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

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/BLA #:	NDA 202-236
Supplement #	S-008
Drug Name:	Dymista® (azelastine hydrochloride and fluticasone propionate) Nasal Spray 137 mcg/50 mcg per spray
Indication(s):	Relief of Symptoms of Seasonal Allergic Rhinitis (SAR) in (b) (4) Patients 6 (b) (4) years of age
Applicant:	Meda Pharmaceutical, Inc.
Date(s):	Submitted: August 22, 2014
	PDUFA: February 22, 2015
Review Priority:	Priority
Biometrics Division:	Division of Biometrics II
Statistical Reviewer:	Yongman Kim, Ph.D.
Concurring Reviewer:	David Petullo, M.S.
Statistical Supervisor	Thomas Permutt (Division Director)
Medical Division:	Division of Pulmonary, Allergy, and Rheumatology Products
Clinical Team:	Kathleen Donohue, M.D. (Clinical Reviewer)
	Banu Karimi-Shah, M.D. (Clinical Team Leader)
Project Manager:	Sadaf Nabavian, Pharm. D.

Keywords: NDA, clinical studies, repeated measures ANCOVA model

Table of Contents

LIST OF TABLES

LIST OF FIGURES

1 EXECUTIVE SUMMARY

In the current supplemental NDA submission, Meda Pharmaceutical, Inc. (Meda) has proposed Dymista (azelastine hydrochloride and fluticasone propionate) nasal spray 137 mcg/50 mcg per spray for the Relief of Symptoms of Seasonal Allergic Rhinitis (SAR) in pediatric patients 6 to 11 years of age. Efficacy was assessed by a single primary endpoint, change from baseline in 12-hour reflective Total Nasal Symptom Score (rTNSS) over 14-day treatment period in a study (MP4008) conducted to meet the Pediatric Research Equity Act (PREA) commitments. Key secondary endpoints included change from baseline in 12-hour reflective Total Ocular Symptom Score (rTOSS) and Pediatric Rhinoconjunctivitis Quality of Life Questionnaire (PRQLQ).

Based on my review of the data from study MP4008, the primary and secondary efficacy endpoints, rTNSS and rTOSS, did not demonstrate a statistically significant treatment effect in favor of Dymista nasal spray. This was supported by exploratory analyses of these endpoints including a pattern mixture model to examine the impact of missing data. Although there was a statistically significant treatment effect was noted for PRQLQ at Day 15, the study failed to reach significance for its primary efficacy endpoint. Therefore, from a statistical perspective, the submitted data did not provide substantial evidence of Dymista's efficacy benefit in the pediatric patients with ages between 6 and 11 years.

2 INTRODUCTION

2.1 Overview

Meda submitted this application on August 22, 2014 to support the approval of Dymista nasal spray for the relief of symptoms associated with SAR in pediatric patients 6 to 11 years of age. Dymista nasal spray consists of a fixed-dose combination of azelastine hydrochloride and fluticasone propionate. Each actuation of the nasal spray pump delivers 137 mcg of azelastine hydrochloride and 50 mcg of fluticasone propionate such that 1 spray per nostril twice daily delivers a total daily dose of 548 mcg of azelastine hydrochloride and 200 mcg of fluticasone propionate.

The submission included the efficacy results from a single phase 3, randomized, double-blind, placebo-controlled study, MP4008. The objective of the study was to evaluate the efficacy and safety of Dymista in pediatric patients with SAR.

Regulatory History and Interactions

Dymista Nasal Spray was approved on May 1st, 2012 for the treatment of SAR in patients 12 years of age and older. Study MP4008 evaluated children 4 to 11 years of age and was a postmarketing requirement under PREA. Pediatric study requirements for children less than four years of age were waived.

In a clinical information request issued on September 6th, 2013, we provided recommendation on the statistical analysis plan for the pediatric efficacy study:

> *Statistical information, including power of study(ies) and statistical assessments:*
> Study 3 must have a pre-specified, detailed statistical analysis plan appropriate to the study design and outcome measure. The statistical analysis plan for Study 3 should be submitted prior to the start of the study. This should include information addressing the issue of missing data, key efficacy endpoint. Reasons for discontinuation of treatment or withdrawal from the study should be recorded, avoiding less informative terms such as 'lost to follow-up', 'patient/investigator decision,' 'withdraw consent', in favor of categories relevant to safety or effectiveness, such as 'treatment ineffective' or 'adverse reaction.' You should refer to "The Prevention and Treatment of Missing Data in Clinical Trials" by the National Research Council.

In a pre-sNDA meeting held on January 13th, 2014, (b) (4) [redacted]

2.1.1 Specific Studies Reviewed

The current submission includes two phase 3 studies, MP4007 and MP4008 that evaluated children between the ages of 4 and 11. However, since study MP007 was an open-label trial designed to evaluate the safety of Dymista nasal spray compared to fluticasone propionate nasal spray, the focus of this review is on the efficacy data from study MP4008. The design of this study, which is also referenced in the label, is described in Table 1.

