[redacted]

STATISTICAL ANALYSIS PLAN

The RENEW Trial: A Multi-Center, Randomized, Double-Masked, Parallel-Group, Vehicle-Controlled, Adaptive Phase 3 Clinical Trial to Assess the Safety and Efficacy of Reproxalap 0.25% Ophthalmic Solution Compared to Vehicle in Subjects with Dry Eye Disease: Part 1

Sponsor: Aldeyra Therapeutics, Inc.

Protocol Number: ADX-102-DED-012 Part 1

Author: [redacted]

Date: 31/OCT/2019

Version: 2.0

[redacted]

The RENEW Trial: A Multi-Center, Randomized, Double-Masked, Parallel-Group, Vehicle-Controlled, Adaptive Phase 3 Clinical Trial to Assess the Safety and Efficacy of Reproxalap 0.25% Ophthalmic Solution Compared to Vehicle in Subjects with Dry Eye Disease: Part 1

Protocol Number:	**ADX-102-DED-012 Part 1**
Version:	**2.0**
Date:	**31/OCT/2019**

Statistical Analysis Plan Approval

[redacted]

[redacted]

Table of Contents

List of Abbreviations

AE	Adverse Event
ANCOVA	Analysis of Covariance
ATC	Anatomical Therapeutic Chemical
BID	*Bis in die* (Twice Daily)
CAC	Conjunctival Allergen Challenge
[redacted]	[redacted]
CI	Confidence Interval
CRO	Contract Research Organization
CS	Clinically Significant
eCRF	Electronic Case Report Form
ETDRS	Early Treatment of Diabetic Retinopathy Study
FSN	Fluorescein Staining of the Nasal Region
ICH	International Conference on Harmonisation
IOP	Intraocular Pressure
IP	Investigational Product
ITT	Intent-to-Treat
ITTFSN	Intent-to-Treat Fluorescein Nasal Score
ITTOD	Intent-to-Treat Ocular Dryness
IWRS	Interactive Web Response System
logMAR	Logarithm of the Minimum Angle of Resolution
LS	Least Squares
MAR	Missing at Random
MCMC	Markov Chain Monte Carlo
MedDRA	Medical Dictionary for Regulatory Activities
mmHg	Millimeters of Mercury
MMRM	Mixed Model for Repeated Measures
MNAR	Missing Not at Random
mOsm	Milliosmoles
NCS	Not Clinically Significant
OD	*Oculus Dexter* (Right Eye)
OS	*Oculus Sinister* (Left Eye)
OSDI	Ocular Surface Disease Index
OU	*Oculus Uterque* (Both Eyes)
PDF	Portable Document Format
PP	Per Protocol

PPOD	Per Protocol Ocular Dryness
PPFSN	Per Protocol Fluorescein Nasal Score
PT	Preferred Term
Q-B	QID to BID dosing
Q-Q	QID to QID dosing
QID	*Quater in Die* (Four Times Daily)
RTF	Rich Text Format
SAE	Serious Adverse Event
SANDE	Symptom Assessment in Dry Eye
SAP	Statistical Analysis Plan
SAS	Statistical Analysis Software
SD	Standard Deviation
[redacted]	[redacted]
SE	Standard Error
SOC	System Organ Class
SOP	Standard Operating Procedure
TEAE	Treatment-Emergent Adverse Event
TFBUT	Tear Film Break-Up Time
VA	Visual Acuity
VAS	Visual Analog Scale
WHODrug	World Health Organization Drug Dictionary

1. Introduction

The purpose of this statistical analysis plan (SAP) is to describe the planned analyses and reporting for Part 1 of protocol ADX-102-DED-012, version 2.0 dated 5 MAR 2019. This statistical analysis plan pertains to Part 1 only. The statistical analysis plan for Part 2 will be finalized after database lock for Part 1.

This SAP is being written with due consideration of the recommendations outlined in the most recent International Conference on Harmonisation (ICH) E9 Guideline entitled Guidance for Industry: Statistical Principles for Clinical Trials and the most recent ICH E3 Guideline, entitled Guidance for Industry: Structure and Content of Clinical Study Reports.

This SAP describes the data that will be analyzed and the subject characteristics, efficacy, and safety assessments that will be evaluated. This SAP provides details of the specific statistical methods that will be used. The statistical analysis methods presented in this document will supersede the statistical analysis methods described in the clinical protocol. If additional analyses are required to supplement the planned analyses described in this SAP, they may be completed and will be identified in the clinical study report.

2. Study Objectives

2.1 Primary Measures

The coprimary efficacy measures for the trial are the following:

- Subject-reported ocular dryness score measured by the Visual Analog Scale (VAS) (0-100 scale). The population used for analysis of ocular dryness will be the intent-to-treat ocular dryness (ITTOD) population, which consists of subjects who have a baseline ocular dryness score of ≥ 3 [redacted]).
- Fluorescein staining of the nasal region (FSN) score graded on the [redacted] (0-4 scale). The population used for analysis of FSN score will be intent-to-treat fluorescein nasal score (ITTFSN) population, which consists of subjects who have a baseline (Day 1) FSN score of ≥ 2.

2.2 Key Secondary Measure

The key secondary efficacy measures for the trial include the following:

- [redacted] Symptom Questionnaire for Dryness

2.3 Secondary Measures

The secondary efficacy measures for the trial include the following:

- Fluorescein staining ([redacted])
- Unanesthetized Schirmer's Test

- Ocular Dryness through Week 4 (0-100 VAS)

2.4 Exploratory Measures

The exploratory efficacy measures for the trial include the following:

- Burning/Stinging, Itching, Foreign Body Sensation, Eye Discomfort, Photophobia, and Pain Score (0-100 VAS)
- [redacted] Discomfort Scale
- [redacted] Symptom Questionnaire [redacted]
- Ocular Surface Disease Index (OSDI)©
- Symptom Assessment in Dry Eye (SANDE) questionnaire
- [redacted] Ocular Itching Scale
- Lissamine green staining ([redacted]
- Tear film break-up time (TFBUT)
- Tear osmolarity

2.5 Safety Measures

The safety measures for the trial include the following:

- Visual acuity (VA) at distance utilizing an Early Treatment of Diabetic Retinopathy Study (ETDRS) chart
- Slit-lamp evaluation
- Adverse event (AE) query (reported, elicited and observed)
- Intraocular Pressure (IOP)
- Dilated fundoscopy

2.6 Statistical Hypotheses

The following hypotheses will be tested comparing reproxalap to vehicle. In Part 1, null hypothesis testing will be conducted for each dosing regimen (four times daily [QID] to twice daily [BID], and QID) separately. Dryness and staining null hypotheses must be rejected for the dosing regimen to claim efficacy. The null and alternative hypotheses, based on the primary variables, are as follows:

> H_{01}: There is no difference between reproxalap and vehicle in the overall mean change from baseline in VAS ocular dryness score in the ITTOD population from Week 2 to Week 12.

H_{11}: The overall mean change from baseline in VAS ocular dryness score is greater with reproxalap than with vehicle in the ITTOD population from Week 2 to Week 12.

H_{02}: There is no difference between reproxalap and vehicle in the overall mean change from baseline in fluorescein staining in the nasal region using the [redacted] scale in the ITTFSN population from Week 2 to Week 12.

H_{12}: The overall mean change from baseline in fluorescein staining in the nasal region using the [redacted] scale is greater with reproxalap than with vehicle in the ITTFSN population from Week 2 to Week 12.

3. Study Design and Procedures

3.1 General Study Design

The clinical trial is an adaptive Phase 3, multi-center, randomized, double-masked, parallel-group, vehicle-controlled design with block enrollment and is comprised of two parts (Part 1 and Part 2). Test article will be self-administered topically to the eye.

Screening

Prior to randomization into either Part 1 or Part 2, subjects will receive 14 consecutive days (± 2) of vehicle, self–administered four times daily (QID) in both eyes, between Visits 1 and 2. Subjects randomized in Part 1 will not be eligible to participate in Part 2.

During the screening period (Visit 1 [Day -14] to Visit 2 [Day 1]), two 90-minute exposures in the [redacted] will be conducted to ascertain eligibility. Subjects who qualify will be randomized to receive study drug in a double-masked fashion for 84 days. At Visits 3 - 9, no [redacted] exposure will occur but signs and symptoms will be assessed. [redacted]

Part 1

Subjects will be randomized 1:1:1:1 into one of four treatment arms (approximately 100 subjects per arm):

A1. Reproxalap administered QID for twelve weeks (Q-Q)

A2. Vehicle administered QID for twelve weeks (Q-Q)

B1. Reproxalap administered QID for four weeks, followed by BID administration for eight weeks (Q-B)

B2. Vehicle administered QID for four weeks, followed by BID administration for eight weeks (Q-B)

The results of A1 will be compared to the results of A2. The results of B1 will be compared to the results of B2.

Subjects, the Sponsor, the Contract Research Organization (CRO), investigators, and site personnel will be masked to treatment assignment. Subjects will not be masked to dosing regimen.

The total number of expected participants in Part 1, including screen failures, is approximately 1,000 subjects. The duration of Part 1 is approximately 14 weeks (98 days), including 12 weeks of treatment.

Part 2

Subjects will be randomized 1:1 to receive treatment with either reproxalap or vehicle for twelve weeks. Dosing regimen and subject number per arm will be determined based on the results of Part 1. The duration of Part 2 is approximately 14 weeks (98 days), including 12 weeks of treatment.

Subjects, the Sponsor, the CRO, investigators, and site personnel will be masked to treatment assignment.

Table 1 shows the scheduled study visits, their planned study day (note that there is no Day 0 and that Day 1 corresponds to the day of randomization), and the acceptable visit window for each study visit:

Table 1. Scheduled Study Visits, Planned Study Days, and Visit Windows

Scheduled Visit	Planned Study Day	Visit Window
Visit 1	Day -14	± 2 Days
Visit 2	Day 1	N/A
Visit 3	Day 8	± 2 Days
Visit 4	Day 15	± 2 Days
Visit 5	Day 29	± 2 Days
Visit 6	Day 43	± 2 Days
Visit 7	Day 57	± 3 Days
Visit 8	Day 71	± 3 Days
Visit 9	Day 85	± 3 Days

3.2 Schedule of Visits and Assessments

The schedule of visits and assessments is provided below for Part 1 of the study.

