Predictors of Placebo Response in Pharmacological Clinical Trials of Negative Symptoms in Schizophrenia: A Meta-regression Analysis

David Fraguas*1, Covadonga M. Díaz-Caneja1, Laura Pina-Camacho1,2, Daniel Umbricht3, and Celso Arango1

1Department of Child and Adolescent Psychiatry, Hospital General Universitario Gregorio Marañón, IISGM, School of Medicine, Universidad Complutense, CIBERSAM. Madrid, Spain; 2Department of Child and Adolescent Psychiatry, Institute of Psychiatry, Psychology & Neuroscience, King’s College London, London, UK; 3Neuroscience, Ophthalmology, Rare Diseases, Roche Pharma Research & Early Development, Roche Innovation Center Basel, F. Hoffmann-La Roche Ltd, Basel, Switzerland

*To whom correspondence should be addressed; Department of Child and Adolescent Psychiatry, Hospital General Universitario Gregorio Marañón, Calle Ibiza 43, Madrid 28009, Spain; tel: +34-914265005, fax: +34-914265004, e-mail: davidfraguas@gmail.com

We conducted a meta-regression analysis of all double-blind, randomized, placebo-controlled clinical trials (DBRCTs) reporting effects of drug and placebo on negative symptoms in people with stable schizophrenia and predominant or prominent negative symptoms to assess predictors of placebo response in these individuals. We used Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guidelines for systematic reviews and meta-analyses to conduct a systematic literature search to identify DBRCTs assessing treatment efficacy on negative symptoms, as primary outcome, in patients with stable schizophrenia and predominant or prominent negative symptoms. We used Cohen’s $d$, with 95% CIs, as the effect size measure for placebo response, based on negative symptom change scores from baseline to endpoint (range 4 to 24 wk) in the placebo-treated group. We included 18 DBRCTs from 17 publications, assessing the effect of 13 drugs vs placebo on negative symptoms and comprising 998 patients, in the meta-regression analyses. Overall, drugs showed greater efficacy than placebo in reducing negative symptoms, with small effect size (Cohen’s $d$: 0.208, $P = .020$). Placebo response was significant ($P < .001$) and clinically relevant (Cohen’s $d$: 2.909), but there was significant heterogeneity and high risk of publication bias. Multivariable meta-regression analyses showed that larger numbers of arms in the trial, larger numbers of study sites and industry sponsorship were significant moderators of placebo response in this population. Our results suggest that some clinical trial design and operational factors affect the level of placebo response in such studies, thus highlighting the need for designs better suited to assess these outcomes.

Key words: psychosis/clinical trial/deficit syndrome/meta-analysis/predictive factors

Introduction

Schizophrenia is a heterogeneous psychiatric syndrome characterized by the presence of positive symptoms (hallucinations, delusions, incoherence of thought), negative symptoms (motivational impairment, social withdrawal, blunted affect), affective dysregulation (depression, mania) and cognitive impairment.1 Severe negative symptoms have greater impact on functioning2 than do positive symptoms, and constitute one of the main predictors of functional impairment and poor outcomes in people with schizophrenia.3

Despite the clinical relevance of negative symptoms, many pharmacological treatments so far have shown only statistically significant but not clinically meaningful efficacy in improving this symptomatology relative to placebo.4 Further, except for some studies with positive results in phase 2,5,6 most studies investigating new treatment targets, such as glutamatergic compounds, have shown no consistent evidence of negative symptom improvement,7 especially in patients with predominant or prominent negative symptoms of schizophrenia.8 Both predominant and prominent negative symptoms constructs identify people with schizophrenia with disabling negative symptoms. “Predominant negative symptoms” refers to cases with high negative symptoms, but mild and stable positive symptoms.9,10 “Prominent negative symptoms” (sometimes also called dominant negative symptoms) refers to cases of relatively high and persistent negative symptoms alongside substantial positive psychotic symptoms.10

A meta-analysis based on 89 samples from 41 longitudinal studies and including 5944 outpatient participants with schizophrenia found that negative symptoms tend to decrease during follow-up, regardless of the intervention...
type. Thus, the reported lack of efficacy of available drugs on negative symptoms in clinical trials could be the consequence of a relatively strong placebo response (defined as an improvement in symptoms in the patients treated with placebo) rather than because the drugs do not improve symptoms. This presumption is supported by findings of significantly increasing placebo responses in schizophrenia drug trials and a relatively stable drug response during the last few years. That large and increasing placebo response may have precluded finding significant drug-placebo differences in negative symptom trials.

Several meta-regression analyses have assessed various predictors of placebo response in trials of antipsychotics in acute and chronic schizophrenia, finding that younger age, shorter duration of illness, greater baseline symptom severity, shorter trial duration, a smaller probability of patients to be assigned to placebo, larger number of sites, greater percentage of female patients and more recent year of publication were significantly associated with greater placebo response. However, these analyses included heterogeneous samples of people with schizophrenia, mixed acute and stable phases and focused on positive or global symptoms. This precludes generalizing their results to samples of patients with stable schizophrenia and predominant/prominent negative symptoms.

