CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

207695Orig1s000

STATISTICAL REVIEW(S)

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 207695
Drug Name:	EUCRISA™ (crisaborole) topical ointment, 2%
Indication(s):	Atopic Dermatitis
Applicant:	Anacor Pharmaceuticals
Date(s):	Letter Date: 1/7/2016
	PDUFA Date: 1/7/2017
Review Priority:	Standard
Biometrics Division:	Division of Biometrics III
Statistical Reviewer:	Matthew Guerra, Ph.D.
Concurring Reviewers:	Mohamed Alosh, Ph.D.
Medical Division:	Division of Dermatology and Dental Products
Clinical Team:	Melinda McCord, M.D. / Snezana Trajkovic, M.D.
Project Manager:	Lydia Springs

Keywords: atopic dermatitis, superiority trial, multiple imputation

Table of Contents

1 EXECUTIVE SUMMARY

The applicant, Anacor Pharmaceuticals, has developed EUCRISA™ (crisaborole) topical ointment, 2% for the treatment of mild to moderate atopic dermatitis in patients 2 years of age and older.

The applicant submitted data from two identically-designed, randomized, multicenter, vehicle-controlled, parallel-group, pivotal Phase 3 trials (Trials 301 and 302). The trials enrolled subjects 2 years of age and older with a clinical diagnosis of atopic dermatitis with body surface area (BSA) involvement ≥ 5% (excluding scalp) and an Investigator's Static Global Assessment (ISGA) score of 2 (mild) or 3 (moderate). The protocol-specified primary efficacy endpoint was the proportion of subjects achieving success in ISGA at Day 29, where success in ISGA was defined as an ISGA score of 0 (clear) or 1 (almost clear) with at least a 2-grade improvement from baseline. The protocol specified the following two secondary efficacy endpoints: the proportion of subjects with an ISGA score of 0 (clear) or 1 (almost clear) at Day 29 and the time to success in ISGA.

Table 1 presents the results of the primary and secondary efficacy endpoints. In both trials, EUCRISA ointment, 2% was statistically superior (p-values ≤ 0.038) to vehicle ointment for all endpoints presented in Table 1. It should be noted that the median time to success in ISGA (i.e., the time at which 50% of the subjects achieved success in ISGA) could not be calculated, as fewer than 50% of subjects achieved success in ISGA.

Table 1: Efficacy Results at Day 29 (ITT, MI[1])

	Trial 301			Trial 302		
Endpoints	**EUCRISA (N=503)**	**Vehicle (N=256)**	**P-Value**	**EUCRISA (N=513)**	**Vehicle (N=250)**	**P-Value**
Primary: Success in ISGA[2]	32.8%	25.4%	0.038[3]	31.4%	18.0%	<0.001[3]
Secondary: ISGA score of Clear or Almost Clear	51.7%	40.6%	0.005[3]	48.5%	29.7%	<0.001[3]
Time to Success in ISGA[2]	NC[4]	NC	<0.001[5]	NC	NC	<0.001[5]

Source: Reviewer's Analysis (same results as Applicant's Analysis)

(1) Missing data was imputed using multiple imputation (MI). The values displayed are the averages over the 140 imputed datasets for Trial 301 and the 135 datasets for Trial 302.
(2) Success is defined as an ISGA score of 0 (clear) or 1 (almost clear) with at least a 2-grade improvement from baseline.
(3) P-value from a logistic regression (using Firth's Penalized Likelihood) with treatment and analysis center as factors.
(4) Median time to success in ISGA could not be calculated because fewer than 50% of subjects achieved success in ISGA.
(5) P-value based on a log-rank test.

2 INTRODUCTION

2.1 Overview

The applicant, Anacor Pharmaceuticals, has developed EUCRISA™ (crisaborole) topical ointment, 2% for the treatment of mild to moderate atopic dermatitis in patients 2 years of age and older.

2.1.1 Regulatory History

On February 13, 2008, the Agency and the applicant met for a Pre-IND meeting. The Agency provided general comments regarding the development of an investigator static global assessment (ISGA) scale.

On February 26, 2014, the Agency and the applicant met for an End-of-Phase 2 (EOP2) meeting. The Agency agreed with the applicant that the primary efficacy endpoint for the Phase 3 trials be the proportion of subjects achieving success in ISGA at Day 29, where success is defined as a score of 0 (clear) or 1 (almost clear) with at least a 2-grade improvement from baseline. For the Phase 3 trials, the applicant proposed the following three secondary efficacy endpoints:

- Proportion of subjects with an ISGA score of 0 (clear) or 1 (almost clear) at Day 29
- Time to improvement in pruritus (defined as a pruritus score of None [0] or Mild [1] with at least a 1-grade improvement from Baseline)
- Time to success in ISGA (i.e., score of 0 (clear) or 1 (almost clear) with at least a 2-grade improvement from baseline)

The Agency commented that secondary endpoints should be clinically meaningful and supportive of the proposed primary efficacy endpoint. In addition, the Agency stated that "the secondary endpoints that the Division recommends include an evaluation of the signs and symptoms of atopic dermatitis (e.g. erythema, induration/population, scaling and oozing/crusting) which should be dichotomized to success/failure a priori in the protocol. These signs should be evaluated globally on a 4-5 point scale and not by body region (as in the EASI score)." During the meeting, the applicant stated that they will include signs and symptoms of atopic dermatitis as exploratory endpoints which are not intended for labeling. In addition, the applicant agreed not to use time to improvement of pruritus as a secondary endpoint. The applicant stated that time to improvement of pruritus will be used as an exploratory endpoint which is not intended for labeling.

On September 23, 2015, the Agency and the applicant met for a Pre-NDA meeting. The Agency provided general comments on how the data should be submitted (data tabulation datasets, data definition files, annotated case report forms, and analysis datasets).

2.1.2 Clinical Studies Overview

The applicant submitted data from a two pivotal Phase 3 trials (Trials 301 and 302). An overview of the trials is presented in Table 2.

Table 2: Clinical Study Overview

Trial	Location	Study Population	Treatment Arms	Number of Subjects	Dates
301	U.S. (47 centers)	Aged 2 years and older, BSA ≥ 5%, and ISGA score of 2 (mild) or 3 (moderate)	EUCRISA Ointment, 2%	503(1)	3/26/2014 – 4/29/2015
			Vehicle Ointment	256	
302	U.S. (42 centers)		EUCRISA Ointment, 2%	513(2)	3/26/2014 – 4/27/2015
			Vehicle Ointment	250	

(1) An additional 4 subjects were randomized to EUCRISA ointment, 2%; however, these subjects were not dispensed study drug.
(2) One additional subject was randomized to EUCRISA ointment, 2%; however, this subject was not dispensed study drug.

