U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

Statistical Review and Evaluation

CLINICAL STUDIES

NDA/BLA Serial Number:	21-999/ S-024
Drug Name:	Paliperidone (INVEGA)
Indication(s):	Schizophrenia
Applicant:	Johnson & Johnson
Date(s):	October 8, 2010
Review Priority:	Priority (Pediatric Exclusivity)
Biometrics Division:	Division of Biometrics I
Statistical Reviewer:	George Kordzakhia
Concurring Reviewers:	Peiling Yang, H.M. James Hung
Medical Division:	Division of Psychiatry Products
Clinical Team:	Jenn Sellers, Jing Zhang
Project Manager:	Ann Sohn, Terry Harrison

Keywords: clinical studies, NDA review

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. EXECUTIVE SUMMARY

The sponsor submitted results of Study R076477-PSZ-3001 in support of efficacy of paliperidone ER (Low, Medium, and High doses) versus placebo for the treatment of schizophrenia (b) (4) in adolescents aged 12 to 17.

In the primary analysis of PANSS Total score, patients in the paliperidone ER medium dose treatment group (3mg daily for patients weighing 29 to less than 51 kg, 6mg daily for patients weighing at least 51 kg) showed statistically significant improvement over patients in the placebo treatment group. Based on the ANCOVA model with treatment and country as factors and baseline PANSS total score as a covariate, using a closed testing procedure with Dunnett's test to control for multiplicity, the LS mean change in the paliperidone ER Medium treatment group was statistically significantly superior to that in the placebo group (p=0.006). The paliperidone ER High and Low doses did not demonstrate statistical superiority over placebo.

This reviewer noted that the observed improvements from baseline in PANSS Total score and treatment differences from placebo were highly inconsistent between the two weight subgroups (see Table 11). For the <51 kg weight subgroup, the observed dose response for paliperidone had an umbrella shape with a large improvement from baseline observed in the Medium dose group, while for the other weight subgroup the observed improvements in the Medium and the High dose groups were large and were close to each other. There is apparently a large placebo response observed in the <51 kg subgroup; such a high placebo response is unusual. It is also not clear why the dose response relationships are very different in the two weight subgroups. Although these inconsistencies could potentially or partially be due to small sample sizes in the <51 kg subgroup, there could be other unknown factors for the inconsistencies. If the drug is effective to adolescents weighing ≥51 kg, it'd be surprising that the drug is not effective to adolescents weighing <51 kg, but the problem is at what dose level. If the dose response relationships are indeed very different in these two subgroups, perhaps the trial shouldn't have been designed as it was and efficacy shouldn't have been analyzed by combining the two weight subgroups. Because of very small sample sizes in the <51 kg subgroup, the observed improvements from baseline in PANSS Total score in the overall population were driven by the ≥51 kg subgroup. Given the substantial inconsistencies in the observed improvements from baseline between these two weight subgroups, the results of subgroup analysis do not support that the drug is effective in the <51 kg subgroup, and another trial would be needed to confirm the dose response relationship in the overall adolescent population.

The sponsor included the following results of exploratory analysis by actual dose in Section 14.1 of the proposed labeling: "*Efficacy was evaluated using PANSS. This study demonstrated the efficacy of INVEGA in adolescent subjects with schizophrenia* (b) (4) " The comparisons of actual doses with placebo were nonrandomized comparisons. The hypotheses associated with comparisons with actual dose were not prespecified as the primary or key secondary objectives, and these post hoc comparisons were not adjusted for multiplicity (in fact, multiplicity adjustment can not be properly made post hoc after data was unblinded and analyzed). The trial was not designed to investigate the effects of actual doses: patients were randomized to 4 treatment arms (Placebo, Low, Medium and High dose arms) within each weight subgroup (<51 kg and ≥ 51 kg). As a result, the paliperidone 3 mg actual-dose group included patients weighing <51 kg only, and the paliperidone 12 mg actual-dose group included patients with weight ≥51 kg only. Hence the actual-dose groups and the placebo arm may not be comparable with respect to baseline covariates (particularly those unobserved). Above all, the results of the weight subgroup analysis were not consistent between the two weight subgroups (see Table 11). For all these reasons, the results of such a post hoc "actual-dose" analysis cannot be properly interpreted and hence cannot be included in the labeling.

2. INTRODUCTION

2.1 Overview

This supplemental New Drug Application (sNDA) includes data from 3 clinical studies of paliperidone in adolescents with schizophrenia: the Phase 1 PK study, PALIOROS-PSZ-1001; the Phase 3 double-blind, placebo-controlled efficacy and safety study, R076477-PSZ-3001; and the Phase 3 long-term, open-label safety and efficacy study, R076477-PSZ-3002.

Results from Study R076477-PSZ-3001 form the basis for the efficacy claims for paliperidone ER tablets in the treatment of schizophrenia in adolescents aged 12 to 17 years. This randomized, double-blind, parallel-group, placebo-controlled, 6-week study used a fixed dose weight-based treatment group design to explore the efficacy and tolerability of paliperidone ER over the dose range of 1.5 to 12 mg/day in adolescent subjects meeting DSM-IV criteria for schizophrenia, with diagnosis confirmation using the Kiddie Schedule for Affective Disorders and Schizophrenia-Present and Lifetime Version (K-SADS-PL)

Eligible subjects were randomly assigned to 1 of 4 treatment groups (placebo, paliperidone ER Low, paliperidone ER Medium, paliperidone ER High) corresponding to nonoverlapping milligram per kilogram groups. Subjects weighing 29 to less than 51 kg at the baseline visit were randomly assigned to receive placebo or 1.5, 3, or 6 mg of paliperidone ER daily, corresponding to the placebo, paliperidone ER Low, paliperidone ER Medium, or paliperidone ER High groups, respectively. Subjects weighing at least 51 kg at the baseline visit were randomly assigned to receive placebo or 1.5, 6, or 12 mg of paliperidone ER daily, corresponding to the placebo, paliperidone ER Low, paliperidone ER Medium, or paliperidone ER High groups, respectively. Overall, two hundred and one patients were randomly assigned to 1 of 4 treatment groups. Fifty one patients were assigned to Placebo, 54 to Low dose, 48 to Medium, and 48 to High dose groups respectively.

This study was conducted at 12 study sites in Russia, 9 sites in the United States, 7 sites in India, 6 sites in the Ukraine, and 1 site in Romania. Overall, 41% of the subjects were from Russia, 23% were from India, 17% were from Ukraine, 15% were from the U.S., and 5% were from Romania.

