U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Pharmacoepidemiology and Statistical Science
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/Serial Number:	21-514
Drug Name:	SPD485, d,l (threo)-methylphenidate, Methylphenidate Transdermal System (MTS)
Indication(s):	Attention Deficit Hyperactivity Disorder (ADHD)
Applicant:	Shire Development Inc. and Noven Pharmaceuticals, Inc.
Date(s):	June 16, 2005
Review Priority:	Priority
Biometrics Division:	Division of Biometrics I
Statistical Reviewer:	Fanhui Kong, PhD
Concurring Reviewers:	Peiling Yang, PhD, H. M. James Hung, PhD
Medical Division:	Division of Psychiatric Products
Clinical Team:	Robert Levin, MD, Thomas Laughren, MD
Project Manager:	Richardae Taylor

Keywords: Methylphenidate Transdermal System (MTS), CONCERTA®, Attention-Deficit/Hyperactivity Disorder (ADHD), Crossover study, Mixed Effects model

Table of Contents

1. EXECUTIVE SUMMARY

There are two pivotal efficacy studies in this submission, Studies SPD485-201 and SPD485-302. The title of Study SPD485-201 is "A Phase II, Randomized, Double-Blind, Multi-Center, Placebo-Controlled, Dose Optimization, Analog Classroom, Crossover Study, Designed to Assess the Time Course of Treatment Effect, Tolerability and Safety of Methylphenidate Transdermal System (MTS) in Pediatric Patients aged 6-12 with Attention-Deficit/Hyperactivity Disorder (ADHD)". The primary objective of this study is to evaluate, under controlled conditions at multiple time points throughout the day, the behavioral effects measured by the SKAMP deportment scale of MTS compared to placebo in children (aged 6-12) diagnosed with ADHD by DSM-IV-TR criteria.

The title of Study SPD485-302 is "A Phase III, Randomized, Double-Blind, Multi-Center, Parallel-Group, Placebo-Controlled, Dose Optimization Study, Designed to Evaluate the Safety and Efficacy of Methylphenidate Transdermal System (MTS) vs. CONCERTA® in Pediatric Patients aged 6-12 with Attention-Deficit/Hyperactivity Disorder (ADHD)". The primary objective of this study is to evaluate, under controlled conditions, the safety and efficacy of SPD485 (MTS) compared to placebo with reference to CONCERTA®, as determined by the change in the clinician completed ADHD-RS-IV, in the symptomatic treatment of children (aged 6-12) diagnosed with ADHD by DSM-IV-TR criteria.

1.1 Conclusions and Recommendations

In this submission, the sponsor conducted two pivotal clinical trial studies, a Phase II, placebo controlled, randomized, crossover study and a Phase III, randomized, placebo controlled study with reference of CONCERTA®. These studies evaluated the efficacy and safety of MTS over placebo on children (aged 6-12) with ADHD. Both studies are evaluated in this review.

In Study SPD485-201, the reviewer's statistical analyses confirm the sponsor's efficacy results and support their claim of the efficacy of MTS in the treatment of children with ADHD. The drug effect seems to have started at the end of the second hour. Despite such positive evidences, we have two major concerns in the conduct of this study that add uncertainty to the validity of the claim of the sponsor. The first concern is that the baseline measurement of the primary endpoint was not taken in the study therefore could not be adjusted in the statistical analyses; the second is that the patients in placebo group did not go through a tapering period before changing to placebo, therefore it's unclear if the observed treatment effect was real or was due to the sudden change of treatment. In Study SPD485-302, the reviewer's statistical analysis results also confirm the sponsor's efficacy results and support their claim of the effectiveness of MTS in the treatment of children with ADHD.

1.2 Brief Overview of Clinical Studies

This submission consists of two pivotal clinical trial studies, a Phase II, placebo controlled, randomized, crossover study and a Phase III, randomized, placebo controlled study with the reference of CONCERTA®. The studies were conducted in 2004-2005.

In Study SPD485-201, 93 subjects aged 6 to 12 were enrolled into the Open-Label Dose Optimization period. Following this period, 80 subjects were randomized in a 1:1 sequence ratio (MTS/PTS: PTS/MTS), into the double-blind crossover Analog Classroom period and 79 (MTS/PTS: 41; PTS/MTS: 38) were avaluable for the primary ITT analysis. In Study SPD485-302, 282 subjects aged 6 to 12 were enrolled and randomized in a 1:1:1 ratio (MTS: CONCERTA: Placebo) into the double-blind dose

optimization/maintenance period and 270 (MTS: 96; CONCERTA: 89; placebo: 85) were evaluable for the primary ITT analysis.

1.3 Statistical Issues and Findings

1.3.1 Study SPD485-201

This was a phase II, randomized, double-blind, multi-center, placebo-controlled, analog classroom, crossover study, to evaluate the efficacy of MTS in treating the children (aged 6-12) diagnosed with ADHD using the SKAMP deportment scale as the primary endpoint. With a sample size of 79 in ITT population, statistical analysis using a mixed effects linear model indicates that MTS is highly statistically significant. The sponsor did not check the model assumptions in the statistical analyses as required in the SAP. There are evidences indicating that some model assumptions are violated. However, results using nonparametric models by the reviewer still support the claim that the treatment MTS is effective in reducing the SKAMP deportment score among children with ADHD.

Further analyses on the SKAMP deportment score at Hours 2 and 3 indicate that the treatment seems to have started the effect at the end of Hour 2, with p-values of 0.0467. Without the data at Hour 1, it's hard to give a better estimate of the real starting time of the drug effect.

Despite the positive efficacy results, the reviewer has two major concerns about the study. The first is that the baseline measurement of the primary endpoint was not taken in the study, therefore it couldn't be adjusted in the statistical model. The baseline measurement is meant to be the measurement at the end of Week 7, before the randomization of the crossover study. The second concern is that right after the dose optimization period, the patients were directly randomized into treatment and placebo groups. Those patients randomized to placebo group did not go through a tapering period before changing to placebo. Therefore it's unclear if the observed treatment effect was real or was due to the sudden change of treatment.

1.3.2 Study SPD485-302

This is a Phase III, randomized, double-blind, placebo-controlled, dose optimization study to compare MTS with placebo in children (aged 6-12) diagnosed with ADHD using the ADHD-RS-IV total score as the primary endpoint. With 270 subjects in the ITT population, the ANCOVA analysis indicates that MTS is highly statistically significant compared to placebo in reducing the ADHD-RS-IV total score. With model assumptions being violated, the reviewer applies the rank ANCOVA model to the data set. This analysis gives p-values of <0.0001 and 0.0156 in LOCF and OC analyses. Given the total patient dropout being about 40%, the reviewer uses the MMRM method, which takes the missingness into consideration using the assumption of non-informative dropout in the analysis of treatment. This analysis gives a $p<0.0001$ in the rest of the efficacy of MTS. All the results support the sponsor's claim of the effectiveness of MTS in treating children with ADHD.

2. INTRODUCTION

2.1 Overview

ADHD is a prevalent psychiatric disorder of childhood. It consists of a variety of behaviors and personality types. The three main symptoms of ADHD include inattention, hyperactivity, and impulsivity.

It is estimated that 3%-7% of school aged children have ADHD. These symptoms must appear before age 7, be present for more than 6 months, and must be adversely affecting social, occupational, or school functioning for the diagnosis of ADHD to be made. ADHD is believed to result from a deficiency of neurotransmission of dopamine and norepinephrine either through the insufficient sensitivity of the receptors or amount of dopamine produced. Some of the functions associated with sufficient levels of these metabolites in the central nervous system include controlling the ability to shift from an open to focused-state of awareness and, indirectly, the sense of time.