Table 2. Part 1 Schedule of Visits and Measurements

Procedure	Visit 1 Day -14±2		Visit 2 Day 1		Visit 3 Day 8±2	Visit 4 Day 15±2	Visit 5 Day 29±2	Visit 6 Day 43±2	Visit 7 Day 57±3	Visit 8 Day 71±3	Visit 9 Day 85 ± 3
	Pre [redacted]	Post [redacted]	Pre [redacted]	Post [redacted]	Non [redacted]	Non [redacted]	Non [redacted]	Non [redacted]	Non [redacted]	Non [redacted]	Non [redacted]
Informed Consent / HIPAA	X										
Medical / Medication History and Demographics	X										
Medical / Medication Update			X		X	X	X	X	X	X	X
[redacted] Exposure	X		X								
Vehicle Run-In Dispensation		X									
Vehicle Run-In Instillation		X	X								
Vehicle Run-in Collection			X								
Randomization				X							
Study Drug Dispensation				X	X[2]	X	X	X[2]	X	X[2]	
Study Drug Instillation				X							
Study Drug Collection					X	X	X	X	X	X	X
Review of Qualification Criteria	X	X	X	X							
Adverse Event Query		X	X	X	X	X	X	X	X	X	X
Pregnancy Test	X[1]										X[1]
Multi-Symptom VAS	X		X		X	X	X	X	X	X	X
SANDE Questionnaire	X		X		X	X	X	X	X	X	X
[redacted] Discomfort Scale	X		X	X	X	X	X	X	X	X	X
[redacted] Symptom Questionnaire	X		X	X	X	X	X	X	X	X	X

Procedure	Visit 1 Day -14±2		Visit 2 Day 1		Visit 3 Day 8±2	Visit 4 Day 15±2	Visit 5 Day 29±2	Visit 6 Day 43±2	Visit 7 Day 57±3	Visit 8 Day 71±3	Visit 9 Day 85 ± 3
	Pre [redacted]	Post [redacted]	Pre [redacted]	Post [redacted]	Non [redacted]	Non [redacted]	Non [redacted]	Non [redacted]	Non [redacted]	Non [redacted]	Non [redacted]
[redacted] Itching Scale	X		X		X	X	X	X	X	X	X
OSDI© Questionnaire	X		X		X	X	X	X	X	X	X
Visual Acuity (ETDRS)	X		X		X	X	X	X	X	X	X
Slit-lamp Biomicroscopy	X	X	X	X	X	X	X	X	X	X	X
Tear Osmolarity			X				X				X
TFBUT	X	X	X	X	X	X	X	X	X	X	X
Fluorescein Staining	X	X	X	X	X	X	X	X	X	X	X
Lissamine Green Staining	X	X	X	X	X	X	X	X	X	X	X
Schirmer's Test	X		X				X				X
Intraocular Pressure		X									X
Dilated Fundoscopy		X									X
Exit Subject from Study											X

[1:] To women of child-bearing potential, as defined.
[2]The previous study drug kit is redispensed.

4. Study Treatments

4.1 Method of Assigning Subjects to Treatment Arms

Before the initiation of study run-in at Visit 1 (Day -14), each subject who provides written and informed consent will be assigned to a screening number. All screening numbers will be assigned in strict numerical sequence at a site and no numbers will be skipped or omitted. Each subject who meets all the inclusion and none of the exclusion criteria at Visit 1 and Visit 2 will be assigned a randomization number at the end of Visit 2 (Day 1). The Interactive Web Response System (IWRS) will be used to assign all randomization numbers.

Subjects are to be stratified by Visit 2 (Day 1, Baseline) Pre[redacted] ocular dryness score (on the Ocular Discomfort & 4-Symptom Questionnaire) and Visit 2 (Day 1, Baseline) Pre-[redacted] fluorescein nasal staining score in the study eye using the following strata:

a. Ocular Dryness Score ≥ 3 and Fluorescein Nasal Staining Score ≥ 2
b. Ocular Dryness Score ≥ 3 and Fluorescein Nasal Staining Score < 2
c. Ocular Dryness Score < 3 and Fluorescein Nasal Staining Score ≥ 2
d. Ocular Dryness Score < 3 and Fluorescein Nasal Staining Score < 2

Randomization and kit numbers will be assigned automatically to each subject as they are entered into the IWRS.

The site staff will dispense kit(s) required until the next visit. Both the randomization number and the dispensed study drug kit number(s) will be recorded on the subject's source document and electronic case report form (eCRF). The Sponsor, investigators, and study staff will be masked during the randomization process and throughout the study to study drug treatment assignment but not dosing regimen.

4.2 Masking and Unmasking

All subjects, investigators, and study personnel involved with the conduct of the study will be masked with regard to study drug treatment assignments. When medically necessary, the investigator may need to determine what treatment has been assigned to a subject. When possible (i.e., in non-emergent situations), [redacted] and/or the Sponsor should be notified before unmasking investigational product (IP). [redacted] and/or the study Sponsor must be informed immediately about any unmasking event.

If an investigator identifies a medical need for unmasking the treatment assignment of a subject, he/she should contact [redacted] and/or the medical monitor prior to unmasking the identity of the IP, if possible. [redacted] will ask the site to complete and send them the Unmasking Request Form. [redacted] will notify Aldeyra and jointly will determine if the unmasking request should be granted. They may consult the medical monitor

as needed. The result of the request will be documented on the Unmasking Request Form. If approval is granted to unmask a subject, written permission via the Unmasking Request Form will be provided to the investigator. The investigator will unmask the subject using IWRS. The investigator will complete the Unmasking Memo form and include it in the subject's study file and provide a copy for the TMF. For each unmasked request, the reason, date, signature, and name of the person who unmasked the subject, must be noted in the subject's study file.

Unmasked subjects will be discontinued from the study.

5. Sample Size and Power Considerations

Sample Size Part 1:

Based on the results of a Phase 2b clinical trial of reproxalap in dry eye disease subjects (ADX-102-DED-009), for the primary outcome measurements of FSN and ocular dryness score, approximately 100 subjects per randomized arm will be evaluated in ADX-102-DED-012. Estimates were obtained for the mean difference between reproxalap and vehicle with all post-treatment visits combined, the intra-subject standard deviation (SD), and the estimated correlation between repeated measures. As a conservative measure the following practices were adopted: (1) the estimated correlation was increased by 25%, (2) the estimated SD was increased by 25%, and (3) four repeated measures (rather than the six planned) were assumed. Sample size estimates were made assuming an alpha of 0.05 and power of 90% for the two efficacy populations. [redacted] the sample size estimated for staining was 39 per treatment arm (78 total); the sample size estimated for dryness score was 54 per treatment arm (108 total). The assumptions for parameters used in sample size calculation are listed below.

Parameters assumption for ITTFSN population:

Targeted Power	Sample Size	Repeated Measures	LS Mean Difference	Standard Deviation	Correlation Coefficient	Type I Error
0.9	78	4	-0.2967	0.53	0.5749	0.05

Parameters assumption for ITTOD population:

Targeted Power	Sample Size	Repeated Measures	LS Mean Difference	Standard Deviation	Correlation Coefficient	Type I Error
0.9	108	4	-0.3952	0.73	0.7783	0.05

Sample Size Part 2:

Estimated treatment differences, SDs, and correlation coefficients from Part 1 will be used to determinate sample sizes and estimate power in Part 2, which is expected to require approximately 200-400 subjects per treatment arm.

6. Data Preparation

Electronic Case Report Forms (eCRF) will be developed by [redacted]. Unless otherwise specified by Aldeyra, the eCRFs will follow SDTM standards. Data from source documents will be entered into the eCRF by site personnel.

The clinical study database will be developed and tested in iMedNet™ v1.191.0 or higher.

After data are entered into the clinical study database, electronic edit checks and data review will be performed. All data validation specifications and procedures are detailed in the Data Validation Manual as a separate document. When the database has been declared to be complete and accurate, the database will be locked. Any changes to the database after data have been locked can only be made with the approval of Aldeyra and [redacted] in consultation with [redacted].

All analyses outlined in this document will be carried out after the following have occurred:

- All data management requirements are met according to [redacted] standard operating procedures (SOP), including data entry, performance of edit and validation checks, documentation and resolution of data queries, and database lock with written authorization provided by appropriate [redacted], Aldeyra and [redacted] personnel.
- Protocol deviations have been identified and status defined (major/minor deviations).
- Analysis populations have been determined.
- Randomized treatment codes have been unmasked.

7. Analysis Populations

7.1 Intent-to-Treat

Intent-to-Treat Population: The intent-to-treat (ITT) population includes all randomized subjects. Subjects in the ITT population will be analyzed as randomized. In addition, two populations are further defined for the primary analysis:

- ITTOD Population – The ITTOD population includes subjects in the ITT population with baseline ocular dryness score ≥ 3 (Ocular Discomfort and 4-Symptom Questionnaire).
- ITTFSN Population – The ITTFSN population includes subjects in the ITT population with baseline fluorescein nasal staining score ≥ 2.

7.2 Per Protocol

Per-Protocol Population: The per-protocol (PP) population includes subjects in the ITT population who do not have significant protocol deviations and who complete the study. Protocol deviations will be assessed prior to database lock and unmasking. Subjects in the PP population will be analyzed as treated. In addition, two populations are further defined for the primary analysis:

- PPOD Population – The PPOD population includes subjects in the PP population with baseline ocular dryness score ≥ 3 (Ocular Discomfort and 4-Symptom Questionnaire).
- PPFSN Population – The PPFSN population includes subjects in the PP population with baseline FSN score ≥ 2.

7.3 Safety

Safety Population: The safety population includes all randomized subjects who have received at least one dose of study drug. Subjects in the safety population will be analyzed as treated.

8. General Statistical Considerations

8.1 Unit of Analysis

Safety endpoints will be analyzed for both eyes. For subject-level efficacy endpoints, the unit of analysis will be the subject. For efficacy endpoints assessed on each eye individually, the unit of analysis will be the study eye as defined by the following:

[redacted]

8.2 Missing or Inconclusive Data Handling

The primary analyses using mixed model repeated measures (MMRM) method will be performed on the ITTOD and ITTFSN population with observed data only. Efficacy analyses may also be conducted using the ITT, ITTOD, and ITTFSN populations with multiple imputation under missing at random (MAR) and missing not at random (MNAR) assumptions. Per-protocol population analysis may also be conducted to assess sensitivity.

No other secondary efficacy endpoints, exploratory endpoints, or safety endpoints will be imputed.

8.3 Definition of Baseline

Baseline measures are defined as the last non-missing measure prior to the initiation of randomized study treatment at Visit 2 (Day 1) prior to [redacted] Change from baseline will be calculated as follow-up

visit value minus baseline value. Intraocular pressure and dilated fundoscopy will have baseline measures recorded at Visit 1 (Day -14).

8.4 Data Analysis Conventions

All data analysis will be performed by [redacted] after the study is completed and the database has been locked and released for unmasking. Statistical programming and analyses will be performed using Statistical Analysis Software® (SAS®) Version 9.4 or higher. Output will be provided in rich text format (RTF) for tables and portable document format (PDF) for tables, listings, and figures using landscape orientation. All study data will be listed by subject, treatment, and visit (as applicable) based on all randomized subjects unless otherwise specified.