2.2 Data Sources

The clinical study report and data sets for Study R076477-PSZ-3001are submitted electronically. The network path for the submission that includes the clinical study report is:
\\Cdsesub1\evsprod\NDA021999\0120.
Additional submissions pertaining to statistics are located at
\\Cdsesub1\evsprod\NDA021999\0138, and \\Cdsesub1\evsprod\NDA021999\0142.
Primary analysis data set kpanss.xpt is located at:
\\Cdsesub1\evsprod\NDA021999\0120\m5\datasets\ped-r076477-psz-3001\analysis.

3. STATISTICAL EVALUATION

3.1 Evaluation of Efficacy

The sponsor conducted Study R076477-PSZ-3001 to evaluate the efficacy, safety, and tolerability of 3 weight-based, fixed-dose groups of paliperidone ER as compared with placebo in adolescent subjects 12 to 17 years of age, inclusive, with schizophrenia.

Study Design and Endpoints

Study R076477-PSZ-3001 was a randomized, double-blind, parallel-group, placebo-controlled, multicenter study with 3 weight-based, fixed dose groups designed to determine the efficacy and safety of paliperidone ER in adolescents 12 to 17 years of age, inclusive, who had a Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV) diagnosis of schizophrenia.

The study consisted of 3 phases: a screening phase (with a possible overlapping washout period), a 6-week double-blind treatment phase with an end-of-study (EOS) or early-withdrawal visit, and a 1-week follow-up visit for subjects who did not enter a separate, optional open-label safety study. *The study allowed early rescue*. Subjects who did not have any response to treatment or whose symptoms worsened could drop out after at least 21 days of the double-blind phase and were eligible to participate in open label safety study R076477-PSZ-3002.

Subjects with schizophrenia who were in a state of acute exacerbation and not doing well on their current antipsychotics, and who met all entry criteria at screening, had their current disallowed psychotropic medications tapered and discontinued if necessary during the screening phase. Eligible subjects were then randomly assigned to 1 of 4 treatment groups (placebo, paliperidone ER Low, paliperidone ER Medium, paliperidone ER High) corresponding to nonoverlapping milligram per kilogram groups. Subjects weighing 29 to less than 51 kg at the baseline visit were randomly assigned to receive placebo or 1.5, 3, or 6 mg of paliperidone ER daily, corresponding to the placebo, paliperidone ER Low, paliperidone ER Medium, or paliperidone ER High groups, respectively. Subjects weighing at least 51 kg at the baseline visit were randomly assigned to receive placebo or 1.5, 6, or 12 mg of paliperidone ER daily, corresponding to the placebo, paliperidone ER Low, paliperidone ER Medium, or paliperidone ER High groups, respectively (see Table 1). Central randomization was implemented in conducting this study. The randomization was balanced by using permuted blocks of treatments and was stratified by study center.

Table 1. Study Treatment Arms and Daily Doses

Weight Group	Placebo	Paliperidone Low	Paliperidone Medium	Paliperidone High
<51 kg	Placebo	1.5 mg	3 mg	6 mg
≥ 51 kg	Placebo	1.5 mg	6 mg	12 mg

Source: Reviewer's Summary

The primary efficacy measure for this study was the change in the PANSS total score (sum of the scores of all 30 PANSS items) from baseline to the last post-randomization assessment of the study (excluding the follow-up visit). A closed testing procedure using Dunnett's test was used to adjust for multiple comparisons in testing the 3 paliperidone ER treatment groups against placebo for the primary efficacy variable.

Patient Disposition, Demographic and Baseline Characteristics

This study was conducted at 12 study sites in Russia, 9 sites in the United States, 7 sites in India, 6 sites in the Ukraine, and 1 site in Romania.

Two hundred and one subjects were randomly assigned to 1 of 4 treatment groups, as shown in Table 2. All randomized subjects received at least 1 dose of study medication and were included in the safety analysis set. All but 1 randomized subject also provided both baseline and at least 1 post-baseline efficacy assessments and were included in the ITT analysis set.

Table 2 and Table 3 present patient's completion/withdrawal information and withdrawal rates by visit. The proportion of subjects who completed the study was higher in each of the paliperidone ER treatment groups than in the placebo group. The proportion of subjects who were withdrawn due to lack of efficacy was higher in the placebo and paliperidone ER Low treatment groups than in the other 2 groups.

Table 2. Subject Disposition and Completion/Withdrawal Information

Population	Placebo	Paliperidone ER		
		Low	Medium	High
All randomized subjects	**51 (100%)**	**54 (100%)**	**48 (100%)**	**48 (100%)**
<51 kg	14 (100%)	19 (100%)	16 (100%)	13 (100%)
≥51 kg	37 (100%)	35 (100%)	32 (100%)	35 (100%)
Safety (at least 1 dose)	**51 (100%)**	**54 (100%)**	**48 (100%)**	**48 (100%)**
Intent-to-treat	**51 (100%)**	**54 (100%)**	**48 (100%)**	**47 (98%)**
Completed	**26 (51%)**	**35 (65%)**	**40 (83%)**	**37 (77%)**
Withdrawn	**25 (49%)**	**19 (35%)**	**8 (17%)**	**11 (23%)**
<51 kg	5 (36%)	5 (26%)	3 (19%)	3 (23%)
≥51 kg	20 (54%)	14 (40%)	5 (16%)	8 (23%)
Lack of Efficacy	**20 (39%)**	**14 (26%)**	**2 (4%)**	**4 (8%)**
<51 kg	3 (21%)	4 (21%)	0 (0%)	2 (15%)
≥51 kg	17 (46%)	10 (29%)	2 (6%)	2 (6%)
Withdrawal Consent	**2 (4%)**	**1 (2%)**	**2 (4%)**	**4 (8%)**
<51 kg	2 (14%)	0 (0%)	1 (6%)	1 (8%)
≥51 kg	0 (0%)	1 (3%)	1 (3%)	3 (9%)
Lost to follow-up	**3 (6%)**	**0 (0%)**	**2 (4%)**	**1 (2%)**
Adverse Event	**0 (0%)**	**1 (2%)**	**1 (2%)**	**1 (2%)**
Other	**0 (0%)**	**3 (6%)**	**1 (2%)**	**1 (2%)**

Source: Clinical Study Report Table 8 (pg. 64), Table 9 (pg. 66) and sponsor's response with letter dated February 25, 2011 available in \\Cdsesub1\evsprod\NDA021999\0138 .

