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/Serial Number	22,037 (Supplement 2, SD-54)
Drug Name:	Intuniv™ (guanfacine) extended release tablets
Indication(s):	ADHD (Attention Deficit Hyperactivity Disorder)
Applicant:	Shire
Dates:	Received Date: April 28, 2010
Review Priority:	Standard
Biometrics Division:	Division of Biometrics I
Statistical Reviewer:	Eiji Ishida, M.S.
Concurring Reviewers:	Peiling Yang, Ph.D.
	Kooros Mahjoob, Ph.D.
Medical Division:	DPP (Division of Psychiatry Products)
Clinical Team:	Silvana Borges, M.D.
	Ni Aye Khin, M.D.
Project Manager:	ShinYe Chang, Pharm. D.

Keywords: Clinical studies; NDA review

Table of Contents

List of Tables

List of Figures

1. EXECUTIVE SUMMARY

1.1 Conclusions and Recommendations[1]

The statistical evidence based on Study SPD503-313 supports Sponsor's claim that Intuniv™ is efficacious as an adjunctive therapy to a long-acting oral psychostimulant in the treatment of children and adolescents aged 6-17 years with a diagnosis of Attention-Deficit Hyperactivity Disorder.

1.2 Brief Overview of Clinical Studies

Sponsor submitted a phase 3 study (SPD503-313) in their new drug application (NDA) to request revisions to the US Prescribing Information (USPI) of Intuniv™ (SPD503). The study results of SPD503-313 were the basis of the proposed clinical efficacy revision. In this study, Sponsor evaluated efficacy of optimized SPD503 (1, 2, 3, and 4 mg/day) as an adjunctive therapy to a long-acting oral psychostimulant in the treatment of children and adolescents aged 6-17 years with a diagnosis of Attention-Deficit Hyperactivity Disorder (ADHD).

In Study SPD503-313, randomized subjects were enrolled in the 5-week dose-optimization phase to reach their optimized dose (1, 2, 3, or 4 mg/day), and then continued onto the 3-week dose-maintenance phase with their optimized dose. Efficacy assessments based on ADHD-RS-IV total score were conducted every week during the 8-week efficacy study, before a follow-up phase. The primary efficacy variable was the change from baseline to the 8th week endpoint (last visit of the dose-maintenance phase) in ADHD-RS-IV total score.

1.3 Statistical Issues and Findings[2]

The reviewer found no major statistical issues regarding the efficacy conclusion of the study. First of all, the primary analysis result reported in Sponsor's clinical study report (CSR) was confirmed and verified. Secondly, this reviewer conducted pre-specified sensitivity analyses and secondary analyses of the primary efficacy variable, and confirmed Sponsor's results when available. The results were consistent with the primary analysis result. Thirdly, subgroup analysis results suggested little deviation from the primary analysis result.

The dropout rate, calculated as the number of subjects (67 subjects) who early terminated prior to the efficacy endpoint divided by that of randomized subjects (455 subjects), was about 15%. This reviewer concluded that there was no evidence suggesting that the dropouts and missing data impacted on the primary efficacy analysis to the extent that the study result should be questioned.

The reviewer found that Sponsor's study analyses and clinical study report (CSR) had some quality issues. As an example, Sponsor neither correctly performed the pre-specified sensitivity analysis (based on the MMRM method), nor discussed the analysis results in the CSR. In addition to this deficiency, the CSR did not sufficiently address the issue of dropouts and missing data. It lacked examinations of the underlying assumptions the adopted statistical method relies on.

[1] Refer to Section 5.2
[2] Refer to Section 5.1

2. INTRODUCTION

2.1 Overview

This review provides a statistical evaluation of Intuniv™ (extended-release guanfacine hydrochloride) as an adjunctive therapy to a psychostimulant in the treatment of Attention-Deficit Hyperactivity Disorder (ADHD). The evaluation is based on the NDA submission of a phase 3, double-blind, randomized, placebo-controlled, multicenter, dose-optimization study (Study SPD503-313). The patient population of this study is children and adolescents aged 6-17-years with a diagnosis of ADHD and a suboptimal, partial response to psychostimulant. The study consists of phases of screening, dose-optimization (5 weeks), dose-maintenance (3 weeks), and dose-tapering (1 week). During the dose-optimization phase, the optimized dose of SPD503 (1, 2, 3, or 4mg/day) was determined for each subject, and its efficacy was compared to placebo at the end of the dose-maintenance phase (8 weeks). All the subjects were given a pre-specified, long-acting, oral psychostimulant[3] for the entire study period.

2.2 Data Sources

Sponsor submitted the NDA Supplement-2 on April 28, 2010. The submission is located at the CDER's electronic document room: \\Cdsesub1\evsprod\NDA022037\0007\m5\datasets\spd503-313. The concomitant medication and adverse event data sets of this submission turned out to be insufficient, and Sponsor resubmitted these raw data sets. They are located at the CDER's electronic document room: \\Cdsesub1\evsprod\NDA022037\0023\m5\datasets\spd503-313.

The protocol, clinical study report and statistical analysis plan are located at the CDER's electronic document room: \\Cdsesub1\evsprod\NDA022037\0007\m5\53-clin-stud-rep\535-rep-effic-safety-stud\adhd\5351-stud-rep-contr\spd503-313.

3. STATISTICAL EVALUATION

Study title:

The title of Study SPD503-313 is "A Phase 3, Double-blind, Randomized, Placebo-controlled, Multicenter, Dose-optimization Study Evaluating the Efficacy and Safety of SPD503 in Combination With Psychostimulants in Children and Adolescents Aged 6-17 Years With a Diagnosis of Attention deficit Hyperactivity Disorder (ADHD)".

3.1 Evaluation of Efficacy

3.1.1 Study Objectives

Primary objective:

To assess the efficacy of optimized SPD503 (1, 2, 3, and 4mg/day), dosed either in the morning or evening, compared to placebo, when co-administered with psychostimulants, in the treatment of children and adolescents aged 6-17 years with a diagnosis of ADHD, with a suboptimal, partial response to

[3] ADDERALL XR® [mixed salts of a single-entity amphetamine product], VYVANSE® [lisdexamfetamine dimesylate], CONCERTA® [methylphenidate HCl], FOCALIN XR® [dexmethylphenidate HCl], RITALIN LA® [methylphenidate HCl extended-release], METADATE CD® [methylphenidate HCl, USP], or FDA-approved generic equivalents.

stimulants, as measured by the Attention-Deficit Hyperactivity Disorder Rating Scale-IV (ADHD-RS-IV) change from Baseline score at Endpoint.

Secondary objective:

To assess the efficacy of optimized SPD503, dosed either in the morning or evening, compared to placebo, when co-administered with psychostimulants, in the treatment of children and adolescents aged 6-17 years with a diagnosis of ADHD, with a suboptimal, partial response to stimulants, as measured by:

- Conners' Global Index – Parent (CGI-P) at morning and evening timepoints
- Clinical Global Impressions – Severity of Illness Scale (CGI-S)
- Clinical Global Impressions – Improvement Scale (CGI-I)
- Parent Global Assessment (PGA)
- Before-school Functioning Questionnaire - Wil-Hammer (BSFQ)
- Oppositional subscale of the Conners' Parent Rating Scale – Revised Long Form (CPRS-R:L).

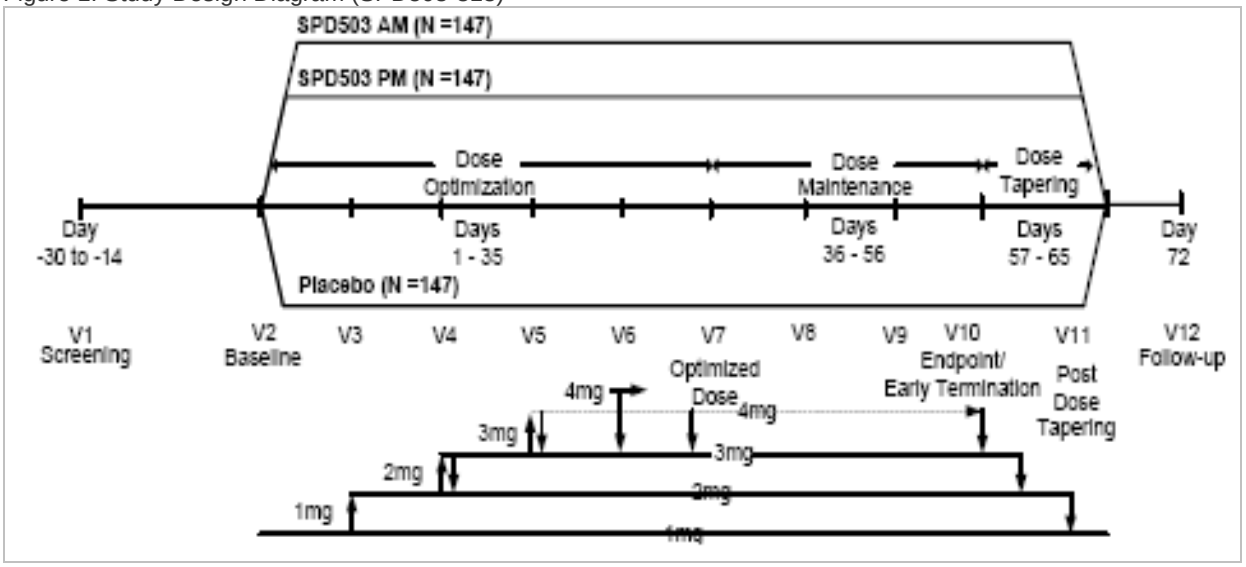
To evaluate sleep parameters, as measured by a Post-Sleep Questionnaire (PSQ), of SPD503 and placebo when co-administered with psychostimulants in children and adolescents with a diagnosis of ADHD.