In the past, the most common therapy for ADHD has been orally dosed stimulants such as methylphenidate (MPH), dextroamphetamine and pemoline. It is believed that these medications may either stimulate the release of dopamine or block its re-uptake. It is felt that increasing dopamine levels results in increasing impulse control and enhancing a more "focused state of awareness." Studies have shown that, in children with ADHD, MPH improves classroom functioning, notably by decreasing disruptive behavior and increasing academic productivity, accuracy and improvement in teacher ratings. In addition, MPH has been shown to improve performance in children's attention and memory. For the treatment of ADHD, Ritalin-IR and Ritalin-SR were developed. In 2000, CONCERTA® was approved, and has grown in popularity due to its effectiveness through 12 hours after dosing.

MTS a transdermal delivery system containing MPH in a multi-polymeric adhesive platform, as a means of providing sustained levels of *d,l*-methylphenidate while the patch is worn. The system is designed to release *d,l*-methylphenidate continuously upon application to intact skin in order to provide greater consistency in therapeutic response, and therefore improve therapeutic efficacy. Transdermal administration of *d,l*-methylphenidate in subjects is intended to result in more stable plasma concentrations over the course of the day that may contribute to a prolonged duration of effect.

The sponsor (Noven) submitted IND 54,732 for MTS on December 12, 1997 and NDA 21-514 on June, 27 2002. On April 25, 2003, the Division issued a not approvable letter. On October 10, 2003 and March 1, 2004, Noven proposed meetings to obtain Division input on its proposed development plan to address the issues raised in the not approvable letter. At the Type C meeting on May 26, 2004, the sponsors (Noven and Shire) gained Division concurrence on the sponsors' proposal to pursue three new Phase II/III studies that could address FDA's concerns. After initiation of these new clinical studies, Noven requested a second Type C meeting with the Division. On April 5, 2005, the sponsors discussed their plans for a Type 2 Resubmission and gained Division concurrence to proceed with a mid-2005 submission.

This report summarizes the review of both studies: SPD485-201 and SPD485-302. Both studies were conducted in 2004-2005. Study SPD485-201 was a 14-week, phase II, randomized, double-blind, multi-center, placebo-controlled, analog classroom, crossover study, with an open-label optimization phase, designed to assess the time course of treatment effect, tolerability and safety of MTS in pediatric subjects diagnosed with ADHD. Study SPD485-302 was a 14-week phase III, randomized, double-blind, multi-center, parallel-group, placebo-controlled, dose optimization study designed to evaluate the safety and efficacy of MTS compared to placebo with reference to CONCERTA® in pediatric subjects diagnosed with ADHD.

In Study SPD485-201, 93 subjects aged 6 to 12 were enrolled into the Open-Label Dose Optimization period. Following this period, 80 subjects were randomized, in a 1:1 sequence ratio (MTS/PTS: PTS/MTS), into the double-blind crossover Analog Classroom period and 79 (MTS/PTS: 41; PTS/MTS: 38) were avaluable for primary ITT analysis. In Study SPD485-302, 282 subjects aged 6 to 12 were enrolled and randomized in a 1:1:1 ratio (MTS: CONCERTA: Placebo) into the double-blind dose optimization/maintenance period and 270 (MTS: 96; CONCERTA: 89; placebo: 85) were evaluable for the primary ITT analysis.

2.2 Data Sources

The applicant study reports for the efficacy and safety of the pivotal Studies SPD485-201 and SPD485-302 are all provided electronically. Individual clinical study reports may be found in Sections 8 and 10. Analysis data sets are provided electronically in \\Cdsesub1\n21514\N_000\2005-06-28\crt\datasets.

3. STATISTICAL EVALUATION

3.1 Evaluation of Efficacy

3.1.1 Study SPD485-201

3.1.1.1 Title and Study Objectives

The title of this study is "A Phase II, Randomized, Double-Blind, Multi-Center, Placebo-Controlled, Dose Optimization, Analog Classroom, Crossover Study, Designed to Assess the Time Course of Treatment Effect, Tolerability and Safety of Methylphenidate Transdermal System (MTS) in Pediatric Patients aged 6-12 with Attention-Deficit/Hyperactivity Disorder (ADHD)."

The primary objective of the study was to evaluate, under controlled conditions at multiple time points throughout the day, the behavioral effects measured by the SKAMP deportment scale of MTS in children (aged 6-12) diagnosed with ADHD by DSM-IV-TR criteria.

The main secondary objective was to assess the duration of the efficacy of MTS in children with ADHD using the PERMP (age-adjusted math test) administered at pre-dose, 2.0, 3.0, 4.5, 6.0, 7.5, 9.0, 10.5 and 12.0 hours post application/dosing in a controlled environment. Additional secondary objectives included the evaluation of the efficacy of MTS in children with ADHD as measured by the SKAMP Total score, the SKAMP sub-scales of attention and quality of work, the clinician completed ADHD-RS-IV, the parent weekly-rated CPRS-R, the Clinical Global Impressions (CGI-S and CGI-I) and Parent Global Assessments (PGA), etc.

3.1.1.2 Study Design and Endpoints

This was a phase II, randomized, double-blind, multi-center, placebo-controlled, Analog Classroom, crossover study, with an open-label optimization phase. Subjects visited the study site nine times during the course of approximately 14 weeks. The study consisted of four periods detailed below:

Screening and Washout Period: Subjects were screened for approximately 2 weeks prior to washout (up to a maximum of 28 days).

Open Label Dose Optimization Period: The objective of this 5-week period was to ensure subjects to be titrated to an optimal dose of MTS (using 12.5cm2, 18.75cm2, 25cm2, and 37.5cm2 patch sizes) based upon investigator review of parent rating forms, TEAEs, and clinical judgment (using the ADHD-RS-IV). All subjects were initiated on the MTS 12.5cm2 size patch (1/day) and were evaluated after one week for tolerability and effectiveness. The approximate duration of MTS patch wear was 9 hours per day starting each morning upon awakening. Subjects were titrated to the next patch size after a minimum of one week on the previous size. Subjects may have been titrated back down to the previous patch size to optimize

tolerability. Subject response was categorized by the investigator into one of the following three conditions:

1. Intolerable condition: (unacceptable safety profile): Subject was tapered to a lower MTS patch size (if available). If the lower patch size was not tolerable, the subject was discontinued from the study.
2. Ineffective condition: (<25% change in ADHD-RS score with acceptable safety profile): The MTS patch size was increased to the next available dose strength followed by weekly evaluation.
3. Acceptable condition: Significant reduction in ADHD symptoms with minimal side effects. Subjects who had not reached an acceptable patch size by Visit 7 were withdrawn from the study.

Double-Blind, Crossover, Analog Classroom Period: Following completion of the Dose Optimization period subjects were randomized to a sequence of one week of treatment with each of MTS and PTS (Placebo Transdermal System). The duration of this period was 2 weeks and at each end of week assessment, included both measurement of behavioral effects and plasma collection, and occurred in the controlled environment of the Analog Classroom.

Follow-up Period: Subjects who did not enroll into the open-label extension study (protocol SPD485-303) at the End of Study/Early Termination Visit (Visit 9) were followed for 30 days (+2 days) after their last dose of study drug.

To be eligible, a subject must be a male or female child aged 6 to 12 years, who must satisfy the inclusion/exclusion criteria including the DSM-IV-TR criteria for a primary diagnosis of ADHD.

From previous studies, the effect size of MTS was about 0.5 compared to placebo in children with ADHD. Assuming that the effect size for the primary efficacy variable between 2 sequence groups (MTS-Placebo, Placebo-MTS) is about 0.7, then approximately 76 subjects were needed to complete the double-blind crossover period of the study with 85% power at the significance level of 0.05 (2-sided).