Table 1. Clinical Trial Reviewed

Trial No.	Phase	Design	Treatment Arms	Number of Patients	Dates
MP4008	3	randomized, double-blind, 14-day trial	Dymista nasal spray (azelastine hydrochloride and fluticasone propionate) Placebo	173 175	7/2013-2/2014

Source: Reviewer

2.2 Data Sources

Information regarding NDA 202-236 can be found in the Center for Drug Evaluation and Research electronic document room (EDR). The study report including protocols, statistical analysis plan, and all referenced literature were submitted. The program codes used in statistical analyses and the electronic data sets with raw and derived variables and data definitions can be found at the following location:

\\cdsesub1\evsprod\NDA202236\0056\m5\datasets

3 STATISTICAL EVALUATION

3.1 Data and Analysis Quality

In general, the submitted efficacy data were acceptable in terms of quality and integrity. I was able to derive the primary and secondary efficacy endpoints for the study reviewed. No noticeable deviations between the case report forms and analysis datasets relevant to primary and secondary endpoints were identified. The statistical analyses of my derived endpoints were consistent with the applicant's analyses.

Based on the information provided in this submission, the study seemed to be conducted properly and was consistent with the history of regulatory interactions, protocol revisions/amendments, study report, and study datasets.

3.2 Evaluation of Efficacy

My evaluation of the efficacy of Dymista nasal spray was based on study MP4008. For simplicity, Dymista nasal spray will be denoted by Dymista.

Study Design and Endpoints

The study consisted of a (1) washout period for prohibited concomitant medications, if needed; (2) placebo lead-in period (up to 7 days) during which subjects must meet eligibility criteria; (3) 15 day treatment period in which qualified subjects were randomized to either Dymista nasal spray or placebo nasal spray. Treatment was administered as one spray per nostril twice daily. Randomization was stratified according to age as follows:

1. Ages ≥4 years to < 6 (Group 1)
2. Ages ≥6 years to < 9 (Group 2)
3. Ages ≥9 years to < 12 (Group 3)

Male and female subjects ≥4 years to <12 years of age with a history of SAR with a positive skin test to a local prevailing pollen, who meet all of the study inclusion and none of the exclusion criteria were eligible for randomization. All subjects must have had symptomatic, moderate-to-severe allergic rhinitis as evidenced by their symptom scores at randomization.

Signs/symptoms were scored both as reflective and instantaneous:

- 12-hour reflective scores (r) - how signs/symptoms were over the previous 12 hours
- Instantaneous scores (i) - how signs/symptoms are at the time of the evaluation

Subjects/caregivers were instructed to record the following assessments twice daily (AM and PM) in a diary:

- AM and PM 12-hour reflective (r) and instantaneous (i) TNSS (consisting of nasal congestion, runny nose, sneezing and nasal itching)
- AM and PM 12-hour reflective (r) and instantaneous (i) TOSS (consisting of itchy eyes, watery eyes, and red eyes)
- AM and PM 12-hour reflective (r) and instantaneous (i) TSS (consisting of TNSS and TOSS)

The individual signs/symptoms of the TNSS and the TOSS were scored on a 0 to 3 point scale as follows:

- 0 = absent symptoms (no sign/symptom evident)
- 1 = mild symptoms (sign/symptom clearly present, but minimal awareness; easily tolerated)
- 2 = moderate symptoms (definite awareness of sign/symptom that is bothersome but tolerable)
- 3 = severe symptoms (sign/symptom that is hard to tolerate; causes interference with activities of daily living and/or sleeping)

The maximum daily scores for were 24 for TNSS, 18 for TOSS, and 42 for TSS.

The primary efficacy was change from baseline in AM+PM 12-hour rTNSS for the entire double-blind period (i.e. Day 2 AM to Day 14 PM). Key secondary endpoints were change from baseline in AM+PM 12-hour rTOSS for the entire double blind period and change from baseline to Days 8 and 15 in the PRQLQ. Other secondary endpoints examined were:

- Change from baseline in AM+PM 12-hour iTNSS for the entire double-blind period
- Change from baseline in AM+PM 12-hour iTOSS for the entire double blind period
- Change from baseline in AM+PM 12-hour rTSS and iTSS for the entire double-blind period

- Change from baseline in AM and PM 12-hour rTNSS, rTOSS and rTSS for the entire double-blind period
- Change from baseline in AM and PM 12-hour iTNSS, iTOSS and iTSS for the entire double-blind period
- Change from baseline in AM+PM, AM and PM 12-hour reflective and instantaneous individual nasal symptom scores (rhinorrhea, itchy nose, nasal congestion and sneezing) for the entire double-blind period
- Change from baseline in AM+PM, AM and PM 12-hour reflective and instantaneous individual ocular sign/symptom scores (itchy eyes, watery eyes, and red eyes) for the entire double-blind period
- Daily change from baseline in AM+PM 12-hour rTNSS, iTNSS, rTSS, iTSS, rTOSS and iTOSS for the entire double-blind period

Statistical Methodologies

The Intent-to-treat (ITT) set was defined as all randomized subjects who had at least one post-baseline observation (AM or PM) of efficacy. Of note, there were no subjects in the ITT set who did not have at least one post-baseline observation. The applicant's primary efficacy analysis population excluded children less than 6 years of age since the primary objective was to evaluate the efficacy and safety of Dymista treatment compared with placebo in pediatric patients between 6 and 11 years of age. I refer to this population as the Full Analysis Set (FAS). Also, the applicant considered the data from the younger children in Group 1 (Ages ≥4 years to < 6) as exploratory. However, since the clinical review team requested the efficacy data of these younger children, I conducted analyses with the ITT dataset.

