International Journal of Neuropsychopharmacology (2011), **14**, 1017–1027. © CINP 2011
doi:10.1017/S146114571100040X

ARTICLE

THEMATIC SECTION
Bipolar Disorder
and its Treatment

CINP

Impact of geographical and cultural factors on clinical trials in acute mania: lessons from a ziprasidone and haloperidol placebo-controlled study

Eduard Vieta[1], Elizabeth Pappadopulos[2], Francine S. Mandel[2] and Ilise Lombardo[2]

[1] *Clinical Institute of Neuroscience, Hospital Clinic, University of Barcelona, IDIBAPS, CIBERSAM, Barcelona, Spain*
[2] *Pfizer Inc., New York, NY, USA*

Abstract

Clinical trials today are conducted in multiple countries to enhance patient recruitment and improve efficiency of trials. However, the demographic and cultural diversity may contribute to variations in study outcomes. Here we conducted *post-hoc* analyses for a placebo-controlled study with ziprasidone and haloperidol for the treatment of acute mania to address the demographic, dosing, and outcome disparities in India, Russia and the USA. We compared the baseline characteristics, outcomes and discontinuations in patients and explored the relationship between the outcome measures across these countries. We found substantial differences in baseline characteristics of subjects, administered dosage and disease severity in India compared to the USA and Russia. Conversely, US subjects had a higher placebo response compared to subjects in Russia and India. These results are probably due to demographic differences in patient populations and psychiatric clinical practice across countries. While we offer initial ideas to address the disparities identified in this analysis, it is clear that further research to improve our understanding of geographical differences is essential to ensure globally applicable results for clinical trials in psychiatry.

Received 4 November 2010; Reviewed 28 November 2010; Revised 17 January 2011; Accepted 3 February 2011; First published online 14 March 2011

Key words: Bipolar mania, clinical design, geography, placebo response.

Introduction

International trials are designed to be conducted consistently in all countries. However, prescribing practices and cultural differences may affect enrolment and outcomes, which could have important implications for drug development and the design of international clinical trials. For instance, in mania, a flexible-dose risperidone trial conducted in India revealed that most patients were given doses very close to the upper limit, despite their low body mass index (BMI) (Khanna *et al*. 2005), and a failed aripiprazole trial indicated differential rater performance across countries as potential reason for study failure (El Mallakh *et al*. 2010).

Address for correspondence: Dr E. Vieta, Clinical Institute of Neuroscience, Hospital Clinic, University of Barcelona, IDIBAPS, CIBERSAM, Barcelona, 08036, Spain.
Tel.: +34 932275400 *Fax*: +34 932275795
Email: evieta@clinic.ub.es

A recent randomized, placebo-controlled trial assessing the efficacy of ziprasidone in acute mania revealed that ziprasidone and haloperidol were independently superior to placebo (Vieta *et al*. 2010). The original trial was a 12-wk, double-blind study in 438 patients with bipolar mania started with a 3-wk comparison of ziprasidone (80–160 mg/d) with placebo or haloperidol (8–30 mg/d). This was followed by a 9-wk extension phase during which subjects continued on ziprasidone or haloperidol, and those originally on placebo were switched (with the appropriate escalation) to ziprasidone. Superior changes from baseline in Mania Rating Scale (MRS) scores were observed for ziprasidone and haloperidol compared to placebo from day 2 to week 3 and efficacy was maintained throughout the 9-wk extension phase. Although haloperidol showed greater efficacy than ziprasidone, the latter showed a better tolerability profile (Vieta *et al*. 2010). We were particularly interested in country variations related to placebo response, dosing, and

tolerability. We hypothesized that the baseline subject characteristics, dosage of study drugs and placebo response differed between the countries. We also queried whether there were any differences by country for overall measured outcome variables or when restricting analyses to severely affected subjects. Further *post-hoc* analysis of the results of this study revealed some variation in outcomes and adverse events (AEs) results between countries, namely India, Russia, and the USA, which are the objective of this report.

Methods

Subjects

Data from a 12-wk, double-blind, two-part study in 438 adults with acute bipolar mania and a MRS score of >14 at screening (with scores of $\geqslant 2$ on at least four items) were analysed. Subjects received flexibly dosed ziprasidone (80–160 mg/d), haloperidol (8–30 mg/d) or placebo during the first 3 wk. During the subsequent 9-wk extension phase patients either continued with ziprasidone (40–160 mg/d) or haloperidol (8–30 mg/d). Due to the study design, some of the subjects were on ziprasidone for a maximum of 9 wk, while others for a maximum of 12 wk. In order to eliminate any confound, the *post-hoc* analyses were restricted to the first 3 wk of the trial for all reported results. For this analysis, subjects with at least a 50% decrease in MRS score from baseline to week 3 were defined as MRS responders.

Statistical analyses

Baseline comparisons

A stepwise discriminant function was used to determine the combination of baseline variables that could distinguish between the three countries (USA, India, Russia). The variables derived from patient history included age of onset of illness, BMI, duration of illness, mixed *vs.* manic at baseline, number of prior hospitalizations, number of prior psychotropic medications and psychotic status at baseline. Other variables based on psychiatric evaluations included baseline scores for MRS, Clinical Global Impression – Severity (CGI-S), Global Assessment of Functioning (GAF), Positive and Negative Syndrome Scale (PANSS) Negative, baseline PANSS Positive, baseline PANSS Total, hallucinations [Schedule for Affective Disorders and Schizophrenia – Change (SADS-C), item 42] and Clinical Global Impression – Improvement (CGI-I).

