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:	NDA209905/S000
Drug Name:	Amphetamine Sulfate (AR11) Immediate Release (IR) Orally Disintegrating Tablets (ODT), 5, 10, 15, 20, _(b) (4)_ mg
Indication(s):	Treatment of Attention Deficit Hyperactivity Disorder (ADHD) in Children (Ages 6 _(b) (4)_)
Applicant:	Arbor Pharmaceuticals, LLC
Date(s):	Date of Document: March 30, 2018 PDUFA Due Date: January 30, 2019
Review Priority:	Standard
Biometrics Division:	Biometrics I, HFD-710
Statistical Reviewer:	Ququan Liu, M.D., M.S.
Concurring Reviewers:	Peiling Yang, Ph.D. H.M. James Hung, Ph.D.
Medical Division:	Division of Psychiatry Products, HFD-130
Clinical Team:	John Umhau, M.D. Bernard Fischer, M.D.
Project Manager:	Keith Kiedrow, Pharm.D.
Keywords:	Cross-Over Design, SKAMP-Combined Score, ADHD, Amphetamine Sulfate (AR11)

Table of Contents

List of Tables

List of Figures

1. EXECUTIVE SUMMARY

This is a 505(b)(2) application that includes a study (AR11.001) to evaluate the safety and efficacy of AR11 in a rigorous classroom setting to establish the onset and duration of effect in pediatric subjects 6 to 12 years old diagnosed with Attention-Deficit Hyperactivity Disorder (ADHD). The study results support the efficacy of AR11 on the reduction of signs and symptoms of ADHD evaluated by Swanson, Kotkin, Agler, M-Flynn, Pelham Rating Scale (SKAMP) scores. The primary efficacy endpoint of the SKAMP-Combined scores at 2 hours post-dose is statistically significantly lower (i.e., better) in the AR11 treatment compared to the placebo (p<0.0001). The onset of the AR11 effect seems to start at the first measurement at 0.75 hours post-dose and last through the final measurement at 10 hours post-dose.

2. INTRODUCTION

2.1   Overview

ADHD is one of the most common neurobehavioral disorders in children, with a prevalence rate among school-age children of about 6%‑7%.

Controlled clinical trials have consistently demonstrated that stimulants substantially reduce the characteristic symptoms and impairment of patients with ADHD. The stimulant medications that have been commonly used in the United States to treat ADHD include various preparations of amphetamine and methylphenidate. The study drug used in the current study, AR11, is a racemic mixture of d-amphetamine and l-amphetamine.

This application relies on FDA's previous findings of safety and effectiveness for the listed drug Amphetamine Sulfate tablets via cross reference to NDA 83901 sponsored by Lannett. Amphetamine Sulfate IR ODT (also referred to as Evekeo ODT™, the conditionally accepted proprietary name) bridges to Arbor's product Evekeo®, the reference drug.

2.2   Data Sources

The applicant's SAS datasets were stored in the directory of \\CDSESUB1\evsprod\NDA209905\0000 the Center's electronic document room.

3. STATISTICAL EVALUATION

3.1   Data and Analysis Quality

A consistent result of the primary efficacy analysis can be generated from both raw and derived data.

3.2   Evaluation of Efficacy

3.2.1   Study Design and Endpoints

Primary Study Objective: The primary objective of this study was to establish that an optimal dose of AR11 would result in a significant reduction in signs and symptoms of ADHD compared with placebo treatment in pediatric subjects 6 to 12 years old with ADHD.

