How Informative Are Open-Label Studies for Youth With Bipolar Disorder? A Meta-Analysis Comparing Open-Label Versus Randomized, Placebo-Controlled Clinical Trials

Joseph Biederman, MD; Carter R. Petty, MA; K. Yvonne Woodworth, BA; Alexandra Lomedico, BA; Katherine B. O'Connor, BA; Janet Wozniak, MD; and Stephen V. Faraone, PhD

ABSTRACT

Objective: To examine the informativeness of open-label trials toward predicting results in subsequent randomized, placebo-controlled clinical trials of psychopharmacologic treatments for pediatric bipolar disorder.

Data Sources: We searched journal articles through PubMed at the National Library of Medicine using *bipolar disorder, mania, pharmacotherapy, treatment* and *clinical trial* as keywords. This search was supplemented with scientific presentations at national and international scientific meetings and submitted manuscripts from our group.

Study Selection: Selection criteria included (1) enrollment of children diagnosed with *DSM-IV* bipolar disorder; (2) prospective assessment of at least 3 weeks; (3) monotherapy of a pharmacologic treatment for bipolar disorder; (4) use of a randomized placebo-controlled design *or* an open-label design for the same therapeutic compound; and (5) repeated use of the Young Mania Rating Scale (YMRS) as an outcome.

Data Extraction: The following information and data were extracted from 14 studies: study design, name of medication, class of medication, dose of medication, sample size, age, sex, trial length, and YMRS mean and standard deviation baseline and follow-up scores.

Results: For both study designs, the pooled effect size was statistically significant (open-label studies, z = 8.88, P < .001; randomized placebo-controlled studies, z = 13.75, P < .001), indicating a reduction in the YMRS from baseline to endpoint in both study designs. In a meta-analysis regression, study design was not a significant predictor of mean change in the YMRS.

Conclusions: We found similarities in the treatment effects between open-label and randomized placebo-controlled studies in youth with bipolar disorder indicating that open-label studies are useful predictors of the potential safety and efficacy of a given compound in the treatment of pediatric bipolar disorder.

J Clin Psychiatry 2012;73(3):358–365 © Copyright 2011 Physicians Postgraduate Press, Inc.

Submitted: August 11, 2010; accepted December 22, 2010. Online ahead of print: November 1, 2011 (doi:10.4088/JCP.10m06490).

Corresponding author: Joseph Biederman, MD, Massachusetts General Hospital, 55 Fruit St, Warren 705, Boston, MA 02114 (jbiederman@partners.org). **B** ipolar disorder in youth is increasingly recognized as a valid diagnosis associated with severe impairment that afflicts a sizeable minority of children and adolescents in clinical and research samples.¹⁻³ This extant literature documents that pediatric bipolar disorder is extremely morbid and commonly associated with significant functional impairment in multiple domains, including increased risks for psychiatric hospitalization, antisocial behaviors, addictions, and suicidal behaviors.⁴⁻⁷ In samples of adults with bipolar disorder, as many as 65% have had an onset of their disorder in childhood and adolescence, indicating that onset in childhood and adolescence is common.² The quest for identifying safe and effective treatments for pediatric bipolar disorder has been complex. It started with the increasing recognition that bipolar disorder in youth is a valid diagnosis associated with severe impairment.^{1,2,8,9}

Early reports on therapeutic approaches for pediatric bipolar disorder relied on clinical observations and chart reviews. One of these chart reviews¹⁰ led to the intriguing finding that traditional mood stabilizers (lithium carbonate, divalproex, and carbamazepine), known to be effective for adults with bipolar disorder, had a limited therapeutic benefit for the treatment of pediatric bipolar disorder. At the same time, more encouraging results began to emerge from case reports and chart reviews^{11,12} suggesting that atypical neuroleptics had more robust therapeutic effects in the treatment of pediatric bipolar disorder compared with the traditional mood stabilizers.

Results from case reports and chart reviews were followed by small, prospective, open-label studies of various therapeutic agents as monotherapy in the treatment of youth with bipolar disorder. Several small open-label studies suggested that traditional mood stabilizers had a relatively modest role in this population and were poorly tolerated,^{13,14} whereas atypical neuroleptics showed better efficacy and tolerability.^{15–17} These open-label trials led to large-scale randomized trials and, eventually, US Food and Drug Administration approval of risperidone and aripiprazole for the treatment of pediatric bipolar disorder (Table 1).

While the double-blind, randomized, placebo-controlled design is clearly the gold standard for clinical trials in humans,¹⁸ the conduct of large-scale, well-controlled, clinical trials is extremely expensive, and the planning process for such trials benefits from prior information about the medication, such as an estimate of its efficacy and the nature and frequency of adverse events.

The open-label study design is a useful tool for planning doubleblind, randomized, placebo-controlled trials, and it has several advantages at the early stages of investigating a new medication. Because the cost is much lower than a randomized trial, the open-label trial can provide preliminary estimates of the efficacy effect size as well as the medication's overall safety and tolerability. However, the extent to which open-label studies in pediatric bipolar disorders provide a valid estimate of treatment efficacy has not been adequately investigated in the extant literature.

