a process ...". OK, this reference was not thought to be added to the paper. But in case it would have been important, this doesn't convince me as an argument - first, because no effort to amend a paper should be spared, and secondly: Did you never ever hear of literature management systems?)