Placebo Response of Non-Pharmacological and Pharmacological Trials in Major Depression: A Systematic Review and Meta-Analysis

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

Citation

Published Version
doi:10.1371/journal.pone.0004824

Citable link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:10246811

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA
Placebo Response of Non-Pharmacological and Pharmacological Trials in Major Depression: A Systematic Review and Meta-Analysis

André Russowsky Brunoni1,2, Mariana Lopes1, Ted J. Kaptchuk3, Felipe Fregni1*

1 Berenson-Allen Center for Noninvasive Brain Stimulation, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts, United States of America, 2 Department and Institute of Psychiatry, University of Sao Paulo, Sao Paulo, Brazil, 3 Osher Research Center, Harvard Medical School, Boston, Massachusetts, United States of America

Abstract

**Background:** Although meta-analyses have shown that placebo responses are large in Major Depressive Disorder (MDD) trials; the placebo response of devices such as repetitive transcranial magnetic stimulation (rTMS) has not been systematically assessed. We proposed to assess placebo responses in two categories of MDD trials: pharmacological (antidepressant drugs) and non-pharmacological (device- rTMS) trials.

**Methodology/Principal Findings:** We performed a systematic review and meta-analysis of the literature from April 2002 to April 2008, searching MEDLINE, Cochrane, Scielo and CRISP electronic databases and reference lists from retrieved studies and conference abstracts. We used the keywords placebo and depression and escitalopram for pharmacological studies; and transcranial magnetic stimulation and depression and sham for non-pharmacological studies. All randomized, double-blinded, placebo-controlled, parallel articles on major depressive disorder were included. Forty-one studies met our inclusion criteria - 29 in the rTMS arm and 12 in the escitalopram arm. We extracted the mean and standard values of depression scores in the placebo group of each study. Then, we calculated the pooled effect size for escitalopram and rTMS arm separately, using Cohen’s d as the measure of effect size. We found that placebo response are large for both escitalopram (Cohen’s d - random-effects model - 1.48; 95%CI 1.26 to 1.6) and rTMS studies (0.82; 95%CI 0.63 to 1).

**Conclusions/Significance:** We confirmed that placebo response in MDD is large regardless of the intervention and is associated with depression refractoriness and treatment combination (add-on rTMS studies). The magnitude of the placebo response seems to be related with study population and study design rather than the intervention itself.

Introduction

Placebo effect plays a significant role in clinical trials of major depressive disorder (MDD); in fact, two recent meta-analyses showed that the mean responder rates in the placebo group in antidepressant trials are 29.7% [1] and that drug-placebo differences might be relatively small in patients with MDD due to the large placebo response [2]. Given the importance of placebo response in MDD trials and the need to develop efficient research designs, it is critical to enhance our understanding on the placebo effects of distinct treatments such as repetitive transcranial magnetic stimulation (rTMS), a novel non-pharmacological intervention for neuropsychiatric diseases.

Several meta-analyses of rTMS clinical trials have been performed in the past ten years, initially showing negative or poor results [3,4]; although two recent studies have demonstrated a greater efficacy of the method [5,6]. However, taking into account the heterogeneity of rTMS trials and the lack of precise predictors of outcome, Herrmann and Ebmeier [7] proposed that non-specific contextual effects - such as the use of a new and relatively unknown technological device and the running of trials in major universities and teaching hospitals - play an important role in rTMS depression improvement. In fact, non-pharmacological treatments might have a large placebo response [8]. Furthermore, despite several meta-analyses assessing the placebo response of pharmacological trials in depression [1,2,9,10] placebo response of transcranial magnetic stimulation has not been sufficiently explored. We therefore decided to assess the placebo response of such intervention and perform an exploratory comparison with a non-pharmacological intervention through a systematic review and meta-analysis of recent clinical trials of major depression.

Aims of the study

This study sought to ascertain the magnitude of placebo response in controlled trials of rTMS and non-pharmacological...
studies using escitalopram as the antidepressant drug. Our secondary aim, given the limitations for such aim, was to
exploratory compare the effect sizes of placebo responses of
rTMS studies and pharmacological studies. The importance of our
study is contribute towards a better understanding of the placebo
effects mechanisms by comparing a traditional pill-taking medical
ritual to a new sham-device healing context.

Methods

We chose escitalopram to estimate the placebo response of
pharmacological treatment as several placebo-controlled trials
have been recently conducted and for non-pharmacological
treatment we chose rTMS as, similarly, several sham-controlled
studies have also been performed recently. We performed a
systematic review on all escitalopram and rTMS trials published
since 2002 and subsequently performed two main analyses: for the
placebo-drug response and for the sham-rTMS response. We then
compared the effect size of these groups. We also performed
exploratory analyses to assess predictors associated with placebo
response.

We choose this time period because the first escitalopram trial
was published in 2002 and we looked for concurrent rTMS and
escitalopram trials to make the studies more comparable
methodologically (i.e., with comparable sample sizes, diagnostic
definitions, rating methods and quality of studies) and also because a
meta-analysis performed in 2003 [3] stated that rTMS trials up
to 2002 had been of low quality.