Dosage and outcome variables

An analysis of variance (ANOVA) model was used to compare the average baseline weight with main effects for each treatment group, country and an interaction between treatment and country. The relationship between the outcome variables and dose of each group was examined using multiple regression models using last dose as a predictor for each country and treatment group separately.

Next, separate multiple regression models were examined to determine if these outcome variables could successfully predict a therapeutic improvement, as measured by MRS change from baseline to endpoint [i.e. the last observation carried forward (LOCF) for each country]. An ANOVA was used to compare the average MRS change from baseline to LOCF endpoint with main effects for treatment and country, and an interaction between treatment and country, while controlling for baseline MRS score. The effects sizes by country were calculated according to Cohen's d, defined as placebo-corrected treatment effect/root mean square.

In order to account for variability in baseline severity between the countries, additional logistic regression analyses were conducted, where the MRS at baseline was $\geqslant 30$.

To determine if geography had any placebo effect, Fisher's exact tests were then used to compare the proportion of responders in the placebo group between countries. This analysis was extended to all treatment groups using a logistic regression model to compare the proportion of responders (MRS change from baseline $\geqslant 50\%$) with main effects for treatment and country and to test for an interaction between treatment and country.

Relationship between the outcome variables

In order to investigate if there was any consistent relationship between the outcome variables, a multiple regression model was used with main effects for country and treatment. Pairwise comparisons were used to determine which countries and treatments were different. To assess the precision of clinical evaluation, MRS score at visit was correlated with CGI-S at the same visit, and MRS change from baseline to visit was correlated with CGI-I at the same visit.

Discontinuations and AEs

Estimates of time until discontinuation were made using Kaplan–Meier analysis and these curves were compared using a log-rank test. Rate of discontinuation

Table 1. Descriptive statistics for the baseline characteristics of the three countries

	India ($n=166$)	Russia ($n=123$)	USA ($n=112$)
Subject baseline characteristics[a]			
Age (yr)	24.82 (8.08)	25.01 (9.55)	21.70 (9.04)
BMI	22.17 (3.66)	24.85 (5.02)	28.38 (5.25)
Duration of illness (yr)	9.61 (8.17)	13.29 (10.02)	17.74 (10.90)
Proportion manic (*vs.* mixed)	0.98 (0.15)	0.83 (0.38)	0.53 (0.50)
No. of prior hospitalizations	3.02 (1.64)	6.04 (8.67)	9.13 (14.19)
No. of prior psychotropic medications	2.96 (1.92)	3.59 (2.17)	4.46 (2.50)
Proportion psychotic	0.45 (0.50)	0.35 (0.48)	0.29 (0.46)
Psychiatric diagnostic scores[a]			
MRS	34.48 (7.03)	30.43 (7.63)	24.07 (5.45)
CGI-S	5.02 (0.75)	4.83 (0.78)	4.40 (0.65)
GAF	35.83 (11.34)	40.94 (11.04)	47.01 (8.72)
PANSS Negative	8.19 (2.46)	9.77 (3.67)	12.82 (4.09)
PANSS Positive	20.50 (6.18)	18.12 (5.39)	16.88 (4.05)
PANSS Total	56.34 (14.31)	57.52 (12.89)	63.27 (11.80)
Hallucinations (SADS-C, item 42)	0.42 (1.02)	0.37 (0.93)	0.71 (1.14)

BMI, Body mass index; MRS, Mania Rating Scale; CGI-S, Clinical Global Impression – Severity; GAF, Global Assessment of Functioning; PANSS, Positive and Negative Syndrome Scale; SADS-C, Schedule for Affective Disorders and Schizophrenia – Change.
Values are mean (S.D.)
[a] $p<0.0001$ for all comparisons in the table, except for the number of prior hospitalizations.

due to lack of efficacy was measured by reported AEs or lack of efficacy, exacerbation of mania or worsening of disease, and was compared across the countries using χ^2 tests of equivalence. The average numbers of AEs experienced per subject were compared using an ANOVA model for main effects of country and treatment and interactions between country and treatment.

Results

Baseline subject characteristics across countries

A total of 437 subjects were available for analysis (safety population). We report the discriminant function analysis, with data for only 401 subjects due to missing data, for predictors related to subject characteristics that included age, BMI, duration of illness, proportion of manic *vs.* mixed episodes, proportion of psychotic subjects, number of prior psychotropic drugs and hospitalizations, and psychiatric diagnostic scales including MRS, CGI-S, GAF, SADC item 42 for hallucinations, PANSS Negative, PANSS Positive and PANSS Total scores. With the exception of the number of prior hospitalizations, all other predictors added significantly to the discrimination among the three countries, with $p<0.0001$ in each case. The USA had the youngest subject population while Russia had the oldest. The mean BMI at baseline was highest for patients from the USA and lowest for patients from India.

US subjects had been diagnosed with their illness for the longest period of time followed by Russia and then India (Table 1). However, the proportion of psychotic subjects at baseline was highest in India with the USA having the lowest numbers. Similarly, the proportion of subjects with manic (*vs.* mixed) symptoms was high in India (98%) and Russia (83%), while in the USA (53%) subjects were evenly split. In contrast, the US group had a longer duration of illness, the most prior psychotropic medications and the greatest number of prior hospitalizations. The least number of prior psychotropic medications and prior hospitalizations were observed in India (Table 1).

