

Adjunctive Oral Ziprasidone in Patients With Acute Mania Treated With Lithium or Divalproex, Part 2: Influence of Protocol-Specific Eligibility Criteria on Signal Detection

Gary S. Sachs, MD; Douglas G. Vanderburg, MD, MPH; Suzanne Edman, BA; Onur N. Karayal, MD, MPH; Sheela Kolluri, PhD; Mary Bachinsky, MSc; and Idil Cavus, MD, PhD

ABSTRACT

Objectives: High failure rates of randomized controlled trials (RCTs) are well recognized but poorly understood. We report exploratory analyses from an adjunctive ziprasidone double-blind RCT in adults with bipolar I disorder (reported in part 1 of this article). Data collected by computer interviews and by site-based raters were analyzed to examine the impact of eligibility criteria on signal detection.

Method: Clinical assessments and a remote monitoring system, including a computer-administered Young Mania Rating Scale (YMRS_{Comp}) were used to categorize subjects as eligible or ineligible on 3 key protocol-specified eligibility criteria. Data analyses compared treatment efficacy for eligible versus ineligible subgroups. All statistical analyses reported here are exploratory. Criteria were considered "impactful" if the difference between eligible and ineligible subjects on the YMRS change scores was ≥ 1 point.

Results: 504 subjects had baseline and ≥ 1 post-randomization computer-administered assessments but only 180 (35.7%) met all 3 eligibility criteria based on computer assessments. There were no statistically significant differences between treatment groups in change from baseline YMRS score on the basis of site-based rater or computer assessments. All criteria tested improved signal detection except the entry criteria excluding subjects with $\geq 25\%$ improvement from screen to baseline.

Conclusions: On the basis of computer assessments, nearly two-thirds of randomized subjects did not meet at least 1 protocol-specified eligibility criterion. These results suggest enrollment of ineligible subjects is likely to contribute to failure of acute efficacy studies.

Trial Registration: ClinicalTrials.gov identifier: NCT00312494

J Clin Psychiatry 2012;73(11):1420–1425
© Copyright 2012 Physicians Postgraduate Press, Inc.

Submitted: September 12, 2011; accepted July 9, 2012
(doi:10.4088/JCP.11m07389).

Corresponding author: Gary S. Sachs, MD, Bipolar Clinic and Research Program, Massachusetts General Hospital, 50 Staniford St, 5th Floor, Boston, MA 02114 (SachsG@aol.com).

The high failure rate of randomized controlled trials is a well-recognized obstacle to drug development but remains poorly understood.^{1–6} This report is a companion to the randomized, placebo-controlled, efficacy study⁷ also presented in this issue. That primary study failed to detect any statistical advantage for adjunctive ziprasidone for treatment of acute mania.⁷ It also included planned exploratory analyses intended to test hypotheses related to clinical trial methodology.⁷ While the primary study report considers placebo response in general and key subgroup analyses,⁷ this companion report focuses on the impact of 3 protocol-specified eligibility criteria on the primary outcome variable, change from baseline in the Young Mania Rating Scale (Δ YMRS) score. The analyses presented here are not intended to support or challenge the results of the primary study.

Since ziprasidone is approved for the treatment of acute mania as monotherapy^{8,9} and the need exists to gain insight into possible causes for the failure of the primary study, this exploratory study takes advantage of independent data collected by a remote monitoring system in tandem with the site-based raters.

METHOD

Study Methodology

The primary study methods are described in detail in the companion article.⁷ Briefly, this was a randomized, double-blind, placebo-controlled study (ClinicalTrials.gov identifier: NCT00312494) to evaluate the efficacy and safety of adjunctive ziprasidone therapy in patients with acute mania treated with mood stabilizers, lithium, or divalproex, conducted at 47 centers in the United States. Participants were men and women aged 18 to 65 years with a primary *DSM-IV* diagnosis of bipolar I disorder, manic or mixed episode.