Literature Search

We searched for published articles from April 2002 to April
2008 (period of 96 months) in the following databases: MEDLINE,
Web of Science, Cochrane, and SCIELO. We also examined
reference lists in systematic reviews and retrieved papers. To check
for unpublished trials, we: (i) consulted the CRISP database and
the websites clinicaltrials.gov and clinicalstudyresults.org; (ii)
contacted experts; (iii) searched for conference poster abstracts;
(iv) searched for studies in the monograph reference lists of
Lexapro® and; (iv) sent e-mails asking for unpublished studies to
Forest Labs and to Lundbeck S/A. Our key search terms were
“depression”, “escitalopram”, and “placebo” in the escitalopram
arm; and “depression”, “transcranial magnetic stimulation” and
“sham” in the rTMS arm.

Selection criteria

The following inclusion criteria were adopted: [i] manuscript
written in English (although there were no manuscripts in other
languages); (ii) randomized, double-blinded, placebo-controlled (or
sham-controlled), parallel studies on major depressive disorder; (iii)
mood effects assessed by a continuous mood scale, such as
Hamilton Depression Rating Scale (HDRS) or Montgomery-
Asberg Depression Rating Scale (MADRS); (iv) studies that
reported mean and standard deviation of the mood scales (or
provided other statistical parameters that could be used to deduce
this values) for the placebo group and; (v) studies published from
April 2002 to April 2008.

Data extraction

Data were extracted independently by the first author (AB) and
double-checked by the second author (ML), using a structured
form. The discrepancies were resolved by consensus and the
corresponding author (FF) consulted if needed. The following
variables were extracted: 1) mean and standard deviation values of
depression rating scales at baseline and end of treatment in active
(active group was used for exploratory analysis) and placebo/sham
groups and; 2) demographic, clinical and treatment characteristics
(e.g. number of patients, age, gender, previous use of medications,
depression-resistant subjects, duration of treatment, sham procedure
utilized).

When the study did not report mean and standard deviation
(SD) values, we either deduced them (using statistical parameters)
or contacted the corresponding author. Many escitalopram studies
did not report SD final scores – in these cases, we calculated SD
from standard error (SE) at end-of-treatment or from SD or SE
difference changes when possible [11,12]. In two studies, SD had
been only reported in graphs and we asked for data from Forest
Research Institute [13,14]. We also received data from Forest
Labs of two posters [15,16] and for an unpublished trial
mentioned in another study [13]. Two authors failed to provide
the requested data [17,18]; in these cases we had to input SD post-
treatment scores based on the mean of the available SD scores of
other trials, a method suggested by The Cochrane Collaboration
was to be applied in such cases [19]. In the rTMS arm, some trials just
reported data in graphs, while others did not report SD post-
treatment scores. We contacted the corresponding authors in these
cases [20,21,22,23]. Many rTMS trials also reported several
depression scores at different times using more than one
depression scale -in such cases, we extracted the data presented
by the authors as the main result. Finally, we used only unadjusted
rating scores in our analysis.

Quality assessment

We looked for the following biases: (1) selection bias - adequate
concealment of treatment (e.g., randomization was performed by
lottery and sealed, opaque envelopes were used); (2) performance
bias – if the study is single-blinded or double-blinded - for rTMS
studies we checked if they were single-blinded studies with external
blind raters and also if blinding of patients and physicians were
assessed; (3) attrition bias – if data are adequately reported in the
study, if there is evidence of intention-to-treat treatment, and if
methods used to handle with missing data (e.g., last observation
carried forward, complete case analysis) were reported.

Quantitative analysis

All of our analyses were performed using STATA statistical
software, version 9.0 (Statacorp, College Station, TX, USA). We
initially calculated the standardized mean difference and the
pooled standard deviation for each comparison –i.e. for each study
we calculated the change of either placebo or sham scores (baseline
minus post-treatment scores) and divided by the standard
deviation of change. We used Cohen’s d as a measure of the
effect size. Then, we measured the pooled weighted effect size
(weighted by the inverse variance of each study) using the random
and fixed effect models. We performed the analyses of placebo
response in escitalopram and rTMS trials separately and further
compared the pooled effect sizes. Heterogeneity was evaluated
with Chi-square test. We also performed sensitivity analysis,
cumulative regression and assessed publication bias using Beggs
modified funnel plot and Egger test [24] for each analysis.