Therefore, we conducted a meta-regression analysis that assesses predictive factors of placebo response of negative symptoms in people with stable schizophrenia and predominant or prominent negative symptoms.

Methods

Search Strategies

Using the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guidelines, we conducted a systematic 2-step literature search to identify relevant studies. First, we searched PubMed and TrialTrove to detect double-blind, randomized, placebo-controlled trials (DBRCTs) assessing, as their primary outcome, efficacy of drugs and placebo on negative symptoms in patients with stable schizophrenia and predominant/prominent negative symptoms. TrialTrove is a curated database that gathers information on registered and non-registered clinical trials from more than 40,000 clinical trial data sources including clinicaltrials.gov and other registries across 150 countries. The search dates were from January 2002 to January 24, 2017. We decided to limit the search to the past 15 years, since predominant/prominent negative symptoms only became a target of investigation in the early 2000s.

The initial search covered the combination of 2 concepts: “predominant negative symptoms” (OR “persistent negative symptoms” OR “prominent negative symptoms” OR “persisting negative symptoms” OR “enduring negative symptoms” OR “dominant negative symptoms” OR “deficit syndrome”) AND DBRCTs. Second, we manually checked the reference list of the selected articles for any studies not identified by the computerized literature search.

Selection Criteria

Figure 1 shows the flowchart of the systematic literature search strategy. The initial literature search yielded a total of 159 studies. After removing duplicates, we evaluated 148 potential studies. All papers were double-screened in 3 phases by 3 independent raters (DF, CDC, and LPC), all consultant psychiatrists with extensive experience designing and conducting DBRCTs in schizophrenia. We resolved discrepancies through discussion and consensus. In phase 1, we screened 148 records. We excluded studies in phase 2 if, based on the title and abstract, they met any of the following hierarchical criteria: (1) they did not include patients with a diagnosis of schizophrenia spectrum disorders (SSD) according to DSM-IV, DSM-IV-TR or DSM-5 criteria; (2) they did not follow a DBRCT methodology; (3) they did not assess the effect of one drug vs placebo on negative symptoms as primary outcome (either as monotherapy or as adjunctive treatment); and (4) they were not published in English as original peer-reviewed articles or had not been presented as a poster in scientific congress or as pilot studies, with a minimum of 5 subjects in both the drug and the placebo group. Of the 148 studies, 108 non-duplicate studies did not meet the inclusion criteria, leaving 40 for phase 2.

In phase 2, we obtained full text articles for comprehensive review. We retained studies that met the following hierarchical criteria: (1) they assessed a sample of patients with SSD with “predominant negative symptoms” (or any of the above-mentioned search terms); (2) they did not assess patients in the acute phase of the illness; (3) they reported data sufficient to obtain effect size (ES); (4) they included nonoverlapping samples. Criterion 4 ensured the independence of the samples in the meta-regression. The following hierarchical criteria determined inclusion of overlapping studies: (1) study with the largest patient sample; and (2) most recent publication. Of the 40 phase 2 studies, 17 publications reporting on 18 independent DBRCTs met the inclusion criteria. Therefore, we included a total of 17 publications reporting on 18 independent DBRCTs in the meta-analysis.

In phase 3, we assessed the quality of the included studies using the Cochrane Collaboration’s tool for assessing risk of bias. Quality assessment coded each study on a scale of 0 to 14 with higher values representing greater quality (details are given in supplementary table 1).
Data Extraction

The first author (D.F.) extracted the following data from each eligible study: first author; location and year of the study; study design; duration of trial; country or countries, number of countries and number of sites where the DBRCT was conducted; academic or industry sponsorship; name of the clinical scale used to assess negative symptoms; study quality; number of patients in each arm (drug and placebo); mean and SD of baseline and follow-up scores in the negative symptoms scales (or any other measures sufficient to obtain an ES of negative symptom change as proportion of maximum score) both for the active drug and placebo arms. Additionally, DF extracted putative moderator variables: mean age, sex (as percentage of female participants), diagnosis (as percentage of patients with a diagnosis of schizophrenia), substance use/abuse (as percentage), mean duration of illness, drop-outs (as percentage), severity of positive symptoms at baseline (as proportion of maximum possible score); severity of depressive symptoms at baseline (as proportion of maximum possible score); severity of extrapyramidal symptoms at baseline (as proportion of maximum possible score); medical comorbidity as an exclusion criterion for the study (yes/no).

We categorized drugs in the selected studies into the following groups based on class axis of multi-axial neuroscience-based nomenclature: glutamatergic, serotonergic, neuromodulator, pro-dopaminergic.20 See supplementary table 2 for details.

Statistical Analysis

Data were entered into an electronic database and analyzed with a quantitative meta-analytical approach using Comprehensive Meta-Analysis (CMA) Software version 3 (Biostat, Inc.).