At each study visit, the remote-site monitoring system collected data from site-based raters and from independently administered computer assessments. The site-based rater-administered YMRS (YMRS_{SBR}) was administered by site-based raters trained and certified on the YMRS and having ≥ 2 years of clinical experience. After completing the YMRS_{SBR}, subjects then completed an interactive computer interview that administered and scored the YMRS,¹⁰ the computer-administered YMRS (YMRS_{Comp}). The interactive computer interview presented probes similar to the sequence of questions in the scripted interview guide that site-based raters are trained to use. A prior validation study¹⁰ demonstrated high intraclass correlation coefficients of 0.91 for the YMRS_{Comp} with YMRS_{SBR} as well as intraclass correlation coefficients of 0.97 for the YMRS_{Comp} with expert consensus scores obtained by review of videotape of the site-based rater interview and no evidence of an effect of the order of administration. Cronbach α obtained in the same study indicates comparable internal consistency for YMRS_{Comp} (0.82) and YMRS_{SBR} (0.83).¹⁰

- Diagnostic uncertainty is a major issue in clinical practice, but has seldom been addressed in clinical trials.
- Among 504 subjects randomized by site-based raters, only 180 (35.7%) met all 3 of the eligibility criteria based on the computer assessments.
- The results support the notion that DSM diagnosis and other eligibility criteria matter for treatment outcome. Eligible subjects (such as those meeting DSM criteria for acute mania or mixed episodes on both assessments) tended to respond better to active treatment, but ineligible subjects tended to respond better to placebo than to active medication.

The interactive computer interview consists of probe questions in multiple-choice format. Subjects select responses to the probe questions but do not assign their own scores; dependent on subjects' responses to each probe question, the computer selects follow-up questions as necessary to map a subject's responses to a unique YMRS anchor point for each scale item. For purposes of rater monitoring, items for which tandem ratings differed by ≤ 1 point were considered to be concordant. Concordant Rater Systems (Boston, Massachusetts) contacted site-based raters by telephone to discuss potential causes for discordant ratings when the tandem YMRS total scores differed by ≥ 6 points or when > 2 items differed by ≥ 3 points. No further action was taken when site-based raters provided information supporting their ratings; however, when the discordance could not be attributed to discrepant reporting, site-based raters were scheduled to receive remediation on use of appropriate probes, scoring conventions, or both for the YMRS. In all cases, site-based raters were instructed not to change their original scores.

Sample Size

The final total primary study sample size was 669 subjects. The accompanying publication provides details of sample size estimation.⁷ The primary study was not designed or powered to detect the impact of the eligibility criteria, and all the clinical trial methodology analyses presented here are considered exploratory.

Eligibility Criteria

Among the primary study eligibility criteria, these exploratory analyses focused on 3 criteria that were considered key to selecting an appropriate sample for the efficacy study: (1) subjects must meet *DSM-IV* criteria for acute mania or a mixed episode at the screening and baseline study visits, (2) subjects must have YMRS score ≥ 18 at both the screening and baseline study visits, and (3) subjects must be excluded if their YMRS score at baseline decreased by $\geq 25\%$ from their YMRS score at the screening visit (obtained at least 3 days apart and included at least 7 days with therapeutic levels of either lithium or divalproex).

Consistent with *DSM-IV* criteria and prior successful ziprasidone monotherapy studies,^{8,9} the diagnosis of acute mania or mixed episode for these analyses was operationally defined as valid if at least 3 items on the YMRS_{Comp} were rated to be of sufficient clinical severity to count toward a *DSM-IV* diagnosis of hypomania/mania (defined as a score > 2 on YMRS items 1, 2, 3, 4, 7, 10, and 11 or a score > 4 on YMRS items 5, 6, 8, and 9). Subjects who met all 3 key criteria were considered to be eligible by the "omnibus" criterion. Subjects failing to meet any 1 or more of the key criteria were considered ineligible by the omnibus criterion.