Meta-regression was performed using the random-effects model
and tau² variance was calculated by the method of the residual
maximum likelihood. We tested the following variables: age
(years), gender (%females), duration of treatment (weeks), and
depression response in the active groups (Cohen’s d pooled effect
size of the active groups) – treated as continuous variables; sham
procedure, treatment resistant patients (defined as more than 50%
of patients failed at last two antidepressant treatments); drug-free
patients; and rTMS as an add-on therapy were treated as categorical
variables. It should be underscored that we classified as “angled coil” studies that described the use of an active rTMS coil in a different angle or position when applied to the scalp; whereas “sham coil” included studies that used a non-active coil associated with a method to preserve blindness (e.g. a study [25] described that sham stimulation was performed with “an identical coil (…) but without any electronic connection. This set-up had a similar sound effect but with no stimulation…”). Three studies used a different sham approach and were not pooled together in this analysis, because either a shielded coil [26] or a special coil generating a small field [23] were used.

Also, we considered as “add-on therapy” when a drug treatment was initiated simultaneously to active or sham rTMS, i.e., patients from sham group were actually starting an active drug treatment—in fact, this is the same concept of an “accelerating” study [27].

For baseline depression, we meta-regressed using either MADRS or HDRS baseline scores in escitalopram and rTMS trials, respectively. For rTMS studies that used MADRS scores as the primary outcome, we used the values of HDRS scores reported in secondary outcomes when this was possible [26,28,29,30,31]; in four studies this was not possible [20,32,33,34] and therefore we imputed missing HDRS scores regressing for other variables. Finally we assessed whether improvement in the active group was correlated with the placebo response – including this variable in our model.

Results

Using the keywords previously mentioned we were able to find 67 citations for escitalopram and 92 for rTMS studies. Only 12 and 29 studies met our inclusion criteria. Reasons for exclusion included: (1) reviews and meta-analyses; (2) studies that assessed other psychiatric diseases; (3) other studies designs (open-label, cross-over designs, quasi-randomized trials); (4) lack of sham or placebo group; (5) other topics. (Fig. 1)

Regarding study quality, all escitalopram studies are multicentric, randomized (although only one study reports the allocation method), double-blinded, and performed an intention-to-treat analysis (ITT), using the last observation carrier forward (LOCF) method. The quality of the rTMS studies is heterogeneous: all studies are randomized (thirteen studies report the allocation method); and single-blinded with external evaluation but only 8 studies addressed the integrity of blinding. Eighteen studies performed an intention-to-treat analysis, while 11 performed a complete-case analysis – mostly, exploratory studies. Finally, only two rTMS studies are multicentric. The quality assessment of each study is reported in Table S1.

The clinical characteristics of the 41 studies are summarized in Table 1. Tables 2 and 3 show characteristics of each study. There were 680 patients in sham group in the 29 rTMS studies (median per study = 16, interquartile range (IQR) = 10–26), while the 12 escitalopram studies enrolled 1714 patients in the placebo group (median per study = 133, IQR = 128–153). Also, all escitalopram studies enrolled non-treatment resistant patients who were drug-free, while most patients in rTMS studies were refractory and using antidepressant drugs – in fact, in 6 studies an antidepressant drug was initiated in both active and sham groups at the beginning of the trial. Conversely, the groups were comparable regarding age (50.7 vs. 43.1 years), gender (59% vs. 61% females) and baseline HDRS (24.73 vs. 21.4) and MADRS (33.1 vs. 29.23) scores.

Our main results show that the pooled effect sizes for placebo response in escitalopram trials are 1.46 (95% CI 1.38 to 1.53).
Using the fixed-effects model and 1.48 (95% CI 1.26 to 1.69) using the random-effects model; and, for rTMS studies, the sham pooled effect size is, in the fixed-effects model, 0.77 (95% CI 0.66 to 0.88) and 0.82 (95% CI 0.63 to 1) in the random-effects model (Fig. 2). Since heterogeneity is significant in both analyses ($\chi^2 = 86.54$, $p<0.001$ and $\chi^2 = 66.37$, $p<0.001$, respectively) subsequent analyses were performed using the random-effects model. For both arms, sensitivity analysis and Begg’s funnel plot show neither change in results after the exclusion of any particular study nor evidence of publication bias and systematic heterogeneity across the studies (Figures S1 and S2).

Subgroup analyses also show that the sham effect size of add-on rTMS studies (1.47, 95% CI 1.24–1.70) differ from studies not adopting such strategy (0.56, 95% CI 0.44–0.7, $p<0.001$), and of non-refractoriness studies (1.24, 95% CI 0.96–1.52) vs. studies with treatment-resistant patients (0.55, 95% CI 0.42 to 0.68, $p<0.001$) (Figure S3).

Exploratory analysis

Simple linear regressions (Table 4) show that some variables are associated with the outcome: (1) a negative association is observed for treatment resistant patients ($\beta$ coefficient = −0.69, $p<0.001$), meaning that refractoriness diminish placebo response; (2) a positive association ($\beta$ coefficient = 0.9, $p<0.001$) is observed for rTMS as add-on therapy, meaning that placebo response increases in accelerating studies and; (3) a positive association is observed for depression improvement in active groups for both escitalopram ($B = 0.7$, $p<0.0001$) and rTMS studies ($B = 0.33$, $p = 0.002$), i.e., studies showing a large depression improvement in active group also showed a large depression improvement in the control group.