Table 3. Patient Discontinuation by Week

Withdrew on or Before	Placebo, N=51	Paliperidone Low, N= 54	Paliperidone Medium, N=48	Paliperidone High, N=47
Completed	26 (51%)	35 (65%)	40 (83%)	37 (79%)
Week 2 (Days 1-14)	1 (2%)	3 (6%)	1 (2%)	3 (6%)
Week 3 (Days 1-21)	4 (8%)	5 (9%)	1 (2%)	5 (10%)
Week 4 (Days 1-28)	18 (35%)	14 (26%)	2 (4%)	8 (17%)
Week 5 (Days1-35)	18 (35%)	18 (33%)	4 (8%)	10 (21%)
Withdrew at the end of DB Phase or before completing DB phase	25 (49%)	19 (35%)	8 (17%)	11 (23%)

Source: Clinical Study Report Attachment 1.1.5 (pg. 194)

A majority of the subjects were male, white, between the ages of 15 and 17, and non-smokers. A majority had a baseline body weight of at least 51 kg. The mean age of the subjects in the ITT analysis set was approximately 15 years. The overall mean baseline PANSS total score was 91.1. The 4 treatment groups were generally comparable with respect to demographic and baseline characteristics (see Table 4). The percentage of males was higher in the paliperidone ER Medium and High treatment groups than in the other 2 groups.

Table 4. Demographic and Baseline Characteristics (ITT Population)

Population	Placebo (N=51)	Paliperidone ER		
		Low (N=54)	Medium (N=48)	High (N=47)
Age				
12-14	9 (18%)	16 (30%)	15 (31%)	13 (28%)
15-17	42 (82%)	38 (70%)	33 (69%)	34 (72%)
Sex				
Male	23 (45%)	30 (56%)	31 (65%)	33 (70%)
Female	28 (55%)	24 (44%)	17 (35%)	14 (30%)
Race				
White	35 (69%)	35 (65%)	34 (71%)	32 (68%)
Black	4 (8%)	5 (9%)	3 (6%)	5 (11%)
Asian	12 (24%)	14 (26%)	11 (23%)	10 (21%)
Baseline Body Weight Category				
<51kg	14 (27%)	19 (35%)	16 (33%)	13 (28%)
≥ 51 kg	37 (73%)	35 (65%)	32 (67%)	34 (72%)
Baseline Body Weight				
Mean (SD)	59.5 (16.47)	60.4 (16.07)	57.7 (14.63)	61.5 (16.08)
<51 kg category Mean (SD)	45.6 (6.17)	44.5 (6.14)	44.6 (4.13)	43.9 (5.98)
≥ 51 kg category Mean (SD)	64.8 (16.1)	69.1 (12.84)	64.2 (13.57)	68.2 (13.34)
Baseline PANSS Total				
Mean (SD)	90.6 (12.13)	91.6 (12.54)	90.6 (14.01)	91.5 (13.86)

Source: Clinical Study Report Table 10 (pg. 69), Table 11 (pg. 71)

Overall, 41% of the subjects were from Russia, 23% were from India, 17% were from Ukraine, 15% were from the U.S., and 5% were from Romania (see Table 5).

Table 5. Distribution of Subjects by Country (ITT Population)

Country	Placebo (N=51)	Paliperidone ER		
		Low (N=54)	Medium (N=48)	High (N=47)
Russia	21 (41%)	22 (41%)	20 (42%)	19 (40%)
India	11 (22%)	13 (24%)	11 (23%)	10 (21%)
Ukraine	8 (16%)	9 (17%)	8 (17%)	9 (19%)
United States	9 (18%)	8 (15%)	6 (13%)	6 (13%)
Romania	2 (4%)	2 (4%)	3 (6%)	3 (6%)

Source: Clinical Study Report Attachment 1.1.3 (pg. 191)

Statistical Methodologies

In Statistical Analysis Plan, the sponsor indicated that the primary efficacy endpoint is the change from baseline to end point (i.e. the final post baseline assessment during the double-blind phase) in PANSS total score. In the study protocol the primary efficacy measure was the change in the total PANSS score from baseline to the last post-randomization assessment of the study.

Treatment effects were estimated based on least-squares (LS) means using an analysis of covariance (ANCOVA) model with treatment and country as factors and baseline PANSS total score as a covariate.

A closed testing procedure using Dunnett's test was used to adjust for multiple comparisons in testing the 3 paliperidone ER treatment groups against placebo for the primary efficacy variable. For each time point for both observed case data and LOCF, descriptive statistics were produced on the PANSS total score and change from baseline. In addition, to explore the treatment effect over time, ANCOVA models (with treatment and country as factors and baseline PANSS total score as covariate) on both observed case data as well as LOCF were performed for each time point.
To assess the sensitivity of the results, a repeated measures mixed effects model was carried out on the observed data. Changes from baseline over time (observed case) were modeled using a mixed effects model with time, country, and treatment as factors and baseline PANSS total score as a covariate. In addition, a treatment-by-visit interaction term was added to evaluate the changes in treatment effect over time. An unstructured variance-covariance matrix was employed.
Central randomization was implemented in conducting this study. The randomization was balanced by using permuted blocks of treatments and was stratified by study center.

Results and Conclusions

The primary efficacy variable was the change from baseline to the final post baseline assessment during the double-blind phase in the PANSS total score. A decrease in PANSS total score indicates improvement in neuropsychiatric symptoms. Based on the ANCOVA model with treatment and country as factors and baseline PANSS total score as a covariate, using a closed testing procedure with Dunnett's test to control for multiplicity, the LS mean change in the paliperidone ER Medium treatment group was statistically significantly superior to that in the placebo group (p=0.006). The paliperidone ER High and Low doses did not demonstrate statistical superiority over placebo. The sponsor's primary efficacy results are presented in Table 6. This reviewer confirmed sponsor's results.

Table 6. PANSS Total Score - Change from Baseline to Week 6 (LOCF) for ITT Population

		Placebo (N=51)	Paliperidone Low (N=54)	Paliperidone Medium (N=48)	Paliperidone High (N=47)
Basline Mean (SD)		90.6 (12.13)	91.6 (12.54)	90.6 (14.01)	91.5 (13.86)
Change from Basline Mean (SD)		-7.9 (20.15)	-9.8 (16.31)	-17.3 (14.33)	-13.8 (15.74)
Difference Compared with Placebo (ANCOVA)	LS Mean (SE)		-2.1 (3.17)	-10.1 (3.27)	-6.6 (3.29)
	p-value (unadj)		0.508	0.002	0.047
	Adj. p-value(Dunnett)		0.508	0.006	0.086
	95% CI (unadj)		(-8.36, 4.16)	(-16.58, -3.67)	(-13.07, -0.09)

Source: Clinical Study Report Table 17 (pg. 80), but the unadjusted p-values were from this reviewer.

This reviewer also conducted exploratory ANCOVA LOCF analysis by visit (see Table 7). Numerically, all paliperidone treatment arms were better than placebo.

