

**CENTER FOR DRUG EVALUATION AND
RESEARCH**

APPLICATION NUMBER:
21-514

STATISTICAL REVIEW(S)



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	3
1.1 CONCLUSIONS AND RECOMMENDATIONS.....	3
1.2 BRIEF OVERVIEW OF CLINICAL STUDIES.....	3
1.3 STATISTICAL ISSUES AND FINDINGS.....	4
1.3.1 Study SPD485-201.....	4
1.3.2 Study SPD485-302.....	4
2. INTRODUCTION	4
2.1 OVERVIEW.....	4
2.2 DATA SOURCES.....	6
3. STATISTICAL EVALUATION	6
3.1 EVALUATION OF EFFICACY.....	6
3.1.1 Study SPD485-201.....	6
3.1.1.1 Title and Study Objectives.....	6
3.1.1.2 Study Design and Endpoints.....	6
3.1.1.3 Primary and Secondary Endpoints.....	7
3.1.1.4 Patient Disposition, Demographic and Baseline Characteristics.....	8
3.1.1.5 Statistical Methodologies Used.....	9
3.1.1.6 Results by the Sponsor.....	9
3.1.1.6.1 Primary Endpoint Results.....	9
3.1.1.6.2 Secondary Endpoint Results.....	10
3.1.1.7 Reviewer's Comments and Findings.....	11
3.1.1.7.1 Efficacy Results.....	11
3.1.1.7.2 Further Statistical Analyses.....	11
3.1.1.7.3 Statistical Issues.....	12
3.1.2 Study SPD485-302.....	12
3.1.2.1 Title and Study Objectives.....	12
3.1.2.2 Study Design and Endpoints.....	12
3.1.2.3 Primary and Secondary Endpoints.....	13
3.1.2.4 Patient Disposition, Demographic and Baseline Characteristics.....	13
3.1.2.5 Statistical Methodologies Used.....	15
3.1.2.6 Results by the Sponsor.....	15
3.1.2.6.1 Primary Endpoint Results.....	15
3.1.2.6.2 Secondary Endpoint Results.....	16
3.1.2.7 Reviewer's Comments and Findings.....	17
3.1.2.7.1 Efficacy Results.....	17
3.1.2.7.2 Further Efficacy Analyses.....	18
3.1.2.7.3 Statistical Issues.....	18
3.2 EVALUATION OF SAFETY.....	19
4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS	19
4.1 GENDER, RACE AND AGE.....	19
4.1.1 Study SPD485-201.....	19
4.1.2 Study SPD485-302.....	20
4.2 OTHER SPECIAL/SUBGROUP POPULATIONS.....	21
5. SUMMARY AND CONCLUSIONS	22
5.1 STATISTICAL ISSUES AND COLLECTIVE EVIDENCE.....	22
5.1.1 Study SPD485-201.....	22
5.1.2 Study SPD485-302.....	22
5.2 CONCLUSIONS AND RECOMMENDATIONS.....	22

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 department 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 available 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 department 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 department score among children with ADHD.

Further analyses on the SKAMP department 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 III/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 evaluable 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 department 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 department 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				Total	
	MTS/PTS		PTS/MTS		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		Overall (N=79)
		MTS/PTS (N=41)	PTS/MTS (N=38)	
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 department 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 department 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 department 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 (\pm SE) SKAMP department 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 department scores was -4.8 , with a 95% confidence interval of $(-5.89, -3.64)$.

Table 3.1.1.3 Analysis of Mean SKAMP Department 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 (\pm SE) PERMP: Number of Math Problems Attempted score for MTS (113.8 \pm 6.39) was significantly higher ($p < 0.0001$) than that for PTS (86.2 \pm 6.39). The LS mean (\pm SE) PERMP: Number of Math Problems Correct score for MTS (109.4 \pm 6.34) was significantly higher ($p < 0.0001$) than that for PTS (80.7 \pm 6.34). The LS mean (\pm SE) PERMP: Sum of Number of Math Problems Attempted and Correct score for MTS (223.2 \pm 12.67) was significantly higher ($p < 0.0001$) than that for PTS (167.0 \pm 12.67).

The mean SKAMP Total scores were improved by the treatment of MTS in the ITT population. The MTS LS mean (\pm SE) (9.4 \pm 0.99) was significantly lower ($p < 0.0001$) than the PTS LS mean (\pm SE) (17.9 \pm 0.99). The mean ADHD-RS-IV Total scores were improved by the treatment of MTS in the ITT population. The MTS LS mean (\pm SE) (16.3 \pm 1.24) was significantly less ($p < 0.0001$) than the PTS LS mean (\pm SE) (32.7 \pm 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 (\pm SE) (20.2 \pm 2.11) was significantly lower ($p < 0.0001$) than the PTS LS mean (\pm SE) (35.3 \pm 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)	9.4 (0.99)	17.9 (0.99)	<0.0001
N	79	79	
ADHD-RS-IV total score			
LS Mean (SE)	16.3 (1.24)	32.7 (1.23)	<0.0001
N	78	79	
CPRS-R total score			
LS Mean (SE)	20.2 (2.11)	35.3 (2.21)	<0.0001
N	67	61	

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 Department 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 department 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%)	0 (0%)	0 (0.0%)	2 (0.7%)
	Other	8 (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 (\pm SE) change from baseline of ADHD-RS-IV total score for MTS (-24.2 \pm 1.45) was significantly lower ($p < 0.0001$) than that for placebo (-10.3 \pm 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 CRPS-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 while 76.5% 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 while 75.3% 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)	-13.89 (-19.24, -8.54)	-10.50 (-15.82, -5.20)	
	p-value	<0.0001	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)	-13.75 (-20.49, -7.01)	-12.91 (-17.54, -8.28)	
	p-value	0.0001	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)	-14.28 (-19.34, -9.22)	-13.47 (-18.61, -8.33)	
	p-value	<0.0001	<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)	-13.30 (-20.71, -5.89)	-7.31 (-14.93, -0.31)	
	p-value	0.0007	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 deopment 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 deopment score among children with ADHD.

Further analyses on the SKAMP deopment 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

**APPEARS THIS WAY
ON ORIGINAL**

**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



DEPARTMENT OF HEALTH AND HUMAN SERVICES
PUBLIC HEALTH SERVICE
FOOD AND DRUG ADMINISTRATION
CENTER FOR DRUG EVALUATION AND RESEARCH

STATISTICAL REVIEW AND EVALUATION

NDA: 21,514 (SE)

DRUG NAME: Methylphenidate Transdermal System (MTS)

INDICATION: ADHD in children

SPONSOR: Noven Pharmaceuticals

STATISTICAL REVIEWER: Tristan Massie, Ph.D. (HFD-710)

DATE OF DOCUMENT: July 10, 2002

HFD-710 Kun Jin, Ph.D., Team Leader
George Chi, Ph.D., Director

HFD-120 Anna Marie H. Weikel, R. Ph., Project Manager
Glenn Mannheim, M.D., Clinical Reviewer
Paul Andreason, M.D., Team Leader
Russell Katz, M.D., Director

HFD-700 Charles Anello, Sc. D., Deputy Director

1. Executive Summary Of Statistical Findings	3
1.1. Conclusion and Recommendations	3
1.2. Overview of the Clinical Program and Studies Reviewed	3
1.3. Principal Findings.....	3
2. Introduction and Background.....	5
2.1. Data Analyzed and Sources.....	5
3. Study N17-018.....	7
3.1. Objectives.....	8
3.2. Efficacy Endpoint.....	8
3.3. Number of Subjects and Analysis Plan.....	8
3.4. Sponsor's Results	9
3.4.1. Patient Disposition and Baseline Information.....	9
3.4.2. Primary Analysis.....	10
3.4.3. Secondary Analysis.....	12
3.4.4. Conclusions	12
3.5. Reviewer's Evaluation and Comments	13
3.5.1. Primary Analysis.....	13
3.5.2. Secondary Analyses	14
3.5.2.1. Observed Cases.....	14
3.5.2.2. Secondary Endpoints.....	15
3.5.3. Subgroup Analyses.....	16
3.5.4. Distribution of Patch Size	17
3.5.5. Exploratory Analyses.....	19
4. Study N17-010.....	19
4.1. Design and Objectives.....	19
4.2. Efficacy Endpoints.....	20
4.3. Number of Subjects and Analysis Plan.....	20
4.4. Sponsor's Results	21
4.4.1. Patient Disposition and Demographics	21
4.4.2. Primary Analysis.....	24
4.4.3. Secondary and Post-Hoc Analyses.....	26
4.5. Reviewer's Evaluation and Comments	27
4.5.1. Primary Analysis.....	27
4.5.2. Secondary Analyses	28
4.5.2.1. Observed Cases.....	28
4.5.2.2. Other Secondary and Post-Hoc Analyses.....	28
4.5.3. Subgroup Analyses.....	30
5. Summary and Conclusion.....	31
5.1. Summary.....	31
5.2. Conclusion	33

1. Executive Summary Of Statistical Findings

1.1. Conclusion and Recommendations

The first phase III study, N17-010, did not demonstrate efficacy on the primary endpoint, Teacher rated Inattentive/Overactivity Scale (however, efficacy was demonstrated on several secondary endpoints such as the Parent I/O and the Clinical Global Impression of Improvement). However, the second phase III study, N17-018, which utilized a wider dose range and an additional week of treatment did demonstrate efficacy on the same primary endpoint ($p < 0.0001$). For these reasons, it seems that most of the strength of evidence for efficacy of the MTS resides in study N17-018. A complicating issue is that N17-018 also had an increased rate of adverse events: 50% experienced anorexia in the MTS group compared to 2% for the placebo group, and 29% experienced insomnia in the MTS group compared to 5% for the placebo group.

