Published in final edited form as:
Sleep Med. 2011 June ; 12(6): 557–564. doi:10.1016/j.sleep.2011.03.008.

Dissection of the Factors Driving the Placebo Effect in Hypnotic Treatment of Depressed Insomniacs

W. Vaughn McCall[1,*], **Ralph D'Agostino Jr**[2], **Peter B. Rosenquist**[1], **James Kimball**[1], **Niki Boggs**[1], **Barbara Lasater**[1], and **Jill Blocker**[2]

[1]Department of Psychiatry and Behavioral Medicine; Wake Forest University School of Medicine

[2]Department of Public Health Sciences; Wake Forest University School of Medicine

Abstract

Objectives—Our prior work has shown that there is improvement in self-reported sleep in persons receiving placebo in hypnotic clinical trials. We examined the components of the 'placebo response' in a hypnotic clinical trial.

Methods—This was an exploratory analysis of a randomized, double-blind clinical trial of eszopiclone versus placebo in the treatment of persons with depression and insomnia who were also receiving fluoxetine at a clinic of a teaching hospital. Sixty adults with both depression and insomnia symptoms, who were free of significant primary sleep disorders, received open-label fluoxetine for 9 weeks. Patients were further randomized 1:1 to receive either masked eszopiclone 3 mg or placebo at bedtime after the first week of fluoxetine. We examined the respective contributions of three factors associated with the "placebo effect": (1) regression to the mean, (2) expectancy, and (3) social desirability.

Results—There was evidence for regression to the mean for the continuous measurement of the Insomnia Severity Index (ISI) and the Hamilton Depression Rating Scale. There was evidence for expectancy in self-reported Wake After Sleep Onset, continuous measurement of ISI, and dichotomous remission/non-remitter measurement of ISI. There was evidence of social desirability affecting self-reported Total Sleep Time.

Conclusions—Factors that have been associated with the "placebo effect" are operating in hypnotic clinical trials. However, the role of each factor differs depending upon which self-reported variable is being considered. The findings have implications for clinical trial design in insomnia.

Keywords

placebo effect; insomnia; depression; clinical trial; regression to the mean; expectancy; social desirability; eszopiclone

[*]Corresponding Author: W Vaughn McCall, Department of Psychiatry and Behavioral Medicine; Wake Forest University Health Sciences, Medical Center Blvd; Winston-Salem, NC 27157, 336-716-2911 (phone), 336-716-6830 (fax), vmccall@wfubmc.edu .

INTRODUCTION

The administration of a placebo, as either an inert substance, or a sham treatment, continues to be an integral part of clinical trials, and may have beneficial effects in routine clinical care.[1;2] Apart from the utility of understanding and controlling the placebo effect as part of clinical care, understanding and *making provisions for* the placebo effect could enhance the likelihood of finding new therapeutics. The difficulty in discerning drug/placebo differences is especially problematic as time progresses, because improvement with placebo continues to accrue over several weeks,[3] undermining the initial differences between treatment conditions. Indeed, only about 40% of new central nervous system drugs successfully make the transition from Phase II to Phase III testing, and only 40% of those compounds progress beyond Phase III.[4] The most common reason for failure of drug development is lack of proven efficacy.[5] Success in identifying the factors associated with the placebo effect, and controlling those factors, might inform new study designs that can better discern drug/placebo differences, resulting in fewer persons exposed to placebo, smaller sample sizes, reduced costs, faster recognition of effective treatments, and decreased likelihood of Type II error (failed detection of real treatment effects).

The understanding of the role of placebo response in insomnia treatment, and the factors controlling it, is poorly developed compared with our understanding of the role of placebo response in pain and depression. Indeed, as recently as a decade ago, some investigators argued that there was no placebo effect in insomnia clinical trials.[6-9] In response, we conducted a meta-analysis showing that indeed there was evidence for improvement in insomnia clinical trials with administration of placebo, especially for subjective reports of sleep latency (SL) and total sleep time (TST), and less so for polysomnographic (PSG) measurement.[10] We subsequently found that this effect builds over time during the course of treatment.[3] These findings were generally supported in a subsequent meta analysis by Belanger et al, showing change from baseline in patient-reported SL, Wake after Sleep Onset (WASO), and TST, and PSG improvement in SL and TST.[11] The insomnia placebo effect may be partly related to the actual ingestion of the inert pill, as we found that TST was greater on nights that primary insomniacs elected to take a pill in an as-needed dosing design.[3] Using PSG, others have found that in a randomized comparison of placebo pill versus no treatment that college students with minor sleeping problems experience less WASO when randomized to a placebo pill as compared with no pill.[12] Belanger et al also reported that there was an advantage for a placebo pill versus “no pill” condition for the self-reported SL and TST.[11] Interestingly, no such differences could be found for real versus placebo (sham) psychological treatment of insomnia,[11] reinforcing the idea that some parts of the placebo effect are indeed related to ingestion of the inert pill.