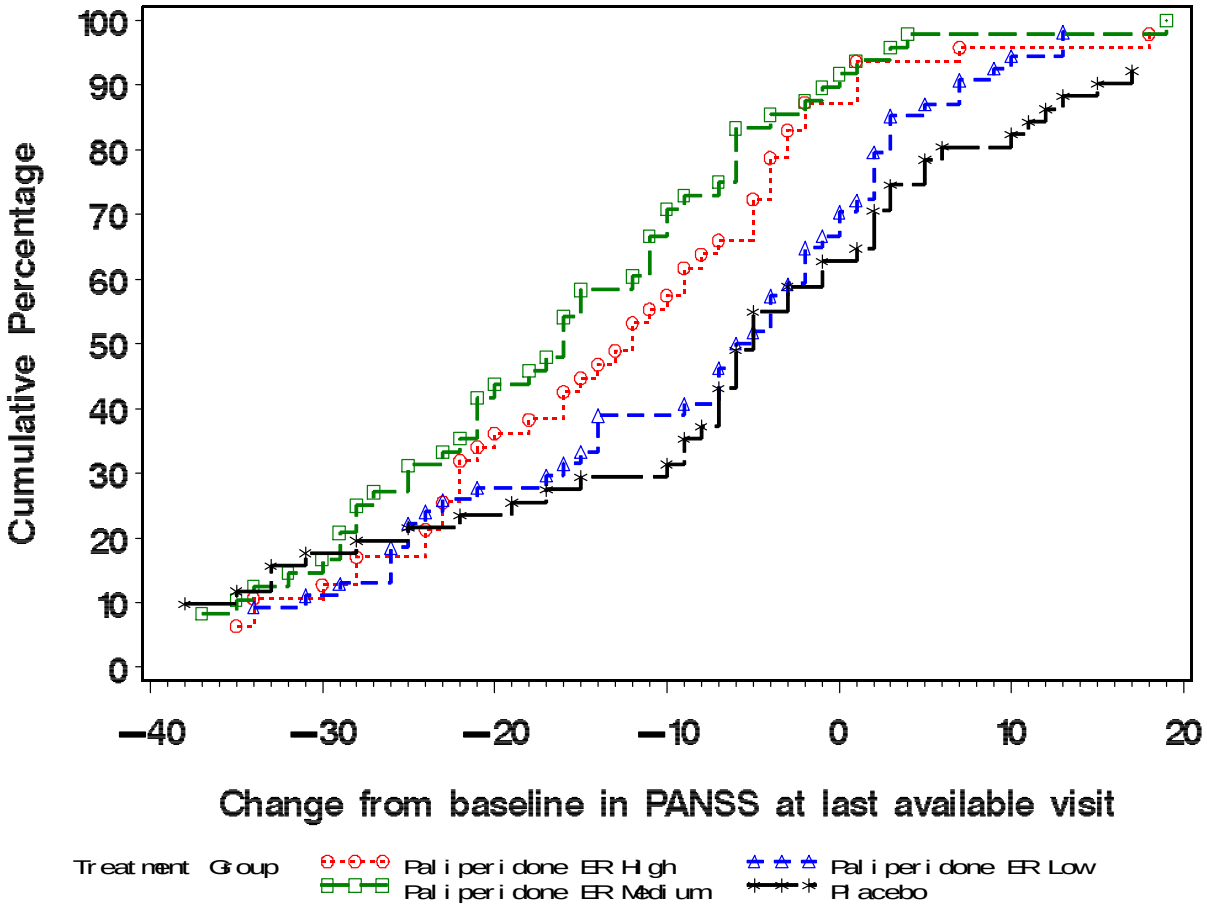
Table 7. PANSS Total Score LS Mean (SE) Change from Baseline by Visit (ANCOVA LOCF Analysis)

Week	**Placebo**	**Paliperidone Low; p-value vs Placebo**	**Paliperidone Med; p-value vs Placebo**	**Paliperidone High; p-value vs Placebo**
Week 1 LS Mean (SE)	-2.4 (1.2)	-3.6 (1.2) p=0.457	-4.1 (1.2) p=0.299	-5.7 (1.2) p=0.038
Week 2 LS Mean (SE)	-4.5 (1.5)	-5.6 (1.5) p=0.599	-8.4 (1.5) p=0.053	-9.7 (1.6) p=0.012
Week 3 LS Mean (SE)	-7.1 (2.0)	-7.4 (2.0) p=0.912	-12.6 (2.0) p=0.039	-11.8 (2.0) p=0.079
Week 4 LS Mean (SE)	-8.1 (2.3)	-9.0 (2.3) p=0.773	-15.6 (2.3) p=0.015	-13.6 (2.3) p=0.075
Week 5 LS Mean (SE)	-8.0 (2.4)	-9.3 (2.3) p=0.678	-17.7 (2.4) p=0.002	-14.3 (2.4) p=0.049
Week 6 LS Mean (SE)	-8.1 (2.5)	-10.2 (2.4) p=0.508	-18.2 (2.5) p=0.002	-14.6 (2.5) p=0.047