The original protocol, Version 1.0, was dated June 24, 2004. On September 16, 2004 and January 28, 2005, the protocol was amended to Versions 2.0 and 3.0. In these amendments, the primary efficacy variable was defined as the mean of the SKAMP deportment scale scores over the course of a day at the 2.0, 3.0, 4.5, 6.0, 7.5, and 9.0 hour. The primary efficacy variable was planned to be assessed by a mixed linear model with sequence, period and treatment as fixed effects, and subject-within-sequence as a random effect. The SKAMP deportment scores at each time point through the day (2.0, 3.0, 4.5, 6.0, 7.5, 9.0, 10.5, and 12.0 hours) were planned to be analyzed using the same model.

More secondary outcome measures were added. These included other SKAMP scores (total, attention, and quality of work) and the PERMP scores which were to be analyzed by the same model for the primary efficacy variable.

Changes to the statistical analysis plan

The original Statistical Analysis Plan (SAP), Version 1.0, was dated January 21, 2005. On February 23, 2005, the SAP was amended to Version 2.0. Major changes were made in the calculation of missing values and in the primary and secondary endpoint analyses. The Shapiro-Wilk's test in the examination of the normality of regression residuals was removed. Major changes to the definition of missing values include: the calculation of missing values in the SKAMP total scale and subscales; the calculation of missing values in the total ADHD-RS scale and subscales and the calculation of missing values in the CPRS-R subscale.

3.1.1.3 Primary and Secondary Endpoints

The primary efficacy variable was the mean SKAMP deportment scale score over the course of the Analog Classroom session days at 2.0, 3.0, 4.5, 6.0, 7.5 and 9.0 hours.

The main secondary outcome measure of the study was the PERMP, measured at pre-dose, 2.0, 3.0, 4.5, 6.0, 7.5, 9.0, 10.5 and 12.0 hours post application of MTS. The PERMP is an age-adjusted math test that is time-sensitive, ADHD medication-sensitive measure to evaluate ADHD subjects across the day. Additional secondary outcomes measures were the clinician-rated ADHD-RS-IV, Clinician Global Impressions of Improvement (CGI-I), Parent Global Assessment (PGA), Conners' Parent Rating Scale – Revised: Short Form (CPRS-R).

3.1.1.4 Patient Disposition, Demographic and Baseline Characteristics

The study was conducted from August 24, 2004 to February 1, 2005. Of the 93 subjects enrolled in this study, 13 subjects were terminated prior to randomization: 7 for AEs, 3 with withdrew consent, and 3 for other reasons. Seventy nine (98.8%) of the 80 randomized subjects completed the study and comprised the ITT population. A total of 56 (70.0%) of subjects were included in the per protocol population.

Table 3.1.1.1 Summary of the End of Study Record (All Enrolled Subjects)

Study Subjects	Treatment Sequence					
	MTS/PTS		PTS/MTS		Total	
	n	%	n	%	n	%
Enrolled	NA	NA	NA	NA	93	NA
Randomized	42	NA	38	NA	80	(100.0)
Discontinued Post-Randomization	1	(2.4)	0	0	1	(1.3)
Completed	41	(97.6)	38	(100.0)	79	(98.8)
Reason for Discontinuation: Post Randomization						
Adverse Events	0	0	1†	(2.6)	0	0
Protocol Violation	1	(2.4)	0	0	1	(1.3)
Analysis Population						
ITT	41	(97.6)	38	(100.0)	79	(98.8)
Per Protocol	31	(73.8)	25	(65.8)	56	(70.0)

† Subject 01-014, completed Visit 9; however, MTS patch was removed prior to maximum wear time.

The mean age of the ITT population was 9.1 years, with 57.0% subjects 6-9 years of age and 43.0% subjects 10-12 years of age. There were 72.2% males and 27.8% females. The majority of subjects were White (69.6%) and of Not Hispanic or Latino (75.9%) ethnicity. The ADHD-RS-IV scores at Baseline ranged from 26-54, with a mean of 41.8.

Table 3.1.1.2 Demographics and Baseline Characteristics: ITT Subjects

Characteristic	Category	Treatment Sequence MTS/PTS (N=41)	Treatment Sequence PTS/MTS (N=38)	Overall (N=79)
Age (years)	**Mean (SD)**	9.3 (1.88)	8.9 (1.56)	9.1 (1.74)
	Median	9.0	8.5	9.0
	Min, Max	6, 12	6, 12	6, 12
Age Category n(%)	**6–9 years**	22 (53.7%)	23 (60.5%)	45 (57.0%)
	10-12 years	19 (46.3%)	15 (39.5%)	34 (43.0%)
Gender n(%)	**Male**	30 (73.2%)	27 (71.1%)	57 (72.2%)

	Female	11 (26.8%)	11 (28.9%)	22 (27.8%)
Ethnicity n(%)	**Hispanic/Latino**	10 (24.4%)	9 (23.7%)	19 (24.1%)
	Not Hispanic/Latino	31 (75.6%)	29 (76.3%)	60 (75.9%)
Race n(%)	**White**	25 (61.0%)	30 (78.9%)	55 (69.6%)
	Black/African American	4 (9.8%)	4 (10.5%)	8 (10.1%)
	Asian	4 (4.9%)	0 (0%)	2 (2.5%)
	Other	10 (24.4%)	4 (10.5%)	14 (17.7%)
Weight (lb)	**Mean (SD)**	72.1 (19.85)	68.3 (13.70)	70.3 (17.17)
	Median	72	65	68.4
	Min-Max	41.0 – 126.5	46.5 -102.0	41.0–126.5
Height (in)	**Mean (SD)**	54.0 (5.21)	53.1 (3.34)	53.6 (4.40)
	Median	53.5	54	54
	Min-Max	43.5 – 65.0	46.0 – 60.0	43.5 – 65.0
ADHD-RS-IV	**Mean (SD)**	41.8 (8.50)	41.8 (6.64)	41.8 (7.61)
	Median	45.0	41.5	43.0
	Min-Max	26 – 53	29 – 54	26 – 54

Protocol violations/deviations recorded for this study included: the subject's average weekly drug compliance was less than 80% or greater than 100%; the subject failed to meet all inclusion/exclusion criteria; the subject took prohibited medications; deviations deemed to affect efficacy and identified at the Blinded Data Review Meeting, held prior to database lock. Major protocol deviations were reported for 23 (29.1%) subjects overall in the ITT population. The number of subjects with deviations was similar for both treatment sequences. Twelve (15.2%) subjects were non-compliant with study medication. Seven (8.9%) subjects had used prohibited medication and 6 (7.6%) subjects had violated inclusion/exclusion criteria

3.1.1.5 Statistical Methodologies Used

All efficacy analyses were based on the ITT population. Statistical testing was performed using a mixed linear model to analyze the mean SKAMP deportment score. The model included sequence (two levels), period (two levels) and treatment (two levels) as fixed effects and subject-within-sequence as a random effect. The two treatment levels were MTS and placebo. The SKAMP deportment scores at each time point through the day (2.0, 3.0, 4.5, 6.0, 7.5, 9.0, 10.5 and 12.0 hours) were also analyzed using the model described above.

3.1.1.6 Results by the Sponsor

3.1.1.6.1 Primary Endpoint Results

The analyses of the efficacy data were conducted in the ITT population for both primary and all secondary efficacy variables. Statistical testing was performed using a mixed linear model to analyze the mean SKAMP deportment score. The model included sequence (two levels), period (two levels) and treatment (two levels) as fixed effects and subject-within-sequence as a random effect. The two treatment levels were MTS and placebo. The LS mean (±SE) SKAMP deportment score for MTS (3.2±0.58) was significantly lower ($p<0.0001$) than that for PTS (8.0±0.58). The LS mean difference in SKAMP deportment scores was -4.8, with a 95% confidence interval of (-5.89, -3.64).