Some of the factors behind the placebo effect include regression to the mean, the Hawthorne effect, expectancy,[13] and social desirability. Regression to the mean refers to the likelihood that an outcome variable will show a significant change depending upon how much baseline values depart from the mean.[14-18] The more extreme the initial value, the more likely it is that subsequent measures will be less extreme, and that greater improvements are likely to occur over time. Regression to the mean might be expected to occur as part of the natural fluctuation of symptom severity, apart from the circumstances of clinical trial participation. Still, to the extent that regression to the mean occurs during administration of placebo in a clinical trial, it may be an aggravating factor in Type II error. The Hawthorne effect refers to the possibility that individual behaviors may be altered because the subjects under observation know they are being studied.[19-22] Expectancy refers to the participant’s belief that clinical improvement will occur, leading to biased reporting in favor of a response to treatment.[23;24] Expectancy can be experimentally manipulated. For example, the more expensive an experimental analgesic is reported to cost, the greater analgesia they are likely

to report from the pill,[25] and the degree of pre-treatment expectancy predicts the degree of reported pain relief in response to placebo.[26;27] The greater the likelihood that a research participant believes that they will be randomized to active antidepressant medication versus placebo (i.e., the greater the number the arms of active treatment), then the greater the antidepressant response rates are for both active treatments and placebo controls.[28] Borkovec and Nau demonstrated that research participants begin to make judgments about the likely efficacy of presumed effective and ineffective treatments before being exposed to these treatments.[29] Thus, expectancy most likely exerts a dose-response effect in clinical trials. Social desirability reflects the inclination of individuals to present themselves in a manner that will be viewed favorably by others.[30] High levels of social desirability may lead to biased reporting in favor of a treatment effect, even when a treatment effect does not exist, potentially narrowing the difference between active and control treatments and leading to Type II error.

The role of regression to the mean, expectancy, and social desirability has never been systematically evaluated in understanding treatment response in hypnotic clinical trials. We recently completed a clinical trial of hypnotic medication in depressed insomniacs, and we examined regression to the mean, expectancy, and social desirability as predictors of insomnia and depression response in both the active treatment and placebo arms. The Hawthorne effect was not measured, because the intensity and frequency of clinical assessments was the same for all patients in both treatment arms, and hence was assumed to result in an equally distributed sense of being observed.

METHODS

Overview

Full details of the main study have been described elsewhere.[31] Briefly, adult patients who simultaneously met criteria for insomnia[32] and Major Depressive Episode[33] were enrolled after providing written informed consent in an experiment approved by the local institutional review board. During the consent process, participants were told they had a 50% chance of receiving placebo versus a 50% chance of receiving eszopiclone at bedtime. Typical side effects of eszopiclone were described in the consent form (appendix).All participants reported a sleep latency of >30 minutes and sleep efficiency <85% at least 4 nights per week, or met RDC criteria for insomnia.[32] All participants demonstrated a Mini Mental State Exam (MMSE) score >24,[34] a 24-item Hamilton Rating Scale for Depression (HRSD) score > 20,[35] and an absence of significant primary sleep disorders, defined as an apnea-hypopnea index <15 and a periodic limb movement arousal index <15.[36] After completion of the screening PSG, participants began one week of open label fluoxetine (FLX), followed by 8 more weeks of open label fluoxetine coupled with 1:1 randomization to blinded eszopiclone (ESZ) 3 mg versus a matched placebo at bedtime. Assessments of self-reported sleep and mood were made at the initial visit, after the PSG, after the first week of open label FLX and at 1, 2, 4, 6, and 8 weeks into randomized treatment.

Measures

The dependent measures in this report are self-reported measures of sleep and mood during randomized treatment. Sleep diaries were collected each morning of study participation, and included reports of SL, WASO, and TST. Insomnia severity was also measured with the Insomnia Severity Index (ISI).[37] The ISI is a 7-item validated scale, with each item scored 0-4 for a maximum of 28 points. We defined remission of insomnia as an ISI score of ≤ 7. Depression was measured with the 24-item HRSD.[35] Response with the HRSD was defined as a decrease of HRSD of ≥ 50% at the final assessment, while remission was defined as a final HRSD ≤ 7.

The independent measures in this study included pre-randomization standardized difference between the individual observations and the mean values of self-reported sleep and mood (to assess regression to the mean), expectancy, and social desirability. Both expectancy and social desirability were assessed one time prior to randomization.

Regression to the mean was measured by computing the standardized difference between each individual observation and the mean value for each sleep and mood variable as assessed immediately prior to randomization, and then again at the end of treatment, using the end of treatment means. For each participant and for each variable, the standardized difference before treatment was plotted against that same individuals standardized difference at the end of treatment, thus producing a Galton squeeze diagram.[38]. Expectancy was measured with an adaptation of the Therapy Evaluation Questionnaire (TEQ) including 5 items, each item scored 1-7 on a Likert scale, with higher scores indicating greater expectancy of a favorable outcome during the clinical trial. The original TEQ has been shown to have high internal consistency (Cronbach α =0.79) and discriminates between sham and active behavioral therapy for insomnia.[39] For the purposes of this study, we removed items which were specific to psychological interventions (appendix). Social desirability was measured with the Social Desirability Scale-16-item (SDS-16), which has a Cronbach's α of 0.80 and test-retest reliability of 0.80 across intervals from 2-6 weeks long.[40] The SDS-16 has high convergent validity with the Eysenck Personality Questionnaire (EPQ) Lie Scale and discriminant validity against the EPQ measures of neuroticism, extraversion, and psychoticism.[40] The SDS-16 has been shown to be equivalent across genders and age ranges up to 80 years old.[40]