2.2 Data Sources

This reviewer evaluated the applicant's clinical study reports, datasets, clinical summaries, and proposed labeling. This submission was submitted in eCTD format and entirely electronic. The datasets in this review are archived at the following locations:
\\cdsesub1\evsprod\NDA207695\0000\m5\datasets\

3 STATISTICAL EVALUATION

3.1 Data and Analysis Quality

The databases for the studies required minimal data management prior to performing analyses and no request for additional datasets were made to the applicant.

3.2 Evaluation of Efficacy

3.2.1 Study Design and Endpoints

The applicant conducted two identically-designed Phase 3 trials (Trials 301 and 302). Both trials were randomized, double-blind, parallel-group, vehicle-controlled, 36-day trials investigating the safety and efficacy of EUCRISA ointment, 2% compared to vehicle ointment for the treatment of atopic dermatitis. For enrollment, the protocol specified the following key inclusion criteria:

- Male or female 2 years of age or older
- Clinical diagnosis of atopic dermatitis according to the criteria of Hanifan and Rajka
- Atopic dermatitis involvement ≥ 5% of treatable body surface area (BSA) excluding scalp
- Investigator's Static Global Assessment (ISGA) score of 2 (mild) or 3 (moderate), see Table 3 for details on the ISGA scale

Each trial was designed to enroll and randomize approximately 750 subjects in a 2:1 ratio to either EUCRISA ointment, 2% (N=500) or vehicle ointment (N=250). Randomization was stratified by center. Subjects applied study product twice daily for 4 weeks. Subjects were evaluated at the following study visits: screening, baseline (Day 1) and Days 8, 15, 22, 29 (end of treatment), and 36 (follow-up / end of study).

Table 3: Investigator's Static Global Assessment (ISGA) Scale

Score	Grade	Definition
0	Clear	Minor residual hypo/hyperpigmentation; no erythema or induration/papulation; no oozing/crusting
1	Almost Clear	Trace faint pink erythema, with barely perceptible induration/papulation and no oozing/crusting
2	Mild	Faint pink erythema with mild induration/papulation and no oozing/crusting
3	Moderate	Pink-red erythema with moderate induration/papulation with or without oozing/crusting
4	Severe	Deep or bright red erythema with severe induration/papulation and with oozing/crusting

The protocol-specified primary efficacy endpoint was the proportion of subjects achieving success in ISGA at Day 29. Success in ISGA was defined as an ISGA score of 0 (clear) or 1 (almost clear) with at least a 2-grade improvement from baseline.

The protocol specified the following two secondary efficacy endpoints:

- Proportion of subjects with an ISGA score of 0 (clear) or 1 (almost clear) at Day 29
- Time to success in ISGA (i.e., score of 0 (clear) or 1 (almost clear) with at least a 2-grade improvement from baseline)

The protocol specified the following two exploratory efficacy endpoints:

- Time to improvement in pruritus (defined as a pruritus score of None [0] or Mild [1] with at least a 1-grade improvement from Baseline)
- Signs of atopic dermatitis (erythema, induration/papulation, exudation [oozing or crusting], excoriation [evidence of scratching], and lichenification [epidermal thickening]) evaluated globally on a 4-point scale and not by body region

Table 4: Severity of Pruritus Scale

Score	Grade	Definition
0	None	No itching
1	Mild	Occasional, slight itching/scratching
2	Moderate	Constant or intermittent itching/scratching which is not disturbing sleep
3	Severe	Bothersome itching/scratching which is disturbing sleep

Table 5: Signs of Atopic Dermatitis

Erythema (Redness)		
Score	**Grade**	**Description**
0	None	No redness
1	Mild	Mildly detectable erythema; pink
2	Moderate	Dull red; clearly distinguishable
3	Severe	Deep, dark red; marked and extensive
Induration/Papulation		
Score	**Grade**	**Description**
0	None	None
1	Mild	Slightly perceptible elevation
2	Moderate	Clearly perceptible elevation but not extensive
3	Severe	Marked and extensive elevation
Exudation (Oozing or Crusting)		
Score	**Grade**	**Description**
0	None	No oozing or crusting
1	Mild	Minor or faint signs of oozing
2	Moderate	Definite oozing or crusting present
3	Severe	Marked and extensive oozing or crusting present
Excoriation (Evidence of Scratching)		
Score	**Grade**	**Description**
0	None	No evidence of excoriation
1	Mild	Mild excoriation present
2	Moderate	Definite excoriation present
3	Severe	Marked, deep, or extensive excoriation present
Lichenification (Epidermal Thickening)		
Score	**Grade**	**Description**
0	None	No epidermal thickening
1	Mild	Minor epidermal thickening
2	Moderate	Moderate epidermal thickening; accentuated skin lines
3	Severe	Severe epidermal thickening; deeply accentuated skin lines

3.2.2 Statistical Methodologies

The primary analysis population specified in the protocol was the intent-to-treat (ITT) population, defined as all randomized subjects. The protocol also specified supportive analyses using the per-protocol (PP). The PP population was defined as all subjects in the ITT population who complete the Day 29 evaluation without any major protocol deviations. Specifically, the protocol specified the following criteria for the PP population:

- Met all of the Inclusion Criteria and none of the Exclusion Criteria
- Have not taken any interfering concomitant medications or therapies during the 29-day study period
- Completed the Day 29 Visit, including the Day 29 efficacy evaluation
- Have applied 80%–120% of the total number of expected doses during the Study Drug Application Period
- Have not missed 6 or more consecutive doses during the Study Drug Application Period
- Were in the visit window (±3 days) for the Day 29 Visit

The protocol specified that the trials were to be conducted in a manner such that a minimum of 12 subjects will be randomized and included in the ITT population for each center. However, the protocol specified a pooling strategy in the event that a center(s) did not enroll at least 12 subjects. The protocol specified combining the center with the smallest enrollment with the largest, restricted to those centers that did not meet the minimum enrollment. Then, the second smallest will be combined with the second largest of those centers that did not meet the minimum enrollment, etc. After the pooling, the centers are referred to as "analysis centers".

The protocol-specified analysis method for the primary efficacy endpoint of success in ISGA at Day 29 was logistic regression with factors of treatment group and analysis center. The protocol specified investigating the consistency of results across analysis center by testing the treatment by analysis center interaction. If the interaction was significant at the 0.10 level, the protocol specified a sensitivity analysis where the data will be analyzed excluding one analysis center at a time to identify the impact of each analysis center on the overall results.

The protocol-specified analysis method for the 1st secondary efficacy endpoint (i.e., proportion of subjects with an ISGA score of 0 or 1 at Day 29) was logistic regression with factors of treatment group and analysis center. For the analysis of the 2nd secondary endpoint (i.e., time to success in ISGA), the protocol specified analyzing this endpoint using the Kaplan-Meier approach and the log-rank test. To control the Type I error rate for testing two secondary efficacy endpoints, the protocol specified analyzing these endpoints sequentially, where the 2nd secondary endpoint would only be tested if the 1st secondary endpoint was significant at the 0.05 level.

