K-161-2.01US

A Phase 2, Prospective, Double-Masked, Randomized, Multi-Center, Vehicle-Controlled, Parallel-group, 4-week Administration Study Assessing the Safety, Efficacy, and Optimum Dosage of K-161 in non-Japanese and Japanese Adult Subjects with Moderate to Severe Dry Eye Disease Both in Environmental and Controlled Adverse Environmental (CAE®) Settings

NCT #: 04084483

Statistical Analysis Plan

Ver. 1.0

03 Jan 2020

STATISTICAL ANALYSIS PLAN

A Phase 2, Prospective, Double-Masked, Randomized, Multi-Center, Vehicle-Controlled, Parallel-group, 4-week Administration Study Assessing the Safety, Efficacy, and Optimum Dosage of K-161 in non-Japanese and Japanese Adult Subjects with Moderate to Severe Dry Eye Disease Both in Environmental and Controlled Adverse Environmental (CAE®) Settings

Sponsor: Kowa Research Institute, Inc. (KRI)

Protocol Number: K-161-2.01US / [redacted]

[redacted]

Date: 03JAN2020

Version: 1.0

A Phase 2, Prospective, Double-Masked, Randomized, Multi-Center, Vehicle-Controlled, Parallel-group, 4-week Administration Study Assessing the Safety, Efficacy, and Optimum Dosage of K-161 in non-Japanese and Japanese Adult Subjects with Moderate to Severe Dry Eye Disease Both in Environmental and Controlled Adverse Environmental (CAE®) Settings

Protocol Number: K-161-2.01US / [redacted]

SAP Version: 1.0

SAP Date: 03JAN2020

Statistical Analysis Plan Approval

[redacted]

A Phase 2, Prospective, Double-Masked, Randomized, Multi-Center, Vehicle-Controlled, Parallel-group, 4-week Administration Study Assessing the Safety, Efficacy, and Optimum Dosage of K-161 in non-Japanese and Japanese Adult Subjects with Moderate to Severe Dry Eye Disease Both in Environmental and Controlled Adverse Environmental (CAE®) Settings

Protocol Number: **K-161-2.01US /** [redacted]