Statistical Analysis

Descriptive statistics were calculated as means and standard deviations or standard errors of continuous variables and percentages and frequencies of discrete outcomes. Means of independent samples were tested with a Student t-test. Bivariate associations between continuous variables were calculated with Pearson's Correlation Coefficient r. Linear models of continuous variables (i.e., SL) which were assessed at multiple time points post-randomization were tested using generalized linear mixed models for repeated measures (Proc Mixed; SAS statistical software). Repeated measures models in which the responses were discrete and correlated were also analyzed using generalized mixed models for binomial data (Proc Glimmix; SAS). In these mixed models, participants were treated as random effects while fixed effects included treatment assignment, time, age, gender, randomization (pre-treatment) value of the outcome and the main variable of interest (TEQ or SD). Models of dichotomous outcomes (i.e., insomnia remission as defined as ISI $\leq$ 7) were tested with logistic regression using generalized estimating equations (Proc Genmod). These models included the same covariates as the mixed models. Multiple interactions were tested in each of the models; if the interaction was non-significant, the model was re-run excluding the interaction term. The interactions included the three combinations of treatment group by time, variable of interest (TEQ or SD) and baseline value. Collinearity was assessed with correlation matrices and variance inflation diagnostics from the regression models. All VIFs were < 1.15. Evidence for regression to the mean was assessed by visual inspection of Galton squeeze diagrams, using SGplots (SAS).[38] Galton squeeze diagrams plotted the value for each subject at the time of randomization on the left vertical axis of a box, with the values standardized to the standard deviation. The standardized end-of-treatment value for each subject was then calculated, and the values for each subject was regressed from pre-treatment to post-treatment and back again, with lines connecting the pre-treatment value on the left vertical axis and the post-treatment value in the right vertical axis. A convergence of the lines into a 'funnel' on the right hand side is evidence of regression to the mean. To further clarify, evidence for regression to the mean was taken

from visual inspection of the Galton squeeze diagrams, and *not* from significance testing of the pre-randomization value of the variables included as co-variates in the linear and logistic models. Significance was accepted at $p<0.05$, with all analyses performed using SAS version 9.2 (Cary, NC, USA).

RESULTS

Fifty-nine participants contributed to this study of placebo effects out of an original sample of 60 persons.[31] The sample was middle-aged and mostly female, with high baseline levels of depression symptoms severity and insomnia symptom severity at initial intake (Table 1). The TEQ and the SDS-16 scales were not significantly different in the two treatment groups, nor were they different by gender. Furthermore neither TEQ nor the SDS-16 scores were significantly related to initial SL, WASO, TST, ISI, or HRSD (all Pearson's $r -0.2<r<0.2$). Similarly, the TEQ and SDS-16 scales were not related ($r= -0.09$, $p>0.5$).

Sleep Latency (SL)

For the repeated measures model of SL that included TEQ, SL at the time of randomization was significantly associated with SL during the randomized period ($t=7.3$, $df=131$, $p<0.0001$), with higher values of SL at randomization associated with higher SL values during treatment. SL fell for all subjects across the period of randomized treatment ($t=-2.3$, $df=52$, $p<0.05$), with the ESZ group taking an average 18 minutes less time to fall asleep ($t= -3.1$, $df=131$, $p<0.005$). Inspection of the Galton Squeeze plots did not suggest regression to the mean for the sample as a whole, and ESZ and placebo showed similar patterns (Figure 1). TEQ was not a significant predictor. The model of SL that included SDS-16 (instead of the TEQ) produced very similar results, and SDS-16 was not a significant predictor of SL.

Wake After Sleep Onset (WASO)

For the repeated measures model of WASO that included TEQ, higher WASO values at the time of randomization were associated with higher WASO values during treatment ($t=5.49$, $df=131$, $p<0.0001$). WASO did not significantly change across the period of randomized treatment. Inspection of Galton squeeze plots did not suggest regression to the mean for the sample as a whole, and ESZ and placebo showed similar patterns (Figure 2). TEQ was a significant predictor of WASO during randomized treatment ($\beta= -2.26$, $SE=0.8$, $p<0.05$), with higher levels of expectancy associated with greater drops in WASO during treatment. The model of WASO that included SDS-16 (instead of the TEQ) produced very similar results, except that SDS-16 was not a significant predictor of SL.