Statistical Analysis

These analyses were conducted independently of the efficacy study analysis by Concordant Rater Systems staff (G.S.S. and S.E.). After completion of the efficacy analysis, unblinded treatment assignments were sent to Concordant Rater Systems and matched with the clinical trial methodology data files. The files were reviewed for accuracy, and analyses were carried out by using Stata version 11.0 statistical software (StataCorp LP, College Station, Texas).

We hypothesized that the separation between active ziprasidone and placebo would be greater in subjects deemed eligible than in subjects considered ineligible on the basis of the computer assessments as measured by both YMRS_{SBR} and YMRS_{Comp}.

Analysis based on site-based rater outcomes. The analysis plan called for 1-way analysis of variance (ANOVA) based on YMRS_{SBR} to compare the 3 treatment groups overall and repeated analysis for comparisons of subgroups dichotomized (as eligible or ineligible) on the basis of the computer assessment of each of the 3 key eligibility criteria, and the omnibus eligibility criteria. If an overall significant treatment effect was found on the basis of ANOVA, further pairwise treatment comparisons to placebo were evaluated.

Analysis based on computer outcomes. The analyses above were repeated on the basis of outcomes as measured by YMRS_{Comp}.

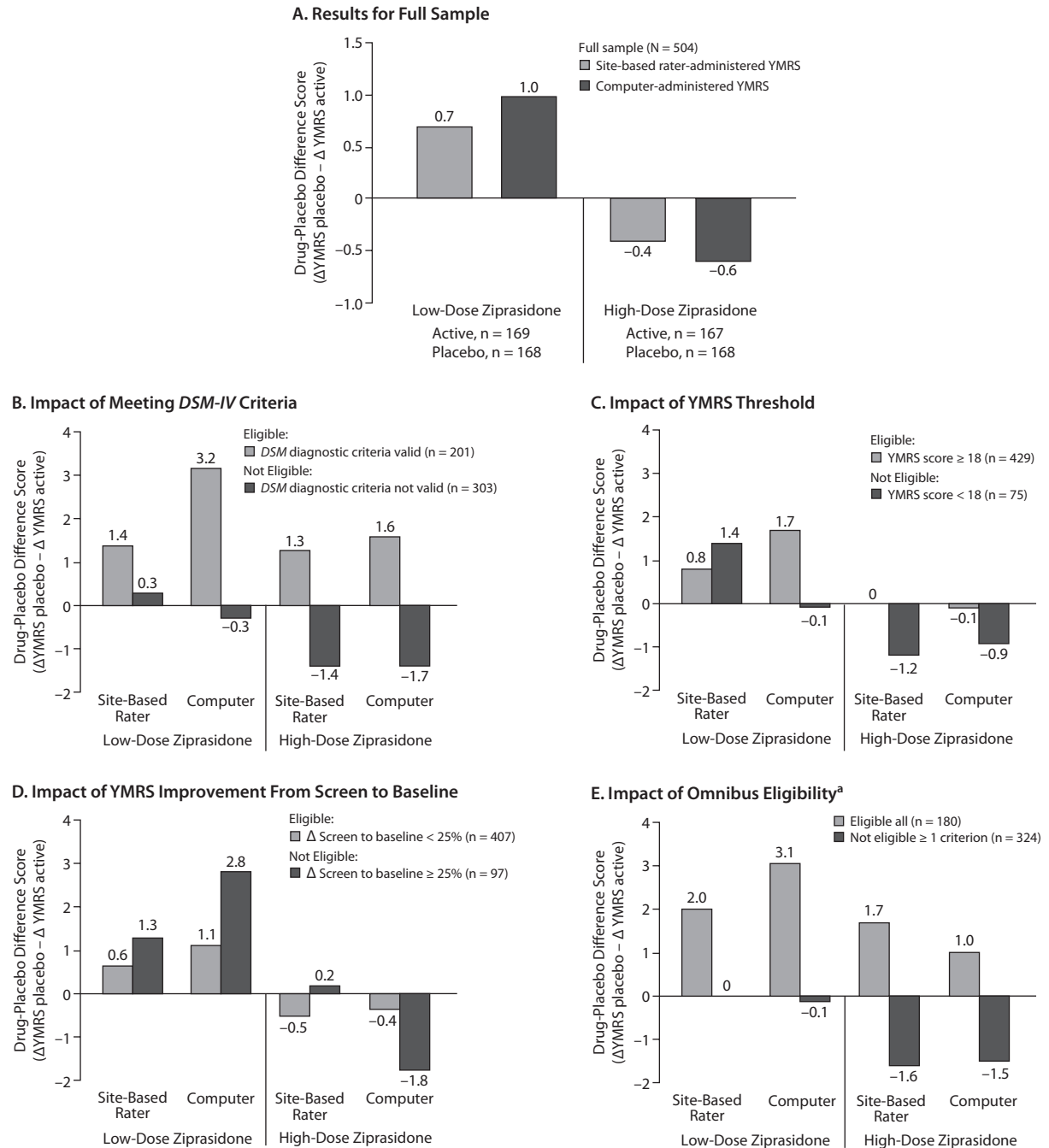
For these exploratory analyses, we operationally defined criteria to be impactful a priori if they resulted in ≥ 1 point difference in signal detection (ziprasidone-placebo difference on mean YMRS score change from baseline) between enrolled subjects meeting study eligibility criteria compared to those enrolled but not meeting the criteria. This definition arbitrarily defined impactful as those criteria that resulted in signal detection differences of $\geq 25\%$ of the estimated effect size in the primary efficacy study (based on Cohen *d*).⁷