Summaries for continuous and ordinal variables will include the number of observations (n), arithmetic mean, SD, median, minimum, and maximum. Effect size relative to baseline for efficacy endpoints will also be reported and will be calculated as the mean change from baseline divided by the SD at baseline. Minima and maxima will be reported with the same precision as the raw values; means, medians and effect sizes will be presented to one additional decimal place than reported in the raw values. Standard deviations will be presented to two additional decimal places than reported in the raw values. Summaries for discrete variables will include frequency counts and percentages. All percentages will be rounded to one decimal place (i.e., XX.X%). Differences between active treatment arms and vehicle will be calculated as Active – Vehicle and change from baseline will be calculated as Follow-up Visit – Baseline. Cohen's d effect size will also be reported and will be calculated as the difference in means divided by the pooled SD. The baseline measure will be defined as the last non-missing measure prior to initiation of investigational treatment.

All statistical tests will be two-sided with a significance level of 0.05 ($\alpha = 0.05$) unless otherwise specified. Confidence intervals (CI) for differences between treatment arms will be two-sided at 95% confidence. All p-values will be rounded to four decimal places; p-values less than 0.0001 will be presented as "<0.0001"; p-values greater than 0.9999 will be presented as ">0.9999."

Unless otherwise specified, summaries will be presented by treatment arm and, where appropriate, visit.

8.5 Adjustments for Multiplicity

The hypothesis testing in Part 1 is performed within each dosing regimen separately. A sequential (closed) testing procedure will be used to control the overall type I error rate within each dosing regimen due to multiple comparisons for the coprimary endpoints. The order of treatment comparisons is as follows:

1. Reproxalap administered QID for twelve weeks versus Vehicle administered QID for twelve weeks (Arm A1 vs Arm A2)

2. Reproxalap administered QID for four weeks, followed by BID administration for eight weeks versus Vehicle administered QID for four weeks, followed by BID administration for eight weeks (Arm B1 vs Arm B2)

The second comparison will be considered valid only if the first comparison is statistically significant for either of the co-primary endpoints.

A stronger type I error control is implemented within each dosing regimen by requiring that the hypothesis testing for both coprimary endpoints must be statistically significant to claim success in Part 1. Hence, multiplicity adjustment will not be applied within each dosing regimen.

9. Disposition of Subjects

Subject disposition will be presented in terms of the numbers and percentages of subjects who were randomized, completed the study, and discontinued from the study. Subjects who are not discontinued from the study will be considered study completers. Disposition will be summarized by treatment arm and for all subjects.

The number of randomized subjects in each of the analysis populations (ITTOD, ITTFSN, ITT, PPOD, PPFSN, PP, and Safety) will be displayed by treatment. The ITTOD, ITTFSN, and ITT populations use treatment as randomized; PPOD, PPFSN, PP, and Safety populations use treatment as treated. Percentages are based on the total number of subjects randomized in each treatment arm.

The number and percentage of subjects prematurely discontinued from the study and the reasons for study discontinuation will be summarized by treatment arm for all randomized subjects. The reasons for study discontinuation that will be summarized include: AE(s), Protocol Violation(s), Administrative Reasons, Sponsor Termination of the Study, Subject Choice, and Other. A subject listing will be provided that includes the date of and reason for premature study discontinuation.

The number and percentage of subjects with major protocol deviations will be summarized by treatment arm for all randomized subjects. The protocol deviations that will be summarized include: Informed Consent, Inclusion/Exclusion and Randomization, Test Article/Study Drug Instillation and Assignment at Site, Improper Protocol Procedures at Site, Site's Failure to Report Serious Adverse Event (SAE)/AE, Visit out of Window (missed, early, late), Subject's Non-compliance with Test Article/Study Drug, Subject's Use of Prohibited Concomitant Medication, Subject's Failure to Follow Instructions, and Other. A subject listing will be provided that includes the date of the deviation, the deviation code, the deviation description, and the classification of whether the deviation was judged to be major or minor.

In addition, subject listings will be provided that include inclusion and exclusion criteria violations and exclusions from the ITT, PP, and Safety populations.

10. Demographic and Pretreatment Variables

10.1 Demographic Variables

The demographic variables collected in this study include age, sex, race, ethnicity, and iris color (OD and OS). Demographic variables will be summarized for the ITT and Safety populations, separately.

Age (years) will be summarized, overall and by treatment, using continuous descriptive statistics. Age will also be categorized as follows: <65 years and ≥65 years. Age will be reported in years and calculated using the following formula:

Age (years) = (Informed Consent Date – Date of Birth) / 365.25, truncated as an integer

The number and percentage of subjects will be presented, overall and by treatment arm, for age category, sex, race, ethnicity, and iris color (right eye [OD] and left eye [OS]).

A subject listing that includes all demographic variables will be provided.

10.2 Pretreatment Variables

Baseline disease characteristics will be summarized by treatment arm for ITTOD, ITTFSN, and ITT populations using continuous descriptive statistics for ocular dryness score (0 – 100 VAS); fluorescein nasal region score using [redacted] scale); fluorescein staining score ([redacted] unanesthetized Schirmer's test; VA; and IOP. The scale for each assessment is provided in the variables' respective subsection in Section 13 and Section 15 of this SAP.

11. Medical History and Concomitant Medications

11.1 Medical History

Ocular medical history will be summarized using discrete summary statistics and presented by treatment arm at the subject and event level by System Organ Class (SOC) and Preferred Term (PT) using the Safety population. Non-ocular medical history will be similarly summarized at the subject and event level. If a subject reports the same PT multiple times within the same SOC, that PT will only be reported once within that SOC. As with the PT, if a subject reports multiple conditions within the same SOC, that SOC will only be reported once. MedDRA 22.0 will be used to code medical history terms.

Listings of medical history will be generated separately for ocular and non-ocular data.

11.2 Concomitant Medications

Ocular and non-ocular concomitant medications will be coded using WHO Drug Dictionary (WHODrug Global, B3, March 2019) and summarized to the therapeutic drug class (Anatomical Therapeutic

Chemical [ATC] 4 classification) and preferred name. If the ATC 4 classification is not provided, then the next lowest classification that is provided in the coding dictionary will be used. The preferred name will be defined as the active ingredient; if the active ingredient is not provided or includes more than two ingredients (e.g., multivitamins), then the drug name will be summarized as the preferred name. Any uncoded terms will be summarized under the ATC classification and preferred name of "Uncoded."

Concomitant medications are defined as those medications listed as having been taken (1) prior to initiation of study drug administration and continuing for any period of time following the first administration of study drug or (2) at any time following the first administration of study drug. Prior medications are reported medications that have been taken prior to initiation of study drug administration but not during the study.

Concomitant medications will be summarized using the ITT population. Medications will be tabulated for each treatment arm using frequencies and percentages. Subjects may have more than 1 medication per ATC text. At each level of subject summarization, a subject will be counted once if he/she reports 1 or more medications. Percentages will be based on the number of subjects in each treatment arm. Listings of concomitant medications will be generated separately for ocular and non-ocular data.

12. Dosing Compliance and Treatment Exposure

12.1 Dosing Compliance

Dosing compliance (% compliance) will be assessed by calculating the number of doses taken and comparing that to the number of expected doses as follows:

$$\text{Compliance (\%)} = \frac{\text{Number of Doses Taken}}{\text{Number of Expected Doses}} \times 100\%$$

The number of actual doses taken will be recorded from the eCRF. The number of expected doses that will be used for calculating compliance will be calculated as follows:

- For subjects in dosing arms Q-Q: 4 x {[Date of Study Completion/Discontinuation – Date of Visit 2 (Day 1)] + 1}, regardless of study completion status
- For subjects in dosing arms Q-B that discontinue prior to Visit 5 (Day 29): 4 x {[Date of Discontinuation – Date of Visit 2 (Day 1)] + 1}
- For subjects in dosing arms Q-B that complete the study or discontinue on or after Visit 5 (Day 29): 4 x {[Date of Visit 5 (Day 29) – Date of Visit 2 (Day 1)] + 1} + 2 x [Date of Study Completion/Discontinuation – Date of Visit 5 (Day 29)]

For subjects who do not receive a dose of study drug per in office instillation, the number of expected doses will be 0 and compliance will be defined as 0.

A categorical dosing compliance variable will also be derived as non-compliant (<80%), compliant (≥80% and ≤125%), and over compliant (>125%).

Dosing compliance (%) will be summarized with continuous descriptive statistics for each treatment arm, using the Safety population. The compliance category defined above will be summarized with discrete summary statistics.

A subject listing of dosing compliance and study drug accountability will also be produced.

A subject listing of run-in, run-in instillation, and run-in replacement will be provided as well as study drug assignment, instillation, and replacement.

12.2 Treatment Exposure

Extent of treatment exposure for completed or discontinued subjects will be calculated in days using the following:

Extent of Exposure (days) = [Date of Study Completion/Discontinuation – Date of Visit 2 (Day 1)] + 1

Extent of treatment exposure for subjects who were lost to follow-up will be calculated in days using the following:

Extent of Exposure (days) = [Date of Last Recorded Visit – Date of Visit 2 (Day 1)] + 1

Extent of treatment exposure (days) for each subject exposed to study drug will be summarized with continuous descriptive statistics for each treatment arm, using the Safety population.

Subject listings of study drug exposure and dosing compliance will be produced.

13. Efficacy Analyses

13.1 Primary Analysis

The co-primary endpoints for symptom and sign are as follows:

- The co-primary symptom endpoint is the overall mean change from baseline in subject-reported ocular dryness score (0 - 100 VAS) from Week 2 to Week 12. The population used for analysis of ocular dryness score will be subjects who have a baseline ocular dryness score of ≥ 3 ([redacted]).
- The co-primary sign endpoint is the overall mean change from baseline in fluorescein nasal region score (0-4 scale) from Week 2 to Week 12. The population used for analysis of fluorescein

nasal region score will be subjects who have a baseline (Day 1) fluorescein nasal region score of ≥ 2 ([redacted]).

Change from baseline for both endpoints will be calculated as Visit – Baseline, where a positive difference indicates a worsening of dry eye signs or symptoms and a negative difference indicates an improvement of dry eye signs or symptoms. Reproxalap and vehicle Q-Q dosing arms will be compared in separate tables from reproxalap and vehicle Q-B dosing arms. Treatment comparisons between active treatment arm and vehicle will be calculated as Active – Vehicle. All primary endpoints will be described in subject listings.

13.1.1 OCULAR DRYNESS SCORE

Subjects will be asked to subjectively rate ocular dryness (both eyes [OU]) by placing a vertical mark on the horizontal line to indicate the level of discomfort. The length of the assessment line is 100 mm; a measure of 0 mm corresponds to “No Discomfort” and 100 mm corresponds to “Maximal Discomfort.”