For each study arm we computed the standardized mean differences (SMD or Cohen’s d) in negative symptom scores from baseline to endpoint (and its 95% CIs). Then, we assessed (1) the effect of drug vs placebo on negative symptom change and (2) “placebo response” (the main focus of our analysis) in the placebo-treated group only (ie, mean negative symptom change score over follow-up in the placebo arm). Based on Cohen’s d values, ES were categorized as small (0.2 to 0.5), moderate (0.5 to 0.8) or large (>0.8).21 Based on the known heterogeneity of studies, we used random-effects models by Der-Simonian and Laird.22 Random-effects models assume that the true ES varies from one study to the next and, accordingly, are more conservative than fixed effect models. We also checked the rigor of the main findings by performing a jackknife sensitivity analysis, which consists of iteratively repeating the meta-analysis, excluding one
We assessed statistical heterogeneity through visual inspection of forest plots and using the $Q$ statistic (a magnitude of heterogeneity) and the $I^2$ statistic (a measure of the proportion of variance in summary ES attributable to heterogeneity). We assessed publication bias by visually inspecting funnel plots and using Orwin's fail-safe $N$, with criterion for a “trivial” standardized difference in means as 0.1 and mean standardized difference in means in missing studies as 0. This generated the number of unpublished studies that would be needed to move estimates to a nonsignificant threshold. Furthermore, we used Egger's linear regression method to quantify the bias captured by the funnel plot. When the funnel plot or test statistics suggested publication bias, we used the Duval and Tweedie trim-and-fill method to estimate an ES corrected for publication bias.

We used meta-regressions with a random-effect model with unrestricted maximum likelihood to test effects of moderators on ES estimates. The slope of the meta-regression line – β coefficient: direct (+) or inverse (−) – indicates the strength of the relationship between moderator and outcome. We performed meta-regressions for moderator variables with at least 5 available studies. To further assess the influence of moderators on placebo response, we conducted multivariable meta-regression models, including those moderators significantly associated with placebo response in the bivariate meta-regressions as covariates. In meta-regression, the outcome variable is the effect estimate such as a mean difference, a risk difference, a log odds ratio or a log risk ratio. The explanatory variables are characteristics of studies that might influence the size of intervention effect. These are often called “potential effect modifiers” or covariates. We used a significance threshold of .05, 2-tailed.

**Results**

We analyzed a total of 18 independent DBRCTs (12 academic and 6 industry-funded) from 17 publications, comprising 998 patients with stable SSD and predominant/prominent negative symptoms. The studies assessed the effect of 13 active drugs vs placebo on negative symptoms as primary outcome. The mean age of patients ranged from 36.6 to 62.7 years (mean of 50.0 y). The percentage of female patients ranged from 0 to 100 (mean of 42.3%). DBRCTs lasted from 4 to 24 weeks (mean of 14.3 wk). Details of the included studies are shown in table 1.

A meta-analysis of the 18 DBRCTs showed that overall drugs had greater efficacy than placebo in reducing negative symptoms although the ES was small (Cohen’s $d$: 0.208, 95% CI: 0.033 to 0.382, $P = .020$). Our jackknife sensitivity analysis revealed stable results. Heterogeneity was significant ($Q = 50.821$, df = 17, $I^2$ 66.549%, $P < .001$) with low risk of publication bias (Orwin’s fail-safe $N = 23$; Egger’s regression intercept = 1.320, 95% CI: −0.502 to 3.142, $P = .144$). Supplementary figure 1 shows the forest plot of the meta-analysis of active drugs vs placebo effect.

The meta-analysis also showed that response to placebo from baseline to endpoint was significantly positive with a large ES value (Cohen’s $d$: 2.909, 95% CI: 2.051 to 3.767, $P < .001$). Jackknife sensitivity analysis revealed stable results. A large and significant study heterogeneity ($Q = 1024.188$, df = 17, $I^2$ 98.340%, $P < .001$) was found. Orwin's fail-safe $N$ was 137, and Egger's regression intercept was 10.333 (95% CI: 4.852 to 15.813), $P = .001$, suggesting high risk of significant publication bias. Duval and Tweedie's trim-and-fill method showed that adjusted point values were significantly greater than observed ones (adjusted values 3.378, 95% CI: 2.200 to 4.557, $Q = 2582.573$ vs observed values 2.909, 95% CI: 2.051 to 3.767, $Q = 1024.188$). Figure 2 shows the forest plot of the meta-analysis of placebo response.

Placebo response was significantly greater in studies with industry funding (Cohen’s $d$: 6.722, 95% CI: 3.952 to 9.942) than in those with academic funding (Cohen’s $d$: 1.010, 95% CI: 0.519 to 1.502).

Meta-regressions showed that a higher placebo response was significantly associated with a higher quality score, a more recent year of publication (range 2003 to 2017), a longer duration of trial, a larger number of study sites and countries in the trial, a larger number of patients (both in the whole sample and in the placebo arm), a younger mean age (in the placebo arm), a lower severity of positive symptoms at baseline (in the placebo arm), and industry funding (vs academic funding). The rest of the covariates had no significant association with placebo response (ie, ES of active drug vs placebo on negative symptoms, duration of illness, percentage of female participants, and severity of negative symptoms at baseline). We could not explore the influence of baseline depressive symptom scores, baseline extrapyramidal symptoms scores, percentage of drop-outs during follow-up, medical comorbidities, type of antipsychotic or substance use/abuse on placebo response because there were not enough studies with available data. Table 2 provides details of the meta-regressions.