For the analyses of the primary and key secondary efficacy endpoints, a repeated measures analysis of covariance (ANCOVA) model with treatment day (Days 2 through 14 only), treatment group and age group (Groups 2 and 3) as fixed effects and baseline as a covariate was utilized. Results of analysis were summarized including the difference between treatment groups in least square means and 95% confidence intervals for the difference. Missing values were not imputed for the primary analysis. Although this was a multi-center trial, a fixed effect due to site was not included in the analysis model due to the central randomization scheme which did not stratify by site. Analysis was conducted using both the ITT and FAS datasets.

In the analysis of the primary endpoint, missing data were not imputed and assumed to be missing-at-random (MAR). In other words, the analysis was conducted on observed cases (OC). Our concern regarding the MAR assumption was conveyed to the applicant and we recommended sensitivity analyses to assess impact of missing data. Subsequently, the applicant proposed the following sensitivity analyses regarding missing data in the SAP:

> A sensitivity analysis of the primary endpoint of AM+PM 12-hour rTNSS and the key secondary endpoint of AM+PM 12-hour rTOSS assessing the impact of missing data will be performed if the amount of missing data for the respective endpoint exceeds 12.5%. A

nearest-neighbor pattern mixture model under the assumption of missing not at random will be used in this event.
Missing data in the quality of life data (PRQLQ) will be handled according to the instrument developer's instructions. A sensitivity analysis of the key secondary endpoint of overall score at the end of treatment will be performed if the amount of missing data exceeds 12.5%. For this analysis, missing data will be imputed using a pattern mixture model which draws random from the placebo group.
For diary data, days with missing severity information will not be imputed. The sums of AM and PM scores will equal the non-missing assessment if one assessment is missing. TNSS will equal the sum of each of the individual nasal symptom scores such that if any individual score is missing, the TNSS will be missing. Similarly for TOSS, if any individual ocular symptom score is missing, the TOSS will be missing. TSS will equal the sum of the TNSS and TOSS such that if either summand is missing, the TSS will equal the non-missing summand. Weekly and overall averages will include only those days with non-missing data and the average will therefore have denominators reduced by the number of days with missing data. Missing data will not be included in the denominators of any endpoint upon which percentages are calculated.

The secondary endpoints related to nasal or ocular symptom scores were analyzed with the same model as for the primary endpoint.

An analysis of covariance (ANCOVA) model was used for analysis of PRQLQ overall score change from baseline at Visit 4 (Day 15). The model included fixed effect terms for treatment, age group (Groups 2 and 3) and baseline score as covariate.

In order to adjust for multiplicity, a gate-keeping strategy was employed. The primary efficacy endpoint of AM+PM 12-hour rTNSS for the entire treatment period is tested at two-sided alpha=0.05. If the two treatment groups are shown to be statistically different, then the following key secondary endpoints are also tested at two-sided alpha=0.05 sequentially:

1. Change from Baseline in AM+PM rTOSS for the entire double-blind period
2. Change from Baseline in PRQLQ overall score at Visit 4.

All other secondary efficacy endpoints are also tested inferentially using the FAS dataset; however, the p-values resulting from these analyses were unadjusted and were therefore to be considered supportive in nature.

Sample Size Calculation

The study was powered for evaluating the primary efficacy outcome variable for Dymista versus placebo. Based on the applicant's sample size calculation, 150 patients in placebo group and 150 patients in Dymista group would provide at least 80% power to detect a treatment difference of 1.0 unit in the change in rTNSS between baseline and Day 14, assuming a standard deviation of 3.0 units at a significance level of 0.05.

Patient Disposition, Demographic and Baseline Characteristics

Three-hundred forty eight subjects were randomized at 35 centers in the United States during seasonal allergy season. Most patients (99%) completed the 14 days of active treatment. Two subjects, one in each treatment group, discontinued from the study due to an adverse event.

The disposition of patients is summarized are shown in Table 2.

Table 2. Patients' Accountability, N (%)

	All Age Strata (N=348)		***Ages ≥6 to <12 (N=304)***	
	Dymista (n=173)	Placebo (n=175)	Dymista (n=152)	Placebo (n=152)
All Randomized Subjects				
N (%)	173 (100)	175 (100)	152 (100)	152 (100)
ITT Population				
N (%)	173 (100)	175 (100)	152 (100)	152 (100)
Subjects Who Completed Study				
N (%)	171 (99)	173 (99)	150 (99)	150 (99)
Subjects Who Discontinued Study				
N (%)	2 (1)	2 (1)	2 (1)	2 (1)
Reason for Discontinuation				
Adverse Event	1 (0.6)	1 (0.6)	1 (0.7)	1 (0.7)
Treatment Failure	1 (0.6)	0	1 (0.7)	0
Other	0	1 (0.6)	0	1 (0.7)

Source: Excerpted from the MP4008 Clinical Study Report (page 46).