RESULTS

Computer-Based Assessments of YMRS

One thousand two hundred three subjects were screened, and 680 were considered eligible on the basis of the site-based rater assessment and were randomized (222 to placebo [mood stabilizer + placebo], 226 to low-dose ziprasidone [mood stabilizer + low-dose ziprasidone, 20–40 mg twice daily], 232 to high-dose ziprasidone [mood stabilizer +

Figure 1. Signal Detection in Full Sample and Individual Eligibility Criteria



^aSubjects who met the omnibus criteria were eligible based on all 3 key criteria (panels B–D). Abbreviation: YMRS = Young Mania Rating Scale.

high-dose ziprasidone, 60–80 mg twice daily]). Of the 680 randomized subjects, 656 received study medication; of these, 152 did not complete the screening and baseline computer interviews. The primary efficacy analyses are detailed in the companion article.⁷ In brief, no significant differences on efficacy outcomes were found between either dose of adjunctive ziprasidone and mood stabilizer + placebo group.

Among the randomized sample, the 504 subjects with at least 1 set of computer-administered assessments prior to randomization and at least 1 postrandomization YMRS_{SBR}

and YMRS_{Comp} score comprised the clinical trial methodology data set.

Overall drug-placebo difference (Δ YMRS). Overall, very little difference was detected between the drug and placebo groups on Δ YMRS, as assessed on the site-based rater and computer-based measures (Figure 1A). The signal (defined as the difference between active drug and placebo in mean YMRS score change from baseline) from both the site-based rater and computer-based assessments was not consistent for the 2 dose groups. Overall, the YMRS score in the mood

Table 1. Impact of Eligibility Criteria Based on Computer Assessments: Comparison of Last Observation Carried Forward Site-Based Rater- and Computer-Administered YMRS Change Scores