Ocular dryness score (0-100 VAS) will be summarized for Visit 3 (Day 8) to Visit 9 (Day 85) by treatment arms using continuous descriptive statistics. Overall mean change from baseline from week 2 to week 12 for each treatment arm in ocular dryness score will be compared between active treatment and vehicle. The primary analysis will use an MMRM adjusted for baseline ocular dryness score (0-100 VAS), treatment arm, visit, and the interaction of treatment arm and visit as fixed effects and subject as the random effect. Least squares (LS) means for each treatment arm and the LS mean difference between treatment arms for the overall mean change from baseline from week 2 to week 12 and change from baseline for each visit will be presented from the model together with standard errors (SE), two-sided p-values, and two-sided 95% CIs. Primary analysis will be performed on the ITTOD population with observed data only.

[redacted]

All post-baseline data from the two treatment groups will be included in the models. Comparisons between two treatment arms (e.g., between reproxalap Q-Q and vehicle Q-Q) for overall mean change from baseline from week 2 to week 12 and change from baseline at a given visit will be obtained by comparing the corresponding LS mean estimates.

Change from baseline will also be analyzed using analysis of covariance (ANCOVA) models adjusted for baseline ocular dryness score and treatment arm. Two-sample t-tests and Wilcoxon rank sum tests will be used as sensitivity analyses. Cohen's d effect size will be reported for differences in change from baseline between treatment arms.

Markov chain Monte Carlo (MCMC) imputation with the ITTOD population will used for the sensitivity analysis under MAR assumption. A single imputation will first impute non-monotone missing data using MCMC to obtain a dataset with a monotone missing pattern. Then, MCMC will be used to perform twenty imputations from the dataset with monotone missingness. The primary endpoint will be re-analyzed to confirm the results from the primary analysis.

As a sensitivity analysis, the primary efficacy endpoint of ocular dryness score will also be analyzed with missing data imputated using a control-based pattern mixture model under the assumption of a missing not at random mechanism. A single imputation will first impute non-monotone missing data using MCMC to obtain a dataset with a monotone missing pattern. Then, a control-based pattern mixture model will used to perform twenty imputations from the dataset with monotone missingness.

The MMRM model outlined in the primary analysis will be used for analysis of the MCMC and control-based pattern mixture model imputed datasets. ANCOVA models adjusted for baseline ocular dryness score (0-100 VAS) and treatment arm and two-sample t-tests will also serve as sensitivity analyses.

Primary efficacy analyses will be performed with the PPOD population for additional sensitivity analysis using an MMRM as described for the ITTOD with observed data only analyses.

The LS Means for ocular dryness score overall mean change from baseline from week 2 to week 12 will be displayed graphically in a bar chart with standard error bars for the active treatment and vehicle arms based on the MMRM analysis of the ITTOD population with observed data only.

13.1.2 FLUORESCEIN NASAL STAINING

[redacted]

Fluorescein nasal staining will be summarized for Visit 3 (Day 8) to Visit 9 (Day 85) by treatment arms for the study eye using continuous descriptive statistics. Overall mean change from baseline from week 2 to week 12 for each treatment arm in fluorescein nasal staining will be compared between active

treatment and vehicle. Primary analysis will use a MMRM as described in section 13.1.1 adjusted for baseline FSN score, treatment arm, visit, and the interaction of treatment arm and visit as fixed effects and subject as the random effect. The LS means for each treatment arm and the LS mean difference between treatment arms for the overall mean change from baseline from week 2 to week 12 and change from baseline at each visit will be presented from the model together with SEs, two-sided p-values, and two-sided 95% CIs. Primary analysis will be performed on the ITTFSN population with observed data only.

Change from baseline will also be analyzed using ANCOVA models adjusted for baseline FSN score and treatment arm. Two-sample t-tests and Wilcoxon rank sum tests will be used as sensitivity analyses. Cohen's d effect size will be reported for differences in change from baseline between treatment arms.

Markov chain Monte Carlo Imputation will be used for sensitivity analysis for the ITTFSN population under MAR assumption. A single imputation will first impute non-monotone missing data using MCMC to obtain a dataset with a monotone missing pattern. Then, MCMC will be used to perform twenty imputations from the dataset with monotone missingness. The primary endpoint will be re-analyzed to confirm the results from the primary analysis.

As a sensitivity analysis, the primary efficacy endpoint of fluorescein nasal staining will also be analyzed with missing data imputated using a control-based pattern mixture model under the assumption of a missing not at random mechanism. A single imputation will first impute non-monotone missing data using MCMC to obtain a dataset with a monotone missing pattern. Then, a control-based pattern mixture model will used to perform twenty imputations from the dataset with monotone missingness.

ANCOVA models adjusted for baseline FSN score and treatment arm and two sample t-tests will also be conducted as sensitivity analyses to assess robustness of the results of the sensitivity analyses.

Further, primary efficacy analyses will be performed with the PPFSN population with observed data only for sensitivity analysis using MMRM.

The LS Means for FSN overall mean change from baseline from week 2 to week 12 will be displayed graphically in a bar chart with standard error bars for the active treatment and vehicle arms based on the MMRM analysis of the ITTFSN population with observed data only.

13.2 Key Secondary Analyses

Key secondary efficacy variables will be summarized using continuous descriptive statistics (n, mean, SD, median, minimum, maximum) by visit and treatment arm. Two-sample t-tests and Wilcoxon rank sum tests will be used to analyze the efficacy variables for visit-based data between treatment dosing schedule Q-Q arms A1 and A2 and, separately, dosing schedule Q-B arms B1 and B2. Cohen's d effect size will be reported for the differences in treatment arms at each visit.

Change from baseline in key secondary efficacy variables will be summarized by visit and treatment arm using continuous descriptive statistics and change from baseline effect size. When applicable, MMRM will be used to analyze overall mean change from baseline from week 2 to week 12 and change from baseline by visit as described in section 13.1.1. Further sensitivity analysis will be conducted using ANCOVA models with terms for baseline value and treatment arm. Two-sample t-tests and Wilcoxon rank sum tests will also be used for sensitivity analysis. Cohen's d effect size will be reported for differences in change from baseline between treatment arms. Paired t-tests will be used to assess change from baseline within each treatment arm.

All key secondary efficacy analyses will be performed on the ITT population with observed data only and PP population with observed data only. No imputation will be performed for secondary analyses. All key secondary endpoints will be described under subject listings.

[redacted]

13.2.1 [redacted] DRYNESS

[redacted]

Visit-based data will be summarized using continuous descriptive statistics for dryness by treatment arm. Change from baseline will also be summarized with continuous descriptive statistics.

Two-sample t-tests and Wilcoxon rank sum tests will be used to analyze differences between treatment arms at each visit between active treatment and vehicle. The differences in means, two-sided 95% CIs for the difference in means, and two-sided p-values will be reported. Cohen's d effect size will be reported for the differences in treatment arms at each visit.

Overall mean change from baseline will be analyzed using MMRM in manner described in section 13.1.1 with adjustment for baseline dryness score, visit, treatment arm, and interaction between treatment arm and visit. Change from baseline by visit will also be analyzed. The LS means, LS mean differences, SEs, two-sided 95% CIs for LS means, two-sided 95% CIs for the difference in means, and two-sided p-values will be reported from the MMRM. Sensitivity analysis will include ANCOVA adjusted for baseline dryness score and treatment arm, two-sample t-tests, and Wilcoxon rank sum test. Cohen's d effect size will be reported for differences in change from baseline between treatment arms. The LS means, LS mean differences, SEs, two-sided 95% CIs for the difference in means, and two-sided p-

values will be reported from the sensitivity analyses where applicable. Paired t-tests will be used to analyze change from baseline within each treatment arm.

Analyses will use ITT population with observed data only and PP population with observed data only.

13.3 Secondary Analyses

Secondary efficacy variables will be summarized using continuous descriptive statistics (n, mean, SD, median, minimum, maximum) by visit and treatment arm. Two-sample t-tests and Wilcoxon rank sum tests will be used to analyze the efficacy variables for visit-based data between treatment dosing schedule Q-Q arms A1 and A2 and, separately, dosing schedule Q-B arms B1 and B2. Cohen's d effect size will be reported for the differences in treatment arms at each visit.

Change from baseline in secondary efficacy variables will be summarized by visit and treatment arm using continuous descriptive statistics and change from baseline effect size. When applicable, MMRM will be used to analyze overall mean change from baseline from week 2 to week 12 and change from baseline by visit as described in section 13.1.1. Further sensitivity analysis will be conducted using ANCOVA models with terms for baseline value and treatment arm. Two-sample t-tests and Wilcoxon rank sum tests will also be used for sensitivity analysis. Cohen's d effect size will be reported for differences in change from baseline between treatment arms. Paired t-tests will be used to assess change from baseline within each treatment arm.

All secondary efficacy analyses will be performed on the ITT population with observed data only and PP population with observed data only. No imputation will be performed for secondary analyses. All secondary endpoints will be described under subject listings.

The following secondary efficacy variables will be tested:

- Fluorescein staining [redacted]
- Unanesthetized Schirmer's test
- Ocular Dryness through Visit 5 (Day 29) (0 - 100 VAS)

13.3.1 FLUORESCEIN STAINING [redacted]

Fluorescein staining will be conducted at all visits and Pre- and Post-[redacted] for Visits 1 (Day -14) and Visits 2 (Day 1). Grading will be conducted [redacted] Comparisons will be made between reproxalap arms and vehicle arms within the same dosing schedule.

Visit-based data will be summarized using continuous descriptive statistics for each region and sum score. Change from baseline will also be summarized using continuous descriptive statistics.

Two-sample t-tests and Wilcoxon rank sum tests will be used to analyze differences between treatment arms at each visit between active treatment and vehicle. The differences in means, two-sided 95% CIs for the difference in means, and two-sided p-values will be reported. Cohen's d effect size will be reported for the differences in treatment arms at each visit.

Overall mean change from baseline from week 2 to week 12 will be analyzed using MMRM in manner described in section 13.1.1 with adjustment for baseline fluorescein staining score, visit, treatment arm, and interaction between treatment arm and visit. Change from baseline by visit will also be analyzed. The LS means, LS mean differences, SEs, two-sided 95% CIs for LS means, two-sided 95% CIs for the difference in means, and two-sided p-values will be reported from the MMRM. Sensitivity analysis will include ANCOVA adjusted for baseline fluorescein staining score and treatment arm, two-sample t-tests, and Wilcoxon rank sum test. Cohen's d effect size will be reported for differences in change from baseline between treatment arms. The LS means, LS mean differences, SEs, two-sided 95% CIs for the difference in means, and two-sided p-values will be reported from the sensitivity analyses where applicable. Paired t-tests will be used to analyze change from baseline within each treatment arm.