Illness severity was measured at baseline using psychiatric diagnostic scales, which included MRS, which was the primary measure of bipolar mania, and CGI-S, GAF and PANSS Total scores, which were secondary measures. Based on these data, it appears

that subjects in the USA had the least severe bipolar manic symptoms. While the study inclusion criteria required MRS scores of >14, the subjects in India had the worst disease severity, as mean baseline MRS scores were considerably higher (34.48) than in Russia (30.43) or the USA (24.07). These findings were confirmed with additional diagnostics where Indian subjects had the lowest GAF scores and US subjects the highest. In contrast, the CGI-S and PANSS Positive scores were highest in Indian subjects and lowest in US subjects. The PANSS Negative scores and the PANSS Total scores were highest in US and Indian subjects. The CGI-S scores were approximately equal for subjects in Russia and India. Finally, US subjects reported stronger hallucinations (SADS-C, item 42) while Russian subjects were the least affected (Table 1). Overall, in this study US subjects had been diagnosed and treated for longer, while Indian subjects had the highest disease severity.

Baseline weight of subjects and dosage analysis

As observed in the previous analyses, the baseline BMI was remarkably different between countries. Specifically, it was expected that there would be a significant difference in the average baseline weight across the countries (each was significantly different from the other at $p<0.001$). As previously indicated, subjects in the USA were the heaviest at baseline and those in India were the lightest. Notably, there was no interaction between treatment and country and there was no difference in average weight within the treatment groups (Table 2).

Given these differences in baseline weight, it could be expected that the doses prescribed were proportional to the baseline weights. However, the mean doses of ziprasidone at week 3 varied by country, with subjects in Russia receiving lower doses of ziprasidone than subjects in India or the USA. Similarly, the mean dose of haloperidol at week 3 was lower in Russia than in India or the USA. There was a significant difference in the modal dose (regardless of adjustment factor) between India and both USA and Russia. The modal doses in India were significantly higher than in both USA and Russia. There was no significant difference between the adjusted modal dose in USA and Russia (Table 2).

Table 2. Baseline weight and dosage analysis by country (intent-to-treat population) at week 3

	India	Russia	USA
N			
Ziprasidone	74	56	47
Haloperidol	69	55	46
Placebo	36	28	24
LS mean weight, kg (S.E.)			
Ziprasidone	57.95 (1.11)	73.50 (1.78)	82.38 (2.52)
Haloperidol	56.89 (1.59)	73.88 (1.79)	82.36 (2.54)
Placebo	57.13 (1.15)	60.04 (2.51)	83.71 (3.56)
Mean dosage at week 3 (mg/d)			
Ziprasidone	128.4	121.8	126.5
Haloperidol	20.7	15.2	15.3
Placebo	–	–	–
Mean modal dose per kg (S.D.)			
Ziprasidone	2.41 (0.72)[ab]	1.53 (0.51)[a]	1.58 (0.59)[b]
Haloperidol	0.39 (0.20)[ab]	0.24 (0.25)[a]	0.22 (0.13)[b]
Placebo	–	–	–
Mean modal dose/BMI (S.D.)			
Ziprasidone	6.36 (2.00)[ab]	4.36 (1.45)[a]	4.62 (1.66)[b]
Haloperidol	1.00 (0.49)[b]	0.69 (0.80)	0.61 (0.35)[b]
Placebo	–	–	–

LS, least squares; BMI, body mass index.
[a] $p<0.001$ India *vs.* Russia.
[b] $p<0.001$ India *vs.* USA.

Outcome variables across countries

Due to the differences in the baseline disease severity and dosage between countries, the next analyses aimed to investigate the relationship between outcome variables, baseline characteristics and dosage. Multiple regression models were used to determine if baseline characteristics could predict change in MRS score from baseline to LOCF endpoint (for each country separately). No baseline characteristic was a significant predictor using a criterion of $p<0.01$. However, in the model for India, psychotic at baseline was a borderline significant predictor ($p=0.03$); in the US model, PANSS Negative score was a borderline predictor ($p=0.04$), but in the model for Russia no variable was significant at $p<0.05$.

With respect to therapeutic effects measured by country, within-group MRS change at week 3 was significantly higher in India (ziprasidone −12.35, haloperidol −19.59, placebo −5.81) than in the USA (ziprasidone −11.10, haloperidol −12.64, placebo −8.37) and Russia (ziprasidone −7.97, haloperidol −13.80, placebo −3.72) (Fig. 1). For ziprasidone compared to placebo, the effect sizes for MRS change from baseline to endpoint were 0.52 in India, 0.39 in Russia and 0.26 in the USA and 0.40 regardless of country. The effect size for haloperidol *vs.* placebo was the largest for Russia (1.21), then India (1.12), then USA (0.44) and overall (0.92). The effect size for MRS