SAP Version: **1.0**

SAP Date: **03JAN2020**

Statistical Analysis Plan Approval

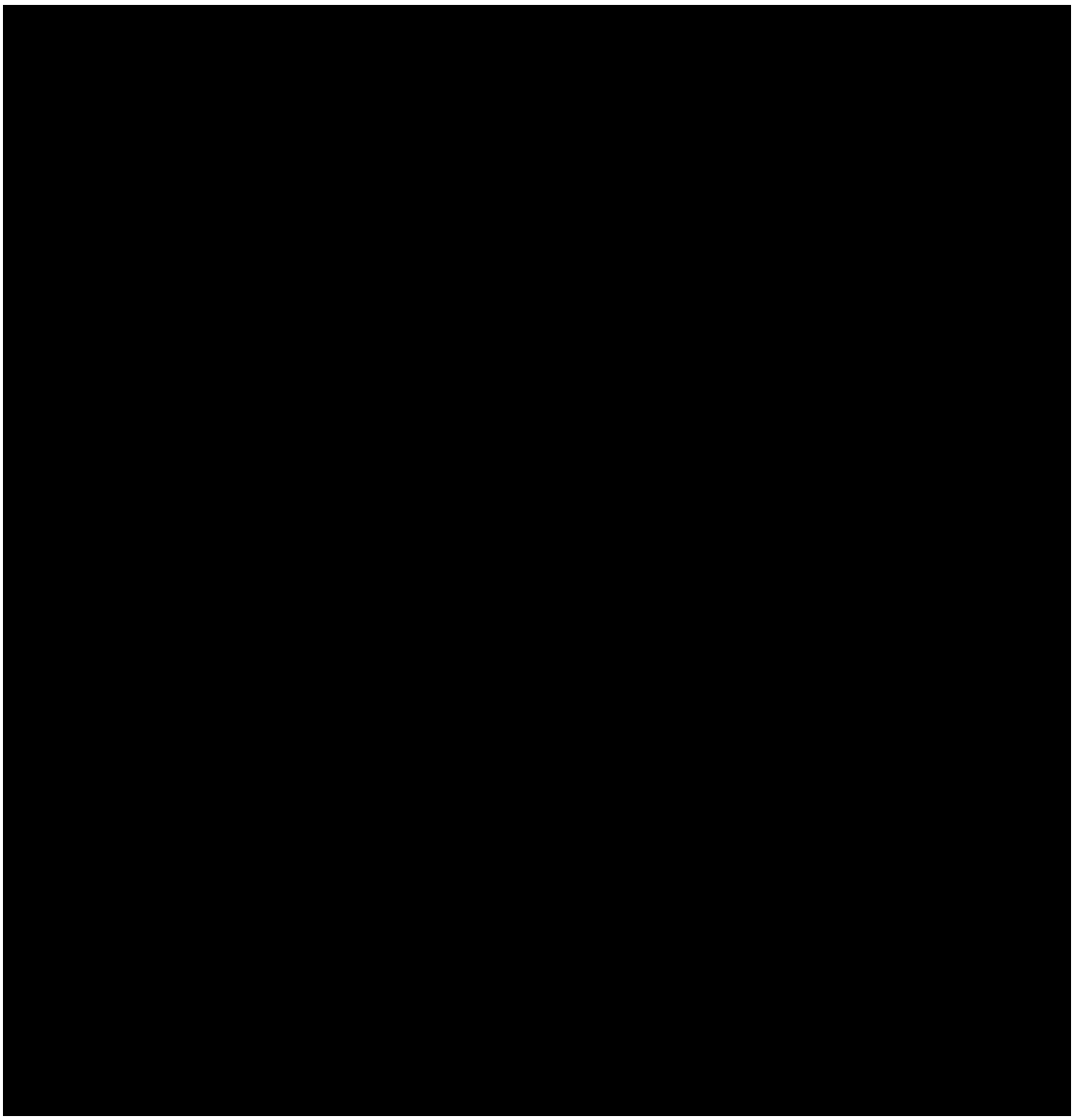


Table of Contents

[redacted]

List of Abbreviations

ADaM	Analysis Data Model
AE	Adverse Event
ANCOVA	Analysis of Covariance
ATC	Anatomical Therapeutic Chemical
[redacted]	
CI	Confidence Interval
CRO	Clinical Research Organization
CS	Clinically Significant
DMP	Data Management Plan
eCRF	Electronic Case Report Form
[redacted]	
IB	Investigator's Brochure
ICF	Informed Consent Form
ICH	International Conference on Harmonisation
[redacted]	
ITT	Intent-to-Treat
IWRS	Interactive Web Response System
LOCF	Last Observation Carried Forward
[redacted]	
LS	Least Squares
MCMC	Markov Chain Monte Carlo
MedDRA	Medical Dictionary for Regulatory Activities
NCS	Not Clinically Significant
[redacted]	
ODS	Ocular Discomfort Score
OSDI©	Ocular Surface Disease Index
OU	*Oculus uterque* (Both Eyes)
PDF	Portable Document Format
[redacted]	
PP	Per Protocol
PT	Preferred Term
[redacted]	
RTF	Rich Text Format
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
[redacted]	
SDC	Statistics & Data Corporation
SDTM	Study Data Tabulation Model
SE	Standard Error
SOC	System Organ Class
TEAE	Treatment-Emergent Adverse Event
TFBUT	Tear Film Break-Up Time
TMF	Trial Master File

[redacted]	
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 protocol K-161-2.01US / [redacted] Amendment 2 dated 26JUL2019

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

The objectives of the study are to assess the safety, efficacy, optimum dosage, and dosing regimen of K-161 compared to its vehicle from Visit 2 (Day 1) to Visit 5 (Day 29) in non-Japanese and Japanese adult subjects with moderate to severe dry eye disease both in environmental and Controlled Adverse Environment (CAE®) settings.

2.1 Primary Efficacy Endpoints

The primary efficacy endpoints are the following:

- Change in inferior corneal staining score from Visit 2 (Day 1) to Visit 5 (Day 29) by comparing [redacted] K-161 [redacted] to its vehicle [redacted] in CAE® setting
- Change in Ora Calibra® Ocular Discomfort Scale from Visit 2 (Day 1) to Visit 5 (Day 29) by comparing [redacted] K-161 [redacted] to its vehicle [redacted] in CAE® setting

2.2 Important Secondary Efficacy Endpoints

The important secondary efficacy endpoints include the following:

- Change in Schirmer's Test value (unanesthetized) from Visit 2 (Day 1) to Visit 5 (Day 29) by comparing [redacted] K-161 [redacted] to its vehicle [redacted] in environmental setting
- Change in tear film break up time (TFBUT) from Visit 2 (Day 1) to Visit 5 (Day 29) by comparing [redacted] K-161 [redacted] to its vehicle [redacted] in CAE® setting

2.3 Secondary Efficacy Endpoints

The secondary efficacy endpoints for comparisons of the vehicle [redacted] and K-161 [redacted] (i.e., [redacted] K-161 [redacted] and [redacted] K-161 [redacted] treatment groups include the following:

