Baseline Factors Predicting Placebo Response to Treatment in Children and Adolescents With Autism Spectrum Disorders: A Multisite Randomized Clinical Trial

Bryan H. King, MD; Kimberly Dukes, PhD; Craig L. Donnelly, MD; Linmarie Sikich, MD; James T. McCracken, MD; Lawrence Scahill, MSN, PhD; Eric Hollander, MD; Joel D. Bregman, MD; Evdokia Anagnostou, MD; Fay Robinson, MPH; Lisa Sullivan, PhD; Deborah Hirtz, MD

**IMPORTANCE** The finding of factors that differentially predict the likelihood of response to placebo over that of an active drug could have a significant impact on study design in this population.

**OBJECTIVE** To identify possible nonspecific, baseline predictors of response to intervention in a large randomized clinical trial of children and adolescents with autism spectrum disorders.

**DESIGN, SETTING, AND PARTICIPANTS** Randomized clinical trial of citalopram hydrobromide for children and adolescents with autism spectrum disorders and prominent repetitive behavior. Baseline data at study entry were examined with respect to final outcome to determine if response predictors could be identified. A total of 149 children and adolescents 5 to 17 years of age (mean [SD] age, 9.4 [3.1] years) from 6 academic centers were randomly assigned to citalopram (n = 73) or placebo (n = 76). Participants had autistic disorder, Asperger syndrome, or pervasive developmental disorder, not otherwise specified; had illness severity ratings that were moderate or more than moderate on the Clinical Global Impression–Severity scale; and scored moderate or more than moderate on compulsive behaviors measured with the modified Children's Yale-Brown Obsessive-Compulsive Scale.

**INTERVENTIONS** Twelve weeks of treatment with citalopram (10 mg/5 mL) or placebo. The mean (SD) maximum dose of citalopram was 16.5 (6.5) mg by mouth daily (maximum dose, 20 mg/d).

**MAIN OUTCOMES AND MEASURES** A positive response was defined as having a score of at least much improved on the Clinical Global Impression–Improvement scale at week 12. Baseline measures included demographic (sex, age, weight, and pubertal status), clinical, and family measures. Clinical variables included baseline illness severity ratings (the Aberrant Behavior Checklist, the Child and Adolescent Symptom Inventory, the Vineland Adaptive Behavior Scales, the Repetitive Behavior Scale–Revised, and the Children's Yale-Brown Obsessive-Compulsive Scale). Family measures included the Caregiver Strain Questionnaire.

**RESULTS** Several baseline predictors of response were identified, and a principal component analysis yielded 3 composite measures (disruptive behavior, autism/mood, and caregiver strain) that significantly predicted response at week 12. Specifically, participants in the placebo group were significantly less likely than participants in the citalopram group to respond at week 12 if they entered the study more symptomatic on each of the 3 composite measures, and they were at least 2 times less likely to be responders.

**CONCLUSIONS AND RELEVANCE** This analysis suggests strategies that may be useful in anticipating and potentially mitigating the nonspecific response in randomized clinical trials of children and adolescents with autism spectrum disorders.

**TRIAL REGISTRATION** clinicaltrials.gov Identifier: NCT00086645

Published online September 23, 2013.
Autism is a neurobehavioral syndrome characterized by impairments in social communication, by unusual preoccupations or interests, and by stereotyped or repetitive behaviors. Although there are no medications yet approved specifically for any of these core deficits, the number of medications targeting behavioral symptoms in this population has increased significantly in recent years, and the majority of children with autism are treated with at least 1 psychotropic medication by the time they reach 8 years of age.1

Virtually every class of psychotropic medication has been explored, typically in an open-label fashion, in search of treating some facet of autistic disorder. Moreover, at least 1 positive case report can be found for an overwhelming number of therapeutics, including lysergic acid diethylamide.2 Over the past decade, increased efforts have been directed toward identifying safe and effective treatments for both core symptoms and associated, severely impairing symptoms such as self-injury, aggression, hyperactivity, and repetitive behaviors. As a result of the number of placebo-controlled studies that have been completed in recent years, one of the consistent findings is the recognition of the importance of the placebo response in this population.

Considerable interest was generated in the administration of secretin following the observation by Horvath and colleagues3 that it appeared to dramatically improve core social deficits in children with autism who had received the peptide in the course of an evaluation of gastrointestinal symptoms. A series of additional case reports quickly followed, and eventually several controlled trials were mounted to examine the effects of secretin.

