### Contribution to the discussion of ‘When should meta-analysis avoid using the normal distribution?'

### Mark Simmonds

### Centre for Reviews and Dissemination, University of York, UK

The paper by Jackson and White [1] neatly highlights the many problems when making normality assumptions in meta-analysis. While these issues are real, we should not lose sight of what a meta-analysis is for: which is to provide a summary of evidence that is, in some sense, better than that available from any individual study. Taking a weighted average is a natural way of summarising studies and, as the authors note, normality assumptions are not required for such an average to be a sensible summary. The standard inverse-variance weighting, for example, is a reasonable choice of weighted average even if the effect estimates do not follow a normal distribution or if the estimated variance is not exactly the true variance.

This holds true even for random-effects analyses. A random-effects analysis merely alters the standard weight used for each study by some factor τ2. There are various estimators for τ2,these maydepend on normality assumptions, but, in the end, they just produce a number used to reweight the weighted average. Since there are many estimators for heterogeneity [2] ideally all meta-analyses should consider the consequences of a range of possible re-weightings.

The need for normality assumptions arises, however, largely because of the fundamentally incorrect assumption about what a meta-analysis is: namely that it is a summary of effect estimates across studies. A meta-analysis should always be interpreted as a re-analysis of the totality of the evidence; specifically, an analysis of all the data from all participants in all studies. Most meta-analyses, of course, do not have access to this complete data, but for binary outcomes [3] and sometimes for survival outcomes [4] we can reconstruct that individual-level data, or use the same form of analysis. As Jackson & White mention, generalised linear mixed models (GLMMs) are one way to achieve this and avoid most normality assumptions. Exactly how they should be used requires more investigation [5], but it should be remembered that they are not the only option.

The important point is to consider how the data should be modelled at the individual level. All meta-analysts should ask themselves how they would analyse the individual data if they had it, and if not, what simplifications and assumptions (including normality assumptions) are required to analyse the available summary data.

Finally, while being aware of the assumptions made in a meta-analysis is important, there appears to be little reason to believe that assuming normality will lead to substantial bias or incorrect conclusions in most cases, although this would benefit from further research. Any problems arising from making normality assumptions are likely to be small compared to the much bigger problems facing a meta-analyst, such as publication bias, outcome reporting bias or the poor quality of included studies. In all meta-analyses we should not strain at the gnat of non-normality while happily swallowing the camel of systematic bias.

[1] Jackson, D., White, I.R. (2018). When should meta-analysis avoid making hidden normality assumptions? *Biometrical Journal* **XX**, zzz-zzz

[2] Langan, D., Higgins, J. P. T., Simmonds, M. (2017). Comparative performance of heterogeneity variance estimators in meta‐analysis: a review of simulation studies. *Res. Syn. Meth.***8**, 181–198.

[3] Simmonds, M.C, Higgins, J.P.T. (2014). A general framework for the use of logistic regression models in meta-analysis. *Statistical Methods in Medical Research* **25**, 2858 – 2877.

[4] Guyot, P., Ades, A.E., Ouwens, M.J., Welton ,N.J. (2012). Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. *BMC Med Res Methodol***12**, 9

[5] Jackson, D., Law, M., Stijnen, T., Viechtbauer, W., White, I.R. (2018). A comparison of seven random‐effects models for meta‐analyses that estimate the summary odds ratio. *Statistics in Medicine***37**, 1059–1085.