The main aim of this study was to address this question using meta-analytic methods. We hypothesized that open-label studies of children and adolescents with bipolar disorder would provide accurate estimates of efficacy. To the best of our knowledge, this is the first study evaluating this issue in psychiatry.

DATA SOURCES AND STUDY SELECTION

We conducted a systematic literature search of all available prospective treatment studies examining the effect of a pharmacologic treatment in pediatric patients with bipolar disorder. We searched journal articles through PubMed at the National Library of Medicine using *bipolar disorder*, *mania*, *pharmacotherapy*, *treatment*, and *clinical trial* as keywords. This search was supplemented with additional data from scientific presentations at national and international scientific meetings and submitted manuscripts from our group.

To be included in the meta-analysis, a study had to satisfy all of the following criteria: (1) enrollment of children (less than 18 years of age) diagnosed with DSM-IV bipolar disorder; (2) prospective assessment of at least 3 weeks (minimum amount of time required for titration and to measure efficacy); (3) monotherapy of a pharmacologic treatment for bipolar disorder; (4) use of a randomized placebo-controlled design or an open-label design for the same therapeutic compound; and (5) repeated (ie, at least twice) use of the Young Mania Rating Scale (YMRS) as an outcome. We used the YMRS as the measure of efficacy because it was the most common measure used across studies of pediatric bipolar disorder. No other measures of efficacy were consistently used. We included medicines that had at least 1 open-label study and 1 placebo-controlled study (eg, oxcarbazepine had no published, open-label studies and therefore was not used). Studies of combined medicines were not included. Case reports were not included. The studies ultimately included in our analyses are summarized in Table 1.

DATA EXTRACTION

Data on the following variables were extracted from each of the studies: study design (ie, randomized placebo-controlled versus open-label), name of medication, class of medication (ie, atypical antipsychotic versus mood stabilizer), dose of medication, sample size, sample age (ie, mean age in years), sample sex (ie, percentage male), trial length (in weeks), YMRS mean and standard deviation baseline scores, and YMRS mean and standard deviation follow-up scores. The baseline YMRS standard deviation if missing. Effect sizes for the YMRS for each study were expressed as standardized mean differences. The standardized mean difference is computed by (1) taking the mean endpoint YMRS score

- While expensive and time consuming, the randomized placebo-controlled design continues to be the gold standard for asserting the safety and efficacy of drugs in humans.
- Open-label studies are useful predictors of the potential safety and efficacy of compounds for the treatment of pediatric bipolar disorder.
- Our results indicate that open-label studies can provide valid estimates of the treatment effects that would be captured in a randomized double-blind clinical trial and help inform if one is worth pursuing.

minus the mean baseline YMRS score and dividing the result by the pooled SD of the groups or (2) taking the mean YMRS change score (endpoint minus baseline) of the active drug group minus the mean YMRS change score of the placebo group and dividing by the pooled SD.

We tested the treatment effect within each study design with a random-effects meta-analysis²⁹ that allows for sampling variability within and between studies. This test has a Gaussian distribution with a null hypothesis stating that the standardized mean difference is equal to 0. To assess heterogeneity between studies, we calculated the Q statistic, which is χ^2 distributed with n-1 degrees of freedom, where *n* equals the number of studies. To assess for the bias associated with the greater likelihood of positive studies to be published compared to null studies, we used the method of Egger.³⁰ Egger's statistic will be significantly greater than 0 in the presence of publication bias. We estimated a metaanalysis regression model with the standardized mean difference as the dependent variable and study design as the independent variable. The study design variable tested whether there is a significant difference between randomized placebo-controlled studies and open-label studies in the magnitude of the YMRS treatment effect. Finally, we examined the relationship between the standardized mean differences of the open-label clinical trials with those of the randomized, placebo-controlled clinical trials by computing the correlation coefficient between the standardized mean differences estimated by the 2 types of trials. Because the standardized mean differences were skewed, we log transformed them prior to computing the correlations.

RESULTS

Sixteen studies satisfied our inclusion criteria, but 2 studies^{13,15} did not provide baseline and endpoint YMRS scores and therefore could not be included in the analysis (Table 1). Therefore, the meta-analysis included 14 studies. Two of the 14 studies (14%) were submitted manuscripts. Five of the studies had 2 arms, yielding a total of 19 observations for our analyses: 11 from open-label designs and 8 from randomized placebo-controlled designs. The total number

 Table 1. Sample and Methodological Characteristics of Randomized, Placebo-Controlled and