From the experience that followed, including some 15 controlled trials and more than 600 participants,4,5 it could be argued that treatment with secretin was associated with an improvement in a significant percentage of participants. However, the placebo condition uniformly delivered equal or better outcomes across all of these trials, and the sum of the evidence is that there is no therapeutic signal for secretin in autism.4,5 And although neither autism nor the symptom targets that have been the focus of clinical trials would seem particularly likely to be placebo responsive, it is clear that improvement is to be expected in many children and adolescents with autism who enter these studies.

Taken together, in controlled trials that have used dichotomous outcomes in children with autism for a variety of clinical targets, placebo response rates have ranged from just under 20% to 50% (Figure 1).6-27 Similar results for childhood depression and for anxiety disorders have been reported.28-30 Indeed, the majority of variability between positive and negative trials in childhood depression is in the range of placebo response rates. Other factors implicated in higher placebo response rates in childhood depression include milder baseline illness severity, younger age, and number of study sites.28 Similarly, a reexamination of depression trials in adults suggest that evidence for the effectiveness of antidepressant medicines is correlated with more severe illness.31-33

Various factors have been advanced to account for placebo or nonspecific response in the population of children with autism, including the heightened expectancy in parents who enter their children in clinical trials against a backdrop of relentless media exposure and seemingly daily claims of curative interventions; a symptom pattern in autism that commonly includes a waxing and waning course; and the attention, care, and structure that comes with participation in a clinical trial.34 In their recent analysis of moderators, mediators, and nonspecific predictors of response in 100 participants from the Research Units on Pediatric Psychopharmacology Autism Network risperidone trial, Arnold et al35 observed that parent education and family income predicted positive outcome in both

---

**Figure 1. Placebo Response in Randomized Clinical Trials of Autism and Related Disorders in Which a Dichotomous Response (eg, Responder vs Nonresponder) Was Reported**

<table>
<thead>
<tr>
<th>Drug</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amantadine</td>
<td>King et al, 2001</td>
</tr>
<tr>
<td>Aripiprazole</td>
<td>Owen et al, 2009</td>
</tr>
<tr>
<td>Atomoxetine</td>
<td>Arnold et al, 2006</td>
</tr>
<tr>
<td>Divalproes</td>
<td>Hollander et al, 2009</td>
</tr>
<tr>
<td>Citalopram</td>
<td>King et al, 2009</td>
</tr>
<tr>
<td>Fluvoxamine</td>
<td>McDougle et al, 1996</td>
</tr>
<tr>
<td>Methylphenidate</td>
<td>RUPP, 2005</td>
</tr>
<tr>
<td>Glutarazine</td>
<td>Hollander et al, 2006</td>
</tr>
<tr>
<td>Oral IG</td>
<td>Handen et al, 2009</td>
</tr>
<tr>
<td>ORG 2766</td>
<td>Buitelaar et al, 1996</td>
</tr>
<tr>
<td>Risperidone</td>
<td>McDougle et al, 1998</td>
</tr>
<tr>
<td>Secretin</td>
<td>Carey et al, 2002</td>
</tr>
<tr>
<td>Valproate</td>
<td>Hellings et al, 2005</td>
</tr>
<tr>
<td></td>
<td>Sondheimer et al, 2002</td>
</tr>
</tbody>
</table>

Oral IG indicates oral human immunoglobulin; ORG 2766, an adrenocorticotropic hormone (4-9) analog; and RUPP, Research Units on Pediatric Psychopharmacology Autism Network.
active and placebo conditions. A host of baseline behavioral ratings derived from the Child and Adolescent Symptom Inventory (CASI), including ratings for anxiety, bipolar symptoms, oppositional-defiant symptoms, stereotypy, and hyperactivity, negatively predicted outcome. Garcia and colleagues were also interested in identifying predictors and moderators of outcome in the Pediatric Obsessive Compulsive Treatment Study, and after an excellent review, they identified demographic factors (sex, age, and household income), severity of illness markers (baseline severity of obsessive-compulsive disorder, functional impairment, and insight), comorbidity (internalizing diagnosis, externalizing diagnosis, anxiety symptoms, and externalizing symptoms), and family factors (parental psychopathology, family history of obsessive-compulsive disorder, family functioning, and accommodation) as relevant predictors or moderators of outcome. In their study of 112 participants, baseline severity of obsessive-compulsive disorder, functional impairment, insight, externalizing symptoms, and family accommodation all emerged as significant predictors.