Study Design: The study was a randomized, double-blind, placebo-controlled, multi-center, two-arm two-period, crossover design trial in the U.S to investigate the safety and efficacy of AR11 in the treatment of ADHD in children from 6 to 12 years of age. The study consisted of:

- A 30-day screening period and baseline evaluation
- Open-label treatment with AR11 for 8 weeks for dose optimization, and an additional 2 weeks of double-blind treatment (1 week of AR11 with no dose adjustments and 1 week of placebo)
- Double-blind randomization to an assigned treatment sequence, A or B (AR11/placebo or placebo/AR11, respectively) occurred at Visit 10
- An abbreviated practice Laboratory Classroom Day was held during Visit 10
- Double-blind study drug was dispensed at Visit 10
- Laboratory classroom testing and final visit assessments were conducted on the second test Laboratory Classroom Day (Visit 12)
- The final safety assessments were completed during the post-withdrawal follow-up phone call approximately 1 week after Visit 12

Reviewer's Note: It's noted that there was no wash-out period between the two treatment periods.

Primary efficacy endpoint: The primary efficacy endpoint is the SKAMP-Combined scores at 2 hours post-dose as measured during the Laboratory Classroom Days (Visit 11 and Visit 12). The SKAMP is a 13-item, independent-observer rating of subject impairment of classroom observed behaviors. Each item is rated on a 7-point impairment scale (0 = normal to 6 = maximal impairment). The combined scores and subscale scores for the SKAMP are obtained by summing the values of corresponding items in the assessment.

Secondary efficacy endpoints:
--Key secondary efficacy endpoints: determined by the SKAMP-Combined scores at 0.75, 4, 6, 8, and 10 hours postdose on each Laboratory Classroom Day (Visits 11 and 12):

- onset of clinical effect
- duration of clinical effect

--Other secondary efficacy endpoints:

- Model-adjusted average of SKAMP-Combined scores over the entire Laboratory Classroom Day
- SKAMP-Combined scores measured at 0.75, 4, 6, 8, and 10 hours post-dose
- Model-adjusted average of SKAMP-Attention and -Deportment scores over the entire Laboratory Classroom Day
- SKAMP-Attention and -Deportment scores measured at 0.75, 2, 4, 6, 8, and 10 hours post-dose
- PERMP scores measured at 0.75, 2, 4, 6, 8, and 10 hours post-dose

3.2.2 Statistical Methodologies

Determination of sample size: Assuming an effect size of 0.45 between AR11 and placebo, an intra-class correlation of 0.2, the study with a sample size of 85 subjects randomized to double-blind treatment, will have 90% power at the significance level of 0.05 (2-sided) using a paired t-test.

Primary efficacy analysis: A mixed-model, repeated-measures analyses was performed on the ITT population (defined as all randomized subjects who received at least 1 dose of double-blind study drug and had at least 1 post-dose assessment of the primary efficacy variable), with study center, period (1 and 2), sequence (AR11/placebo and placebo/AR11), and time point (0.75, 2, 4, 6, 8, and 10 hours post-dose)-by-treatment (AR11 and placebo) interaction as fixed effects and subject's intercept as a random effect. An unstructured covariance matrix was used to model within-subject/within-treatment variability. The treatment difference was estimated using least-squares (LS) means from the mixed-effects repeated-measures model. The treatment comparison was conducted as a 2-sided test at the 5% level of significance.

Secondary efficacy analyses: The same mixed-model repeated-measures analysis performed on the primary efficacy variable was performed on the ITT population for:
- Onset of clinical effect defined as the earliest post-dose time point at which the difference between the 2 treatments was statistically significant ($p < 0.05$).
- Duration of clinical effect defined as the difference between the onset time and the latest consecutive time point at which the difference between the 2 treatments was still statistically significant ($p < 0.05$).

Additional secondary efficacy analyses: The primary efficacy analysis was repeated on the clinically evaluable population (defined as all ITT subjects who received the morning dose of double-blind study drug, as determined during the dose optimization period, at both test laboratory classroom sessions; completed all laboratory classroom tests; did not miss more than 2 days of therapy during the double-blind treatment period; and did not use prohibited medication during the double-blind treatment period) for:
- SKAMP-Combined scores at 0.75, 4, 6, 8, and 10 hours post-dose,
- SKAMP subscale scores,
- Model-adjusted average of SKAMP-Combined and SKAMP subscale scores over the entire Laboratory Classroom Day, and
- PERMP scores.