- Fluorescein staining by region: central, superior, inferior, temporal, nasal, and corneal sum as assessed by the Ora Calibra® Corneal and Conjunctival Staining Scale
- Lissamine green staining by region: central, superior, inferior, temporal, nasal, and conjunctival sum as assessed by the Ora Calibra® Corneal and Conjunctival Staining Scale
- Conjunctival Redness as assessed by the Ora Calibra® Scale
- TFBUT
- Tear Osmolarity
- Schirmer's Test (unanesthetized)
- Blink Rate
- Ora Calibra® Ocular Discomfort Scale
- Ora Calibra® Ocular Discomfort and 4-Symptom Questionnaire
- Visual Analog Scale (VAS): Burning/Stinging, Itching, Foreign Body Sensation, Blurred Vision, Eye Dryness, Photophobia, and Pain
- Ocular Surface Disease Index (OSDI©)

2.4 [redacted]

[redacted]

- [redacted]
- [redacted]
- [redacted]

2.5 Safety Variables

The safety variables include the following:

- [redacted]
- [redacted]
- Adverse Event (AE) query
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- Maximum dose tolerability by comparing the [redacted] K-161 [redacted] arm to the [redacted] and [redacted] K-161 [redacted] arms

2.6 Other Measures

The other measures being evaluated are:

- [redacted]
- [redacted]
- [redacted]
- [redacted]

2.7 Statistical Hypotheses

The following primary hypotheses will be tested against their respective two-sided alternative hypotheses in the order:

H_{01}: There is no difference in the change from baseline in the post-CAE® inferior corneal staining score after 28 days of treatment of [redacted] K-161 [redacted] compared to vehicle.

H_{02}: There is no difference in the change from baseline in the post-CAE® ocular discomfort score after 28 days of treatment of [redacted] K-161 [redacted] compared to vehicle.

Upon rejecting both H_{01} and H_{02}, the secondary hypotheses will be tested in the following order:

H_{03}: There is no difference in the change from baseline in the pre-CAE® Schirmer's test value after 28 days of treatment of [redacted] K-161 [redacted] compared to vehicle.

H_{04}: There is no difference in the change from baseline in the post-CAE® TFBUT after 28 days of treatment of [redacted] K-161 [redacted] compared to vehicle.

These four hypotheses will be tested hierarchically, each at the alpha = 0.05 level to maintain a study-wise Type I error rate at 0.05.

3. Study Design and Procedures

3.1 General Study Design

This is a Phase 2, multicenter, randomized, double-masked, vehicle controlled, parallel-group design with block enrollment. Subjects will be randomized to one of the following treatment groups at Visit 2 (Day 1):

- [redacted] K-161; [redacted] (N=~60)
- [redacted] K-161; [redacted] (N=~60)
- [redacted] K-161; [redacted] (N=~60)
- Vehicle Ophthalmic Solution (Vehicle); [redacted] (N=~60)

Approximately 240 subjects will be randomly assigned to one of the four treatment groups (1:1:1:1). [redacted] Subjects, Sponsor, Contract research organization (CRO), and site personnel will be masked to treatment assignment. To ensure masking, a dedicated unmasked staff member/technician will

be delegated to dispense and observe instillation of randomized study drug and collect randomized study drug from a subject for drug accountability. This person cannot perform any other study-related procedures. Additionally, the informed consent form (ICF) will not indicate any relationship between dosing regimen and treatment. Subjects will be instructed not to discuss their assigned dosing regimen or perceived treatment effects with other study participants.

During the screening period, two 90-minute exposures to the CAE® will be conducted to ascertain eligibility to enter the study. Subjects who qualify after the initial screening visit will enter the run-in phase, where they will self-administer vehicle [redacted] for approximately 14 days. Those who qualify at Visit 2 (Day 1) will be randomized to receive study drug in a double-masked fashion for 28 days. [redacted] Subjects will self-administer drops either [redacted] or [redacted]

At Visit 4 (Day 15) and Visit 5 (Day 29), CAE® exposure will occur, with pre-CAE®, during CAE® (symptoms only) and post-CAE® assessments of ocular signs and symptoms. At Visit 3, no CAE® exposure will occur but signs and symptoms will be assessed. Study drug will be discontinued at Visit 5. Subjects will exit from the study at this visit. The follow-up phone call will occur about 7 days after the final day of randomized study drug treatment.