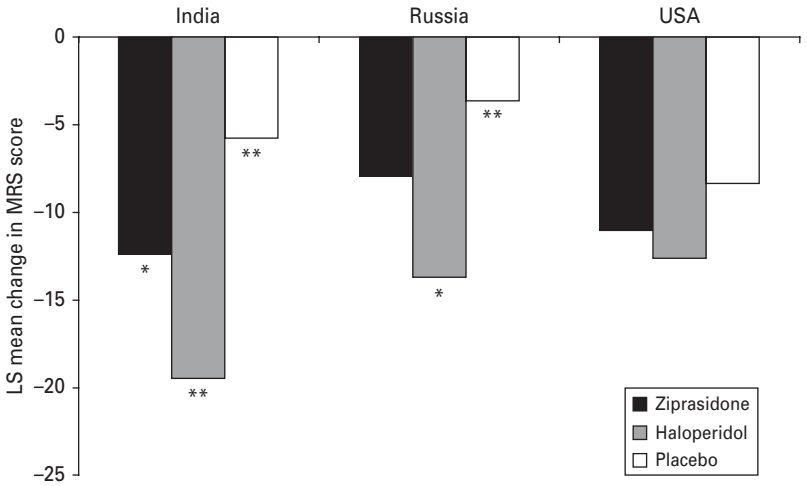

Fig. 1. Least squares (LS) mean change in Mania Rating Scale (MRS) score at week 3 following treatment with ziprasidone or haloperidol by country. * $p<0.005$, ** $p<0.0001$.

Table 3. Relationships between therapeutic effect and dose at week 3

	India	Russia	USA
Relationship between MRS change and mean modal dose, R^2 (p value)			
Ziprasidone	0.0001 (0.93)	0.001 (0.82)	0.013 (0.44)
Haloperidol	0.001 (0.78)	0.003 (0.70)	0.001 (0.81)
Relationship between MRS change and last dose of active drug, R^2 (p value)			
Ziprasidone	0.0001 (0.94)	0.003 (0.73)	0.028 (0.35)
Haloperidol	0.04 (0.15)	0.012 (0.44)	0.002 (0.43)

MRS, Mania Rating Scale.

change from baseline to endpoint for haloperidol *vs.* ziprasidone was 0.60 for India, 0.83 for Russia, 0.18 for the USA and 0.52 regardless of country.

Finally, there was virtually no relationship between mean modal dose or last dose of active drug and change in MRS score from baseline to LOCF end point (Table 3), indicating that the observed higher therapeutic effect in India was not related to the higher administered doses.

Sub-analysis restricted to severely ill at baseline

As the severity of illness differed between subjects in the three countries, we investigated if this difference impacted the treatment outcome. In subjects who were severely ill (MRS $\geqslant$30 at baseline), there was a significant difference in change from baseline to LOCF endpoint in MRS score when controlling for baseline MRS score, treatment, country and treatment × country. There was a significant difference between ziprasidone and placebo in both India ($p=0.001$) and the USA ($p=0.007$), but there was no difference in Russia ($p=0.15$). In each case, the mean change from baseline to LOCF endpoint MRS score was larger in the ziprasidone group compared to the placebo group. However, the differences between ziprasidone and haloperidol were inconsistent between countries. In India and Russia, the haloperidol group showed a larger improvement than the ziprasidone group ($p=0.001$ and $p=0.007$, respectively); whereas in the USA, the ziprasidone group showed a non-significantly larger change from baseline compared to the haloperidol group ($p=0.14$). Results were similarly inconsistent when comparing the mean change between countries (regardless of treatment group). The mean change was significantly greater in both India and the USA than in Russia ($p=0.004$ and $p=0.03$, respectively) but the mean change was not significantly different when comparing the USA to India ($p=0.65$). Finally, the mean change in MRS scores across the treatment groups, controlling for country, baseline MRS score and the country × treat-treatment interaction, showed a significant difference between ziprasidone and placebo ($p<0.001$) favouring ziprasidone and between haloperidol and placebo ($p<0.001$), but no overall difference between ziprasidone and haloperidol ($p=0.40$).

Using the more stringent criteria for treatment response ($\geqslant$50% decrease in MRS score from baseline to LOCF endpoint), there was no significant difference

between countries in the proportion of severely ill subjects (baseline MRS scores $\geqslant 30$, $p = 0.07$). However, US subjects responded with the lowest frequency overall (8/116, 6.9%) compared to Russia (23/139, 16.5%) and India (60/179, 33.5%). Notably, almost five times as many severely ill subjects in India responded, compared to those in the USA; more than twice as many as Russia. When comparing treatments for the above-mentioned category of subjects, there was a significant difference between groups ($p = 0.002$). In the ziprasidone group, 17.0% (30/176) of the subjects responded, compared to 32.9% (56/170) of the haloperidol group and 5.7% (5/88) of the placebo group. There was no interaction between country and treatment ($p = 0.55$).

Placebo response across countries

Several reports of differences in outcomes between countries and in drug *vs.* placebo effects across sites and over time, underscore the importance of this issue in relation to the precision of clinical trials (Sysko & Walsh, 2007; Walsh *et al*. 2002; Watsky *et al*. 2009). Placebo response has been observed increasingly in trials of major depressive disorders (Walsh & Sysko, 2005; Walsh *et al*. 2002) and bipolar mania (Sysko & Walsh, 2007). Intriguingly, a recent report identified a positive correlation between year of publication and placebo response rate, with placebo response increasing over time (Sysko & Walsh, 2007). While the cause remains unclear, this observation could be attributed to a number of factors, such as trial design and duration, severity of illness and additional rescue medication (Sysko & Walsh, 2007). Publication bias, and particularly the prioritization of positive trials for publication submission, has been reported to be the most likely reason for these findings (Vieta & Cruz, 2008).