To evaluate the safety and tolerability of SPD503 based on TEAEs, clinical laboratory tests, physical examinations, vital signs, and ECGs.

3.1.2 Study Design

Figure 1 illustrates the schedule of study phases (screening - baseline, dose-optimization, dose-maintenance, dose-tapering, follow-up). The primary efficacy assessment was based on the data of ADHD-RS-IV total score collected from the 5-week dose-optimization and 3-week dose-maintenance phases.

Figure 1: Study Design Diagram (SPD503-313)



[Source: *Figure 1. Study Design Diagram* of Sponsor's CSR (page 24)]

In this study, there were three treatment groups (SPD503-AM, SPD503-PM, and Placebo)[4]. An optimized dose was determined during the dose-optimization phase, and study subjects were expected to continue on their optimized dose during the dose-maintenance phase at the end of which the primary efficacy was assessed. All the subjects continued on their pre-specified, long-acting, oral psychostimulant[5] for the entire study period.

Study phases:

With reference to the baseline visit (Visit 2), visits were scheduled 7 days apart during the dose-optimization and the dose-maintenance phases of the study (±2 days).
At the baseline visit, eligible subjects were randomized to SPD503 AM (SPD503 in the morning, placebo in the evening), SPD503 PM (placebo in the morning, SPD503 in the evening), or placebo (placebo morning and evening) and received their first dose of SPD503/placebo at bedtime in the evening of their baseline visit. During the dose optimization phase (Days 1-35; Visits 3-7), subjects initialized treatment with 1 mg (SPD503 or matching placebo) and received 1 tablet every morning (upon awakening) and evening (at bedtime), while maintaining their current (baseline), stable dose of psychostimulant treatment taken each morning. Subjects returned to the site weekly for evaluation of ADHD symptoms, and were titrated to their optimal dose.
During the dose-maintenance phase (Days 36-56; Visits 8-10), subjects were maintained on their optimal dose for an additional 21 days, visiting the study site on a weekly basis for efficacy and safety evaluations. The final evaluation for the primary efficacy was scheduled to occur at Visit 10.

Sample size calculation:

The primary efficacy measurement for this study is the change from baseline score at endpoint on the ADHD-RS-IV scale. To detect an effect size of at least 0.4 between either SPD503 group and placebo (equivalent to a standard deviation of 10 points and a difference between active and placebo of approximately 4 points) at 90% power and a significance level of 0.05 (2-sided) using a 2-sample t-test with a 1:1:1 (SPD503 AM: SPD503 PM: placebo) allocation ratio, Sponsor found it necessary to have approximately 399 subjects (133 subjects for each treatment group - SPD503 AM, SPD503 PM, and placebo).
To account for subjects who drop out without providing post-baseline ADHD-RS-IV data, Sponsor planned to enroll a total of 441 subjects (147 subjects for each treatment group - SPD503 AM, SPD503 PM, and placebo). The actual total number of enrolled subjects was 461.

3.1.3 Statistical Method and Analysis

Definition of study population in primary analysis:

The full analysis set (FAS) was defined as the set of subjects who received at least one dose of any study drug during the study. There were 455 subjects in the FAS, as shown in Table 1. The *primary efficacy analysis set* consisted of all randomized subjects who took at least one dose of study drug, and provided at least one post baseline efficacy assessment. There were 449 such subjects, as discussed in the next section.

[4] In this review, the treatment groups are denoted by SPD503-AM, SPD503-PM, and Placebo. All subjects of the treatment groups, SPD503-AM, SPD503-PM, and Placebo, were treated for the entire study with a psychostimulant determined prior to the first administration of the investigational drug.
[5] ADDERALL XR® [mixed salts of a single-entity amphetamine product], VYVANSE® [lisdexamfetamine dimesylate], CONCERTA® [methylphenidate HCl], FOCALIN XR® [dexmethylphenidate HCl], RITALIN LA® [methylphenidate HCl extended-release], METADATE CD® [methylphenidate HCl, USP], or FDA-approved generic equivalents.

Primary endpoint and analyses:

The primary efficacy measure was defined as the change from baseline of the ADHD-RS-IV total score. The last on-therapy, post-randomization treatment week, prior to any dose tapering, at which a valid ADHD-RS-IV total score was collected, is defined as the primary efficacy endpoint[6]. The LOCF (last observation carried forward) approach was used as an imputation method for missing data of the primary efficacy endpoint.

The primary efficacy analysis was performed on the *primary efficacy analysis set*[7], using the LOCF ANCOVA method. The analysis model included treatment group and psychostimulant type (amphetamine or methylphenidate) as model factors, and baseline score as a covariate. The null hypothesis was that there was no difference between optimized SPD503 AM and placebo, or between optimized SPD503 PM and placebo. The hypothesis test was conducted for 2-sided, overall 0.05% type I error rate. Dunnett's adjustment for multiple comparisons was used to control the overall significance level in the analysis.

The following secondary analyses of the primary efficacy measure were performed on the primary analysis set:

- The primary efficacy analysis was repeated for each of Hyperactivity/Impulsivity and Inattentiveness subscale total scores.
- Cochran-Mantel-Haeszel tests were performed on the following responses adjusting for psychostimulant type to compare each of the treatment groups with the placebo group.
    1. Responder analysis: a responder was defined as a subject who had a reduction from baseline in the ADHD-RS-IV total score of at least 25%.
    2. Remission analyses:
        - Symptomatic remission: a responder was defined as a subject who had ADHD-RS-IV total score of less than or equal to 18.
        - Syndromal remission: a responder was defined as a subjects who had ADHD-RS-IV total score of less than or equal to 18 and CGI-S of less than or equal to 2.

Secondary endpoints and analyses:

Other secondary analyses were performed on the following secondary measurements: Clinical Global Impressions-Severity (CGI-S), Clinical Global Impressions-Improvement (CGI-I), Clinical Global Index – Parent (CGI-P), Parent Global Assessment (PGA), Before-school Functioning Questionnaire (BSFQ), Oppositional Subscale of the Conners' Parent Rating Scale Revised: Long Form (CPRS-R:L), and Post-Sleep Questionnaire (PSQ).

No key secondary endpoint was pre-specified.

[6] Refer to Page 40 of CSR. This definition gives ambiguity and was inconsistent with the actual analysis data set. According to this definition, subjects who had an efficacy assessment at Visit 10 will be included in the analysis set, as long as the first tapering dose date occurs after the assessment. However, subjects who met this criterion were not included in the actual primary analysis set, if the assessment was performed after the last non-tapering dose date.

[7] The CSR only defines the primary analysis set as the FAS (N=455), but in fact the primary analysis was performed on the FAS subjects who had at least one post-baseline measure (N=449).

3.1.4 Efficacy Results

3.1.4.1 Subject Disposition and Baseline Demographic Characteristics

3.1.4.1.1 Subject Disposition

In Table 1, a total of 461 subjects (all subjects who were informed consented) were randomly assigned in a 1:1:1 ratio to SPD503 AM (N=154), SPD503 PM (N=153), or Placebo (N=154). The full analysis set (FAS), all subjects who received at least one dose of any investigational product during this study, is reported to consist of 455 subjects, 98.7% of the randomized subjects. Out of the 461 randomized subjects, 6 subjects[8] (1.3%) were not included in the FAS (Table 1.2.2 of CSR of the submission (page 133)). These 6 subjects terminated before receiving any investigational product (Table 2).

Table 1: Subject Dispositions of Study SPD503-313

	Placebo + Stimulant (N=154)	SPD503 AM + Stimulant (N=154)	SPD503 PM + Stimulant (N=153)	Total (N=461)
Subjects				
Randomized	154 (100.0)	154 (100.0)	153 (100.0)	461 (100.0)
Safety Population[a]	153 (99.4)	150 (97.4)	152 (99.3)	455 (98.7)
Full Analysis Set[a]	153 (99.4)	150 (97.4)	152 (99.3)	455 (98.7)
Completed through Visit 10[b]	131 (85.1)	125 (81.2)	130 (85.0)	386 (83.7)
Study Completers[c]	129 (83.8)	121 (78.6)	128 (83.7)	378 (82.0)
Early Termination[d]	25 (16.2)	33 (21.4)	25 (16.3)	83 (18.0)
Reasons for Early Termination				
Adverse Event	1 (0.6)	4 (2.6)	6 (3.9)	11 (2.4)
Protocol non-adherence/subject non-compliance	3 (1.9)	8 (5.2)	6 (3.9)	17 (3.7)
Refused further participation in the study	11 (7.1)	7 (4.5)	8 (5.2)	26 (5.6)
At or Before Visit 10	10 (6.4)	7 (4.5)	7 (4.5)	24 (5.2)
After Visit 10	1 (0.7)	0 (0.0)	1 (0.7)	2 (0.4)
Lost to follow-up	5 (3.2)	9 (5.8)	3 (2.0)	17 (3.7)
At or Before Visit 10	5 (3.2)	8 (5.1)	2 (1.3)	15 (3.3)
After Visit 10	0 (0.0)	1 (0.7)	1 (0.7)	2 (0.4)
Lack of efficacy	5 (3.2)	3 (1.9)	2 (1.3)	10 (2.2)
Other	0 (0.0)	2 (1.3)	0 (0.0)	2 (0.4)

a Includes all subjects who received at least 1 dose of any study drug during this study.
b Visit 10 was the last visit before taper and is considered the Endpoint for statistical purposes, provided that subjects were still on study drug.
c Includes subjects who completed through Visit 12 (final follow-up visit).
d Early termination includes any subject who did not complete all visits through Visit 12.
* Numbers in cells are numbers of subjects and those of apprentices are percentages. *N* denotes the number of subjects.
[Source: *Table 2 Summary of Subject Disposition* of Sponsor's CSR (page 49)]

In addition, there were 6 subjects[9] who had no post-baseline efficacy assessment. In total, 12 subjects were not included in Sponsor's analysis set for the primary analysis. Sponsor's primary efficacy analysis was conducted on the rest, i.e., the set of the 449 subjects (See Table 2.1.1.1 of CSR of the submission (page 195)). The sample size calculation was based on an expectation that about 10% of enrolled subjects will not have a post-baseline efficacy assessment. The actual rate was 2.6% (12 out of 461 subjects), which was much lower than the expected rate.

