Detailed Information on Quantitative Data Synthesis and Moderator Analyses

Intragroup change effect size (standardized mean difference) was calculated using the following formula:

$$d = \left( \frac{Y_1 - Y_2}{S_{\text{within}}} \right)$$

where $Y_1$ is the pretreatment sample mean, $Y_2$ is the posttreatment sample mean, and $S_{\text{within}}$:

$$S_{\text{within}} = \sqrt{\frac{\text{SD}_1^2 + \text{SD}_2^2 - 2r \times \text{SD}_1 \times \text{SD}_2}{2(1-r)}}$$

where $\text{SD}_1$ is the standard deviation of the pretreatment sample mean, $\text{SD}_2$ is the standard deviation of the posttreatment sample mean, and $r$ is the correlation between pretreatment and posttreatment scores.

For studies reporting difference in means, standard deviation of difference and sample size, the intragroup change effect size was calculated using the following formula:

$$d = \frac{\overline{Y}_1}{\sqrt{2(1-r)}}$$

where $\overline{Y}_1$ is the given paired difference in means, $\text{SD}_1$ is the given standard deviation of the paired difference, and $r$ is the estimated correlation between pretreatment and posttreatment scores.

For studies reporting difference in means, confidence limits, sample size, and confidence level, the intragroup change effect size was calculated using the following formula:

$$d = \overline{Y}_1 \times \sqrt{2 \times (1 - R)}$$

where $\overline{Y}_1$ is the standardized paired difference in means and $R$ is the imputed $R$-value (given as 0.50).

Hedges $g$ can be computed by multiplying $d$ by correction factor:

$$J = 1 - \frac{3}{4df - 1}$$

where $df$ is the degrees of freedom to estimate the intragroup standard deviation.

$Q$ is determined by the following formula:

$$Q = \sum_{i=1}^{k} W_i Y_i^2 - \left( \frac{\sum_{i=1}^{k} W_i Y_i}{\sum_{i=1}^{k} W_i} \right)^2$$

with $W_i$ being the weight of the study, $Y_i$ the effect size of the study, and $k$ the number of studies included. To determine the expected value of $Q$, we used the degrees of freedom ($df = k - 1$), with $k$ being the number of studies included. A significant $Q$ test ($P$ value less than alpha set at 0.05) indicates heterogeneity in effect sizes.

We estimated the variance of the true effect between the studies ($T^2$) using the following formula:

$$T^2 = \frac{Q - df}{C}$$

where:

$$C = \sum W_i - \frac{\sum W_i^2}{\sum W_i}$$

$F^2$ is determined by using the following formula:

$$F^2 = \left( \frac{Q - df}{Q} \right) \times 100\%$$

$F^2$ is expressed as a ratio with a range of 0 to 100% and describes what proportion of the observed variance reflects real differences in effect sizes. Higgins and Thompson\(^1\) suggest that values of 25%, 50%, and 75% can be considered as low, moderate, and high, respectively.

We computed the fail-safe $N$ using the following formula:

$$X = \frac{K(K\overline{Z}^2 - 2.706)}{2.706}$$

where $K$ is the number of studies in the meta-analysis and is $\overline{Z}$ the mean $Z$ obtained from the $K$ studies. The effect size can be considered to be robust if the required number of studies ($X$) to reduce the overall effect size to a nonsignificant level exceeds $5K + 10.2$.

We used the Trim and Fill method, which examines whether negative or positive trials are overrepresented or underrepresented, accounting for the sample size. This information can then be used to recalculate the effect size estimates if the funnel plot is asymmetric. The divergence of the original effect size and the recalculated effect size reveal how robust the results are.

REFERENCES