Handling of dropouts/missing data:
- Missing individual items in the SKAMP scale:
  - If 3 or more individual items in the SKAMP have missing or invalid data, the SKAMP-Combined score will be set to missing
  - If 1 or 2 individual items in the SKAMP are missing or invalid, the values for the missing individual items will be imputed using the mean of the non-missing individual items for the particular patient at that visit
  - If any item within a SKAMP subscale is missing or invalid, the entire subscale score will be set to missing
- Missing SKAMP composite scores at individual time points:
  - Primary Analysis: No imputation of missing SKAMP–Combined scores will be done for the primary efficacy analysis. The mixed model repeated measures methods used to analyze the primary endpoint has been designed to utilize all available data and provides valid estimates under the assumption of data which are missing completely at random or missing at random.
  - Sensitivity Analyses: Analyses using two imputation methods may be

conducted—a) LOCF; and b) multiple imputation.

3.2.3   Patient Disposition, Demographic and Baseline Characteristics

Patient disposition: A total of 107 subjects were enrolled and 97 were randomized, 50 in the sequence of placebo/AR11 and 47 in the sequence of AR11/placebo. Only 2 (2.1%) subjects discontinued the study, 1 subject from each treatment sequence (Table 1).

Table 1   Subject Disposition

	Not Randomized N=10 n (%)	Placebo/AR11 N=50 n (%)	AR11/Placebo N=47 n (%)	Total N=107 n (%)
Enrolled	10 (100.0)	50 (100.0)	47 (100.0)	107 (100.0)
Randomized		50 (100.0)	47 (100.0)	97 (90.7)
Study Completion				
Completed		49 (98.0)	46 (97.9)	95 (88.8)
Discontinued	10 (100.0)	1 (2.0)	1 (2.1)	12 (11.2)
Reason for Premature Discontinuation[a]				
Adverse event	6 (60.0)	0	0	6 (50.0)
Withdrawal of consent	3 (30.0)	1 (100.0)	0	4 (33.3)
Lost to follow-up	0	0	1 (100.0)	1 (8.3)
Unable to achieve stable dose	0	0	0	0
Other	1 (10.0)	0	0	1 (8.3)

(Source: Applicant's Table 4, confirmed by the reviewer's analysis)

Patient demographic and baseline characteristics: Demographic and baseline characteristics ware similar in both the treatment sequence groups. The majority of subjects was males (61%), white (60%) and had combined inattentive and hyperactive/impulsive ADHD (81.4%). The age ranged from 6 to 12 years with a mean age of 9.6 years.   All 97 subjects were in the ADHD 90th percentile at baseline. Mean height at screening was 55.1 inches and mean weight at baseline was 84.6 lbs. (Table 2).

Table 2   Summary of Demographics and Baseline Characteristics, ITT

Characteristics	Placebo/AR11 N=50	AR11/Placebo N=47	Total N=97
Gender – n (%)			
Male	30 (60.0)	29 (61.7)	59 (60.8)
Female	20 (40.0)	18 (38.3)	38 (39.2)
Age (years)			
n	50	47	97
Mean	9.6	9.6	9.6
SD	1.78	1.97	1.86
Age Categories – n (%)			
6 – 7 Years	9 (18.0)	7 (14.9)	16 (16.5)
8 – 10 Years	22 (44.0)	22 (46.8)	44 (45.4)
11 – 12 Years	19 (38.8)	18 (38.3)	37 (38.1)
Race – n (%)			
White	28 (56.0)	30 (63.8)	58 (59.8)
Black/African American	20 (40.0)	13 (27.7)	33 (34.0)
Asian	0	1 (2.1)	1 (1.0)
Native Hawaiian/Pacific Islander	0	1 (2.1)	1 (1.0)
Other	2 (4.0)	2 (4.3)	4 (4.1)
ADHD Type – n (%)			
Inattentive	9 (18.0)	9 (19.1)	18 (18.6)
Hyperactive/Impulsive	0	0	0
Combined	41 (82.0)	38 (80.9)	79 (81.4)
ADHD-RS 90th Percentile at Baseline – n (%)			
Yes	50 (100.0)	47 (100.0)	95 (97.9)
No	0	0	0
Height (in) (Screening)			
n	50	47	97
Mean	55.49	54.80	55.16
SD	5.428	5.406	5.400
Weight (lbs) (Baseline)			
n	50	47	97
Mean	86.85	82.20	84.60
SD	31.985	30.219	31.068