In the study, the demographic and baseline disease characteristics were generally well balanced and comparable between the treatment groups (Table 3). The mean age was 9 years in FAS population and 8.5 years in ITT population. The mean height (weight) was 54 inches (82 pounds) in FAS population and 53 inches (77 pounds) in ITT population. Majority of patients were Caucasian and approximately half of patients were male. The overall proportion of black patients was 30%. The overall mean baseline AM+PM rTNSS score was 18 and the mean duration of SAR history was 6 years.

Table 3. Patients' Demographic and Baseline Characteristics by Treatment, N (%)

	All Age Strata (N=348)		***Ages ≥6 to <12 (N=304)***	
Demographic parameter	Dymista (n=173)	Placebo (n=175)	Dymista (n=152)	Placebo (n=152)
Age at Randomization (yrs)				
Mean (SD)	8.5 (2.1)	8.4 (2.1)	9.0 (1.6)	9.0 (1.6)
Sex				
Male	96 (55)	90 (51)	86 (57)	80 (53)
Female	77 (45)	85 (49)	66 (43)	72 (47)
Race*				
White	119 (69)	130 (74)	105 (69)	112 (74)
Black	53 (31)	49 (28)	47 (31)	44 (29)
Asian	4 (2)	0 (0)	3 (2)	0 (0)
Other	5 (3)	1 (1)	3 (2)	1 (1)
Height (in)				
Mean (SD)	53.3 (6.0)	53.0 (5.9)	54.7 (4.9)	54.2 (5.2)
Weight (lb)				
Mean (SD)	77.0 (28.3)	77.7 (31.7)	81.7 (26.8)	81.6 (31.4)
AM+PM rTNSS				
Mean (SD)	18.2 (3.4)	18.0 (3.2)	18.4 (3.5)	18.0 (3.2)
Duration of SAR History (Years)				
Mean (SD)	5.7 (2.7)	5.9 (2.4)	6.1 (2.6)	6.2 (2.4)

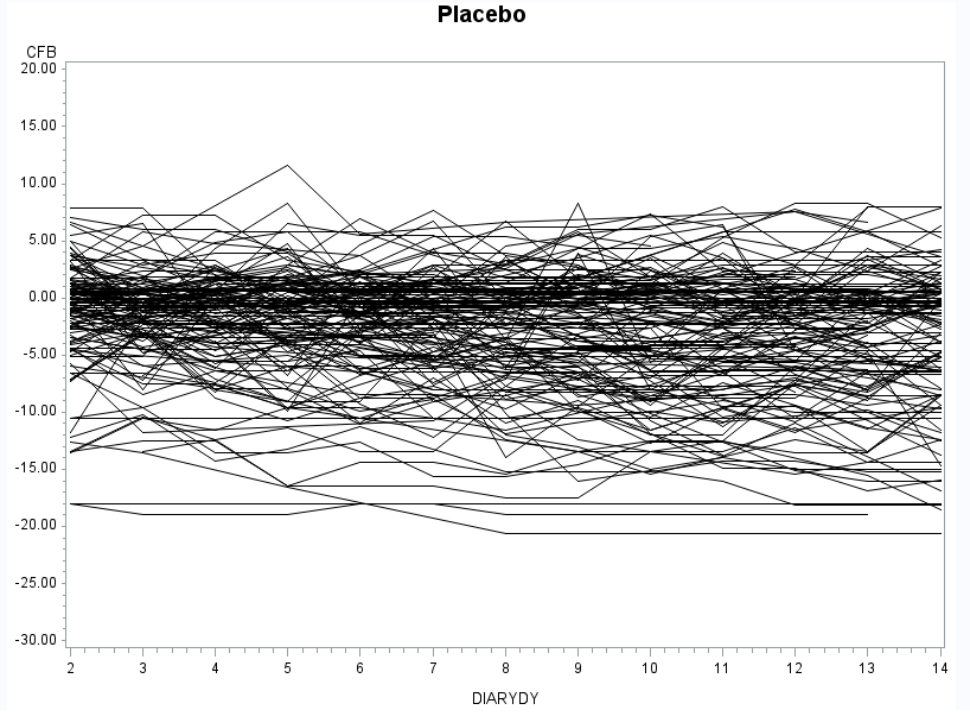
Source: Excerpted from the MP4008 Clinical Study Report (pages 50-51).
* More than one choice could be selected so percentages may total greater than 100%.