Table 3.1.1.3 Analysis of Mean SKAMP Deportment Score during Patch Application (Hours 2.0 – 9.0): ITT Population

	MTS (N=79)	Placebo (N=79)	p-value
Mean (SD)	3.2 (3.64)	8.0 (6.33)	
LS Mean (SE)	3.2 (0.58)	8.0 (0.58)	<0.0001[a]
Difference and 95% CI of LS Means (MTS-Placebo)	-4.8 (-5.89, -3.63)	NA	

[a]: The p-value is obtained using the mixed effects model.

3.1.1.6.2 Secondary Endpoint Results

The treatment of MTS improved the student PERMP scores compared to placebo in the Analog Classroom Period. The LS mean (±SE) PERMP: Number of Math Problems Attempted score for MTS (113.8±6.39) was significantly higher (p<0.0001) than that for PTS (86.2±6.39). The LS mean (±SE) PERMP: Number of Math Problems Correct score for MTS (109.4±6.34) was significantly higher (p<0.0001) than that for PTS (80.7±6.34). The LS mean (±SE) PERMP: Sum of Number of Math Problems Attempted and Correct score for MTS (223.2±12.67) was significantly higher (p<0.0001) than that for PTS (167.0±12.67).

The mean SKAMP Total scores were improved by the treatment of MTS in the ITT population. The MTS LS mean (±SE) (9.4±0.99) was significantly lower (p<0.0001) than the PTS LS mean (±SE) (17.9±0.99). The mean ADHD-RS-IV Total scores were improved by the treatment of MTS in the ITT population. The MTS LS mean (±SE) (16.3±1.24) was significantly less (p<0.0001) than the PTS LS mean (±SE) (32.7±1.23).

The CGI-I scores were improved by the treatment of MTS in the ITT population. A significantly larger (p<0.0001) number of MTS subjects than PTS subjects were rated as improved. For Period 1 (V8), 33 (80.5%) MTS subjects and 6 (15.8%) PTS subjects were rated as showing improvement. For Period 2 (V9), 30 (78.9%) MTS subjects and 3 (7.3%) PTS subjects were rated as showing improvement. The PGA scores were improved by the treatment of MTS in the ITT population. A significantly larger (p<0.0001) number of MTS subjects than PTS subjects were rated as showing improvement. For Period 1 (V8), 27 (65.9%) MTS subjects and 9 (24.3%) PTS subjects were rated as showing improvement. For Period 2 (V9), 29 (76.3%) MTS subjects and 3 (7.3%) PTS subjects were rated as showing improvement.

The mean CPRS-R Total scores were improved by the treatment of MTS in the ITT population. The MTS LS mean (±SE) (20.2±2.11) was significantly lower (p<0.0001) than the PTS LS mean (±SE) (35.3±2.21). The 95% confidence interval for the LS mean difference (MTS-PTS) of -15.1 was (-20.5, -9.66).

Table 3.1.1.4 Secondary Efficacy Endpoints at Analog Classroom Period (ITT Population)

	MTS (N=79)	PTS (N=79)	p-value
PERMP Measure: Number of math problems attempted			
LS Mean (SE)	113.8 (6.39)	86.2 (6.39)	<0.0001
N	79	79	
PERMP Measure: Number of math problems correct			
LS Mean (SE)	109.4 (6.34)	80.7 (6.34)	<0.0001

N	79	79	
SKAMP total score LS Mean (SE) N	9.4 (0.99) 79	17.9 (0.99) 79	<0.0001
ADHD-RS-IV total score LS Mean (SE) N	16.3 (1.24) 78	32.7 (1.23) 79	<0.0001
CPRS-R total score LS Mean (SE) N	20.2 (2.11) 67	35.3 (2.21) 61	<0.0001

3.1.1.7 Reviewer's Comments and Findings

3.1.1.7.1 Efficacy Results

Using the ITT data set provided by the sponsor, the reviewer duplicated the testing results for the primary endpoint and derived the same p-values. The results are depicted in Table 3.1.1.5.

Table 3.1.1.5 Analysis of Mean SKAMP Deportment Score during Patch Application (Hours 2.0 – 9.0): ITT Population

	MTS (N=79)	**Placebo (N=79)**	**p-value**
Mean (SD)	3.2 (3.64)	8.0 (6.33)	
LS Mean (SE)	3.2 (0.58)	8.0 (0.58)	<0.0001[a]
Difference and 95% CI of LS Means (MTS-Placebo)	-4.8 (-5.89, -3.63)	NA	

[a]: The p-value is obtained using the mixed effects model.

3.1.1.7.2 Further Statistical Analyses

According to the SAP, if there is strong evidence that the model assumptions are not met, the non-parametric method for 2x2 crossover design may be performed in support of the primary analysis. The difference between responses of Treatment Period 1 and Period 2 will be assessed by Wilcoxon Rank-Sum test. Treatment by period interaction will be assessed and if a significant interaction is found (at the 10% level), a parallel comparison of treatment groups will be carried out for data measured in Treatment Period 1.

However, the standard of "strong evidence against the model assumptions" is not clearly defined. As a verification of the primary efficacy result, a nonparametric test for treatment efficacy is performed using the Wilcoxon Rank-Sum test and it gives a p-value below 0.0001. The difference of primary responses between Treatment Periods 1 and 2 was also tested using Wilcoxon Rank-Sum test. This test gives a nonsignificant p-value of 0.27. The treatment by period interaction test in the mixed effects model gives a p-value of 0.38. A nonparametric parallel comparison of treatment groups using data measured in Treatment Period 1 gives a p-value below 0.0001. Therefore, the nonparametric analyses support the primary analysis results.

To see when the treatment had started the effect, in addition to the statistical testing on the average SKAMP score, the same method is also performed on the SKAMP deportment score at Hours 2 and 3.

The p-values for the treatment efficacy at Hours 2 and 3 are 0.0467 and 0.0035. This indicates that the treatment effect seems to be borderline significant at Hour 2. Without further the data at Hour 1, it's hard to determine the real starting time of the treatment effect.

3.1.1.7.3 Statistical Issues

In the primary analysis, a major concern is that the baseline SKAMP was not adjusted in the mixed linear model for the treatment efficacy. The baseline measure should have been taken at the end of Week 7, before the randomization of the crossover study. However, such a measurement was not taken. Therefore, it could not be adjusted in the analysis model. The sponsor claimed that the SAP was written and finalized prior to study database lock. Appendix 1.9 indicates that the Version 2 of the SAP was signed off on February 23, 2005. However, the Final Report did not give the data unblinding date. The study was finished on February 1, 2005 and the Final Reported was finished on May 11, 2005. But no specific date of database blocking was given in the Final Report.

Another concern of the study design is that right after the dose optimization period, the patients entered the crossover period in which patients were randomized into treatment and placebo groups. Those patients randomized to placebo group did not go through a tapering period before changing to placebo. Therefore it's unclear if the observed treatment effect was real or was due to the sudden change of treatment.

3.1.2 Study SPD485-302

3.1.2.1 Title and Study Objectives

The title of this study is "A Phase III, Randomized, Double-Blind, Multi-Center, Parallel-Group, Placebo-Controlled, Dose Optimization Study, Designed to Evaluate the Safety and Efficacy of Methylphenidate Transdermal System (MTS) vs. CONCERTA® in Pediatric Patients aged 6-12 with Attention-Deficit/Hyperactivity Disorder (ADHD)".

The primary objective of this study was to evaluate, under controlled conditions, the safety and efficacy of SPD485 (MTS) compared to placebo with reference to CONCERTA®, as determined by the change in the clinician completed ADHD-RS-IV, in the treatment of children (aged 6-12) diagnosed with ADHD by DSM-IV-TR criteria.

