CrossMark

Epidemiology and Psychiatric Sciences (2017), **26**, 369–370. © Cambridge University Press 2017
doi:10.1017/S204579601700004X

COMMENTARY TO SPECIAL ARTICLE

Every effect size has its place: a commentary on the avoidance of pre–post effect sizes

Received 15 January 2017; Accepted 27 January 2017; First published online 16 February 2017

Key words: Meta-analyses, pre–post effect sizes, trial design.

Commentary on: Cuijpers *et al*. Pre–Post effect sizes should be avoided in meta-analyses. *Epidemiology and Psychiatric Sciences*. (doi: 10.1017/S2045796016000809)

Along with the increasing popularity of meta-analysis, their methods have very much improved in the last decades, leading to highly sophisticated methods, for example multiple treatment comparisons and advanced multiple regression models (Sutton & Higgins, 2008). However, these improvements of meta-analytic methods should not obstruct one's view of the basic decisions and assumptions, because these may have a significant impact on result and interpretation of a meta-analysis, as was nicely illustrated by Cuijpers *et al*. (2016) in this journal for the choice of the effect size.

In their paper, the authors illustrate that pre–post effect sizes do not reflect a specific treatment effect only, but will also be affected by various processes, for example spontaneous remission. I think this cannot be questioned, but I am less sure if I am in support of a general statement that suggests avoiding pre–post effect sizes in meta-analysis. Thus, I will play the devil's advocate in defence of the pre–post effect sizes and will give some thoughts to their usefulness.

As already mentioned by the authors, pre–post effect sizes can for example be informative in a comparison of efficacy and effectiveness trials (van der Lem *et al*. 2012). Most effectiveness trials use single group designs, and therefore it is impossible to calculate comparative effect sizes. Consequently, excluding these trials by focusing on controlled effect sizes only may miss a substantial part of the evidence. This may be less important for trials of single interventions, but it is of major importance when summarising evidence in areas where randomised controlled trials are rarely conducted, for example in research on complex interventions (Girlanda *et al*. 2017). Besides this rather pragmatic view, pre–post effects may be more interesting for clinicians than the comparative effects, because I suppose many clinicians are interested in an overall improvement of their patient, irrespective of whether the effect is a specific or unspecific effect of the medication. The pre–post effects may therefore be more interesting for clinicians comparing their treatment outcomes to the literature. This may be particularly true for data derived from placebo controlled trials, which are testing against a treatment option, which is not available for clinicians in daily practice. Due to the drug approval requirement by the drug agencies (Barbui & Bighelli, 2013), the basis for the body of evidence for drug treatments is derived from these trials, and one may argue, that the focus on reporting of relative drug placebo differences may contribute to the observation of the efficacy-effectiveness gap.

In their paper, Cuijpers *et al*. (2016) also highlight the methodological problems connected to the pre–post effect sizes, and I very much appreciate that the authors address the frequently ignored statistical dependency of the pre–post measurements and illustrate the use of common assumptions with a real data example. As illustrated by their example, the pre–post effect size estimates can be misleading if the assumption of the pre–post correlation is incorrect. However, for a wide range of assumed correlation (e.g., $r = 0.0$ to $r = 0.5$ compared a true correlation 0.18 for Hamilton Depression Rating Scale) the effect size estimations were quite robust and may be within an acceptable error margin. Furthermore, the authors point out that heterogeneity can be very high in meta-analyses using pre–post effects. Although very high heterogeneity of effects can also be found in comparative effectiveness research (Koslowski *et al*. 2016), it is plausible to assume that meta-analyses using pre–post effect sizes show higher heterogeneity of the treatment effects. These analyses will more often include non-controlled studies and may have less restrictive inclusion criteria. However, the heterogeneity itself contains some important information, thus

Address for correspondence: M. Kösters, Department of Psychiatry II, Ulm University, Germany.
(Email: Markus.Koesters@uni-ulm.de)

https://doi.org/10.1017/S204579601700004X Published online by Cambridge University Press

the exploration of the heterogeneity may be valuable in understanding moderators of the treatment effects.

As shown by Morris & DeShon (2002), every effect size can be biased and the sources of bias differ between effect sizes. Cuijpers *et al*. (2016) mention that the pre–post effect size is affected, for example by time effects. However, the comparative effect size may be affected by group-related effects, which may not be directly related to the treatments. In a double-blind placebo controlled trial, adverse events of the medication may debunk group assignment, and this may influence the expectation and ratings of clinicians and patients (Basoglu *et al*. 1997; Moncrieff *et al*. 1998). Although this bias related to the study design and conduct, it will affect a comparative effect size and may result in an overestimation of the group difference. A pre–post effect size is also prone to biased ratings, but because the effect sizes can be derived at group level, they may be more helpful in evaluating such bias. Pre–post effect sizes could be used to identify unusual low effect size in the placebo group, which is hard to see in analysis of comparative effect sizes.

In conclusion, pre–post effect sizes may not be a good choice for summarising experimental research, and it is difficult to use them properly. However, every effect size has its place, and the decision in favour of a specific effect size measure should depend primarily on the research question.

M. Kösters
Department of Psychiatry II, Ulm University, Germany

References

Barbui C, Bighelli I (2013). A new approach to psychiatric drug approval in Europe. *PLoS Medicine* **10**, e1001530.

Basoglu M, Marks I, Livanou M, Swinson R (1997). Double-blindness procedures, rater blindness, and ratings of outcome. Observations from a controlled trial. *Archives of General Psychiatry* **54**, 744–748.

Cuijpers P, Weitz E, Cristea IA, Twisk J (2016). Pre-post effect sizes should be avoided in meta-analyses. *Epidemiology and Psychiatric Sciences*, 1–5.

Girlanda F, Fiedler I, Becker T, Barbui C, Koesters M (2017). The evidence-practice gap in specialist mental healthcare. Systematic review and meta-analysis of guideline implementation studies. *British Journal of Psychiatry* **210**, 24–30.

Koslowski N, Klein K, Arnold K, Kösters M, Schützwohl M, Salize HJ, Puschner B (2016). Effectiveness of interventions for adults with mild to moderate intellectual disabilities and mental health problems. Systematic review and meta-analysis. *British Journal of Psychiatry* **209**, 469–474.

Moncrieff J, Wessely S, Hardy R (1998). Meta-analysis of trials comparing antidepressants with active placebos. *British Journal of Psychiatry: the Journal of Mental Science* **172**, 227–231; discussion 232-4.

Morris SB, DeShon RP (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods* **7**, 105–125.

Sutton AJ, Higgins JPT (2008). Recent developments in meta-analysis. *Statistics in Medicine* **27**, 625–650.

van der Lem R, van der Wee NJA, van Veen T, Zitman FG (2012). Efficacy versus effectiveness: a direct comparison of the outcome of treatment for mild to moderate depression in randomized controlled trials and daily practice. *Psychotherapy and Psychosomatics* **81**, 226–234.

https://doi.org/10.1017/S204579601700004X Published online by Cambridge University Press