[8] The subjects are: 20-007, 30-006, 40-022, 46-004 (SPD503 AM), 19-007 (SPD503 PM), 58-004 (Placebo).
[9] The subjects are: 19-006 (SPD503 AM), 07-007, 23-009, 28-014, 56-001 (SPD503 PM), 62-003 (Placebo).

Early terminated subjects reported in Table 1 (a table of subject dispositions of the CSR of the submission) are defined as those who did not complete through Visit 12. In Table 1, 386 subjects are reported as completers through Visit 10. However, 382 subjects were included as completers through Visit 10 in the actual analysis data set that Sponsor used for the primary efficacy analysis (the 449 evaluable subjects of the FAS). This can be confirmed in Table 2.1.1.1 of CSR (page 199).

Sponsor decided that 7 of the 386 subjects be excluded from the set of Visit 10 completers for the reason that they had their Visit 10 efficacy assessment after their last non-tapering dose date[10]. In addition, 3 subjects who terminated their study early at Visit 10 were included in the efficacy analysis set of completers through Visit 10[11]. They had a valid efficacy assessment for Visit 10 (See also Table 2). Therefore, in total 382 subjects were valid as complete cases. Since 449 subjects were included in the efficacy analysis set, 67 subjects (449 minus 382) were included as subjects who early terminated prior to Visit 10 in Sponsor's efficacy assessment (Table 8).

Table 2: Subject Disposition for Efficacy Assessment

Analysis set	Number of subjects		Subjects who:	Note
FAS (full analysis set)	455		had baseline assessment	6 of 461 randomized subjects were not exposed to any treatment.
Set of efficacy analysis (LOCF subjects)	449		had baseline assessment and at least one post-baseline assessment	There were 6 subjects who had no post-baseline ADHD-RS-IV total score. (The primary efficacy analysis set consists of the 449 subjects).
ET* subjects	60		early terminated before Visit 10	Subjects were early terminated at one of the visits: Visits 4, 5, 6, 7, 8, or 9.
Visit 10 completers (Complete Cases)	382	3	had Visit 10 but early terminated for Visit 10	An efficacy assessment at Visit 10 was performed on or prior to last non-tapering dose date**. (382 such subjects were included in Complete case analysis).
		379	had Visit 10	
ET* subjects	7		had Visit 10 but early terminated for Visit 10	Visit 10 efficacy assessment was performed after their last non-tapering dose date. (These patients' Visit 10 efficacy data were treated as missing).

* *ET* denotes "early terminated".
** A last non-tapering dose date necessarily occurs prior to a on-tapering dose date.
Note: In Table 1, Sponsor reported 386 subjects as "Completed through Visit 10". However, Sponsor conducted the complete case efficacy analysis based on 382 subjects ("Visit 10 completers" in Table 2).
[Source: Reviewer's analysis]

Regarding *reasons for early termination* recorded in Table 1, disparities among the treatment groups in the distributions of early terminated subjects were negligible for the primary efficacy assessment. Fewer subjects terminated early in the SPD503 treatment groups in *Refused further participation in the study (at or before Visit 10)* and *Lack of efficacy* than in the placebo group. It does not appear that early termination was associated with efficacy of the investigational drug.

[10] The subjects are: 25-015, 33-005, 33-006, 57-003 (SPD503 AM), 23-007, 43-014, 50-011 (SPD503 PM). These subjects, in fact, had efficacy assessment before the first on-tapering dose date.
[11] "Completers through Visit 10" includes subjects who were early terminated at Visit 10 but had an efficacy assessment at Visit 10 before taking on-tapering dose (and prior to last non-tapering dose date). The subjects are: 23-015 (Placebo), 23-016 (SPD503 AM), 30-002 (Placebo).

3.1.4.1.2 Demographic Characteristics

As shown in Table 3, study subjects of the full analysis set (FAS) were distributed fairly evenly among the treatment groups on baseline demographic characteristics of subgroups in Gender, Age, Height, Weight, BMI in terms of sample means/median (continuous variables), and in Age group, Sex, Race, Concomitant Psychostimulants, in terms of sample proportions (categorical variables). Randomization was stratified by concomitant psychostimulants. It can be observed from Table 3 that subject allocations were well balanced among the psychostimulant groups.

Table 3: Demographic Characteristics by Treatment Group

		Placebo + Stimulant (N=153)	SPD503 AM + Stimulant (N=150)	SPD503 PM + Stimulant (N=152)	Total (N=455)
Age, years					
	Mean (SD)	10.8 (2.3)	11.0 (2.6)	10.6 (2.3)	10.8 (2.4)
	Median	11.0	11.0	10.0	11.0
	Min, Max	6, 17	6, 17	6, 17	6, 17
	Categories, n (%)				
	6-12 years	123 (80.4)	114 (76.0)	124 (81.6)	361 (79.3)
	13-17 years	30 (19.6)	36 (24.0)	28 (18.4)	94 (20.7)
Sex, n (%)					
	Male	112 (73.2)	108 (72.0)	106 (69.7)	326 (71.6)
	Female	41 (26.8)	42 (28.0)	46 (30.3)	129 (28.4)
Race, n (%)					
	White	102 (66.7)	104 (69.3)	102 (67.1)	308 (67.7)
	Black or African American	35 (22.9)	28 (18.7)	37 (24.3)	100 (22.0)
	Native Hawaiian or other Pacific Islander	1 (0.7)	1 (0.7)	1 (0.7)	3 (0.7)
	Asian	1 (0.7)	2 (1.3)	3 (2.0)	6 (1.3)
	American Indian or Alaskan Native	0 (0.0)	1 (0.7)	0 (0.0)	1 (0.2)
	Other	14 (9.2)	14 (9.3)	9 (5.9)	37 (8.1)
Height, in					
	Mean (SD)	57.65 (5.5)	58.10 (6.0)	57.05 (5.4)	57.60 (5.6)
	Median	57.00	57.80	56.60	57.00
	Min, Max	47.6, 70.0	46.0, 73.7	46.4, 71.0	46.0, 73.7
Weight, lbs					
	Mean (SD)	89.14 (27.9)	90.76 (29.7)	85.40 (26.5)	88.43 (28.1)
	Median	85.50	83.00	76.20	81.40
	Min, Max	55.0, 164.0	55.0, 175.0	55.0, 164.0	55.0, 175.0
BMI					
	Mean (SD)	18.39 (3.0)	18.37 (2.8)	18.06 (2.9)	18.27 (2.9)
	Median	17.62	17.84	17.41	17.62
	Min, Max	13.7, 27.0	14.2, 28.2	13.4, 28.3	13.4, 28.3
Concomitant Psychostimulant, n (%)					
	Adderall XR	27 (17.6)	26 (17.3)	28 (18.4)	81 (17.8)
	Concerta	69 (45.1)	69 (46.0)	68 (44.7)	206 (45.3)
	Focalin XR	9 (5.9)	9 (6.0)	9 (5.9)	27 (5.9)
	Metadate CD	2 (1.3)	2 (1.3)	1 (0.7)	5 (1.1)
	Ritalin LA	1 (0.7)	1 (0.7)	0 (0.0)	2 (0.4)
	Vyvanse	45 (29.4)	43 (28.7)	46 (30.3)	134 (29.5)

* *SD* denotes standard deviation.

[Source: *Table 3 Summary of Demographic and other Baseline Characteristics* of Sponsor's CSR (page 50 - 51)]

3.1.4.2 Sponsor's Efficacy Analysis Results

3.1.4.2.1 Primary Efficacy Measure

Primary analysis:

The analysis method for the primary efficacy was pre-specified as the ANCOVA (LOCF) model with a covariate of baseline score and factors of treatment and psychostimulant type.