On the contrary, baseline HDRS scores, baseline MADRS scores, depression scale utilized, gender and age are not associated with placebo response for both escitalopram and rTMS studies - except for baseline HDRS scores in rTMS studies ($B = 0.52$, $p = 0.04$), i.e., depression severity associates with a large placebo response.

Finally, variables that were associated with the outcome ($p<0.1$) were included in multiple linear regressions (Table 5). We observe that in models 1a (covariates: active-rTMS treatment and add-on therapy) and 1b (covariates: HDRS baseline scores and add-on therapy) only rTMS as add-on therapy remains associated with the outcome. Similarly, the variable baseline HDRS scores looses significance when meta-regressed together with treatment resistant (model 2b); however, in model 2a, both depression improvement in active-rTMS group ($B = 0.2$, $p = 0.02$) and treatment resistant ($B = −0.57$, $p<0.0001$) associates with placebo response. In model 3 all mentioned variables are regressed together; results show that only rTMS as add-on therapy still positively associates with the outcome ($B = 0.53$, $p = 0.02$), although there is still a trend for a negative association for treatment resistant variable ($B = −0.31$, $p = 0.08$), meaning that such variables still modify placebo response even when controlled by other significant variables.

Discussion

This meta-analysis includes data from 12 escitalopram and 29 rTMS trials, assessing 2394 subjects in placebo/sham groups. Our main result shows that placebo response is large in major depression trials, regardless of the placebo method. Exploratory analyses found that patients with severe depression and with treatment-resistant depression present a lower placebo response; while in trials that rTMS is initiated concomitantly with an antidepressant drug, the placebo response is larger.

The main finding of our study is that both placebo interventions are associated with a large effect size in major depressive disorder, which is in line with previous studies; Walsh et al. [1] reviewed 75 depression trials and concluded that placebo response is substantial and increasing over years; Stein et al. [35] in a pooled analysis of five escitalopram trials showed that placebo response ranged from 31.6% to 45.9%; and Kirsch et al. [2], reviewing 35 published and unpublished trials, showed that placebo response ranged from 0.7 to 1.1 Cohen’s d. Therefore our study confirms that placebo response is substantial in pharmacological and non-pharmacological trials in major depression.

Another finding of our study is that placebo-drug response appears to be larger than sham-rTMS response – even controlling for treatment refractoriness. Even considering that the small difference might not be meaningful, this finding is contrary to conventional wisdom that sham devices would have a higher

<table>
<thead>
<tr>
<th>Table 1. General characteristics of the studies.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>rTMS</td>
</tr>
<tr>
<td>Escitalopram</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Number of studies</td>
</tr>
<tr>
<td>Patients in active group</td>
</tr>
<tr>
<td>Patients in sham/placebo group</td>
</tr>
<tr>
<td>Age (mean ± SD) *</td>
</tr>
<tr>
<td>Gender (%female) *</td>
</tr>
<tr>
<td>Refractory to Antidepressants (%)</td>
</tr>
<tr>
<td>Concomitant antidepressant use in placebo group(%)</td>
</tr>
<tr>
<td>Studies that used HDRS in the primary outcome</td>
</tr>
<tr>
<td>HDRS baseline scores (mean ± SD) *</td>
</tr>
<tr>
<td>HDRS post-treatment scores (mean ± SD) *</td>
</tr>
<tr>
<td>Studies that used MADRS in the primary outcome</td>
</tr>
<tr>
<td>MADRS baseline scores (mean ± SD) *</td>
</tr>
<tr>
<td>MADRS post-treatment scores (mean ± SD) *</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>SD = standard deviation; MADRS = Montgomery-Asberg Depression Rating Scale; HDRS = Hamilton Depression Rating Scale; rTMS = repetitive transcranial magnetic stimulation. (*) In sham/placebo group.</td>
</tr>
<tr>
<td>doi:10.1371/journal.pone.0004824.t001</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
In fact, a non-pill intervention showed increased response than a placebo pill in a prospective sham device vs. inert pill trial \[37\] and in a meta-analysis comparing subcutaneous placebo with oral placebo from acute migraine \[38\]. Our finding does however agree with a smaller acute care study that found no difference between parenteral medication and oral medication \[39\]. These differences could be related to the concept that placebo response is very heterogeneous and influenced by many variables. In our study, this finding might be explained by several factors:

1. Study populations are different: 73% of rTMS trials enrolled refractory MDD patients; whereas no escitalopram trials enrolled refractory patients - in fact, STAR*D (Sequenced Treatment Alternatives to Relieve Depression) trial shows that remission rates decay at each time an antidepressant drug fail, being only 13% for refractory patients - i.e. patients who failed to remit after two trials \[40\]; and, since antidepressant drug effect is partially composed by a non-specific, placebo effect, placebo response might also decay in refractory patients.