The main secondary objective was to assess the efficacy of MTS in an academic setting using the change in CTRS-R, completed by the subject's teacher in the morning and afternoon, 2 days per week during the study. Other secondary objectives included: To assess the efficacy of MTS in the home environment as rated by parent using the CPRS-R administered weekly; to assess global impressions of ADHD severity and improvement of MTS using CGI-S and CGI-I, PGA; to evaluate the safety and tolerability of MTS; to assess the relationship between plasma exposure and the safety and efficacy measures of MTS and CONCERTA® via sparse sampling, etc.

3.1.2.2 Study Design and Endpoints

This was a phase III, randomized, double-blind, multi-center, parallel-group, placebo-controlled, dose optimization study designed to evaluate the safety and efficacy of MTS (12.5cm2, 18.75cm2, 25cm2, and 37.5cm2 patch sizes) compared to placebo with reference to CONCERTA® in pediatric subjects with ADHD. Subjects visited the study site nine times during the course of approximately 14 weeks.

Subjects were screened approximately 2 weeks prior to washout. Washout was up to 28 days depending upon the half-life of the subject's medication requiring washout. Then the patients entered the double-blind dose optimization/maintenance period. In this period, eligible subjects were randomized in a 1:1:1 ratio to MTS, CONCERTA®, or matching placebo and entered the double-blind stepwise dose optimization period. The objective of this period was to ensure subjects were titrated to at least an acceptable dose of MTS or CONCERTA® based upon investigator review of parent and teacher rating forms, TEAEs, and clinical judgment (using the ADHD-RS-IV). The duration of this period was five weeks to allow for titration up to the highest dose and one titration down to a prior dose level, if necessary. No further titration up or down was permitted once subjects had been titrated down.

The duration of MTS/PTS patch was nine hours per day. All subjects were evaluated after 1 week (7±2 days) for tolerability and effectiveness. Titration to the next patch size/dosage strength was allowed after a minimum of 1 week on the previous size/dose based on the overall response of the subject. Additionally, subjects may have been titrated back down to the previous patch size/dosage strength (once) to optimize tolerability and effectiveness. As in Study 201, subject response was categorized by the investigator into 1 of 3 conditions and associated actions were taken: intolerable condition, ineffective condition and acceptable condition. Subjects who did not reach at least an acceptable dose by Visit 7 were withdrawn from the study. Following the successful titration by Visit 7, subjects maintained the dose through the maintenance period. Double-blind assessment of the safety and efficacy of MTS/CONCERTA®/Placebo proceeded for two weeks. At the end of study visit (Visit 9), eligible subjects had the option to enroll into an open-label extension study (protocol SPD485-303).

A total of 258 subjects (86 per group) was designed to detect an effect size of 0.5 (mean difference of 2.5 and standard deviation of 5.0) with 90% power at a significance level of 0.05. Assuming a dropout rate of 14%, 300 subjects were to be randomized to treatment in ITT group (approximately 100 subjects per treatment group). A total of 282 subjects were enrolled into the study. Following completion of screening and washout, subjects were randomized, in a 1:1:1 ratio (MTS: CONCERTA®: Placebo), into the double-blind dose optimization/maintenance period.

3.1.2.3 Primary and Secondary Endpoints

The primary efficacy variable was the ADHD-RS-IV change from baseline score at the endpoint. The null hypothesis was that there was no difference between MTS and placebo. The main secondary efficacy assessment was the CTRS-R total scores. The other secondary efficacy assessments included the CPRS-R, CGI-I and PGA. The endpoint of these secondary efficacy assessments was defined as the last post-baseline assessment for which a valid value was obtained.

3.1.2.4 Patient Disposition, Demographic and Baseline Characteristics

The study was conducted from August 23, 2004 to February 2, 2005. Of the 282 patients randomized, 270 remained in the ITT population (MTS: 96; Concerta: 89; Placebo: 85) and 141 were in the PP population (MTS: 60; Concerta: 55; Placebo: 26). A total of 113 (40.1%) randomized subjects did not complete the study.

Data from site 44 was eliminated from efficacy population due to incomplete data documentation and GCP noncompliance issues. There were two subjects, of the five CRFs submitted, included in the safety population due to documentation of receiving at least one dose of study medication.

Table 3.1.2.1 shows the incidence of and reasons for premature withdrawal from the study in the enrolled population. Of the 282 subjects in the enrolled population, 113 subjects prematurely withdrew from the study. The most common reason for withdrawal was Change to Study 303, which accounted for 22% of subjects. Other common reasons of withdrawal were Other (6%) and Parental Withdraw Consent (5%).

Table 3.1.2.1 Summary of the End of Study Record (All Enrolled Subjects)

Study Completion or Discontinuation	MTS (N=100)	Concerta (N=94)	Placebo (N=88)	Total (N=282)
Intent-to-Treat (ITT)	96 (96%)	89 (94.7%)	85 (96.6%)	270 (95.7%)
Total Discontinuation	29 (29%)	28 (30%)	56 (63.6%)	113 (40.1%)
Reasons for Discontinuation				
Adverse Event	5 (5%)	3 (3.2%)	1 (1.1%)	9 (3.2%)
Protocol Violation	1 (1%)	1 (1.1%)	3 (3.4%)	5 (1.8%)
Parental Withdraw Consent	3 (3%)	5 (5.3%)	6 (6.8%)	14 (5%)
Subject Lost to Follow-up	2 (2%)	0 (0%)	2 (2.3%)	4 (1.4%)
Other	4 (4%)	1 (1.1%)	11 (12.5%)	16 (5.7%)
Continued to Study 303	12 (12%)	17 (18.1%)	32 (36.4%)	61 (21.6%)

For the ITT population, there were no significant differences between treatment group and placebo group regarding gender, race, age, weight and height. The average patient across all treatment groups was approximately 9 years old. Majority (77%) were Caucasian. The overall percentage of male patients was 66%.

Table 3.1.2.2 Demographic Demographics and Baseline Characteristics of All Randomized Subjects

Characteristic	Category	MTS (N=100)	CONCERTA (N=94)	Placebo (N=88)	Total (N=282)
Age (years)	**Mean**	8.9	8.8	8.5	8.8
	SD	1.96	1.94	1.91	1.94
Age Category n(%)	**6–9 years**	61 (61.0%)	60 (63.8%)	62(70.5%)	183 (64.9%)
	10-12 years	39 (39.0%)	34 (36.2%)	26(29.5%)	99 (35.1%)
Gender n(%)	**Male**	60 (60.0%)	62 (66.0%)	65(73.9%)	187 (66.3%)
	Female	40 (40.0%)	32 (34.0%)	23(26.1%)	95 (33.7%)
Ethnicity n(%)	**Hispanic/Latino**	16 (16.0%)	11 (11.7%)	8 (9.1%)	35 (12.4%)
	Not Hispanic/Latino	84 (84.0%)	83 (88.3%)	79(89.8%)	246 (87.2%)
	Missing			1 (1.1%)	1 (0.4%)
Race n(%)	**White**	79 (79.0%)	75 (79.8%)	64(72.7%)	218 (77.3%)
	Black/African American	11 (11.0%)	13 (13.8%)	17(19.3%)	41 (14.5%)
	Asian	2 (2.0%) 8	0 (0%)	0 (0.0%)	2 (0.7%)
	Other	(8.0%)	6 (6.4%)	7 (8.0%)	21 (7.4%)
Weight (lb)	**Mean**	72.9	73.0	68.7	71.6
	SD	24.09	20.89	19.18	21.60
	Median	68.2	69.8	62.5	67.2
	Min-Max	37.0 – 148.3	41.0 -144.5	40.0 – 35.0	37.0–148.3
Height (in)	**Mean**	53.4	53.2	52.4	53.1
	SD	5.39	4.97	5.14	5.17
	Median	54.0	52.5	52.3	52.6

	Min-Max	42.3 – 68.0	42.9 – 66.5	39.2 – 65.8	39.2 – 68.0
ADHD-RS-IV	**Mean**	43.1	43.4	42.1	42.9
	SD	7.39	7.11	7.41	7.30
	Median	44.0	45.0	43.0	44.0
	Min-Max	28 – 54	19 – 54	27 – 54	19 – 54

The major protocol violation/deviation was non-compliance. A total of 36 (13.3%) patients in the ITT population were considered as having non-compliance. The incidence of non-compliance was similar in the three groups. There was no notable difference between the treatment groups. For all randomized subjects, the mean (SD) age at ADHD onset was 7.07 (2.33) years, the mean (SD) duration of ADHD diagnosis was 1.64 (2.28) years, and the combined ADHD sub-type was the most common (227 subjects, 80.5%). The characteristics of the ITT and PP populations were similar. The primary outcome variable at baseline (randomization) was comparable between the MTS group and the placebo group.