1.2. Overview of the Clinical Program and Studies Reviewed

A total of 18 clinical studies were conducted to evaluate the efficacy and safety of the (Methylphenidate Transdermal System or MTS) for the treatment of Attention Deficit Hyperactivity Disorder (ADHD). All of the clinical studies with efficacy assessments were conducted in pediatric patients (6 - 13 years old) with ADHD and were carried out in the United States. The two studies most relevant to the assessment of efficacy were N17-010 and N17-018. These were Phase III, flexible dose titration studies conducted in the community classroom setting (Studies N17-010 and N17-018). Each included about 20 centers and randomized about 210 patients. Study N17-010 was a 3-week, multicenter, randomized, double-blind, placebo-controlled, parallel-group, dose titration trial. N17-009 and N17-015 were phase II crossover studies (N17-009 and N17-015) conducted in summer treatment day camps and had 36 and 27 patients respectively.

1.3. Principal Findings

Although efficacy was demonstrated on secondary endpoints in N17-010, MTS did not differentiate statistically from placebo TS on the primary efficacy endpoint [Teacher Inattention/Overactivity (I/O) Factor on the IOWA Conners Rating Scale]. The sponsor believes this failure was likely due to the use of an inadequate maximum dose and a subtherapeutic starting dose. Therefore, the following study (N17-018) was conducted with 1) higher starting doses dependent on prior ADHD treatment and body weight, 2) inclusion of higher maximal doses, and 3) an extra week of double-blind treatment, to allow for modification of dose and wear time. The latter study demonstrated the efficacy of MTS over placebo TS in the primary endpoint and nearly all secondary efficacy endpoints.

The protocols for N17-009 and N17-015 were written as if these studies were exploratory (e.g. no primary endpoint was specified for either protocol). Both of the phase II studies had complex crossover designs in the interests of having more observations from fewer

patients. N17-009 had 8 periods, 4 treatments (6.25 cm², 12.5 cm², 25 cm² MTS and placebo TS), and 2 times of administration. N17-015 (part A) had 24 periods and 4 treatments (12.5 cm², 25 cm², 37.5 cm² MTS and placebo TS). These studies are difficult to analyze because period (or carryover) effects can occur in crossover trials and the correlation between repeated observations from a single patient must be accounted for. These complications increase with the number of periods, which is high in this case. Also, both trials had some missing data and in crossover trials missing data makes results difficult to interpret. These difficulties were compounded by the fact that the behavior ratings were skewed towards the low end of the scale and, thus, violated the usual assumption of normality (particularly in N17-009). Therefore, although there was some evidence that the MTS groups were superior to placebo, several features of these studies make it less convincing. First, the lack of prior specification of a primary endpoint and the associated inflation of type I error due to the resulting need for many multiple comparisons. There were actually two layers of multiplicity because of the need to compare multiple doses to placebo for multiple endpoints. Second, the difficulties of analysis caused by the complex crossover designs. For these reasons, when the protocols for N17-009 and N17-015 were submitted for review the agency commented that these trials would be considered exploratory.

**APPEARS THIS WAY
ON ORIGINAL**

2. Introduction and Background

Attention Deficit Hyperactivity Disorder (ADHD) is the most prevalent psychiatric disorder in children. ADHD consists of a variety of behaviors and personality types. However, the principal characteristics associated with ADHD include distractibility, short attention span, disorganization, impulsivity, disorientation of time, and more commonly in children, hyperactivity. Methylphenidate (Ritalin[®] and generic equivalents) is the most commonly used agent to treat ADHD in children. The immediate release form of Methylphenidate, Ritalin[®], is widely used. However, it has a half life of only 2 to 3 hours so it must be administered three times daily to maintain efficacy. A sustained release formulation Ritalin-SR[®] was developed but it may not be as effective as the immediate release form and has not been widely adopted. Recently, several new sustained release formulations have been approved but whether any of these will become favored over multiple doses of Ritalin IR[®] remains to be seen.

Noven Pharmaceuticals, Inc. (Noven) has developed a transdermal delivery system containing methylphenidate in a multi-polymeric adhesive platform as an alternative means of delivering systemic methylphenidate with once-daily application of a patch.

The clinical development program for the methylphenidate transdermal system (MTS) consisted of 18 studies. The primary objectives of the MTS clinical development program were to demonstrate the safety and efficacy of MTS for the treatment of ADHD in pediatric patients and to establish appropriate titration-to-effect dosing guidelines. Two of the four adequate and well-controlled studies were Phase III, flexible dose titration studies conducted in the community classroom setting (Studies N17-010 and N17-018) and the other two were crossover studies (N17-009 and N17-015). These studies are summarized below.

2.1. Data Analyzed and Sources

This review focuses on studies N17-010 and N17-018.

Table 2.1.1 Efficacy Studies

Study	Study Design	Population	Patch Sizes	Primary Efficacy Endpoint
N17-018	Phase III, 4 week flexible dose titration study (multi-site, double blinded, placebo controlled, parallel group)	Age 6-12 106 patients MTS 105 patients placebo TS	6.25, 12.5, 18.75, 25, 37.5, 50 cm ² MTS and matching placebo	Teacher I/O Factor of IOWA-Conners Rating Scale at end of 4 th treatment week.
N17-010	Phase III, 3 week flexible dose titration study (multi-site, double blinded, placebo controlled, parallel group)	Age 6-12 101 patients MTS 109 patients placebo TS	6.25, 12.5, and 25 cm ² MTS and matching placebo	Teacher I/O Factor of IOWA-Conners Rating Scale at end of 3 rd treatment week.
N17-015	Phase II, Part A: (Monday-Thursday) 4 treatment daily crossover Part B: (Alternate Fridays) 3 treatment daily crossover	Age 6-12 27 patients	Part A: 12.5, 25, and 37.5 cm ² MTS, and placebo TS Part B: 18.75, 37.5 cm ² and placebo TS	none specified
N17-009	8 treatment, 8-day daily crossover	Age 6-13 36 patients		none specified
N17-002	single center, double blinded, randomized, placebo controlled, 3-treatment, 3 period crossover study	Age 6-9 11 patients	MTS 20 cm ² , Ritalin 10 mg TID, and placebo (TS and capsule)	none designated
N17-003	Single center, double blinded, randomized, placebo controlled, 5-treatment, 5-period, dose ranging crossover	Age 6-10 13 patients	MTS 2.5 cm ² , 5 cm ² , 10 cm ² , 20 cm ² , and placebo TS	none designated

The data for these studies is accessible at the following network location:
\\CDSESUB1\N21514\N_000\2002-06-27

3. Study N17-018

This was a multicenter, randomized, double-blind, parallel-group, dose titration, placebo-controlled study conducted between October 23, 2001 and March 5, 2002. A total of 21 sites participated in the study. The MTS or placebo TS was applied once daily to the hip upon awakening in the morning and were to remain in place for up to 12 hours.

Six visits were scheduled over approximately a 6-week period: a Screening Visit and a Baseline Visit to establish and then confirm patients' eligibility for the study, establish a diagnosis of ADHD, begin ADHD medication wash-out (Screening, Visit 1), and assign double-blind treatment (Baseline, Visit 2); and four visits to evaluate double-blind treatment (Evaluation, Visits 3 through 6).

Each patient was screened over a 2-week period to evaluate eligibility for the study. The screening (Visit 1) was followed by a 1-week wash-out period (5 to 7 days) for patients who were previously on methylphenidate (MPH) or other stimulant medications. Parents were asked to complete the computerized version of the National Institute of Mental Health Diagnostic Interview Schedule for Children (C-DISC Interview) unless the patient was already diagnosed with ADHD in the previous 12 months and was taking a stable dose of methylphenidate (MPH). At baseline (Visit 2), eligibility was confirmed, the teacher/parent Disruptive Behavior Disorder Rating Scale (DBD) was assessed, and in conjunction with the results from the screening C-DISC Interview, if performed, the investigator provided a final ADHD diagnosis. Patients then were assigned (1:1), based on a randomization schedule generated by ~~;~~, to one of two parallel arms: 1) either the 12.5 cm² or 18.75 cm² MTS (depending on current MPH regimen, or on body weight if naïve to MPH, i.e., <4 weeks on oral MPH at a stable daily dose) or 2) the matching 12.5 cm² or 18.75 cm² placebo TS. For patients who were controlled on MPH (taking a stable dose for ≥4 weeks) at ≤20 mg per day, the initial patch size was 12.5 cm² (delivering approximately 11 mg for a 12 hour wear period) and for those taking >20 mg but not more than 60 mg/day, the initial patch size was 18.75 cm² (delivering approximately 16 mg for a 12 hour wear period). If a patient was naïve to MPH, the starting patch size of MTS was 12.5 cm² or 18.75 cm² depending on whether or not the patient's body weight was <25 kg.

At the completion of the first week of double-blind medication (Visit 3), patients were evaluated to determine whether the dosage needed to be titrated up or down, or remain the same. At the completion of the second and third weeks of double-blind medication (Visits 4 and 5), patients were again evaluated to determine whether the dosage needed to be titrated up, titrated down, or remain the same as the previous week. The range of allowable patch sizes was 6.25 to 50 cm². Weekly titration of dose was based mostly on Teacher Inattention/Overactivity (I/O ratings), the Clinical Global Impression (CGI) scores, and parent and clinician assessment of safety and patch tolerance.