Total Sleep Time (TST)

For the repeated measures models of TST that included TEQ or SDS-16, higher TST at the time of randomization was strongly related to higher TST during randomized treatment ($t=8.4$, $df=130$, $p<0.0001$). On average, TST increased for all subjects across the period of randomized treatment ($t=2.0$, $df=52$, $p<0.05$), although the increase was greater in the ESZ group ($t=2.0$, $df=130$, $p=0.05$) by an average of 27.5 minutes. Inspection of Galton squeeze plots did not suggest regression to the mean for the sample as a whole, and ESZ and placebo showed similar patterns (Figure 3). TEQ was not a significant predictor. SDS-16 was a significant predictor of TST during the randomized treatment period, with a significant interaction between TST at time of randomization and SDS-16 ($t=-3.2$, $df=86$, $p<0.005$). The interaction revealed that a high SDS-16 augmented increases in TST during randomized treatment only when baseline TST was very low.

Insomnia Severity Index (ISI)

For the repeated measures model of ISI that included TEQ, higher values of ISI at randomization were associated with higher ISI values during treatment ($t=2.1$, $df=141$, $p<0.05$). ISI fell for all subjects across the period of randomized treatment ($t=-4.2$, $df=54$, $p<0.0001$), although the drop was greater in the ESZ group ($t=-2.6$, $df=141$, $p<0.05$). The average advantage in the decrease in ISI over time for the ESZ group was approximately 3.6 compared with the placebo group. Inspection of Galton squeeze plots suggested the presence of regression to the mean for the sample as a whole, with similar effects for both the ESZ and placebo groups (Figure 4). Higher values of TEQ were associated with greater drops in ISI during randomized treatment ($\beta= -0.38$, $SE=0.1$, $p<0.01$). The repeated measures model of ISI that included SDS-16 (instead of the TEQ) produced very similar results, except SDS-16 was not a significant predictor of ISI. Examination of remission of insomnia (final ISI ≤ 7) with logistic regression demonstrated that ISI values before randomization did not predict remission status, but age was a predictor, with younger persons more likely to remit ($OR=0.87$, $\chi=7.5$, $df=1$, $p<0.01$). Also, there was a significant interaction between treatment assignment and TEQ ($OR=-0.54$, $\chi=5.1$, $df=1$, $p<0.05$), such that higher TEQ scores were associated with greater likelihood of insomnia remission in the placebo group, but not the ESZ group (Figure 5). No similar relationships were seen between SDS-16 and remission from insomnia.

Hamilton Rating Scale for Depression-24 item (HRSD-24)

For the repeated measures models of HRSD-24 that included TEQ or SDS-16, HRSD-24 values just prior to randomization were associated with higher HRSD-24 values during treatment ($t=4.8$, $df=139$, $p<0.0001$). HRSD-24 fell for all subjects across the period of randomized treatment ($t=-4.7$, $df=54$, $p<0.0001$), although the drop was greater in the ESZ group ($p<0.05$) with an average superior decrease of 3.3 over time. Inspection of Galton squeeze plots did suggest the presence of regression to the mean for the entire sample, with similar effects seen for both the ESZ and placebo groups (Figure 6). TEQ was not significant. The model of HRSD-24 that included SDS-16 (instead of the TEQ) produced very similar results, and SDS-16 was also not a significant predictor of HRSD-24. Consideration of the HRSD-24 as a dichotomous outcome of either response or remission still failed to show that either TEQ or SDS-16 were significant predictors.

DISCUSSION

This secondary analysis suggests that several factors implicated with the “placebo effect” are operating in insomnia and depression outcomes in this clinical trial. Of note, expectancy and social desirability appeared to represent separate constructs as there was minimal correlation between them. Further, although expectancy might have been predicted to vary inversely with severity of depression symptoms, this was not the case, supporting the idea that expectancy is separate from mood state. Overall, results showed evidence for the action of all three factors: regression to the mean, expectancy, and social desirability. However, the relevance of each factor varied with the specific outcome under consideration. Significant findings were in the expected direction. The finding that regression to the mean was present for the ISI, but not present for SL, WASO, or TST, might be explained by the study-entry requirements for a baseline reported sleep latency > 30 minutes and sleep efficiency <85%. In other words, the range of values of baseline SL, WASO, and TST were constrained and therefore might reduce regression to the mean. No such constraints were present for ISI at baseline.

Of note, these “placebo effect” factors were seen in both the ESZ and the placebo groups, raising questions about the appropriateness of calling these “placebo effects” as opposed to

"non-specific effects." The sole exception to this observation is for the relationship of expectancy to ISI-remission, where were found an effect that was indeed specific to the placebo group. Remarkably, similar finds were reported by Manber et al, who found in a cognitive behavior therapy for insomnia (CBT-I) trial for insomnia in depressed patients receiving escitalopram that higher degrees of pretreatment expectancy boosted depression remission rates in the control group, but not the CBT-I group.[41] Further, they found that across randomized treatment groups, greater improvement in ISI scores were related to higher pretreatment expectancy (Pearson's r=0.49). Our findings and those of Manber et al should be re-confirmed, as they have implications both for clinical trials and clinical care.

