Effect of Placebo Conditions on Polysomnographic Parameters in Primary Insomnia: A Meta-Analysis

Alexander Winkler, Dipl.-Psych; Winfried Rief, Prof.

University of Marburg, Department for Clinical Psychology and Psychotherapy, Marburg, Germany

Study Objectives: Little is known about the role of placebo response in the pharmacotherapy of primary insomnia, especially about the effect of placebo intake on objectively assessed outcome variables. Our aim was therefore to conduct an effect-size analysis of placebo conditions in randomized controlled drug trials addressing primary insomnia also including polysomnography.

Design: We conducted a comprehensive literature search using PubMed, PsycINFO, PSYNDEx, PQDT OPEN, OpenGREY, ISI Web of Knowledge, Cochrane Clinical Trials, and the World Health Organization International Clinical Trials Registry Platform. The meta-analysis used a random effects model and was based on 32 studies reporting 82 treatment conditions covering a total of 3,969 participants. Special emphasis was given to the comparison of objective and subjective outcomes and the proportion of the placebo response to the drug response.

Measurements and Results: Effect sizes estimates (Hedges’ g) suggest that there is a small to moderate yet significant and robust placebo response reducing the symptoms of insomnia in terms of sleep onset latency (−0.35), total sleep time (0.42), wake after sleep onset (−0.29), sleep efficiency (0.31), subjective sleep onset latency (−0.29), subjective total sleep time (0.43), subjective wake after sleep onset (−0.32), subjective sleep efficiency (0.25) and sleep quality (0.31). Thus, the placebo response was also evident in objective, physiological (polysomnographic) variables. Our results indicate that 63.56% of the drug responses are achieved even in the placebo groups.

Conclusions: In light of these strong placebo responses, future studies should investigate how to exploit placebo mechanisms in clinical practice.

Keywords: insomnia, meta-analysis, placebo, polysomnographic, review, treatment

Citation: Winkler A, Rief W. Effect of placebo conditions on polysomnographic parameters in primary insomnia: a meta-analysis. SLEEP 2015;38(6):925–931.

INTRODUCTION

Primary insomnia is a frequent health complaint defined by difficulty in initiating or maintaining sleep, or nonrestorative sleep, for at least 1 month causing clinically significant distress or impairment in social, occupational, or other important areas of functioning. Symptoms do not occur exclusively during the course of narcolepsy, breathing-related sleep disorder, circadian rhythm sleep disorder, parasomnia, another mental disorder, or due to a drug’s direct physiological effects.

In the American Insomnia Survey, prevalence rates range from 3.9% under International Classification of Diseases (ICD)-10 to 22.1% under Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR). Because most primary studies applied diagnostic criteria based on DSM-IV-TR, we refer to DSM-IV-TR instead of DSM-V criteria.

Although almost half the individuals suffering from sleep problems never see a physician to address their complaints, most who consult a physician receive pharmacotherapy (approximately 50% in Western Europe or the United States, and up to 90% in Japan) to address their sleep problems. Although there is solid evidence of the efficacy of sleeping pills, the effect size only appears moderate, and it is unclear how large the placebo response is in relation to the drug response. Moreover, the benefit-risk ratio is frequently critical; for instance, the modest efficacy of hypnotics is accompanied by the risk of serious side effects and dangers such as cognitive effects, daytime fatigue, tolerance, addiction, risk of falls, fractures, depression, suicide, and increased mortality.

Other areas of research reveal evidence that the placebo response accounts for up to 75% of the treatment effect in antidepressant trials and up to 50% in pain or generalized anxiety disorder trials. Although pain or depression research is mainly based on subjective outcome variables, the subject of insomnia enables us to compare subjective and objective outcome parameters in the placebo group. It is frequently postulated that placebo responses are mainly detected in subjective scores, whereas the placebo groups in insomnia trials allow the comparison of subjective placebo responses with objective polysomnographic outcome variables.