Variable	n	Mood Stabilizer + Placebo, YMRS Change From Baseline, Mean (SD)		n	Mood Stabilizer + Low-Dose Ziprasidone, YMRS Change From Baseline, Mean (SD)		n	Mood Stabilizer + High-Dose Ziprasidone, YMRS Change From Baseline, Mean (SD)		n	Overall Treatment Effect From ANOVA, P Value	
		Site-Based Rater	Computer		Site-Based Rater	Computer		Site-Based Rater	Computer		Site-Based Rater	Computer
Full sample	504	-9.4 (8.5)	-6.9 (9.3)	168	-10.1 (9.0)	-7.9 (10.8)	169	-9.0 (8.9)	-6.3 (10.5)	167	>.51	>.33
Computer-administered YMRS score \geq 18 at screening and baseline												
Eligible ^a	429	-9.6 (8.5)	-7.8 (9.2)	149	-10.4 (9.2)	-9.5 (10.6)	141	-9.6 (9.1)	-7.7 (9.9)	139	>.71	>.21
Not eligible ^a	75	-7.4 (8.1)	0.1 (7.1)	19	-8.8 (8.3)	0.2 (8.5)	28	-6.2 (7.6)	1.0 (10.7)	28	>.47	>.93
DSM-IV criteria for acute manic or mixed episode at screening and baseline												
Eligible ^a	201	-9.4 (9.5)	-8.6 (10.1)	70	-10.8 (10.1)	-11.8 (10.3)	68	-10.7 (10.0)	-10.2 (9.2)	63	>.64	>.17
Not eligible ^a	303	-9.4 (7.7)	-5.6 (8.5)	98	-9.7 (8.2)	-5.3 (10.5)	101	-8.0 (8.1)	-3.9 (10.6)	104	>.26	>.43
YMRS score improved \geq 25% from screening to baseline												
No (eligible) ^a	407	-9.5 (8.7)	-8.2 (9.1)	139	-10.1 (9.5)	-9.3 (11.1)	130	-9.0 (9.3)	-7.8 (9.8)	138	>.62	>.47
Yes (not eligible) ^a	97	-9.0 (7.4)	-0.6 (7.7)	29	-10.3 (7.6)	-3.4 (8.5)	39	-9.2 (7.0)	1.2 (11.0)	29	>.74	>.12
All 3 eligibility criteria												
Yes (eligible) ^a	180	-9.1 (9.9)	-9.4 (10.2)	61	-11.1 (10.5)	-12.5 (10.4)	60	-10.8 (10.0)	-10.4 (9.4)	59	>.51	>.22
No (not eligible) ^a	324	-9.6 (7.5)	-5.5 (8.5)	107	-9.6 (8.1)	-5.4 (10.3)	109	-8.0 (8.1)	-4.0 (10.5)	108	>.25	>.48

^aOn the basis of computer-based rating.

Abbreviations: ANOVA = analysis of variance, YMRS = Young Mania Rating Scale.

stabilizer + low-dose ziprasidone group decreased slightly more than the placebo group, but the mood stabilizer + high-dose ziprasidone treatment group improved less than the placebo group.

Eligibility Based on Computer Assessments

On the basis of the operational criteria (derived from the YMRS_{Comp} item scores), 303 subjects did not meet the key study eligibility criteria requiring fulfillment of DSM-IV mania or mixed episodes criteria at screening and baseline. For 75 subjects, YMRS_{Comp} scores at screening or baseline were less than the threshold score of 18 required for study entry. YMRS_{Comp} score improvement of 25% or more between screening and baseline would have excluded 97 subjects. The omnibus criteria for study eligibility were satisfied by 180 of the 504 subjects (35.7%) deemed eligible by the site-based raters. While the majority of randomized subjects (60.1%) did not satisfy the minimal requirement for meeting DSM-IV criteria for a manic or mixed episode at screening and baseline, all 201 subjects who did satisfy the DSM-IV diagnosis requirement also scored \geq 18 at screening and baseline on the YMRS_{Comp}.

Impact of Eligibility Criteria on Efficacy Outcome

Although none of the Δ YMRS comparisons for either of the ziprasidone dose groups to placebo reached statistical significance, the key eligibility criteria on signal detection met the operational definition of impactful for most comparisons (see Figure 1B–D).

When examining the impact of meeting DSM-IV criteria for mania (see Figure 1B), all comparisons were impactful, and a numerically larger YMRS_{Comp} change was seen among those patients with a computer-validated acute mania diagnosis (mood stabilizer + low-dose ziprasidone vs mood stabilizer + placebo = 3.2; mood stabilizer + high-dose ziprasidone vs mood stabilizer + placebo = 1.6) compared

with those without a valid diagnosis (mood stabilizer + low-dose ziprasidone vs mood stabilizer + placebo = -0.3; mood stabilizer + high-dose ziprasidone vs mood stabilizer + placebo = -1.7). Among subjects without a computer-validated diagnosis, a small signal was observed for low-dose ziprasidone-treated subjects, but the site-based rater and computer assessments generated a modest signal favoring placebo over high-dose ziprasidone treatment.