To further evaluate the influence of covariates on placebo response, we conducted 3 multivariable meta-regression models assessing the effect of potential moderators on placebo response. The first model used the number of sites and the number of countries as covariates, and revealed that only the number of sites was significantly associated with placebo response. The second model included the quality, the year of publication, the duration of trial, the number of arms in the trial, the number of study sites, the number of patients (whole sample), the mean age (placebo arm), the severity of positive...
Table 1. Characteristics of the Studies

<table>
<thead>
<tr>
<th>First Author, Year</th>
<th>Active Drug, (Mean Daily Dose)</th>
<th>Drug Added on Antipsychotic Medication</th>
<th>Duration of Trial, Weeks</th>
<th>Funding</th>
<th>Negative Symptom Scale</th>
<th># Arms</th>
<th># Sites (Countries)</th>
<th>N (Total Sample)</th>
<th>Mean Age (PBO Arm), Years</th>
<th>% Female (PBO Arm)</th>
<th>Severity of Negative Symptoms at Baseline</th>
<th>Severity of Positive Symptoms at Baseline</th>
<th>Effect Size (ES) (Cohen’s d), Drug vs PBO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buchanan, 2007</td>
<td>D-cycloserine, (50 mg) and Glycine (60 mg)</td>
<td>Yes</td>
<td>16</td>
<td>Academic</td>
<td>SANS total</td>
<td>3</td>
<td>5 (2)</td>
<td>157</td>
<td>52</td>
<td>43.4</td>
<td>N/A</td>
<td>0.46</td>
<td>N/A</td>
</tr>
<tr>
<td>Buchanan, 2015</td>
<td>Rasagiline (1 mg)</td>
<td>Yes</td>
<td>12</td>
<td>Academic</td>
<td>Modified SANS</td>
<td>2</td>
<td>1 (1)</td>
<td>54</td>
<td>28</td>
<td>45.9</td>
<td>75.9</td>
<td>0.42</td>
<td>0.33</td>
</tr>
<tr>
<td>Bugarski Kirola, 2017 (DayLyte study)</td>
<td>Bitopertine (5 mg) and Bitopertine (10 mg)</td>
<td>Yes</td>
<td>24</td>
<td>Industry</td>
<td>PANSS Negative SFS</td>
<td>3</td>
<td>113 (14)</td>
<td>597</td>
<td>207</td>
<td>42.3</td>
<td>34</td>
<td>0.50</td>
<td>0.21</td>
</tr>
<tr>
<td>Bugarski Kirola, 2017 (FlashLyte study)</td>
<td>Bitopertine (10 mg) and Bitopertine (20 mg)</td>
<td>Yes</td>
<td>24</td>
<td>Industry</td>
<td>PANSS Negative SFS</td>
<td>3</td>
<td>122 (14)</td>
<td>574</td>
<td>190</td>
<td>38.7</td>
<td>28.6</td>
<td>0.48</td>
<td>0.22</td>
</tr>
<tr>
<td>Dunayevich, 2017</td>
<td>D-cycloserine (50 mg), D-cycloserine (50 mg), Folic acid (1 mg)</td>
<td>Yes</td>
<td>12</td>
<td>Industry</td>
<td>NSA-16</td>
<td>4</td>
<td>58 (N/A)</td>
<td>226</td>
<td>74</td>
<td>43.4</td>
<td>32.9</td>
<td>0.61</td>
<td>0.47</td>
</tr>
<tr>
<td>Duncan, 2004</td>
<td>D-cycloserine (50 mg)</td>
<td>Yes</td>
<td>4</td>
<td>Academic</td>
<td>SANS total</td>
<td>2</td>
<td>1 (1)</td>
<td>22</td>
<td>12</td>
<td>54.4</td>
<td>0</td>
<td>0.58</td>
<td>0.66</td>
</tr>
<tr>
<td>Goff, 2005</td>
<td>D-cycloserine (50 mg), D-cycloserine (50 mg), Folic acid (1 mg)</td>
<td>Yes</td>
<td>24</td>
<td>Academic</td>
<td>SANS total</td>
<td>2</td>
<td>3 (1)</td>
<td>55</td>
<td>12</td>
<td>47.0</td>
<td>28.6</td>
<td>0.43</td>
<td>0.28</td>
</tr>
<tr>
<td>Hill, 2011</td>
<td>Citalopram (30 mg) and Reboxetine (8 mg), Folic acid (1 mg)</td>
<td>Yes</td>
<td>12</td>
<td>Academic</td>
<td>Modified SANS</td>
<td>2</td>
<td>1 (1)</td>
<td>28</td>
<td>14</td>
<td>46.5</td>
<td>13.33</td>
<td>0.51</td>
<td>0.36</td>
</tr>
<tr>
<td>Hinkelmann, 2013</td>
<td>Modafinil (200 mg)</td>
<td>Yes</td>
<td>8</td>
<td>Academic</td>
<td>Modified SANS</td>
<td>2</td>
<td>1 (1)</td>
<td>20</td>
<td>10</td>
<td>49.8</td>
<td>10</td>
<td>0.48</td>
<td>0.32</td>
</tr>
<tr>
<td>Pierre, 2007</td>
<td>Org25935 (4–8 mg/d mg) and Org25935 (12–16 mg/d mg)</td>
<td>Yes</td>
<td>12</td>
<td>Industry</td>
<td>SANS total</td>
<td>3</td>
<td>25 (7)</td>
<td>214</td>
<td>62</td>
<td>38.1</td>
<td>34.30</td>
<td>0.53</td>
<td>0.27</td>
</tr>
<tr>
<td>Schoemaker, 2014</td>
<td>PGM (40 mg)</td>
<td>Yes</td>
<td>16</td>
<td>Industry</td>
<td>NSA-16</td>
<td>2</td>
<td>N/A (4)</td>
<td>164</td>
<td>82</td>
<td>42.8</td>
<td>21.95</td>
<td>0.62</td>
<td>0.34</td>
</tr>
<tr>
<td>Stauffer 2013</td>
<td>DHEA (100 mg)</td>
<td>Yes</td>
<td>6</td>
<td>Academic</td>
<td>SANS</td>
<td>2</td>
<td>1 (1)</td>
<td>27</td>
<td>12</td>
<td>36.6</td>
<td>75</td>
<td>N/A</td>
<td>0.26</td>
</tr>
<tr>
<td>Strzelecki 2015</td>
<td>Sarcosine (2 mg)</td>
<td>Yes</td>
<td>24</td>
<td>Academic</td>
<td>PANSS Negative SFS</td>
<td>2</td>
<td>1 (1)</td>
<td>50</td>
<td>25</td>
<td>40</td>
<td>48</td>
<td>0.53</td>
<td>0.22</td>
</tr>
<tr>
<td>Umbricht, 2014</td>
<td>Bitopertine (10 mg), Bitopertine (30 mg), Bitopertine (60 mg)</td>
<td>Yes</td>
<td>8</td>
<td>Industry</td>
<td>Negative SFS</td>
<td>4</td>
<td>66 (4)</td>
<td>235</td>
<td>61</td>
<td>39</td>
<td>44</td>
<td>0.53</td>
<td>0.36</td>
</tr>
</tbody>
</table>
symptoms at baseline (placebo arm), and sponsorship (academic vs industry) as covariates, and revealed that larger number of arms in the trial, larger number of study sites, and industry funding were significantly related to placebo response. The third model included these 3 variables (number of arms in the trial, number of study sites and sponsorship) and confirmed that all of them significantly contributed to the model (placebo response). Table 3 shows details of the multivariate meta-regression models. Supplementary figure 2 shows the scatterplots of univariate meta-regressions of these 3 variables.