 Open-Label Clinical Trials of Medications for the Treatment of Youth With Bipolar Disorder

open Laber ennear mais o		weels Circo Anno Mala Duration of					
Study	Medication	Drug Class	Sample Size (baseline), n	Age, Moon v	Male, %	Duration of Trial, wk	
1	Medication	Drug Class	(Dasenne), n	Mean, y	70	111al, wK	
Randomized controlled studies							
Haas et al, ¹⁹ 2009	Risperidone	Atypical	61	13.0	43	3	
Haas et al, ¹⁹ 2009	Risperidone	Atypical	50	13.0	56	3	
Study A1281132, 2009 ²⁰	Ziprasidone	Atypical	149	13.6	56	4	
Study 149, ²¹ 2009	Quetiapine	Atypical	95	13.2	56	3	
Study 149, ²¹ 2009	Quetiapine	Atypical	98	13.2	56	3	
Tohen et al, ²² 2007	Olanzapine	Atypical	107	15.1	57	3	
Tramontina et al, ²³ 2009	Aripiprazole	Atypical	18	11.7	33	6	
Wagner et al, ²⁴ 2009	Divalproex	Mood stabilizer	74	12.9	59	4	
All controlled studies, mean			81.5	13.2	52	3.6	
Open-label studies							
Biederman et al, ¹⁶ 2005	Olanzapine	Atypical	15	5.0	67	8	
Biederman et al, ¹⁶ 2005	Risperidone	Atypical	16	5.3	75	8	
Biederman et al, ²⁵ 2005	Risperidone	Atypical	30	10.1	73	8	
Biederman et al, ²⁶ 2007	Aripiprazole	Atypical	19	11.6	58	8	
Biederman et al, ²⁷ 2007	Ziprasidone	Atypical	21	10.3	81	8	
Joshi, 2010ª	Quetiapine	Atypical	30	5.2	84	8	
Joshi, 2010 ^a	Quetiapine	Atypical	19	9.9	58	8	
DelBello et al, ¹⁷ 2008	Ziprasidone	Atypical	31	13.8	77	27	
DelBello et al, ¹⁷ 2008	Ziprasidone	Atypical	15	13.2	47	27	
Frazier et al, ¹⁵ 2001 ^b	Olanzapine	Atypical	23	10.3	57	8	
Kowatch et al, ¹³ 2000 ^b	Divalproex	Mood stabilizer	15	11.4	62	6	
Wozniak et al, ²⁸ 2009	Olanzapine	Atypical	17	10.2	67	8	
Wozniak, 2010 ^c	Divalproex	Mood stabilizer	18	8.9	83	8	
All open-label studies, mean ^d	*		21.0	9.4	70	11.5	

^aG. Joshi, MD, unpublished data, 2010.

^bNot included in meta-analysis due to unavailable data.

^cJ. Wozniak, MD, unpublished data, 2010.

dIncludes only studies in the meta-analysis.

of subjects was 883, with 231 from open-label designs and 652 from randomized placebo-controlled designs.

The samples from open-label designs were 70% male, while those of the randomized placebo-controlled designs were 52% male. The mean \pm SD age of open-label designs was 9.4 \pm 3.1 years, and the mean \pm SD age of randomized placebo-controlled designs was 13.2 \pm 0.9 years (ie, mean and standard deviation of observed mean ages). The mean length of open-label studies was more than 3 times as large as that of the randomized placebo-controlled designs (Table 1).

Open-Label Studies

The pooled estimate of the standardized mean difference effect size for the YMRS was statistically significant (standardized mean difference = 1.72; *z* = 8.88, *P* < .001 [Figure 1]), indicating a significant reduction in symptoms associated with pharmacotherapy. The test of between-study heterogeneity was also significant (χ^2_{10} = 29.76, *P* < .001, *I*² = 66.4%), indicating significant variability in the magnitude of response to different compounds.

We conducted a sensitivity analysis for the open-label studies in which the pooled estimate was repeatedly computed after omitting 1 data point (eg, 1 study) at a time. The goal of this analysis was to determine whether the significance of the combined estimate could be attributed to a single study or whether a single study had an undue influence on the overall estimate. The pooled estimate of the YMRS outcomes ranged from 1.59 to 1.81 (all 95% confidence intervals [CIs] indicated statistical significance), suggesting that no one study was heavily influencing the combined estimate. The Egger test was not statistically significant ($t_{10} = 2.21, P = .054$), which suggests no evidence for publication bias.

Baseline-to-Endpoint Standardized Mean Differences for Randomized Placebo-Controlled Studies

Next, we estimated the pooled baseline-to-endpoint effect of medication on subjects from randomized placebocontrolled studies. The pooled estimate for the YMRS was statistically significant (standardized mean difference = 2.04; z=13.75, P < .001 [Figure 1]), indicating a significant reduction in symptoms associated with pharmacotherapy. The test of between-study heterogeneity was also significant ($\chi^2_7 = 35.96$, P < .001, $I^2 = 80.5\%$), indicating significant variability in the magnitude of response to different compounds.

The sensitivity analysis for the randomized placebocontrolled studies showed the pooled estimate of the YMRS outcomes ranged from 2.21 to 2.42 (all 95% CIs indicated significance), again suggesting that no one study was heavily influencing the pooled estimate. As with the open-label studies, the Egger test was not statistically significant (t_7 = 2.34, P = .06), again failing to indicate that there was evidence of publication bias.

