

## PLACEBO EFFECTS ON PSG PARAMETERS IN PRIMARY INSOMNIA: A META-ANALYSIS

# 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, PSYINDEX, 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.<sup>1</sup>

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).<sup>2</sup> Because most primary studies applied diagnostic criteria based on DSM-IV-TR, we refer to DSM-IV-TR<sup>1</sup> instead of DSM-V<sup>3</sup> 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.<sup>4</sup> Although there is solid evidence of the efficacy of sleeping pills,<sup>5</sup> 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<sup>6</sup> 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.<sup>7–9</sup>

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.<sup>10</sup> 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.<sup>7,11,12</sup> 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.

Submitted for publication July, 2014

Submitted in final revised form October, 2014

Accepted for publication October, 2014

Address correspondence to: Alexander Winkler, University of Marburg, Department for Clinical Psychology and Psychotherapy, Gutenbergstraße 18, 35032 Marburg, Germany; Tel: +49 6421 28 23740; Fax: +49 6421 28 28904; Email: Alexander.Winkler@staff.uni-marburg.de

## METHODS

For this meta-analysis we adhered to Meta-Analysis Reporting Standards (MARS) guidelines.<sup>13</sup> 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<sup>5</sup> to determine the efficacy of drug treatment of primary insomnia.

### Search Procedure

We identified studies by searching PubMed, PsycINFO, PSYINDEX, 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 *insomnia\** and *placebo\** combined with the term *polysomno\**.

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.<sup>14</sup> We therefore used the Jadad quality scale,<sup>15</sup> 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.”<sup>16</sup> We analyzed complete 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  $d$  that corrects for bias due to small sample sizes.<sup>17</sup> The magnitude of Hedges  $g$  can be interpreted using Cohen’s recommendation for small (0.20), medium (0.50), and large (0.80).<sup>18</sup> We followed Rosenthal’s recommendation<sup>19</sup> and used a conservative estimate of  $r = 0.70$  for the correlation between pretreatment and posttreatment measures.

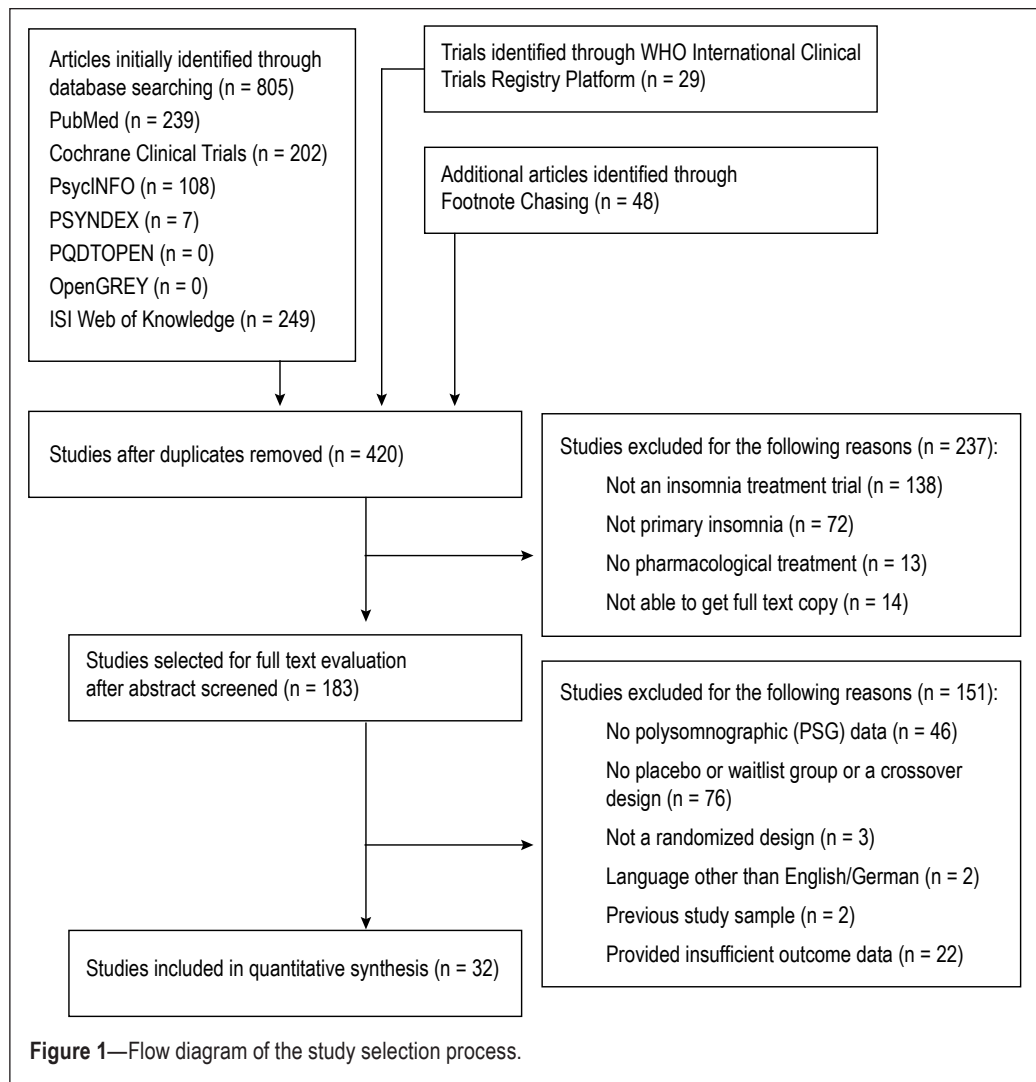
We used a test of significance based on the  $Q$  statistic to identify heterogeneity in effect sizes. Furthermore, we estimated the variance of the true effect between the studies ( $T^2$ ) to quantify heterogeneity in effect sizes. In addition, we used the ratio of true heterogeneity to total observed variation  $I^2$ .<sup>20</sup> These methods are described in more detail in Borenstein, Hedges<sup>21</sup>