**Discussion**

This meta-analysis showed that placebo response in DBRCTs assessing treatment efficacy on negative symptoms (as primary outcome) in people with stable schizophrenia and predominant/prominent negative symptoms was significant ($P < .001$) and clinically relevant (Cohen's $d$: 2.909, 95% CI: 2.051 to 3.767), but in a context of significant study heterogeneity and high risk of publication bias. Notwithstanding, our findings suggest that the inclusion of putatively omitted studies would have increased Cohen's $d$ of placebo response. This result of a large placebo response is congruent with that of previous empirical studies and meta-analyses addressing placebo response in clinical trials of pharmacological treatment strategies for schizophrenia.12,13,15,43 Further, we found that overall, drugs (as add-on strategies) had greater efficacy than placebo (also as add-on to antipsychotics) in reducing negative symptoms although with a small ES (Cohen's $d$: 0.208, 95% CI: 0.033 to 0.382, $P = .020$) and significant heterogeneity. We found a low risk of publication bias as well as an association at trend level, and with small ES value, between greater placebo response and smaller ES of active drug vs placebo (Cohen's $d$: −1.684, 95% CI: −3.500 to 0.132, $P = .069$), thus suggesting that greater placebo response may be associated with a reduced signal for active drugs vs placebo in add-on studies of negative symptoms, which could lead to the loss of potentially effective drugs during drug development.

The use of add-on strategies to antipsychotics, unlike direct comparisons of drug therapy to placebo, involves no placebo monotherapy and both treatment groups receive antipsychotic treatment, thus potentially leading to an increase in secondary negative symptoms. In a monotherapy setting one would expect if anything a higher placebo response as patients came off their antipsychotics.44 All the clinical trials included in our study used add-on designs. A more recent phase 2 study testing MIN-101, a novel cyclic amide derivative with high equipotent affinities for sigma-2 and 5-hydroxytryptamine 2A receptors, showed statistically significant efficacy vs placebo, with a moderate ES (Cohen’s $d$: 0.540), in reducing negative symptoms in stable patients with schizophrenia when used as monotherapy.45 The placebo response