Drug Versus Placebo Standardized Mean Differences for Randomized Placebo-Controlled Studies

Next, we estimated the pooled drug versus placebo effect on subjects from randomized placebo-controlled studies (Figure 2). The pooled estimate for the YMRS was statistically significant (standardized mean difference = 0.71; z=6.27, P<.001 [Figure 2]), indicating a significant reduction

Figure 1. Meta-Analysis of Young Mania Rating Scale (YMRS) Change Scores by Open-Label and Randomized Placebo-Controlled Designs^a

Study	Drug	Drug Class	
Open-Label			
Biederman et al, ²⁶ 2007	Aripiprazole	Atypical antipsychotic	· · · · · · · · · · · · · · · · · · ·
Wozniak, 2010 ^b	Divalproex	Mood stabilizer	· · · · · · · · · · · · · · · · · · ·
Biederman et al, ¹⁶ 2005	Olanzapine	Atypical antipsychotic	• • • • • • • • • • • • • • • • • • •
Wozniak et al, ²⁸ 2009	Olanzapine	Atypical antipsychotic	•
Joshi, 2010 ^c	Quetiapine	Atypical antipsychotic	• · · ·
Joshi, 2010 ^c	Quetiapine	Atypical antipsychotic	
Biederman et al, ¹⁶ 2005	Risperidone	Atypical antipsychotic	
Biederman et al, ²⁵ 2005	Risperidone	Atypical antipsychotic	
Biederman et al, ²⁷ 2007	Ziprasidone	Atypical antipsychotic	
Delbello et al, ¹⁷ 2008	Ziprasidone	Atypical antipsychotic	
Delbello et al, ¹⁷ 2008	Ziprasidone	Atypical antipsychotic	
Subtotal			
Randomized Placebo-Controlled			
Tramontina et al, ²³ 2009	Aripiprazole	Atypical antipsychotic	
Wagner et al, ²⁴ 2009	Divalproex	Mood stabilizer	
Tohen et al, ²² 2007	Olanzapine	Atypical antipsychotic	•
Study 149, ²¹ 2009	Quetiapine	Atypical antipsychotic	
Study 149, ²¹ 2009	Quetiapine	Atypical antipsychotic	
Haas et al, ¹⁹ 2009	Risperidone	Atypical antipsychotic	
Haas et al, ¹⁹ 2009	Risperidone	Atypical antipsychotic	· · · · · · · · · · · · · · · · · · ·
Efficacy results, study A1281132, 2009 ²⁰	Ziprasidone	Atypical antipsychotic	
Subtotal			
Overall			
			0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5
			Standardized Mean Difference and 95% Cl

^aNo significant difference was observed in baseline-to-endpoint YMRS standardized mean differences between the 2 study designs (t=-2.08, P=.053). ^bJ. Wozniak, MD, unpublished data, 2010.

^cG. Joshi, MD, unpublished data, 2010.

Figure 2. Meta-Analysis of Drug Versus Placebo Young Mania Rating Scale Change Scores in Randomized Placebo-Controlled **Studies**^a

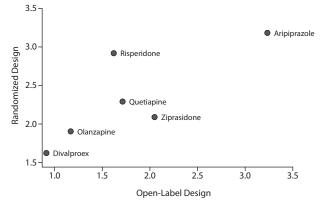
Study	Drug	Drug Class							
Tramontina et al, ²³ 2009	Aripiprazole	Atypical antipsychotic							_
Wagner et al, ²⁴ 2009	Olanzapine	Mood stabilizer ———	+			1 			
Tohen et al, ²² 2007	Quetiapine	Atypical antipsychotic			•	 	_		
Study 149, ²¹ 2009	Quetiapine	Atypical antipsychotic						•	
Study 149, ²¹ 2009	Quetiapine	Atypical antipsychotic							
Haas et al, ¹⁹ 2009	Risperidone	Atypical antipsychotic					•		
Haas et al, ¹⁹ 2009	Risperidone	Atypical antipsychotic							
Efficacy results, study A1281132, 2009 ²⁰	Ziprasidone	Atypical antipsychotic		-		•			
Overall					<	-	>		
		-0.25	0	0.25 ().50	0.75	1.00	1.25	1.50
		Standardized Mean Difference and 95% CI							

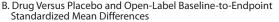
^aThe pooled standardized mean difference was statistically significant (z = 6.27, P < .001), indicating a significant reduction in symptoms associated with pharmacotherapy.

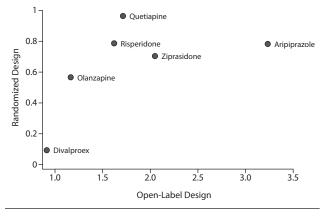
© form Psychiatry 73:3, March 2012 POSTGRADUATE PRESPSYCHIAT RISTCOMENT 2011 Physicians Postgraduate Press, IN361

Figure 3. Scatterplots of (A) Baseline-to-Endpoint Young Mania Rating Scale (YMRS) Standardized Mean Differences From Open-Label and Randomized Placebo-Controlled Designs and (B) Drug Versus Placebo and Open-Label Baseline-to-Endpoint YMRS Standardized Mean Differences

A. Baseline-to-Endpoint Standardized Mean Differences







in symptoms associated with pharmacotherapy. The test of between-study heterogeneity was also significant ($\chi^2_7 = 22.56$, P = .002, $I^2 = 69.0\%$), indicating significant variability in the magnitude of response to different compounds.