The protocol specified the primary imputation method to be the multiple imputation (MI) approach. For each treatment arm separately, missing data was imputed using the Markov Chain Monte Carlo (MCMC) method. The protocol specified the following two sensitivity analyses for the handling of missing data:

- Repeated-measures logistic regression model (GEE), with dichotomized ISGA success as the dependent variable and treatment, analysis center, and visit (i.e., Days 8, 15, 22, and 29) as independent factors. In this analysis, data from all post-baseline visits will be included with no imputation for missing data.
- Model-based multiple imputation method to impute missing data for the dichotomized ISGA data. The imputation model (i.e., logistic regression) will include treatment and analysis center.

3.2.3 Patient Disposition, Demographics and Baseline Characteristics

Trial 301 enrolled and randomized a total 763 subjects (507 to EUCRISA and 256 to vehicle) from 47 centers in the United States. Four of the 507 subjects randomized to EUCRISA ointment were not dispensed study drug and are not included in the ITT population. Trial 302 enrolled and randomized a total of 764 subjects (514 to EUCRISA and 250 to vehicle) from 42 centers in the United States. One of the 514 subjects randomized to EUCRISA ointment was not dispensed study drug and is not included in the ITT population. In both trials, the discontinuation rate was

higher in the vehicle arm compared to EUCRISA arm. The reasons for discontinuation are presented in Table 6.

Table 6: Disposition of Subjects (ITT)

	Trial 301		Trial 302	
	EUCRISA (N=503)	**Vehicle (N=256)**	**EUCRISA (N=513)**	**Vehicle (N=250)**
Discontinued	30 (6%)	31 (12%)	31 (6%)	37 (15%)
Adverse Event	7 (1%)	2 (1%)	5 (1%)	4 (2%)
Lost to Follow-Up	5 (1%)	4 (2%)	4 (1%)	4 (2%)
Other	3 (1%)	1 (<1%)	2 (<1%)	6 (2%)
Withdrawal by Parent/Guardian	12 (2%)	18 (7%)	14 (3%)	20 (8%)
Withdrawal by Subject	3 (1%)	6 (2%)	6 (1%)	3 (1%)

Source: Reviewer's Analysis (same results as Applicant's Analysis)

The demographics and baseline disease characteristics are presented in Table 7. The demographics and baseline disease characteristics were generally balanced across the treatment arms within each trial and were similar between each trial. Approximately 38% and 39% of subjects had an ISGA score of 2 (mild) at baseline in Trials 301 and 302, respectively.

Table 7: Demographics and Baseline Disease Characteristics (ITT)

	Trial 301		Trial 302	
	EUCRISA (N=503)	**Vehicle (N=256)**	**EUCRISA (N=513)**	**Vehicle (N=250)**
Age (years)				
Mean (SD)	12.0 (11.6)	12.4 (10.7)	12.6 (12.7)	11.8 (12.6)
Median	9.0	10.0	9.0	8.5
Range	2 – 65	2 – 63	2 – 79	2 – 79
Categories				
2-6	162 (32%)	78 (30%)	173 (34%)	93 (37%)
7-11	155 (31%)	73 (29%)	137 (27%)	71 (28%)
12-17	121 (24%)	67 (26%)	126 (25%)	57 (23%)
18+	65 (13%)	38 (15%)	77 (15%)	29 (12%)
Gender				
Male	219 (44%)	113 (44%)	231 (45%)	112 (45%)
Female	284 (56%)	143 (56%)	282 (55%)	138 (55%)
Race				
White	308 (61%)	162 (63%)	309 (60%)	144 (58%)
Black	138 (27%)	61 (24%)	147 (29%)	78 (31%)
Asian	26 (5%)	17 (7%)	26 (5%)	10 (4%)
American Indian or Alaska Native	8 (2%)	3 (1%)	3 (1%)	2 (1%)
Native Hawaiian or Other Pacific Islander	0	4 (2%)	7 (1%)	4 (2%)
Other	23 (5%)	9 (4%)	21 (4%)	12 (5%)
ISGA				
2 – Mild	196 (39%)	93 (36%)	197 (38%)	100 (40%)
3 – Moderate	307 (61%)	163 (64%)	316 (62%)	150 (60%)

Source: Reviewer's Analysis (same results as Applicant's Analysis)
SD: Standard Deviation

3.2.4 Primary Efficacy Results

Table 8 presents the results for the primary efficacy endpoint at Day 29 for both trials in the ITT population. EUCRISA ointment, 2% was statistically superior (p-values ≤ 0.038) to vehicle ointment on the primary efficacy endpoint in both trials. The success proportions for the EUCRISA arm were similar between the two trials (i.e., 32.8% for Trial 301 and 31.4% for Trial 302); however, the success proportion for the vehicle arm was higher in Trial 301 compared to Trial 302 (i.e., 25.4% vs. 18.0%). The results for the primary efficacy endpoint at Day 29 in the PP population are presented in Table 9. While the success proportions in the PP population are generally similar to those in the ITT population, the comparison in Trial 301 is no longer statistically significant (p-value = 0.088); however, this could be due to the combination of a smaller sample size and a slightly higher success proportion in the vehicle arm.

Table 8: Results for the Primary Efficacy Endpoint at Day 29 (ITT, MI[1])

	Trial 301			Trial 302		
Endpoint	**EUCRISA (N=503)**	**Vehicle (N=256)**	**P-Value[2]**	**EUCRISA (N=513)**	**Vehicle (N=250)**	**P-Value[2]**
Success in ISGA[3]	32.8%	25.4%	0.038	31.4%	18.0%	<0.001

Source: Reviewer's Analysis (same results as Applicant's Analysis)
(1) Missing data was imputed using multiple imputation (MI). The values displayed are the averages over the 140 imputed datasets for Trial 301 and the 135 datasets for Trial 302.
(2) P-value from a logistic regression (using Firth's Penalized Likelihood) with treatment and analysis center as factors.
(3) Success is defined as an ISGA score of 0 (clear) or 1 (almost clear) with at least a 2-grade improvement from baseline.

Table 9: Results for the Primary Efficacy Endpoint at Day 29 (PP)

	Trial 301			Trial 302		
Endpoint	**EUCRISA (N=435)**	**Vehicle (N=201)**	**P-Value[1]**	**EUCRISA (N=454)**	**Vehicle (N=208)**	**P-Value[1]**
Success in ISGA[2]	32.4%	26.9%	0.088	32.2%	18.3%	<0.001

Source: Reviewer's Analysis (same results as Applicant's Analysis)
(1) P-value from a logistic regression (using Firth's Penalized Likelihood) with treatment and analysis center as factors.
(2) Success is defined as an ISGA score of 0 (clear) or 1 (almost clear) with at least a 2-grade improvement from baseline.