### Table 1. Continued

<table>
<thead>
<tr>
<th>First Author, Year</th>
<th>Active Drug, (Mean Daily Dose)</th>
<th>Drug Added on Antipsychotic Medication</th>
<th>Duration of Trial, Weeks</th>
<th>Funding</th>
<th>Negative Symptom Scale</th>
<th># Sites (Countries)</th>
<th># Arms</th>
<th># Arms (Arm)</th>
<th>N (Total Sample)</th>
<th>N (PBO Arm)</th>
<th>Mean Age (PBO Arm), Years</th>
<th>% Female (PBO Arm)</th>
<th>Effect Size (ES) (Cohen’s $d$), Drug vs PBO, Negative and Positive Syndrome Scale</th>
<th>Severity of Negative Symptoms at Baseline</th>
<th>Severity of Positive Symptoms at Baseline</th>
<th>Note: DHEA, dehydroepiandrosterone; ES, effect size; N/A not available; NSA-16, Negative Symptom Assessment; PANSS, Positive and Negative Syndrome Scale; SANS, Scale for the Assessment of Negative Symptoms; SFS, Symptom Factor Scale; PGM, pomaglumetad methyl.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Usall, 2011</td>
<td>Yes</td>
<td>Acute Raloxifene (60 mg)</td>
<td>12</td>
<td>Academic</td>
<td>PANSS</td>
<td>2</td>
<td>2 (1)</td>
<td>33 (1)</td>
<td>100</td>
<td>47 (2)</td>
<td>62.7</td>
<td>100</td>
<td>0.25 0.73</td>
<td>0.25 0.73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Usall, 2016</td>
<td>Yes</td>
<td>Acute Raloxifene (60 mg)</td>
<td>24 (1)</td>
<td>Academic</td>
<td>PANSS</td>
<td>2</td>
<td>3 (1)</td>
<td>70 (1)</td>
<td>100</td>
<td>57 (2)</td>
<td>61.3</td>
<td>100</td>
<td>0.47 0.35 0.64</td>
<td>0.47 0.35 0.64</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weiser, 2012</td>
<td>Yes</td>
<td>Acute D-serine (2 mg)</td>
<td>16</td>
<td>Academic</td>
<td>SANS</td>
<td>2</td>
<td>10 (1)</td>
<td>39 (1)</td>
<td>195</td>
<td>98 (2)</td>
<td>39.8</td>
<td>100</td>
<td>0.49 0.27 -0.09</td>
<td>0.49 0.27 -0.09</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: DHEA, dehydroepiandrosterone; ES, effect size; N/A not available; NSA-16, Negative Symptom Assessment; PANSS, Positive and Negative Syndrome Scale; PGM, pomaglumetad methyl.
in the MIN-101 study (Cohen's $d$: 3.255, 95% CI: 2.601 to 3.910, $P < .001$) was comparable to that found in our meta-analysis of add-on strategies (Cohen's $d$: 2.909, 95% CI: 2.051 to 3.767, $P < .001$). Future studies should ascertain whether the type of design (add-on or monotherapy) is a significant predictor of placebo response in clinical trials assessing strategies for negative symptoms in stable schizophrenia and whether the determinants of placebo response are the same in both kinds of design.

Among potential predictors of placebo response, we found that larger numbers of arms in the trial, larger numbers of study sites, and industry funding were significant moderators of placebo response in these patients. Interestingly, these moderators are not substantially different from those found in global studies of treatment of patients with schizophrenia that did not focus on stable patients with predominant/prominent negative symptoms.$^{14-17}$

Several things might explain our finding of these potential moderators of placebo response. First, we found that the larger the number of arms in the trial, the higher the probability of being assigned to active treatment. This suggests that a DBRCT with more than 2 active arms may induce a greater expectation of improvement in patients and raters, which is a significant mediator of placebo response. This issue has been previously highlighted in trials of schizophrenia$^{17}$ and major depressive disorder (MDD).$^{46,47}$ Thus, researchers should
consider expectations of symptom changes when interpreting results and designing clinical trials.

Second, we found that the larger the number of study sites, the greater placebo response, even after controlling for the number of countries and number of participants. Some previous meta-regression analyses have found equivalent results in antipsychotic trials in patients with schizophrenia. As a result, some researchers recommend limiting to 40 the number of study sites in clinical trials, especially in nonacademic settings. Clinical trials of antidepressants for MDD have also found an association between placebo response and number of study sites, suggesting that this is not specific for schizophrenia trials. Study sites in schizophrenia trials significantly increased from a median of 2 sites before 1990 to 38 sites in the 2005–2010 interval, so there might be a confounding effect of study year on these findings. As found in an analysis of predictors of drug effect in schizophrenia trials, contrary to some criticisms and expectations, industry sponsorship was associated with smaller drug effect sizes, probably in part as a result of greater placebo response. Industry studies often involve multiple countries and sites, leading to great heterogeneity in the participant profile, due to factors such as cultural differences in the interpretation of psychopathology, pressure for swift recruitment, which may promote inclusion of so-called “professional patients” and may involve a change in the patient population enrolling in antipsychotic clinical trials or differences in rates of recruitment between centers (in the majority of industry sponsored studies, 10% of the sites recruit more than 50% of the patients). These factors generate variability, decreasing the effect sizes of active drugs vs placebo. Further, industry-sponsored trials might be associated with an inflation of baseline symptom scores to enable enrollment of the patient into the study. In our case, industry studies reported greater severity of baseline negative symptoms than academic studies (P = .068). In such a scenario, we would expect symptom scores to decrease to more accurate levels upon follow-up visits (possibly due to regression toward the mean), leading to significant improvement in both the placebo and active arms. We and other interested researchers should investigate this supposition. Additional factors likely include the fact that industry studies are conducted mostly by commercial research sites that get the majority of patients through referrals or advertisements and therefore may not have the same knowledge of the patient as investigators in an academic center may have. Similarly, the personal involvement and attention to quality and detail may be higher in academic studies as they are usually funded by direct grants to