The sensitivity analysis for the drug versus placebo effect in randomized placebo-controlled studies showed the pooled estimate of the YMRS outcomes ranged from 0.64 to 0.80 (all 95% CIs indicated significance), again suggesting that no one study was heavily influencing the pooled estimate. As with the open-label studies, the Egger test was not statistically significant (t_7 =0.04, P=.97), again failing to indicate that there was evidence of publication bias.

Meta-Analysis Regression

We ran meta-analytic regression models to test the effect of study design (open-label versus randomized placebocontrolled designs) as an independent variable on the mean change in YMRS score. We found no evidence of a significant difference in baseline-to-endpoint YMRS standardized mean differences between open-label and randomized placebocontrolled studies (t=-2.08, P=.053). Although not reaching statistical significance, the randomized placebo-controlled studies had a larger pooled standardized mean difference (2.32; 95% CI, 1.99–2.65) compared to the open-label studies (1.72; 95% CI, 1.34–2.10). When we added sample mean age, drug class, baseline YMRS score, and duration of follow-up as covariates, the group difference was still not significant (P=.38).

Individual Medications

The standardized mean difference for risperidone was significantly larger in randomized placebo-controlled studies (2.92; 95% CI, 2.54–3.30) compared to open-label studies (1.62; 95% CI, 1.15–2.10). There were no other significant differences between open-label and randomized placebocontrolled studies by individual medication.

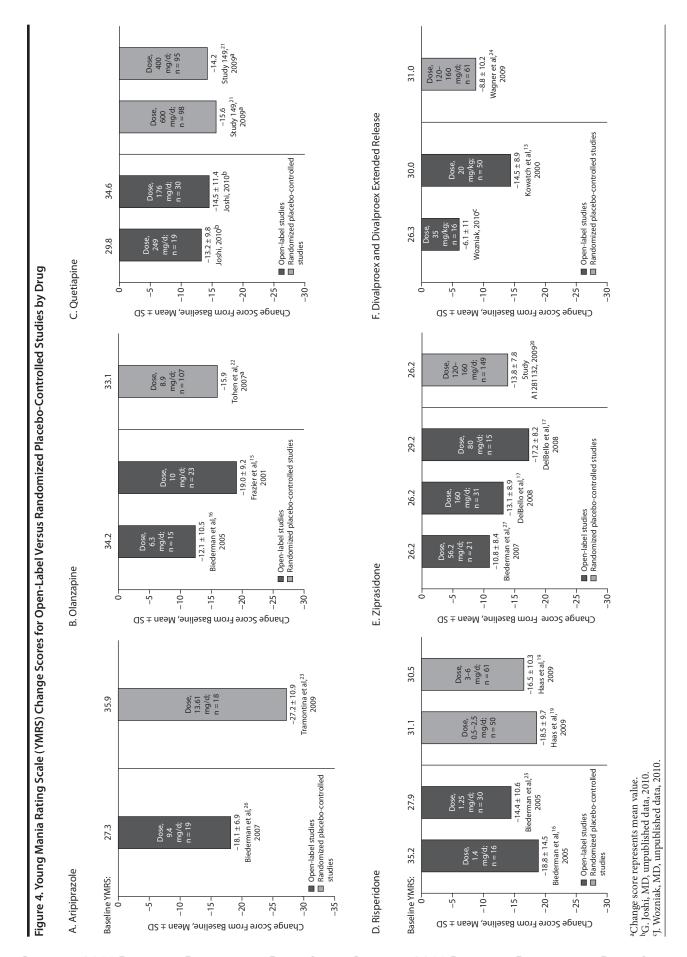
The similarity between the baseline-to-endpoint standardized mean differences from the 2 types of design can be seen in Figure 3A, which plots, for each drug, the estimated baseline-to-endpoint standardized mean differences from randomized studies against the estimated baseline-to-endpoint standardized mean differences from open-label studies along with the regression line predicting the randomized design standardized mean differences. As the figure shows, there is a good correspondence between the 2 types of estimates, with the correlation between them being 0.82 (P=.04).

We ran additional meta-analytic regression models to test the difference between the baseline-to-endpoint open-label design standardized mean differences and the drug versus placebo randomized placebo-controlled design standardized mean differences. As expected, the open-label studies had significantly larger baseline-to-endpoint standardized mean differences compared to the drug versus placebo standardized mean differences of the randomized placebo-controlled studies (t=4.09, P=.001).

The similarity between the baseline-to-endpoint openlabel standardized mean differences and the drug versus placebo standardized mean differences can be seen in Figure 3B, which plots, for each drug, the drug versus placebo standardized mean differences from randomized studies against the estimated baseline-to-endpoint standardized mean differences from open-label studies. The correlation between the 2 types of estimates was 0.70 (P=.12) (Figure 3A).

Figure 4 shows the YMRS change scores for open-label versus randomized placebo-controlled studies by drug (includes all studies from Table 1). Each panel (A through F) compares open-label (in light gray) and randomized (in dark gray) design YMRS change scores for 1 drug. The raw change scores confirm the findings from our meta-analysis, indicating a good correspondence between open-label and randomized studies. Details on dose, sample size, and baseline YMRS scores are also provided.