Study visits will be referred to in all tables and listings as the expected study day corresponding to the visit to enable reviewers to understand the assessment timing without referring to the protocol visit schedule. 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. Study 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	± 1 Days
Visit 4	Day 15	± 2 Days
Visit 5	Day 29	± 2 Days
Follow-up Phone Call	Day 36	± 2 Days

3.2 Schedule of Visits and Assessments

The schedule of visits and assessments is provided on Table 2.

Table 2. Schedule of Visits and Assessments

Procedure	Visit 1 Day -14 ± 2		Visit 2 Day 1		Visit 3 Day 8 ± 1	Visit 4 Day 15 ± 2		Visit 5 Day 29 ± 2		Follow-up Phone Call (Day 36 ± 2)
	Pre CAE	Post CAE	Pre CAE	Post CAE	Non CAE	Pre CAE	Post CAE	Pre CAE	Post CAE	
Informed Consent / HIPAA	X									
Medical / Medication History and Demographics	X									
[redacted]										
Medical / Medication Update			X		X	X		X		
Adverse Event Query		X	X	X	X	X	X	X	X	X
[redacted]										
Ora Calibra® Ocular Discomfort Scale	X	X	X	X	X	X	X	X	X	
Ora Calibra® Ocular Discomfort & 4-Symptom Questionnaire	X	X	X	X	X	X	X	X	X	
OSDI© Questionnaire	X		X		X	X		X		
Visual Analog Scale	X	X	X	X	X	X	X	X	X	
[redacted]										
Conjunctival Redness	X	X	X	X	X	X	X	X	X	
Tear Osmolarity			X		X	X		X		
[redacted]										
TFBUT	X	X	X	X	X	X	X	X	X	
Fluorescein Staining	X	X	X	X	X	X	X	X	X	
Lissamine Green Staining	X	X	X	X	X	X	X	X	X	
Schirmer's Test	X		X		X	X		X		
Blink Rate			X		X	X		X		
[redacted]										

Procedure	Visit 1 Day -14 ± 2		Visit 2 Day 1		Visit 3 Day 8 ± 1	Visit 4 Day 15 ± 2		Visit 5 Day 29 ± 2		Follow-up Phone Call (Day 36 ± 2)
	Pre CAE	Post CAE	Pre CAE	Post CAE	Non CAE	Pre CAE	Post CAE	Pre CAE	Post CAE	
[redacted]										
Vehicle Run-In Dispensation		X								
Vehicle Run-In Instillation		X								
Vehicle Run-in Collection			X							
Randomization				X						
Study Drug Dispensation				X	X		X			
[redacted]										
Study Drug Collection					X	X		X		
Exit Subject from Study										X

[redacted]

[redacted]

[redacted]

4. Study Treatments

4.1 Method of Assigning Subjects to Treatment Groups

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

[redacted]

For the non-Japanese subjects, randomization will be further stratified by the following factors and cut-offs:

[redacted]

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).

4.2 Masking and Unmasking

Subjects, Sponsor, CRO, and site personnel (with the exception of the following person) will be masked to treatment assignment. To ensure masking, a dedicated unmasked staff member/technician will be delegated to dispense and observe instillation of randomized study drug and collect randomized study drug from a subject for drug accountability. This person cannot perform any other study-related procedures. Prior to database lock, select eCRF pages will only be viewable by unmasked study members as detailed in the Data Management Plan (DMP) to ensure masking.

When medically necessary, the Investigator may need to determine what treatment group has been assigned to a subject. When possible (i.e., in non-emergent situations), [redacted] and/or the Sponsor will be notified before unmasking the Investigator. [redacted] and/or the Sponsor must be informed immediately about any unmasking event.

If the Investigator identifies a medical need that requires unmasking the treatment assignment of a subject, [redacted] and/or the Medical Monitor will be contacted prior to unmasking the identity of the IP, if possible. [redacted] will ask the staff to complete and send them the Unmasking Request Form. [redacted] will notify the Sponsor and jointly determine if the unmasking request should be granted. In addition, 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 will be provided on the Unmasking Request Form. The Investigator will unmask the subject using the IWRS, complete the Unmasking Memo form, include it in the subject's study file and make a copy for the trial master file (TMF). For each unmasked request, the reason, date, signature, and name of the person unmasking the subject must be noted in the subject's study file.

Unmasked subjects will be discontinued from the study and followed for safety monitoring, until resolution of the AE or study completion, whichever occurs last.

5. Sample Size and Power Considerations

This study is expected to enroll 240 subjects in a 1:1:1:1 ratio across four treatment groups, or 60 subjects per treatment group. In addition, [redacted]

The standard deviation (SD) for change from baseline in post-CAE® inferior corneal fluorescein staining ranges between 0.68 and 0.81 units. Assuming a common SD of 0.77, a sample size of 60 subjects per group will have approximately 90% power to detect a difference of 0.46 units between the active treatment group and the vehicle group using a two-sample t-test at a two-sided significance level of 0.05. Again, it is assumed that adjusting for baseline and site will further reduce variability and provide >90% power.