Results and Conclusions

Primary Efficacy Endpoint

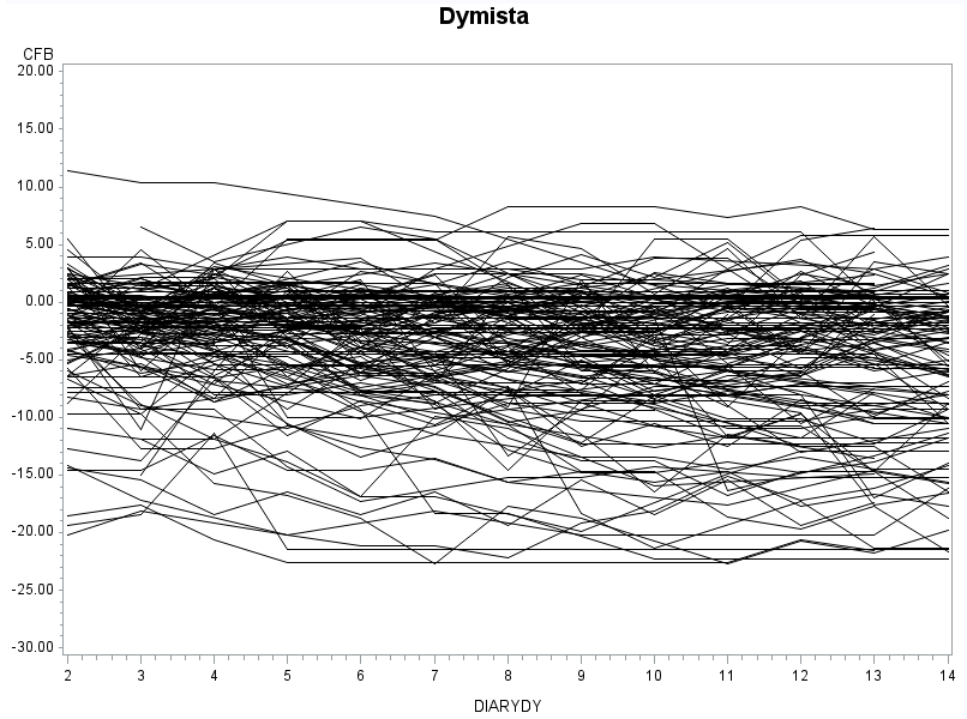
Figures 1and 2 describe the rTNSS change from baseline over double-blind period in each individual patient by treatment group. The majority of patients seem to experience a slight decline in rTNSS although degree of decline appears similar between treatment groups. In figure 3, the group mean along with standard error bars are shown. The decline in rTNSS for the Dymista group is consistently greater than that of placebo group, but the difference was not statistically significant.

Figure 1. AM+PM rTNSS change over double-blind period in individuals randomized to placebo



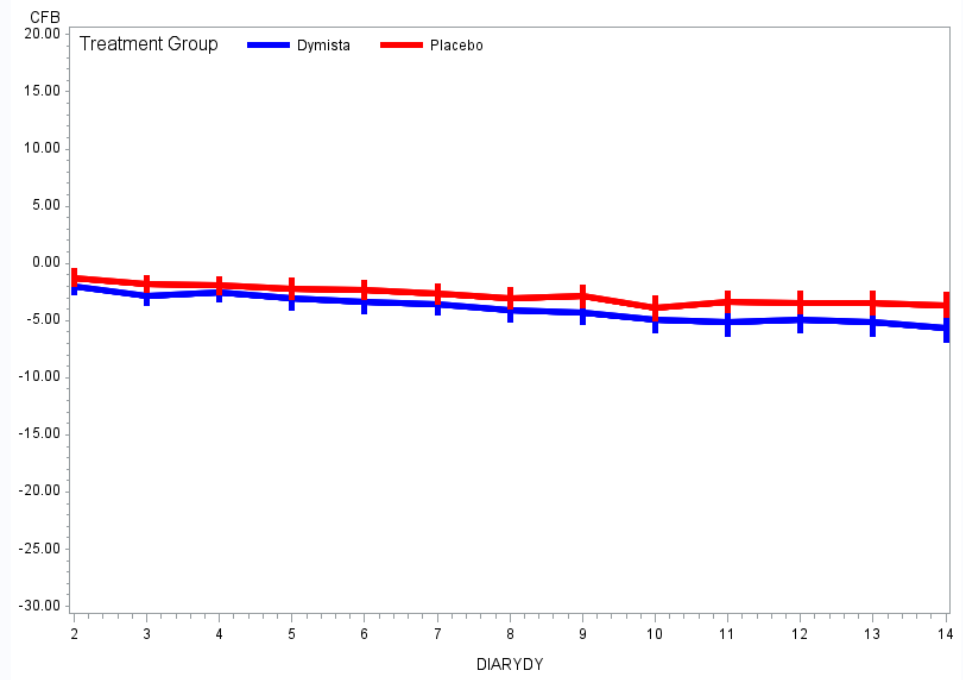
Source: Reveiwer

Figure 2. AM+PM rTNSS change over double-blind period in individuals randomized to Dymista



Source: Reveiwer

Figure 3. Mean AM+PM rTNSS change over double-blind period by treatment group



Source: Reviewer

The analysis of the primary endpoint was a repeated measures ANCOVA model without imputation for missing data. However, since less than 2% of patients discontinued the study before the double blind treatment period of 14 days, impact of missing data due to dropout was not a concern even though the statistical model assumed missing-at-random (MAR) mechanism for missing data. A sensitivity analysis using the nearest neighbor multiple imputation as specified in the statistical analysis plan was conducted by the applicant to examine the potential effect of missing data on the reliability of the primary analysis. See the appendix for detail of the analysis. I conducted a sensitivity analysis with average scores for individual patients without imputation for missing data using ANCOVA model with terms for treatment, age group and baseline score as covariate.