Source: Reviewer's Results
Remark: Reported p-values are unadjusted p-values

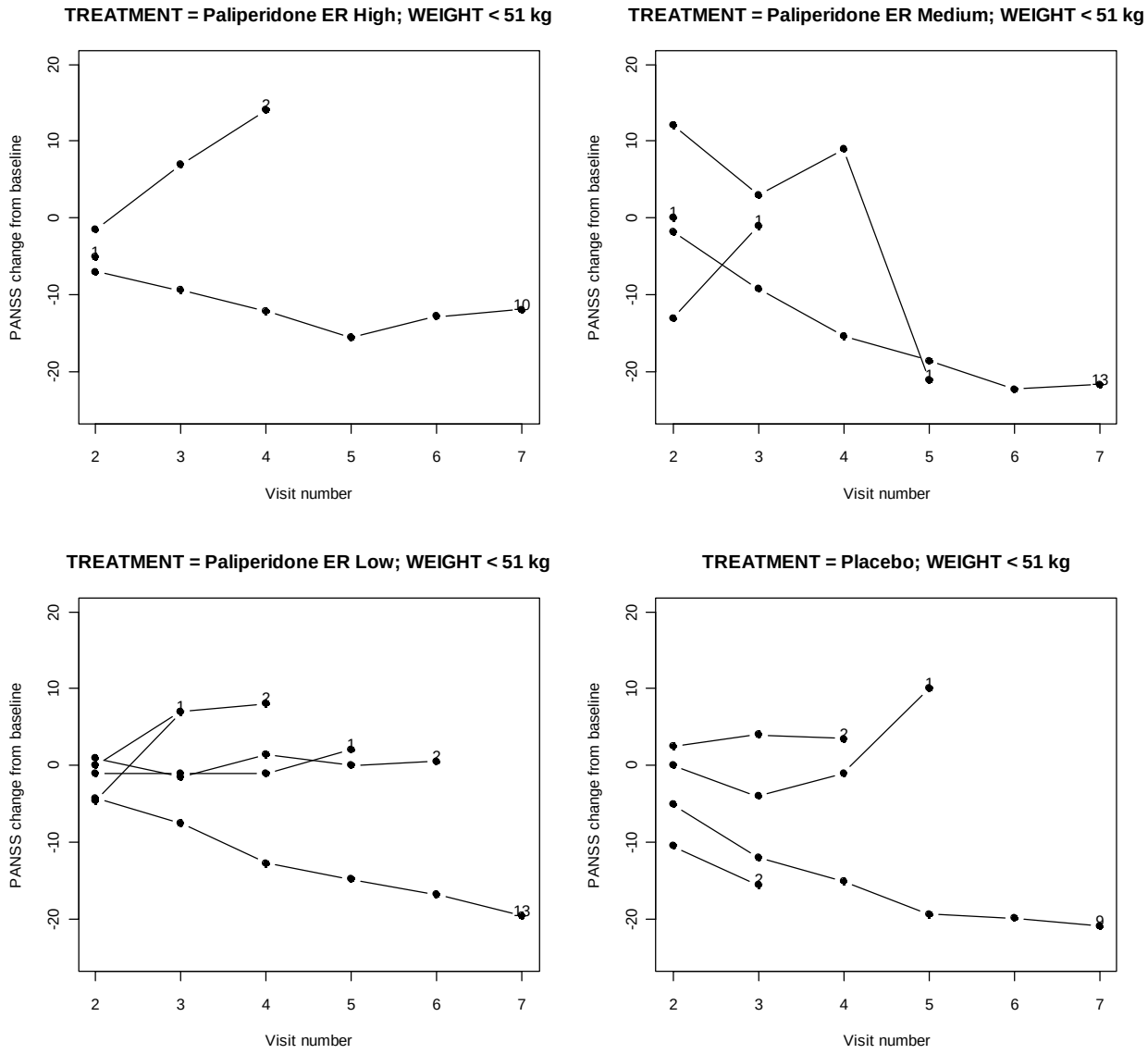
Figure 1 displays empirical cumulative distribution functions (CDF) of the primary endpoint, change from baseline in PANSS at week 6 (LOCF), for the four treatment arms in Study R076477-PSZ-3001. Negative values of the primary endpoint represent improvement. The cumulative distribution functions describe the percentage of patients (vertical axis) in each treatment arm with primary endpoint values (horizontal axis) equal to or less than a given number x where x varies from -40 to 20. About 15% of patients in the placebo arm had improvement of more than 30 units (i.e. to the left of -30) in PANSS Total score. For values of x larger than -25, the CDFs for the paliperidone Medium and High dose treatment arms separate from the CDF for the placebo arm. Numerically, in the paliperidone Medium and High dose treatment arms larger proportions of patients had negative value of the primary endpoint compared with placebo arm.

Figure 1. Empirical Cumulative Distribution Functions for Change from Baseline in PANSS Total Score



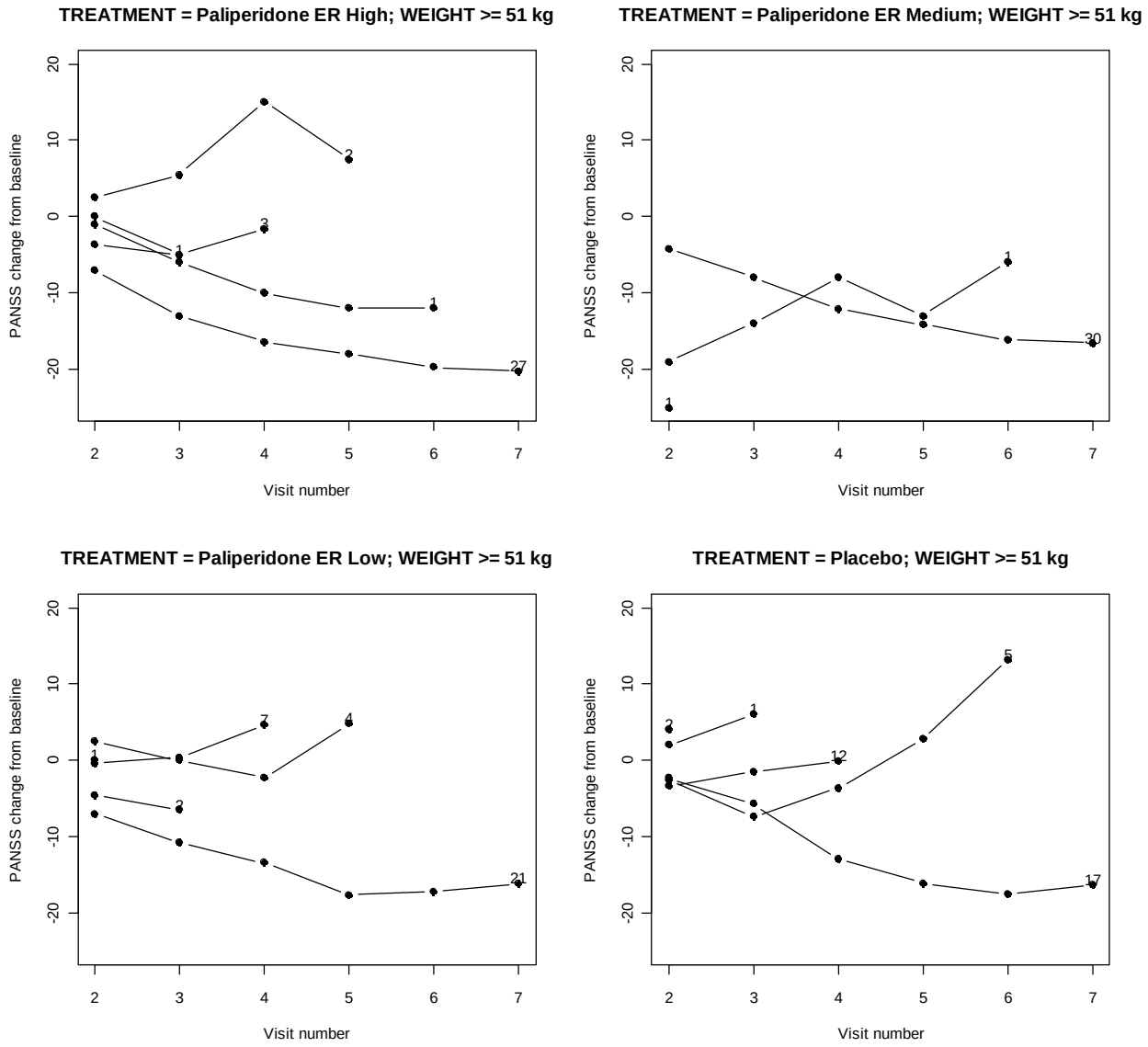
Source: Reviewer's results

Figure 2. PANSS Total Score Response Profiles by Treatment Group (weight < 51 kg)



Source: Reviewer's results
Remark: The numbers at the end of each curve represent the numbers of patients in each group.