Table 10 provides the number of subjects with missing data for the primary efficacy endpoint by week, treatment arm, and trial. The missing data profiles were similar between the two trials. For the primary time-point (i.e., Day 29), the proportion of subjects with missing data was approximately two times higher in the vehicle arm compared to the EUCRISA arm in both trials.

Table 10: Missing Data for the Primary Efficacy Endpoint by Day (ITT)

	Trial 301		Trial 302	
	EUCRISA (N=503)	**Vehicle (N=256)**	**EUCRISA (N=513)**	**Vehicle (N=250)**
Baseline	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Day 8	8 (2%)	7 (3%)	5 (1%)	6 (2%)
Day 15	13 (3%)	19 (7%)	25 (5%)	11 (4%)
Day 22	25 (5%)	28 (11%)	26 (5%)	13 (5%)
Day 29	25 (5%)	28 (11%)	27 (5%)	26 (10%)

Source: Reviewer's Analysis

The primary imputation method was the multiple imputation (MI) approach using the Markov Chain Monte Carlo (MCMC) method to impute the missing data. The protocol specified the following two sensitivity analyses for the handling of missing data: (i) not impute missing data and analyze using a repeated-measures logistic regression (GEE) with treatment, analysis center, and visit (i.e., Days 8, 15, 22, and 29) in the model and (ii) impute missing data using model-based multiple imputation with treatment and analysis center in the imputation model. In the study reports, the applicant conducted an additional sensitivity analysis where missing data was imputed using the last observation carried forward (LOCF). This reviewer conducted an additional sensitivity analysis where missing data is imputed as failures. The results for the sensitivity analyses as well as the primary imputation method are presented in Table 11. For both trials, the results were generally similar across the various methods for handling missing data.

Table 11: Comparison of Different Approaches for Handling Missing Data for Success in ISGA(1) at Day 29 (ITT)

	Trial 301			Trial 302		
Endpoint	**EUCRISA (N=503)**	**Vehicle (N=256)**	**P-Value**	**EUCRISA (N=513)**	**Vehicle (N=250)**	**P-Value**
MI-MCMC (primary)(2)	32.8%	25.4%	0.038	31.4%	18.0%	<0.001
Observed Data(3)	33.5%	26.8%	0.039	31.9%	19.2%	<0.001
MI-Reg(4)	33.4%	26.8%	0.070	31.9%	19.2%	<0.001
LOCF(5)	32.2%	23.8%	0.014	30.6%	17.6%	<0.001
Failure(6)	31.8%	23.8%	0.018	30.2%	17.2%	<0.001

Source: Reviewer's Analysis

(1) Success is defined as an ISGA score of 0 (clear) or 1 (almost clear) with at least a 2-grade improvement from baseline.
(2) Multiple imputation (MI) where missing data is imputed using the MCMC method. The values displayed are the averages over the 140 imputed datasets for Trial 301 and the 135 datasets for Trial 302. P-value from a logistic regression (using Firth's Penalized Likelihood) with treatment and analysis center as factors.
(3) Missing data is not imputed. P-value based on GEE analysis with treatment, analysis center, and visit (i.e., Days 8, 15, 22, and 29) in the model.
(4) Multiple imputation (MI) where missing data is imputed using a regression model with treatment and analysis center as factors. The values displayed are the averages over the 140 imputed datasets for Trial 301 and the 135 datasets for Trial 302. P-value from a logistic regression (using Firth's Penalized Likelihood) with treatment and analysis center as factors.
(5) Missing data imputed using the last observation carried forward. P-value based on a Cochran-Mantel-Haenszel (CMH) test stratified by analysis center. The p-value from a logistic regression (using Firth's Penalized Likelihood) with treatment and analysis center as factors would be 0.015 for Trial 301 and <0.001for Trial 302.
(6) Missing data imputed as failures. P-value based on a Cochran-Mantel-Haenszel (CMH) test stratified by analysis center. The p-value from a logistic regression (using Firth's Penalized Likelihood) with treatment and analysis center as factors would be 0.020 for Trial 301 and for <0.001 Trial 302.

3.2.5 Secondary Efficacy Results

Tables 12 and 13 present the results for the secondary efficacy endpoints in both trials for the ITT and PP populations, respectively. In both trials, EUCRISA ointment, 2% was statistically superior (p-values ≤ 0.005) to vehicle ointment on both secondary efficacy endpoints. It should be noted that the median time to success in ISGA (i.e., the time at which 50% of the subjects achieved success in ISGA) could not be calculated, as fewer than 50% of subjects achieved success in ISGA. Figure 1 presents the proportion of subjects who achieve success in ISGA over time (i.e., Days 1, 8, 15, 22 and 29) for Trials 301 and 302.

Table 12: Results for the Secondary Efficacy Endpoints at Day 29 (ITT, MI[1])

Endpoint	Trial 301			Trial 302		
	EUCRISA (N=503)	Vehicle (N=256)	P-Value	EUCRISA (N=513)	Vehicle (N=250)	P-Value[2]
ISGA score of Clear or Almost Clear	51.7%	40.6%	0.005[2]	48.5%	29.7%	<0.001[2]
Time to Success in ISGA[3] Median[4]	NC	NC	<0.001[5]	NC	NC	<0.001[5]

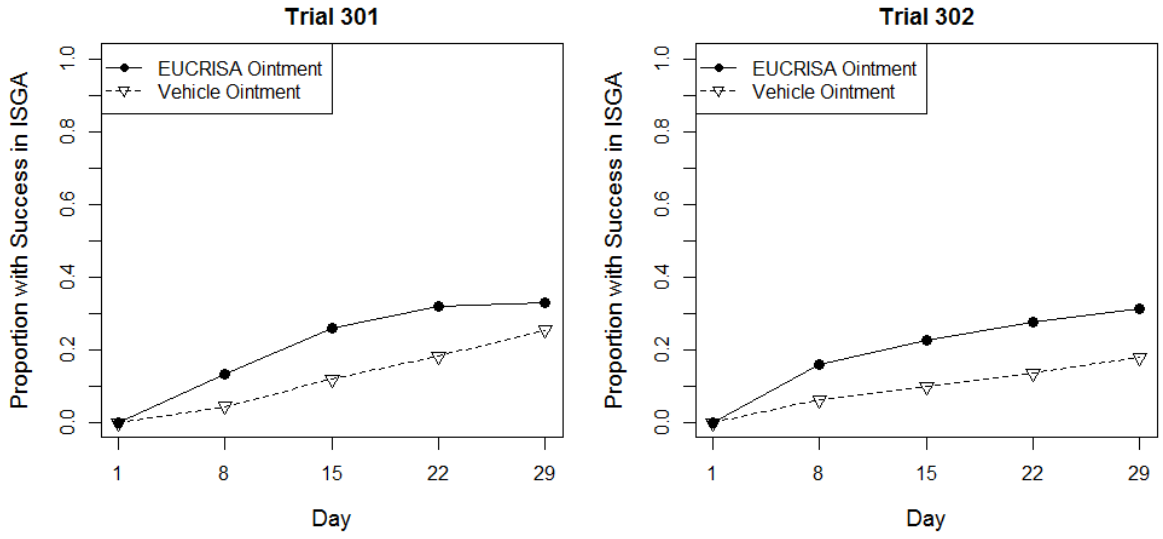
Source: Reviewer's Analysis (same results as Applicant's Analysis)
(1) Missing data was imputed using multiple imputation (MI). The values displayed are the averages over the 140 imputed datasets for Trial 301 and the 135 datasets for Trial 302.
(2) P-value from a logistic regression (using Firth's Penalized Likelihood) with treatment and analysis center as factors.
(3) Success is defined as an ISGA score of 0 (clear) or 1 (almost clear) with at least a 2-grade improvement from baseline.
(4) Median time to success in ISGA could not be calculated because fewer than 50% of subjects achieved success in ISGA.
(5) P-value based on a log-rank test.