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.<sup>22</sup> Instead of conducting a power analysis, we report the observed effect size with its confidence interval.<sup>21</sup> 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 Sarpstein<sup>23</sup> 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<sup>24</sup> resulted in statistically significant recalculated effect sizes, and the fail-safe  $N^{19}$  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<sup>14,25</sup> due to differences in methodological quality across studies, which is known in the literature as the garbage in/garbage out problem.<sup>21</sup> 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<sup>26–57</sup>

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 neuro-peptide 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

**Table 1—Pooled Within-Group Effect Sizes for Placebo Treatment.**

Outcome	k	g	95% CI	z	P	I <sup>2</sup>	Fail-Safe N	Obj. vs. Sub.	
								Q	P
Objective outcomes									
SOL	26	-0.35**	-0.42, -0.28	-9.69	< 0.001	52.52	1,324	0.832	0.362
TST	25	0.42**	0.29, 0.55	6.38	< 0.001	83.18	1,765	0.019	0.891
WASO	22	-0.29**	-0.38, -0.19	-5.67	< 0.001	71.87	636	0.209	0.648
SE	25	0.31**	0.19, 0.43	5.11	< 0.001	82.91	1,079	0.397	0.529
Subjective outcomes									
sSOL	12	-0.29**	-0.39, -0.19	-5.62	< 0.001	53.39	213	0.832	0.362
sTST	15	0.43**	0.34, 0.52	8.97	< 0.001	57.55	771	0.019	0.891
sWASO	6	-0.32**	-0.41, -0.22	-6.61	< 0.001	33.01	95	0.209	0.648
sSE	4	0.25**	0.11, 0.40	3.50	< 0.001	0.00	10	0.397	0.529
sQuality	11	0.31**	0.16, 0.46	4.07	< 0.001	79.79	289	-	-

\*P < 0.05. \*\*P < 0.01. CI, confidence interval; k, number of treatment conditions in the analysis; I<sup>2</sup>, ratio (0 to 100%) indicating the proportion of the observed variance that reflects real differences in effect sizes (values of 25%, 50%, and 75% can be considered low, moderate, and high, respectively); Fail-Safe N, indicates the number of studies with a treatment effect of 0 that would be needed to lead to a nonsignificant overall result; Obj., objective; Q, measure of heterogeneity to compare objective and subjective outcomes; SE, sleep efficiency; SOL, sleep onset latency; sQuality, subjective quality; sSE, subjective sleep efficiency; 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.