### Table 3. Multivariable Meta-regression Models of the Effect of Moderators on Placebo Response

<table>
<thead>
<tr>
<th>Model</th>
<th>Number of Studies in the Analysis</th>
<th>Moderator Variable</th>
<th>Coefficient (95% CI)</th>
<th>Z-value</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>16</td>
<td>Number of sites</td>
<td>0.143 (0.071 to 0.214)</td>
<td>3.91</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Number of countries</td>
<td>−0.335 (−0.943 to 0.273)</td>
<td>−1.08</td>
<td>.280</td>
</tr>
<tr>
<td>2*</td>
<td>17</td>
<td>Study quality score</td>
<td>−0.149 (−0.614 to 0.316)</td>
<td>−0.63</td>
<td>.530</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Year of publication</td>
<td>−0.147 (−0.326 to 0.031)</td>
<td>−1.61</td>
<td>.106</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Duration of trial</td>
<td>0.027 (−0.097 to 0.151)</td>
<td>0.43</td>
<td>.670</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Number of arms in the trial</td>
<td>2.729 (0.737 to 4.721)</td>
<td>2.68</td>
<td>.007</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Number of sites</td>
<td>0.124 (0.041 to 0.206)</td>
<td>2.94</td>
<td>.003</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Number of patients (whole sample)</td>
<td>0.003 (−0.012 to 0.019)</td>
<td>0.43</td>
<td>.666</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean age (placebo arm)</td>
<td>0.055 (−0.044 to 0.155)</td>
<td>1.09</td>
<td>.275</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Severity of positive symptoms at baseline (placebo arm)</td>
<td>−3.323 (−10.292 to 3.647)</td>
<td>−0.93</td>
<td>.350</td>
</tr>
<tr>
<td>3*</td>
<td>17</td>
<td>Number of arms in the trial</td>
<td>5.313 (1.131 to 9.496)</td>
<td>2.49</td>
<td>.013</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Number of sites</td>
<td>1.965 (0.790 to 3.139)</td>
<td>3.28</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Number of patients (whole sample)</td>
<td>0.135 (0.110 to 0.159)</td>
<td>10.86</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Industry funding</td>
<td>5.491 (2.923 to 8.060)</td>
<td>4.19</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

Note: *Proportion of variance explained by Model 2: 66% (P < .001).

Proportion of variance explained by Model 3: 82% (P < .001).
the investigator. These factors may all contribute to the recruitment of patients who truly have persistent negative symptoms whereas in commercial centers the rate of patients in whom the persistence of negative symptoms cannot be truly ascertained may be higher. Lastly it is important to note that the observed placebo response is composed both of the actual reporting of improvement by patients themselves but also of an expectation bias by the rater. The attention to detail and the quality of ratings may be better in academic trials as their investigators have more personal investment in a study. Thus, the biasing of assessment by rater expectations may be reduced.

Fourth, contrary to prior studies assessing predictors of placebo response in schizophrenia trials, we did not find duration of trial to be a significant moderator of placebo response. However, our study differs from prior studies because it focuses on a specific subgroup of patients with prominent negative symptoms and focuses on negative symptom change. Although we do not have a definitive explanation for this finding, it is congruent with previous studies showing that negative symptoms in chronic, stable, outpatient patients improve over time through clinical trials, unrelated to the duration of trial. Our results do not allow us to speculate about the duration of the placebo effect, such as whether it persists or decreases over time. This would be a relevant topic for future research on placebo response.

A recent review of schizophrenia cognition trials showed that greater severity of depression/anxiety and more motivation were moderators of placebo response. A study on neurochemical mechanisms underlying placebo response in patients with MDD also found that placebo-induced activation of the mu-opioid system is implicated in the placebo antidepressant effect. However, except for these cases, studies on predictors of placebo response in schizophrenia (and other mental disorders) have failed to identify individual biological or psychological predictive factors of placebo response. Distinguishing the placebo effect from the effect of medication, or even from that of surgical intervention, is crucial in interpreting the results of clinical trials. Thus, a personal precision medicine approach may help identify individual markers of placebo response and thereby implement treatment better suited to each patient.