Concerning insomnia trials, there have been three meta-analyses finding evidence of significant improvements under placebo conditions. Recent studies either did not report any effects in the placebo groups on objective outcome parameters (but they mainly focused on subjective aspects of sleep quality), or they suffer from small sample sizes and the lack of objective polysomnographic (PSG) data. Additionally, the proportion of the placebo response on drug response to different drug classes remains unclear.

We therefore conducted a meta-analysis of placebo conditions in PSG randomized controlled drug trials to examine the efficacy of placebo treatment for primary insomnia, to compare its efficacy on objective versus subjective outcome measures, and to determine its proportion in the response to pharmacological treatments. We conducted moderator analyses to identify potential treatment moderators.
METHODS
For this meta-analysis we adhered to Meta-Analysis Reporting Standards (MARS) guidelines. In addition to the assessment of within group changes in the placebo conditions, the identified primary literature was furthermore used for an analysis of between group comparisons to determine the efficacy of drug treatment of primary insomnia.

Search Procedure
We identified studies by searching PubMed, PsycINFO, PSYNDEx, PQDT OPEN, OpenGrey, ISI Web of Knowledge, and the Cochrane Clinical Trials Library. We conducted extensive searches for studies published between the first available year and April 5, 2013 using the terms insomn* and placebo* combined with the term polysomn*. In addition, we searched the World Health Organization (WHO) International Clinical Trials Registry Platform (ICTR), a manual review of relevant journals, and did a manual review of reference lists of relevant articles and review papers extracted from the database searches. We adopted comprehensive search strategies in order to identify both published and unpublished articles (including asking the contact persons in all clinical trials at the ICTR for data from their unpublished trials).

Determination of Outcome Variables
We chose “Sleep Onset Latency” (SOL) and “subjective Sleep Onset Latency” (sSOL) as core outcome variables. We also included “Total Sleep Time” (TST), “Wake After Sleep Onset” (WASO), and “Sleep Efficiency” (SE) as additional objective outcome variables (assessed with PSG recordings) and “subjective TST” (sTST), “subjective Wake After Sleep Onset” (sWASO) and “Sleep Quality” (SQ) as additional subjective outcome variables (assessed with sleep diaries and sleep questionnaires). These are established outcome variables in insomnia treatment trials.

Study Selection
Only pharmacological treatment trials addressing primary insomnia were considered via title and abstract screening. Studies were excluded after full text screening if no PSG data or insufficient data to perform an effect-size analysis were reported or a waitlist control condition was used instead of a placebo control condition. Studies were also excluded if the sample overlapped, either partially or wholly, with the sample of another study already included in the meta-analysis. Moreover, the study had to have used a double-blind randomized parallel group design using a placebo control condition, and the entire text had to be available in the English or German language.

We made no restrictions on sample size, treatment duration, or publication date because of the anticipated small number of studies using PSG data. We also made no geographical or cultural restrictions because we were interested in a global perspective on insomnia and its treatments.

Because our focus was on changes in the placebo control conditions instead of changes in the drug condition of primary insomnia trials, we made no restrictions on drug classes and decided to include trials assessing drugs not established for insomnia therapy if the study reported data for a placebo control condition separately.

Each identified article was further examined by two independent, experienced researchers for potential inclusion in the meta-analysis. Disagreements were resolved by discussion.

Validity Assessment
Only studies using a randomized controlled parallel group design were included. Nevertheless, we rated the quality of each study and analyzed study quality as a moderator to control for possible confounds. We therefore used the Jadad quality scale, which consists of seven dichotomous items with a maximum score of five and assesses aspects of validity. Each study’s quality was assessed independently by two trained researchers, and interrater reliability was calculated. Disagreements were resolved through discussion.

Data Extraction
For each study, data and the following study characteristics were extracted from each study collectively by two independent trained experts: total N, N of treatment group, N of control group, drug in treatment group, dose of treatment drug, class of drug in treatment group, duration of treatment, average age in placebo group, and percentage of female participants in placebo group. In case of missing data on age or percentage of female participants in subgroups, we used age and percentage of female participants in the total sample as an estimator. In cases of missing data on individual moderator variables, the relevant study was excluded only from the analysis of that moderator variable. Disagreements were resolved through discussion.