All subjects with YMRS_{Comp} scores < 18, also failed to meet the study eligibility criteria based on the DSM-IV mania or mixed criteria. The YMRS threshold criteria appear impactful for Δ YMRS_{SBR} in the high-dose ziprasidone subgroup and for Δ YMRS_{Comp} in the low-dose ziprasidone group. In contrast, the trend observed in the signal for low-dose ziprasidone versus placebo from site-based rater ratings was of slightly higher magnitude in ineligible subjects compared to the signal for eligible subjects (Figure 1C).

The effect of eligibility criteria excluding subjects with > 25% improvement in YMRS screening to baseline is shown in Figure 1D. This criterion was impactful, but, in the low-dose ziprasidone group, the signal from ineligible subjects was numerically higher than the signal from subjects considered eligible. In the high-dose ziprasidone group, little difference was observed in signal detection based on YMRS_{SBR}, but, in detection based on YMRS_{Comp}, a modest trend was found favoring better response to placebo among subjects considered ineligible on the basis of this criterion.

The impact of the omnibus eligibility criteria is shown in Figure 1E. Among the 180 subjects meeting all 3 computer-based eligibility criteria, subjects receiving either dose of ziprasidone demonstrated greater numeric improvement than those receiving placebo. In contrast, in the subsample of 324 ineligible subjects, negligible differences were seen between low-dose ziprasidone versus placebo, but the signal from the high-dose ziprasidone versus placebo comparison

was numerically more favorable for placebo. The omnibus criteria was impactful in all comparisons (see Figure 1E).

Table 1 details the impact of applying an individual criterion or the omnibus eligibility criteria based on computer-assessment on the YMRS change scores as rated by the site-based rater and the computer for all treatment groups.

DISCUSSION

The failure to demonstrate efficacy of adjunctive ziprasidone in the primary study suggests that ziprasidone may be ineffective as adjunct therapy for acute mania. Lack of differentiation between drug and placebo effects in clinical trials can, however, have many other causes, including poor study design, failure of randomization to balance variables moderating response between groups, inappropriate enrollment, and rating reliability.^{4,6,11}

To our knowledge, this is the first report using independent computer-administered assessments to evaluate the impact of eligibility criteria in a randomized clinical trial. The consistency and lack of rater bias in the computer assessments is the main strength of this article. There are, however, important limitations to consider. First, the primary study failed and, as noted in the companion article, many factors other than those considered here may have contributed to the lack of drug-placebo difference in the study. Second, the efficacy study was neither designed nor powered to detect the impact of the eligibility criteria. Therefore, all the analyses here should be considered exploratory. Furthermore, the computer assessments and the remote site monitoring system were used to manage all study sites and may have altered the rater and subject behaviors in a manner that obscured effects of some practices in clinical trials without such monitoring and management interventions. Hence, the data here cannot be interpreted as indicating any benefit of the site monitoring system itself. In addition, computer-administered assessments for subjects with psychiatric disorders are a relatively new methodology, which might be subject to the influence of unforeseen respondent adaptation in the future. Despite these limitations, we believe the results of this study can be instructive to the design of future clinical trials.

Perhaps the most striking finding in these exploratory analyses is that the majority (nearly 65%) of subjects randomized on the basis of computer assessments did not meet the key protocol-specified eligibility requirements related to diagnosis and severity of illness. Although none of the subgroups defined by the computer-based criteria were associated with statistically significant separation between active ziprasidone and placebo, the data suggest that the inclusion of many ineligible patients contributed to the failure of the efficacy study. Among the 180 subjects meeting all the computer-based eligibility criteria, subjects receiving either dose of ziprasidone demonstrated greater numeric improvement compared with those receiving placebo. Although not statistically significant, the observed signal is of a magnitude similar to the estimated effect size on which the primary

study was powered. Despite the relatively small sample deemed eligible on the basis of the computer assessments, an effect of this size is potentially important. In contrast, in the subsample of 324 ineligible subjects, placebo treatment was numerically superior to treatment with either dose of ziprasidone. Inclusion of subjects likely to have better response to placebo than active drug is a serious handicap for any clinical trial.