Abbreviations: ADHD = Attention Deficit Hyperactivity Disorder; ADHD-RS = Attention Deficit Hyperactivity Disorder Rating Scale; ITT = intent-to-treat; SD = standard deviation

(Source: Applicant's Table 6, confirmed by the reviewer's analysis)

3.2.4   Results and Conclusion

Applicant's Primary Efficacy Result: The primary analysis result shows that SKAMP-Combined Scores at 2 hours post-dose was statistically significantly lower (i.e., better) in the AR11 treatment compared to the placebo (a LS mean difference of -7.9, 95% CI (-10.1, -5.6), p<0.0001).   The mean change from pre-dose at 2 hours post-dose in SKAMP-Combined scores was also numerically lower in the AR11 treatment compared to the placebo (a LS mean difference of -10.5, 95% CI (-13.2, -7.8)) (Table 3).

Table 3   Primary Analysis of SKAMP-Combined Scores at 2 Hours Post-Dose: ITT

Statistic	AR11 N = 97	Placebo N = 97	Treatment Difference (AR11 – Placebo) N = 97
n	95	97	
Mean (SD)	10.0 (8.24)	17.8 (11.94)	
Median (Q1, Q3)	8.0 (5, 12)	15.0 (8, 24)	
Range (min, max)	(0, 48)	(1, 52)	
LS mean (SE)	10.3 (1.09)	18.1 (1.09)	−7.9 (1.14)
95% CI	(8.1, 12.4)	(16.0, 20.2)	−10.1, −5.6
p-value			< 0.0001

Abbreviations: CI = confidence interval; ITT = intent-to-treat; LS = least-squares; max = maximum; min = minimum; Q = quartile; SD = standard deviation; SE = standard error; SKAMP = Swanson, Kotkin, Agler, M-Flynn, and Pelham Rating Scale.

(Source: The applicant's Table 7, confirmed by the reviewer's analysis)

Applicant's Key Secondary Efficacy Results:

1) Onset and duration: The onset of AR11 effect was seen at 0.75 hours post-dose (a LS mean difference of -5.5, 95% CI (-7.5, -3.5), p<0.0001).   The duration of AR11 effect lasted through the final measurement at 10 hours post-dose (a LS mean difference of -4.3, 95% CI (-6.4, -2.3), p<0.0001)). The treatment differences are statistically significant at all post-dose timepoints (Table 4 & Figure 1).

Table 4   SKAMP-Combined Scores Post-Dose Over Time: ITT

Time point	AR11 LS Mean (SE) N=97	Placebo LS Mean (SE) N=97	Treatment Difference (AR11 – Placebo)		
			LS Mean (SE) N=97	p-value	95% CI
0.75 hours postdose	12.1 (1.03)	17.6 (1.02)	−5.5 (1.01)	<0.0001	−7.5, −3.5
2 hours postdose[a]	10.3 (1.09)	18.1 (1.09)	−7.9 (1.14)	<0.0001	−10.1, −5.6
4 hours postdose	11.9 (1.05)	20.2 (1.05)	−8.3 (1.06)	<0.0001	−10.4, −6.2
6 hours postdose	14.8 (1.00)	20.5 (1.00)	−5.7 (0.96)	<0.0001	−7.6, −3.8
8 hours postdose	16.3 (1.01)	22.3 (1.00)	−6.1 (0.96)	<0.0001	−8.0, −4.2
10 hours postdose	16.8 (1.05)	21.1 (1.04)	−4.3 (1.05)	<0.0001	−6.4, −2.3
Average over all postdose time points	13.7 (0.89)	20.0 (0.88)	−6.3 (0.69)	<0.0001	−7.6, −4.9

Abbreviations: CI = confidence interval; LS = least-squares; SE = standard error; SKAMP = Swanson, Kotkin, Agler, M-Flynn, and Pelham Rating Scale.