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<sup>58</sup> of  $\kappa = 0.794$ .

### 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-to-medium and 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<sup>7</sup> 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 McCall, D'Agostino<sup>11</sup> reporting a significant improvement in sSOL and sTST in the placebo groups of five drug trials and Belanger, Vallieres<sup>12</sup> 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<sup>7</sup> reporting significant effect sizes for both subjective and objective SOL, whereas McCall, D'Agostino<sup>11</sup> did not report significant changes with respect to objective (polysomnographic) data. Belanger, Vallieres<sup>12</sup> 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, sWASO, 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.<sup>59–61</sup>

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,<sup>62</sup> actual ingestion of the inert pill,<sup>63</sup> the Hawthorne effect, cognitive dissonance, participation in research, and physiologic changes produced by placebos<sup>64</sup> are discussed as contributors to the placebo response.<sup>62–64</sup> In their review on the placebo response in medicine, Enck, Bingel<sup>65</sup> reported several strategies to optimize placebo responses via the management of patients' expectations, the use of conditioning strategies (e.g., placebo-controlled dose reduction<sup>66</sup>), 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.<sup>67</sup> However,

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

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

\*P < 0.05. \*\*P < 0.01. CI, confidence interval; k, number of treatment conditions in the analysis; Placebo Response, proportion of placebo effect to pre-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.

natural history seems less likely in the case of primary insomnia, because insomnia symptoms tend to become chronic.<sup>68</sup> 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.<sup>69</sup> 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

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

1. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th edition, text revision. Washington, DC: American Psychiatric Association, 2000.
2. Roth T, Coulouvrat C, Hajak G, et al. Prevalence and perceived health associated with insomnia based on DSM-IV-TR; International Statistical Classification of Diseases and related health problems, Tenth Revision; and Research Diagnostic Criteria/International Classification of Sleep Disorders, Second Edition criteria: Results from the America Insomnia Survey. *Biol Psychiatry* 2011;69:592–600.
3. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, 5th edition. Arlington, VA: American Psychiatric Publishing, 2013.
4. Leger D, Poursain B, Neubauer D, Uchiyama M. An international survey of sleeping problems in the general population. *Curr Med Res Opin* 2008;24:307–17.
5. Winkler A, Auer C, Doering B, Rief W. Drug treatment of primary insomnia: a meta-analysis of polysomnographic randomized controlled trials. *CNS Drugs* 2014;28:799–816.
6. Buysse DJ. Insomnia. *JAMA* 2013;309:706–16.
7. Huedo-Medina TB, Kirsch I, Middlemass J, Klonizakis M, Siriwardena AN. Effectiveness of non-benzodiazepine hypnotics in treatment of adult insomnia: meta-analysis of data submitted to the Food and Drug Administration. *BMJ* 2012;345.
8. Siriwardena A, Apekey T, Tilling M, Dyas JV, Middleton H, Orner R. General practitioners' preferences for managing insomnia and opportunities for reducing hypnotic prescribing. *J Eval Clin Pract* 2010;16:731–7.
9. Glass J, Lancotot KL, Herrmann N, Sproule BA, Busto UE. Sedative hypnotics in older people with insomnia: meta-analysis of risks and benefits. *BMJ* 2005;331:1169–73.