3.1.2.5 Statistical Methodologies Used

The primary efficacy analysis was performed on the ITT population. The null hypothesis was tested using the analysis of covariance (ANCOVA) model with treatment as a factor and baseline ADHD-RS-IV score as a covariate. The same ANCOVA model was used for continuous secondary endpoints. The CGI-I and PGA were analyzed by a Chi-square test. Prior to the analysis, this variable were dichotomized to two categories, with 'very much improved' and 'much improved' into one category and the remaining levels into the other.

3.1.2.6 Results by the Sponsor

3.1.2.6.1 Primary Endpoint Results

The analyses of the efficacy data were conducted in the ITT population as well as PP population for both the primary and all secondary efficacy variables. Treatment efficacy was analyzed using ANCOVA model for the change from baseline of ADHD-RS-IV total score with treatment as factor and baseline ADHD-RS-IV total score as the covariate. In the ITT population, with LOCF data set, the LS mean (±SE) change from baseline of ADHD-RS-IV total score for MTS (-24.2±1.45) was significantly lower ($p<0.0001$) than that for placebo (-10.3±1.54). The LS mean difference between MTS and Placebo in the change of the total ADHD-RS-IV scores was -13.9, with a 95% confidence interval of (-18.1, -9.7). The magnitude of such difference in the PP population is much smaller, -5.6 (-10.6, -0.6), and less significant, with p-value being 0.029.

Table 3.1.2.3 Analysis of the Change from Baseline of ADHD-RS-IV Total Score (ITT Population)

	MTS (N=96)	**Concerta (N=89)**	**Placebo (N=85)**
LOCF analysis			
N	96	89	85
Mean (SD)	-24.2 (14.55)	-22.0 (14.91)	-9.9 (14.06)
LS Mean (SE)	-24.2 (1.45)	-21.6 (1.51)	-10.3 (1.54)
Difference and 95% CI of LS Means (Active-Placebo)	-13.89 (-18.06, -9.72)	-11.32 (-15.58, -7.06)	
p-value	<0.0001	<0.0001	
OC Analysis			

N	70	64	31
Mean (SD)	-29.8 (10.40)	-28.0 (11.13)	-22.4 (13.67)
LS Mean (SE)	-30.1 (1.21)	-27.2 (1.27)	-23.5 (1.83)
Difference and 95% CI of LS Means (Active-Placebo)	-6.58 (-10.91, -2.24)	-3.77 (-8.19, 0.66)	
p-value	0.0032	0.095	

Table 3.1.2.4 Analysis of the Change from Baseline of ADHD-RS-IV Total Score (PP Population)

	MTS (N=60)	**Concerta (N=55)**	**Placebo (N=26)**
LOCF analysis			
N	60	55	26
Mean (SD)	-28.4 (10.72)	-29.2 (11.18)	-21.5 (15.0)
LS Mean (SE)	-28.8 (1.39)	-28.0 (1.47)	-23.2 (2.13)
Difference and 95% CI of LS Means (Active-Placebo)	-5.61 (-10.62, -0.60)	-4.85 (-10.02, 0.31)	
p-value	0.029	0.065	
OC Analysis			
N	59	53	25
Mean (SD)	-28.7 (10.44)	-29.7 (10.13)	-22.8 (13.92)
LS Mean (SE)	-29.1 (1.30)	-28.4 (1.38)	-24.5 (2.01)
Difference and 95% CI of LS Means (Active-Placebo)	-4.65 (-9.36, 0.07)	-3.99 (-8.87, 0.89)	
p-value	0.053	0.11	

3.1.2.6.2 Secondary Endpoint Results

Significant differences were also found between MTS and placebo groups in the mean changes from baseline in the secondary endpoints. These secondary endpoints include ADHD-RS-IV subscale for hyperactivity/impulsivity, ADHD-RS-IV subscale for inattentiveness, CTRS-R total score, CRPS-R total score at 11:00 am, 3:00 pm and the endpoint, and finally the CGI and PGA scales.

At Endpoint in the ITT population, the LS mean change in the ADHD-RS-IV hyperactivity/impulsivity score in MTS group was statistically significantly different ($p<0.0001$) from the corresponding score in placebo group, with LS means (SE) of -11.8 (0.73) and -5.2 (0.78), respectively, and an LS mean difference (95% CI) of -6.65 (-8.86, -4.53). The LS mean change in the ADHD-RS-IV inattentiveness score in MTS group was statistically significantly different ($p<0.0001$) from the corresponding score in placebo group, with LS means (SE) of -12.4 (0.78) and -5.2 (0.83), respectively, and an LS mean difference (95% CI) of -7.25 (-9.49, -5.01). The LS mean change in the CTRS-R total score in MTS group was statistically significantly different ($p<0.0001$) from the corresponding score in placebo group, with LS means (SE) of -15.3 (1.69) and -5.1 (1.78), respectively, and an LS mean difference (95% CI) of -10.19 (-15.03, -5.35). The LS change in the mean CPRS-R total score in MTS group was statistically significantly different ($p<0.0001$) from the corresponding score in placebo group, with LS means (SE) of -27.8 (2.08) and -14.4 (2.22), respectively, and an LS mean difference (95% CI) of -13.42 (-19.42, -7.42). The percentage of improvement of CGI scale in MTS group was statistically significantly different ($p<0.0001$) from the corresponding percentage of improvement in placebo group. In the MTS group, 71.9% improved while 28.1% did not improve, and in the placebo group 23.5% improved while76.5% did not improve. The percentage of improvement of PGA scale in MTS group was statistically significantly different ($p<0.0001$) from the corresponding percentage of improvement in placebo group. In the MTS

group, 69.8% improved while 30.2% did not improve, and in the placebo group 24.7% improved while75.3% did not improve.

Table 3.1.2.4 Secondary Efficacy Endpoints at the End of Study Relative to Baseline (LOCF Analysis, ITT Population)

Mean Change from Baseline	MTS (N=96)	Concerta (N=89)	Placebo (N=85)
ADHD-RS-IV subscale for hyperactivity/impulsivity			
LS Mean (SE)	-11.8 (0.73)	-10.9 (8.06)	-4.8 (6.86)
p-value	<0.0001	<0.0001	
N	96	89	85
ADHD-RS-IV subscale for Inattentiveness			
LS Mean (SE)	-12.4 (0.78)	-11.0 (0.81)	-5.2 (0.83)
p-value	<0.0001	<0.0001	
N	96	89	85
CTRS-R total score			
LS Mean (SE)	-15.3 (1.69)	-17.5 (1.75)	-5.1 (1.78)
p-value	<0.0001	<0.0001	
N	82	76	74
CPRS-R daily mean total score at endpoint			
LS Mean (SE)	-27.8 (2.08)	-23.0 (2.10)	-14.4 (2.22)
p-value	<0.0001	0.0053	
N	85	83	75
CGI scale			
Improvement at the end (%)	69 (71.9%)	59 (66.3%)	20 (23.5%)
p-value	<0.0001	<0.0001	
PGA scale			
Improvement at the end (%)	67 (69.8%)	54 (60.7%)	21 (24.7%)
p-value	<0.0001	<0.0001	

3.1.2.7 Reviewer's Comments and Findings

3.1.2.7.1 Efficacy Results

Using both the ITT and PP data sets provided by the sponsor, the reviewer duplicates the testing results for the primary endpoint using both the LOCF and OC data sets and derives the same p-values. Only the results of ITT population are given in the following Table 3.1.2.5.