Table 13: Results for the Secondary Efficacy Endpoints at Day 29 (PP)

Endpoint	Trial 301			Trial 302		
	EUCRISA (N=435)	Vehicle (N=201)	P-Value	EUCRISA (N=454)	Vehicle (N=208)	P-Value
ISGA score of Clear or 1 Almost Clear	51.7%	43.8%	0.032[1]	50.0%	29.8%	<0.001[1]
Time to Success in ISGA[2] Median[3]	NC	NC	0.003[4]	NC	NC	<0.001[4]

Source: Reviewer's Analysis (same results as Applicant's Analysis)
(1) P-value from a logistic regression (using Firth's Penalized Likelihood) with treatment and analysis center as factors.
(2) Success is defined as an ISGA score of 0 (clear) or 1 (almost clear) with at least a 2-grade improvement from baseline.
(3) Median time to success in ISGA could not be calculated because fewer than 50% of subjects achieved success in ISGA.
(4) P-value based on a log-rank test.

Figure 1: Success in ISGA over Time (ITT, MI)



Source: Reviewer's Analysis

3.3 Evaluation of Safety

3.3.1 Extent of Exposure

The extent of exposure to study product is presented in Table 14. In both trials, the number of applications and the number of dosing days was on average slightly higher in the EUCRISA arm compared to the vehicle arm; however, this is probably due to the higher dropout rate in the vehicle arm.

Table 14: Extent of Exposure (Safety Population)

	Trial 301		Trial 302	
	EUCRISA	**Vehicle**	**EUCRISA**	**Vehicle**
Number of Applications				
N	503	256	513	250
Mean (SD)	55.2 (8.8)	52.4 (13.4)	54.3 (9.5)	52.3 (11.2)
Median	56.0	56.0	56.0	56.0
Range	1 – 101	1 – 88	1 – 76	1 – 73
Number of Dosing Days				
N	503	256	513	250
Mean (SD)	28.2 (4.3)	26.8 (6.7)	27.8 (4.6)	26.9 (5.6)
Median	29.0	28.0	28.0	28.0
Range	1 – 52	1 – 46	1 – 38	1 – 28
Amount of Drug Used (grams)				
N	478	240	482	237
Mean (SD)	171.6 (194.2)	167.1 (187.9)	167.0 (168.5)	170.6 (154.7)
Median	109.8	106.4	120.0	121.0
Range	2 – 1602	1 – 1356	2 – 1328	7 – 991

Source: pg. 93 of Study Report for Trial 301 and pg. 94 of Study Report for Trial 302.

3.3.2 Adverse Events

Table 15 presents an overview of the adverse events reported during both trials. The treatment-emergent adverse events (TEAE) reported in at least 1% of subjects within either treatment group for trials are presented in Table 16.

Table 15: Overview of Adverse Events Reported (Safety Population)

	Trial 301		Trial 302		Pooled Trials	
Subjects With:	**EUCRISA (N=502)**	**Vehicle (N=252)**	**EUCRISA (N=510)**	**Vehicle (N=247)**	**EUCRISA (N=1012)**	**Vehicle (N=499)**
Any TEAEs	147 (29%)	50 (20%)	150 (29%)	79 (32%)	297 (29%)	129 (26%)
Maximum Severity of TEAE						
Mild	77 (15%)	26 (10%)	88 (17%)	44 (18%)	165 (16%)	70 (14%)
Moderate	62 (12%)	20 (8%)	52 (10%)	33 (13%)	114 (11%)	53 (11%)
Severe	8 (2%)	4 (2%)	10 (2%)	2 (1%)	18 (2%)	6 (1%)
Any Serious AEs	4 (1%)	1 (<1%)	3 (1%)	0	7 (1%)	1 (<1%)
Any TEAEs Leading to Discontinuation	7 (1%)	2 (1%)	5 (1%)	4 (2%)	12 (1%)	6 (1%)

Source: pg. 93 of Summary of Clinical Safety.

Table 16: Treatment-Emergent Adverse Events (TEAE) Reported by ≥1% of Subjects Within Any Treatment Group and Trial (Safety Population)

System Organ Class / Preferred Term	Trial 301		Trial 302		Pooled Trials	
	EUCRISA (N=502)	Vehicle (N=252)	EUCRISA (N=510)	Vehicle (N=247)	EUCRISA (N=1012)	Vehicle (N=499)
Gastrointestinal disorders						
Diarrhea	3 (1%)	0 (0%)	6 (1%)	2 (1%)	9 (1%)	2 (<1%)
Vomiting	8 (2%)	3 (1%)	7 (1%)	2 (1%)	15 (1%)	6 (1%)
General disorders and administration site conditions						
Application site pain	31 (6%)	3 (1%)	14 (3%)	3 (1%)	45 (4%)	6 (1%)
Application site pruritus	4 (1%)	3 (1%)	1 (<1%)	3 (1%)	5 (<1%)	6 (1%)
Application site urticaria	2 (<1%)	0 (0%)	0 (0%)	3 (1%)	2 (<1%)	3 (1%)
Pyrexia	12 (2%)	3 (1%)	7 (1%)	4 (2%)	19 (2%)	7 (1%)
Infections and infestations						
Nasopharyngitis	9 (2%)	0 (0%)	9 (2%)	6 (2%)	18 (2%)	6 (1%)
Staphylococcal skin infection	0 (0%)	1 (<1%)	1 (<1%)	4 (2%)	1 (<1%)	5 (1%)
Upper respiratory tract infection	14 (3%)	10 (4%)	16 (3%)	5 (2%)	30 (3%)	15 (3%)
Nervous System disorders						
Headache	5 (1%)	0 (0%)	6 (1%)	1 (<1%)	11 (1%)	1 (<1%)
Respiratory, thoracic and mediastinal disorders						
Cough	5 (1%)	1 (<1%)	7 (1%)	7 (3%)	12 (1%)	8 (2%)
Nasal congestion	7 (1%)	0 (0%)	1 (<1%)	2 (1%)	8 (1%)	2 (<1%)
Oropharyngeal pain	4 (1%)	0 (0%)	7 (1%)	2 (1%)	11 (1%)	2 (<1%)
Skin and subcutaneous tissue disorders						
Dermatitis atopic	3 (1%)	2 (1%)	4 (1%)	6 (2%)	7 (1%)	8 (2%)
Eczema	1 (<1%)	0 (0%)	3 (1%)	3 (1%)	4 (<1%)	3 (1%)
Pruritus	2 (<1%)	0 (0%)	4 (1%)	3 (1%)	6 (1%)	3 (1%)