Figure 3. PANSS Total Score Response Profiles by Treatment Group (weight ≥51 kg)

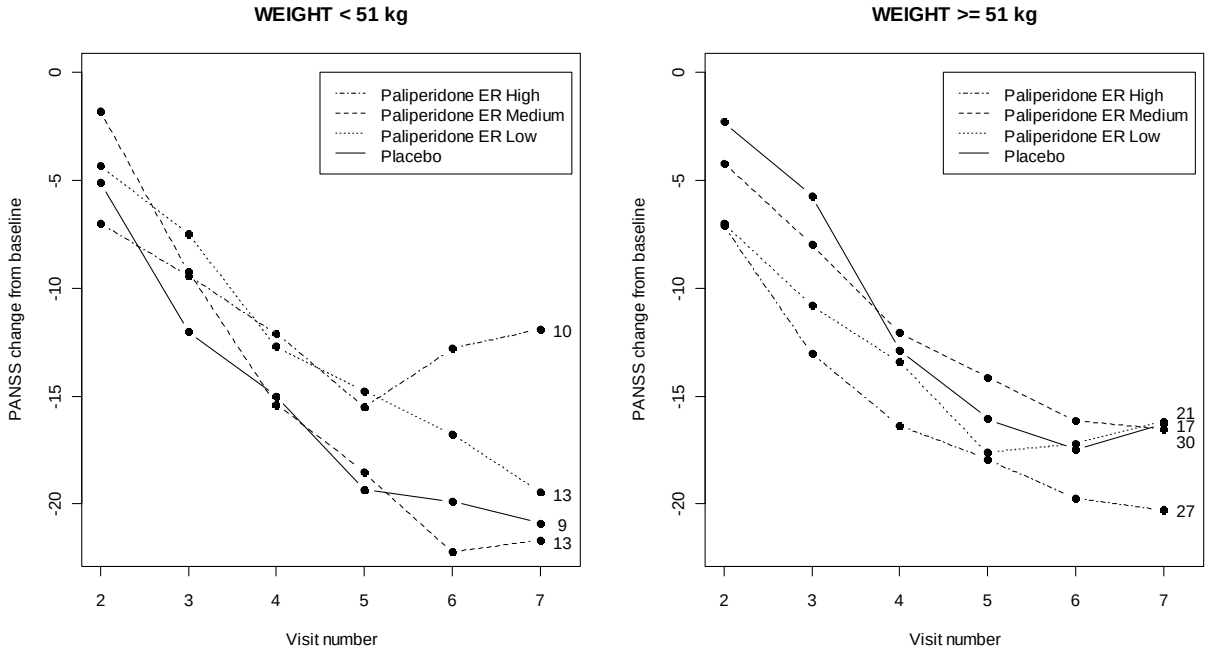


Source: Reviewer's results
Remark: The numbers at the end of each curve represent the numbers of patients in each group.

Each curve on the plots of the Figure 2 (patients weighing less than 51 kg) and Figure 3 (patients weighing at least 51 kg) shows the mean change of the PANSS total score by visit for subgroups of patients, grouped by their last visit before dropout or study completion (visits 2, 3, 4, 5, 6, and 7). The numbers at the end of each curve are the numbers of patients in each subgroup. The plots illustrate the general tendency of the patients to drop out from the study as their PANSS total score increases (getting worse). For all treatment arms, those patients who stayed to the end of the double-blind phase tend to have larger improvement from the beginning to the last visit than those who dropped out earlier. The Figure 4 compares the responses of treatment arms for the patients who stayed to the Week 6 of the double-blind phase. It appears that, within each weight subgroup, the treatment arms showed similar response at Week 6, except for Paliperidone ER High arm in the <51 kg weight subgroup. The plots of the Figure 2 and Figure 3 suggest a violation of the missing data mechanism MCAR (missing completely at

random), which is required for the LOCF imputation approach. Although the LOCF ANCOVA is the prespecified primary analysis, the strength of evidence is weakened because of the MCAR assumption, so results of the ANCOVA LOCF have to be interpreted with caution.

Figure 4. Response profiles for patients with PANSS Total Score Assessment at Week 6.



Source: Reviewer's results
Remark: The numbers at the end of each curve represent the numbers of patients in each group.

Sensitivity Analysis for The Primary Endpoint

This reviewer conducted sensitivity analysis on the primary endpoint. Change from baseline in PANSS total score was analyzed by mixed model with repeated measures (MMRM) using PROC MIXED in SAS. The model included country, treatment group, visit, and treatment group-by-visit interaction as factors and baseline PANSS total score as a covariate. The findings supported the primary analysis results and were consistent with the MMRM sensitivity analysis conducted by sponsor. The pailiperidone ER Medium dose was statistically significantly better than placebo at Week 6 (see Table 8).

Table 8. PANSS Total Score LS Mean Change from Baseline (MMRM analysis)

Week	**Placebo**	**Paliperidone Low**	**Paliperidone Medium**	**Paliperidone High**
Week 1 LS Mean(SE)	-2.3 (1.2) N=51	-3.5 (1.2) N=54	-4.0 (1.2) N=48	-5.7 (1.2) N=47
Week 2 LS Mean(SE)	-4.6 (1.5) N=49	-5.5 (1.5) N=53	-8.3 (1.5) N=46	-9.5 (1.5) N=46
Week 3 LS Mean (SE)	-7.0 (2.0) N=46	-7.5 (1.9) N=50	-12.6 (2.0) N=45	-11.6 (2.0) N=44
Week 4 LS Mean (SE)	-8.3 (2.4) N=32	-9.4 (2.3) N=41	-15.6 (2.4) N=45	-13.4 (2.4) N=40
Week 5 LS Mean (SE)	-7.6 (2.6) N=31	-9.6 (2.4) N=36	-17.9 (2.5) N=42	-13.9 (2.5) N=38
Week 6 LS Mean (SE)	-7.3 (2.8) N=26	-10.6 (2.6) N=34	-17.8 (2.6) N=43	-14.4 (2.6) N=37
LS Mean Difference from Placebo at Week 6		**-3.3 (3.76)**	**-10.4 (3.74)**	**-7.0 (3.80)**
p-value		**0.388**	**0.006**	**0.066**
95% Confidence Interval		**(-10.7, 4.2)**	**(-17.8, -3.0)**	**(-14.5, 0.5)**

Source: Reviewer's Results
Remark: The reported p-values and CI's are unadjusted

Exploratory Analysis by Actual Dose Group

The sponsor conducted exploratory analysis of the primary efficacy variable by actual dose group. The results are presented in Table 9. Numerically, all paliperidone ER dose groups (1.5, 3, 6, and 12 mg) were superior to the placebo group (nominal p-values of 0.507, 0.016, 0.044, and 0.014, respectively, with no multiplicity adjustment).