In this study, a higher placebo response was observed in the USA (Fig. 1). However, these findings differed from other reports where placebo response was higher in subjects at non-US sites (Watsky *et al*. 2009). The conflicting views in the literature regarding the potential impact of geography on precision in international clinical trials (Watsky *et al*. 2009) prompts a more in-depth investigation into the differences in patient populations and outcomes.

Relationship between outcome variables

In order to determine if severity of illness was consistently measured, MRS score at visit was correlated with the investigator's rating of severity (CGI-S) using a multiple regression model both at the beginning of the study and post-baseline. The model also included main effects for treatment, country and treatment × country with baseline MRS score as a covariate. Overall, although there was a significant correlation ($p < 0.001$) for all comparisons, there was no clear trend with respect to country or treatment. In the post-baseline comparisons of MRS and CGI-S scores, while within the USA and India the correlations were close between treatment groups, in Russia the correlations were highest in the placebo group and lowest in the haloperidol group (0.82 *vs.* 0.41, respectively).

When this analysis was extended to compare the MRS score to another diagnostic scale (CGI-S), there was a significant relationship between the two scores ($p < 0.0001$). There was a significant effect for treatment ($p < 0.0001$) and country ($p = 0.007$), and no interaction between country and treatment ($p = 0.96$). The relationship was significantly different between haloperidol and both placebo and ziprasidone ($p < 0.0001$ for each), and ziprasidone and placebo ($p = 0.053$). Furthermore, there was not a significant difference between India and Russia ($p = 0.10$), but there was a significant difference between India and the USA ($p = 0.002$), and Russia and the USA ($p = 0.05$).

In order to determine if improvement was consistently measured by the investigator at each site, MRS change from baseline to visit was correlated with the investigator's rating of improvement (CGI-I) using a multiple regression model. The model also included main effects for treatment, country and treatment × country with baseline MRS as a covariate. Overall, there was a significant relationship between MRS change from baseline and CGI-I ($p < 0.0001$). While there was no interaction between country and treatment ($p = 0.13$), there was a significant effect for treatment ($p < 0.0001$) and country ($p < 0.0001$). The relationship was significantly different between haloperidol and both placebo and ziprasidone ($p < 0.0001$ for each), but ziprasidone and placebo were not significantly different ($p = 0.73$). Furthermore, there was no significant difference between India and Russia ($p = 0.09$), but there was a significant difference between India and the USA ($p < 0.0001$), and Russia and the USA ($p < 0.0001$).

In all the above analyses, we observed an inconsistency between the global rating (CGI) and the MRS based upon which country the rating was conducted in and the treatment that the subject received.

Safety

The next step was to ascertain the differences in safety measures by country, as measured by the reported

Table 4. Comparison of adverse events (AEs) per subject by treatment group and country

	India	Russia	USA
N			
Ziprasidone	74	56	48
Haloperidol	69	55	47
Placebo	36	28	24
LS mean number of AEs (S.E.)			
Ziprasidone	1.51 (0.21)	0.61 (0.24)	3.27 (0.26)
Haloperidol	2.12 (0.22)	1.31 (0.24)	2.66 (0.26)
Placebo	0.78 (0.30)	0.46 (0.34)	1.38 (0.36)
LS mean number of AEs difference			
Ziprasidone			
Haloperidol	−0.60[a]	−0.70[a]	0.61
Placebo	0.74[a]	0.14	1.90[b]

LS, Least squares.
[a] $p<0.05$.
[b] $p<0.001$.

Table 5. Total discontinuations and discontinuations due to adverse events (AEs) by country and treatment

	India	Russia	USA
N			
Ziprasidone	74	56	48
Haloperidol	69	55	47
Placebo	36	28	24
Total discontinuations, *n* (%)			
Ziprasidone	26 (35.1)	11 (19.6)	6 (12.5)
Haloperidol	11 (15.9)	6 (10.9)	11 (23.4)
Placebo	21 (58.3)	7 (25.0)	5 (20.8)
p value[a]	<0.0001	0.2318	0.1876
Discontinuations due to AEs related to study drug, *n* (%)			
Ziprasidone	3 (4.1)	1 (1.8)	3 (6.3)
Haloperidol	3 (4.3)	2 (3.6)	8 (17.0)
Placebo	1 (2.8)	0	0
p value[a]	0.9053	0.6151	0.0237
Discontinuations due to lack of efficacy, *n* (%)			
Ziprasidone	23 (31.1)	10 (17.9)	3 (6.3)
Haloperidol	8 (11.6)	4 (7.3)	3 (6.4)
Placebo	20 (55.6)	7 (25.0)	5 (20.8)
p value[a]	<0.0001	0.0916	0.1783

[a] The *p* values represent a test between all three treatment groups.

AEs. Table 4 shows the pairwise comparisons between the ziprasidone group and the other treatment groups within each country. In all three countries, the placebo group had the fewest AEs, although the pattern for ziprasidone and haloperidol varied by country (Table 4). In Indian subjects, compared to the ziprasidone group, both placebo and haloperidol groups had a significantly different number of AEs ($p=0.04$ for both comparisons). Nearly twice as many AEs were reported in the ziprasidone-treated subjects as the placebo-treated subjects, while nearly 1.5 times as many AEs were reported in the haloperidol group as in the ziprasidone group. In Russia, the pattern was slightly different with similar numbers of AEs in the ziprasidone and placebo groups. The ziprasidone group had half as many AEs as the haloperidol group ($p=0.04$). In contrast to India and Russia, USA subjects treated with ziprasidone reported almost 2.5 times as many AEs compared to those on placebo ($p<0.001$), although there was no difference compared to the haloperidol group ($p=0.10$, n.s.).