The Pittsburgh Modified Conners Rating Scale, which incorporates the IOWA Conners Rating Scale, was completed by the classroom teacher once per week at the end of the school day on Friday and by the parent once per week on Thursday or Friday evenings prior to the next scheduled visit. This score was based on the child's performance over the course of that week.

Patients who qualified for randomization were assigned to a treatment group according to a center-specific randomization number. _____ provided the pre-determined randomization schedule. A block size of four was used for each center-specific randomization schedule.

3.1. Objectives

The primary objective was to assess the safety and efficacy of Noven Methylphenidate Transdermal System (MTS), in comparison to placebo, in children 6 to 12 years of age diagnosed with Attention Deficit Hyperactivity Disorder (ADHD). The secondary objective was to assess skin tolerance and patch adhesiveness of MTS.

3.2. Efficacy Endpoint

The primary efficacy endpoint was the teacher rated scores on the I/O Factor of the IOWA-Conners Rating Scale obtained at the end of the school day on Friday of the last week of the double-blind period. Five items comprise the I/O Factor and the possible scores range from 0 ("not at all") to 3 ("very much") for each item. Secondary efficacy was based on the parent rated scores on the I/O Factor of the IOWA-Conners Rating Scale (which were taken on thursday or friday evenings before the next visit); the Abbreviated Conners Rating Scale, Peer Relations Factors, Effective Normalization Factors (Teacher Only) and the teacher and parent rated scores on the O/D Factor of the IOWA-Conners Rating Scale; and the Clinical Global Impression (CGI) ratings.

3.3. Number of Subjects and Analysis Plan

A total of 172 patients (86 per treatment) were sought to be enrolled. The sample size calculation was based on the following assumptions: a mean difference of 2.5 units between the groups on the I/O Factor is clinically meaningful; standard deviation of 5 (supported by the literature), two sided significance level of 0.05; 90% power.

All efficacy analyses were to be performed on the intent-to-treat (ITT) population. The ITT population was defined as all randomized subjects with baseline data who received at least one patch of double-blind study drug and provided at least one post-baseline efficacy assessment.

Change from baseline to the last visit of the Teacher I/O Factor was to be compared between the placebo and MTS groups using the analysis of covariance (ANCOVA) technique. Baseline I/O score and treatment were to be included in the model. For analysis purposes, the last observation was to be carried forward (LOCF) for patients who withdrew prior to the end of the study.

Analyses of all secondary endpoints except the global impression scale were to be carried out in a similar fashion. The global impression scale was to be analyzed by techniques

appropriate for ordered categorical data, including a Cochran-Mantel-Haenszel test adjusting for other factors, if necessary.

3.4. Sponsor's Results

3.4.1. Patient Disposition and Baseline Information

Table 3.4.1 gives the details about the population loss and retention. A total of 212 patients were randomized to the double-blind treatment phase. 72 of these did not complete the study. The most frequent reason for withdrawal was lack of efficacy. The placebo TS group had a substantially larger number of such withdrawals (49 (46%) to 8 (8%)). Patient (19/06), randomized to placebo, and patient (19/07), randomized to the drug, accidentally had their medications switched after the first treatment week. However, the sponsor considered them both in the MTS group for analysis purposes.

Table 3.4.1 Disposition of Screened Patients

Patient Accounting	Randomized Treatment		Total
	MTS	TS	
Screened			268
Screen Failures			56
Randomized	106	106	212
Treated	106	105	211
Completed all visits	91	49	140
Discontinued due to:	15	56	71
Adverse Events	4	3	7
Protocol Violation	2	0	2
Lack of Efficacy	8	49	57
Lost to Follow-up	0	1	1
Other	1	3	4
# of patients for assessment			
Efficacy (ITT)	106	105	211

Table 3.4.2 contains demographic and baseline information for the ITT population. The distribution of age, gender, race, weight, height, diagnosis, comorbid conditions, and baseline CGI-S appear comparable across the treatment groups. Some notable demographic statistics are: 71% of the population was male, 69% was Caucasian, and 86% had combined subtypes of ADHD.

Table 3.4.2 Baseline Demographic Characteristics – Intent-to-Treat Patients

Demographic characteristic	Treatment Group		Test for Difference p-value	Total (211)
	MTS (106)	TS (105)		
Age (yr):	8.5 ± 1.8	8.8 ± 1.8	0.7860	8.7 ± 1.8
Gender: Male	76 (69.8%)	76 (72.4%)	0.6170	150 (71.1%)
Female	32 (30.2%)	29 (27.6%)		61 (28.9%)
Race: Caucasian	72 (67.9%)	74 (70.5%)	0.7349	146 (69.2%)
Black	20 (18.9%)	20 (19.0%)		40 (19.0%)
Asian	0 (0.0%)	1 (1.0%)		1 (0.5%)
Hispanic	8 (7.5%)	6 (5.7%)		14 (6.6%)
Other	6 (5.7%)	4 (3.8%)		10 (4.7%)
Weight (lb) mean	72.7	76.6		0.3059
Height (in) mean	52.4	53.6		53.0
Diagnosis:				
Inattentive	11 (10.4%)	14 (13.3%)	0.9116	25 (11.8%)
Hyperactive/Impulsive	2 (1.9%)	3 (1.9%)		4 (1.9%)
Combined	93 (87.7%)	89 (84.8%)		182 (86.3%)
Comorbid condition:				
Oppositional Defiant	43 (40.6%)	31 (29.5%)	0.5261	74 (35.1%)
Conduct Disorder	1 (0.9%)	1 (1.0%)		2 (0.9%)
Combined (OD/C)	25 (23.6%)	26 (24.8%)		51 (24.2%)
Did not meet criteria	37 (34.9%)	47 (44.8%)		84 (39.8%)
CGI-S				
Normal	0 (0.0%)	0 (0.0%)	0.6781	0 (0.0%)
Borderline	0 (0.0%)	0 (0.0%)		0 (0.0%)
Mildly III	0 (0.0%)	2 (1.9%)		2 (0.9%)
Moderately III	57 (53.8%)	54 (51.4%)		111 (52.6%)
Markedly III	35 (33.0%)	36 (34.3%)		71 (33.6%)
Severely III	12 (11.3%)	11 (10.5%)		23 (10.9%)
Most Extremely III	1 (0.9%)	0 (0.0%)		1 (0.5%)
Missing	1 (0.9%)	2 (1.9%)		3 (1.4%)
Teacher I/O Factor				
N	101	103		204
Mean	10.9	10.5		10.7
s.d.	3.2	3.8		3.5
Median	11.0	11.0		11.0
Minimum	0.0	0.0		0.0
Maximum	15.0	15.0	0.3249	15.0

3.4.2. Primary Analysis

The average baseline teacher I/O was 10.9 for the MTS group and 10.5 for the placebo group. Table 3.4.3 shows that group differences in the change from baseline in the Teacher I/O were significant as early as the end of the first week of treatment and persisted through the final week of treatment. The average group difference progressed

from 2.4 after the first week of treatment to 4.1 points after the fourth week. The differences at the end of the study exceed 2.5, which was deemed clinically relevant.

Table 3.4.3 Change From Baseline in Teacher I/O Factor by Visit (ITT-E Population)

Teacher I/O Factor Score	Treatment Group		Test for Difference p-value
	MTS (n=103)	TS (n=104)	
Change from Baseline to Visit 3			
N	99	98	
Mean	-3.2	-0.8	
s.d.	3.8	2.7	
Median	-3.0	0.0	
Minimum	-11	-9	
Maximum	5	5	
Within-Group p-value	<0.0001	0.0035	
95% Confidence Interval	-3.9 -2.4	-1.4 -0.3	
LSMEANS	-3.131	-0.9554	<0.0001
Change from Baseline to Visit 4			
N	98	95	
Mean	-4.5	-1.1	
s.d.	4.4	2.8	
Median	-4.5	-1.0	
Minimum	-12	-10	
Maximum	5	4	
Within-Group p-value	<0.0001	0.0003	
95% Confidence Interval	-5.4 -3.6	-1.7 -0.5	
LSMEANS	-4.515	-1.231	<0.0001
Change from Baseline to Visit 5			
N	92	68	
Mean	-4.8	-1.3	
s.d.	4.4	3.4	
Median	-4.0	-2.0	
Minimum	-14	-9	
Maximum	7	5	
Within-Group p-value	<0.0001	0.0029	
95% Confidence Interval	-5.7 -3.9	-2.1 -0.4	
LSMEANS	-4.363	-1.485	<0.0001
Change from Baseline to Visit 6			
N	78	47	
Mean	-6.1	-2.0	
s.d.	3.5	3.1	
Median	-6.0	-2.0	
Minimum	-14	-10	
Maximum	0	5	
Within-Group p-value	<0.0001	<0.0001	
95% Confidence Interval	-6.9 -5.3	-2.9 -1.1	
LSMEANS	-5.776	-2.082	<0.0001

Teacher I/O Factor Score	Treatment Group		Test for Difference p-value
	MTS (n=103)	TS (n=104)	
Change from Baseline to Visit Final (LOCF)			
N	101	102	
Mean	-5.3	-1.1	
s.d.	4.0	3.2	
Median	-5.0	-1.0	
Minimum	-14	-10	
Maximum	5	5	
Within-Group p-value	<0.0001	0.0006	
95% Confidence Interval	-6.1 -4.5	-1.8 -0.5	
LSMEANS	-5.153	-1.106	<0.0001

3.4.3. Secondary Analysis

Group differences were also apparent for most secondary endpoints, including the Parent I/O, as seen in Table 3.4.4.