Analyses will use ITT population with observed data only and PP population with observed data only.

13.3.2 UNANESTHETIZED SCHIRMER'S TEST

Unanesthetized Schirmer's test will be conducted at Visits 1 (Day -14) Pre-[redacted] 2 (Day 1) Pre-[redacted] 5 (Day 29), and 9 (Day 85). [redacted] Lower values indicate less tears produced in the eye. Comparisons will be made between reproxalap arms and vehicle arms within the same dosing schedule. Analyses will only be produced for study eye.

Visit-based data will be summarized using continuous descriptive statistics by treatment arm. Change from baseline will also be summarized with continuous descriptive statistics.

Two-sample t-tests and Wilcoxon rank sum tests will be used to analyze differences between treatment arms are each visit between active treatment and vehicle. The differences in means, two-sided 95% CIs for the difference in means, and two-sided p-values will be reported. Cohen's d effect size will be reported for the differences in treatment arms at each visit.

Overall mean change from baseline from week 4 to week 12 will be analyzed using MMRM in manner described in section 13.1.1 with adjustment for baseline unanesthetized Schirmer's test reading, visit, treatment arm, and interaction between treatment arm and visit. Change from baseline by visit will also

be analyzed. The LS means, LS mean differences, SEs, two-sided 95% CIs for LS means, two-sided 95% CIs for the difference in means, and two-sided p-values will be reported from the MMRM. Sensitivity analysis will include ANCOVA adjusted for baseline unanesthetized Schirmer's test reading and treatment arm, two-sample t-tests, and Wilcoxon rank sum test. Cohen's d effect size will be reported for differences in change from baseline between treatment arms. The LS means, LS mean differences, SEs, two-sided 95% CIs for the difference in means, and two-sided p-values will be reported from the sensitivity analyses where applicable. Paired t-tests will be used to analyze change from baseline within each treatment arm.

Analyses will use ITT population with observed data only and PP population with observed data only.

13.3.3 OCULAR DRYNESS THROUGH VISIT 5 (DAY 29)

Ocular dryness (OU) will recorded at Visit 5 (Day 29) using the VAS. Subject will be asked to subjectively rate ocular dryness by placing a vertical mark on the horizontal line to indicate the level of discomfort. The length of the assessment line is 100 mm; a measure of 0 mm corresponds to "No Discomfort" and 100 mm corresponds to "Maximal Discomfort." Comparisons will be made between reproxalap arms and vehicle arms within the same dosing schedule.

Visit-based data will be summarized using continuous descriptive statistics for each region and sum score by treatment arms by visit. Change from baseline will also be summarized with continuous descriptive statistics.

Two-sample t-tests and Wilcoxon rank sum tests will be used to analyze differences between treatment arms at each visit between active treatment arms and vehicle arms. The differences in means, two-sided 95% CIs for the difference in means, and two-sided p-values will be reported. Cohen's d effect size will be reported for the differences in treatment arms at each visit.

Overall mean change from baseline from week 1 to week 4 will be analyzed using MMRM adjusted for baseline ocular dryness score, treatment arm, visit, the interaction of treatment arm and visit as fixed effects and subject as the random effect. Change from baseline by visit will also be analyzed. The LS means, LS mean differences, SEs, two-sided 95% CIs for LS means, two-sided 95% CIs for the difference in means, and two-sided p-values will be reported from the MMRM. Sensitivity analysis will include ANCOVA adjusted for baseline ocular dryness score and treatment arm, two-sample t-tests, and Wilcoxon rank sum test. Cohen's d effect size will be reported for differences in change from baseline between treatment arms. The LS means, LS mean differences, SEs, two-sided 95% CIs for LS means, two-sided 95% CIs for the difference in means, and two-sided p-values will be reported from the sensitivity analyses where applicable. Paired t-tests will be used to analyze change from baseline within each treatment arm.

Analyses will use ITT population with observed data only and PP population with observed data only.

14. Exploratory Analyses

Exploratory efficacy variables will be summarized using continuous descriptive statistics (n, mean, SD, median, minimum, maximum, and Cohen's d effect size relative to baseline) by visit and treatment arm. Two-sample t-tests and Wilcoxon rank sum tests will be used to analyze the efficacy variables for visit-based data between active treatment arms and vehicle. Cohen's d effect size will be reported for the differences in treatment arms at each visit.

Change from baseline in exploratory efficacy variables will be summarized by visit and treatment arm using continuous descriptive statistics. When applicable, MMRM will be used to analyze overall mean change from baseline and change from baseline by visit as described in section 13.1.1. Further sensitivity analysis will be conducted using ANCOVA models with terms for baseline value and treatment arm. Two-sample t-tests and Wilcoxon rank sum tests will also be used for sensitivity analysis. Cohen's d effect size will be reported for differences in change from baseline between treatment arms. Paired t-tests will be used to assess change from baseline within each treatment arm.

All exploratory efficacy analyses will be performed on the ITT population with observed data only. No imputation will be performed for exploratory analyses. All exploratory endpoints will be described under subject listings.

The following exploratory efficacy variables will be tested for Part 1:

- Burning/Stinging, Itching, Foreign Body Sensation, Eye Discomfort, Photophobia, and Pain Score (0 – 100 VAS)
- [redacted] Discomfort [redacted]
- [redacted] Symptom Questionnaire
- OSDI©
- SANDE Questionnaire
- [redacted] Itching [redacted]
- Lissamine green staining [redacted]
- TFBUT
- Tear osmolarity

14.1.1 Visual Analog Scale

Ocular symptoms of the VAS will be recorded at all visits. Subject will be asked to subjectively rate ocular symptom (OU) by placing a vertical mark on the horizontal line to indicate the level of discomfort. The length of the assessment line is 100 mm; a measure of 0 mm to "No Discomfort" and 100 mm

corresponds to “Maximal Discomfort.” The ocular symptoms used for exploratory analysis are Burning/Stinging, Itching, Foreign Body Sensation, Eye Discomfort, Photophobia, and Pain Score. Comparisons will be made between reproxalap arms and vehicle arms within the same dosing schedule.

Visit-based data will be summarized using continuous descriptive statistics for each symptom by treatment arm. Change from baseline will also be summarized with continuous descriptive statistics.

Two-sample t-tests and Wilcoxon rank sum tests will be used to analyze differences between treatment arms at each visit between active treatment and vehicle. The differences in means, two-sided 95% CIs for the difference in means, and two-sided p-values will be reported. Cohen’s d effect size will be reported for the differences in treatment arms at each visit.

Overall mean change from baseline from week 2 to week 12 will be analyzed using MMRM in manner described in section 13.1.1 with adjustment for baseline ocular symptom score, visit, treatment arm, and interaction between treatment arm and visit. Change from baseline by visit will also be analyzed. The LS means, LS mean differences, SEs, two-sided 95% CIs for LS means, two-sided 95% CIs for the difference in means, and two-sided p-values will be reported from the MMRM. Sensitivity analysis will include ANCOVA adjusted for baseline ocular symptom score and treatment arm, two-sample t-tests, and Wilcoxon rank sum test. The LS means, LS mean differences, SEs, two-sided 95% CIs for the difference in means, and two-sided p-values will be reported from the sensitivity analyses where applicable. Cohen’s d effect size will be reported for differences in change from baseline between treatment arms. Paired t-tests will be used to analyze change from baseline within each treatment arm.

Analyses will be performed on the ITT population with observed data only.

14.1.2 OCULAR DISCOMFORT SCALE [redacted]

Ocular discomfort scores will be subjectively graded by the subjects using the [redacted] Discomfort Scale at all post-baseline visits and Visits 1 (Day -14) and 2 (Day 1) at Pre-[redacted]. [redacted] Comparisons will be made between reproxalap arms and vehicle arms within the same dosing schedule.

Visit-based data will be summarized using continuous descriptive statistics for each symptom by treatment arm. Change from baseline will also be summarized with continuous descriptive statistics.

Two-sample t-tests and Wilcoxon rank sum tests will be used to analyze differences between treatment arms at each visit between active treatment and vehicle. The differences in means, two-sided 95% CIs for the difference in means, and two-sided p-values will be reported. Cohen’s d effect size will be reported for the differences in treatment arms at each visit.

Overall mean change from baseline from week 2 to week 12 will be analyzed using MMRM in manner described in section 13.1.1 with adjustment for baseline discomfort score, visit, treatment arm, and interaction between treatment arm and visit. Change from baseline by visit will also be analyzed. The LS means, LS mean differences, SEs, two-sided 95% CIs for LS means, two-sided 95% CIs for the difference in means, and two-sided p-values will be reported from the MMRM. Sensitivity analysis will include ANCOVA adjusted for baseline discomfort score and treatment arm, two-sample t-tests, and Wilcoxon rank sum test. Cohen's d effect size will be reported for differences in change from baseline between treatment arms. The LS means, LS mean differences, SEs, two-sided 95% CIs for the difference in means, and two-sided p-values will be reported from the sensitivity analyses where applicable. Paired t-tests will be used to analyze change from baseline within each treatment arm.

Analyses will be performed on the ITT population with observed data only.

14.1.3 [redacted] SYMPTOM QUESTIONNAIRE [redacted]

[redacted] Questionnaires will be recorded at all visits including Visit 1 (Day -14) at Pre-[redacted]® and Visit 2 (Day 1) at Pre-[redacted] and Post-[redacted] Comparisons will be made between reproxalap arms and vehicle arms within the same dosing schedule.

Visit-based data will be summarized using continuous descriptive statistics for each symptom by treatment arm. Change from baseline will also be summarized with continuous descriptive statistics.

Two-sample t-tests and Wilcoxon rank sum tests will be used to analyze differences between treatment arms at each visit between active treatment and vehicle. The differences in means, two-sided 95% CIs for the difference in means, and two-sided p-values will be reported. Cohen's d effect size will be reported for the differences in treatment arms at each visit.

Overall mean change from baseline will be analyzed using MMRM in manner described in section 13.1.1 with adjustment for baseline symptom score, visit, treatment arm, and interaction between treatment arm and visit. Change from baseline by visit will also be analyzed. The LS means, LS mean differences, SEs, two-sided 95% CIs for LS means, two-sided 95% CIs for the difference in means, and two-sided p-values will be reported from the MMRM. Sensitivity analysis will include ANCOVA adjusted for baseline symptom score and treatment arm, two-sample t-tests, and Wilcoxon rank sum test. Cohen's d effect size will be reported for differences in change from baseline between treatment arms. The LS means, LS mean differences, SEs, two-sided 95% CIs for the difference in means, and two-sided p-values will be reported from the sensitivity analyses where applicable. Paired t-tests will be used to analyze change from baseline within each treatment arm.

Analyses will be performed on the ITT population with observed data only.