For multiple reasons, including ethical, methodological, and financial, it is important to minimize the chances of including children in a failed clinical trial (ie, one that does not provide conclusive evidence for or against efficacy of the active treatment). One consequence of a higher-than-expected placebo response in a trial is to increase the chances of a signal being lost owing to lack of statistical power, resulting in a potentially useful treatment being passed over or discarded. Trials are extraordinarily expensive and time consuming to complete, and it is important that research efforts are undertaken with the greatest likelihood of success in detecting true differences between placebo and active treatment conditions.

The Studies to Advance Autism Research and Treatment (STAART) psychopharmacology network completed a trial of citalopram hydrobromide for the treatment of repetitive behavior in children with autism. We hypothesized that citalopram would produce significant global improvement, either by reducing interference associated with repetitive behaviors or by reducing anxiety and improving frustration tolerance or mood. In this large clinical trial, we found that there were just as many responders to placebo as to citalopram.

In an effort to better understand the placebo response in autism, we were motivated to determine if there were predictors at the time of study entry that might distinguish participants who are more likely to respond generally and also whether there were factors that specifically influenced response to placebo vs citalopram. Identification of participant-specific attributes that contribute to placebo responsiveness in autism will inform the design and interpretation of future studies to advance effective treatments for this complex disorder.

Methods

A detailed description of the STAART citalopram trial has been published elsewhere. The clinical trial (NCT00086645) was registered at www.clinicaltrials.gov prior to initiation and conducted at 6 academic medical centers: the Mount Sinai School of Medicine, New York, New York; the North Shore–Long Island Jewish Health System, Great Neck, New York; the Dartmouth-Hitchcock Medical Center, Lebanon, New Hampshire; UCLA; the University of North Carolina, Chapel Hill; and Yale University, New Haven, Connecticut. Each site’s institutional review board approved the study, and informed consent was obtained from all study participants and/or legal representatives.

The National Institute of Mental Health convened a data and safety monitoring board that monitored the trial. The primary analyses showed that there was no significant difference in response on the Clinical Global Impression–Improvement subscale between the citalopram group (32.9% response rate) and the placebo group (34.2% response rate); the mean dose was 0.49 mg/kg (range, 0.07–1.26 mg/kg) at week 12. There were no significant differences between citalopram and placebo groups on any of the secondary outcome measures.

Participants

A total of 149 children and adolescents between 5 and 17 years of age (mean [SD] age, 9.4 [3.1] years) were randomly assigned to receive either citalopram (n = 73) (maximum up to 20 mg/d) or placebo (n = 76) for 12 weeks.

Main Outcome Measures

Primary Outcome

The primary outcome was the same as the outcome used in the primary study by King et al. A responder (someone who had a positive clinical response) was defined by a Clinical Global Impression–Improvement score of 1 or 2 (much or very much improved compared with baseline) at week 12 by an evaluating clinician.

Baseline Predictors

The baseline predictors contain many of the same measures used as secondary outcomes (at week 12) in the primary study. Prior to randomization, the following symptom measures, including subscales for the child, were obtained: the Aberrant Behavior Checklist (ABC), Child and Adolescent Symptom Inventory (CASI), Repetitive Behavior Scale–Revised, Children’s Yale-Brown Obsessive-Compulsive Scale, and Vineland Adaptive Behavior Scale subscale scores. All of the symptom measures were reported by the child’s guardian and reviewed by the evaluating clinician, masked to drug dose and side effects. The following data were obtained for the guardian: age, education, employment status, race, parity, and Caregiver Strain Questionnaire.

Confounders

Potential baseline confounders included age, sex, race, severity of autism symptoms overall (using the Clinical Global Impression–Severity scale), vital signs, body mass index, pubertal status, IQ, medication history and concomitant services, study site, and presence of adverse events elicited from the Side-Effect Monitoring Uniform Report Form at study entry. Potential confounders that could affect response, particularly for those children randomly assigned to citalopram, included various measures of adverse effects. We created ad-
verse effect measures that reflected phenotype of disease, burden of disease, or any adverse effect that worsened by week 12 from baseline.