In the applicant's primary analysis using the FAS population, patients receiving Dymista had a numerically, but not statistically, greater mean change from baseline in AM+PM 12-hour rTNSS for the entire double-blind period compared to those receiving placebo. An estimated absolute difference was -0.8 units between the two treatment groups. Sensitivity analyses to examine the impact of missing data gave consistent results and agreed with the primary analysis. Additionally, an analysis with the ITT population including all age groups gave almost same mean changes for two treatment groups, -3.9 units for Dymista and -3.7 units for placebo. Results are shown in Table 4.

Table 4. Analyses of change from baseline in AM+PM rTNSS

Treatment	N	LS Mean (SE)	Difference vs. PBO		
			LS Mean Difference	95% CI	P-value
Applicant's Primary Analysis: FAS Set					
Dymista	152	-3.7 (0.4)	-0.8 (0.5)	(-1.8, 0.2)	0.099
Placebo	152	-2.9 (0.4)			
Applicant's Sensitivity Analysis with Nearest Neighbor Multiple Imputation: FAS Set					
Dymista	152	-3.7 (0.4)	-0.7 (0.5)	(-1.7, 0.2)	0.142
Placebo	152	-3.0 (0.4)			
My Sensitivity Analysis with Average Scores for Individual Patients : FAS Set					
Dymista	152	-3.8 (0.4)	-1.0 (0.6)	(-2.1, 0.1)	0.086
Placebo	152	-2.8 (0.4)			
Applicant's Primary Analysis: ITT Set					
Dymista	173	-3.9 (0.4)	-0.2 (0.5)	(-1.1, 0.8)	0.716
Placebo	175	-3.7 (0.4)			

Source: Excerpted from the MP4008 Clinical Study Report (page 56) & Reviewer

In summary, the results from study MP4008 did not show sufficient evidence in favor of Dymista on the change in rTNSS (primary efficacy endpoint). Due to a very low dropout rate, there was no impact of missing data on the primary analysis. The results from sensitivity analyses were consistent with the primary analysis. Additionally, an analysis on the ITT population with all age groups included gave consistent results with the primary analysis; there was no significant treatment effect noted for Dymista.

Secondary Efficacy Endpoints

I was able to confirm the results of the applicant's analyses of the key secondary endpoints. A review of these two pre-specified secondary efficacy endpoints is included in my review even though the primary endpoint failed to reach significance. To control for multiplicity, these endpoints were to be tested only if the primary endpoint was significant at an alpha of 0.05.

Change from baseline in AM+PM rTOSS for the entire double-blind period

Results from the applicant's analysis using the FAS population demonstrated that on average, patients receiving Dymista had a numerically, but not statistically, greater mean change from baseline in AM+PM rTOSS for the entire double-blind treatment period compared to those

receiving placebo. This is consistent with the primary endpoint, rTNSS. An estimated absolute mean difference was -0.5 units between the two treatment groups. Sensitivity analysis with respect to missing data by the applicant gave consistent results from the pre-specified analysis. Additionally, as shown in Table 5, an analysis using the ITT population which included all age groups gave consistent results.

Table 5. Analyses on change from baseline in AM+PM rTOSS

Treatment	N	LS Mean (SE)	Difference vs. PBO		
			LS Mean Difference	95% CI	P-value
Applicant's Analysis: FAS Set					
Dymista	152	-2.0 (0.3)	-0.5 (0.4)	(-1.2, 0.2)	0.143
Placebo	152	-1.5 (0.2)			
Applicant's Sensitivity Analysis with Nearest Neighbor Multiple Imputation: FAS Set					
Dymista	152	-2.1 (0.3)	-0.6 (0.4)	(-1.3, 0.1)	0.116
Placebo	152	-1.5 (0.2)			
Applicant's Analysis: ITT Set					
Dymista	173	-2.3 (0.3)	-0.1 (0.4)	(-0.8, 0.6)	0.568
Placebo	175	-2.2 (0.3)			

Source: Excerpted from the MP4008 Clinical Study Report (page 72) & Reviewer

The Change from baseline in PRQLQ score at Visit 4 (Day 15)

The results from the analyses of the mean change from baseline in PRQLQ score are summarized in Table 6. This endpoint was analyzed using an ANCOVA model with fixed effects for treatment, age group and baseline PRQLQ score as covariate.

The mean change in PRQLQ score for patients treated with Dymista was statistically significantly lower when compared to patients treated with placebo (-0.95 vs. -0.66, respectively; difference of -0.29, p=0.027). Sensitivity analyses with the same ANCOVA model after imputing missing data with pattern mixture model which drew random values from the placebo group were consistent with results from the applicant's pre-specified ANCOVA analyses.

Table 6. Analyses on change from baseline in PRQLQ score at Visit 4

Treatment	N	LS Mean (SE)	Difference vs. PBO		
			LS Mean Difference	95% CI	P-value
Applicant's Analysis: FAS Set					
Dymista	152	-0.95 (0.10)	-0.29 (0.13)	(-0.55, -0.03)	0.027
Placebo	152	-0.66 (0.09)			
Applicant's Sensitivity Analysis with Pattern Mixture Multiple Imputation: FAS Set					
Dymista	152	-0.95 (0.09)	-0.29 (0.13)	(-0.33, -0.03)	0.027
Placebo	152	-0.65 (0.09)			

Source: Excerpted from the MP4008 Clinical Study Report (page 83) & Reviewer

3.3 Evaluation of Safety

The assessment of the safety of the study drug was mainly conducted by the reviewing medical team. The reader is referred to Dr. Kathleen Donohue's review for information regarding the safety profile of the drug.

