JOURNAL OF Evolutionary Biology

doi: 10.1111/jeb.12945

COMMENTARY

Visualizing unbiased and biased unweighted meta-analyses

S. NAKAGAWA*† & M. LAGISZ*

*Evolution & Ecology Research Centre, School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, NSW, Australia

†Diabetes and Metabolism Division, Garvan Institute of Medical Research, Sydney, NSW, Australia

Most meta-analyses in ecology and evolution are concerned with an overall effect (average effect sizes) of a biological phenomenon. However, some meta-analyses are interested in average magnitudes or the mean of absolute effect sizes. For example, in a meta-analysis of selection gradients where signs of such gradients can be arbitrary, one uses absolute values of selection gradients as effect size (e.g. Kingsolver *et al.*, 2001). Michael Morrissey's Target Review (2016) clearly demonstrates that when average magnitudes are of interest, one must correct for the statistical noise (sampling error) of effect sizes. Otherwise, meta-analytic results, that is average magnitudes, will be biased. This is in contrast to a classical situation of overall average of effect sizes, which is unbiased when not accounting for sampling error (variance). Morrissey terms meta-analyses that do not deal with sampling error as 'informal meta-analysis'. Such meta-analysis is also referred to as unweighted meta-analysis, as effect sizes are not weighted by the inverse of corresponding sampling error variances (e.g. Jennions *et al.*, 2001).

Many may believe that informal (unweighted) meta-analysis provides biased estimates, whereas formal (weighted) meta-analysis provides unbiased estimates (cf. Mengersen & Gurevitch, 2013). However, it has been long been pointed out that unweighted mean is unbiased (Gurevitch & Hedges, 1999). What is new is that Morrissey (2016) shows that unweighted mean is unbiased only when it comes to average values but not average magnitudes. Morrissey mathematically describes why this is the case. However, we are slightly concerned many potential readers may skip his mathematically derivations (cf. Fawcett & Higginson, 2012), and they may not truly understand why we get unbiased or biased averages in different meta-analyses. Therefore, we present an intuitive visualization of Morrissey's main results in this commentary.

Correspondence: Shinichi Nakagawa, Evolution & Ecology Research Centre, School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, NSW 2052, Australia.
Tel.: +61-(0)2-9385-9138; fax: +61-(0)2-9385-1558;
e-mail: s.nakagawa@unsw.edu.au

The upper panels (a and b) of Fig. 1 depict idealized normal distributions of effect sizes. The lower panels (c and d) depict the folded normal distributions corresponding to a and b, respectively; they are a visualization of absolute values of effect sizes, such as selection gradients. The left panels (a and c) are based on the normal distributions with the mean of 0, whereas the right panels are based on the normal distributions with the mean of some positive value. The distributions in red are without sampling error variance (i.e. visualization of formal meta-analysis), whereas the ones in blue are with sampling error variance (i.e. visualization of informal meta-analysis). Note that the blue distributions have heaver tails as they include sampling error variance. Figure 1 conveys three important messages.

First, the mean values of the two normal distributions superimpose each other in Panels a and b. In other words, the average values are the same between the red and blue distributions (formal and informal meta-analyses, respectively). Put it in yet another way, the mean from informal meta-analysis will be unbiased. On the other hand, the means of the two folded normal distributions (Panels c and d) are different, with the blue distributions having larger means. That is, the distributions with sampling error provide upwardly biased estimates.

Second, relating to the first point, the upward bias in informal meta-analysis will be worse when the original (unfolded) normal distribution has a mean closer to zero than otherwise (compare Panel c with Panel d). One can easily visualize that if the mean of a normal distribution gets far enough from 0, then the corresponding folded normal distribution will be just that normal distribution. Then, the bias described here (Panels c and d) will be negligible. However, in most of cases when we are interested in absolute values (magnitudes), the mean of the original values will probably be close to zero (e.g. the examples in Morrissey, 2016). Thus, not controlling for statistical noise will overestimate the average magnitude of interest.

Third, the upper panels (a and b) actually give an intuitive explanation for why formal (weighted) meta-analysis is often preferred over informal (unweighted) meta-analysis. The red normal distributions (formal meta-analysis) are narrower than the blue normal distributions (informal meta-analysis, a and b). In other words, formal meta-analysis is more precise or statistically more powerful than informal meta-analysis (Gurevitch & Hedges, 1999).

Although we are very much in support of Morrissey's main thesis in his article, we would like to express some difference in opinion on his concluding paragraph. There, he warns about the danger of using meta-analysis in general, stating that meta-analysis can dilute a handful of very good studies, or that meta-