2. Study designs are different: although escitalopram and rTMS trials present comparable quality, they mainly differ in blinding quality, as adequate blinding is more difficult to obtain in non-pharmacological interventions \[44\]. The rTMS trials assessed used an approach in which patients and raters were blinded to the treatment group allocated; however, it is possible that rTMS appliers unconsciously behave different when applying real and sham stimulation as well as that patients discover in which intervention they were allocated. Unsuccessful blinding biases the results as expectation effects and intervention confidence will be lost \[45,46,47,48\], therefore diminishing placebo response.

---

**Table 2. Characteristics of each rTMS study included.**

<table>
<thead>
<tr>
<th>Author and Year</th>
<th>Patients in sham group</th>
<th>Age (mean)</th>
<th>N of female</th>
<th>Depression scale</th>
<th>Baseline sham scores</th>
<th>Post-tto sham scores</th>
<th>Add-on therapy</th>
<th>Treatment resistant</th>
<th>Concomitant AD use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boutros [57]</td>
<td>9</td>
<td>49.5</td>
<td>1</td>
<td>HDRS</td>
<td>35.44</td>
<td>26.42</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Hoppner [28]</td>
<td>9</td>
<td>56</td>
<td>6</td>
<td>MADRS</td>
<td>37.5</td>
<td>29</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Fitzgerald [32]</td>
<td>20</td>
<td>49.15</td>
<td>11</td>
<td>MADRS</td>
<td>35.75</td>
<td>35.4</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Loo [33]</td>
<td>10</td>
<td>54.9</td>
<td>6</td>
<td>MADRS</td>
<td>33.1</td>
<td>27</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Herwig [21]</td>
<td>12</td>
<td>47.8</td>
<td>8</td>
<td>HDRS</td>
<td>23.1</td>
<td>14.5</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Koeseleman [58]</td>
<td>26</td>
<td>52</td>
<td>17</td>
<td>HDRS</td>
<td>25.9</td>
<td>20.2</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Poulet [20]</td>
<td>9</td>
<td>N/A</td>
<td>4</td>
<td>MADRS</td>
<td>36.22</td>
<td>18.125</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Holtzheimer [59]</td>
<td>7</td>
<td>45.4</td>
<td>4</td>
<td>HDRS</td>
<td>20.8</td>
<td>15.3</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Jorge [60]</td>
<td>10</td>
<td>66.5</td>
<td>5</td>
<td>HDRS</td>
<td>20.8</td>
<td>20</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Mosimann [61]</td>
<td>9</td>
<td>64.4</td>
<td>5</td>
<td>HDRS</td>
<td>24.5</td>
<td>20.4</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Hausmann [62]</td>
<td>13</td>
<td>47</td>
<td>9</td>
<td>HDRS</td>
<td>33.7</td>
<td>20.2</td>
<td>Yes</td>
<td>N/A</td>
<td>Yes</td>
</tr>
<tr>
<td>Rossini (J Clin Psych) [63]</td>
<td>49</td>
<td>46.4</td>
<td>40</td>
<td>HDRS</td>
<td>25.1</td>
<td>16.8</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Rumi [64]</td>
<td>24</td>
<td>38.9</td>
<td>20</td>
<td>HDRS</td>
<td>29.71</td>
<td>22.1</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Rossini (Psych Res), [22]</td>
<td>17</td>
<td>56.3</td>
<td>11</td>
<td>HDRS</td>
<td>28.7</td>
<td>24.9</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Su [65]</td>
<td>10</td>
<td>42.6</td>
<td>7</td>
<td>HDRS</td>
<td>22.7</td>
<td>19</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Januel [66]</td>
<td>16</td>
<td>37.19</td>
<td>12</td>
<td>HDRS</td>
<td>22.5</td>
<td>16.69</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Fitzgerald [31]</td>
<td>27</td>
<td>43.7</td>
<td>16</td>
<td>MADRS</td>
<td>34</td>
<td>30.9</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Avery [67]</td>
<td>33</td>
<td>44.2</td>
<td>16</td>
<td>HDRS</td>
<td>23.5</td>
<td>20</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Garcia-Toro [68]</td>
<td>10</td>
<td>47.2</td>
<td>7</td>
<td>HDRS</td>
<td>25.1</td>
<td>23.67</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Loo [29]</td>
<td>19</td>
<td>45.7</td>
<td>8</td>
<td>MADRS</td>
<td>32.6</td>
<td>27.1</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Herwig [30]</td>
<td>65</td>
<td>49</td>
<td>32</td>
<td>MADRS</td>
<td>27.1</td>
<td>16.3</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>O’Reardon [26]</td>
<td>146</td>
<td>48.7</td>
<td>74</td>
<td>MADRS</td>
<td>33.9</td>
<td>30</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Anderson [34]</td>
<td>16</td>
<td>48</td>
<td>9</td>
<td>MADRS</td>
<td>27.7</td>
<td>21.9</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Stern [69]</td>
<td>10</td>
<td>53.3</td>
<td>6</td>
<td>HDRS</td>
<td>27.4</td>
<td>26.7</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Botolomasi [70]</td>
<td>7</td>
<td>55.6</td>
<td>4</td>
<td>HDRS</td>
<td>22</td>
<td>19</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Jorge1 [23]</td>
<td>15</td>
<td>66.1</td>
<td>8</td>
<td>HDRS</td>
<td>19.9</td>
<td>16.8</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Jorge2 [23]</td>
<td>29</td>
<td>62.1</td>
<td>17</td>
<td>HDRS</td>
<td>17.6</td>
<td>14.8</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Mogg [71]</td>
<td>30</td>
<td>52</td>
<td>21</td>
<td>HDRS</td>
<td>21.6</td>
<td>19.4</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Bretlau [25]</td>
<td>23</td>
<td>57.8</td>
<td>13</td>
<td>HDRS</td>
<td>24.7</td>
<td>19.1</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