(Source: The applicant's Table 8, confirmed by the reviewer's analysis)

Figure 1   SKAMP-Combined Scores Over Time (LS Mean ± SE) by Treatment: ITT
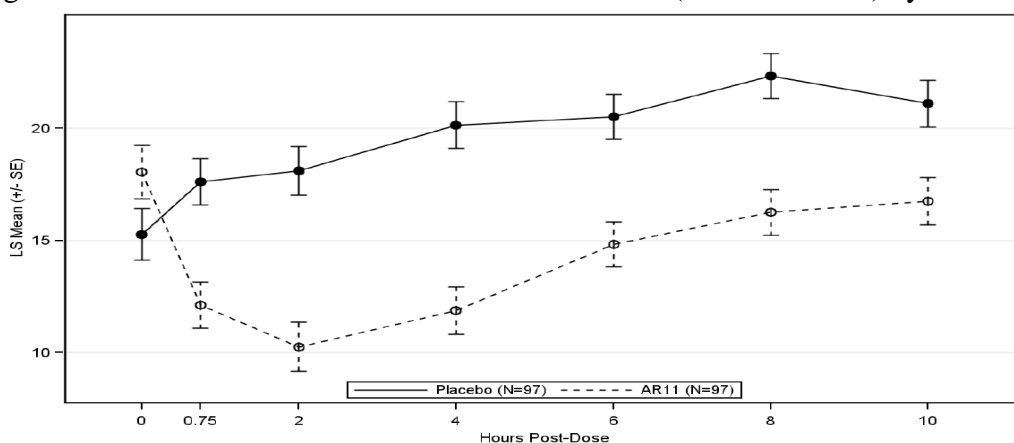
Abbreviations: ITT = intent to treat; LS = least-squares; SE = standard error; SKAMP = Swanson, Kotkin, Agler, M-Flynn, and Pelham Rating Scale

(Source: The applicant's Figure 1, confirmed by the reviewer's analysis)

Reviewer's Note: The scores at hour zero are raw means of pre-dose scores in Figure 1.

Reviewer's Results:

There was no wash-out period between the two treatment periods and a concern that the estimate of a true treatment effect is biased if differential carryover effects are present.   This reviewer performed the following analyses to verify the primary efficacy result:

1) Based on this reviewer's analysis, there was no evidence of the substantial treatment by period interaction (i.e., sequence effect) (p=0.1658).

2) The plots in Figure 2 suggest that AR11 treatment is numerically superior to placebo in each treatment period.

3) The period 1 data analyses result in the same conclusion that SKAMP-Combined Scores were statistically significantly lower in the AR11 treatment compared to placebo at each post-dose time point (Table 5).

In conclusion, it does not seem to be a concern to draw statistical inferences about efficacy using data from both periods.