Among the individual eligibility criteria examined, the data suggest that the criterion with the most robust effect on drug-placebo separation was whether subjects met *DSM-IV* criteria for a current mixed or manic episode based on YMRS_{Comp}. As seen in Figure 1B, the difference in separation between drug (mood stabilizer + high-dose ziprasidone group) and placebo for subjects satisfying this criterion versus subjects not meeting the *DSM-IV* criteria as assessed by the computer was as much as 2.7 points based on the site-based rater scores and 3.3 points based on the computer scores.

These post hoc analyses show that 75 subjects randomized by site personnel fell below the threshold YMRS score necessary for inclusion when the same test was administered by computer. Notably, the samples were not balanced with respect to this factor, and proportionally more ineligible subjects were randomized to the adjunctive ziprasidone groups. In general, subjects with more severe symptomatology respond better to the treatment, and, thus, the inclusion of low-severity subjects might be expected to reduce the likelihood of demonstrating an effect.

The findings among low-dose ziprasidone subjects, with baseline YMRS_{Comp} scores below the eligibility threshold (YMRS score of 18) seem contradictory. On the basis of the site-based rater outcomes, a signal with somewhat higher amplitude was found in the ineligible sample compared to the signal found in eligible subjects. On the basis of the computer assessments, however, the signal has higher amplitude for the eligible sample. This apparent contradiction may be a simple consequence of defining eligibility at baseline in terms of the computer scores. Since among ineligible subjects, baseline YMRS_{SBR} score (mean = 23.8) was by definition higher than YMRS_{Comp} score (mean = 14.4), it is not possible to discern the extent to which the contradictory findings are an artifact of bias related to baseline inflation by site-based raters or an expected insensitivity to change among subjects with low baseline YMRS_{Comp} scores. The latter is consistent with the very low placebo response observed with YMRS_{Comp} for subjects below the severity threshold at baseline.

Importantly, these data do not establish that computer assessments of any of the eligibility criteria are superior to those of a well-trained site-based rater. Our study lacked data on subjects that might have been found eligible by the computer and considered ineligible by the site-based rater. Therefore, our exploratory conclusions should be understood as limited to the context of tandem ratings. In this context, the data suggest better signal detection in subjects for whom eligibility has been established with higher confidence (agreement between site-based rater and computer-administered assessments).

Data from computer-administered ratings may provide a consistent metric across subjects, sites, and time for quantification of disease severity and for evaluating eligibility criteria such as the diagnostic criteria for current episodes of mania. Tandem assessment of manic symptoms, including independent evaluations by the site-based rater and computer, offers a potential means of evaluating study performance. Further refinements of the concordance analyses may facilitate progress in testing hypotheses about causes of study failure, cultural influences, and other potentially interesting questions about rater training and novel clinical trial methodology.

In conclusion, comparison of computer and site-based ratings suggests that contributing factors related to eligibility impacted the results of the primary study. Not meeting operational criteria for *DSM-IV* diagnosis based on the computer assessments had a particularly marked impact. The findings suggest areas in which improvements in methodology may enhance the ability of future studies to detect true drug effects.

Drug names: divalproex (Depakote and others), lithium (Lithobid and others), ziprasidone (Geodon and others).

Author affiliations: Massachusetts General Hospital (Dr Sachs) and Concordant Rater Systems (Dr Sachs and Ms Edman), Boston; Medicines Development Group (Dr Vanderburg), Specialty Neuroscience (Dr Karayal), and Statistics (Specialty Care) (Dr Kolluri), Pfizer Inc, New York, New York; and Specialty Care Neuroscience (Dr Cavus) and Medicines Development Group (Ms Bachinsky) Pfizer Inc, Groton, Connecticut.