Comparisons of safety measures between open-label and randomized placebo-controlled studies were very limited due to the inconsistency of reported measures. Open-label studies of olanzapine showed significantly (all *P* values < .05) larger changes in weight, ^{15,28} glucose, ^{16,28} and cholesterol^{16,28}



© Com Psychiatry 73:3, Marchi 2012 Postgraduate Prespsychiatrostconght 2011 Physicians Postgraduate Press, IN363

compared to the randomized placebo-controlled study of olanzapine.²² There were no other significant differences between the 2 study designs on weight (for olanzapine and divalproex), glucose (for olanzapine and risperidone), or cholesterol (for olanzapine and risperidone).

DISCUSSION

The main aim of this study was to investigate the extent to which open-label studies in pediatric bipolar disorder can provide useful estimates of treatment effects for planning randomized, double-blind, clinical trials. We found 2 types of correspondence between the efficacy estimates from open-label and randomized trials: (1) the baseline-toendpoint effect sizes from the 2 types of trials were very similar to one another and (2) although, as expected, the drug versus placebo effect sizes from randomized trials were smaller than the open-label baseline-to-endpoint effect sizes, the 2 types of effect sizes were sufficiently correlated with one another such that they would be useful in planning clinical trials. These findings were confirmed by our examination of raw YMRS change scores (Figure 4). We found no evidence of publication biases for either type of study. These results indicate that open-label studies are useful predictors of the potential efficacy of medications for the treatment of pediatric bipolar disorder.

Figure 3A shows the high degree of correspondence between the baseline-to-endpoint effect sizes from the 2 types of studies. The open-label studies predicted the greatest efficacy for aripiprazole; modest efficacy for risperidone, quetiapine, and ziprasidone; and the lowest efficacy for divalproex and olanzapine. One would draw the same conclusion from the randomized trials. The 2 types of trials disagreed only in the relative efficacy of the 3 modestly effective medications.

Figure 3B shows that the nature of the correspondence between the 2 types of trials is similar when using the open-label baseline-to-endpoint effect sizes to predict the randomized design drug versus placebo effect sizes. For example, the very low efficacy of divalproex extended release observed in the open-label trial was consistent with the drug's failure to separate from placebo in a randomized clinical trial. Likewise, the potent effects observed for atypical neuroleptics in open-label studies were fully replicated in their robust separation from placebo observed in the randomized clinical trials.

The drug versus placebo effect sizes were smaller than the baseline-to-endpoint effect sizes. This finding was expected because the baseline versus placebo effect sizes are not corrected for placebo effects, which, as Figure 3A shows, are not negligible. Moreover, Figure 3A, like Figure 2, shows that, although the correspondence between the 2 types of designs is substantial, there is not complete agreement as to the order of efficacy of the various compounds. Thus, although our results highlight the value of open-label designs for planning randomized studies, they also emphasize that open-label designs cannot replace randomized trials. They can only be

viewed as providing preliminary evidence for efficacy along with observations on safety and tolerability.

The current results provide some quantitative insight into the transition from open-label to randomized placebocontrolled designs. By documenting that early observations gleaned from open-label studies were highly predictive of independent results observed subsequently in large-scale, multisite, placebo-controlled trials, our work highlights the value of open-label studies as a critical step in the drug development process. While clearly the randomized placebo-controlled design continues to be the gold standard for asserting the safety and efficacy of drugs in humans,¹⁸ the conduct of such large-scale clinical trials is extremely expensive and time consuming. Thus, whether a large-scale, randomized, clinical trial is worth pursuing can be informed by the results of open-label study designs.

Our results should be viewed in light of some methodological limitations. We are limited to data that have been published or made available to us. Although this sample is potentially skewed, it is reassuring that we found no evidence of publication bias or the undue influence of a single study in our results. The number of pairs of open-label studies that also had a subsequent randomized clinical trial was relatively small. Although this constrains the generalizability of our analyses, our power was reasonable because it derives from the numbers of subjects enrolled in the trials, not the number of trials. Our results showing correspondence of open-label and randomized trials were specific to drugs treating pediatric bipolar disorder, and this correspondence may differ for other drugs or other psychiatric disorders (eg, pediatric depression or autism/pervasive developmental disorder).

Despite these considerations, we found a high degree of similarities in the magnitude of treatment effects, and the relative ordering of these effects among drugs, between open-label and randomized placebo-controlled studies in youth with bipolar disorder. These results indicate that openlabel studies are useful predictors of the potential safety and efficacy of compounds for the treatment of pediatric bipolar disorder and that they can provide valid estimates of the treatment effects that would be captured in a randomized, double-blind, clinical trial.

Drug names: aripiprazole (Abilify), carbamazepine (Carbatrol, Equetro, and others), divalproex (Depakote and others), lithium (Lithobid and others), olanzapine (Zyprexa), oxcarbazepine (Trileptal and others), quetiapine (Seroquel), risperidone (Risperdal and others), ziprasidone (Geodon).