14.1.4 OCULAR SURFACE DISEASE INDEX©

The OSDI© is assessed on a scale of 0 to 4, where 0 = None of the Time, 1 = Some of the Time, 2 = Half of the Time, 3 = Most of the Time, and 4 = All of the Time. The OSDI© asks the following 12 questions at the subject level:

Have you experienced any of the following during the last week?

1) Eyes that are sensitive to light?
2) Eyes that feel gritty?
3) Painful or sore eyes?
4) Blurred vision?
5) Poor vision?

Have problems with your eyes limited you in performing any of the following during the last week?

6) Reading?
7) Driving at night?
8) Working with a computer or bank machine (ATM)?
9) Watching TV?

Have your eyes felt uncomfortable in any of the following situations during the last week?

10) Windy conditions?
11) Places or areas with low humidity (very dry)?
12) Areas that are air conditioned?

The total OSDI© score is calculated by the following:

$$\text{OSDI©} = \frac{(\text{Sum of Scores}) \times 25}{\#\text{ of Questions Answered}}$$

Note that the number of questions answered in the denominator should exclude those questions with a response of "N/A."

Ocular Surface Disease Index© will be assessed pre-[redacted] at Visit 1 (Day -14), pre-[redacted] at Visit 2 (Day 1), and all post-baseline visits.

Comparisons will be made between reproxalap arms and vehicle arms within the same dosing schedule.

Visit-based data will be summarized using continuous descriptive statistics for each question score and total OSDI© score by treatment arm. Change from baseline will also be summarized with continuous descriptive statistics.

Two-sample t-tests and Wilcoxon rank sum tests will be used to analyze differences between treatment arms at each visit between active treatment and vehicle. The differences in means, two-sided 95% CIs for the difference in means, and two-sided p-values will be reported. Cohen's d effect size will be reported for the differences in treatment arms at each visit.

Overall mean change from baseline from week 2 to week 12 will be analyzed using MMRM in manner described in section 13.1.1 with adjustment for baseline OSDI question score, visit, treatment arm, and interaction between treatment arm and visit. Change from baseline by visit will also be analyzed. The LS means, LS mean differences, SEs, two-sided 95% CIs for LS means, two-sided 95% CIs for the difference in means, and two-sided p-values will be reported from the MMRM. Sensitivity analysis will include ANCOVA adjusted for baseline OSDI question score and treatment arm, two-sample t-tests, and Wilcoxon rank sum test. Cohen's d effect size will be reported for differences in change from baseline between treatment arms. The LS means, LS mean differences, SEs, two-sided 95% CIs for the difference in means, and two-sided p-values will be reported from the sensitivity analyses where applicable. Paired t-tests will be used to analyze change from baseline within each treatment arm.

Analyses will be performed on the ITT population with observed data only.

14.1.5 SYMPTOM ASSESSMENT IN DRY EYE

Subjects will be asked to complete the SANDE Questionnaire at all post-baseline visits including Pre-[redacted] Visit 1 (Day –14) and Pre-[redacted] Visit 2 (Day 1). First, the subject will be asked to place a vertical line to indicate how often, on average, their eyes feel dry and/or irritated. Second, the subject will be asked to place a vertical line to indicate how severe, on average, they feel their symptoms of dryness and/or irritation are. The length of the assessment line is 100 mm. For frequency, 100 mm corresponds to "All the Time" and 0 mm corresponds to "Rarely." For severity, 100 mm corresponds to "Very Severe" and 0 mm corresponds to "Very Mild." The total SANDE questionnaire score is calculated by multiplying the frequency of symptoms score with the severity of symptoms score and obtaining the square root.

Visit-based data will be summarized using continuous descriptive statistics for frequency, severity, and total SANDE score by treatment arm. Change from baseline will also be summarized with continuous descriptive statistics. Comparisons will be made between reproxalap arms and vehicle arms within the same dosing schedule.

Two-sample t-tests and Wilcoxon rank sum tests will be used to analyze differences between treatment arms at each visit between active treatment and vehicle. The differences in means, two-sided 95% CIs

for the difference in means, and two-sided p-values will be reported. Cohen's d effect size will be reported for the differences in treatment arms at each visit.

Overall mean change from baseline from week 2 to week 12 will be analyzed using MMRM in manner described in section 13.1.1 with adjustment for baseline SANDE score, visit, treatment arm, and interaction between treatment arm and visit. Change from baseline by visit will also be analyzed. The LS means, LS mean differences, SEs, two-sided 95% CIs for LS means, two-sided 95% CIs for the difference in means, and two-sided p-values will be reported from the MMRM. Sensitivity analysis will include ANCOVA adjusted for baseline SANDE score and treatment arm, two-sample t-tests, and Wilcoxon rank sum test. Cohen's d effect size will be reported for differences in change from baseline between treatment arms. The LS means, LS mean differences, SEs, two-sided 95% CIs for the difference in means, and two-sided p-values will be reported from the sensitivity analyses where applicable. Paired t-tests will be used to analyze change from baseline within each treatment arm.

Analyses will be performed on the ITT population with observed data only.

14.1.6 [redacted] ITCHING [redacted]

Subjects will rate the severity of their ocular itching symptom in both eyes, according to the following scale from 0 to 4 where 0 = None and 4 = An incapacitating itch with an irresistible urge to rub. Half (0.5) grade increments may be used. Questionnaires will be recorded at Visit 1 (Day -14) at Pre-[redacted], Visit 2 (Day 1) at Pre-[redacted], and at all post-baseline visits. Comparisons will be made between reproxalap arms and vehicle arms within the same dosing schedule. Analyses will be provided for study eye only.

Visit-based data will be summarized using continuous descriptive statistics by treatment arm. Change from baseline will also be summarized with continuous descriptive statistics.

Two-sample t-tests and Wilcoxon rank sum tests will be used to analyze differences between treatment arms at each visit between active treatment and vehicle. The differences in means, two-sided 95% CIs for the difference in means, and two-sided p-values will be reported. Cohen's d effect size will be reported for the differences in treatment arms at each visit.

Overall mean change from baseline from week 2 to week 12 will be analyzed using MMRM in manner described in section 13.1.1 with adjustment for baseline CAC score, visit, treatment arm, and interaction between treatment arm and visit. Change from baseline by visit will also be analayzed. The LS means, LS mean differences, SEs, two-sided 95% CIs for LS means, two-sided 95% CIs for the difference in means, and two-sided p-values will be reported from the MMRM. Sensitivity analysis will include ANCOVA adjusted for baseline CAC score and treatment arm, two-sample t-tests, and Wilcoxon rank sum test. Cohen's d effect size will be reported for differences in change from baseline between treatment arms. The LS means, LS mean differences, SEs, two-sided 95% CIs for the difference in

means, and two-sided p-values will be reported from the sensitivity analyses where applicable. Paired t-tests will be used to analyze change from baseline within each treatment arm.

Analyses will be performed on the ITT population with observed data only.

14.1.7 LISSAMINE GREEN STAINING [redacted]

Subjects will undergo lissamine green staining at all visits including Pre-[redacted] and Post-[redacted] at Visits 1 (Day -14) and 2 (Day 1). [redacted] Analyses will be provided for study eye only. Comparisons will be made between reproxalap arms and vehicle arms within the same dosing schedule.

Visit-based data will be summarized using continuous descriptive statistics by treatment arm. Change from baseline will also be summarized with continuous descriptive statistics.

Two-sample t-tests and Wilcoxon rank sum tests will be used to analyze differences between treatment arms at each visit between active treatment and vehicle. The differences in means, two-sided 95% CIs for the difference in means, and two-sided p-values will be reported. Cohen's d effect size will be reported for the differences in treatment arms at each visit.

Overall mean change from baseline from week 2 to week 12 will be analyzed using MMRM in manner described in section 13.1.1 with adjustment for baseline lissamine green staining score, visit, treatment arm, and interaction between treatment arm and visit. Change from baseline by visit will also be analyzed. The LS means, LS mean differences, SEs, two-sided 95% CIs for LS means, two-sided 95% CIs for the difference in means, and two-sided p-values will be reported from the MMRM. Sensitivity analysis will include ANCOVA adjusted for baseline lissamine green staining score and treatment arm, two-sample t-tests, and Wilcoxon rank sum test. Cohen's d effect size will be reported for differences in change from baseline between treatment arms. The LS means, LS mean differences, SEs, two-sided 95% CIs for the difference in means, and two-sided p-values will be reported from the sensitivity analyses where applicable. Paired t-tests will be used to analyze change from baseline within each treatment arm.

Analyses will be performed on the ITT population with observed data only.

14.1.8 TEAR FILM BREAK-UP TIME

The TFBUT will be recorded at all visits including Pre-[redacted] and Post-[redacted] at Visits 1 (Day -14) and 2 (Day 1). [redacted]

[redacted]

[redacted] Comparisons will be made between reproxalap arms and vehicle arms within the same dosing schedule.

Visit-based data will be summarized using continuous descriptive statistics by treatment arm. Change from baseline will also be summarized with continuous descriptive statistics at each post-baseline visit.

Two-sample t-tests and Wilcoxon rank sum tests will be used to analyze differences between treatment arms at each visit between active treatment and vehicle. The differences in means, two-sided 95% CIs for the difference in means, and two-sided p-values will be reported. Cohen's d effect size will be reported for the differences in treatment arms at each visit.

Overall mean change from baseline from week 2 to week 12 will be analyzed using MMRM in manner described in section 13.1.1 with adjustment for baseline TFBUT score, visit, treatment arm, and interaction between treatment arm and visit. Change from baseline by visit will also be analyzed. The LS means, LS mean differences, SEs, two-sided 95% CIs for LS means, two-sided 95% CIs for the difference in means, and two-sided p-values will be reported from the MMRM. Sensitivity analysis will include ANCOVA adjusted for baseline TFBUT score and treatment arm, two-sample t-tests, and Wilcoxon rank sum test. Cohen's d effect size will be reported for differences in change from baseline between treatment arms. The LS means, LS mean differences, SEs, two-sided 95% CIs for the difference in means, and two-sided p-values will be reported from the sensitivity analyses where applicable. Paired t-tests will be used to analyze change from baseline within each treatment arm.

Analyses will be performed on the ITT population with observed data only.

14.1.9 TEAR OSMOLARITY

Tear osmolarity will be measured at Visits 2 (Day 1) Pre-[redacted], 5 (Day 29), and 9 (Day 85). Tear osmolarity will be taken once from the temporal canthus of each eye and the measurement will be recorded. A second reading may be taken if the first reading is out of range. A maximum of 2 attempts will be made per eye. Tear osmolarity will be measured in milliosmoles per liter (mOsm/L). Analyses will be provided for study eye only. Comparisons will be made between reproxalap arms and vehicle arms within the same dosing schedule.