Implications for Clinical Trials

As already noted, failure to demonstrate efficacy is the main reason that new therapeutics do not succeed in pushing through the process of development towards availability for patients.[5] Excessive placebo responding may explain why many drugs fail to separate from placebo.[42-45] Success in taming the placebo effect could improve the chances of useful therapeutics going forward. Our findings showed that regression to the mean is relevant for the ISI and HRSD as continuous measures. This implies that stipulating fixed-decreases in an outcome variable as an indicator of treatment success, such as a 6 point drop in the ISI,[46] could result in unexpectedly large placebo response if extreme baseline values were included in the selection of research participants. This problem could be addressed by stipulating both ceiling and floor values for the ISI during subject recruitment.

The problem of regression to the mean for ISI might be side-stepped by choosing an absolute end point (remission as defined by ISI ≤ 7) as the primary outcome measure. However, we found that expectancy had effects on WASO, ISI repeated measures, and ISI-remission status. Remarkably, expectancy had discernable effects for ISI remission status only for the placebo group, meaning that the presence of a high degree of expectancy would narrow the difference between drug and placebo and make drug efficacy more difficult to discern. Conceivably, a measure like the TEQ could be used as a recruitment screening tool to exclude persons with high TEQ scores who are likely to be placebo-remitters and thus enhance the chance for a positive clinical trial result. An alternative solution would be to stipulate the ISI as a secondary endpoint, rather than a primary endpoint in hypnotic trials. Expectancy in hypnotic trials could also be reduced by limiting the number of active treatment arms and by limiting the amount of positive information shared with the participants regarding the intervention, as both of these have been positively correlated with the degree of placebo responding.[26-28]

The finding of social desirability as a predictor of TST is difficult to interpret and does not lend itself to any easy recommendations about how to consider social desirability in a clinical trial.

Implications for Clinical Care

The application of our findings to a clinical trials setting suggests that attempts to modify patients' expectancy in a positive direction may improve the clinical outcome, at least in the short term. Also, as initial ISI severity was not related to final ISI remission status, clinicians need not worry that greater intensities of insomnia severity necessarily mean a lesser likelihood of achieving remission from insomnia.

Limitations

The study had a number of limitations, including small sample size, the exploratory nature of the hypotheses and analyses with multiple comparisons, and the inability to take into account other important factors in the "placebo effect," such as the Hawthorne effect. If

expectancy is related to the number of arms in a clinical trial, then given that this trial had only 2 arms, different results might have been seen with different numbers of arms. The sample included persons with both depression and insomnia, and these results might not apply to persons with primary insomnia. We chose not to examine the role of regression to the mean, expectancy, and social desirability in PSG variables, as we and others have found that, in general, PSG sleep measures do not change with placebo as much as self-reported measures.[10;11] A major limitation is our inclusion of regression to the mean as a "placebo effect," and we acknowledge that all of the factors we have described have the potential to occur with or without treatment, and with or without placebo, and hence could be argued to be more accurately termed as "non-specific effects" rather than placebo effects. As this study did not include a no-treatment arm, we are unable to disentangle how much of the effects we observed were placebo effects per se, versus non-specific effects.

In summary, "placebo effects" appear to be operating in this clinical trial of depressed insomniacs. A clearer understanding of these effects may help study design strategy in choosing the primary outcome variable least likely to be impacted by placebo effects, and designing exclusion criteria to identify persons with a high likelihood of selective responding to placebo.

Acknowledgments

NIH MH70821 and M01-RR07122, as well as funding and medications from Sepracor, and funding and material support from Mini Mitter

ClinicalTrials.gov Identifier: NCT00247624

Dr McCall is on the Speaker Bureau for Sepracor and Merck and has received research support from Sealy, Sepracor, Mini Mitter, and Corcept.

Dr Rosenquist has received research support from Cyberonics, Aspect Medical Systems, and Astra-Zeneca

Dr Kimball has received research support from Corcept, and is on the speaker's bureau for Astra Zeneca, Pfizer, and Merck

Reference List

1. Miller F, Rosenstein DL. Varience and Dissent: The Nature and Power of the Placebo Effect. J Clin Exp Neuropsychol. 2006; 59:331–335.
2. Beecher HK. Surgery as Placebo. A Quantitative Study of Bias. JAMA. 1961; 176:1102–1107. [PubMed: 13688614]
3. McCall V, Perlis M, Tu X, et al. A comparison of placebo and no-treatment during a hypnotic clinical trial. Int J Clin Pharm Ther. 2005; 43:355–359.
4. DiMasi J. New drug development in the United States from 1963 to 1999. Clin Pharmacol Ther. 2001; 69:286–296. [PubMed: 11371996]
5. DiMasi J. Risks in new drug development: Approval success rates for investigational drugs. Clin Pharmacol Ther. 2001; 69:297–307. [PubMed: 11371997]
6. 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; 62:453–463. [PubMed: 11465523]
7. Monti JM, Alvarino F, Cardinali D, et al. Polysomnographic study of the effect of melatonin on sleep in elderly patients with chronic primary insomnia. Archives of Gerontology and Geriatrics. 1999; 28:85–98. [PubMed: 15374088]
8. Bixler, EO.; Vgontzas, AN.; Kales, A. Pharmacology of Sleep. Springer; Berlin: 1995. Methodological issues in pharmacological studies of sleep; p. 323-343.
9. Adam K, Adamson L, Brezinova V, et al. Do placebos alter sleep? Br Med J. 1976; 1:195–196. [PubMed: 1247770]