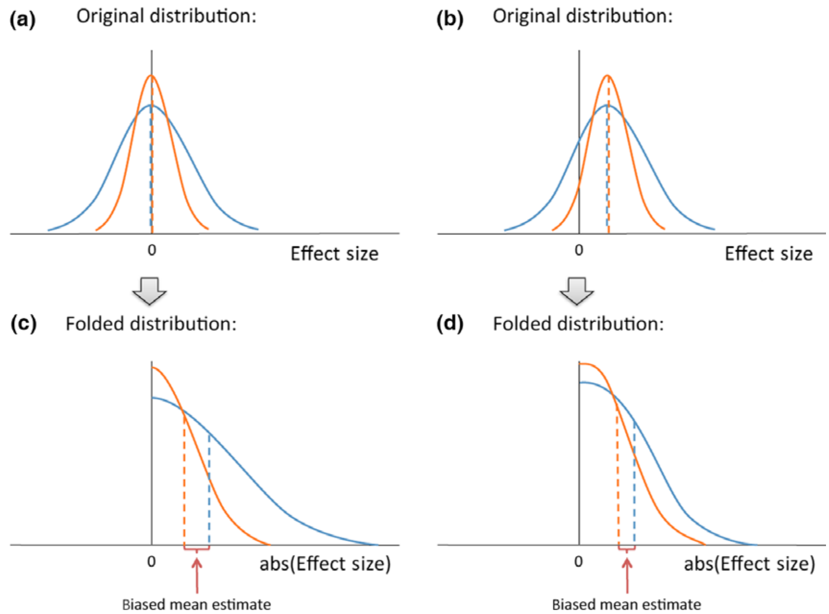

Fig. 1 A conceptual visualization of formal (red lines) and informal (blue lines) meta-analyses of average values and average magnitudes: (a) meta-analyses of average values with the distribution mean of 0, (b) meta-analyses of average values with the distribution mean larger than 0, (c) meta-analyses of average magnitudes with the distribution mean of 0, (d) meta-analyses of average magnitudes with the distribution mean larger than 0.

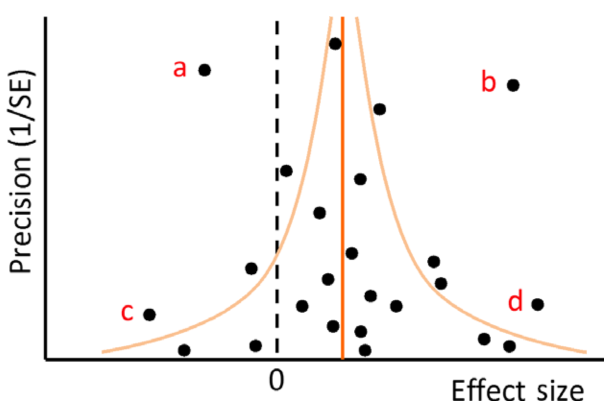

Fig. 2 A funnel plot of an imaginary meta-analytic data. Points a–d illustrate different types of potentially unusual effect sizes that are outside the confidence intervals of the funnel (orange lines). Each effect size is plotted against its precision (the inverse of standard error). Red vertical line indicates the meta-analytic mean (average).

analysis can even exclude creative studies, which use nonstandard methodologies. Thus, he emphasizes the importance of qualitative synthesis. These are indeed possibilities, and we do not necessarily disagree. However, we also want to point out that meta-analysis is equipped with tools that can detect unique studies.

Figure 2 shows a funnel plot (Egger *et al.*, 1997) of an imaginary meta-analytic dataset; a funnel plot has the meta-analytic mean (average) at the centre of the funnel with each effect size being plotted according to its precision (the inverse of standard error). Points a and b are more unusual than points c and d, although all the points are outside of the funnel (95% confidence intervals around the mean). This is because points a and b both have high precisions, but are far away from the meta-analytic mean. Points d and c, in contrast, have low precision, and thus, their large deviation from the mean may be expected. Also, point a is probably more remarkable than point b, because the sign of point a is the opposite of the mean. Our main message here is that graphical tools developed for meta-analysis can be used to identify special and unique studies that deserve more attention in a quantitative rather than qualitative manner. Nonetheless, these tools cannot completely replace researchers carefully reading every empirical study, evaluating the quality and importance of each work.

Acknowledgments

We are grateful to Rose O'Dea for her comments. SN is supported by a Future Fellowship from the Australian Research Council (FT130100268).

References

Egger, M., Smith, G.D., Schneider, M. & Minder, C. 1997. Bias in meta-analysis detected by a simple, graphical test. *Br. Med. J.* **315**: 629–634.

Fawcett, T.W. & Higginson, A.D. 2012. Heavy use of equations impedes communication among biologists. *Proc. Natl. Acad. Sci. U.S.A.* **109**: 11735–11739.

Gurevitch, J. & Hedges, L.V. 1999. Statistical issues in ecological meta-analyses. *Ecology* **80**: 1142–1149.

Jennions, M.D., Moller, A.P. & Petrie, M. 2001. Sexually selected traits and adult survival: a meta-analysis. *Q. Rev. Biol.* **76**: 3–36.

Kingsolver, J.G., Hoekstra, H.E., Hoekstra, J.M., Berrigan, D., Vignieri, S.N., Hill, C.E. *et al.* 2001. The strength of phenotypic selection in natural populations. *Am. Nat.* **157**: 245–261.

Mengersen, K. & Gurevitch, J. 2013. Using other metrics of effect size in meta-analysis. In: *The Handbook of Meta-Analysis in Ecology and Evolution* (J. Koricheva, J. Gurevitch & K. Mengersen, eds), pp. 72–85. Princeton University Press, Princeton.

Morrissey, M. 2016. Meta-analysis of magnitudes, differences, and variation in evolutionary parameters. *J. Evol. Biol.* in press.

Received 15 March 2016; accepted 26 March 2016