Figure 2   Treatment Effect Over Time by Period
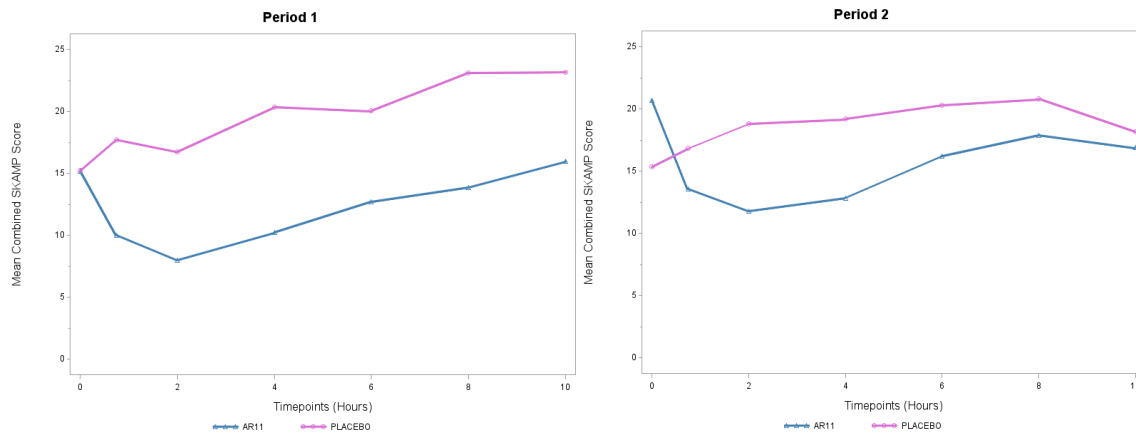
(Source: The reviewer's analyses)

Table 5   Analysis of SKAMP-Combined Scores Over Time: Period 1 vs. Period 2

Time Point (hour)	First Period				Second Period			
	SKAMP-Combined Score LS Mean (SE)		Difference in LSMean (SE) (AR11-Placebo)	p-value	SKAMP-Combined Score LS Mean (SE)		Difference in LSMean (SE) (AR11-Placebo)	p-value
	AR11	Placebo			AR11	Placebo		
Pre-dose*	15.2 (1.4)	15.2 (1.7)			20.7 (1.9)	15.3 (1.6)		
0.75 post-dose	10.3 (1.4)	17.7 (1.3)	-7.4 (1.9)	0.0001	13.8 (1.5)	17.3 (1.5)	-3.5 (2.1)	0 1059
2 post-dose	8.3 (1.4)	16.7 (1.4)	-8.5 (1.9)	<0.0001	12.1 (1.4)	19.3 (1.4)	-7.2 (2.0)	0 0003
4 post-dose	10.5 (1.5)	20.4 (1.5)	-9.9 (2.1)	<0.0001	13.1 (1.4)	19.6 (1.4)	-6.5 (2.0)	0.001
6 post-dose	13 (1.3)	20 (1.3)	-7.1 (1.8)	0.0001	16.5 (1.5)	20.8 (1.5)	-4.3 (2.1)	0 0385
8 post-dose	14.2 (1.4)	23.1 (1.4)	-8.9 (1.9)	<0.0001	18.2 (1.6)	21.3 (1.6)	-3.1 (2.2)	0 1675
10 post-dose	16.3 (1.6)	23.2 (1.6)	-6.9 (2.2)	0.002	17.2 (1.4)	18.6 (1.4)	-1.5 (1.9)	0 4512
Average all post-dose time points	12.1 (1.2)	20.2 (1.2)	-8.1 (1.7)	<0.0001	15.1 (1.2)	19.5 (1.3)	-4.3 (1.7)	0 0133

*Raw mean score at pre-dose.
(Source: The reviewer's analysis)

Reviewer's Note: *The IR was sent to request the details on assessment of differential carryover effect. The applicant investigated carryover effect using different approaches and their results agree with the reviewer's results (\\CDSESUB1\evsprod\NDA209905\0012):*

- *A hypothesis is tested that the carryover effects from period 1 were equivalent to the carryover effects from period 2 at 2 hours post-dose. Intrasubject sums were analyzed via a t-test in order to make inferences regarding a difference between the two carryover effects (Fleiss, p. 268). The result indicates that there is no statistical evidence of a differential carryover effect (p=0.63) at the significance level of 0.10 proposed by Grizzle.*
- *There was no evidence of substantial treatment by period interactions over the course of the entire classroom day (p=0.16).*
- *The analysis using period 1 data only also concluded that the primary efficacy outcome was statistically significantly different between AR11 and placebo at all post-dose time points.*

3.3   Evaluation of Safety

Please refer to clinical review for safety assessment.