10. McCall V, D'Agostino RB Jr. Dunn A. A meta analysis of sleep changes associated with placebo in hypnotic clinical trials. Sleep Medicine. 2003; 4:57–62. [PubMed: 14592361]
11. Belanger L, Vallieres A, Ivers H, et al. Meta-analysis of sleep changes in control groups of insomnia treatment trials. J Sleep Res. 2007; 16:77–84. [PubMed: 17309766]
12. Fratello F, Curcio G, Ferrar M, et al. Can an inert sleeping pill affect sleep? Effect on polysomnographic, behavioral and subjective measures. Psychopharmacology. 2005; 181:761–770. [PubMed: 15986193]
13. Perlis M, McCall V, Jungquist C, et al. Placebo effects in primary insomnia. Sleep Medicine Reviews. 2005; 9:381–389. [PubMed: 16109495]
14. Hick J, Feldman S. Eligibility creep: A cause for placebo group improvement in controlled trials of psoriasis treatments. J Am Acad Dermatol. 2007; 57:972–976. [PubMed: 17884244]
15. Haeckel R, Wosniok W. The phenomenon of regression-toward-the-mean. Sleep. 2002; 25:A240.
16. Rejeski W, Brawley L, Norris J. When study designs control for regression to the mean. Medicine and Science in Sports and Exercise. 2003; 35:887.
17. Shephard R. A threat to exercise science? Sports Medicine. 2003; 33:575–584. [PubMed: 12797839]
18. Zhang X, Tomblin J. Explaining and controlling regression to the mean in longitudinal research designs. Journal of Speech Language and Hearing Research. 2003; 46:1340–1351.
19. Adair J. The Hawthorne Effect - A Reconsideration of the Methodological Artifact. Journal of Applied Psychology. 1984; 69:334–345.
20. Diaper G. The Hawthorne Effect - A Fresh Examination. Educational Studies. 1990; 16:261–267.
21. Levine R, Cohen E. Hawthorne Effect. Educational Studies. 1974; 22:111–112.
22. Wickstrom G, Bendix T. The "Hawthorne effect" - What did the original Hawthorne studies actually show? Scandinavian Journal of Work Environment & Health. 2000; 26:363–367.
23. Bootzin, R. The role of expectancy in behavior change. In: White, L., editor. Clinical Phenomenon and New Insights. Guilford Press; 1985. p. 196-210.SGTB
24. Lick J, Bootzin R. Expectancy, demand characteristics, and contact desensitization in behavior change. Behavior Therapy. 1970; 1:176–183.
25. Waber RL, Shiv B, Carmon Z. Commercial Features of Placebo and Therapeutic Efficacy. JAMA. 2008; 299:1016–1017. [PubMed: 18319411]
26. Scott DJ, Stohler CS, Egnatuk CM, et al. Placebo and Nocebo Effects are Defined by Opposite Opioid and Dopaminergic Responses. Arch Gen Psychiatry. 2008; 65:220–231. [PubMed: 18250260]
27. Pollo A, Amanzio M, Arslanian A, et al. Response Expectancies in Placebo Analgesia and Their Clinical Relevance. Pain. 2001; 93:77–84. [PubMed: 11406341]
28. Sinyor M, Levitt A, Cheung A, et al. Does inclusion of a placebo arm influence response to active antidepressant treatment in randomized controlled trials: Results from pooled and meta-analyses. Journal of Clinical Psychiatry. 2010; 71:270–279. [PubMed: 20122371]
29. Borkovec T, Nau SD. Credibility of analogue therapy rationales. Journal of Behavior Therapy and Experimental Psychiatry. 1972; 3:247–260.
30. Barger S. The Marlowe-Crowne affair: short forms, psychometric structure, and social desirability. J Pers Assess. 2002; 79:286–305. [PubMed: 12425392]
31. McCall WV, Blocker JN, D'Agostino RB Jr. et al. Treatment of insomnia in depressed insomniacs: effects on health-related quality of life, objective and self-reported sleep, and depression. Journal of Clinical Sleep Medicine. 2010; 6:322–329. [PubMed: 20726279]
32. Edinger JD, Bonnet MH, Bootzin RR, et al. Derivation of research diagnostic criteria for insomnia: Report of an American Academy of Sleep Medicine work group. Sleep. 2004; 27:1567–1596. [PubMed: 15683149]
33. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. Fourth ed.. Washington, DC: 1994.
34. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res. 1975; 12:189–198. [PubMed: 1202204]