Author affiliations: Clinical and Research Program in Pediatric Psychopharmacology and Adult ADHD, Massachusetts General Hospital, Boston (Drs Biederman and Wozniak; Mr Petty; and Mss Woodworth, Lomedico, and O'Connor); Department of Psychiatry, Harvard Medical School (Drs Biederman and Wozniak), Boston, Massachusetts; and Departments of Psychiatry and Neuroscience and Physiology, SUNY Upstate Medical University, Syracuse, New York (Dr Faraone). Potential conflicts of interest: Dr Biederman is currently receiving research support from ElMindA, Janssen, McNeil, Next Wave Pharmaceuticals, and Shire. In 2011, he gave a single unpaid talk for Juste Pharmaceutical Spain; received honoraria from the Massachusetts General Hospital (MGH) Psychiatry Academy for a tuition-funded continuing medical education (CME) course and from Cambridge University Press for a chapter publication; and received departmental royalties-which were paid by Eli Lilly, Shire, and AstraZeneca and paid to the Department of Psychiatry at MGH-from a copyrighted rating scale used for attention-deficit/hyperactivity disorder diagnoses. In 2010, he received a speaker's fee from a single talk given at Fundación Dr Manuel Camelo A.C. in Monterrey Mexico; provided single consultations, which were paid to the Department of Psychiatry at MGH, for Shionogi Pharma and Cipher Pharmaceuticals; and received honoraria from the MGH Psychiatry Academy for a tuition-funded CME course. In 2009, he received a speaker's fee from Fundacion Areces (Spain), Medice Pharmaceuticals (Germany), and the Spanish Child Psychiatry Association. In previous years, he received research support, consultation fees, or speaker's fees for/ from Abbott, Alza, AstraZeneca, Bristol-Myers Squibb, Celltech, Cephalon, Eli Lilly, Esai, Forest, Glaxo, Gliatech, Janssen, McNeil, Merck, National Alliance for Research on Schizophrenia and Depression (NARSAD), National Institute on Drug Abuse, New River, National Institute of Child Health and Development, National Institute of Mental Health (NIMH), Novartis, Noven, Neurosearch, Organon, Otsuka, Pfizer, Pharmacia, Prechter Foundation, Shire, Stanley Foundation, UCB Pharma, and Wyeth. Dr Wozniak is the author of the book Is Your Child Bipolar? published May 2009, Bantam Books. She has been a consultant to Pfizer, Shire, and Eli Lilly; has received research funding from Eli Lilly and NIMH; and has served on speakers bureaus of Eli Lilly and Janssen. Her spouse, John Winkelman, MD, PhD, has been on the speakers bureau for Cephalon, Sanofi-Aventis, Sepracor, GlaxoSmithKline, and Takeda; has served on advisory boards for Pfizer, GlaxoSmithKline, Sepracor, Schwarz-Pharma, Takeda, and Boehringer-Ingelheim, and has received research support from Pfizer, GlaxoSmithKline, UCB Pharma, Boehringer-Ingelheim, and Schwarz-Pharma. Dr. Faraone receives royalties from a book published by Guilford Press: Straight Talk about Your Child's Mental Health. In the past year, he has received consulting fees and was on advisory boards for Shire Development and has received research support from Shire and National Institutes of Health (NIH). In previous years, he has received consulting fees, was on advisory boards, or participated in CME programs sponsored by Shire, McNeil, Janssen, Novartis, Pfizer, and Eli Lilly and has received research support from Eli Lilly, Shire, Pfizer, and NIH. Mr Petty and Mss Woodworth, Lomedico, and O'Connor report no conflicts of interests. Funding/support: This work was supported in part by the Pediatric Psychopharmacology Council Fund.

REFERENCES

- 1. Mick E, Biederman J, Faraone SV, et al. Defining a developmental subtype of bipolar disorder in a sample of nonreferred adults by age at onset. *J Child Adolesc Psychopharmacol.* 2003;13(4):453–462.
- Perlis RH, Miyahara S, Marangell LB, et al; STEP-BD Investigators. Longterm implications of early onset in bipolar disorder: data from the first 1000 participants in the systematic treatment enhancement program for bipolar disorder (STEP-BD). *Biol Psychiatry*. 2004;55(9):875–881.
- Geller B, Tillman R. Prepubertal and early adolescent bipolar I disorder: review of diagnostic validation by Robins and Guze criteria. *J Clin Psychiatry*. 2005;66(suppl 7):21–28.
- Biederman J, Mick E, Faraone SV, et al. A prospective follow-up study of pediatric bipolar disorder in boys with attention-deficit/hyperactivity disorder. J Affect Disord. 2004;82(suppl 1):S17–S23.
- Birmaher B, Axelson D, Strober M, et al. Clinical course of children and adolescents with bipolar spectrum disorders. *Arch Gen Psychiatry*. 2006;63(2):175–183.
- Geller B, Bolhofner K, Craney JL, et al. Psychosocial functioning in a prepubertal and early adolescent bipolar disorder phenotype. J Am Acad Child Adolesc Psychiatry. 2000;39(12):1543–1548.
- Wozniak J, Biederman J, Kiely K, et al. Mania-like symptoms suggestive of childhood-onset bipolar disorder in clinically referred children. *J Am Acad Child Adolesc Psychiatry*. 1995;34(7):867–876.
- 8. Biederman J. Pediatric bipolar disorder coming of age. *Biol Psychiatry*. 2003;53(11):931–934.
- Youngstrom EA, Birmaher B, Findling RL. Pediatric bipolar disorder: validity, phenomenology, and recommendations for diagnosis. *Bipolar Disord*. 2008;10(1, pt 2):194–214.
- Biederman J, Mick E, Prince J, et al. Systematic chart review of the pharmacologic treatment of comorbid attention deficit hyperactivity disorder in youth with bipolar disorder. *J Child Adolesc Psychopharmacol.* 1999;9(4):247–256.