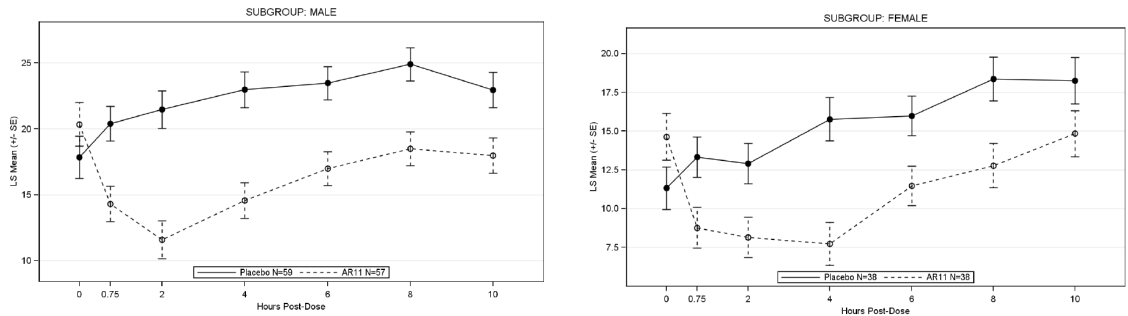## 4.   FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1   Gender, Race, Age, and Geographic Region

Exploratory subgroup analyses of primary endpoint were performed (by the applicant) on age, gender, and race. Although the applicant explored the age impact on the primary efficacy outcome, it is noted that subgroup analysis by age in this trial may not provide clinically relevant information because all enrolled subjects were children and the cutoffs to define subgroups appear to be arbitrary, thus the result of subgroup analysis by age is not presented here. The study was conducted in the United States (7 investigational sites) only and thus no analysis by geographic region was

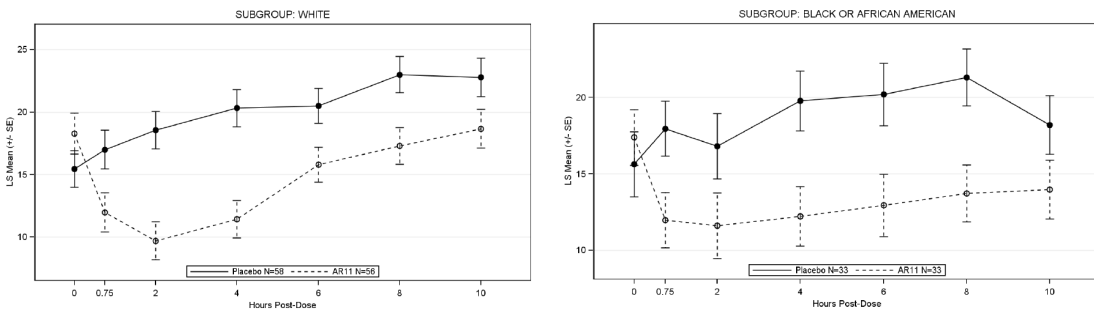performed. The numerical results seem to trend in favor of AR11 treatment for all subgroups (Figure 3).

Figure 3   Subgroup Analyses by Gender/Race

1. Subgroup by Gender:

(Source: Applicant's figure 14.2.5, confirmed by the reviewer's analysis)

2. Subgroup by Race:

(Source: Applicant's figure 14.2.7, confirmed by the reviewer's analysis)

4.2   Other Special/Subgroup Populations

Exploratory subgroup analysis of primary endpoint was also performed by the applicant on final dose used in the double-blind phase of the study.   The results are consistent across all doses trending in favor of AR11 treatment (Figure 4).