35. Hamilton M. A rating scale for depression. J Neurol Neurosurg Psychiatry. 1960; 23:56–62. [PubMed: 14399272]
36. McCall WV, Kimball J, Boggs N, et al. Prevalence and prediction of primary sleep disorders in a clinical trial of depressed patients with insomnia. Journal of Clinical Sleep Medicine. 2009; 5:454–458. [PubMed: 19961031]
37. Bastien CH, Vallieres A, Morin CM. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. Sleep Medicine. 2001; 2:297–307. [PubMed: 11438246]
38. Burrell G, Thoemmes F, MacKinnon D. Visual displays of regression toward the Mean using SAS SGplot. Paper 271-2010 ed 2010.
39. Edinger JD, Wohlgemuth WK, Radtke RA, et al. Cognitive behavioral therapy for treatment of chronic primary insomnia: a randomized controlled trial. JAMA. 2001; 285:1856–1864. [PubMed: 11308399]
40. Stober J. The Social Desirability Scale-17 (SDS-17). European Journal of Psychological Assessment. 2001; 17:222–232.
41. Manber R, Edinger Jack D, Gress JL, et al. Cognitive behavioral therapy for insomnia enhances depression outcomes in patients with comorbid major depressive disorder and insomnia. Sleep. 2008; 31:489–495. [PubMed: 18457236]
42. Sysko R, Walsh T. A Systematic Review of Placebo Response in Studies of Bipolar Mania. J Clin Psychiatry. 2007; 68:1213–1217. [PubMed: 17854245]
43. Khan A, Kolts R, Rapaport M, et al. Magnitude of placebo response and drug-placebo differences across psychiatric disorders. Psychol Med. 2005; 35:743–749. [PubMed: 15918351]
44. Schweizer E, Rickels K. Placebo response in generalized anxiety: its effect on the outcome of clinical trials. J Clin Psychiatry. 1997; 58:30–38. [PubMed: 9363046]
45. Khan A, Detke M, Khan S, et al. Placebo response and antidepressant clinical trial outcome. J Nerv Ment Dis. 2003; 191:211–218. [PubMed: 12695731]
46. Yang M, Morin C, Schaefer K, et al. Interpreting score differences in the Insomnia Severity Index: using health-related outcomes to define the minimally important difference. Current Medical Research and Opinions. 2009; 25:2487–2494.