rTMS = repetitive transcranial magnetic stimulation; AD = antidepressant drug; HDRS = Hamilton Depression Rating Scale; MADRS = Montgomery-Asberg Depression Rating Scale; N/A = data not available.

doi:10.1371/journal.pone.0004824.t002

---
Table 3. Characteristics of each escitalopram study included.

<table>
<thead>
<tr>
<th>Author and Year</th>
<th>Patients in placebo group</th>
<th>Age (mean)</th>
<th>N of female</th>
<th>Depression scale</th>
<th>Baseline placebo scores</th>
<th>Post-tto placebo scores</th>
<th>Weeks of treatment</th>
<th>Treatment resistant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Burke [11]</td>
<td>119</td>
<td>40</td>
<td>71</td>
<td>MADRS</td>
<td>29.5</td>
<td>20.1</td>
<td>8</td>
<td>No</td>
</tr>
<tr>
<td>Wade [72]</td>
<td>189</td>
<td>40</td>
<td>147</td>
<td>MADRS</td>
<td>28.7</td>
<td>16.7</td>
<td>8</td>
<td>No</td>
</tr>
<tr>
<td>Lepola [73]</td>
<td>154</td>
<td>43</td>
<td>111</td>
<td>MADRS</td>
<td>28.7</td>
<td>16.2</td>
<td>8</td>
<td>No</td>
</tr>
<tr>
<td>Ninan [15]</td>
<td>153</td>
<td>39</td>
<td>99</td>
<td>MADRS</td>
<td>30.5</td>
<td>20.5</td>
<td>8</td>
<td>No</td>
</tr>
<tr>
<td>Rapaport † [13]</td>
<td>127</td>
<td>42.2</td>
<td>74</td>
<td>MADRS</td>
<td>28.8</td>
<td>17.5</td>
<td>8</td>
<td>No</td>
</tr>
<tr>
<td>Alexopoulos [16]</td>
<td>132</td>
<td>N/A</td>
<td>75</td>
<td>MADRS</td>
<td>30.7</td>
<td>18.4</td>
<td>8</td>
<td>No</td>
</tr>
<tr>
<td>Kasper [17]</td>
<td>180</td>
<td>75</td>
<td>137</td>
<td>MADRS</td>
<td>28.6</td>
<td>14.6</td>
<td>8</td>
<td>No</td>
</tr>
<tr>
<td>Clayton2 [12]</td>
<td>126</td>
<td>37</td>
<td>73</td>
<td>HDRS</td>
<td>23.3</td>
<td>11.4</td>
<td>8</td>
<td>No</td>
</tr>
<tr>
<td>Clayton1 [12]</td>
<td>130</td>
<td>35</td>
<td>81</td>
<td>HDRS</td>
<td>23.2</td>
<td>11.1</td>
<td>8</td>
<td>No</td>
</tr>
<tr>
<td>Wagner [74]</td>
<td>133</td>
<td>12.4</td>
<td>69</td>
<td>CDRS</td>
<td>56.6</td>
<td>36.4</td>
<td>8</td>
<td>No</td>
</tr>
<tr>
<td>Nierenberg [18]</td>
<td>137</td>
<td>42.5</td>
<td>49</td>
<td>HDRS</td>
<td>17.7</td>
<td>11.7</td>
<td>8</td>
<td>No</td>
</tr>
<tr>
<td>Bose [14]</td>
<td>134</td>
<td>68.5</td>
<td>79</td>
<td>MADRS</td>
<td>28.4</td>
<td>17.8</td>
<td>12</td>
<td>No</td>
</tr>
</tbody>
</table>

Data of unpublished study. HDRS = Hamilton Depression Rating Scale; MADRS = Montgomery-Asberg Depression Rating Scale; CDRS = Children’s Depression Rating Scale; N/A = data not available.

doi:10.1371/journal.pone.0004824.t003

Figure 2. Forest plots showing placebo response in control groups of escitalopram (A) and rTMS (B) studies. Forest plots show effect sizes from the random effects model. A negative effect indicates that endpoint depression scores in control groups are higher than baseline scores. Effect sizes are Cohen’s d (standardized mean difference), error bars represent the 95% confidence interval.

doi:10.1371/journal.pone.0004824.g002
that study designs influence outcomes, since Woods et al. [49] showed that, in schizophrenia controlled-trials, improvement was larger in trials having no placebo arm; and Trivedi et al. [50] showed that response raters were different in depression controlled-trials regarding using or not a placebo run-in phase.