10. Mora MS, Nestoriuc Y, Rief W. Lessons learned from placebo groups in antidepressant trials. *Philos Trans R Soc Lond B Biol Sci* 2011;366:1879–88.
11. McCall WV, D'Agostino R, Dunn A. A meta-analysis of sleep changes associated with placebo in hypnotic clinical trials. *Sleep Med* 2003;4:57–62.
12. Belanger L, Vallieres A, Ivers H, Moreau V, Lavigne G, Morin CM. Meta-analysis of sleep changes in control groups of insomnia treatment trials. *J Sleep Res* 2007;16:77–84.
13. JARS. Reporting standards for research in psychology: why do we need them? What might they be? *Am Psychol* 2008;63:839–51.
14. Glass G. Primary, secondary and meta-analysis of research. *Health Educ Res* 1976;5:3–8.
15. Jadad AR, Moore RA, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials* 1996;17:1–12.
16. Borenstein M, Hedges L, Higgins J, Rothstein H. Comprehensive meta-analysis, version 2. Engelwood, NJ: Biostat Inc., 2005.
17. Hedges LV, Olkin I. Nonparametric estimators of effect size in meta-analysis. *Psychol Bull* 1984;96:573–80.
18. Cohen J. Statistical power analysis for the behavioral sciences, 2nd edition. Hillsdale, NJ: Lawrence Erlbaum Associates Inc, 1988.
19. Rosenthal R. Meta-analytic procedures for social research. Newbury Park, CA: Sage Publications, 1993.
20. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;327:557–60.
21. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Introduction to meta-analysis. Chichester, UK: Wiley, 2009.
22. Hedges LV, Vevea JL. Fixed- and random-effects models in meta-analysis. *Psychol Methods* 1998;3:486–504.
23. Kirsch I, Sapirstein G. Listening to Prozac but hearing placebo: a meta-analysis of antidepressant medication. *Prevention & Treatment* 1998;1(2).
24. Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics* 2000;56:455–63.
25. Johnson BT, Eagly AH. Quantitative synthesis of social psychological research. In: Reis HT, Judd CM, eds. Handbook of research methods in social and personality psychology. London: Cambridge University Press, 2000:496–528.
26. Bes F, Hofman W, Schuur J, Van Boxtel C. Effects of delta sleep-inducing peptide on sleep of chronic insomniac patients. A double-blind study. *Neuropsychobiology* 1992;26:193–7.
27. Buckley T, Duggal V, Schatzberg AF. The acute and post-discontinuation effects of a glucocorticoid receptor (GR) antagonist probe on sleep and the HPA axis in chronic insomnia: a pilot study. *J Clin Sleep Med* 2008;4:235–41.
28. Fleming J, Moldofsky H, Walsh JK, Scharf M, Nino MG, Radonjic D. Comparison of the residual effects and efficacy of short term zolpidem, flurazepam and placebo in patients with chronic insomnia. *Clin Drug Invest* 1995;9:303–13.
29. Hajak G, Rodenbeck A, Voderholzer U, et al. Doxepin in the treatment of primary insomnia: a placebo-controlled, double-blind, polysomnographic study. *J Clin Psychiatry* 2001;62:453–63.
30. Herrmann WM, Kubicki ST, Boden S, Eich FX, Attali P, Coquelin JP. Pilot controlled double-blind study of the hypnotic effects of zolpidem in patients with chronic 'learned' insomnia: psychometric and polysomnographic evaluation. *J Intern Med Res* 1993;21:306–22.
31. Krystal AD, Durrence HH, Scharf M, et al. Efficacy and safety of doxepin 1 mg and 3 mg in a 12-week sleep laboratory and outpatient trial of elderly subjects with chronic primary insomnia. *Sleep* 2010;33:1553–61.
32. Krystal AD, Lankford A, Durrence HH, et al. Efficacy and safety of doxepin 3 and 6 mg in a 35-day sleep laboratory trial in adults with chronic primary insomnia. *Sleep* 2011;34:1433–42.
33. Luthringer R, Muzet M, Zisapel N, Staner L. The effect of prolonged-release melatonin on sleep measures and psychomotor performance in elderly patients with insomnia. *Int Clin Psychopharmacol* 2009;24:239–49.
34. Mayer G, Wang-Weigand S, Roth-Schechter B, Lehmann R, Staner C, Partinen M. Efficacy and safety of 6-month nightly ramelteon administration in adults with chronic primary insomnia. *Sleep* 2009;32:351–60.