Quantitative Data Synthesis
All analyses were completed by using the software program “Comprehensive Meta-analysis, version 2.” We analyzed completer data in all cases. Separate within-group effect sizes for the continuous variables SOL, TST, WASO, SE, sSOL, sTST, sWASO, and SQ were calculated using within-group changes of placebo and drug conditions (for detailed information see supplemental material). We calculated effect sizes using Hedges g and its 95% confidence interval. Hedges g is a variation of Cohen’s d that corrects for bias due to small sample sizes. The magnitude of Hedges g can be interpreted using Cohen’s recommendation for small (0.20), medium (0.50), and large (0.80). We followed Rosenthal’s recommendation and used a conservative estimate of between-treatment and posttreatment measures.

We used a test of significance based on the Q statistic to identify heterogeneity in effect sizes. Furthermore, we estimated the variance of the true effect between the studies (I²) to quantify heterogeneity in effect sizes. In addition, we used the ratio of true heterogeneity to total observed variation F. These methods are described in more detail in Borenstein, Hedges.

Effect size estimates for SOL, TST, WASO, SE, sSOL, sTST, sWASO, and SQ were pooled across studies to obtain a summary statistic. The effect size estimates were calculated using a random effects model. Instead of conducting a power analysis, we report the observed effect size with its confidence interval. For the purposes of conducting subgroup analyses, we chose a random effects model and used the Q test for heterogeneity across studies to compare the effects of different subgroups.
We used a method described by Kirsch and Sapirstein and subtracted the mean placebo response rates from mean drug response rates to determine the proportion of placebo response to drug response to pharmacological treatment.

**Sensitivity Analysis**

To minimize publication bias, we conducted a careful literature search that included strategies to find published and unpublished studies. The results of our meta-analysis were considered to be unbiased and robust if the funnel plot for the effect sizes was symmetrical, the trim and fill method resulted in statistically significant recalculated effect sizes, and the fail-safe N exceeded 5K+10 (with K representing the number of studies included). We treated effect sizes as outliers if the distance to the average value of all effect sizes was 1.5 times the interquartile range or more.

**Moderator Analyses**

The moderating effect of study quality was tested to address the problem of possible confounds of effect sizes due to differences in methodological quality across studies, which is known in the literature as the garbage in/garbage out problem. Year of publication was chosen as a potential moderator because we wanted to know whether the methodological and technical developments in primary PSG studies moderate the treatment effect. Duration of treatment was chosen as a potential moderator to examine whether participants who received treatment for a longer period of time gained more or less benefit from the placebo treatment. Average age and percentage of female participants were chosen as potential moderators to examine (1) whether men and women or (2) younger and older participants gained the same benefit from the placebo treatment. Moderator effects were examined using meta-regression analyses (95% confidence intervals).

**RESULTS**

**Study Selection**

As Figure 1 shows, our initial search of databases identified 420 unique articles examined for relevance. After screening titles and abstracts, we selected 183 articles for full text evaluation. None of the 32 included studies fulfilling our selection criteria reported unusually high effect sizes with respect to the placebo group. The studies included in the meta-analysis included 82 treatment conditions and covered a total of 3,969 participants. None of the unpublished studies we found met our selection criteria. A table providing descriptive information on each included study can be requested from the corresponding author.

**Study Characteristics**

The 82 pharmacological treatment conditions include 17 hypnotic drugs (851 participants), 6 antidepressants (351 participants), 8 antiepileptics (349 participants), 7 benzodiazepine conditions (152 participants), 1 antihistamine condition (60 participants), 2 gamma-aminobutyric acid GABA receptor modulator conditions (105 participants), 1 hormone condition (20 participants), 3 melatonin receptor agonist conditions (433 participants), 1 narcotic condition (64 participants), 1 neuropeptide condition (8 participants), 1 progesterone receptor antagonist condition (5 participants), 2 valerian conditions (67 participants), and 32 placebo conditions (1,504 participants).