(3) Study sites and approaches are different: whereas drug trials are conducted along 8 weeks, with weekly returns, tRMS trials are conducted in 2 to 4 weeks – therefore longer exposure might be associated with a larger placebo response. On the other hand, tRMS treatment is associated with an intensive 10-day treatment (as opposed to weekly or bi-weekly interaction in drug trials) and this could potentiate placebo response in the tRMS trials.

Our results show that sham-response is smaller in trials that tRMS is not used as add-on therapy (0.56 vs. 1.24), suggesting that such device might not be associated with a large placebo effect, a finding that was also observed in meta-analyses of Parkinson’s disease [51] and of refractory MDD [41]. Also, add-on tRMS trials improve response in placebo arm even when controlled for other variables, which could point to a synergistic effect between sham-tRMS and the drug, since there is no association between placebo response and previous use of antidepressant drugs. Finally, sham method (sham coil vs. angled coil) does not change placebo response – perhaps because both approaches, in fact, do not guarantee blinding.

Limitations
There was significant between-study heterogeneity in our meta-analysis, suggesting that the variation of effect size estimates in the studies were more than expected by chance. To address this limitation, we (1) used a random-effects model, which is a more conservative pooled analysis that take into account the between-study heterogeneity; (2) performed sensitivity analyses, to address whether the exclusion of an study could affect the pooled effect; (3) performed a conservative pooled analysis that take into account the between-study heterogeneity; (2) performed sensitivity analyses, to address whether the exclusion of an study could affect the pooled effect size; (3) assessed the quality of each study, looking for potential biases; and (4) checked for publication biases using Begg’s funnel plot. Another limitation is that, for pharmacological studies, we only included escitalopram studies; consequently, it is possible that the placebo response of other drugs is different. However, our study is in line with previous meta-analyses that showed similar placebo responses in major depression studies [1,52], assessed a significant number (1714) of patients and included unpublished studies; therefore this hypothesis is less likely.

Finally, it should be emphasized that the secondary analyses performed are exploratory and might be underpowered; in fact, since ten linear regressions have been performed in each pooled analysis, there is a 50% probability of observing one false positive association merely due to chance.

### Table 4. Meta regression results in which several variables were analyzed through simple linear regressions.

<table>
<thead>
<tr>
<th>Explanatory variables</th>
<th>Escitalopram</th>
<th>tRMS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>d.f.</td>
<td>Coef (B)</td>
</tr>
<tr>
<td>Baseline MADRS/HDRS</td>
<td>6</td>
<td>-0.12</td>
</tr>
<tr>
<td>Depression scale (HDRS vs MADRS)</td>
<td>10</td>
<td>-0.15</td>
</tr>
<tr>
<td>Gender (n Female)</td>
<td>10</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Age (years)</td>
<td>9</td>
<td>0.07</td>
</tr>
<tr>
<td>Patients using ADs (Y/N)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Tto Resistant (Y/N)</td>
<td>(*)</td>
<td>26</td>
</tr>
<tr>
<td>Week of post-tto scores</td>
<td>(-)</td>
<td>27</td>
</tr>
<tr>
<td>Sham method (Coil angled vs. Sham coil)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>rTMS as add-on therapy (Y/N)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Active Group change (Cohen’s d)</td>
<td>10</td>
<td>0.70</td>
</tr>
</tbody>
</table>

Coef. (B) is the regression coefficient of each regression, representing the slope of each model. Significant observations (p<0.05) are highlighted in bold. D.f. = degrees of freedom; HDRS = Hamilton Depression Rating Scale; MADRS = Montgomery-Asberg Depression Rating Scale; AD = antidepressant drug; MDD = Major Depressive Disorder; Y/N = yes or no. (*) There are no escitalopram studies that enrolled patients with refractory MDD. (**) All escitalopram studies except one assessed post-treatment scores at week 8.

doi:10.1371/journal.pone.0004824.t004

### Table 5. Exploratory regression models for tRMS studies in which significant results (obtained from simple linear regressions) were forced into several models.