**Statistical Analysis**

The goal was to determine the baseline characteristics that influenced or predicted response at week 12. Because there were many candidate variables and overlap between variables with respect to substantive meaning, we performed a principal component analysis to reduce the number of tests performed (to minimize the effect of multiple testing by controlling the experimental \( P \) value) and to group like items for ease of interpretation. First, we investigated the relationship between each baseline measure/predictor and response at week 12 based on responder status (ie, dichotomized Clinical Global Impression–Improvement score), in all participants and stratified by treatment group using either a \( t \) test or \( \chi^2 \) test, based on the type of baseline measure under investigation (continuous or categorical, respectively). Second, to ensure that there was no imbalance between treatment groups at baseline, the same analytic approach described in step 1 was used to compare each baseline predictor variable and each potential baseline confounder by treatment group at baseline. Third, a principal component analysis was performed using the significant baseline predictors from step 1. The number of components retained in the principal component analysis was determined by a combination of the eigenvalue greater than 1 rule and the percent variance–explained rule, as well as by substantive interpretation after both varimax and promax rotation. Fourth, based on the results of the principal component analysis, composite variables were created by standardizing each predictor variable and taking the mean value of the standardized predictor variables that uniquely loaded on a given component. Fifth, analysis of each composite variable predicting 12-week response, adjusted for relevant adverse effect confounders (after randomization through the study period), was performed using logistic regression analysis.

The effect modification (statistical interaction) between each composite measure (continuous and dichotomized at the median) and treatment group on response was investigated using multivariable logistic regression analysis. If there was significant statistical interaction, analyses were stratified by treatment group. Relative risks and 95% CIs were computed within the treatment group for each of the composite measures to summarize effect. Our study has 80% power to detect a relative risk as small as 2.4 in stratified analysis (within treatment groups), assuming a response as small as 20% in 1 group.

All analyses were performed in SAS version 9.2 (SAS Institute Inc). No adjustment was made for multiple testing because the analyses were exploratory in nature. Statistical significance was based on 2-sided tests and an \( \alpha \) level of .05.

**Results**

Several of the individual baseline predictors were statistically related to response at week 12, primarily placebo response (eTable 1 in Supplement). Baseline predictors included ABC Hyperactivity and Irritability subscales; the CASI Attention-Deficit/Hyperactivity Disorder, Autism, Depression, Panic, Schizophrenia, Somatization, and Tic subscales; the Caregiver Strain Internalizing and Externalizing subscales; and the Vineland Socialization Domain Standard Score. For all but the CASI Panic item score, the higher the symptom burden as reflected in the particular measure, the less likely the participant would be a responder to citalopram or placebo. There were no statistically significant differences between the proportions of responders for each of the baseline confounders, with the exception of body mass index (a lower body mass index predicted response), and we subsequently adjusted for body mass index in our analyses.

By including all of the baseline predictor variables listed in our principal component analyses, we found that 3 distinct components emerged, and from these components, we created 3 composite measures (Table). The first composite, Disruptive Behavior, is composed of ABC Hyperactivity, ABC Irritability, and CASI Attention-Deficit/Hyperactivity Disorder scales; the second composite, AutismSeverity/Mood, is composed of the CASI Autism and Depression scales; and the third composite, Caregiver Strain, is composed of the Caregiver Strain QuestionnaireExternalized and Internalized scales. Our Table provides descriptive statistics for each of the composite measures by treatment group. Each composite measure was approximately normally distributed, and the theoretic range for each measure was between 0 and 100, in which a higher score indicates more symptoms. Each of the 3 composite measures was dichotomized at the median. The interaction between each composite measure and treatment group was significant or trended toward significance; thus, stratified analysis was performed. Placebo responders are statistically significantly less symptomatic on all 3 composite measures at baseline than nonresponders (all \( P < .013 \)), and there are no statistically significant differences between each composite measure at baseline for those receiving citalopram (Table). Participants in the placebo group were less likely to respond if they entered the study with scores exceeding the median of Disruptive Behavior (relative risk, 2.0 [95% CI, 1.0-4.0]), Mood/Autism (relative risk, 2.5 [95% CI, 1.2-5.2]), or Caregiver Strain (relative risk, 2.5 [95% CI, 1.2-5.3]) composite measures (Figure 2). Adjustment for adverse effects during the trial did not modify the results.

An additional exploratory analysis was performed to determine whether response at week 12 would be reflected differently if there was a threshold that would have limited the population of participants enrolled. Because our baseline predictors largely measured disruptive behaviors, we chose the Caregiver Strain composite as a hypothetical screener; thus, we retained participants whose guardians scored above the median on the Caregiver Strain composite. There are no statistically significant differences between the treatment and placebo groups in standard measures of repetitive behavior on response from baseline to week 12, although there is a trend in ABC Irritability subscale: most likely an issue with power (eTable 2 in Supplement).
Discussion