35. McCall WV, Erman M, Krystal AD, et al. A polysomnography study of eszopiclone in elderly patients with insomnia. *Curr Med Res Opin* 2006;22:1633–42.
36. Monti JM, Alvarino F, Monti D. Conventional and power spectrum analysis of the effects of zolpidem on sleep EEG in patients with chronic primary insomnia. *Sleep* 2000;23:1075–84.
37. Monti JM, Attali P, Monti D, Zipfel A, de la Giclais B, Morselli PL. Zolpidem and rebound insomnia--a double-blind, controlled polysomnographic study in chronic insomniac patients. *Pharmacopsychiatry* 1994;27:166–75.
38. Monti JM, Monti D, Estevez F, Giusti M. Sleep in patients with chronic primary insomnia during long-term zolpidem administration and after its withdrawal. *International Clin Psychopharmacol* 1996;11:255–63.
39. Morin CM, Colecchi C, Stone J, Sood R, Brink D. Behavioral and pharmacological therapies for late-life insomnia: a randomized controlled trial. *JAMA* 1999;281:991–9.
40. Morin CM, Koetter U, Bastien C, Ware JC, Wooten V. Valerian-hops combination and diphenhydramine for treating insomnia: a randomized placebo-controlled clinical trial. *Sleep* 2005;28:1465–71.
41. Randall S, Roehrs TA, Roth T. Efficacy of eight months of nightly zolpidem: a prospective placebo-controlled study. *Sleep* 2012;35:1551–7.
42. Riemann D, Voderholzer U, Cohrs S, et al. Trimipramine in primary insomnia: results of a polysomnographic double-blind controlled study. *Trimipramin bei primärer Insomnie: Ergebnisse einer polysomnographischen kontrollierten Doppel-Blind-Studie*. *Pharmacopsychiatry* 2002;35:165–74.
43. Roth T, Soubirane C, Titeux L, Walsh JK. Efficacy and safety of zolpidem-MR: a double-blind, placebo-controlled study in adults with primary insomnia. *Sleep Med* 2006;7:397–406.
44. Roth T, Wright KP, Walsh J. Effect of tiagabine on sleep in elderly subjects with primary insomnia: a randomized, double-blind, placebo-controlled study. *Sleep* 2006;29:335–41.
45. Roth TG, Roehrs TA, Koshorek GL, Greenblatt DJ, Rosenthal LD. Hypnotic effects of low doses of quazepam in older insomniacs. *J Clin Psychopharmacol* 1997;17:401–6.
46. Scharf MB, Roth T, Vogel GW, Walsh JK. A multicenter, placebo-controlled study evaluating zolpidem in the treatment of chronic insomnia. *J Clin Psychiatry* 1994;55:192–9.
47. Schulz H, Stolz C, Muller J. The effect of valerian extract on sleep polygraphy in poor sleepers - a pilot-study. *Pharmacopsychiatry* 1994;27:147–51.
48. Sivertsen B, Omvik S, Pallesen S, et al. Cognitive behavioral therapy vs zopiclone for treatment of chronic primary insomnia in older adults: a randomized controlled trial. *JAMA* 2006;295:2851–8.
49. Walsh JK, Perlis M, Rosenthal M, Krystal A, Jiang J, Roth T. Tiagabine increases slow-wave sleep in a dose-dependent fashion without affecting traditional efficacy measures in adults with primary insomnia. *J Clin Sleep Med* 2006;2:35–41.
50. Walsh JK, Salkeld L, Knowles LJ, Tasker T, Hunneyball IM. Treatment of elderly primary insomnia patients with EVT 201 improves sleep initiation, sleep maintenance, and daytime sleepiness. *Sleep Med* 2010;11:23–30.
51. Walsh JK, Soubirane C, Roth T. Efficacy and safety of zolpidem extended release in elderly primary insomnia patients. *Am J Geriatr Psychiatry* 2008;16:44–57.
52. Walsh JK, Vogel GW, Scharf M, et al. A five week, polysomnographic assessment of zaleplon 10 mg for the treatment of primary insomnia. *Sleep Med* 2000;1:41–9.
53. Ware J, Walsh JK, Scharf MB, Roehrs T, Roth T, Vogel GW. Minimal rebound insomnia after treatment with 10-mg zolpidem. *Clin Neuropharmacol* 1997;20:116–25.
54. Wu RG, Bao JF, Zhang CA, Deng J, Long CL. Comparison of sleep condition and sleep-related psychological activity after cognitive-behavior and pharmacological therapy for chronic insomnia. *Psychother Psychosom* 2006;75:220–8.
55. Xu ZQ, Jiang XJ, Li W, Gao D, Li XJ, Liu J. Propofol-induced sleep: efficacy and safety in patients with refractory chronic primary insomnia. *Cell Biochem Biophys* 2011;60:161–6.