Table 3.4.4 Change From Baseline in Parent I/O Factor by Visit (ITT-E Population)

Parent I/O Factor Score	Treatment Group		Test for Difference p-value
	MTS (n=103)	TS (n=104)	
Change from Baseline to Final Visit (LOCF)			
N	102	104	
Mean	-6.5	-1.8	
s.d.	3.8	3.6	
Median	-6.5	-1.0	
Minimum	-15	-11	
Maximum	3	7	
Within-Group p-value	<0.0001	<0.0001	
95% Confidence Interval	-7.3 -5.8	-2.5 -1.1	
LSMEANS	-6.513	-2.138	<0.0001

3.4.4. Conclusions

In contrast to the efficacy results, the safety profile was somewhat troubling. There were notable numbers of insomnia and anorexia related adverse events and group differences in their rates of occurrence in study N17-018:

50% experienced anorexia in the MTS group compared to 2% for the placebo group, and 29% experienced insomnia in the MTS group compared to 5% for the placebo group.

The sponsor suggests that, in an effort to reduce the incidence of AEs, the initial starting dose of MTS should be limited to 12.5 cm². This recommendation is based on the observation that there were fewer AEs reported in study N17-010 and fewer AEs in the

present study at Week 1 when the 12.5 cm² patch was the starting dose. The sponsor claims that these results suggest that patients should be started on the lower dose (12.5 cm²) and gradually titrated upward to achieve maximum balance between efficacy and safety.

3.5. Reviewer's Evaluation and Comments

3.5.1. Primary Analysis

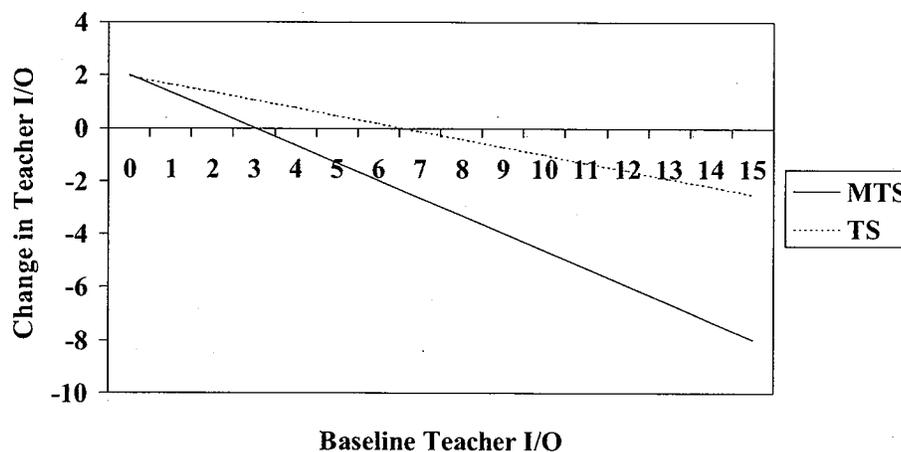
Eighty-six patients per group were sought to allow for 80% power to detect a group mean difference of 2.5 units in the Teacher I/O, assuming a standard deviation of 5 and a two sided significance level of 0.05. However, 212 were randomized and 205 were able to be included in the primary analysis. In addition, the observed standard deviation was smaller (in both studies) than initially estimated. This suggests that the power to detect the group mean difference of 2.5 units may have been 95% or more.

This reviewer performed the sponsor's primary analysis according to the protocol (two way ANCOVA model using change from baseline in the Teacher's I/O factor and adjusting for baseline Teacher I/O factor) and obtained the significant overall treatment effect ($p < 0.0001$). A slight inconsistency in the sponsor's derived data and the reviewer's created data was found but the results were not affected.

This reviewer found that there is a significant interaction ($p = 0.0063$) between baseline Teacher I/O score and treatment, i.e. the slope of the line relating change in Teacher I/O to baseline depends on the treatment group. This implies that differences between the groups vary depending on the baseline score as seen in Figure 3.5.1 and the treatment is more effective for those most severely affected at baseline. The differences in the model based mean changes are significant in favor of MTS for baseline Teacher I/O scores ≥ 5 . The difference in slopes might be explained by the limited range of the Teacher I/O factor scale, i.e., the lower the baseline score the less room there is for improvement and for a baseline score of 0 there is no room for improvement.

**APPEARS THIS WAY
ON ORIGINAL**

Figure 3.5.1



To further investigate, this reviewer performed a Wilcoxon test for the percent change from baseline in the Teacher I/O. The Wilcoxon test, a nonparametric method, was chosen over an analysis of covariance (ANCOVA) because percent change frequently does not satisfy the assumption of normality that is needed for an ANCOVA. The p-value associated with the Wilcoxon test for a treatment effect is $p < 0.0001$. Thus, the treatment effect shows some robustness and the interaction is not a cause for concern.

3.5.2. Secondary Analyses

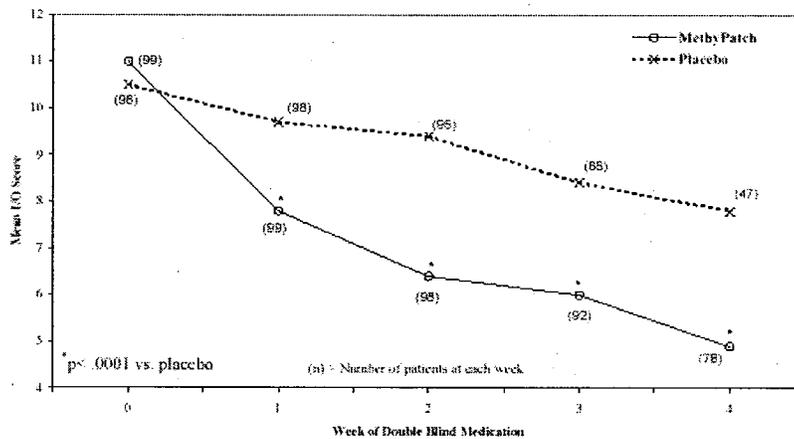
3.5.2.1. Observed Cases

Of the 211 treated patients, 140 (66.4%) completed study N17-018. Substantially more patients withdrew from the placebo group (56/105 (53.3%) compared to 15/106 (14.2%)). The primary reason given for withdrawal from the placebo group was lack of efficacy (49/105 (46.7%)). This would mean that more patients in the placebo group had their last observations carried forward in the LOCF analysis, which, considering the observed trend of improvement over time (seen in the figure below), might benefit the MTS group. On the other hand, the subset of patients who completed the study could be biased in favor of the placebo group because it would have a higher proportion of placebo responders than the original group. This reviewer found that only 119 (77 MTS; 42 TS) of the 211 treated patients had Teacher I/O evaluations for all weeks. However, the treatment difference is still highly significant for the observed cases ($p < 0.0001$). The changes from baseline were -5.87 for MTS and -2.26 for placebo. The agreement of the conclusions drawn from the LOCF and Observed Cases analyses suggest that the treatment effect is robust despite the substantial number of dropouts.

Figure 3.5.2 Comparison of Results for LOCF and Observed Cases analyses

	MTS		Placebo TS		Test for Difference p-value
	N	Mean	N	Mean	
LOCF	103	-6.51	104	-2.13	<0.0001
Observed Cases	77	-5.87	42	-2.26	<0.0001

Figure 3.5.3 Mean Teacher I/O Score Over Time



3.5.2.2. Secondary Endpoints

This reviewer also verified the secondary analyses involving the Parent I/O and the Clinical Global Impression of Improvement for the ITT population. The MTS group was rated significantly more improved with respect to both of these endpoints at the end of the study (p-values were both < 0.0001). For the Parent I/O the mean changes were -5.3 for MTS and -1.1 for Placebo TS. The mean CGI-Is were 2.3 for MTS and 3.8 for placebo TS (81.5% of MTS were at least minimally improved compared to 27.2% for placebo TS).

3.5.3. Subgroup Analyses

The following table gives the baseline Teacher I/O and the change in the Teacher I/O at the end of the 4th treatment week for several subgroups. If we include group and group by treatment interaction effects in the ANCOVA model and test whether the interaction effect is needed we obtain the p-value shown in the table. The treatment effect does not appear to depend on gender or race. However, the p-value for the treatment by age group interaction (p=0.07) suggests that the treatment effect may be larger for individuals in the 6-9 age group than individuals in the 7-10 age group. Also, the p-value for the previous methylphenidate use by treatment interaction (p=0.066) suggests that the treatment effect may be less for patients naïve to MPH. Note though that both of these p-values are only marginally significant and that because the randomization was not stratified these subgroups may be unbalanced with respect to other important factors.