Specific cognitive appraisals and beliefs play a role in the expression and persistence of negative symptoms, so the construct of negative symptom must be relevant to the patients’ beliefs and expectations. Expectations of success have been found to predict significant reduction in negative symptoms, so psychosocial interventions targeting expectations and beliefs might play a role in the management of negative symptoms, augmenting the interventions tested. However, in spite of this and other demonstrations of the relevance of measuring patients’ expectations most schizophrenia clinical trials do not include any assessment of expectation of improvement, making it impossible to assess the effect of expectations on the placebo response in meta-regressions. Since expectations might drive placebo response both on the side of the patient and the clinician, future clinical trials in negative symptoms in schizophrenia might consider measuring expectations and beliefs in both participants and clinicians and implementing strategies for clinicians to manage these variables.

Some limitations of our analysis should be considered. The definition of predominant or prominent negative symptoms varied across studies. There is no consensus on the definition of predominant negative symptoms, so we decided to include only those trials with at least one specific inclusion criteria of severity of negative symptoms (see supplementary table 3 for details of negative symptoms criteria used in each study). Nonetheless, the patient samples may have differed to some extent in their degree and predominance of negative symptoms. Also, the number of available studies fulfilling defined criteria (N = 18) was quite small. Nonetheless, our analyses were powerful enough to find significant placebo response across the trials and significant predictors of this response. We also limited the number of studies by deciding to restrict the time frame of our search to the last 15 years. We chose this time frame since predominant/prominent negative symptoms only became a target of investigation in the early 2000s. Nor could be explored the influence of several moderators on placebo response (eg, baseline depressive symptom scores, baseline extrapyramidal symptoms scores, percentage of drop-outs during follow-up, medical comorbidities, type of antipsychotic or substance use/abuse), because they appeared in fewer than 5 studies. Further, even for those variables that could be assessed, the number of studies with available data could have influenced the statistical power to detect significant associations, considering that, with the exception of the number of countries, no other variable with data from fewer than 18 studies reached significance. Nonetheless, meta-regressions showed the effect on placebo response of significant moderator variables, such as quality score, year of publication, duration of trial, number of drug active treatment arms in the trial, number of study sites and countries in the trial, number of patients, mean age, severity of positive symptoms at baseline, or industry funding.

In summary, this meta-regression analysis identified relevant clinical trial design and operational factors, such as the number of study arms, the number of study sites and funding source, as factors associated with placebo response in DBRCTs assessing treatment efficacy of negative symptoms in schizophrenia patients with predominant/prominent negative symptoms. It is conceivable that the large and increasing placebo response observed may have obscured true effects of some of the novel treatments for negative symptoms. That could lead to the health system denying potentially useful treatment.
strategies to patients. These findings might help guide the design of clinical trials better suited to testing the efficacy of novel pharmacological strategies for treating negative symptoms in this population, perhaps by limiting the number of arms and study sites.

Supplementary Material

Supplementary data are available at Schizophrenia Bulletin online.

Funding

This work was supported by the Spanish Ministry of Economy, Industry and Competitiveness. Carlos III Health Institute (PI12/01303, PI13/02112, PI14/00397), co-financed by ERDF Funds from the European Commission, “A way of making Europe,” CIBERSAM. Madrid Regional Government (S2010/BMD-2422 AGES and B2017/BMD- 3740 AGES-2-CM), European Union Structural Funds, European Union Seventh Framework Program under grant agreements FP7-HEALTH-2009-2.2.1-2-241909 (Project EU-GEI), FP7-HEALTH-2009-2.2.1-3-242114 (Project OPTiMISE), FP7-HEALTH-2013-2.2.1-2-603196(ProjectPSYSCAN), FP7-HEALTH-2013-2.2.1-2-602478 (Project METSY), and FP7-HEALTH-F4-2010-241959 (Project PERs), and European Union H2020 Program under the Innovative Medicines Initiative 2 Joint Undertaking (grant agreement No 115916; Project PRISM), Fundación Alicia Koplowitz and Fundación Mutua Madrileña. ERA-NET NEURON (Network of European Funding for Neuroscience Research).

Acknowledgments

D.F. has been a consultant and/or has received fees from AstraZeneca, Bristol-Myers-Squibb, Janssen, Lundbeck, Otsuka and Pfizer. He also received a fellowship from Instituto de Salud Carlos III (Ministry of Economy and Competitiveness). C.M.D-C. and L.P-C. have previously held grants from Instituto de Salud Carlos III (Spanish Ministry of Economy and Competitiveness) and from the Alicia Koplowitz Foundation. D.U. is full-time employee of F. Hoffmann-La Roche. C.A. has been a consultant to or has received honoraria or grants from Acadia, Abbot, AMGEN, AstraZeneca, Bristol-Myers Squibb, Caja Navarra, CIBERSAM, Dainippon Sumitomo Pharma, Fundación Alicia Koplowitz, Forum, Instituto de Salud Carlos III, Gedeon Richter, Janssen Cilag, Lundbeck, Merck, Ministerio de Ciencia e Innovación, Ministerio de Sanidad, Ministerio de Economía y Competitividad, Mutua Madrileña, Otsuka, Pfizer, Roche, Servier, Shire, Schering Plough, Sunovio and Takeda.

References