Table 4: Sponsor Primary Efficacy Analysis (ANCOVA – LOCF)

Primary analysis (Endpoint – Analysis set)	Treatment Group	Sample Size (N=449)	Change from baseline in ADHD-RS-IV Total score			
			Mean (SD*)	LS mean (SE*)	Placebo-adjusted difference** in LS mean (95% CI)	p-value
ANCOVA (LOCF – FAS)	SPD503 AM	149	-20.4 (12.77)	-20.3 (0.97)	-4.5 (-7.5, -1.4)	.002
	SPD503 PM	148	-21.0 (12.39)	-21.2 (0.97)	-5.3 (-8.3, -2.3)	< .0001
	Placebo	152	-16.0 (11.77)	-15.9 (0.96)		

* *SD* denotes standard deviation, and *SE* standard error.
** *LS mean* and *p-value* for *placebo-adjusted difference* were based on type III sum of squares from an ANVOVA model for the change from baseline with treatment group and psychostimulant type as fixed effects, and baseline value as a covariate. The reported p-values were based on Dunnett's multiple comparison procedure.
Note: Precisely, the analysis set was not the FAS but the set of efficacy assessment (See Table 2).
[Source: *Table 5 Summary of Change from Baseline in ADHD-RS-IV Total Score at Endpoint=LOCF (FAS)* of Sponsor's CSR (page 56)]

The primary endpoint was change from baseline score in ADHD-RS-VI total score. The primary efficacy assessment consisted of two comparisons, SPD503 AM vs. Placebo and SPD503 PM vs. Placebo, and their multiplicity adjustment in the statistical hypothesis tests was performed using Dunnett's procedure.

Table 4 displays the sponsor's primary analysis results. Both the SPD503 AM and SPD503 PM groups had a significantly larger improvement from baseline in ADHD-RS-IV total score than the placebo group. This reviewer reproduced and confirmed these results.

Secondary analysis:

1) Subscales of ADHD-RS-IV total score

The ADHD-RS-IV scale, where the primary endpoint was derived, consists of two subscales: Hyperactivity/Impulsivity and Inattention. Sponsor conducted the efficacy analyses for each subscale using the same ANCOVA approach as in the primary analysis. The results are displayed in Table 5 and Table 6. They are consistent with the primary analysis results (Table 4).

2) Response and remission

Based on Sponsor's response analysis and remission analyses, this reviewer found that both the SPD503 AM and PM groups showed numerical evidence of a higher efficacy than the placebo group. (The results are found in Table 2.1.10, Table 2.1.11, and Table 2.1.12 of Sponsor's CSR).

Table 5: Sponsor Secondary Analysis (ANCOVA – LOCF) for Hyperactivity/Impulsivity

Analysis (Endpoint – Analysis set)	Treatment Group	Sample Size (N=449)	Change from baseline in ADHD-RS-IV Total score			
			Mean (SD*)	LS mean	Placebo-adjusted difference** in LS mean (95% CI)	p-value
ANCOVA (LOCF - FAS)	SPD503 AM	149	-9.6 (6.82)	-9.6	-2.1 (-3.4, -0.7)	.002
	SPD503 PM	148	-9.9 (6.53)	-9.8	-2.3 (-3.6, -0.9)	< .001
	Placebo	152	-7.5 (6.01)	-7.6		

* *SD* denotes standard deviation.
** *LS mean* and *p-value* for *placebo-adjusted difference* were based on type III sum of squares from an ANVOVA model for the change from baseline with treatment group and psychostimulant type as fixed effects, and baseline value as a covariate. The reported p-values were based on Dunnett's multiple comparison procedure.
[Source: *Table 2.1.8.1* (page 294) and *Table 2.1.8.2* (page 295) of Sponsor's CSR]

Table 6: Sponsor Secondary Analysis (ANCOVA – LOCF) for Inattentiveness

Analysis (Endpoint – Analysis set)	Treatment Group	Sample Size (N=449)	Change from baseline in ADHD-RS-IV Total score			
			Mean (SD*)	LS mean	Placebo-adjusted difference** in LS mean (95% CI)	p-value
ANCOVA (LOCF - FAS)	SPD503 AM	149	-10.8 (7.24)	-10.7	-2.4 (-3.9, -0.9)	.002
	SPD503 PM	148	-11.1 (7.01)	-11.4	-3.1 (-4.6, -1.5)	< .001
	Placebo	152	-8.4 (6.64)	-8.3		

* *SD* denotes standard deviation.
** *LS mean* and *p-value* for *placebo-adjusted difference* were based on type III sum of squares from an ANVOVA model for the change from baseline with treatment group and psychostimulant type as fixed effects, and baseline value as a covariate. The reported p-values were based on Dunnett's multiple comparison procedure.
[Source: *Table 2.1.9.1* (page 300) and *Table 2.1.9.2* (page 301) of Sponsor's CSR]

Sensitivity analysis:

1) Repeated measure analysis via linear mixed model (MMRM)

Sponsor planned to conduct a sensitivity analysis using MMRM on the change from baseline in the ADHD-RS-IV total score. A mixed effects linear model for repeated measures (PROC MIXED) was used to analyze the observed change from baseline scores at all post-randomization on-treatment visits (Visits 3 - 10). The model included a covariate of baseline score, a factor of treatment, a stratification factor of psychostimulant type (categorized as amphetamine or methylphenidate), and a term for treatment-by-visit interaction. An unstructured covariance matrix was used with a random subject effect. The analysis result was based on the three dose maintenance visits (Visits 8, 9, and 10). The MMRM analysis Sponsor conducted was inappropriate, since averaged treatment effects over those three visits between treatment groups were adopted[12] This reviewer conducted his sensitivity analysis via the MMRM method. The corrected results are presented in Section *3.1.4.3 Reviewer's Assessments*.

[12] The results for repeated measures analyses presented in Table 2.1.1.2 of CSR (pages 201 and 202), i.e., the estimates of 'Difference in LS Mean', were reported as -4.4 (SPD503 AM vs. Placebo) and -5.0 (SPD503 PM vs. Placebo), computed over the three measurements from the dose-maintenance period.

2) Complete case analysis for the effect of early terminated subjects

Sponsor planned to explore the effect of withdrawals on the primary analysis by repeating the primary analysis on subjects who completed through Visit 10 (efficacy endpoint).
A completer is a subject who completed visit 10 efficacy assessment on or before the last non-tapering dose date. There were 382 subjects in Sponsor's Complete Case (CC) analysis set. Three subjects in this dataset recorded as early terminated at visit 10 were included as complete cases, as their efficacy assessments had been performed on the last non-tapering dose date (See also Table 2).
As can be seen in Table 7, Sponsor's ANCOVA results based on complete cases were similar to those of the ANCOVA LOCF analysis in the primary efficacy assessment.

Table 7: Sponsor Sensitivity Analysis (ANCOVA – Complete Cases)

Primary analysis (Endpoint – Analysis set)	Treatment Group	Sample Size (N=382)	Change from baseline in ADHD-RS-IV Total score			
			Mean (SD*)	LS mean (SE*)	Placebo-adjusted difference** in LS mean (95% CI)	p-value
ANCOVA (Complete Cases)	SPD503 AM	122	-22.3 (11.64)	-22.3 (1.01)	-5.4 (-8.5, -2.3)	.0003
	SPD503 PM	127	-22.0 (11.96)	-22.2 (0.99)	-5.3 (-8.4, -2.3)	.0003
	Placebo	133	-17.1 (11.79)	-16.9 (0.97)		

* *SD* denotes standard deviation, and *SE* standard error.
** *LS mean* and *p-value* for *placebo-adjusted difference* were based on type III sum of squares from an ANVOVA model for the change from baseline with treatment group and psychostimulant type as fixed effects, and baseline value as a covariate. The reported p-values were based on Dunnett's multiple comparison procedure.
[Source: *Table 2.1.1.1* (page 199) and *Table 2.1.1.2* (pages 201 and 202) of Sponsor's CSR]

Sponsor planned to perform analysis using the primary analysis method on discontinued subjects, but this reviewer was unable to find Sponsor's results in the submitted CSR. This reviewer conducted the sensitivity analysis. The placebo-adjusted differences in LS mean estimates were -3.0 (SPD503 AM vs. Placebo) and -6.9 (SPD503 PM vs. Placebo). The placebo-adjusted difference of SPD 503 AM (-3.0) is numerically smaller in magnitude than that in the primary analysis (-4.5), while placebo-adjusted difference of SPD 503 PM (-6.9) is numerically larger in magnitude than that in the primary analysis (-5.3). The primary analysis results can be found in Table 4.

Table 8: Endpoint LOCF Change from Baseline: Early Terminated Subjects

Treatment Group	Sample Size (N = 67)	Mean	SD*	Median
SPD503 AM	27	-11.9	14.28	-10.0
SPD503 PM	21	-14.9	13.49	-14.0
Placebo	19	-7.9	8.01	-5.0

* *SD* denotes standard deviation.
[Source: Reviewer's result]

This reviewer also obtained sample statistics (mean, standard deviation, median) of LOCF change from baseline in ADHD-RS-IV total score of each treatment group (Table 8). As reported in this table, the number of subjects of the early terminated subjects in each treatment group was small to give a reliable estimate of the mean. The difference in mean between each SPD503 treatment group and the Placebo group was not too far from the corresponding placebo-adjusted difference in LS mean estimates of the primary efficacy analysis. The observed discrepancies from the primary analysis result do not seem to suggest that the effect of early terminated subjects on the primary analysis raises concern in the interpretation of the primary efficacy analysis.

3.1.4.2.2 Secondary Efficacy Measures

Sponsor reported that analysis results on secondary efficacy measures (CGI-S, CGI-I, CGI-P, PGA, BSFQ, CPRS-R: L, PSQ) were supportive in the proposed efficacy claim, and consistent with the primary efficacy analysis result.