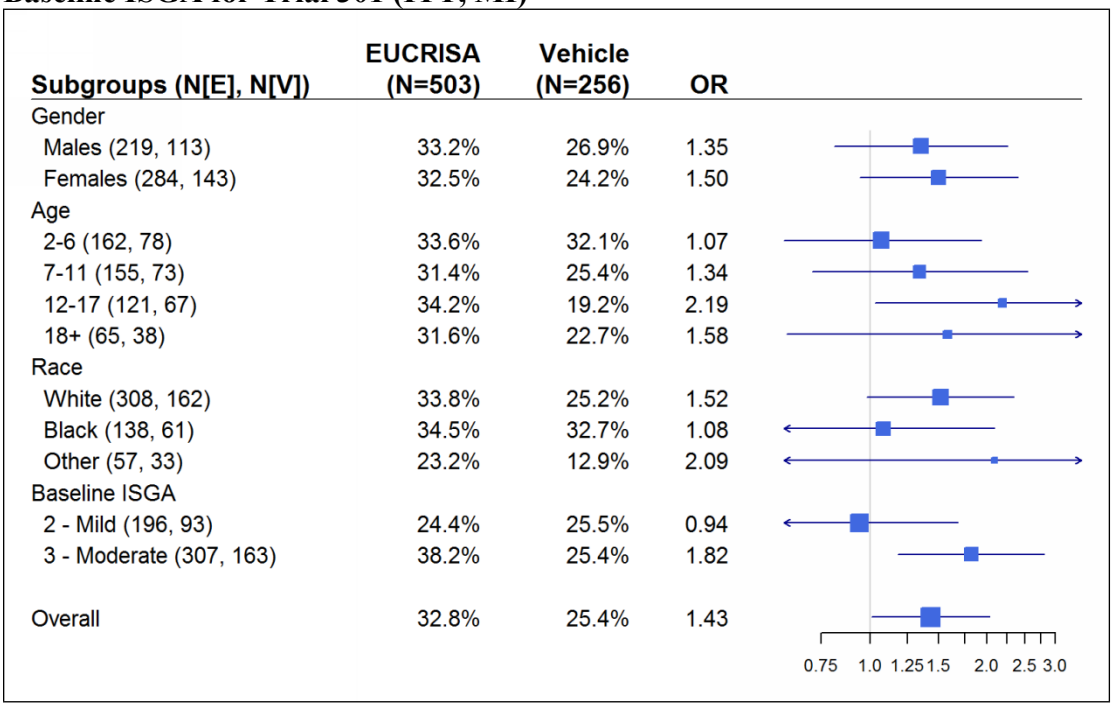
Source: pg. 383- of Study Report for Trial 301 and pg. 239-251 of Study Report for Trial 302.

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Race, Age, and Baseline Disease Severity

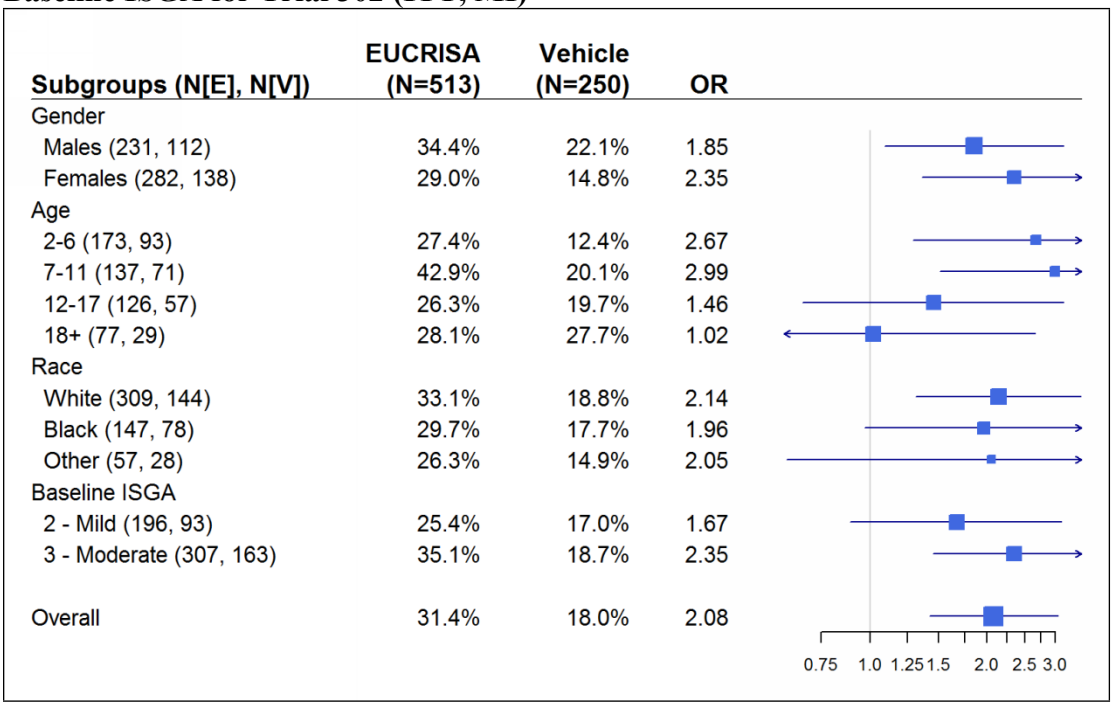
Figures 2 and 3 present the results for the primary efficacy endpoint at Day 29 by gender, age (2-6, 7-11, 12-17, and 18+ years), race (white, black, and other) and baseline ISGA score. For gender, the treatment effect was greater in females in both trials. The response rate for the EUCRISA ointment arm was consistently higher than the vehicle ointment arm across the age subgroups in both trials; however, the treatment effect was variable across the age subgroups and trials. For race, the treatment effect was higher in whites compared to blacks in both trials. In Trial 301, the response rate for the EUCRISA ointment arm was slightly smaller than the vehicle ointment arm in subjects with a baseline ISGA score of 2 (mild); however, this is not the case in Trial 302.