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Subgroup Analyses by Age Strata

In a subgroup analysis by age group, there seemed to be a qualitative interaction of treatment by age group. The youngest group of patients, between 4 and 5 years of age (Group 1) favored placebo over Dymista whereas each of the two older groups of patients, between 6 and 11 years of age (Groups 2 and 3) favored Dymista over placebo. Results are shown in Table 7.

Table 7. Analyses on change from baseline in AM+PM rTNSS by age group

Treatment	N	LS Mean (SE)	Difference vs. PBO		
			LS Mean Difference	95% CI	P-value
≥4 to <6 Years of Age (Group 1)					
Dymista	21	-2.4 (0.6)	5.3 (1.1)	(3.0, 7.5)	<0.001
Placebo	23	-7.7 (1.0)			
≥6 to <9 Years of Age (Group 2)					
Dymista	59	-3.7 (0.6)	-0.7 (0.8)	(-2.2, 0.8)	0.363
Placebo	60	-3.0 (0.6)			
≥9 to <12 Years of Age (Group 3)					
Dymista	93	-3.8 (0.4)	-1.1 (0.5)	(-2.2, -0.1)	0.040
Placebo	92	-2.6 (0.4)			

Source: Excerpted from the MP4008 Clinical Study Report (page 65) & Reviewer

5. SUMMARY AND CONCLUSIONS

5.1 Statistical Issues and Collective Evidence

The single study that evaluated the efficacy of Dymista nasal spray for the relief of symptoms of SAR failed to achieve statistical significance with respect to the primary efficacy endpoint, change from baseline in AM+PM rTNSS. There were no statistical issues that could not be addressed. Missing data was not a concern as this study had a 99% completion rate. Therefore, the mixed model approach with missing-at-random assumption utilized by the applicant was acceptable. The applicant's primary analysis excluded patients younger than 6 years of age. Per clinical request, I conducted analyses for the primary and secondary endpoints using patients between 4 and 12 years of age. The results from these analyses were consistent with the pre-specified analyses that evaluated children between 6 and 12 years of age.

In terms of multiplicity, the applicant proposed a hierarchical testing strategy for the key secondary endpoints, change from baseline in AM+PM rTOSS and change from baseline in PRQLQ. If the primary efficacy endpoint was significant, then rTOSS was to be tested, if significant, PRQLQ was to be tested. Even though the analysis of PRQLQ was statistically significant in favor of Dymista, since the primary endpoint failed to reach the statistical significance, the proposed sequential test stopped after the primary analysis.

5.2 Conclusions and Recommendations

Analyses of the efficacy data from the study MP4008 failed to provide substantial evidence that Dymista nasal spray relieves symptoms associated with SAR. Although the data showed improvement in quality of life based on PRQLQ, since the primary endpoint, rTNSS and one of the key secondary endpoints, rTOSS did not reach the statistical significance the significance of PRQLQ does not provide substantial evidence of efficacy.

5.3 Labeling Recommendations

Following is an excerpt from section 14 in the proposed label. I generally agree with the study description, but I do not agree with their interpretation of the efficacy data that results of the trial provided data supportive of efficacy in children 6-11 years of age. Further, based on the current draft guidance for including pediatric information in product labeling, the results from pediatric studies that do not provide substantial evidence to warrant a pediatric indication must appear only in the Pediatric Use subsection of the label, section 8.4. This section should also state that (b) (4).

(b) (4)

APPENDICES

Sensitivity analysis to the primary analysis proposed by the applicant (excerpted from the attachment to Statistical Analysis Plan)

Sensitivity Analysis – Multiple Imputation based on Nearest Neighbor Approach

Introduction:

In this study, 7 individual allergy symptom scores, 4 nasal and 3 ocular, were recorded in electronic diaries (eDiary) twice daily in the morning (AM) and evening (PM) approximately 12 hours apart. They were rated as reflective over the last 12 hours ("reflective" abbreviated as "r") as well as how they felt at the moment of recording ("instantaneous" abbreviated as "i"). Reflective and instantaneous symptom scores were summarized separately. The daily combined AM+PM score for each symptom was calculated as the sum of the AM and PM scores such that if either summand was missing, the daily total was missing. Total scores of nasal and ocular symptoms were calculated (TNSS and TOSS, respectively). The daily combined AM+PM TNSS was based on the sum of the AM+PM scores of each of the 4 individual nasal symptoms such that if any individual score was missing, the TNSS was missing. Similarly for TOSS, if any individual ocular symptom score was missing, the TOSS was missing.

The primary analysis model included data captured on study days 2 through 14 inclusive. While only 2 subjects discontinued from the study early (1% of the total number of subjects) and were therefore missing some days of diary data, 69% of subjects were missing the AM+PM rTNSS and rTOSS on one or more days during Days 2 through 14. Days where the AM+PM rTNSS and rTOSS were missing was caused primarily by subjects not completing their eDiary symptom assessments in either the AM or the PM on a given day (12% of the total expected number of days of data for Days 2 through 14 were missing either an AM or a PM score). Relatively few subjects were missing both the AM and PM score for any given day (2% of the total expected number of days of data for Days 2 through 14 were missing both the AM and the PM score).