Because the STAART citalopram trial was among the largest multisite clinical trials in autism to date, the number of participants randomly assigned to each condition was sufficient to enable an analysis of factors at baseline that may predispose a participant to a placebo or nonspecific response. Interestingly, we found that several factors significantly predicted response to placebo but not to citalopram. These factors included the severity of certain disruptive behaviors (primarily hyperactivity at baseline), additional measures of mood and autism symptom burden, and a measure of caregiver strain. In each case, higher scores were associated with lower placebo response. The items on the internalizing subscale of the Caregiver Strain Questionnaire include questions about caregiver mood and fatigue, and also questions that assess worry about the child’s and the family’s future. This subscale may thus serve as a proxy for reduced hopefulness or optimism at the time of study entry. In addition, the absence of a relatively higher frequency or intensity of challenging behaviors at study entry may also raise the bar for placebo response.

Figure 2. Relationship Between Baseline Composite Predictor Measures Dichotomized at the Median and Response to Treatment at Week 12

The arrows, the relative risk (RR), and the corresponding 95% CI pertain to the placebo group and are interpreted as the likelihood of response if the participant entered the study with a composite score below the median value of Disruptive Behavior (A), Mood/Autism (B), and Caregiver Strain (C). Response to citalopram hydrobromide was not affected by the baseline predictor composite score.
exclusive to the placebo condition and not a more general predictor of response overall, and that the conditions that predisposed to a favorable response rate in the placebo condition seemed not to carry over to the citalopram condition.

Thus, in the lower half of the Disruptive Behavior, Mood/Autism, and Caregiver Strain component measures, 50% of participants responded at 12 weeks when treated with placebo, whereas only about 30% of participants responded at 12 weeks when treated with citalopram. It appeared as if citalopram actually got in the way of the placebo response for some participants. However, although the participants in the citalopram group had a significantly higher number of adverse effects than did the participants in the placebo group, the adverse effects were not significantly different above or below the median splits. It remains possible that a family’s tolerance for adverse effects is somehow different in the setting of milder or more severe behaviors or their associated strain, but we are unable to address that possibility in the present study.

Placebo response is known to vary by type of disorder. Different placebo responses among clinical trials with the same disorder may provide important information leading to better selection of study participants, thus enhancing both the validity and the precision of identifying effective treatments. Interestingly, not all randomized clinical trials with children with autism show evidence of a strong placebo responder effect (Figure 1). Rather than undermining the previous assertions, these exceptions compel us to better understand the factors at play. Improvements in our understanding of these effects will lead directly to potential mitigation strategies in the design and conduct of randomized clinical trials and, possibly, to improved and more efficient assessments of treatments and their availability for this patient population.

There was no benefit to citalopram (34% response rate) relative to placebo (32% response rate) regarding the primary and secondary outcome measures in the primary analyses. We investigated whether the subpopulation of participants in the citalopram group who scored in the upper half of the Caregiver Strain composite were different from the subpopulation of participants in the placebo group who also scored in the upper half of the Caregiver Strain composite, and it is interesting to highlight that none of the secondary outcomes at week 12 from the primary analysis, with the possible exception of ABC Irritability, suggested a specific signal. This finding may indicate that the global improvement that was identified in this subpopulation was unlikely to have been associated with a distinct change in repetitive behaviors and may have been due to changes that were not adequately captured in our ratings (eg, mood or anxiety symptoms). The fact that selective serotonin reuptake inhibitors such as citalopram are so widely prescribed for the population with autism may indeed suggest that clinicians and families are seeing benefits in areas other than repetitive behavior, which was the primary focus for recruitment into this trial.

To our knowledge, this is the first study of autism and child psychopharmacology to demonstrate factors that differentially predict the likelihood of response to placebo over that of an active drug. Replication of these findings may be applied to other large autism trials. The power of future trials may be significantly enhanced by taking these factors into account, and studies designed to attempt to replicate these findings should be a priority.

In conclusion, there is a clear recognition that children and adolescents with autism spectrum disorders have serious behavioral problems and psychiatric symptoms that may be appropriate targets for pharmacotherapy. Large-scale trials to guide clinical practice are expensive, time consuming, and relatively sparse. The results of our secondary analysis do not change the finding that citalopram appeared not to be specifically advantageous compared with placebo for the treatment of repetitive behaviors. However, this analysis does underscore the need to conduct trials that are adequately powered to look for informative subgroups, and it highlights the ongoing need for placebo-controlled trials of medications commonly used for children with autism spectrum disorders to determine whether the benefits of specific drugs substantially outweigh their risks.
Baseline Factors Predicting Response to Treatment

Original Investigation Research

REFERENCES