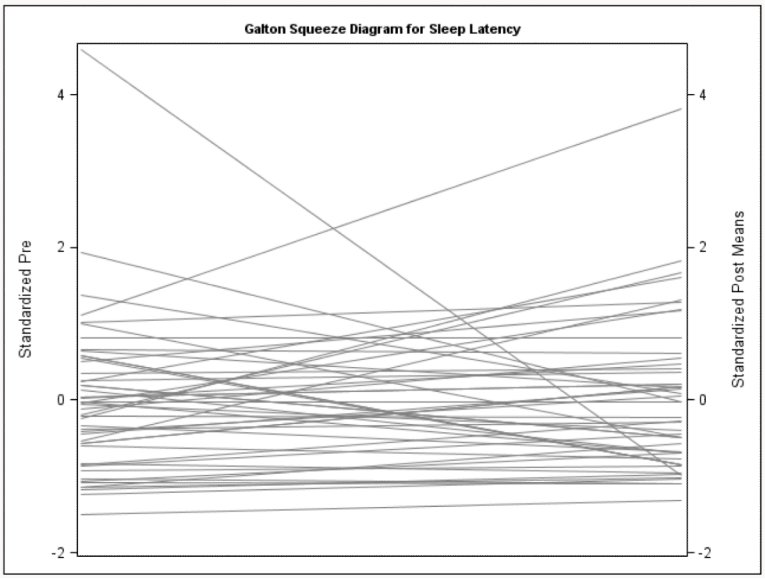

Figure 1.
scale on verticle axes are in Standard Deviation Units

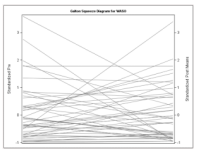

Figure 2.
scale on verticle axes are in Standard Deviation Units

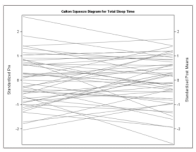

Figure 3.
scale on verticle axes are in Standard Deviation Units

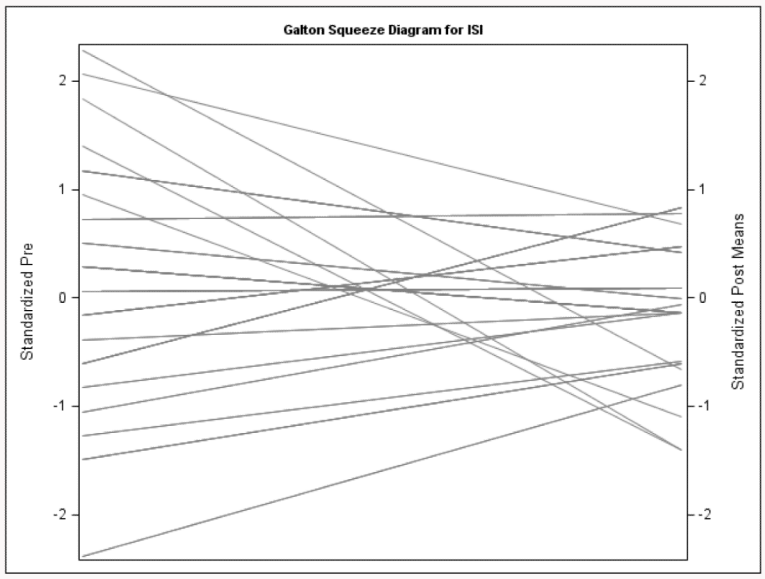

Figure 4.
scale on verticle axes are in Standard Deviation Units

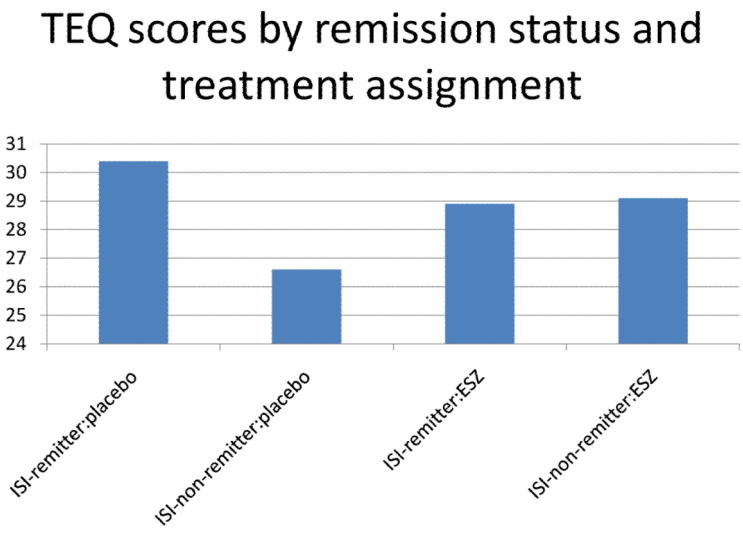

Figure 5.

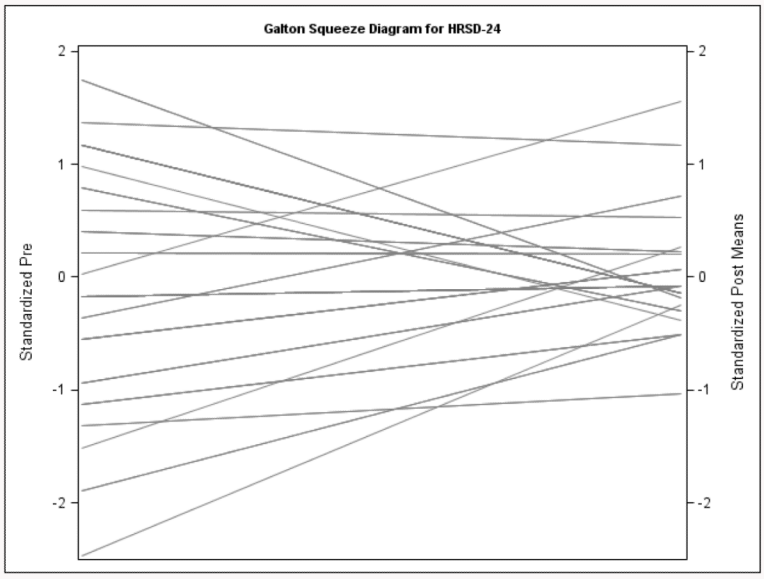

Figure 6.
scale on verticle axes are in Standard Deviation Units

Table 1

Demographics and Clinical Characteristics by Treatment Group
N (col %) or Mean [std]

	Total sample(N=59)	ESZ(N=29)	Placebo(N=30)
Gender			
Male	20 (33.9)	10 (34.4)	10 (33.3)
Female	39 (66.1)	19 (65.5)	20 (66.7)
Age	41.4 [12.6]	45.0 [12.0]	38.0 [12.5]
Body Mass Index (BMI)	27.7 [4.9]	26.8 [4.6]	28.6 [5.2]
Mini Mental State Exam	29.4 [0.9]	29.4 [0.8]	29.5 [1.0]
24-item HRSD at initial visit	27.1 [3.9]	27.2 [3.4]	26.9 [4.5]
24-item HRSD at randomization	22.0 [5.1]	21.8 [4.7]	22.2 [5.5]
ISI Total Score at initial visit	20.8 [4.0]	21.3 [3.8]	20.2 [4.1]
ISI Total Score at randomization	17.5 [4.7]	17.2 [5.4]	17.7 [3.9]
Sleep Latency at baseline	70.2 [41.0]	68.5 [50.5]	71.8 [30.3]
Sleep Latency at randomization	55.4 [39.8]	56.4 [48.9]	54.5 [29.3]
WASO at baseline	62.5 [55.7]	62.9 [47.2]	62.2 [63.4]
WASO at randomization	65.0 [63.3]	67.5 [70.7]	62.6 [56.2]
TST at baseline	354.4 [66.6]	342.1 [65.3]	366.1 [66.8]
TST at randomization	366.9 [77.4]	364.9 [85.9]	368.8 [69.6]
TEQ	28.2 [4.5]	30.0 [3.9]	27.4 [5.0]
SDS-16	7.7 [3.1]	8.2 [3.3]	7.2 [3.0]