Visit-based data will be summarized using continuous descriptive statistics by treatment arm. Change from baseline will also be summarized with continuous descriptive statistics at each post-baseline visit.

Two-sample t-tests and Wilcoxon rank sum tests will be used to analyze differences between treatment arms at each visit between active treatment and vehicle. The differences in means, two-sided 95% CIs for the difference in means, and two-sided p-values will be reported. Cohen's d effect size will be reported for the differences in treatment arms at each visit.

Overall mean change from baseline from week 4 to week 12 will be analyzed using MMRM in manner described in section 13.1.1 with adjustment for baseline tear osmolarity score, visit, treatment arm, and interaction between treatment arm and visit. Change from baseline by visit will also be analyzed. The LS means, LS mean differences, SEs, two-sided 95% CIs for LS means, two-sided 95% CIs for the difference in means, and two-sided p-values will be reported from the MMRM. Sensitivity analysis will include ANCOVA adjusted for baseline tear osmolarity and treatment arm, two-sample t-tests, and Wilcoxon rank sum test. Cohen's d effect size will be reported for the differences in treatment arms at each visit. The LS means, LS mean differences, SEs, two-sided 95% CIs for the difference in means, and two-sided p-values will be reported from the sensitivity analyses where applicable. Paired t-tests will be used to analyze change from baseline within each treatment arm.

Analyses will be performed on the ITT population with observed data only.

15. Safety Analyses

All safety analyses will be conducted using the Safety population.

15.1 Adverse Events

An AE is defined as any untoward medical occurrence associated with the use of an IP in humans, whether or not considered IP-related. An AE can be any unfavorable and unintended sign (e.g., an abnormal laboratory finding), symptom, or disease temporally associated with the use of an IP, without any judgment about causality. An AE can arise from any use of the IP (e.g., off-label use, use in combination with another drug or medical device) and from any route of administration, formulation, or dose, including an overdose.

All AEs will be coded using the MedDRA 22.0.

Treatment-emergent adverse events (TEAE) are defined as any event that occurs or worsens on or after the day that randomized study treatment is initiated. Adverse events recorded in the eCRF which began prior to treatment will not be included in the summary tables but will be included in the AE data listings.

An overall summary will be presented that includes the number and percentage of subjects who experienced at least one AEs, ocular AEs, non-ocular AEs, SAEs, AEs by maximal severity, AEs by relationship to study drug, AEs leading to treatment discontinuation, and AEs resulting in death by treatment arm for the Safety population. In addition, overall TEAEs and the number and percentage of subjects who experienced at least one TEAE, ocular TEAEs, non-ocular TEAEs, TE-SAEs, TEAEs by maximal severity, TEAEs by relationship to study drug, TEAEs leading to treatment discontinuation, and TEAEs resulting in death by treatment arm for the Safety population.

Separate summaries will be provided for the following categories of AEs:

- Ocular AEs by SOC and PT
- Non-ocular AEs by SOC and PT
- Ocular TEAEs by SOC and PT
- Non-ocular TEAEs by SOC and PT
- Ocular TEAEs by SOC, PT, and maximal severity
- Non-ocular TEAEs by SOC, PT, and maximal severity
- Ocular TEAEs by SOC, PT, and strongest relationship to study drug
- Non-ocular TEAEs by SOC, PT, and strongest relationship to study drug
- Ocular TEAEs by SOC, PT, maximal severity, and strongest relationship to study drug
- Non-ocular TEAEs by SOC, PT, maximal severity, and strongest relationship to study drug
- TEAEs That Led to Premature Discontinuation
- SAEs

Adverse Events and TEAEs will be summarized using discrete summary statistics and presented by treatment arm and all subjects for the the Safety population. If a subject reports the same PT multiple times within the same SOC, that PT will only be reported once within that SOC. As with the PT, if a subject reports multiple conditions within the same SOC, that SOC will only be reported once. In the summary, SOC will be listed in ascending alphabetical order; PTs will be listed in order of descending frequency for all subjects within each SOC.

Severity of an AE is defined as a qualitative assessment of the degree of intensity of an AE as determined by the investigator or reported to him/her by the subject. The assessment of severity is made irrespective of relationship to IP or seriousness of the event and should be evaluated according to the following scale:

- Mild: Event is noticeable to the subject, but is easily tolerated and does not interfere with the subject's daily activities.
- Moderate: Event is bothersome, possibly requiring additional therapy, and may interfere with the subject's daily activities.
- Severe: Event is intolerable, necessitates additional therapy or alteration of therapy, and interferes with the subject's daily activities.

Subjects experiencing more than one AE within a given SOC or PT are counted once within that SOC or PT for the maximal severity.

The relationship of each AE to the study drug should be determined by the investigator using these explanations:

- Definitely Related: Relationship exists when the AE follows a reasonable sequence from the time of IP administration, follows a known response pattern of the drug class, is confirmed by improvement on stopping the IP and no other reasonable cause exists.
- Probably Related: Relationship exists when the AE follows a reasonable sequence from the time of IP administration, follows a known response pattern of the drug class, is confirmed by improvement on stopping the IP and the suspect IP is the most likely of all causes.
- Possibly Related: Relationship exists when the AE follows a reasonable sequence from the time of administration, but could also have been produced by the subject's clinical state or by other drugs administered to the subject.
- Unlikely to be related: Relationship between the AE and IP is more than likely to be unrelated to IP administration but the relationship cannot be definitely attributed to another cause.
- Not Related: Concurrent illness, concurrent medication, or other known cause is clearly responsible for the AE, the administration of the IP and the occurrence of the AE are not reasonably related in time, OR exposure to IP has not occurred.

Subjects experiencing more than one AE within a given SOC or PT are counted once within that SOC or PT for the maximum relationship. All AEs, ocular AEs, non-ocular AEs, and SAEs will be presented in subject listings.

15.2 Visual Acuity (ETDRS)

The logarithm of the minimum angle of resolution (logMAR) VA must be assessed using an ETDRS chart. The procedure used will be consistent with the recommendations provided for using the ETDRS eye chart. Visual acuity will be evaluated at the beginning of each visit in the study (i.e., prior to slit-lamp examination). Visual acuity testing should be done with most recent correction.

The observed and change from baseline VA will be summarized for each eye (study eye and fellow eye) using continuous descriptive statistics by visit for each treatment arm and for all actively treated subjects. A subject listing of VA will also be produced.

15.3 Slit-Lamp Biomicroscopy Examination

Slit lamp biomicroscopic observations will be graded as Normal or Abnormal. Abnormal findings will be categorized as clinically significant (CS) (i.e., findings that may interfere with study parameters or otherwise confound the data as determined by the investigator) or not clinically significant (NCS). Slit lamp biomicroscopic observations will be recorded at each visit, including Pre-[redacted] and Post-[redacted] for Visits 1 (Day -14) and 2 (Day 1).The following will be examined:

- Cornea
- Conjunctiva

- Anterior Chamber
- Iris
- Lens
- Lid

The results will be summarized using counts and percentages for each treatment arm at each visit for each eye (study eye and fellow eye). Percentages will be based on the number of subjects in each treatment arm with responses. Shift tables for the slit-lamp biomicroscopy parameters will also be provided comparing each follow-up visit to baseline. A subject listing of the slit-lamp biomicroscopy parameters will also be produced.

15.4 Intraocular Pressure

Intraocular pressure will be measured in each eye by contact tonometry by the examiner at Visit 1 (Day -14) Post-[redacted] and Visit 9 (Day 85). A single measurement is made to obtain a determination of IOP. The same tonometer employing the investigator's standard technique will be used throughout the study. In addition, all reasonable efforts will be made to have the same examiner obtain all IOP measurements for a given subject. Intraocular pressure will be assessed by non-contact tonometry at each visit. Results will be taken from a single measurement and will be recorded in millimeters of Mercury (mmHg).

The IOP values and changes from baseline for each eye (study eye and fellow eye) will be summarized using continuous descriptive statistics by visit and eye for each treatment arm. A subject listing of IOP will also be produced.

15.5 Dilated Fundoscopy Examination

Dilated fundus exams will be performed using indirect ophthalmoscopy at Visit 1 (Day -14) Post[redacted] and Visit 9 (Day 85). The investigator will make observations of the vitreous, retina, macula, choroid and optic nerve.

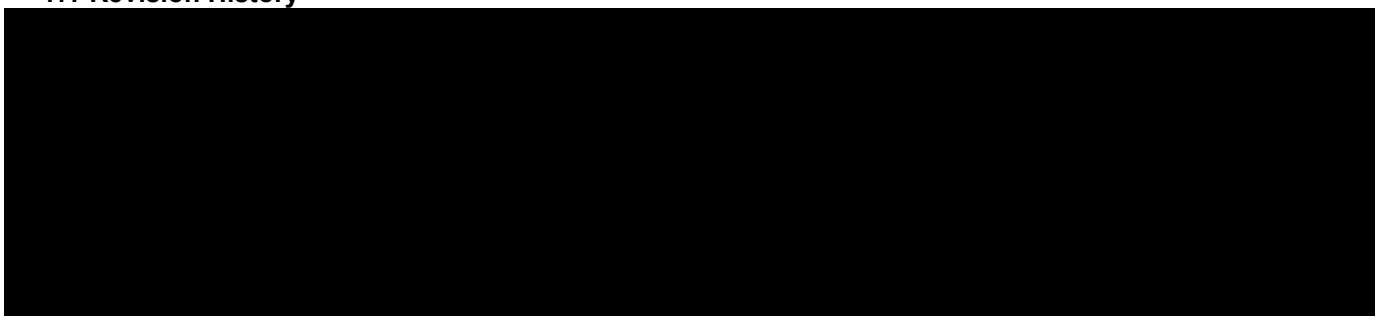
Observations will be graded as Normal or Abnormal. Abnormal findings that are CS (as determined by the investigator that may interfere with study parameters or otherwise confound the data) and those that are NCS will be described. An indirect fundoscopy examination should be performed if retinal disease is detected.

The results will be summarized using counts and percentages for each treatment arm and for all actively treated subjects at each visit for each eye (study eye and fellow eye). Percentages will be based on the number of subjects in each treatment arm with responses. Shift tables for the undilated fundoscopy parameters will also be provided comparing each follow-up visit to baseline. A subject listing of the dilated fundoscopy parameters will also be produced.

16. Changes from Protocol-Stated Analyses

The description of the testing strategy for multiple dosing regimens in Section 8.5 was revised from the protocol to specify that the QID for four weeks followed by BID for eight weeks dosing regimen statistical comparisons will be considered valid only if the QID for twelve weeks dosing regimen comparisons are statistically significant for either of the co-primary endpoints.