Table 9. PANSS Total Score - Change from Baseline to Week 6 by Actual Dose (LOCF ANCOVA)

		Placebo (N=51)	Paliperidone 1.5 mg (N=54)	Paliperidone 3 mg (N=16)	Paliperidone 6 mg (N=45)	Paliperidone 12 mg (N=34)
Basline Mean (SD)		90.6 (12.13)	91.6 (12.54)	92.1 (16.88)	90.8 (13.66)	91.0 (13.00)
Change from Basline Mean (SD)		-7.9 (20.15)	-9.8 (16.31)	-19.0 (15.45)	-13.8 (14.75)	-16.3 (15.41)
Difference compared with placebo (ANCOVA)	LS Mean (SE)		-2.1 (3.18)	-11.5 (4.75)	-6.8 (3.34)	-9.0 (3.64)
	p-value		0.507	0.016	0.044	0.014
	95% CI		(-8.39; 4.16)	(-20.85; -2.13)	(-13.35;-0.19)	(-16.23;-1.86)

Source: Clinical Study Report Table 18 (pg. 82)
Remark: The reported p-values and CI's are unadjusted p-values.

Reviewer's Comments:
The comparisons of actual doses with placebo were nonrandomized comparisons. The hypotheses associated with comparisons with actual dose were not prespecified as the primary or key secondary objectives, and these post hoc comparisons were not adjusted for multiplicity (in fact, multiplicity adjustment can not be properly made post hoc after data was unblinded and analyzed). The trial was not designed to investigate the effects of actual doses: patients were randomized to 4 treatment arms (Placebo, Low, Medium and High dose arms) within each weight subgroup (<51 kg and ≥ 51 kg). As a result, the paliperidone 3 mg actual-dose group included patients weighing <51 kg only, and the paliperidone 12 mg actual-dose group included patients with weight ≥51 kg only. Hence the actual-dose groups and the placebo arm may not be comparable with respect to baseline covariates (particularly those unobserved). Above all, the results of the weight subgroup analysis were not consistent between the two weight subgroups (see Table 11). For all these reasons, the results of such a post hoc "actual-dose" analysis cannot be properly interpreted and hence cannot be included in the labeling.

3.2 Evaluation of Safety

Not evaluated by this reviewer.

3.3 Benefit:Risk Assessment

Not evaluated by this reviewer.

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Race, Age, and Geographic Region

Exploratory subgroup analysis was conducted by this reviewer on the primary efficacy variable (change from baseline in PANSS Total score at Week 6), using LOCF ANCOVA models, including terms for treatment and baseline score. The subgroups of interest included gender, race, and geographic region. For all subgroups, except the black race subgroup, the 12-14 age subgroup, and the North America geographic region subgroup, the treatment effect appeared to be numerically in favor of paliperidone when compared with placebo.

Table 10. Subgroup Analysis (Change from Baseline in PANSS Total Score)

Subgroup	Placebo	Paliperidone ER		
		Low	Medium	High
Gender				
Male	N=23	N=30	N=31	N=33
Change from Bsln: Mean (SD)	-9.4 (21.8)	-10.4 (14.8)	-13.4 (11.5)	-14.1 (14.3)
LS Mean difference from Placebo	NA	-0.7 (4.3)	-3.6 (4.3)	-4.7 (4.2)
Female	N=28	N=24	N=17	N=14
Change from Bsln: Mean (SD)	-6.6 (19.0)	-9.0 (18.3)	-24.5 (16.4)	-13.2 (19.2)
LS Mean difference from Placebo	NA	-2.7 (5.1)	-19.4 (5.7)	-5.7 (6.0)
Race				
White	N=35	N=35	N=34	N=32
Change from Bsln: Mean (SD)	-3.9 (13.8)	-6.2 (12.4)	-16.9 (13.9)	-12.8 (11.7)
LS Mean difference from Placebo	NA	-2.1 (3.1)	-12.8 (3.1)	-8.8 (3.2)
Black	N=4	N=5	N=3	N=5
Change from Bsln: Mean (SD)	-43.8 (12.9)	-19.6 (22.6)	-12.3 (11.0)	-36.0 (18.4)
LS Mean difference from Placebo	NA	20.8 (11.5)	23.9 (13.7)	1.7 (11.9)
Asian	N=12	N=14	N=11	N=10
Change from Bsln: Mean (SD)	-7.3 (26.3)	-15.2 (20.6)	-20.0 (16.9)	-6.0 (17.3)
LS Mean difference from Placebo	NA	-7.9 (8.2)	-12.6 (8.7)	0.2 (9.0)
Geographical Region				
US	N=9	N=8	N=6	N=6
Change from Bsln: Mean (SD)	-24.1 (24.1)	-15.4 (23.7)	-22.7 (21.0)	-30.8 (20.8)
LS Mean difference from Placebo	NA	10.8 (9.2)	-0.4 (10.0)	-5.4 (10.0)
Non US	N=42	N=46	N=42	N=41
Change from Bsln: Mean (SD)	-4.4 (17.6)	-8.8 (14.8)	-16.5 (13.3)	-11.3 (13.5)
LS Mean difference from Placebo	NA	-4.5 (3.2)	-12.2 (3.3)	-7.1 (3.3)

Source: Reviewer's Results

4.2 Other Special/Subgroup Populations

This reviewer conducted exploratory weight subgroup analysis (<51 kg, ≥51 kg) on the primary efficacy variable (change from baseline in PANSS Total score at week 6), using ANCOVA LOCF model, including terms treatment and the baseline score as a covariate. The patient randomization to the treatment arms was stratified by the weight groups.

For all paliperidone arms, the observed treatment differences from placebo were highly inconsistent between the two weight subgroups. For patient subgroup with weight <51 kg, the placebo arm was numerically much superior to the paliperidone ER Low and the High dose treatment arms. For the paliperidone ER Medium dose, based on ANCOVA LOCF and MMRM analyses, the observed treatment difference from placebo in the <51 kg weight subgroup was much lower than that in the ≥51 kg subgroup.

Also, the observed improvement from baseline in PANSS Total score in the paliperidone ER High dose group for patients weighing <51 kg was very different than that in the ≥51 kg subgroup, with larger improvement observed in the ≥51 kg subgroup. The observed improvement from baseline of the High dose was numerically much smaller than that of the Medium dose in the <51 kg subgroup (i.e., the observed dose response is in an umbrella shape), whereas in the ≥51 kg group the observed improvements in the High dose and Medium doses were numerically similar.

Given the substantial inconsistencies in the improvements from baseline in PANSS Total score between these two weight subgroups, the results of subgroup analysis do not support that the drug is effective in the <51 kg subgroup, although one should also keep it in mind that there were not many patients in this lighter-weight subgroup. Because of very small sample sizes in the <51 kg subgroup, the observed improvements from baseline in the overall population were driven by the ≥51 kg subgroup.