All studies were published between 1992 and 2012. The number of days of intervention ranges from 2 to 224 (mean $M = 31.72$, standard deviation $SD = 42.35$). The total number of patients across all studies was 3,969 with 2,465 patients in treatment and the remaining 1,504 in control groups. The

---

**Figure 1**—Flow diagram of the study selection process.
samples were predominantly female (63.23%). The average age of participants ranges from 35 to 72 (M = 51.24, SD = 12.15 for all patients, M = 51.58, SD = 11.46 for patients in placebo groups). The Jadad quality scores ranged from 2 to 5 points (out of a maximum of 5 points; M = 3.73, SD = 0.54). We used two independent quality ratings, with Cohen kappa interrater reliability of κ = 0.794.

Quantitative Data Synthesis

Table 1 shows that the pooled within-group effect sizes (Hedges g) of the placebo conditions for SOL (26 studies), TST (25 studies), WASO (22 studies), SE (25 studies), sSOL (12 studies), sTST (15 studies), sWASO (6 studies), sSE (4 studies), and sQuality (11 studies) were significant. According to Cohen’s interpretation recommendations, all effects were small-to-medium with confidence intervals suggesting small-medium-to-large effects for TST and sTST, respectively.

Sensitivity Analysis

Table 1 also illustrates that all fail-safe Ns (with the exception of sSE) exceeded 5K+10 and, accordingly, we considered these effect sizes to be robust regarding this analysis. Trim and Fill method results suggest that the effect size estimates for all considered outcome variables were unbiased.

Moderator Analysis

To take into account the variance of effect sizes from study to study (see Table 1) and to explore possible predictors of placebo treatment outcome, we conducted a moderator analysis for all pooled effect sizes. None of the chosen potential moderators (study quality, year of publication, duration of treatment, average age, and percentage of female participants) showed appreciable and significant moderation of the placebo treatment effect.

Comparison of Objective Outcomes with Subjective Outcomes

As Table 1 shows, the confidence intervals of objective and subjective outcomes overlapped in each comparison, and results from the Q tests for heterogeneity between subgroups yielded nonsignificant results from each comparison, which indicates no significant differences in the efficacy of improving insomnia between objective and subjective outcome measures.

Proportion of the Placebo Response to the Drug Response

Table 2 shows that subtracting the mean placebo response rates from mean drug response rates revealed that 39% (sSE) to 100% (sWASO) of the response to the medications under investigation are reported in the placebo group as well. In fact, one outcome variable (sWASO) placebo treatment was even more effective than the pharmacological therapy. The pooled proportion of the placebo response to the drug response was 63.56% (SD = 20.92).

DISCUSSION

Results indicated that the pooled effect sizes of placebo treatment for all outcome variables were small to medium, but significant and robust. Moreover, we detected no significant differences in the efficacy of placebo treatment between objective (PSG) and subjective (sleep diary and questionnaires) assessments. Thus placebo responses were also detectable in association with objective variables like the PSG parameters. With respect to the proportion of the placebo response to the drug response, our results reveal that 63.56% (SD = 20.92) of the response to the medications are achieved even in the placebo group.

The finding of a significant placebo response in pharmacological interventions for primary insomnia stands in line with Huedo-Medina, Kirsch reporting that the placebo response is a major contributor to the efficacy of nonbenzodiazepine hypnotics, but it ought to be generalized to all pharmacological
treatments for insomnia. It also supports findings by McCaUll, D'Agostino reporting a significant improvement in sSOL and sTST in the placebo groups of five drug trials and Belanger, Vallieres reporting significant improvements in 23 placebo conditions compared to seven waitlist conditions from different trials with respect to subjective parameters (sSOL and sTST).

The conclusion whether placebo responses were also detectable in association with objective variables was equivocal in earlier studies, with Huedo-Medina, Kirsch reporting significant effect sizes for both subjective and objective SOL, whereas McCaUll, D'Agostino did not report significant changes with respect to objective (polysomnographic) data. Belanger, Vallieres detected no significant group differences in their between-group comparison's objective data, although they did report a significant within-group improvement in subjective outcomes (sSOL, sTST, sTST, sSQ) and in objective outcomes (SOL, SE). These heterogeneous findings may be attributable to the limited number of studies included that assessed objective data in previous reviews.