3.1.4.2.3 Sponsor's Conclusion on Efficacy

Sponsor summarized the efficacy assessments as follows[13]:

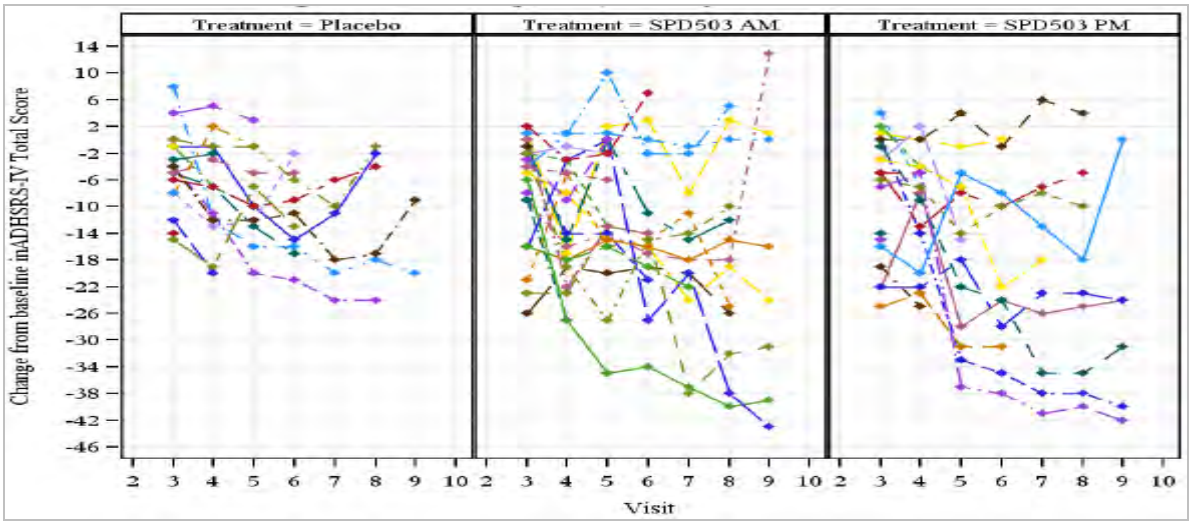
for subjects with ADHD who were suboptimal responders to a long-acting, oral psychostimulant, SPD503 in combination with a psychostimulant resulted in significant improvement as assessed by multiple measures compared to psychostimulant use alone (ie, placebo). The significant improvement was noted on measures of symptoms and global assessments, evaluated by both clinicians and parents and was shown on both morning and evening behaviors. Clinically meaningful improvement was shown for both morning and evening administration of SPD503.

3.1.4.3 Reviewer's Assessments

3.1.4.3.1 Subject Dropouts Profiles

Figure 2 displays subject-wise spaghetti plots of change from baseline score in the primary efficacy measure of 67 discontinued subjects over the dose-optimization and dose-maintenance phases.

Figure 2: Efficacy Profiles of Dropouts (Observed Change from Baseline Score)



[Source: Reviewer's result]

It does not appear that in the dropout subjects, there are systematic patterns in the profiles of change from baseline scores over both the dose-optimization (Visits 3 – 7) and dose-maintenance (Visits 8 – 10) phases, although there were relatively larger variations in both SPD503 AM and PM groups in later visits than in the placebo group. There appears to be no evidence suggesting that the efficacy profiles over time were associated with efficacy of the investigational drug.

[13] Refer to Sponsor's CSR (page 74)

3.1.4.3.2 Distributions of Primary Efficacy Measure at Baseline

Baseline Primary Efficacy Measure:

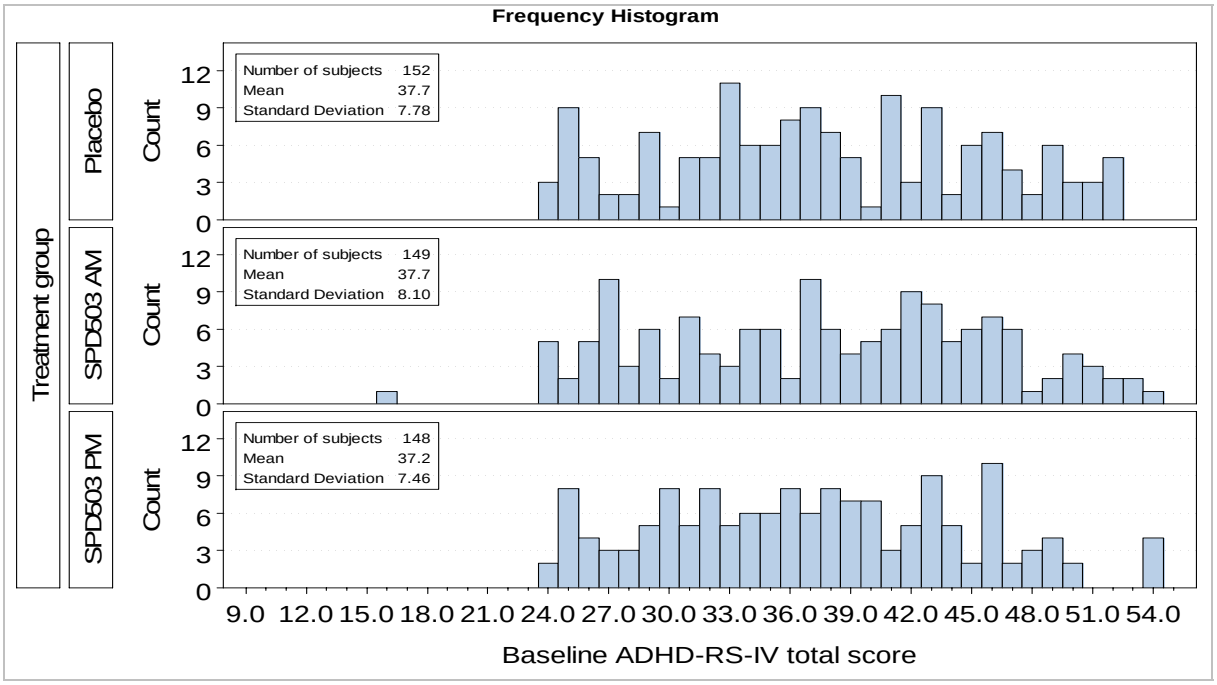
Table 9 provides mean, standard deviation and median of baseline ADHD-RS-IV total scores from each treatment group (Placebo, SPD503 AM and SPD503 PM). Figure 3 displays histograms of baseline ADHD-RS-IV total scores of the three treatment groups. Along with simple statistics given in Table 9, a visual inspection of the histograms suggests that the baseline ADHD-RS-IV total scores of the three groups were comparable.

Table 9: Baseline ADHD-RS-IV Total Score by Treatment Group

Treatment Group	Sample Size (N = 449)	Mean	SD*	Median
SPD503 AM	149	37.7	8.10	38.0
SPD503 PM	148	37.2	7.46	37.0
Placebo	152	37.7	7.78	37.0

* *SD* denotes standard deviation.
Note: There were 455 subjects who had baseline scores. Out of the 455subjects (FAS), there were 449 who had at least one post-baseline score. (Refer also to Table 2).
[Source: Reviewer's analysis]

Figure 3: Distribution of Baseline ADHDRS-IV Total Score by Treatment



[Source: Reviewer's analysis]

3.1.4.3.3 Primary Efficacy Endpoint

This reviewer conducted the primary efficacy analysis using Sponsor's analysis data set and confirmed the result based on the data set he created from the submitted tabulation (raw) data sets. The results were identical to those of Table 4. The psychostimulant type as a factor of the primary efficacy analysis model was not found statistically significant in the primary analysis. This reviewer also examined the effect of the interaction of treatment and baseline score and that of the study sites, by adding each of the

independent variables to the primary analysis model. Neither the interaction of treatment and baseline score nor the study site was found statistically significant.

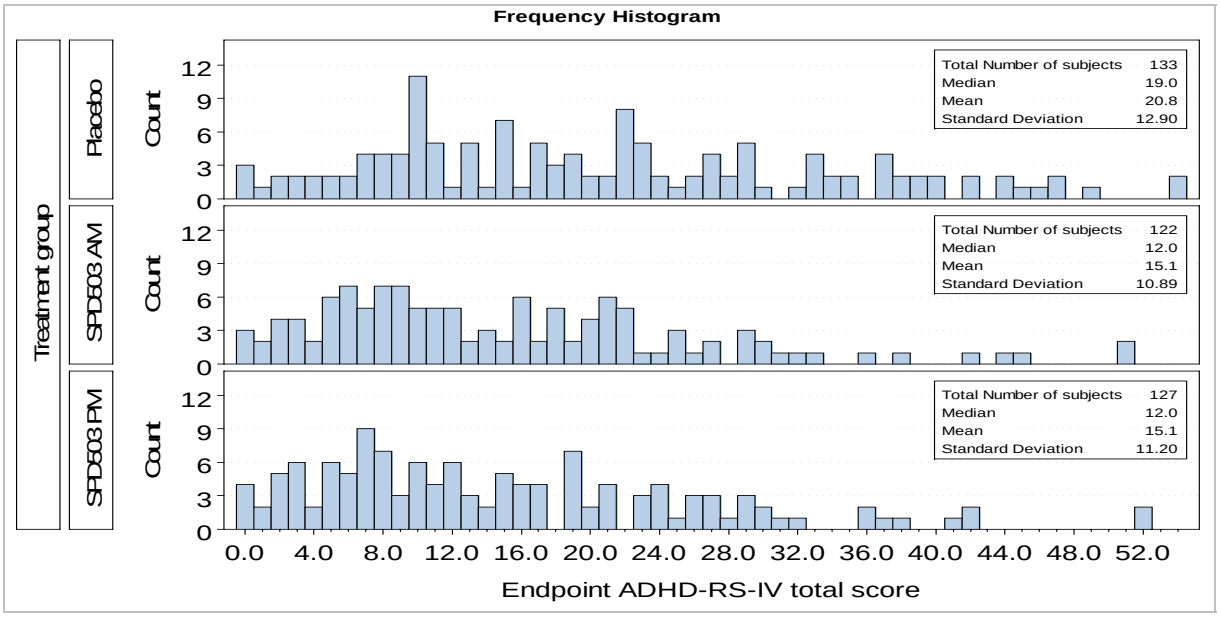
In addition, this reviewer examined observed endpoint primary efficacy measure, by obtaining the histogram of endpoint data (ADHD-RS-IV total score), endpoint sample statistics (mean, standard deviation, and median) and by drawing a figure of mean endpoint efficacy profiles (with confidence limits) over time. These are shown in Table 10, Figure 4 and Figure 5.