Table 11 Subgroup Analysis by Weight Subgroup (Change from Baseline in PANSS Total Score)

Subgroup	Placebo	Paliperidone ER		
		Low	Medium	High
Weight Group				
ITT Population <51 kg	N=14	N=19	N=16	N=13
Change from Bsln: Mean (SD)	-14.4 (19.1)	-11.9 (17.5)	-19.0 (15.4)	-7.4 (15.3)
ANCOVA LOCF				
LS Mean (SE)	-14.6 (4.6)	-11.8 (3.9)	-19.0 (4.3)	-7.5 (4.7)
LS Mean difference with Placebo	NA	2.8 (6.1)	-4.5 (6.2)	7.1 (6.6)
MMRM				
LS Mean (SE)	-17.3 (4.9)	-15.6 (4.2)	-22.0 (4.4)	-10.1 (5.1)
LS Mean difference with Placebo	NA	1.8 (6.2)	-4.7 (6.4)	7.3 (6.8)
ITT Population ≥51 kg	N=37	N=35	N=32	N=34
Change from Bsln: Mean (SD)	-5.4 (20.2)	-8.6 (15.8)	-16.5 (13.9)	-16.3 (15.4)
ANCOVA LOCF				
LS Mean (SE)	-5.7 (2.7)	-8.0 (2.8)	-16.7 (2.9)	-16.3 (2.8)
LS Mean difference with Placebo	NA	-2.3 (3.9)	-11.0 (4.0)	-10.6 (3.9)
MMRM				
LS Mean (SE)	-3.1 (3.4)	-8.2 (3.5)	-16.1 (3.2)	-16.3 (3.2)
LS Mean difference with Placebo	NA	-5.1 (4.8)	-13.0 (4.6)	-13.3 (4.6)

Source: Reviewer's Results

5. SUMMARY AND CONCLUSIONS

5.1 Statistical Issues and Collective Evidence

In the primary analysis of PANSS Total score, patients in the paliperidone ER medium dose treatment group (3mg daily for patients weighing 29 to less than 51 kg, 6mg daily for patients weighing at least 51 kg) showed statistically significant improvement over patients in the placebo treatment group. Based on the ANCOVA model with treatment and country as factors and baseline PANSS total score as a covariate, using a closed testing procedure with Dunnett's test to control for multiplicity, the LS mean change in the paliperidone ER Medium treatment group was statistically significantly superior to that in the placebo group (p=0.006). The paliperidone ER High and Low doses did not demonstrate statistical superiority over placebo. Sensitivity MMRM analysis supported the results of ANCOVA analysis.

This reviewer noted that the observed improvements from baseline in PANSS Total score and treatment differences from placebo were highly inconsistent between the two weight subgroups (see Table 11). For the <51 kg weight subgroup, the observed dose response for paliperidone had an umbrella shape with a large improvement from baseline observed in the Medium dose group, while for the other weight subgroup the observed improvements in the Medium and the High dose groups were large and were close to each other. There is apparently a large placebo response observed in the <51 kg subgroup; such a high placebo response is unusual. It is also not clear why the dose response relationships are very different in the two weight subgroups. Although these inconsistencies could potentially or partially be due to small sample sizes in the <51 kg subgroup, there could be other unknown factors for the inconsistencies. If the drug is effective to adolescents weighing ≥51 kg, it'd be surprising that the drug is not effective to adolescents weighing <51 kg, but the problem is at what dose level. If the dose response relationships are indeed very different in these two subgroups, perhaps the trial shouldn't have been designed as it was and efficacy shouldn't have been analyzed by combining the two weight subgroups. Because of very small sample sizes in the <51 kg subgroup, the observed improvements from baseline in PANSS Total score in the overall population were driven by the ≥51 kg subgroup. Given the substantial inconsistencies in the observed improvements from baseline between these two weight subgroups, the results of subgroup analysis do not support that the drug is effective in the <51 kg subgroup, and another trial would be needed to confirm the dose response relationship in the overall adolescent population.

The sponsor included the following results of exploratory analysis by actual dose in Section 14.1 of the proposed labeling: "*Efficacy was evaluated using PANSS. This study demonstrated the efficacy of INVEGA in adolescent subjects with schizophrenia* (b) (4) The comparisons of actual doses with placebo were nonrandomized comparisons. The hypotheses associated with comparisons with actual dose were not prespecified as the primary or key secondary objectives, and these post hoc comparisons were not adjusted for multiplicity (in fact, multiplicity adjustment can not be properly made post hoc after data was unblinded and analyzed). The trial was not designed to investigate the effects of actual doses: patients were randomized to 4 treatment arms (Placebo, Low, Medium and High dose arms) within each weight subgroup (<51 kg and ≥ 51 kg). As a result, the paliperidone 3 mg actual-dose group included patients weighing <51 kg only, and the paliperidone 12 mg actual-dose group included patients with weight ≥51 kg only. Hence the actual-dose groups and the placebo arm may not be comparable with respect to baseline covariates (particularly those unobserved). Above all, the results of the weight subgroup analysis were not consistent between the two weight subgroups (see Table 11). For all these reasons, the results of such a post hoc "actual-dose" analysis cannot be properly interpreted and hence cannot be included in the labeling.

5.2 Conclusions and Recommendations

In the primary analysis of PANSS Total score, adolescent patients with schizophrenia in the paliperidone ER medium dose treatment group (3 mg daily for patients weighing 29 to less than 51 kg, 6 mg daily for patients weighing at least 51 kg) showed statistically significant improvement over patients in the placebo treatment group. The paliperidone ER High and Low doses did not demonstrate statistical superiority over placebo.

The observed improvements from baseline in PANSS Total score and treatment differences from placebo were highly inconsistent between the two weight subgroups (see Table 11). Because of very small sample sizes in the <51 kg subgroup, the observed improvements from baseline in the overall population were driven by the ≥51 kg subgroup. The results of subgroup analysis by weight group do not support that the drug is effective in the <51 kg subgroup, and another trial would be needed to confirm the dose response relationship in the overall adolescent population.

The sponsor included the following results of exploratory analysis by actual dose in Section 14.1 of the proposed labeling: "Efficacy was evaluated using PANSS. This study demonstrated the efficacy of INVEGA in adolescent subjects with schizophrenia (b) (4) . The minimum effective dose for INVEGA in this population was 3 mg/day." In this reviewer's opinion, the exploratory efficacy results by actual dose can not be properly interpreted and hence they cannot be included in the labeling (see reviewer's comments in Section 5.1, page 18).

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

GEORGE KORDZAKHIA
03/08/2011

PEILING YANG
03/08/2011
I concur.

HSIEN MING J J HUNG
03/08/2011
I concur