Discontinuation

Discontinuation rates for those on an active drug were highest in India (ziprasidone 35.1%, haloperidol 15.9%), followed by the USA (ziprasidone 12.5%, haloperidol 23.4%) and lowest in Russia (ziprasidone 19.6%, haloperidol 10.9%) (Table 5). More ziprasidone- and placebo-treated subjects discontinued due to lack of efficacy in all three countries ($p=0.0098$). The haloperidol group had similar incidences of discontinuation due to lack of efficacy across all countries. Comparing rates of discontinuation due to treatment-related AEs, only the haloperidol group showed a significant difference between countries ($p=0.0007$).

When comparing the treatment groups within a country, there was a significant difference between the treatments for discontinuation due to lack of efficacy in India ($p<0.0001$). In Russia, there was no difference among the treatments and in the USA there was a difference in the rate of discontinuation due to treatment-related AEs ($p=0.02$)

The median time until discontinuation in ziprasidone-treated subjects was longer than the placebo group in India (it was not estimable in the USA and Russia because <50% of the patients discontinued). However, the median time until discontinuation from haloperidol was only 14 d in the USA, compared to 24 d in Russia (Table 6). Finally, in the ziprasidone and placebo groups, there was no significant difference in time to discontinuation across countries (Fig. 2*a*, *c*); however it was significantly different across countries in the haloperidol groups ($p<0.001$) (Fig. 2*b*).

Table 6. Time to discontinuation by country and treatment

	India	Russia	USA	*p* value[a]
Ziprasidone	22 (22–23)	n.e. (24 – n.e.)	n.e. (n.e. – n.e.)	0.024
Haloperidol	n.e. (n.e. – n.e.)	24 (24 – n.e.)	14 (9 – n.e.)	< 0.0001
Placebo	16.5 (11–20)	n.e. (n.e. – n.e.)	n.e. (14 – n.e.)	0.0094

Values given are median in days (95% CI).
CI, confidence interval around median; n.e., not estimable (e.g. not enough patients discontinued to estimate the parameter of interest, median or upper limit of the median).
[a] *p* value is derived from the comparison of time to discontinuation within treatment groups across countries.

Discussion

As more clinical trials are conducted at multiple sites globally, it is necessary to identify potential sources of variability in patient groups across countries that may impact trial outcomes. In these *post-hoc* analyses, we systematically compared baseline characteristics of patients with bipolar mania, the treatment outcome variables and safety data between countries (USA, Russia, India) and treatment groups (ziprasidone, haloperidol, placebo). Some of our analyses revealed unexpected differences between countries and here we discuss the source and likely impact of these differences. Awareness of these distinctions can instruct a smarter design of global psychiatric trials in the future.

Geographical differences in baseline characteristics

Nearly all baseline characteristics differed in the patient population, with the greatest difference between US and Indian subjects, and the Russian group closely aligned with the US pool. Access to quality, early psychiatric diagnosis and intervention in the USA may be why US subjects were younger and had tried more psychotropic medications for treatment options. Indeed the high percentage of Indian subjects with psychotic manic episodes suggests that patients may be prioritized for clinical intervention based on disease severity by their physicians. This was also the case in a risperidone placebo-controlled study (Khanna *et al*. 2005). It is possible that the higher number of discontinuations in Indian subjects across all treatment groups may be because the subjects were more severely ill. Identifying the reasons for differences in discontinuation rates is beyond the scope of the present analyses. One explanation for the higher discontinuation rates for haloperidol in the USA could be that these subjects had lower tolerance to the drug.

Dosage disparities across countries

Our analyses confirmed the disparity in dosing between countries. The drug dosage did not appear to have been adjusted for weight, so subjects in India who weighed the least received the highest doses. Furthermore, the doses of haloperidol and ziprasidone were not equivalent, since a dose of 2 mg/d haloperidol is equivalent to 60 mg/d ziprasidone (Vieta & Cruz, 2008; Woods, 2003). This led to subjects in India and Russia receiving doses of haloperidol higher than that associated with optimal efficiency and tolerability, and all subjects receiving doses of ziprasidone lower than that associated with maximum efficacy (Vieta & Cruz, 2008; Woods, 2003). In the Khanna *et al*. (2005) study conducted in India, risperidone was also dosed much higher than usual. This may be because in some cultures, similar to India, efficacy is prioritized over safety and in these countries trial patients tend to be more severely affected, whereas in the USA safety is prioritized over efficacy.

The lack of consistency in dosage between countries appears to be driving a difference in outcomes. However, the differential response may also be attributed to the severity of mania at baseline, where a greater response is observed with increasing baseline severity of mania. There is also growing acceptance in the medical community that age, sex and BMI can all impact basal metabolic systems, such as renal clearance, and thereby impact clinical pharmacodynamics (Han *et al*. 2007; Woods, 2003). These factors may contribute to the difference in response profiles amongst subjects in India compared to those in Russia and the USA.