Figure 2: Results for the Primary Efficacy Endpoint at Day 29 by Gender, Age, Race and Baseline ISGA for Trial 301 (ITT, MI)



Source: Reviewer's Analysis
*N[E] = subgroup sample size in EUCRISA arm, N[V] = subgroup sample size in vehicle arm, 95% CI for forest plot

Figure 3: Results for the Primary Efficacy Endpoint at Day 29 by Gender, Age, Race and Baseline ISGA for Trial 302 (ITT, MI)



Source: Reviewer's Analysis
*N[E] = subgroup sample size in EUCRISA arm, N[V] = subgroup sample size in vehicle arm, 95% CI for forest plot

4.2 Center

Trial 301 enrolled subjects from 47 centers in the United States and Trial 302 enrolled subjects from 42 centers in the United States. The protocol specified a pooling strategy for centers that enrolled less than 12 subjects. These centers were pooled by ordering and combining the smallest with the largest. The process repeated until all centers had at least 12 subjects. For Trial 301, 19 of the 47 centers enrolled less than 12 subjects and the pooling process yielded 37 analysis centers (29 unpooled and 8 pooled). For Trial 302, 17 of the 42 centers enrolled less than 12 subjects and the pooling process yielded a total of 34 analysis centers (28 unpooled and 6 pooled).

Figures 4 and 5 present the results for the primary efficacy endpoint at Day 29 by analysis centers. In both trials, the efficacy rates varied among centers. Some centers had higher efficacy with vehicle ointment than with EUCRISA ointment. The applicant investigated the consistency of results across analysis centers by testing the treatment by analysis center interaction in the logistic regression. If the interaction was significant at the 0.10 level, the protocol specified a sensitivity analysis where the data will be analyzed excluding one analysis center at a time to identify the impact of each analysis center on the overall results. The p-values for the treatment by analysis interaction were 0.992 and 0.951 for Trials 301 and 302, respectively.

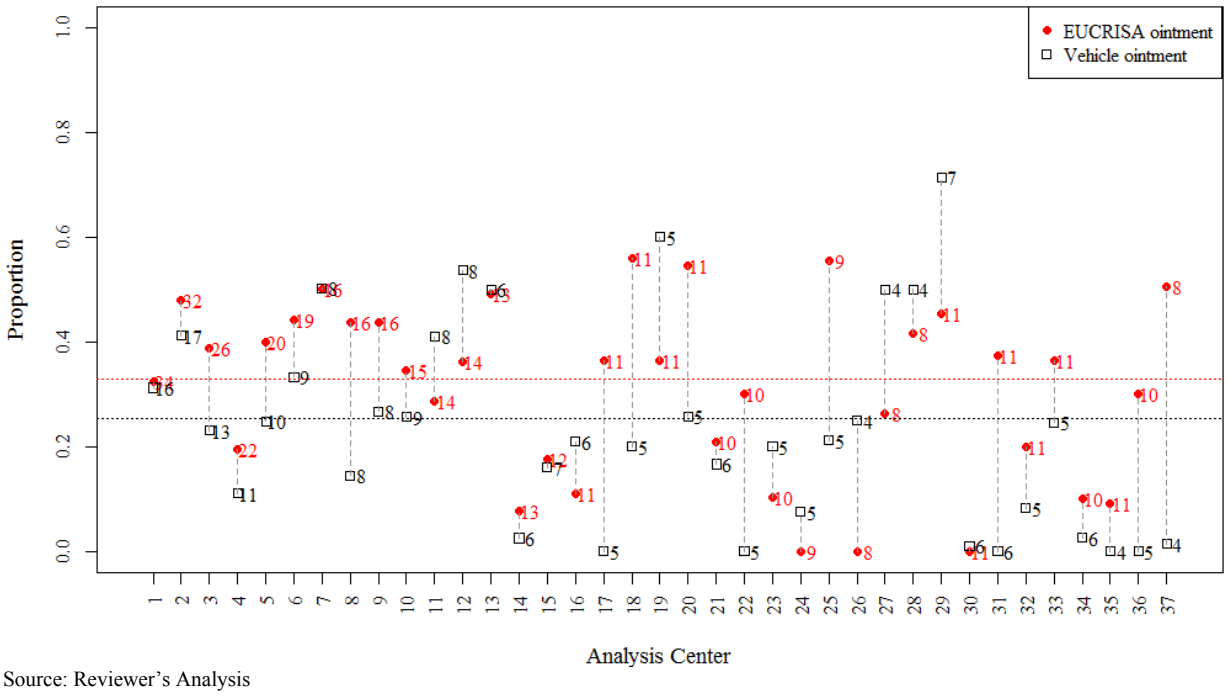
As the pooling process could mask center effects, this reviewer conducted a sensitivity analysis where each center (prior to pooling) was removed and the primary efficacy endpoint was then analyzed. For Trial 301, while the sensitivity analysis identified 10 centers where the removal of any one of them resulted in a p-value greater than 0.05, there was only a very slight change in the response rates with the removal of each center, see Table 17. Therefore, the change in the p-value is primarly driven by the decrease in sample size with the removal of the identified centers. For Trial 302, the removal of any one center did not affect the overall conclusions (p-values <0.001).

Table 17: Centers Identified by Reviwer's Sensitivity Anlysis for Trial 301

	EUCRISA (N=503)	Vehicle (N=256)	P-value
Overall	32.8%	25.4%	0.038
Without (N_E, N_V)*:			
Center 138 (477, 243)	32.5%	25.5%	0.057
Center 150 (483, 246)	32.5%	25.4%	0.051
Center 126 (487, 248)	32.4%	25.8%	0.064
Center 106 (492, 251)	32.7%	25.9%	0.060
Center 110 (492, 251)	32.3%	25.5%	0.058
Center 145 (492, 251)	32.3%	25.4%	0.053
Center 144 (493, 251)	32.9%	25.9%	0.055
Center 152 (494, 251)	32.4%	25.5%	0.056
Center 142 (496, 252)	32.5%	25.8%	0.061
Center 122 (497, 252)	32.8%	25.8%	0.053