Table 3.5.1 Change in Teacher I/O for Subgroups

	Treatment Group								Inter- action p-value
	MTS				Placebo TS				
	baseline N	change Mean	baseline N	change Mean	baseline N	change Mean	baseline N	change Mean	
Sex									0.523
FEMALE	33	10.21	32	-4.97	29	8.69	29	-0.93	
MALE	69	11.39	68	-5.42	76	11.33	74	-1.34	
Race									0.462
AFRICAN-AMERICAN	19	11.11	18	-4.56	19	10.16	19	-1.37	
ASIAN					1	7.00	1	2.00	
CAUCASIAN	70	10.94	69	-5.24	75	10.67	73	-1.25	
HISPANIC	7	11.29	7	-6.57	6	9.67	6	-1.08	
OTHER	6	11.17	6	-6.33	4	13.75	4	-1.25	
Age Group									0.077
06-09	65	11.77	64	-5.97	50	11.36	50	-1.23	
10-13	37	9.68	36	-4.03	55	9.91	53	-1.23	
Any Pr ADHD Meds Taken									0.066
NO	50	11.06	49	-5.29	51	10.61	50	-2.11	
YES	52	10.96	51	-5.25	54	10.59	53	-0.40	

3.5.4. Distribution of Patch Size

As mentioned earlier this was not a fixed dose study and at the end of each week the investigator could decide to increase or decrease the patch size by one step for the following week. Such a decision was to be based on the weekly value of the Teacher I/O and the global impression of change. The table shows how many patients in the MTS group were using a particular patch size at each week of treatment. As the weeks went on, the patch size tended to be increased. As we might expect, this trend was even more dramatic in the placebo group. Because of the flexible titration it is not possible to determine an effect of dose beyond the first week of treatment. In the first week the patients were assigned to receive 12.5 or 18.75 cm² depending on previous use of methylphenidate and body weight, so a dose effect in the first week would be confounded with these factors. Change in the Teacher I/O at the end of the first week was significant ($p < 0.0001$) and the treatment effect was comparable in both dose groups. Therefore one might argue for the efficacy of these doses. Yet no effect was seen in study N17-010, which had 6.25, 12.5, and 25 cm² patch sizes, and testing the differences at the end of week 1 was not prespecified and therefore increases the overall type I error.

Table 3.5.2 Patch Size in MTS group by Week

Week	Patch Size N (%)					
	6.25	12.5	18.75	25	37.5	50
1	0 (0.0)	41 (38.7)	65 (61.3)	0 (0.0)	0 (0.0)	0 (0.0)
2	0 (0.0)	9 (8.7)	21 (20.4)	37 (35.9)	36 (35.0)	0 (0.0)
3	0 (0.0)	6 (6.2)	16 (16.5)	30 (30.9)	32 (33.0)	13 (13.4)
4	1 (1.1)	5 (5.4)	11 (12.0)	29 (31.5)	27 (29.3)	19 (20.7)

Table 3.5.3 presents the change in the Teacher I/O from baseline to last observation (LOCF) by last patch size. Note that the number of patients with each patch size in week 4 may differ from that in Table 3.5.2 because of the use of the last observation carried forward algorithm. When considering this table it should be kept in mind that patients were not randomized to these dose groups so any conclusions drawn from it are suspect. Also, the placebo group had significantly many more patients near the high end of the patch size range so the groups don't have equal numbers of patients at each patch size. The group difference is largest for a final patch size of 25 cm². In addition, MTS patients using the largest patch at the end had the smallest improvement, but this could be because lack of efficacy was the basis for increasing the patch size and, also, these individuals had a smaller baseline average and therefore less room for improvement.

Table 3.5.3 Change in Teacher I/O from baseline by Last Patch Size

	TREAT							
	MTS				Placebo TS			
	baseline		change		baseline		change	
	N	Mean	N	Mean	N	Mean	N	Mean
Last Patch								
6.25	1	13.00	1	-6.00	0		0	
12.5	5	12.80	5	-8.20	0		0	
18.75	14	10.86	14	-4.88	12	11.04	12	-2.38
25	34	11.60	34	-6.22	17	11.29	17	-0.24
37.5	29	11.13	29	-5.75	34	9.85	34	-1.34
50	20	9.60	20	-2.80	41	10.61	41	-1.00

It is also important to consider how the safety profile depends on the patch size. The flexible dose design makes this difficult to evaluate because there are different numbers of patients using each patch size. For example, more adverse events occurred when the patch size was 18.75 cm² than 12.5 cm², but most of the patients started with the 18.75 cm² patch (61.3%).

Week 1 of the study allowed for a comparison of the incidence of anorexia and insomnia with starting doses of 12.5 versus 18.75 cm². The incidence of insomnia was 5% (2/40) and 27% (18/66) with starting doses of 12.5 and 18.75 cm², respectively. For anorexia, the corresponding values were 27% (11/40) and 33% (22/66). The sponsor concludes that the starting dose should be at the lower end of the dose range for transdermal administration.

For anorexia, insomnia, emotional lability, nervousness, twitching, and somnolence, the time to first occurrence was calculated. Most tended to occur in the first of week of treatment and all occurred within 2 weeks of initiation of treatment. For anorexia, approximately 40% resolved on study and 60% were ongoing at study end. For insomnia, approximately 60% resolved on study and 40% were ongoing at study end.

3.5.5. Exploratory Analyses

A patient's initial patch size was dependent on two factors: previous use of ADHD medication and weight. In particular, if a patient was on a stable dose and, therefore, not naïve to ADHD medication, then the initial size was 12.5 or 18.75 depending on whether or not the pre-study dose was less than 20 mg. If a patient was naïve to ADHD medication then the initial patch size was 12.5 or 18.75 depending on whether or not the patient's weight was less than 25 kg. In each treatment group about 63 % of the patients started on the 18.75 cm² size. Therefore, it is important to determine whether either of these baseline factors is related to the change in the Teacher I/O from baseline. However, neither initial patch size or weight > 25 kg had an effect on the change in the Teacher I/O from baseline to the last visit.

For each patient the teacher and parent use the same questionnaire to determine the Teacher I/O and the Parent I/O so these measures should ideally be correlated. However, for the ITT patients at baseline the correlation between the measures was only 0.18 (and R²=0.03). This week correlation might be explained by the different relationships with the patient, different observation periods, and the fact that the questionnaires are filled out at different times of the day and different days of the week.

4. Study N17-010

4.1. Design and Objectives

This was a multi-center, randomized, double-blind, parallel-group, multi-dose, placebo-controlled, titration study (one week for Screening and ADHD medication wash-out plus three weeks for titration) that compared d1-threo-methylphenidate (administered daily as individually titrated dosages of the NovenTM 6.25, 12.5, and 25 cm² MTS) to matching placebo patches in ADHD patients. Utilizing a weekly parallel-arm, titration design, the study sought to compare the NovenTM Methylphenidate Transdermal System (MTS) – applied once daily to the hip in the morning as 6.25, 12.5, and 25 cm² patches, to matching placebo transdermal system (TS)-applied once-daily to the hip in the morning as 6.25, 12.5, and 25 cm² patches, in children diagnosed with ADHD. The parents were to remove the patch at bedtime.

A total of five visits were scheduled over approximately a four week period: a screen visit and a baseline visit to establish and then confirm patients' eligibility for the study (entry criteria), establish diagnosis of ADHD, begin ADHD medication wash-out (Screening, Visit 1), and assign double-blind treatment (Baseline, Visit 2); and three visits to evaluate double-blind treatment (Evaluation, Visits 3 to 5).

At Baseline (Visit 2), the patient was randomly assigned, on an equal basis, to one of two parallel arms for a one-week period: either the 6.25 cm² MTS or the matching 6.25 cm² placebo TS. At the completion of the first week (Visit 3), patients were evaluated to

determine whether the dosage needed to be titrated up or remain the same as the week prior. Likewise, at the completion of the second week of treatment (Visit 4), patients were evaluated to determine whether the dosage needed to be titrated up, titrated down, or remain the same as the week prior. At the completion of the third week of treatment (Visit 5), patients completed the study. Weekly evaluation of dosage was an overall clinical decision based on parent, teacher and site assessment of safety, efficacy, and patch tolerance. Weekly titration occurred in both the active methylphenidate and the placebo patch arms of the study by assigning patients to matching 6.25, 12.5, and 25 cm² patches. These patch sizes contain 13.8, 27.5, and 55.0 mg of methylphenidate. Study N17-010 started on 09-12-2000 and ended on 02-16-2001.

4.2. Efficacy Endpoints

The primary efficacy measure was the Teacher rated Inattentive/Overactivity (I/O) Factor of the IOWA-Conners Rating Scale. This scale was designed to measure the several components (factors) that make up the diagnostic category ADHD. The Inattention/Overactivity Factor was measured by items 1 through 5, which were ratings of the following behaviors: fidgeting, makes odd noises, excitable/impulsive, inattentive/distractible, and fails to finish/short attention span. The items are scored from 0 (not at all) to 3 (very much), so the I/O factor scores range between 0 and 15.

Secondary efficacy was assessed based on the parent rated I/O Factor of the IOWA-Conners Rating Scale; the teacher and parent rated Oppositional/Defiant (O/D) Factor of the IOWA-Conners Rating Scale; the Abbreviated Conners Rating Scale, Peer Relations Factors, and Effective Normalization Factors as administered by the Pittsburgh Modified Conners Rating Scale; and the Clinical Global Impression (CGI) ratings.

The IOWA-Conners Rating Scale was to be completed by the teacher on Thursday or Friday before the next scheduled visit and was to be based on the child's performance over that school week. The parent also was to complete the IOWA-Conners Rating Scale on Wednesday, Thursday, or Friday, before the next scheduled visit based on the child's performance over that school week. The CGI severity was to be assessed at baseline and the CGI improvement, based on improvements from baseline, was to be assessed at subsequent weekly visits.