Table 10: ADHD-RS-IV Total Score and Change from Baseline Score at Endpoint (Visit 10) for Completers

Treatment Group	Observed efficacy measure at Endpoint	Sample Size (N=382)	Mean	SD*	Median
SPD503 AM	ADHD-RS-IV total score	122	15.1	10.89	12.0
	Change from baseline		-22.3	11.64	-23.0
SPD503 PM	ADHD-RS-IV total score	127	15.1	11.20	12.0
	Change from baseline		-22.0	11.96	-22.0
Placebo	ADHD-RS-IV total score	133	20.8	12.90	19.0
	Change from baseline		-17.1	11.79	-18.0

* *SD* denotes standard deviation.
Note: 382 subjects from the FAS were included in the primary analysis as completers through Visit 10.
[Source: Reviewer's analysis]

Figure 4: Distribution of Endpoint ADHDRS-IV Total Score by Treatment

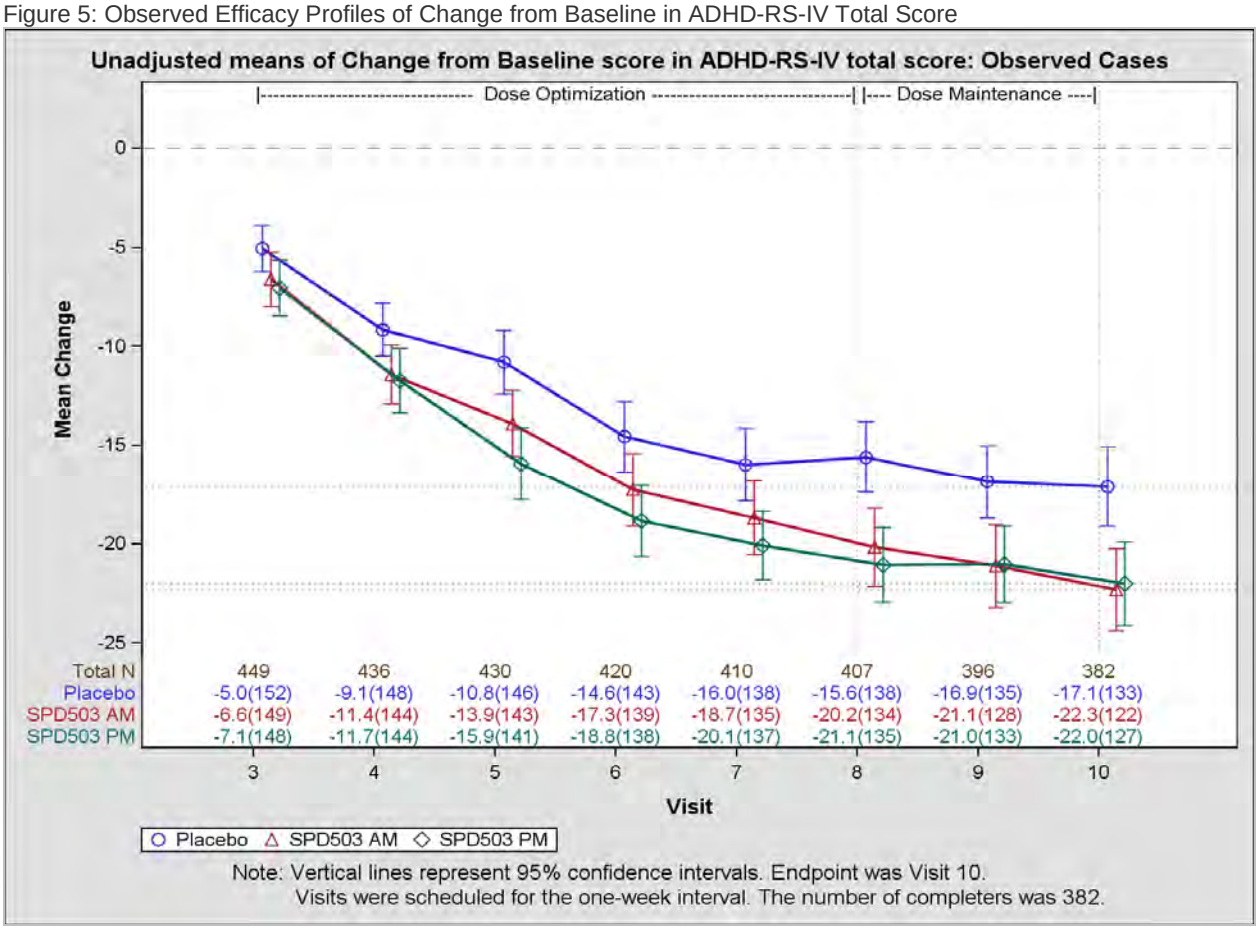


[Source: Reviewer's analysis]

Table 10 provides mean, standard deviation and median of endpoint ADHDRS-IV total scores from each treatment group (Placebo, SPD503 AM and SPD503 PM). Figure 4 displays histograms of endpoint ADHDRS-IV total scores of the three treatment groups. The sample statistics and the histogram suggest that the (unadjusted) means of ADHD-RS-IV total score and its change from baseline score are lower for both SPD503 groups than for the placebo group. The distributions of SPD503 groups appear to be similar, and relatively more slanted toward the left than that of the placebo.

Figure 5 displays the profiles of unadjusted means of observed change from baseline scores in ADHD-RS-IV total score over the dose-optimization and dose-maintenance phases. The profiles show an apparent separation of the unadjusted mean of each of the SPD503 treatment groups from that of the placebo over the three visits of the dose maintenance period. As an adjunctive therapy to psychostimulants, the SPD503 AM and PM groups seem to have improved more than the placebo group at the end of the 8-week efficacy evaluation period, suggesting that the means in the SPD503 treatment population may be better (lower in endpoint change from baseline in the primary efficacy measure) than in the placebo population.

Figure 5: Observed Efficacy Profiles of Change from Baseline in ADHD-RS-IV Total Score



[Source: Reviewer's analysis]

3.1.4.3.4 Sensitivity Analysis

This reviewer conducted a mixed model for repeated measures (MMRM) analysis as a sensitivity analysis, in order to check the robustness of the sponsor's efficacy analysis result based on the LOCF ANCOVA approach. The MMRM model included baseline score as a fixed covariate, treatment, psychostimulant type, visit and the treatment by visit interaction as fixed factors. The method of estimation was restricted maximum likelihood (REML). The within subject covariance matrix was unstructured. The degree of freedom of the denominator was approximated by the Kenward-Roger's method.

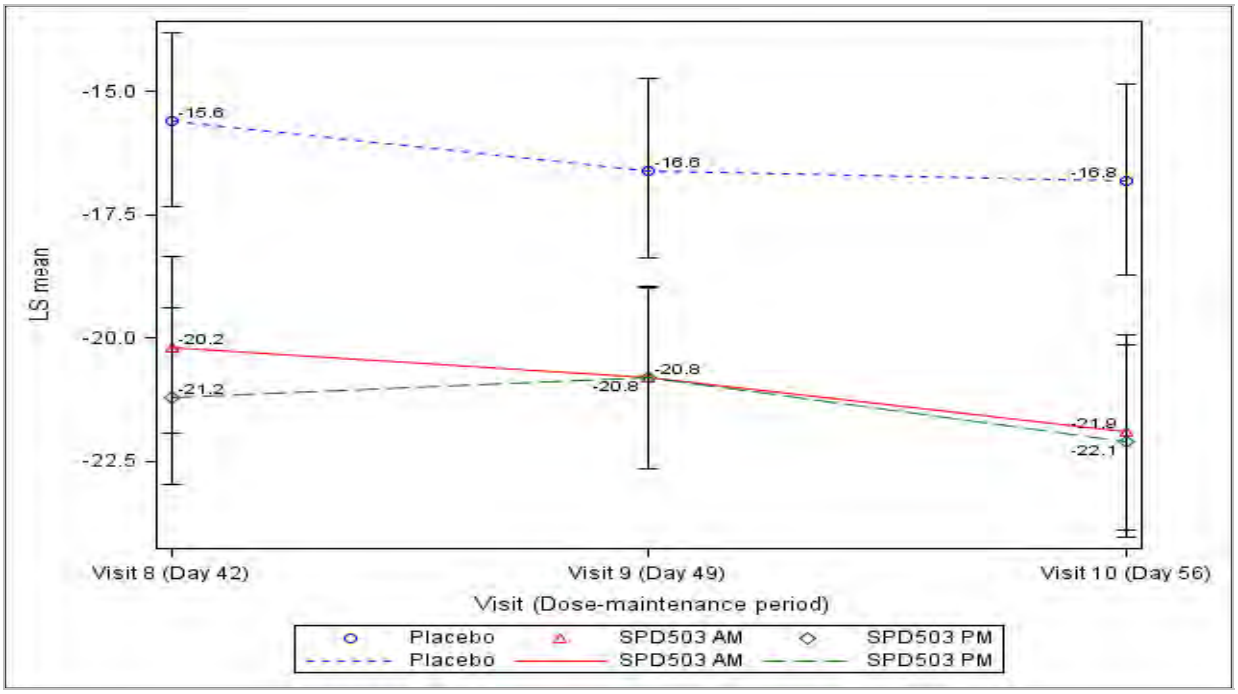
Table 11: Reviewer's MMRM Analysis for Dose-Maintenance Period