Furthermore, even though subjects in India received higher mean doses of ziprasidone, they tended to stay on the drug longer than subjects in Russia or the USA. Overall it is not clear whether the high dosing was

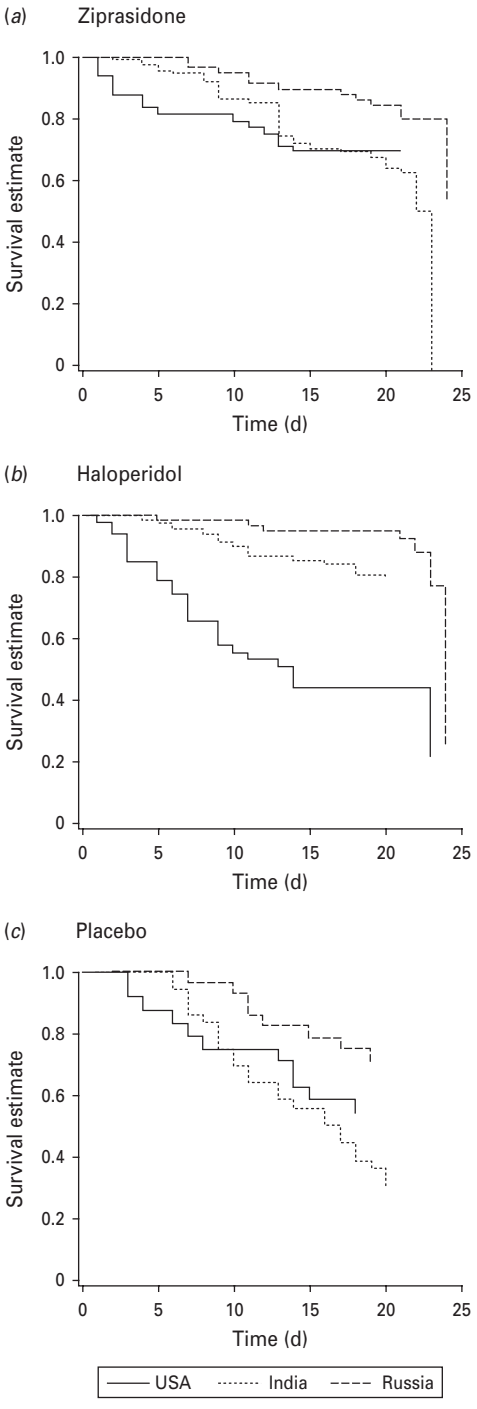

Fig. 2. Kaplan–Meier curves for discontinuations over time by country and treatment.

necessary because the disease severity was higher, because the therapeutic threshold in the Indian subjects was higher or whether potentially unequal investigator training led to different prescribing practices. In any circumstance, further investigation is warranted to understand these clinical and perhaps cultural factors.

Geographical differences in placebo response

Subjects in different countries may also respond differently to participating in a clinical trial, as evidenced by the disproportionately high placebo response observed in the USA. Our results further substantiate recent reports of differences in outcomes between countries and variability in drug–placebo differences across sites and over time (Sysko & Walsh, 2007; Vieta & Cruz, 2008; Walsh & Sysko, 2005; Walsh *et al*. 2002; Watsky *et al*. 2009).

Conversely, our results differed from the observations from a phase 2a clinical trial in schizophrenia that used a response criterion of $\geqslant 30\%$ change in PANSS from baseline and found a significant active control/placebo difference ($p < 0.1$) in the USA, but not outside the USA (Watsky *et al*. 2009). A higher response criterion of $\geqslant 50\%$ on the MRS or Young MRS (YMRS) or a score of 1 or 2 on the CGI-I scale was employed in the review of placebo response in bipolar mania (Sysko & Walsh, 2007). This may account for the different findings in the schizophrenia study that defined response as a $\geqslant 30\%$ change in PANSS, although it could be related to the nature of the illness as well (Watsky *et al*. 2009). These conflicting views in the literature, along with the potential impact of geography on precision in clinical trials remain debatable. Lower severity, a well-known factor that increases placebo response (Vieta & Carne, 2005), might have played a role in the higher placebo response of US subjects. Furthermore, in the USA, the treatment of psychiatric illness is well established and the role of clinical trials appears to be well comprehended across potential subjects. This awareness might actually favour the inclusion of subjects with lower severity and greater insight into clinical trials, indirectly fostering placebo response. Such cultural considerations could have an impact on the recruitment and outcomes of clinical trials in Russia and India (Platonov, 2003; Raja *et al*. 2010; Shah *et al*. 2010). Hence, to understand the difference in placebo response, further research is necessary to examine the relative contribution of the cultural differences across subjects and inherent methodological factors in the trial design (Sysko & Walsh, 2007; Vieta & Cruz, 2008).

Implications for global psychiatric clinical trials

With the continuing move towards the globalization of clinical trials, the difference between subject groups that we report here may have considerable

implications for the design of clinical trials. When designing multinational trials it may be beneficial to provide a smarter protocol that considers guidelines that directly tackle these differences at the outset, rather than discovering them after the trial completion.

One option would be to create study protocols that assess differences in patient baseline characteristics, such as severity of disease, duration of illness, weight/BMI and treatment history prior to the ending of the trial. However, this approach is problematic as most psychiatric trials are relatively short and do not allow for quick correction of the protocol to compensate for differences seen in patient populations across countries. Alternatively, a consensus approach prior to the start of the trial to monitor patient recruitment is a possibility where investigators from all sites make a judgement on whether or not to include patients in the trial. This latter strategy could prevent potential drift of a single trial site and maintain uniformity in the trial.