Our results reinforce the evidence that placebo responses were also detectable in conjunction with objective variables—an important contribution to the current pool of evidence in placebo research, because most studies investigating placebo mechanisms have addressed placebo analgesia without having evaluated objective outcomes. Beyond the PSG parameters in insomnia research, there are few clinical examples (e.g., Parkinson disease and hypertension) enabling comparison of such a placebo response in objective and subjective data.

Our results indicate that 63.56% of the response to the medications examined may have been a placebo response. That is a key finding, because a great proportion of the therapeutic effect could also be achieved by optimizing placebo mechanisms. Regression to the mean, expectancy, social desirability, actual ingestion of the inert pill, the Hawthorne effect, cognitive dissonance, participation in research, and physiologic changes produced by placebos are discussed as contributors to the placebo response. In their review on the placebo response in medicine, Enck, Bingel reported several strategies to optimize placebo responses via the management of patients' expectations, the use of conditioning strategies (e.g., placebo-controlled dose reduction), and improving the physician-patient relationship. Against the background of our results, those strategies may also improve outcomes in the treatment of primary insomnia.

Nevertheless, a number of limitations should be noted. Examining intragroup changes in our analysis may have led to a biased estimate of the effect size due to additional influences such as natural history and regression to the mean. However, natural history seems less likely in the case of primary insomnia, because insomnia symptoms tend to become chronic. Furthermore, there is a lack of studies including both a placebo and a waitlist condition in the same trial. Therefore, limiting our analysis to intergroup comparisons would have ruled out all the studies we included, making it impossible to determine the current state of evidence.

To compute the proportion of placebo response to the drug response, we subtracted the placebo condition's pooled effect size from that of the drug condition. This approach depends on assuming the additivity of natural course effects, placebo effects, and drug effects, a model that is being increasingly questioned. Therefore, other options to analyze genuine placebo responses should also apply (e.g., "hidden application designs" and the further experimental manipulation of placebo mechanisms).

The methods we used to test the potential effect of publication bias are no equivalent alternative to including unpublished studies. Unfortunately, we were unable to find unpublished studies meeting our inclusion criteria even though we did an extensive and systematic search for unpublished data. However, in contrast to drug conditions, regarding placebo conditions in clinical trials it is highly unlikely whether a publication bias exists, as that would mean that studies with larger effect sizes in the placebo condition tend to be published more frequently.

We used the same assessment periods for objective (PSG) and subjective sleep parameters whenever the same time points were reported in the primary literature. In a minority of studies subjective estimates were derived from different and longer assessment periods (up to 2 w before the intervention to 2-w follow-up) than objective estimates limiting their comparability.

The strength of our study is a comprehensive search of the literature to preventively minimize publication bias. In comparison with previous reviews, we could identify several

### Table 2—Pooled Within-Group Effect Sizes for Drug Treatment and Proportion of Placebo Response to Drug Response.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>k</th>
<th>g</th>
<th>95% CI</th>
<th>P</th>
<th>Placebo response (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Objective outcomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOL</td>
<td>41</td>
<td>-0.55**</td>
<td>-0.63, -0.46</td>
<td>&lt; 0.001</td>
<td>64</td>
</tr>
<tr>
<td>TST</td>
<td>38</td>
<td>0.79**</td>
<td>0.69, 0.88</td>
<td>&lt; 0.001</td>
<td>53</td>
</tr>
<tr>
<td>WASO</td>
<td>35</td>
<td>-0.55**</td>
<td>-0.66, -0.44</td>
<td>&lt; 0.001</td>
<td>53</td>
</tr>
<tr>
<td>SE</td>
<td>41</td>
<td>0.64**</td>
<td>0.55, 0.73</td>
<td>&lt; 0.001</td>
<td>48</td>
</tr>
<tr>
<td>Subjective outcomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sSOL</td>
<td>22</td>
<td>-0.45**</td>
<td>-0.57, -0.34</td>
<td>&lt; 0.001</td>
<td>64</td>
</tr>
<tr>
<td>sTST</td>
<td>25</td>
<td>0.54**</td>
<td>0.42, 0.65</td>
<td>&lt; 0.001</td>
<td>80</td>
</tr>
<tr>
<td>sWASO</td>
<td>11</td>
<td>-0.29**</td>
<td>-0.42, -0.16</td>
<td>&lt; 0.001</td>
<td>100</td>
</tr>
<tr>
<td>sSE</td>
<td>5</td>
<td>0.64**</td>
<td>0.39, 0.89</td>
<td>&lt; 0.001</td>
<td>39</td>
</tr>
<tr>
<td>sQuality</td>
<td>22</td>
<td>0.51**</td>
<td>0.37, 0.65</td>
<td>&lt; 0.001</td>
<td>61</td>
</tr>
</tbody>
</table>