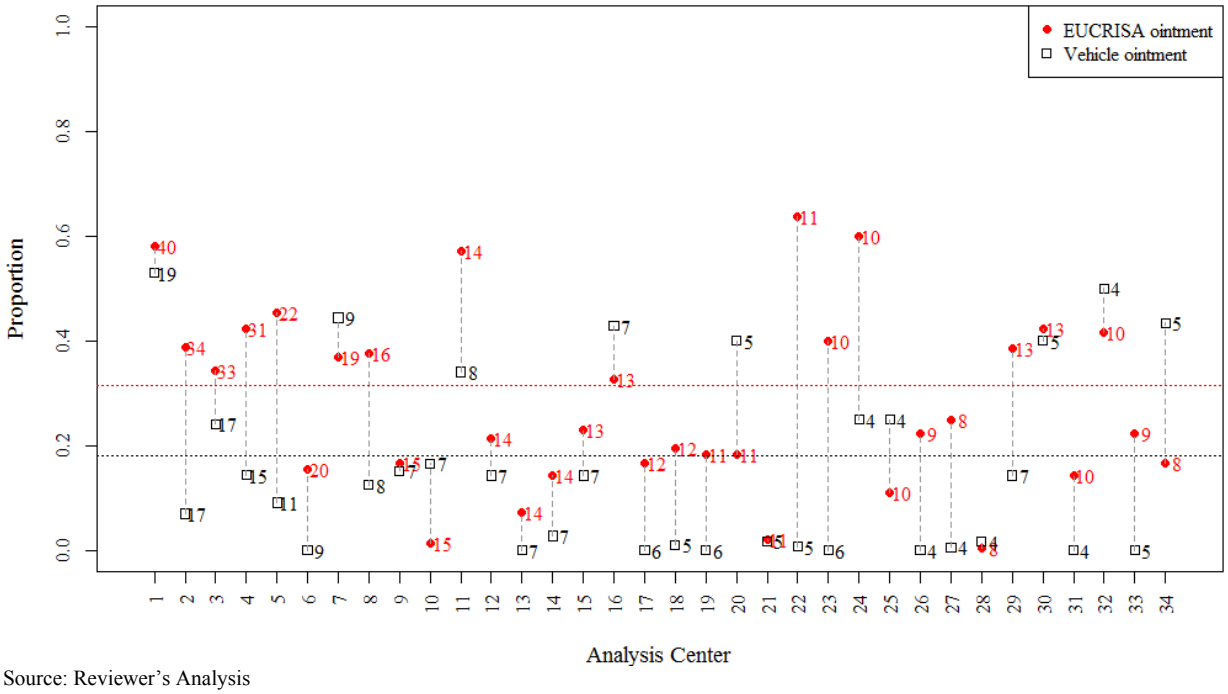
Source: Reviewer's Analysis
*Sample size for the EUCRISA arm (N_E) and the vehicle arm (N_V) after removal of the specified center.

Figure 4: Results for the Primary Efficacy Endpoint at Day 29 by Analysis Centers in Trial 301 (ITT, MI)



Source: Reviewer's Analysis

Figure 5: Results for the Primary Efficacy Endpoint at Day 29 by Analysis Centers in Trial 301 (ITT, MI)



Source: Reviewer's Analysis

5 SUMMARY AND CONCLUSIONS

5.1 Statistical Issues and Collective Evidence

The applicant submitted data from two identically-designed, randomized, multicenter, vehicle-controlled, parallel-group, pivotal Phase 3 trials (Trials 301 and 302). The trials enrolled subjects 2 years of age and older with a clinical diagnosis of atopic dermatitis with body surface area (BSA) involvement ≥ 5% (excluding scalp) and an Investigator's Static Global Assessment (ISGA) score of 2 (mild) or 3 (moderate). The protocol-specified primary efficacy endpoint was the proportion of subjects achieving success in ISGA at Day 29, where success in ISGA was defined as an ISGA score of 0 (clear) or 1 (almost clear) with at least a 2-grade improvement from baseline. The protocol specified the following two secondary efficacy endpoints: the proportion of subjects with an ISGA score of 0 (clear) or 1 (almost clear) at Day 29 and the time to success in ISGA.

Table 18 presents the results of the primary and secondary efficacy endpoints. In both trials, EUCRISA ointment, 2% was statistically superior (p-values ≤ 0.038) to vehicle ointment for all endpoints presented in Table 18. It should be noted that the median time to success in ISGA (i.e., the time at which 50% of the subjects achieved success in ISGA) could not be calculated, as fewer than 50% of subjects achieved success in ISGA.

Table 18: Efficacy Results at Day 29 (ITT, MI[1])

	Trial 301			Trial 302		
Endpoints	**EUCRISA (N=503)**	**Vehicle (N=256)**	**P-Value**	**EUCRISA (N=513)**	**Vehicle (N=250)**	**P-Value**
Primary: Success in ISGA[2]	32.8%	25.4%	0.038[3]	31.4%	18.0%	<0.001[3]
Secondary: ISGA score of Clear or Almost Clear	51.7%	40.6%	0.005[3]	48.5%	29.7%	<0.001[3]
Time to Success in ISGA[2]	NC[4]	NC	<0.001[5]	NC	NC	<0.001[5]

Source: Reviewer's Analysis (same results as Applicant's Analysis)

(1) Missing data was imputed using multiple imputation (MI). The values displayed are the averages over the 140 imputed datasets for Trial 301 and the 135 datasets for Trial 302.
(2) Success is defined as an ISGA score of 0 (clear) or 1 (almost clear) with at least a 2-grade improvement from baseline.
(3) P-value from a logistic regression (using Firth's Penalized Likelihood) with treatment and analysis center as factors.
(4) Median time to success in ISGA could not be calculated because fewer than 50% of subjects achieved success in ISGA.
(5) P-value based on a log-rank test.

For the handling of missing data, the results for the primary efficacy endpoint were similar between the primary imputation method (i.e., multiple imputation using MCMC) and the applicant's sensitivity analyses. This reviewer conducted an additional sensitivity analysis where missing data was imputed as failures and the results were similar to those of the primary imputation method and the applicant's sensitivity analyses.

In both trials, the efficacy rates varied among centers. Some centers had higher efficacy with vehicle ointment than with EUCRISA ointment. The applicant's investigation of the treatment by center interaction focused on the effects after pooling. As the pooling process could mask center effects, this reviewer conducted a sensitivity analysis where each center (prior to pooling) was removed and the primary efficacy endpoint was then analyzed. For Trial 301, while the sensitivity analysis identified 10 centers where the removal of any one of them resulted in a p-value greater than 0.05, there was only a very slight change in the response rates with the removal of each center. Therefore, the change in the p-value is primarly driven by the decrease in sample size with the removal of the identified centers. For Trial 302, the removal of any one center did not affect the overall conclusions (p-values <0.001).

5.2 Conclusions and Recommendations

Efficacy findings from two pivotal Phase 3 trials (Trials 301 and 302) established the efficacy of EUCRISA ointment, 2% for the treatment of mild to moderate atopic dermatitis in patients 2 years of age and older.

SIGNATURES/DISTRIBUTION LIST

Primary Statistical Reviewer: Matthew Guerra, Ph.D.
Date: August 19, 2016

Statistical Team Leader: Mohamed Alosh, Ph.D.
Date: August 19, 2016

cc:
DDDP/Marcus
DDDP/Lindstrom
DDDP/Trajkovic
DDDP/McCord
DDDP/Phillips
DDDP/Springs
OBIO/Patrician
DBIII/Wilson
DBIII/Alosh
DBIII/Guerra

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

/s/

MATTHEW W GUERRA
08/19/2016

MOHAMED A ALOSH
08/19/2016