	SPD503 AM vs. Placebo		SPD503 PM vs. Placebo	
Dose-Maintenance visits	**Difference in LS mean**	**P-value***	**Difference in LS mean**	**P-value***
Visit 8 (42 days)	-4.6	.0004	-5.6	<.0001
Visit 9 (49 Days)	-4.2	.0013	-4.2	.0013
Visit 10 (56 Days)	-5.1	.0003	-5.3	.0002

* For each visit, Dunnett's multiplicity adjustment was applied for two comparisons (SPD503 AM vs. Placebo and SPD503 PM vs. Placebo).
Note: No adjustment was performed for multiple visits. [Source: Reviewer's analysis]

The analysis results and profiles of LS means of treatment groups are presented in Table 11 and Figure 6. The primary analysis results based on the LOCF ANCOVA analysis are supported by those from the MMRM analysis.

Figure 6: LS means and 95% Confidence Limits in Change from Baseline Score in ADHD-RS-IV Total Score by Treatment Group from Reviewer's MMRM analysis (Dose-Maintenance Period: Visits 8 – 10)



* Values in the plots are means. Vertical lines represent 95% confidence intervals.
[Source: Reviewer's analysis]

3.1.4.3.5 Effects of Study Sites

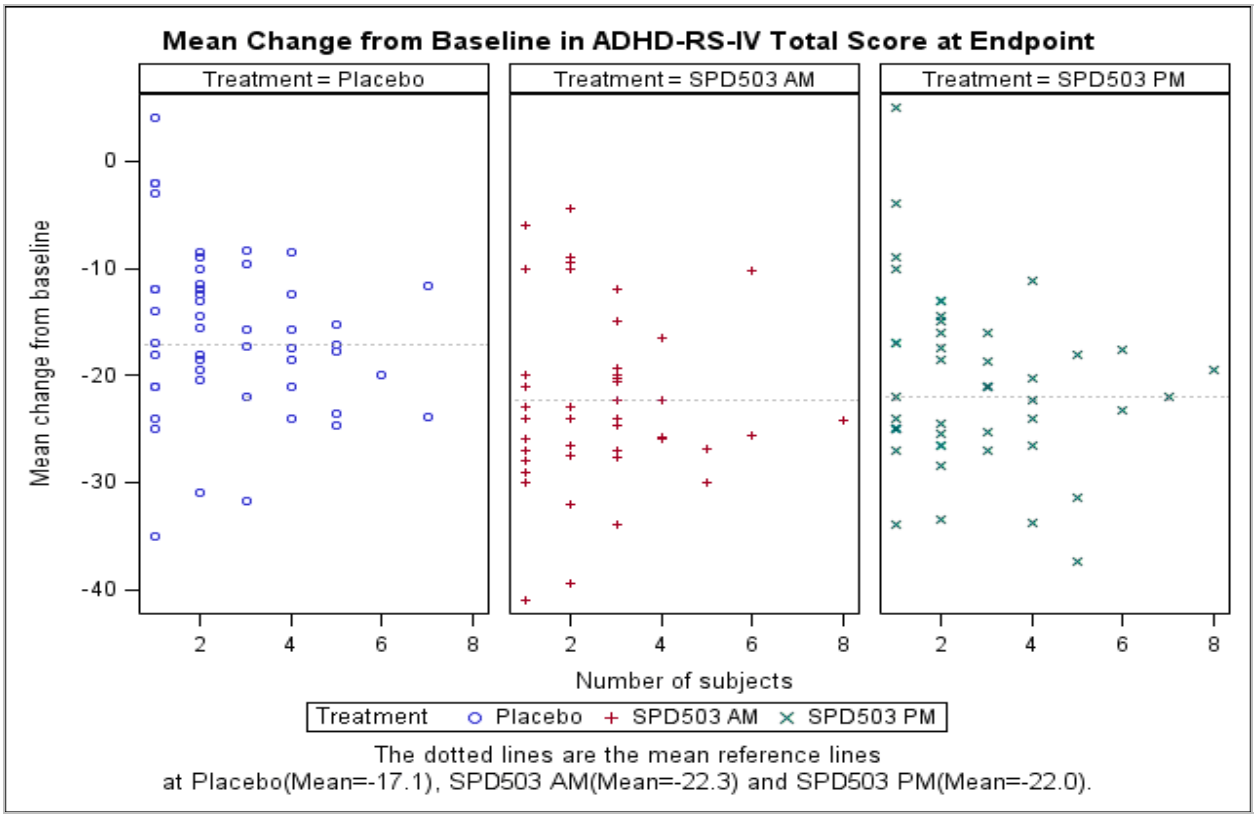
This reviewer plotted endpoint mean change from baseline scores in ADHD-RS-IV total score against the number of subjects in study sites, by treatment group (Figure 7).

It appears from the graphs that in all the treatment groups, the spread of scatter plots narrows around the horizontal mean reference line, as the number of site subjects increases. The trend of the placebo group was parallel to the horizontal mean reference line at -17.1. A downward trend relative to the respective mean reference lines (-22.3 and -22.0) may be observed in both SPD503 treatment groups, but it is not

steep at all. In conclusion, for each treatment group, there is no evidence suggesting that the number of subjects was influential.

At baseline, there were 58 study sites in this study and the size of subjects varied from 1 to 23 among study sites. Two study sites, site#40 (20 subjects) and site #07 (18 subjects), which were the second and third largest sites, were selected for the FDA site inspection by the Division of Scientific Inspection (DSI). The DSI recommended that three subjects from site #07 be excluded from the primary efficacy analysis. The reviewer checked on the efficacy data of these subjects, concluding that there was almost no impact of these subjects on the analysis results.

Figure 7: Effect of Site Number of Subjects by Treatment Group



* Means and site sizes are based on observed values of Endpoint (Visit 10).
** The horizontal dotted lines are mean reference lines for each treatment group.
Note: 57 sites had ADHD-RS-IV total score recorded at Visit 10 (Endpoint).
[Source: Reviewer's analysis]

3.2 Evaluation of Safety

(The evaluation of safety is deferred to the clinical team.)

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

In this section all the subgroup analyses were exploratory, and conducted for the purpose of assessing consistency across subgroups with regard to the primary analysis result. The same LOCF ANCOVA model as in the primary efficacy analysis was used for the subgroup exploratory analyses for gender, race, age group, and psychostimulant type. The mean and standard deviation of the observed primary efficacy variable at baseline and at endpoint (Visit 10) were also obtained for each subgroup. The results tabulated in the tables appearing in the following sections will be compared with the overall results. The overall results, based on Table 4 and Table 9, are provided in Table 12 for the reader's convenience.

Table 12: Overall Efficacy Evidence

Overall	ADHD-RS-IV Total score							
	Observed**						LOCF	
	SPD 503 AM		SPD 503 PM		Placebo		SPD 503 AM vs. Placebo	SPD 503 PM vs. Placebo
	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)	LS mean difference	
Baseline	149	37.7 (8.10)	148	37.2 (7.46)	152	37.7 (7.78)	-4.5	-5.3
Mean change	122	-22.3 (11.64)	127	-22.0 (11.96)	133	-17.1 (11.79)		

* *Mean change* is mean change from baseline, and *LS mean difference* is least square mean difference in change from baseline between the treatment and placebo. *SD* denotes standard deviation, *N* sample size.
** Complete cases through Visit 10 for change from baseline for *Mean change*
[Source: Table 4 and Table 9 of this review]

4.1 Gender, Race and Age

This reviewer compared the results of Table 13, Table 14 and Table 15 to those of Table 12. In all the subgroup analyses for gender, race and age group, baseline ADHD-RS-IV total score and change from baseline score of each treatment group were similar to those of the overall results, except that the Placebo group had more variation across subgroups. This variation seems to have reflected on the variation in the LS mean estimates of the Placebo group. For instance, in the comparison of the SPD503 PM and the Placebo groups for non-white subgroup, the difference in the LS mean estimates was relatively smaller in magnitude, i.e., -2.1, which seems to be a result of a higher efficacy response (lower observed mean of -20.7) of non-white subgroup in the Placebo group.
This reviewer concluded that the differences in the LS mean estimate of change from baseline score of each SPD503 treatment group in comparison with the placebo group were consistent with the primary efficacy result. A higher variation in the LS means of subgroups of the Placebo group than in the SPD treatment groups may be reasonable, and its magnitude did not effect the overall conclusion, since the variation in the differences in the LS means of subgroups did not appear to be surprisingly large.