Given our findings of differences in dosing approaches between countries, better training and a clearly defined dosing algorithm during the trial can ensure that patients receive comparable doses. Furthermore, this will also take into consideration any apprehensions that investigators have regarding newer drugs. Last, documenting and understanding the factors influencing the diagnosis, severity of the disease and the treatment approach in different countries should be an important consideration in global clinical psychiatric trials.

Conclusions

Within each treatment group we found significantly different baseline characteristics, treatment response (including placebo response) and discontinuation by country, indicating a need for further research to determine whether this is the result of cultural differences, baseline disease severity or differing healthcare practices among these countries. These differences need to be fully examined and explained in order to have international and intercontinental clinical trials designed to ensure globally applicable results. Furthermore, in the future it may be necessary to analyse clinical data on a country-by-country basis to account for any geographical differences.

Acknowledgements

The original study was conducted in compliance with the ethical principles originating from the Declaration of Helsinki. The authors are responsible for the scientific content presented in this article. This study was sponsored by Pfizer Inc. Editorial support was provided by Hajira Koeller, Ph.D., of PAREXEL and was funded by Pfizer Inc.

Statement of Interest

Dr Vieta is a consultant for AstraZeneca, Bial, Bristol–Myers Squibb, Eli Lilly and Company, Forest, Geodon Richter, Janssen Pharmaceutica, Lundbeck, Merck Sharp and Dohme, Novartis, Pierre Fabre, Pfizer Inc., Otsuka, Sanofi-Aventis, Servier, Takeda, and UCB Pharmaceuticals. Dr Vieta has received research grants from AstraZeneca, GlaxoSmithKline, Bial, Eli Lilly and Company, Janssen Pharmaceutica, Novartis, Pfizer Inc., Sanofi-Aventis and Servier, and he has participated at advisory boards or has been a speaker for AstraZeneca, Bristol–Myers Squibb, Eli Lilly and Company, Forest, Geodon Richter, Janssen Pharmaceutica, Lundbeck, Merck Sharp and Dohme, Novartis, Pfizer Inc., and Takeda.

Dr Pappadopulos, Dr Mandel and Dr Lombardo are employees of Pfizer Inc.

References

El Mallakh RS, Vieta E, Rollin L, Marcus R, *et al*. (2010). A comparison of two fixed doses of aripiprazole with placebo in acutely relapsed, hospitalized patients with bipolar disorder I (manic or mixed) in subpopulations (CN138-007). *European Neuropsychopharmacology: The Journal of the European College of Neuropsychopharmacology* **20**, 776–783.

Han PY, Duffull SB, Kirkpatrick CM, Green B (2007). Dosing in obesity: a simple solution to a big problem. *Clinical Pharmacology and Therapeutics* **82**, 505–508.

Khanna S, Vieta E, Lyons B, Grossman F, *et al*. (2005). Risperidone in the treatment of acute mania: double-blind, placebo-controlled study. *British Journal of Psychiatry* **187**, 229–234.

Platonov P (2003). Clinical trials in Russia and Eastern Europe: recruitment and quality. *Interational Journal of Clinical Pharmacology and Therapeutics* **41**, 277–280.

Raja S, Wood SK, de M V, Mannarath SC (2010). Mapping mental health finances in Ghana, Uganda, Sri Lanka, India and Lao PDR. *International Journal of Mental Health System* **4**, 11.

Shah JY, Phadtare A, Rajgor D, Vaghasia M, *et al*. (2010). What leads Indians to participate in clinical trials? A meta-analysis of qualitative studies. *PLoS One* **5**, e10730.

Sysko R, Walsh BT (2007). A systematic review of placebo response in studies of bipolar mania. *Journal of Clinical Psychiatry* **68**, 1213–1217.

Vieta E, Carne X (2005). The use of placebo in clinical trials on bipolar disorder: a new approach for an old debate. *Psychotherapy and Psychosomatics* **74**, 10–16.

Vieta E, Cruz N (2008). Increasing rates of placebo response over time in mania studies. *Journal of Clinical Psychiatry* **69**, 681–682.

Vieta E, Ramey T, Keller D, English P, *et al*. (2010). Ziprasidone in the treatment of acute mania: a 12-week, placebo-controlled, haloperidol-referenced study. *Journal of Psychopharmacology* **24**, 547–558.

Walsh BT, Seidman SN, Sysko R, Gould M (2002). Placebo response in studies of major depression: variable, substantial, and growing. *Journal of the American Medical Association* **287**, 1840–1847.

Walsh BT, Sysko R (2005). Placebo control groups in trials of major depressive disorder among older patients. *Journal of Clinical Psychopharmacology* **25** (4 Suppl. 1), S29–S33.

Watsky EJ, Schwartz JH, Kremer CME (2009). The impact of geography on precision in clinical trials: experience with placebo and an active control in a phase 2a study in patients with acute symptoms of schizophrenia. Presented at the International Society for CNS Clinical Trials and Methodology (ISCTM) 2009 Autumn Conference, 5–6 October 2009, San Diego, CA.

Woods SW (2003). Chlorpromazine equivalent doses for the newer atypical antipsychotics. *Journal of Clinical Psychiatry* **64**, 663–667.