4.3. Number of Subjects and Analysis Plan

The planned sample size for this study was 172 patients (86 patients per treatment arm). The sample size was estimated based on the Inattention/Overactivity (I/O) subscale of the IOWA-Conner Rating scale. Estimates of variability from the available literature were used in the sample size calculation and a standard deviation of 5 units was selected. Assuming clinical significance would be indicated between the MTS treatment group and the placebo group by a mean difference of 2.5 units or more, 86 patients per treatment group would be sufficient for a two-sided test at the 0.05 level with a power of at least 90% to detect this difference.

All efficacy analyses were to be performed on the intent-to-treat (ITT) population. The ITT population was defined as all randomized subjects with baseline data who received at least one patch of double-blind study drug and provided at least one post-baseline efficacy assessment. The secondary population involved the per-protocol population.

Change from baseline to the last visit of the Teacher I/O Factor was to be compared between the placebo and MTS group using the analysis of covariance (ANCOVA) technique. Baseline I/O score and treatment would be included in the model to compare the treatment differences. For patients who withdrew prior to the end of the double-blind treatment, the last observation was to be carried forward in the analysis (LOCF).

Analyses of all secondary endpoints except the global impression scale was to be carried out in a similar fashion. The global impression scale was to be analyzed by techniques appropriate for ordered categorical data, including a Cochran-Mantel-Haenszel test adjusting for other factors if necessary.

4.4. Sponsor's Results

4.4.1. Patient Disposition and Demographics

The primary efficacy analysis was to be based on all randomized patients who received at least one double-blind patch and provided at least one post-Baseline efficacy assessment. 213 of the 251 patients screened were randomized as shown Table 4.4.1. The remaining 38 screened patients did not meet the eligibility criteria. Of the 213 patients randomized, 210 received at least one study medication patch and were therefore included in the ITT population. 101 of these patients were assigned to MTS and 109 to placebo TS.

Table 4.4.1 Disposition of Screened Patients

	MTS	TS	Total
Number of Patients:			
Screened ^a			251
Screen Failures			38
Randomized	103	110	213
Treated (ITT Patients) ^b	101	109	210

a: Screened patients gave informed written consent.

b: Treated patients received at least one study medication patch.

Table 4.4.2 shows how many of the ITT patients did not complete all visits and compiles the various reasons for discontinuation from the study. Although the placebo group had a slightly higher dropout rate because of the complaint of lack of efficacy (5.5% to 1.0%), the treatment groups were fairly comparable with respect to percentages of dropouts and reasons for dropouts.

Table 4.4.2 Disposition of Treated Patients – Intent-to-Treat Patients (n, %)

	MTS (n=101)	TS (n=109)	Total (n=210)
Patients Evaluated for:			
Adverse Events	101 (100)	109 (100)	210 (100)
Routine Laboratories	93 (92.1)	96 (88.1)	189 (90.0)
Efficacy			
Intent-to-Treat (ITT-E)	101 (100)	108 (99.1)	209 (99.5)
Per Protocol (PPE)	91 (90.1)	95 (87.2)	186 (88.6)
Patient Completion Status:			
Completed All Visits	92 (91.1)	97 (89.0)	189 (90.0)
Discontinued Due to:	9 (8.9)	12 (11.0)	21 (10.0)
Adverse Event	3 (3.0)	2 (1.8)	5 (2.4)
Protocol Violation	2 (2.0)	3 (2.8)	5 (2.4)
Administrative	0 (0.0)	0 (0.0)	0 (0.0)
Lack of Efficacy	1 (1.0)	6 (5.5)	7 (3.3)
Lost to Follow-up	0 (0.0)	0 (0.0)	0 (0.0)
Death	0 (0.0)	0 (0.0)	0 (0.0)
Other	3 (3.0)	1 (0.9)	4 (1.9)

Table 4.4.3 compares baseline and demographic information for the two treatment groups. There were no significant differences between the two groups in terms of age, height, weight, gender, race, age of onset, IQ, comorbid diagnoses, Clinical Global Impressions-Severity, or baseline teacher I/O factor.

Table 4.4.3 Baseline Demographic Characteristics – Intent-to-Treat Patients

Demographic Characteristic	MTS (n=101)	TS (n=109)	p-value	Total (n=210)
Age (yr)				
Mean	8.7	8.6		8.7
s.d.	1.7	1.6		1.7
Median	9.0	8.0		8.0
Minimum	6	6		6
Maximum	12	12	0.9274	12
Body Weight (lb)				
Mean	77.1	78.4		77.8
s.d.	25.52	25.00		25.20
Median	71.0	73.0		72.0
Minimum	37.0	39.0		37.0
Maximum	160.0	177.0	0.7249	177.0
Height (in)				
Mean	53.3	53.1		53.2
s.d.	4.9	4.3		4.6
Median	53.0	53.0		53.0
Minimum	39.0	43.0		39.0
Maximum	66.2	64.0	0.8237	66.2
Age at ADHD Onset (yr)				
Mean	5.5	5.7		5.6
s.d.	1.8	2.0		1.9
Median	5.0	5.0		5.0
Minimum	1	<1		<1
Maximum	11	12	0.8090	12
Sex, n (%)				
Male	75 (74.3)	84 (77.1)		159 (75.7)
Female	26 (25.7)	25 (22.9)	0.6858	51 (24.3)
Ethnic Background, n (%)				
Caucasian	76 (75.2)	76 (69.7)		152 (72.4)
African-American	9 (8.9)	17 (15.6)		26 (12.4)
Asian	1 (1.0)	0 (0.0)		1 (0.5)
Hispanic	11 (10.9)	11 (10.1)		22 (10.5)
Other	4 (4.0)	5 (4.6)	0.5948	9 (4.3)
IQ				
N	101	108		209
Mean	103.4	101.1		102.2
s.d.	15.7	15.9		15.8
Median	103.0	100.0		101.0
Minimum	71	73		71
Maximum	142	146	0.3073	146
Subtype of ADHD^b, n (%)				
Inattentive	10 (9.9)	16 (14.7)		26 (12.4)
Hyperactive/Impulsive	2 (2.0)	5 (4.6)		7 (3.3)
Combined (Inattentive & Hyperactive/Impulsive)	89 (88.1)	88 (80.7)	0.5935	177 (84.3)
Comorbid Behavior Disorders^b, n (%)				
ODD	38 (37.6)	39 (35.8)		77 (36.7)
CD	1 (1.0)	0 (0.0)		1 (0.5)
Combined (ODD/CD)	25 (24.8)	20 (18.3)		45 (21.4)
Did Not Meet Criteria	37 (36.6)	50 (45.9)	0.3821	87 (41.4)

Table 4.4.3 (continued)

Demographic Characteristic	MTS (n=101)	TS (n=109)	p-value ^a	Total (n=210)
CGI Severity of Illness ^c , n (%)				
Not Assessed	1 (1.0)	0 (0.0)		1 (0.5)
Mildly ill	2 (2.0)	5 (4.6)		7 (3.3)
Moderately ill	63 (62.4)	66 (60.6)		129 (61.4)
Markedly ill	28 (27.7)	34 (31.2)		62 (29.5)
Severely ill	7 (6.9)	4 (3.7)	0.4521	11 (5.2)
Teacher I/O Factor				
N	98	103		201
Mean	9.5	9.4		9.4
s.d.	3.93	3.78		3.85
Median	10.0	10.0		10.0
Minimum	1.0	0.0		0.0
Maximum	15.0	15.0	0.6887	15.0

a: p-Values were based on ANOVA (continuous data) and Cochran-Mantel-Haenszel test (categorical data), with treatment group and center main effects, and treatment group-by-center interactive effects.

The p-values for the Teacher I/O Factor were based on two-way ANOVA models that included treatment group and center main effects and treatment group-by-center interaction effects.

b: ADHD subtypes and comorbid behavior disorders were based on teacher and parent DBD Rating Scales.

c: No patients were rated "Normal, not at all ill," Borderline mentally ill," or "Among the most extremely ill."

4.4.2. Primary Analysis

According to the protocol the primary analysis would compare the two groups with respect to change from baseline in the teacher I/O at endpoint (or LOCF) using an analysis of covariance model with baseline Teacher I/O and treatment effects. According to the statistical analysis plan, which dates from before the end of the study, center and center x treatment interaction effects were to be included in the ANCOVA. A significance level of 0.05 was adopted for main effects and 0.10 for interaction effects. Additionally, centers with less than two patients per treatment arm were to be pooled for the analysis. This led to a pooling of 7 of the 20 centers.

Table 4.4.4 summarizes the sponsor's primary analysis. The mean scores at baseline were very similar, 9.5 for MTS and 9.4 for TS, and the average reduction in the Teacher I/O at the last visit (or LOCF) was 2.3 for MTS and 1.5 for placebo. The p-value for the treatment x center interaction (0.0146) suggests that the treatment effect varied significantly among the centers. After adjusting for the interaction, the p-value associated with the test for a difference in the least squares group means is 0.7927. This indicates that we cannot reject the hypothesis that the mean change from baseline in the Teacher

I/O factor is the same for both groups. Thus, there was not a statistically significant difference between the MTS and placebo groups in change from baseline in ADHD symptoms, as measured by the primary endpoint. Nor were differences between the treatment groups detected for any of the secondary teacher rated measures.