Author contributions: Clinical trial methodology analyses were conducted by Dr Sachs and Ms Edman.

Potential conflicts of interest: Dr Sachs is an employee of Concordant Rater Systems (Bracket) and Massachusetts General Hospital; is a consultant to Astellas, AstraZeneca, Bristol-Myers Squibb, DSP, Otsuka, Pfizer, Sepracor, Takeda and Wyeth; has received grant/research support from Repligen; has served on speakers or advisory boards of Astellas, Bristol-Myers Squibb, GlaxoSmithKline, Sanofi, Pfizer, Sepracor, Takeda, and Wyeth; and is a stock shareholder in Concordant Rater Systems. Drs Vanderburg, Karayal, Kolluri, and Cavus are employees of and stock

shareholders in Pfizer Inc. Ms Edman is an employee of Concordant Rater Systems. Ms Bachinsky is an employee of Pfizer.

Funding/support: This study was sponsored by Pfizer. The article was written by the authors with editorial assistance from Hilary Bennett, MSc, Karen Vondy, PhD, and Hajira Koeller, PhD, of PAREXEL, which was funded by Pfizer Inc.

Previous presentation: Concepts included in this article were previously presented in poster form at the 48th Annual New Clinical Drug Evaluation Unit Meeting; May 27–30, 2008; Phoenix, Arizona.

REFERENCES

1. Sysko R, Walsh BT. A systematic review of placebo response in studies of bipolar mania. *J Clin Psychiatry*. 2007;68(8):1213–1217.
2. Merlo-Pich E, Gomeni R. Model-based approach and signal detection theory to evaluate the performance of recruitment centers in clinical trials with antidepressant drugs. *Clin Pharmacol Ther*. 2008;84(3):378–384.
3. Merlo-Pich E, Alexander RC, Fava M, et al. A new population-enrichment strategy to improve efficiency of placebo-controlled clinical trials of antidepressant drugs. *Clin Pharmacol Ther*. 2010;88(5):634–642.
4. Fava M, Evins AE, Dorer DJ, et al. The problem of the placebo response in clinical trials for psychiatric disorders: culprits, possible remedies, and a novel study design approach. *Psychother Psychosom*. 2003;72(3):115–127.
5. Quitkin FM, Petkova E, McGrath PJ, et al. When should a trial of fluoxetine for major depression be declared failed? *Am J Psychiatry*. 2003;160(4):734–740.
6. Greist JH, Mundt JC, Kobak K. Factors contributing to failed trials of new agents: can technology prevent some problems? *J Clin Psychiatry*. 2002;63(suppl 2):8–13.
7. Sachs GS, Vanderburg DG, Karayal ON, et al. Adjunctive oral ziprasidone in patients with acute mania treated with lithium or divalproex, pt 1: results of a randomized, double-blind, placebo-controlled trial. *J Clin Psychiatry*. 2012;73(11):1412–1419.
8. Keck PE Jr, Versiani M, Potkin S, et al; Ziprasidone in Mania Study Group. Ziprasidone in the treatment of acute bipolar mania: a three-week, placebo-controlled, double-blind, randomized trial. *Am J Psychiatry*. 2003;160(4):741–748.
9. Potkin SG, Keck PE Jr, Segal S, et al. Ziprasidone in acute bipolar mania: a 21-day randomized, double-blind, placebo-controlled replication trial. *J Clin Psychopharmacol*. 2005;25(4):301–310.
10. Reilly-Harrington NA, DeBonis D, Leon AC, et al. The interactive computer interview for mania. *Bipolar Disord*. 2010;12(5):521–527.
11. Greist J, Mundt J, Jefferson J, et al. Comments on “why do clinical trials fail? the problem of measurement error in clinical trials: time to test new paradigms?” *J Clin Psychopharmacol*. 2007;27(5):535–537.