4.1.1 Gender

As can be seen in Table 13, the difference in the LS means of the female subgroup was smaller in magnitude than the overall result in both comparisons (SPD503 AM vs. Placebo and SPD503 PM vs. Placebo), -3.4 and -4.4 respectively. This seems to be due to the fact that the observed mean in change from baseline score (-18.3) was relatively larger in magnitude than in the overall result (-17.1). The observed mean change estimates of the SPD503 treatment groups were very close to the overall result.

Table 13: Gender Subgroup Analysis

Gender		ADHD-RS-IV Total score							
		Observed**						LOCF	
		SPD 503 AM		SPD 503 PM		Placebo		SPD 503 AM vs. Placebo	SPD 503 PM vs. Placebo
		N	Mean (SD)	N	Mean (SD)	N	Mean (SD)	LS mean difference	
Female	Baseline	42	37.9 (8.48)	46	35.7 (7.52)	41	37.9 (7.10)	-3.4	-4.4
	Mean change	36	-22.8 (12.97)	38	-22.3 (11.93)	35	-18.3 (10.98)		
Male	Baseline	108	37.5 (8.03)	106	37.6 (7.67)	112	37.6 (8.01)	-4.6	-5.5
	Mean change	86	-22.1 (11.12)	89	-21.9 (12.04)	98	-16.7 (12.09)		

* *Mean change* is mean change from baseline, and *LS mean difference* is least square mean difference in change from baseline between the treatment and placebo. *SD* denotes standard deviation, *N* sample size.
** Complete cases through Visit 10 for change from baseline for *Mean change*
[Source: *Table 2.1.6.1* and *2.1.6.2* of CSR (pages 267 – 278)]

4.1.2 Race

As can be seen in Table 14, the observed mean of change from baseline score in the Placebo non-white subgroup, -20.7, was larger in magnitude than in the overall result, -17.1. This reflected upon a smaller difference, -2.1, in the LS mean of the SPD503 PM group versus the Placebo group. Similarly, the observed mean of change from baseline score the Placebo white subgroup, -15.3, was smaller in magnitude than in the overall result. This reflected upon a larger difference, -6.7, in the LS mean of the SPD503 PM group versus the Placebo group when compared to the overall result (-5.3).
We also notice that the observed mean of change from baseline score in the SPD503 non-white subgroup, -25.7, was relatively larger in magnitude than in the overall result, -22.3.

Table 14: Race Subgroup Analysis

Race		ADHD-RS-IV Total score							
		Observed**						LOCF	
		SPD 503 AM		SPD 503 PM		Placebo		SPD 503 AM vs. Placebo	SPD 503 PM vs. Placebo
		N	Mean (SD)	N	Mean (SD)	N	Mean (SD)	LS mean difference	
White	Baseline	104	37.6 (7.86)	102	36.6 (6.81)	102	37.6 (7.61)	-4.8	-6.7
	Mean change	87	-21.0 (10.87)	86	-21.7 (11.15)	89	-15.3 (11.15)		
Non-white	Baseline	46	37.7 (8.81)	50	37.9 (9.16)	51	37.8 (8.10)	-4.3	-2.1
	Mean change	35	-25.7 (12.95)	41	-22.7 (13.62)	44	-20.7 (12.35)		

* *Mean change* is mean change from baseline, and *LS mean difference* is least square mean difference in change from baseline between the treatment and placebo. *SD* denotes standard deviation, *N* sample size.
** Complete cases through Visit 10 for change from baseline for *Mean change*
Note: Black or African American were about 68% of non-white subgroup,
[Source: *Table 2.1.7.1* and *2.1.7.2* of CSR (pages 279 – 290)]

4.1.3 Age

As can be seen in Table 15, the difference in the LS means of the 13-17 year-old Age subgroup was larger in magnitude than the overall result in both comparisons (SPD503 AM vs. Placebo and SPD503 PM vs. Placebo), -8.2 and -6.3 respectively. This seems to be caused mainly by the fact that an estimate of observed mean change from baseline score in the Placebo 13-17 year-old Age subgroup, -12.1, was smaller in magnitude than that of the overall result, -17.1.

Table 15: Age Subgroup Analysis

Age group		ADHD-RS-IV Total score							
		Observed**						LOCF	
		SPD 503 AM		SPD 503 PM		Placebo		SPD 503 AM vs. Placebo	SPD 503 PM vs. Placebo
		N	Mean (SD)	N	Mean (SD)	N	Mean (SD)	LS mean difference	
6-12 year-old	Baseline	114	38.2 (8.10)	124	37.6 (7.64)	123	38.6 (7.46)	-3.6	-5.1
	Mean change	93	-22.3 (12.12)	104	-22.3 (12.42)	105	-18.4 (12.05)		
13-17 year-old	Baseline	36	35.5 (7.98)	28	34.4 (7.26)	30	34.1 (8.01)	-8.2	-6.3
	Mean change	29	-22.4 (10.16)	23	-20.5 (9.72)	28	-12.1 (9.32)		

* *Mean change* is mean change from baseline, and *LS mean difference* is least square mean difference in change from baseline between the treatment and placebo. *SD* denotes standard deviation, *N* sample size.
** Complete cases through Visit 10 for change from baseline for *Mean change*
[Source: *Table 2.1.4.1* and *2.1.4.2* of CSR (pages 243 – 254)]

4.2 Other Special/Subgroup Populations

4.2.1 Psychostimulant type

As can be seen in Table 16, in both psychostimulant-type groups, baseline ADHD-RS-IV total score and change from baseline score of each treatment group were similar to those of the overall results (Table 12). The difference in the LS mean estimates of change from baseline score of each SPD503 treatment group in comparison with the placebo group was consistent with the primary efficacy result (Table 12).

Table 16: Psychostimulant-type Subgroup Analysis

Psychostimulant-type		ADHD-RS-IV Total score							
		Observed**						LOCF	
		SPD 503 AM		SPD 503 PM		Placebo		SPD 503 AM vs. Placebo	SPD 503 PM vs. Placebo
		N	Mean (SD)	N	Mean (SD)	N	Mean (SD)	LS mean difference	
Amphetamine	Baseline	69	36.8 (7.87)	74	36.5 (7.32)	72	38.1 (7.40)	-3.5	-5.1
	Mean change	51	-21.6 (9.61)	62	-21.4 (11.42)	63	-17.3 (12.61)		
Methylphenidate	Baseline	81	38.2 (8.34)	78	37.6 (7.97)	81	37.4 (8.08)	-5.2	-5.4
	Mean change	71	-22.8 (12.95)	65	-22.6 (12.50)	70	-16.9 (11.09)		

* *Mean change* is mean change from baseline, and *LS mean difference* is least square mean difference in change from baseline between the treatment and placebo. *SD* denotes standard deviation, *N* sample size.
** Complete cases through Visit 10 for change from baseline for *Mean change*
[Source: *Table 2.1.5.1* and *2.1.5.2* of CSR (pages 255 – 266)]

5. SUMMARY AND CONCLUSIONS

5.1 Statistical Issues and Collective Evidence

The reviewer found no major statistical issues regarding the efficacy conclusion of the study. First of all, the primary analysis result reported in Sponsor's clinical study report (CSR) was confirmed and verified. Secondly, this reviewer conducted pre-specified sensitivity analyses and secondary analyses of the primary efficacy variable, and confirmed Sponsor's results when available. The results were consistent with the primary analysis result. Thirdly, subgroup analysis results suggested little deviation from the primary analysis result.

The dropout rate, calculated as the number of subjects (67 subjects) who early terminated prior to the efficacy endpoint divided by that of randomized subjects (455 subjects), was about 15%. This reviewer concluded that there was no evidence suggesting that the dropouts and missing data impacted on the primary efficacy analysis to the extent that the study result should be questioned.

The reviewer found that Sponsor's study analyses and clinical study report (CSR) had some quality issues. As an example, Sponsor neither correctly performed the pre-specified sensitivity analysis (based on the MMRM method), nor discussed the analysis results in the CSR. In addition to this deficiency, the CSR did not sufficiently address the issue of dropouts and missing data. It lacked examinations of the underlying assumptions the adopted statistical method relies on.

5.2 Conclusions and Recommendations

The statistical evidence based on Study SPD503-313 supports Sponsor's claim that Intuniv™ is efficacious as an adjunctive therapy to a long-acting oral psychostimulant in the treatment of children and adolescents aged 6-17 years with a diagnosis of Attention-Deficit Hyperactivity Disorder.

SIGNATURES/DISTRIBUTION LIST

Primary Statistical Reviewer: Eiji Ishida, M.S.
Date: December 21, 2010

Concurring Reviewer(s): Peiling Yang, Ph.D., Kooros Mahjoob, Ph.D.

Statistical Team Leader: Peiling Yang, Ph.D.

Biometrics Division Deputy Director: Kooros Mahjoob, Ph.D

cc:
HFD-130/Dr. S. Chang
HFD-130/Dr. T. Laughren
HFD-130/Dr. M. Mathis
HFD-130/Dr. S. Borges
HFD-130/Dr. N. A.Khin
HFD-710/Dr. P. Yang
HFD-710/Dr. K. Mahjoob
HFD-710/Dr. J. Hung
HFD-700/Ms. L. Patrician

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

/s/

EIJI ISHIDA
12/23/2010

PEILING YANG
12/23/2010

KOOROS MAHJOOB
12/28/2010
Review was discussed with primary and secondary reviewer. My views/comments are incorporated in this version and I concur with it.