Figure 4   Subgroup Analysis by Final Dose
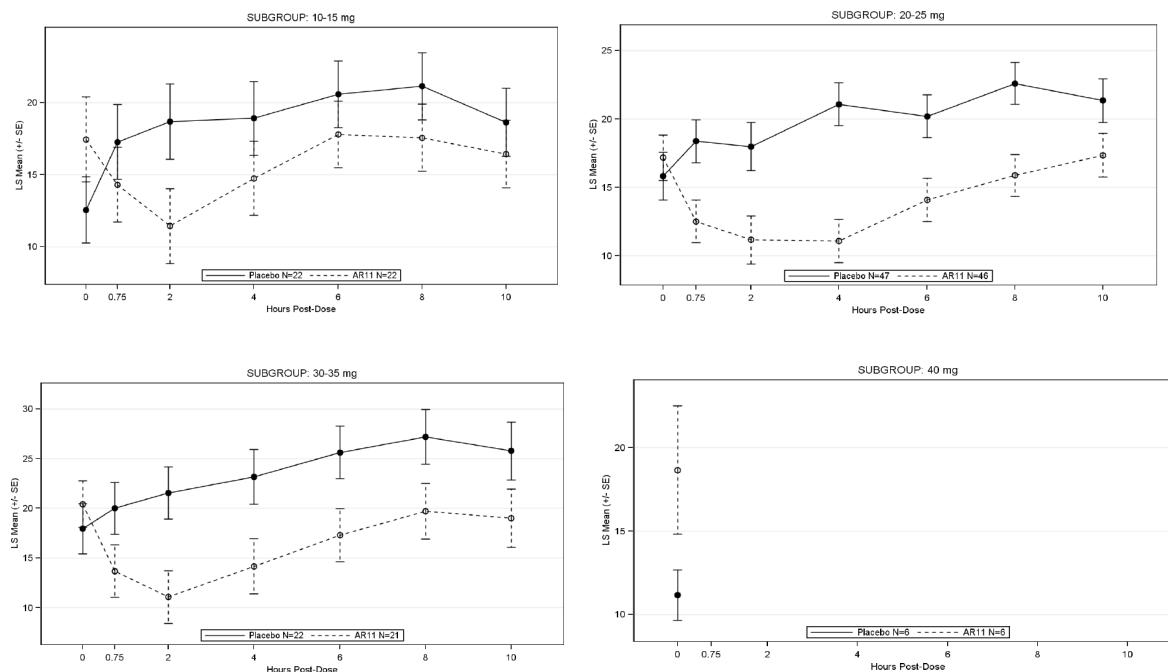
(Source: Applicant's figure 14.2.3, confirmed by the reviewer's analysis)

Reviewer' Notes: *Results from this exploratory analysis should be interpreted with extreme caution because the final dose was outcome driven.*

5.   SUMMARY AND CONCLUSIONS

5.1 Statistical Issues

The only clinical study conducted in this NDA was a randomized, double-blind, placebo-controlled, two-arm two-period, crossover design trial. A potential issue is that there was no wash-out period between the two treatment periods and possible differential carryover effect (or substantial treatment by period interaction) can bias the estimate of treatment effect. This reviewer conducted analyses to address this issue.

5.2 Collective Evidence

This reviewer's analysis does not show evidence of substantial treatment by period interaction (i.e., sequence effect) (p=0.1658). The plots also support that AR11treatment is numerically superior to placebo in each treatment period and when both treatment periods are combined over time.

5.3 Conclusions and Recommendations

The study supports the efficacy of AR11 on the reduction of signs and symptoms of ADHD in pediatric subjects 6 to 12 years old diagnosed with ADHD.

5.4 Labeling Recommendations

No additional recommendation.

This is a representation of an electronic record that was signed electronically. Following this are manifestations of any and all electronic signatures for this electronic record.

/s/

QUQUAN LIU
12/20/2018

PEILING YANG
12/20/2018

HSIEN MING J HUNG
12/20/2018