Methodology

Per the SAP, a sensitivity analysis of the primary endpoint of AM+PM rTNSS and the key secondary endpoint of AM+PM rTOSS assessing the impact of missing data was to be performed if the amount of missing data for the respective endpoint exceeded 12.5%. Given the total amount of missing data due to dropouts and other sources unknown, the planned sensitivity analyses for rTNSS and rTOSS were performed.

A nearest-neighbor pattern mixture model was used in accordance with the SAP. The details of the approach to the data imputation are given below. Once the missing data were imputed, analysis of covariance (ANCOVA) using the primary analysis model (described in Section 9.7.1.4) was performed and the parameter estimates were derived. This process was completed 1000 times. The results of these 1000 iterations of the analysis were combined using SAS version 9.3 and SAS/STAT version 12.1 PROC MIANALYZE to derive the final estimates of the parameters and a test of treatment difference. The output from the combined analysis is attached to this document (Table 51 for AM+PM rTNSS and Table 51.1 for AM+PM rTOSS). The results for AM+PM rTNSS are summarized in Section 11.4.1.2.1 Table 10 of the CSR and in Section 11.4.1.2.6 for AM+PM rTOSS.

Data Imputation

1. Revise the derived DIARY analysis dataset, to calculate the change from previous score for every record (AM/PM timepoint and diary day) that has a non-missing score.
    a. The change will be calculated from the most recent previous timepoint, if that timepoint's score is also non-missing.

For example, for a Day 3 AM score, the change from previous score will be calculated as (Day 3 AM score – Day 2 PM score). If the Day 2 PM score is missing, the change is missing.

b. This CHANGE variable will be used in imputation. The CHANGE variable will always be calculated from real scores, never from imputed scores.

2. LOOP 1000 TIMES – Create 1000 iterations of a complete set of diary data.
    a. Impute missing scores for each timepoint, in consecutive time order (i.e. Day 2 AM, Day 2 PM, Day 3 AM, Day 3 PM) as follows:
        i. For all subjects with a missing score, impute scores for a given day/timepoint:
            1. Determine how many scores are missing for this day/timepoint.
            2. Obtain a random selection of that number of change-from-previous values, from among the subjects that have non-missing scores at that day/timepoint.
                a. Note: random selection is performed without regard to treatment. This reduces treatment differences and is conservative.
            3. For every subject that has a missing score, calculate an imputed score as (the subject's most recent previous score) + (a randomly selected change-from-previous value). Truncate to ensure score is between 0 and 12, inclusive.
                a. Note: the subject's most recent previous score will be selected as the most recent previous timepoint (for example, for a missing Day 4 PM score, the Day 4 AM score will be selected); the score selected may be an imputed score.
                    i. *Exception: 3 subjects (824-004, 825-005, 827-004) have a missing Day 2 AM score, and also a missing Day 1 PM score. For these 3 subjects, the Day 1 AM score will be used as the most recent previous score in calculating the imputed score.*
                b. Note: the randomly selected change-from-previous value is always to be derived from real scores, never from imputed scores.

            At end of this step (2.a.i), we have a complete set of scores for all subjects, for a given day and time (AM or PM).

        ii. Repeat for each consecutive day and time, from Day 2 AM through Day 14 PM.

        At the end of this step (2.a.ii), we have a complete set of scores for all days 2-14, for both AM and PM.

    b. Using the imputed AM and PM scores, calculate the AM+PM combined scores, and the Change from Baseline.

At the end of this loop (2), we have 1000 sets of complete diary data.

Example:

Consider timepoint X, and calculate imputed score = A + B.

- A is the subject's "most recent previous score", i.e. from the timepoint immediately prior to this time (X-12 hrs); an imputed score from the previous step will be used, if a real score is not available.
- B is a randomly selected "change from most recent previous score", from the other subjects that have a non-missing score at this timepoint. An imputed score will not be used when calculating "change from most recent previous score" for the other subjects.

Determining change from most recent previous score:

Subject	2AM	2PM	3AM	3PM	4AM	4PM	Change from most recent previous score
001	x	x			x	x	missing (no Day 3 PM score)
002	x	x	x	x		x	missing (no Day 4 AM score)
003	x		x		x	x	missing (no Day 3 PM score)
004	x	x		x	x	x	Day 4 AM - Day 3 PM
005	x	x	x	x	x	x	Day 4 AM - Day 3 PM
006	x	x	x		x	x	missing (no Day 3 PM score)
007	x	x	x	x		x	missing (no Day 4 AM score)
008	x	x	x	x			missing (no Day 4 AM score)
009	x	x		x	x	x	Day 4 AM - Day 3 PM

For subject 002 at Day 4 AM,
A=Day 3 PM score; B=random selection of change from previous, among subjects 004, 005 and 009.

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

/s/

YONGMAN KIM
01/22/2015

DAVID M PETULLO
01/23/2015
I concur.