Table 4.4.4 Summary of Patient Baseline, By Visit, and Change from Baseline in the Teacher I/O Factor Score – ITT Patients

Visit / Descriptor	Treatment Groups		Test for Difference - p-value ^a	Test for Interaction - p-value ^a
	MTS (n=98)	TS (n=101)		
Final Visit (LOCF)^c				
N	98	101		
Baseline				
Mean	9.5	9.4		
s.d.	3.9	3.8		
Median	10.0	10.0		
Minimum	1	0		
Maximum	15	15		
LOCF Assessment				
Mean	7.2	7.9		
s.d.	4.3	4.1		
Median	6.5	8.0		
Minimum	0	0		
Maximum	15	15		
Change from Baseline ^b				
Mean	-2.3	-1.5		
s.d.	4.0	3.4		
Median	-2.0	-1.0		
Minimum	-13	-13		
Maximum	8	5		
Lsmeans ^d	-1.738	-1.881	0.7927	0.0146

a: Test for difference and Test for interaction p-values were derived from the ANCOVA model which included baseline Teacher I/O, treatment, center, and center by treatment interaction effects.

b: Lower factor scores represent more acceptable behavior. Efficacy is assessed from change from baseline. (Change = Post-Baseline Value – Baseline Value). A reduction from baseline (negative change) represents improvement in behavior after treatment.

c: The 95% confidence intervals were based on the change from baseline to the follow-up assessment using the t-distribution.

d: Least Squares Means (Lsmeans) change is derived from the ANCOVA model and is the basis for group comparison.

4.4.3. Secondary and Post-Hoc Analyses

In Table 4.4.5 we see that, in contrast to the results for the primary endpoint (Teacher I/O factor), a significant difference was found between the treatment groups with respect to change from baseline in the Parent rated I/O factor ($p < 0.0001$). The baseline Parent I/O factors were similar for the two groups, 10.8 for MTS and 10.5 for TS. By the last visit (LOCF), the Parent I/O was reduced on average by 4.4 for the MTS group and 2.3 for the placebo TS group.

Table 4.4.5 Summary of Patient Baseline, By Visit, and Change from Baseline in the Parent I/O Factor Score – ITT-E Patients

Visit / Descriptor	Treatment Groups		Test for Difference - p-value ^a	Test for Interaction - p-value ^a
	MTS (n=99)	TS (n=105)		
Final Visit (LOCF)^c				
N	99	105		
Baseline Comparison				
Mean	10.8	10.5		
s.d.	2.7	3.1		
Median	11.0	10.0		
Minimum	4	2		
Maximum	15	15		
LOCF Assessment				
Mean	6.4	8.3		
s.d.	3.7	3.7		
Median	5.0	9.0		
Minimum	0	0		
Maximum	15	15		
Change from Baseline				
Mean	-4.4	-2.3		
s.d.	4.2	3.7		
Median	-5.0	-2.0		
Minimum	-13	-12		
Maximum	7	8		
Lsmeans ^d	-4.308	-2.315	0.0001	0.4860

a: Test for difference and Test for interaction p-values were derived from the ANCOVA model which included baseline Teacher I/O, treatment, center, and center by treatment interaction effects.

b: Lower factor scores represent more acceptable behavior. Efficacy is assessed from change from baseline. (Change = Post-Baseline Value – Baseline Value). A reduction from baseline (negative change) represents improvement in behavior after treatment.

c: The 95% confidence intervals were based on the change from baseline to the follow-up assessment using the t-distribution.

d: Least Squares Means (Lsmeans) change is derived from the ANCOVA model and is the basis for between-group comparison.

e: LOCF is the last observation carried forward or the final assessment during the double-blind treatment phase.

Likewise, the MTS group was found to be significantly more improved than the placebo group with respect to the Clinical Global Impression of Improvement score, which ranges from 1 (very much improved) to 7 (very much worse).

The sponsor also conducted several post-hoc analyses. First, the sponsor reports that if we consider the subgroup of patients who had a baseline Teacher I/O score ≥ 10 (the median value for both treatment groups), then we find a significant difference ($p=0.0399$) between the treatment groups, in favor of the MTS group, in terms of the change in the Teacher I/O from baseline to LOCF. The LS means are -3.760 for MTS and -2.272 for placebo. The sponsor argues that patients in the subgroup with baseline Teacher I/O score < 10 had less room for improvement due to the lower limit of the scale.

Second, the sponsor looked at subgroups defined by baseline teacher I/O score ≥ 10 and final patch size. Despite the small size of the subgroup (12 MTS; 6 TS) the sponsor claims that a significant difference ($p=0.021$) between the treatment groups is found for patients with a baseline teacher I/O ≥ 10 and a final patch size of 12.5 cm². However, no significance was found for the larger subset of patients with a baseline score ≥ 10 and a final patch size of 25 cm² ($p=0.13$). Nor were any differences found for any of the 3 final patch size subgroups of patients with baseline score < 10 .

Third, the sponsor investigated the effects of prior use of ADHD medication (Yes or No) on the change from baseline in the teacher I/O for the subgroups of patients with baseline score ≥ 10 and < 10 respectively. Significant prior use effects ($p=0.012$ in favor of No) were found for patients who had a baseline score > 10 . Thus, it is claimed that the positive effects of treatment were most apparent in naïve MTS patients who had higher teacher I/O Factor scores (10-15) at baseline than in any other patient group.

4.5. Reviewer's Evaluation and Comments

4.5.1. Primary Analysis

The primary analysis specified in the original protocol (ANCOVA with baseline Teacher I/O as the covariate) was changed in the statistical analysis plan to include center and treatment x center interaction effects in the ANCOVA model. This latter analysis was the one reported. The test for a group difference based on an ANCOVA model which included center and group by center interaction effects was $p=0.7927$. This reviewer performed both analyses. In the protocol specified analysis which ignores center no significant difference between the treatment groups was observed ($p=0.14$). This reviewer also confirmed the sponsor's analysis which incorporated center and center x treatment interactions in the ANCOVA model. As the sponsor reported, a significant center x treatment interaction was found ($p=0.015$). This indicates that the treatment effect varied from center to center and is substantiated by the fact that in 8 of the 20 centers the mean reduction in the Teacher I/O from baseline was numerically larger for the placebo TS group. Two of these eight centers ranked in the top three in terms of total number of patients treated. In summary, we cannot conclude that MTS is more effective than placebo TS in reducing ADHD symptoms as measured by the Teacher I/O.

4.5.2. Secondary Analyses

4.5.2.1. Observed Cases

The same conclusions were reached for the per-protocol group ($n_{TS}=95$; $n_{MTS}=91$). Table 4.5.1 presents the sponsor's analysis of the change in Teacher I/O from baseline to last visit (LOCF) for the per-protocol population (and the ITT population for comparison). Notice that 24 patients were not included in the analysis. This reviewer was not able to determine the identities of these patients and found a p-value of 0.096 for the test of a treatment difference in the Observed Cases population based on 87 MTS patients and 89 placebo patients (ignoring center and group x center interaction effects which were not specified in the protocol). Thus, the results in the per-protocol and observed cases populations agree with those found for the ITT population, i.e., the hypothesis of no treatment difference for change in the Teacher I/O from baseline to last visit is not rejected in any of these populations.

Table 4.5.1 Sponsor's ITT (LOCF) and Per Protocol Analyses for Teacher I/O

		MTS	Placebo TS	Test for Difference p-value ^a	Test for Interaction p-value ^a
LOCF	n	98	101		
	Mean Change	-2.3	-1.5		
	STD	4.0	3.4		
	LS Means	-1.738	-1.881	0.7927	0.0146
Per Protocol	n	78	84		
	Mean Change	-2.5	-1.8		
	STD	3.7	3.5		
	LS Means	-2.379	-1.865	0.3514	0.1134

a: Test for difference and Test for interaction p-values were derived from the ANCOVA model which included baseline Teacher I/O, treatment, center, and center by treatment interaction effects.

4.5.2.2. Other Secondary and Post-Hoc Analyses

The significant treatment difference ($p<0.0001$) for the change in the Parent I/O from baseline to last visit (LOCF) was also verified. The average change was -4.4 for the MTS group and -2.3 for the placebo TS group. Note that the group difference is less than the 2.5 unit difference which was the basis for the sample size calculation.

It is interesting to note that the average baseline Parent I/O was slightly more than 1 point higher than the baseline Teacher I/O. Thus, there was slightly more room for improvement in the Parent I/O. For each patient the teacher and parent used the same questionnaire to determine the Teacher I/O and the Parent I/O so these measures would ideally be correlated. However, for the ITT patients at baseline the correlation between

the measures was only 0.14 (and $R^2=0.02$). The different periods of observation and different relationships with the child of the teacher and parent might explain this. Also, the questionnaires were administered at different times of the day and potentially on different days of the week. As will be seen later, a similar correlation was observed in study N17-018 in which a significant treatment effect was found on change from baseline for both the teacher and parent I/O factors.

The sponsor found a significant treatment effect ($p=0.04$) in the subgroup of patients with baseline Teacher I/O ≥ 10 . This reviewer verified this subgroup analysis. The sponsor argued that patients with baseline score < 10 had less room for improvement. In light of this, the percent change from baseline would be more sensitive for patients with baseline Teacher I/O < 10 . Yet, if we perform a Wilcoxon test of a group difference in the percent change we still find no treatment effect ($p=0.13$).