<table>
<thead>
<tr>
<th>Regression model</th>
<th>Variable</th>
<th>D.f.</th>
<th>Coef (B)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1a</td>
<td>Active rTMS Group improvement</td>
<td>26</td>
<td>0.12</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>rTMS as add-on Therapy</td>
<td>0.77</td>
<td>&lt;0.0001</td>
<td></td>
</tr>
<tr>
<td>Model 1b</td>
<td>Baseline HDRS scores</td>
<td>26</td>
<td>0.15</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td>rTMS as add-on Therapy</td>
<td>0.86</td>
<td>&lt;0.0001</td>
<td></td>
</tr>
<tr>
<td>Model 2a</td>
<td>Treatment resistant</td>
<td>26</td>
<td>-0.57</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>Active rTMS Group improvement</td>
<td>0.20</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>Model 2b</td>
<td>Baseline HDRS scores</td>
<td>25</td>
<td>0.03</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>Treatment resistant</td>
<td>-0.49</td>
<td>&lt;0.0001</td>
<td></td>
</tr>
<tr>
<td>Model 3</td>
<td>Baseline HDRS scores</td>
<td>23</td>
<td>0.01</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>Treatment resistant</td>
<td>-0.31</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Active rTMS Group improvement</td>
<td>0.11</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td></td>
<td>rTMS as add-on Therapy</td>
<td>0.53</td>
<td>0.02</td>
<td></td>
</tr>
</tbody>
</table>

Coef. (B) is the regression coefficient of each regression, representing the slope of each model. Significant observations (p<0.05) are highlighted in bold. D.f. = degrees of freedom; HDRS = Hamilton Depression Rating Scale; rTMS = repetitive transcranial magnetic stimulation.

doi:10.1371/journal.pone.0004824.t005
Clinical implications

Because we addressed the influence of several variables in sham-response, our results have some implications for future rTMS trial designs, such as: (1) sham device method is not associated with placebo response; therefore this factor seems less relevant than currently considered by the researchers in this field [53]; (2) age and gender are probably not related with placebo response - although age seems to be related to depression response in some studies [54,55]; (3) refractoriness is associated with a lower placebo response – and, in fact, a lower depression response [41,54] [56]; perhaps indicating that such patients are very unresponsive to any intervention at all and therefore rTMS studies should focus on non-refractory patients or, on the contrary, the positive results of rTMS trials might be due to a lower placebo response that increases active-sham difference – therefore, future rTMS trials should quantify the degree of refractoriness of each patient, and; (4) placebo response is high in add-on rTMS trials – this could indicate there is a synergistic effect with the drug and, therefore, future trials could use a two-way factorial design (i.e., sham vs. real-rTMS and placebo vs. active drug) to address the relationship among rTMS and drug interventions.

Our study also stresses the heterogeneity of placebo response in different contexts and interventions; therefore, the lower placebo response observed in sham trials could be explored by using a qualitative approach to understand patient’s expectancies regarding rTMS intervention or, perhaps, by a sham-device vs. inert pill trial, in the same fashion of a prior placebo study [37].

Final remarks

In summary, our study shows that placebo response in rTMS and escitalopram trials is large and appears to be lower for rTMS trials. The sham response is negatively associated with refractoriness and positively associated with rTMS add-on studies; whereas sham method utilized, age and gender are not associated with a greater sham response. It is possible that design issues such as the lack of adequate blinding associate with lower placebo responses; however, we cannot measure in which extent such difference is explained by other cultural factors, as pill-taking healing is a mainstream medical ritual, while sham devices are not. The sham response of rTMS significantly varies among studies and can influence the results of a clinical trial as it will determine the effect size of a given sham-controlled trial, therefore, further studies are needed to explore its effects as to design appropriate sham-controlled randomized clinical trials.

Supporting Information

Table S1 The file contains the quality assessment of all the studies included. Found at: doi:10.1371/journal.pone.0004824.s001 (0.57 MB PPT)

Figure S1 (A) shows the sensitivity analysis, assessing the individual influence of a particular study by showing the resulting effect size and 95% confidence interval (CI) after its exclusion. (B) shows the funnel plot of the effect sizes (Cohen’s d) according to their standard errors. Cohen’s d is the standardized mean difference, error bars represent the 95% CI. Found at: doi:10.1371/journal.pone.0004824.s002 (3.00 MB TIF)

Figure S2 (A) shows the sensitivity analysis, assessing the individual influence of a particular study by showing the resulting effect size and 95% confidence interval (CI) after its exclusion. (B) shows the funnel plot of the effect sizes (Cohen’s d) according to their standard errors. Cohen’s d is the standardized mean difference, error bars represent the 95% CI. Found at: doi:10.1371/journal.pone.0004824.s003 (3.00 MB TIF)

Figure S3 (A) shows the influence of the variable add-on rTMS in the pooled analysis of the studies, by pooling together only studies in which this variable is present (top) or absent (bottom) and thereby comparing the resulting effect sizes (Cohen’s d, standardized mean difference). (B) shows the influence of the variable treatment-resistant depression, when it is present (top) or absent (bottom) in the resulting effect sizes. Found at: doi:10.1371/journal.pone.0004824.s004 (3.00 MB TIF)

Acknowledgments

We wish to thank Forest Research Institute and all contacted authors that kindly provided us the data to perform the present study.

Author Contributions

Conceived and designed the experiments: FF. Wrote the paper: ARB ML. TK FF.

References