* P < 0.05. ** P < 0.01. CI, confidence interval; k, number of treatment conditions in the analysis; Placebo Effect, proportion of placebo effect to post-pharmacological effect in percent; SE, sleep efficiency; SOL, sleep onset latency; sSE, subjective sleep efficiency; sQuality, subjective quality; sSOL, subjective sleep onset latency; sTST, subjective total sleep time; sWASO, subjective wake after sleep onset; Sub., subjective; TST, total sleep time; WASO, wake after sleep onset.
additional treatment studies, especially additional studies assessing objective (PSG) outcomes. This enabled us to investigate whether placebo responses were also detectable in objective variables and to compare objective and subjective data based on an adequate sample of studies. Additionally, we were able to determine the proportion of the placebo response to the drug response to pharmaceuticals in different drug classes.

To conclude, further research on insomnia treatment should retain placebo control conditions and add waitlist conditions in the same clinical trial. Further studies on placebo mechanisms should utilize options independent from the assumption of additivity to analyze the proportion of placebo response to drug response. Most notably, attempts should be undertaken to exploit placebo mechanisms in clinical practice.

**DISCLOSURE STATEMENT**

This was not an industry supported study. The study was prepared in the context of the FOR1328 research unit on placebo and nocebo mechanisms and was supported by a grant from the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG). The authors have indicated no financial conflicts of interest. The study was prepared in the context of the FOR1328 research unit on placebo and nocebo mechanisms and was supported by a grant from the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG). A Winkler and W Rief have no conflicts of interest including any financial, personal or other relationships with other people or organizations to declare that could inappropriately influence, or be perceived to influence, the present work. This study did not require ethics approval.

**REFERENCES**


44. Roth T, Wright KP, Walsh J. Effect of tiagabine on sleep in elderly primary insomnia patients with EVT 201 improves sleep initiation, sleep maintenance, and daytime sleepiness. Sleep Med 2010;11:23–30.


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 \(S_{\text{within}}\) is:

\[
S_{\text{within}} = \frac{\sqrt{SD_1^2 + SD_2^2 - 2r \times SD_1 \times SD_2}}{\sqrt{2(1 - r)}}
\]

where \(SD_1\) is the standard deviation of the pretreatment sample mean, \(SD_2\) 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:

\[
d = \frac{Y_1}{SD_1 \sqrt{2(1 - r)}}
\]

where \(Y_1\) is the given paired difference in means, \(SD_1\) 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:

\[
d = Y_1 \times \sqrt{2} \times (1 - R)
\]

where \(Y_1\) 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:

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

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:

\[
C = \frac{\sum W_i - \sum W_i^2}{\sum W_i}
\]

\(F^2\) is determined by using the following formula:

\[
F^2 = \left(\frac{Q - df}{Q}\right) 	imes 100\%
\]

\(F^2\) 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:

\[
X = \frac{K (KZ^2 - 2.706)}{2.706}
\]

where \(K\) is the number of studies in the meta-analysis and \(Z\) 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 5\(K + 10\).\(^2\)

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