Table 3.1.2.5 Analyses of the Change from Baseline of ADHD-RS-IV Total Score (ITT Population)

	MTS (N=96)	Concerta (N=89)	Placebo (N=85)
LOCF analysis			
N	96	89	85
Mean (SD)	-24.2 (14.55)	-22.0 (14.91)	-9.9 (14.06)
LS Mean (SE)	-24.2 (1.45)	-21.6 (1.51)	-10.3 (1.54)

Difference and 95% CI of LS Means (Active-Placebo)	-13.89 (-18.06, -9.72)	-11.32 (-15.58, -7.06)	
p-value	<0.0001	<0.0001	
OC Analysis			
N	70	64	31
Mean (SD)	-29.8 (10.40)	-28.0 (11.13)	-22.4 (13.67)
LS Mean (SE)	-30.1 (1.21)	-27.2 (1.27)	-23.5 (1.83)
Difference and 95% CI of LS Means (Active-Placebo)	-6.58 (-10.91, -2.24)	-3.77 (-8.19, 0.66)	
p-value	0.0032	0.095	

3.1.2.7.2 Further Efficacy Analyses

According to SAP, the assumptions of the ANCOVA model will be confirmed regarding normality of the distributions and homogeneity of variance. The residuals will be examined through histograms, normal q-q plots, and plots of residuals versus fitted values. If there is strong evidence that the assumptions are not met, a rank ANCOVA will be performed in support of the primary model. The rank ANCOVA (non-parametric approach) will be conducted using the following method. The change from baseline to endpoint and baseline are first ranked and then the change from baseline to endpoint is regressed on the baseline. The residuals from this linear regression model are finally compared for two treatment groups using the Mantel-Haenszel mean score Chi-Square test.

To test the normality assumption, the ANCOVA model is performed by the reviewer in both LOCF and OC analyses. The normality of the residuals is tested using the Shapiro-Wilk test and the p-values are 0.007 and <0.0001, respectively, indicating a strong evidence against the normality assumption. Both the q-q plots and the histograms indicate that the residuals are not normally distributed. Among them the residuals for the LOCF analysis are more symmetrically distributed than that for the OC analysis. On the other hand, the scatter plots of the residuals against the predicted values do not indicate the non-homogeneity of the variances. Based on these results, we perform the rank ANCOVA according to the SAP. The rank ANCOVA analyses give p-values of <0.0001 and 0.0156 in LOCF and OC analysis data for the treatment effect of MTS versus placebo. Therefore these results support the primary analyses.

3.1.2.7.3 Statistical Issues

The reviewer also notices that the actual percentage of patients who did not complete the efficacy study was almost 40% (105 in number) in the ITT population rather than the estimated 14% in the computation of sample size. In fact, there were 27% dropout in the MTS group, 28% dropout in the Concerta group and 64% dropout in the placebo group. The difference is highly significant with $p< 0.0001$ according to Fisher's exact test.

Although the LOCF analysis was accepted as the primary analysis by the agency, the shortcoming was obvious. It requires the outcome values to be stable over time. This is obviously not the case given the mean total ADHD-RS-IV score changed from 42.9 at the baseline to 14.8 at the last visit. Alternatively, the reviewer uses the mixed effects model, namely the MMRM method to test the treatment effect which takes the missingness into consideration based on the assumption of non-informative dropout. Although such an assumption is hard to verify, it seems to be a much more reasonable and acceptable one compared to LOCF analysis. This analysis gives a $p<0.0001$ for the treatment effect of MTS over placebo at the last visit time. This supports the sponsor's claim.

3.2 Evaluation of Safety

See medical review for detail.

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Race and Age

4.1.1 Study SPD485-201

During the statistical review, the effect of sex on the treatment effect is evaluated by first testing the significance of sex as a factor and then testing the treatment effect after the adjustment of sex in the mixed effects model. The significance test of sex in the model gives a p-value of 0.047. But MTS is still highly significant after the adjustment of sex. To see if the treatment effects are the same in these two groups, we perform subgroup analyses in the two gender groups separately and the results are depicted in Table 4.1.1.

Table 4.1.1 Subgroup Analyses of the Treatment Effect on the Primary Endpoint in Two Sex Groups (LOCF Analysis)

Sex		MTS (N=79)	Placebo (N=79)	p-value
Male		**N=57**	**N=57**	
	LS Mean (SE)	2.74 (0.62)	7.51 (0.62)	<0.0001
	Difference and 95% CI of LS Means	-4.77 (-5.88, -3.67)		
Female		**N=22**	**N=22**	
	LS Mean (SE)	2.11 (0.95)	5.89 (0.95)	0.0021
	Difference and 95% CI of LS Means	-3.79 (-5.96, -1.62)		

The sample size of the male group is about three times as large as the female group. The above table shows that efficacy results are similar in both groups. They are also similar to the whole population.

To see if age affects treatment effect, patients are separated into two age groups which are age groups of 6-9 and 10-12. Subgroup analyses on age groups are conducted using the mixed effects model. The effect of age on the treatment effect is first evaluated by testing the significance of age group as a factor and testing the treatment effect after the adjustment of age in the mixed effects model. The significance test of age group in the model gives a p-value of 0.0006. MTS is highly significant after the adjustment of age group. The treatment effects of MTS in the two age groups are given in Table 4.1.2.

Table 4.1.2 Subgroup Analyses of the Treatment Effect on the Primary Endpoint in Two Age Groups (LOCF Analysis)

Age Group		MTS (N=79)	Placebo (N=79)	p-value
6-9 Years		**N=45**	**N=45**	
	LS Mean (SE)	3.80 (0.80)	10.41 (0.80)	<0.0001

	Difference and 95% CI of LS Means	-6.61 (-8.09, -5.13)		
10-12 Years		**N=34**	**N=34**	
	LS Mean (SE)	2.48 (0.69)	4.86 (0.69)	0.0004
	Difference and 95% CI of LS Means	-2.38 (-3.56, -1.2)		

The sample size of the age group of 6-9 is larger than the age group of 10-12. The younger group also has a larger treatment effect than the older one as indicated in the above table.

To see if race affects treatment effect, patients are separated into two race groups: White and Non-white. There are 55 Whites (70%) and 24 Non-whites (30%). The effect of race group on the treatment effect is first evaluated by testing the significance of race group as a factor and then testing the treatment effect of MTS after the adjustment of race group in the mixed effects model. The significance test for race group in the model gives a p-value of 0.012. MTS is highly significant ($p<0.0001$) after the adjustment of race group. The treatment effects of MTS in the two race groups are given in Table 4.1.3.

Table 4.1.3 Subgroup Analyses of the Treatment Effect on the Primary Endpoint in Two Race Groups (LOCF Analysis)

Race Group		**MTS (N=79)**	**Placebo (N=79)**	**p-value**
White		**N=55**	**N=55**	
	LS Mean (SE)	2.82 (0.64)	6.72 (0.64)	<0.0001
	Difference and 95% CI of LS Means	-3.91 (-5.06, -2.76)		
Non-White		**N=24**	**N=24**	
	LS Mean (SE)	4.31 (1.24)	10.74 (1.24)	<0.0001
	Difference and 95% CI of LS Means	-6.43 (-8.97, -3.89)		

As Table 4.1.3 indicates that the white group has smaller primary outcome values both in the treatment and placebo groups. The treatment effect is smaller. The Non-white group has larger primary outcome values and also a larger treatment effect.