The other subgroup analyses were verified except for the claim of significance in the subgroup with baseline Teacher I/O > 10 and 12.5 cm² final patch size. Although the means appear different in this subgroup the result is questioned because of the small sample size (12 MTS, 6 TS) and the disparity of the standard deviations (5.4 MTS and 1.5 TS). This reviewer also verified the sponsor's finding that prior use of ADHD medication was important in the subgroup of patients with baseline > 10 (and for all patients). Treatment naïve patients improved more for both treatment groups. Although there was not a significant interaction effect, a larger difference between the treatment groups was found for non-naïve patients.

Table 4.5.2 Change from Baseline in Teacher I/O as a function of Weight

Weight	Change from Baseline in Teacher I/O	
	Placebo	MTS
35.0 - 56.0	-2.75	-4.67
56.0 - 66.0	-0.68	-3.83
66.0 - 77.8	-1.42	-1.61
77.8 - 95.0	-1.05	-0.77
95.0 - 178.0	-1.54	0.14

Table 4.5.2 suggests that in the MTS group change in Teacher I/O is an increasing function of weight, i.e., the change is more negative for lower weights than for higher weights. The weight groups were defined so as to have approximately the same number of patients in each group. If we assume a linear relationship between change and weight within each treatment group the estimated slopes are 0.053 for MTS and 0.003 for Placebo. The hypothesis for equality of these slopes is rejected ($p=0.0058$). It turns out that the slope for the Placebo group is not significantly different from 0, so we conclude that only in the MTS group is there an effect of weight on change in the Teacher I/O from baseline. The existence of this effect might imply that the doses in this study were too small for the heavier individuals. Although this relationship was found post-hoc and therefore increases the type I error rate, it is intuitively reasonable because of the wide range of weights (37 to 177 lbs).

4.5.3. Subgroup Analyses

Table 4.5.3 presents the change in Teacher I/O within subgroups of interest. No differences were found in treatment effects for gender or prior use of ADHD medications. However, there is a significant treatment by race interaction. Caucasians appear to have benefited the most from the treatment while Hispanics actually did better on placebo (however, note the small sample size for this group). There is also some evidence that older individuals did not benefit as much from the treatment (however, note that they had less room to improve because they had lower values at baseline).

Table 4.5.3 Change in Teacher I/O within Subgroups

	Treatment Group								Inter- action p-value
	MTS				Placebo TS				
	baseline N	Mean	change N	Mean	baseline N	Mean	change N	Mean	
Sex									0.304
FEMALE	24	7.13	24	-1.04	24	7.71	24	-1.46	
MALE	74	10.32	74	-2.67	80	9.91	77	-1.52	
Race									0.043
AFRICAN-AMERICAN	8	12.94	8	-0.94	15	8.90	13	-0.37	
ASIAN	1	15.00	1	-1.00					
CAUCASIAN	74	9.02	74	-2.60	73	9.54	72	-1.41	
HISPANIC	11	10.00	11	-1.73	11	8.91	11	-3.82	
OTHER	4	9.75	4	-0.75	5	10.00	5	-0.80	
Age Group									0.084
06-09	59	10.06	59	-3.24	66	10.22	65	-1.88	
10-13	39	8.76	39	-0.81	38	7.98	36	-0.82	
Any Pr ADHD Meds Taken									0.537
NO	53	8.84	53	-2.55	53	8.79	52	-2.17	
YES	45	10.37	45	-1.94	51	10.04	49	-0.80	

5. Summary and Conclusion

5.1. Summary

Study N17-018 was conducted after the completion of study N17-010. Study N17-010 started on 09-12-2000 and ended on 02-16-2001. Study N17-018 started on 10-23-2001 and was completed on 03-05-2002.

Study Similarities:

- Flexible Dosing - At the end of each week the investigator was allowed to titrate the dose up or down for the following week for lack of efficacy or safety reasons.
- Sample Size – about 210 randomized patients

Study Differences:

- In study N17-018, patients were treated and followed for 4 weeks as compared to 3 weeks in study N17-010.
- On average, doses tended to be higher in study N17-018 than study N17-010. In study N17-010 patients started with 6.25 cm² patches, whereas in study N17-018 patients started with 12.5 or 18.75 cm² patches (depending on their pre-study dose of methylphenidate or their body weight if they were not taking methylphenidate previously). In addition, the highest allowable dose in study N17-018 (50 cm²) was twice the corresponding limit in study N17-010 (25 cm²). In fact, the average dose in the first week of study N17-018 is only slightly smaller than the average dose for the last week in study N17-010 and the average dose in N17-018 increases in later weeks.
- The mean baseline Teacher I/O was higher in study N17-018 (10.7 vs. 9.4) so there was more room for improvement.

In study N17-010 no treatment group difference was found in the change from baseline to last observation in the Teacher I/O. There was a significant treatment by center interaction ($p=0.01$) though. The fact that the average reduction in the Teacher I/O was larger for the placebo group in 8 of the 20 centers (and smaller in 12) is evidence of this interaction. If we ignore center and treatment by center effects then the ANCOVA model based mean reduction in the change in the Teacher I/O is 2.29 for the MTS group and 1.55 for the placebo group. The associated p-value is 0.1356. Thus, there is still no treatment difference. However, significant group differences were found for some secondary endpoints including the change in the Parent I/O from baseline to last visit, and the Clinical Global Impression of Improvement. The ANCOVA based mean reductions in the Parent I/O were 4.3 for the MTS group and 2.3 for the placebo group ($p=0.0001$). Although the Parent I/O result is significant the group difference is less than the 2.5 points used in the sample size calculation.

In the subgroup of 109 patients who had a baseline Teacher I/O greater than the median, 10, the MTS group had significantly better improvement ($p=0.0399$). The mean reductions were 3.87 for the MTS group and 2.28 for the placebo group. The sponsor pointed out that this is important because patients who started with higher scores had more room for improvement. The 1.59 group difference still seems modest. We should also remember that subgroup findings should be regarded with caution because of the increased possibility of a type I error that comes with multiple testing.

In study N17-018 the MTS group had significantly more improvement than the placebo group in the Teacher I/O from baseline to last observation ($p=0.0001$). The ANCOVA model based mean reductions from baseline to last visit were 5.16 for the MTS group and 1.33 for the placebo group. A treatment by baseline interaction was observed which suggests that the effectiveness of the MTS group was a function of the baseline Teacher I/O. In particular, the interaction suggests that the MTS group had significantly better improvement only for baseline scores greater than 5. On the other hand, the more the baseline exceeded 5 the more dramatic the improvement was. However, the interaction could be caused by the bounded range of the I/O scale, i.e., the closer the baseline is to 0 the less room there is for improvement. One way to eliminate this problem would be to examine the change in the Teacher I/O relative to the baseline score. When this was done this reviewer found that the treatment effect was significant ($p=0.0001$). So the treatment by baseline interaction seen for the change in the Teacher I/O may not be a concern. The MTS group had significantly more improvement on all secondary endpoints as well.

For study N17-018, the effectiveness of MTS in reducing the Teacher I/O was further investigated in several patient subgroups: age 6-9 vs. 10-13, males and females, treatment naïve vs. not naïve, and race. No significant treatment by gender or treatment by race interaction was found, suggesting that the treatment effect is constant over gender and race categories. On the other hand, there was some evidence that the treatment effect was smaller for the 10-13 age group than the 6-9 age group ($p=0.06$) and smaller for treatment naïve patients than for non-naïve patients ($p=0.06$). However, the treatment was still deemed effective in these groups.

The efficacy evidence in N17-018 would need to be weighed with the safety profile (substantially more adverse events were seen in study N17-018 than in N17-010).

It is interesting to note that in both phase III studies the Parent I/O at the last visit correlated more strongly with the CGI-I than the Teacher I/O did.

	Correlation		
	Teacher I/O CGI-I	Parent I/O CGI-I	Teacher I/O Parent I/O
N17-010	0.42	0.56	0.40
N17-018	0.56	0.71	0.55

5.2. Conclusion

Study N17-010 did not demonstrate efficacy on the primary endpoint (however, efficacy was demonstrated on several secondary endpoints such as the Parent I/O and the Clinical Global Impression of Improvement). For these reasons, it seems that most of the strength of evidence for efficacy of the MTS resides in study N17-018.

N17-018 demonstrated efficacy on the primary endpoint and many secondary endpoints. The results for the primary endpoint, Teacher I/O, were not changed by adjusting for gender or race. There was some suggestion, though, that the treatment effect was slightly diminished for both older children (10-13 as opposed to 6-9) and treatment naïve children. In contrast to the efficacy results, the safety profile in study N17-018 was somewhat troubling. There were notable numbers of insomnia and anorexia related adverse events and group differences in their rates of occurrence: 50% experienced anorexia in the MTS group compared to 2% for the placebo group, and 29% experienced insomnia in the MTS group compared to 5% for the placebo group.

Tristan Massie, Ph.D.
Statistical Reviewer

Concur:

Kun Jin, Ph.D.
Team Leader

George Chi, Ph.D.
Director, Division of Biometrics I

CC: NDA 21,514
HFD-120/Dr. Katz
HFD-120/Dr. Andreason
HFD-120/Dr. Mannheim
HFD-120/Ms. Weikel
HFD-700/Dr. Anello
HFD-710/Dr. Chi
HFD-710/Dr. Jin
HFD-710/Dr. Massie

This document was saved in c:/documents and settings/massiet/my documents/NDAs/nda021514.doc. This review consists of 33 pages.

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

/s/

Tristan Massie
2/14/03 10:43:55 AM
BIOMETRICS

Kun Jin
2/14/03 01:54:04 PM
BIOMETRICS

George Chi
2/14/03 03:50:50 PM
BIOMETRICS