56. Zammit G, Erman M, Wang-Weigand S, Sainati S, Zhang J, Roth T. Evaluation of the efficacy and safety of ramelteon in subjects with chronic insomnia. *J Clin Sleep Med* 2007;3:495–504.
57. Zammit GK, McNabb LJ, Caron J, Amato DA, Roth T. Efficacy and safety of eszopiclone across 6-weeks of treatment for primary insomnia. *Curr Med Res Opin* 2004;20:1979–91.
58. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960;20:37–46.
59. de la Fuente-Fernández R, Ruth TJ, Sossi V, Schulzer M, Calne DB, Stoessl AJ. Expectation and dopamine release: mechanism of the placebo effect in Parkinson's disease. *Science* 2001;293:1164–6.
60. Lidstone SC, Schulzer M, Dinelle K, et al. Effects of expectation on placebo-induced dopamine release in Parkinson disease. *Arch Gen Psychiatry* 2010;67:857–65.
61. Preston RA, Materson BJ, Reda DJ, Williams DW. Placebo-associated blood pressure response and adverse effects in the treatment of hypertension: observations from a Department of Veterans Affairs Cooperative Study. *Arch Intern Med* 2000;160:1449–54.
62. McCall W, D'Agostino R Jr, Rosenquist PB, et al. Dissection of the factors driving the placebo effect in hypnotic treatment of depressed insomniacs. *Sleep Med* 2011;12:557–64.
63. McCall WV, Perlis ML, Xu T, Groman AE, Krystal A, Walsh JK. A comparison of placebo and no-treatment during a hypnotic clinical trial. *Int J Clin Pharmacol Ther* 2005;43:355–9.
64. Perlis ML, McCall WV, Jungquist CR, Pigeon WR, Matteson SE. Placebo effects in primary insomnia. *Sleep Med Rev* 2005;9:381–9.
65. Enck P, Bingel U, Schedlowski M, Rief W. The placebo response in medicine: minimize, maximize or personalize? *Nat Rev Drug Discov* 2013;12:191–204.
66. Doering BK, Rief W. Utilizing placebo mechanisms for dose reduction in pharmacotherapy. *Trends Pharmacol Sci* 2012;33:165–72.
67. Barnett AG, van der Pols JC, Dobson AJ. Regression to the mean: what it is and how to deal with it. *Int J Epidemiol* 2005;34:215–20.
68. Morin C, Belanger L, LeBlanc M, et al. The natural history of insomnia: a population-based 3-year longitudinal study. *Arch Intern Med* 2009;169:447–53.
69. Enck P, Klosterhalfen S. The placebo response in clinical trials-the current state of play. *Complement Ther Med* 2013;21:98–101.

**Detailed Information on Quantitative Data Synthesis and Moderator Analyses**

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

$$d = \left( \frac{\bar{Y}_1 - \bar{Y}_2}{S_{\text{within}}} \right)$$

where  $\bar{Y}_1$  is the pretreatment sample mean,  $\bar{Y}_2$  is the posttreatment sample mean, and  $S_{\text{within}}$ :

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

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:

$$d = \left( \frac{\bar{Y}_1}{\frac{SD_1}{\sqrt{2(1-r)}}} \right)$$

where  $\bar{Y}_1$  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:

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

where  $\bar{Y}_1$  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:

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

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:

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

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

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

$I^2$  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<sup>1</sup> 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:

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

where  $K$  is the number of studies in the meta-analysis and  $\bar{Z}$  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  $5K + 10$ .<sup>2</sup>

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**

1. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;327:557–60.
2. Rosenthal R. *Meta-analytic procedures for social research*. Newbury Park, CA: Sage Publications, 1993.