4.1.2 Study SPD485-302

During the statistical review, the effect of sex on the treatment effect is first evaluated by testing the significance of sex as a factor and testing the treatment effect after the adjustment of sex in the ANCOVA model. The significance test of sex in the model gives a p-value of 0.51. Both MTS and Concerta are still highly significant after the adjustment of sex. So sex does not seem to affect the significance of the treatment. To see if the treatment effects are the same in these two groups, we did a subgroup analysis in the two gender groups separately and the results are depicted in Table 4.2.1.

Table 4.2.1 Subgroup Analyses of the Treatment Effect on the Primary Endpoint in Two Sex Groups (LOCF Analysis)

Sex		**MTS (N=96)**	**Concerta (N=89)**	**Placebo (N=85)**
Male		**N=58**	**N=59**	**N=63**

	LS Mean (SE)	-24.1 (1.96)	-20.72 (1.95)	-10.2 (1.89)
	Difference and 95% CI of LS Means (Active-Placebo) **p-value**	-13.89 (-19.24, -8.54) <0.0001	-10.50 (-15.82, -5.20) 0.0002	
Female		**N=38**	**N=30**	**N=22**
	LS Mean (SE)	-24.3 (2.08)	-23.5 (2.36)	-10.6 (2.75)
	Difference and 95% CI of LS Means (Active-Placebo) **p-value**	-13.75 (-20.49, -7.01) 0.0001	-12.91 (-17.54, -8.28) 0.0007	

The sample size is twice as large in the male as in the female group. The above table shows that statistical significance effects are about the same in both groups. They are also about the same as the whole population.

To see if age affects treatment effect, patients are separated into two age groups which are age groups 6-9 and 10-12. Subgroup analyses on age groups are conducted using ANCOVA model. Age group as a factor in the overall ANCOVA model has a p-value of 0.36. The treatment effects of MTS in the two age groups are given in Table 4.2.2.

Table 4.2.2 Subgroup Analyses of the Treatment Effect on the Primary Endpoint in Two Age Groups (LOCF Analysis)

Age Group		**MTS (N=96)**	**Concerta (N=89)**	**Placebo (N=85)**
6-9 Years		**N=61**	**N=58**	**N=60**
	LS Mean (SE)	-24.7 (1.81)	-23.8 (1.86)	-10.4 (1.83)
	Difference and 95% CI of LS Means (Active-Placebo) **p-value**	-14.28 (-19.34, -9.22) <0.0001	-13.47 (-18.61, -8.33) <0.0001	
10-12 Years		**N=35**	**N=31**	**N=25**
	LS Mean (SE)	-23.4 (2.44)	-17.4 (2.60)	-10.1 (2.89)
	Difference and 95% CI of LS Means (Active-Placebo) **p-value**	-13.30 (-20.71, -5.89) 0.0007	-7.31 (-14.93, -0.31) 0.064	

The sample size of the age group of 6-9 is twice as large as the age group of 10-12. The two groups have similar treatment effect size of MTS. However, the treatment effect of Concerta is only significant in younger group which also has a larger treatment effect size than the older group.

To see if race affects treatment effect, patients are separated into two race groups: White and Non-white. There are 209 Whites (77%) and 61 Non-whites (23%). Treatment effect of MTS is analyzed using ANCOVA model with the race group as a factor which has a p-value of 0.70 in the ANCOVA analysis. The significance level of MTS and Concerta are all below 0.0001 after the adjustment of race group. Subgroup analysis is avoided due to such results.

4.2 Other Special/Subgroup Populations

Not Available.

5. SUMMARY AND CONCLUSIONS

5.1 Statistical Issues and Collective Evidence

5.1.1 Study SPD485-201

This was a phase II, randomized, double-blind, multi-center, placebo-controlled, analog classroom, crossover study, to evaluate the efficacy of MTS in treating the children (aged 6-12) diagnosed with ADHD using the SKAMP deportment scale as the primary endpoint. With a sample size of 79 in ITT population, statistical analysis using a mixed effects linear model indicates that MTS is highly statistically significant. The sponsor did not check the model assumptions in the statistical analyses as required in the SAP. There are evidences indicating that some model assumptions are violated. However, results using nonparametric models by the reviewer still support the claim that the treatment MTS is effective in reducing the SKAMP deportment score among children with ADHD.

Further analyses on the SKAMP deportment score at Hours 2 and 3 indicate that the treatment seems to have started the effect at the end of Hour 2, with p-values of 0.0467. Without the data at Hour 1, it's hard to give a better estimate of the real starting time of the drug effect.

Despite the positive efficacy results, the reviewer has two major concerns about the study. The first is that the baseline measurement of the primary endpoint was not taken in the study, therefore it couldn't be adjusted in the statistical model. The baseline measurement is meant to be the measurement at the end of Week 7, before the randomization of the crossover study. The second concern is that right after the dose optimization period, the patients were directly randomized into treatment and placebo groups. Those patients randomized to placebo group did not go through a tapering period before changing to placebo. Therefore it's unclear if the observed treatment effect was real or was due to the sudden change of treatment.

5.1.2 Study SPD485-302

This is a Phase III, randomized, double-blind, placebo-controlled, dose optimization study to compare MTS with placebo in children (aged 6-12) diagnosed with ADHD using the ADHD-RS-IV total score as the primary endpoint. With 270 subjects in the ITT population, the ANCOVA analysis indicates that MTS is highly statistically significant compared to placebo in reducing the ADHD-RS-IV total score. With model assumptions being violated, the reviewer applies the rank ANCOVA model to the data set. This analysis gives p-values of <0.0001 and 0.0156 in LOCF and OC analyses. Given the total patient dropout being about 40%, the reviewer uses the MMRM method, which takes the missingness into consideration using the assumption of non-informative dropout in the analysis of treatment. This analysis gives a $p<0.0001$ in the test of the efficacy of MTS. All the results support the sponsor's claim of the effectiveness of MTS in treating children with ADHD.

5.2 Conclusions and Recommendations

In this submission, the sponsor conducted two pivotal clinical trial studies, a Phase II, placebo controlled, randomized, crossover study and a Phase III, randomized, placebo controlled study with reference of CONCERTA®. These studies evaluated the efficacy and safety of MTS over placebo on children (aged 6-12) with ADHD. Both studies are evaluated in this review.

In Study SPD485-201, the reviewer's statistical analyses confirm the sponsor's efficacy results and support their claim of the efficacy of MTS in the treatment of children with ADHD. The drug effect seems to have started at the end of the second hour. Despite such positive evidences, we have two major concerns in the conduct of this study that add uncertainty to the validity of the claim of the sponsor. The first concern is that the baseline measurement of the primary endpoint was not taken in the study therefore could not be adjusted in the statistical analyses; the second is that the patients in placebo group did not go through a tapering period before changing to placebo, therefore it's unclear if the observed treatment effect was real or was due to the sudden change of treatment. In Study SPD485-302, the reviewer's statistical analysis results also confirm the sponsor's efficacy results and support their claim of the effectiveness of MTS in the treatment of children with ADHD.

C:\Data\My Documents\NDA Review\NDA 2005\21514

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

/s/

Fanhui Kong
10/24/2005 12:45:04 PM
BIOMETRICS

Peiling Yang
10/24/2005 12:46:32 PM
BIOMETRICS

James Hung
10/24/2005 01:08:03 PM
BIOMETRICS