17. Revision History



18. Tables

Table Number	Title	Population
14.1.1	SUBJECT DISPOSITION	ALL RANDOMIZED SUBJECTS
14.1.2.1	DEMOGRAPHIC CHARACTERISTICS	ITT
14.1.2.2	DEMOGRAPHIC CHARACTERISTICS	SAFETY
14.1.3.1	BASELINE DISEASE CHARACTERISTICS	ITTOD
14.1.3.2	BASELINE DISEASE CHARACTERISTICS	ITTFSN
14.1.3.3	BASELINE DISEASE CHARACTERISTICS	ITT
14.1.4.1	OCULAR MEDICAL HISTORY	SAFETY
14.1.4.2	NON-OCULAR MEDICAL HISTORY	SAFETY
14.1.5.1	OCULAR CONCOMITANT MEDICATIONS	ITT
14.1.5.2	NON-OCULAR CONCOMITANT MEDICATIONS	ITT

Table Number	Title	Population
14.2.1.1.1	OCULAR DRYNESS SCORE	ITTOD WITH OBSERVED DATA ONLY
14.2.1.1.2	OCULAR DRYNESS SCORE	ITTOD WITH MCMC
14.2.1.1.3	OCULAR DRYNESS SCORE	ITTOD WITH PMM
14.2.1.1.4	OCULAR DRYNESS SCORE	PPOD WITH OBSERVED DATA ONLY
14.2.1.2.1	FLUORESCEIN STAINING OF THE NASAL REGION	ITTFSN WITH OBSERVED DATA ONLY
14.2.1.2.2	FLUORESCEIN STAINING OF THE NASAL REGION	ITTFSN WITH MCMC
14.2.1.2.3	FLUORESCEIN STAINING OF THE NASAL REGION	ITTFSN WITH PMM
14.2.1.2.4	FLUORESCEIN STAINING OF THE NASAL REGION	PPFSN WITH OBSERVED DATA ONLY
14.2.2.1.1	[redacted] DRYNESS	ITT POPULATION WITH OBSERVED DATA ONLY
14.2.2.1.2	[redacted] DRYNESS	PP POPULATION WITH OBSERVED DATA ONLY

Table Number	Title	Population
14.2.3.1.1	FLUORESCEIN STAINING [redacted]	ITT WITH OBSERVED DATA ONLY
14.2.3.1.2	FLUORESCEIN STAINING [redacted]	PP WITH OBSERVED DATA ONLY
14.2.3.2.1	UNANESTHETIZED SCHIRMER'S TEST	ITT WITH OBSERVED DATA ONLY
14.2.3.2.2	UNANESTHETIZED SCHIRMER'S TEST	PP WITH OBSERVED DATA ONLY
14.2.3.3.1	OCULAR DRYNESS THROUGH VISIT 5 (DAY 29)	ITT WITH OBSERVED DATA ONLY
14.2.3.3.2	OCULAR DRYNESS THROUGH VISIT 5 (DAY 29)	PP WITH OBSERVED DATA ONLY
14.2.4.1	VISUAL ANALOG SCALE	ITT WITH OBSERVED DATA ONLY
14.2.4.2	[redacted] DISCOMFORT SCALE	ITT WITH OBSERVED DATA ONLY
14.2.4.3	[redacted] SYMPTOM QUESTIONNAIRE	ITT WITH OBSERVED DATA ONLY
14.2.4.4	OCULAR SURFACE DISEASE INDEX (OSDI)	ITT WITH OBSERVED DATA ONLY

Table Number	Title	Population
14.2.4.5	SYMPTOM ASSESSMENT IN DRY EYE (SANDE)	ITT WITH OBSERVED DATA ONLY
14.2.4.6	[redacted] ITCHING SCALE	ITT WITH OBSERVED DATA ONLY
14.2.4.7	LISSAMINE GREEN STAINING [redacted]	ITT WITH OBSERVED DATA ONLY
14.2.4.8	TEAR FILM BREAK-UP TIME (SEC)	ITT WITH OBSERVED DATA ONLY
14.2.4.9	TEAR OSMOLARITY (MOSM/L)	ITT WITH OBSERVED DATA ONLY
14.2.5	SUMMARY OF EFFICACY ANALYSES	
14.3.1.1	OVERALL SUMMARY OF ADVERSE EVENTS BY TREATMENT ARM	SAFETY
14.3.2.1	OCULAR ADVERSE EVENTS BY SYSTEM ORGAN CLASS AND PREFERRED TERM	SAFETY
14.3.2.2	NON-OCULAR ADVERSE EVENTS BY SYSTEM ORGAN CLASS AND PREFERRED TERM	SAFETY
14.3.3.1	OCULAR TREATMENT-EMERGENT ADVERSE EVENTS BY SYSTEM ORGAN CLASS AND PREFERRED TERM	SAFETY
14.3.3.2	NON-OCULAR TREATMENT-EMERGENT ADVERSE EVENTS BY SYSTEM ORGAN CLASS AND PREFERRED TERM	SAFETY
14.3.4.1	OCULAR TREATMENT-EMERGENT ADVERSE EVENTS BY SYSTEM ORGAN CLASS, PREFERRED TERM, AND MAXIMAL SEVERITY	SAFETY

Table Number	Title	Population
14.3.4.2	NON-OCULAR TREATMENT-EMERGENT ADVERSE EVENTS (TEAES) BY SYSTEM ORGAN CLASS, PREFERRED TERM, AND MAXIMAL SEVERITY	SAFETY
14.3.5.1	OCULAR TREATMENT-EMERGENT ADVERSE EVENTS BY SYSTEM ORGAN CLASS, PREFERRED TERM, AND STRONGEST RELATIONSHIP TO STUDY DRUG	SAFETY
14.3.5.2	NON-OCULAR TREATMENT-EMERGENT ADVERSE EVENTS BY SYSTEM ORGAN CLASS, PREFERRED TERM, AND STRONGEST RELATIONSHIP TO STUDY DRUG	SAFETY
14.3.6.1	OCULAR TREATMENT-EMERGENT ADVERSE EVENTS BY SYSTEM ORGAN CLASS, PREFERRED TERM, SEVERITY, AND STRONGEST RELATIONSHIP TO STUDY DRUG	SAFETY
14.3.6.2	NON-OCULAR TREATMENT-EMERGENT ADVERSE EVENTS BY SYSTEM ORGAN CLASS, PREFERRED TERM, SEVERITY, AND STRONGEST RELATIONSHIP TO STUDY DRUG	SAFETY
14.3.7	TREATMENT-EMERGENT ADVERSE EVENTS THAT LED TO PREMATURE DISCONTINUATION	SAFETY
14.3.8	SERIOUS ADVERSE EVENTS	SAFETY
14.3.9	VISUAL ACUITY (ETDRS) – LOGMAR	SAFETY
14.3.10.1	SLIT LAMP BIOMICROSCOPY	SAFETY
14.3.10.2	SHIFT IN SLIT LAMP BIOMICROSCOPY	SAFETY
14.3.11	INTRAOCULAR PRESSURE (MMHG)	SAFETY
14.3.12.1	DILATED FUNDOSCOPY	SAFETY
14.3.12.2	SHIFT IN DILATED FUNDOSCOPY	SAFETY
14.3.13	COMPLIANCE TO STUDY DRUG	SAFETY
14.3.14	EXPOSURE TO STUDY DRUG	SAFETY

19. Listings

Listing Number	Title
16.1.7	RANDOMIZATION SCHEDULE
16.2.1.1	SUBJECT DISPOSITION
16.2.1.2	INCLUSION/EXCLUSION
16.2.2.1	PROTOCOL DEVIATIONS
16.2.3	STUDY POPULATION INCLUSION
16.2.4.1	DEMOGRAPHICS
16.2.4.2	OCULAR MEDICAL HISTORY
16.2.4.3	NON-OCULAR MEDICAL HISTORY
16.2.4.4	OCULAR CONCOMITANT MEDICATIONS
16.2.4.5	NON-OCULAR CONCOMITANT MEDICATIONS
16.2.5.1	RUN-IN, RUN-IN INSTILLATION, AND RUN-IN REPLACEMENT
16.2.5.2	STUDY DRUG ASSIGNMENT, INSTILLATION, AND REPLACEMENT
16.2.5.3	STUDY DRUG EXPOSURE AND DOSING COMPLIANCE
16.2.6.1	VISUAL ANALOG SCALE
16.2.6.2	FLUORESCEIN CORNEAL AND CONJUNCTIVAL STAINING [redacted]
16.2.6.3	UNANESTHETIZED SCHIRMER'S TEST
16.2.6.4	[redacted] OCULAR DISCOMFORT [redacted]
16.2.6.5	[redacted] CULAR DISCOMFORT [redacted]
16.2.6.6	[redacted] SYMPTOM QUESTIONNAIRE
16.2.6.7	OCULAR SURFACE DISEASE INDEX (OSDI)
16.2.6.8	SYMPTOM ASSESSMENT IN DRY EYE (SANDE)
16.2.6.9	[redacted]

Listing Number	Title
16.2.6.10	LISSAMINE GREEN CORNEAL AND CONJUNCTIVAL STAINING [redacted]
16.2.6.11	TEAR FILM BREAK-UP TIME (TFBUT)
16.2.6.12	TEAR OSMOLARITY
16.2.7.1	ALL ADVERSE EVENTS
16.2.7.2	OCULAR ADVERSE EVENTS
16.2.7.3	NON-OCULAR ADVERSE EVENTS
16.2.7.4	SERIOUS ADVERSE EVENTS
16.2.8.1	VISUAL ACUITY (ETDRS) - LOGMAR
16.2.8.2	SLIT-LAMP BIOMICROSCOPY
16.2.8.3	INTRAOCULAR PRESSURE (IOP)
16.2.8.4	DILATED FUNDOSCOPY
16.2.8.5	PREGNANCY TEST

20. Figures

Figure Number	Title	Population
14.2.1.1.1	OCULAR DRYNESS SCORE OVERALL MEAN CHANGE FROM BASELINE FROM WEEK 2 TO WEEK 12, QID ARMS	ITTOD WITH OBSERVED DATA ONLY
14.2.1.2.1	FLUORESCEIN STAINING OF THE NASAL REGION OVERALL MEAN CHANGE FROM BASELINE FROM WEEK 2 TO WEEK 12, QID ARMS	ITTFSN WITH OBSERVED DATA ONLY
14.2.2.1.1	OCULAR DRYNESS SCORE OVERALL MEAN CHANGE FROM BASELINE FROM WEEK 2 TO WEEK 12, QID TO BID ARMS	ITTOD WITH OBSERVED DATA ONLY

Figure Number	Title	Population
14.2.2.2.1	FLUORESCEIN STAINING OF THE NASAL REGION OVERALL MEAN CHANGE FROM BASELINE FROM WEEK 2 TO WEEK 12, QID TO BID ARMS	ITTFSN WITH OBSERVED DATA ONLY