- Biederman J, McDonnell MA, Wozniak J, et al. Aripiprazole in the treatment of pediatric bipolar disorder: a systematic chart review. *CNS Spectr.* 2005;10(2):141–148.
- Biederman J, Mick E, Bostic JQ, et al. The naturalistic course of pharmacologic treatment of children with maniclike symptoms: a systematic chart review. *J Clin Psychiatry*. 1998;59(11):628–637, quiz 638.
- Kowatch RA, Suppes T, Carmody TJ, et al. Effect size of lithium, divalproex sodium, and carbamazepine in children and adolescents with bipolar disorder. J Am Acad Child Adolesc Psychiatry. 2000;39(6): 713–720.
- Wagner KD, Weller EB, Carlson GA, et al. An open-label trial of divalproex in children and adolescents with bipolar disorder. *J Am Acad Child Adolesc Psychiatry*. 2002;41(10):1224–1230.
- Frazier JA, Biederman J, Tohen M, et al. A prospective openlabel treatment trial of olanzapine monotherapy in children and adolescents with bipolar disorder. J Child Adolesc Psychopharmacol. 2001;11(3):239–250.
- Biederman J, Mick E, Hammerness P, et al. Open-label, 8-week trial of olanzapine and risperidone for the treatment of bipolar disorder in preschool-age children. *Biol Psychiatry*. 2005;58(7):589–594.
- DelBello MP, Versavel M, Ice K, et al. Tolerability of oral ziprasidone in children and adolescents with bipolar mania, schizophrenia, or schizoaffective disorder. *J Child Adolesc Psychopharmacol*. 2008;18(5):491–499.
- Friedman L, Furberg C, DeMets D. Fundamentals of Clinical Trials. 3rd Edition. New York, NY: Springer-Verlag; 1998.
- Haas M, Delbello MP, Pandina G, et al. Risperidone for the treatment of acute mania in children and adolescents with bipolar disorder: a randomized, double-blind, placebo-controlled study. *Bipolar Disord*. 2009;11(7):687–700.
- Briefing Document for Psychopharmacologic Drugs Advisory Committee (PDAC) Meeting of June 9–10, 2009. Safety and efficacy supplement to NDA 20-825 regarding the treatment of pediatric patients with bipolar I disorder. Study A1281132. http://www.fda.gov/ downloads/AdvisoryCommittees/Committees/MeetingMaterials/Drugs/ PsychopharmacologicDrugsAdvisoryCommittee/UCM170736.pdf. Updated May 4, 2009. Accessed September 1, 2011.
- Advisory Committee Briefing Document for Psychopharmacologic Drugs Advisory Committee Meeting of June 9–10, 2009: Efficacy and safety of Study 149 (D1441C00149) http://www.fda.gov/ downloads/advisorycommittees/committeesmeetingmaterials/drugs/ psychopharmacologicdrugsadvisorycommittee/ucm164280.pdf. Updated May 12, 2009. Accessed September 1, 2011.
- Tohen M, Kryzhanovskaya L, Carlson G, et al. Olanzapine versus placebo in the treatment of adolescents with bipolar mania. *Am J Psychiatry*. 2007;164(10):1547–1556.
- Tramontina S, Zeni CP, Ketzer CR, et al. Aripiprazole in children and adolescents with bipolar disorder comorbid with attention-deficit/ hyperactivity disorder: a pilot randomized clinical trial. J Clin Psychiatry. 2009;70(5):756–764.
- Wagner KD, Redden L, Kowatch RA, et al. A double-blind, randomized, placebo-controlled trial of divalproex extended-release in the treatment of bipolar disorder in children and adolescents. J Am Acad Child Adolesc Psychiatry. 2009;48(5):519–532.
- Biederman J, Mick E, Wozniak J, et al. An open-label trial of risperidone in children and adolescents with bipolar disorder. J Child Adolesc Psychopharmacol. 2005;15(2):311–317.
- Biederman J, Mick E, Spencer T, et al. An open-label trial of aripiprazole monotherapy in children and adolescents with bipolar disorder. *CNS Spectr.* 2007;12(9):683–689.
- Biederman J, Mick E, Spencer T, et al. A prospective open-label treatment trial of ziprasidone monotherapy in children and adolescents with bipolar disorder. *Bipolar Disord*. 2007;9(8):888–894.
- Wozniak J, Mick E, Waxmonsky J, et al. Comparison of open-label, 8-week trials of olanzapine monotherapy and topiramate augmentation of olanzapine for the treatment of pediatric bipolar disorder. *J Child Adolesc Psychopharmacol.* 2009;19(5):539–545.
- DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986;7(3):177–188.
- Egger M, Davey Smith G, Schneider M, et al. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. 1997;315(7109):629–634.