Likewise, the SD for change from baseline scores in the post-CAE® ODS at Day 29 is assumed a common SD of 0.82, a sample size of 60 subjects per group will have approximately 90% power to detect a difference of 0.49 units between the active treatment group and the vehicle group using a two-sample t-test at a two-sided significance level of 0.05. It is assumed that adjusting for baseline and site will further reduce variability and provide >90% power.

6. Data Preparation

6.1 Input Data

Electronic Case Report Forms will be developed by Statistics & Data Corporation (SDC). SDC will utilize the Electronic Data Capture (EDC) system, iMedNet™ (iMedNet v1.192.1) for this study. Data from source documents will be entered into the eCRF by site personnel.

In addition, the following study data which is not captured directly within the RDC system but is obtained from external vendors will also be included for analysis. These data sources are described in detail in data transfer agreements developed between data management and the respective external laboratory or reading center:

- [redacted]
- [redacted]

[redacted]

When all prerequisites for database lock have been met, including availability of all masked external data, the database will be locked. Following database lock, approval will be obtained from the Sponsor to unmask the study. Once the study has been unmasked, unmasked laboratory data will be sent to SDC. Any changes to the database after data have been locked can only be made with the approval of the Sponsor in consultation with SDC.

Final analysis will be carried out after the following have occurred:

- Database lock has occurred, including receipt of all final versions of external vendor data, with written authorization provided by appropriate SDC and Sponsor personnel.
- Protocol deviations have been identified and status defined (major/minor deviations).
- Analysis populations have been determined.
- Randomized treatment codes have been unblinded.

6.2 Output Data

Data from EDC and external data will be transferred to Biostatistics and incorporated into standard formats following the Study Data Tabulation Model (SDTM). Data will then be mapped to analysis datasets using the Analysis Data Model (ADaM). Both SDTM- and ADaM-formatted data will be used to create the subject listings, while all tables and figures will be based on the ADaM-formatted data.

SDTM will follow the SDTM version 1.4 model and will be implemented using the SDTM Implementation Guide version 3.2 and the SDTM Controlled Terminology version 2019-06-28. ADaM data will follow the ADaM version 2.1 model and will be implemented using the ADaM Implementation Guide version 1.1. Both SDTM and ADaM will be validated using Pinnacle 21 version 3.0.1. Any discrepancies in the validation will be noted in reviewer's guides accompanying the final data transfers.

Define.xml will be created for SDTM and ADaM using the Define-XML version 2.0 model.

7. Analysis Populations

7.1 Intent-to-Treat

The intent-to-treat (ITT) population includes all randomized subjects. Subjects in the ITT population will be analyzed as randomized.

7.2 Per Protocol

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.

7.3 Safety

The safety population includes all randomized subjects who have received at least one dose of the investigational product. Subjects in the safety population will be analyzed as treated.

7.4 [redacted]

[redacted]

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, the unit of analysis will be the study eye, or the "worst eye," as defined by the following:

Study (worst) Eye: Eyes are eligible for analysis if they meet all of the inclusion criteria. At least one eye (the right eye or the left eye) must meet all of the criteria. In the case that both eyes are eligible for analysis, the worst eye will be selected as the eye with the worst post-CAE® inferior corneal staining at Visit 2 (Day 1). If the post-CAE® inferior corneal staining at Visit 2 (Day 1) is the same in both eyes then the eye with the worst post-CAE® Ocular Discomfort Score (ODS) at Visit 2 (Day 1) will be worst eye. If the post-CAE® inferior corneal staining and ocular discomfort are the same in both eyes then the right eye will be selected as the worst eye.

8.2 Missing or Inconclusive Data Handling

Missing data will be imputed using multiple imputation with a control-based pattern mixture model on the ITT population for the primary analyses, the important secondary, other secondary, and exploratory comparisons.

Sensitivity analyses of the primary analyses and important secondary comparisons will include the following in order to provide a robust understanding of the impact of missing and spurious data:

- Using complete case data (i.e., observed data only) on the ITT population
- Using last observation carried forward (LOCF) on the ITT population
- Multiple imputation under missing at random assumption with Markov Chain Monte Carlo (MCMC) techniques on the ITT population
- Using complete case data on the PP population.

No imputation will be used for safety endpoints.

8.3 Definition of Baseline

Baseline is defined as the last measurement prior to the first dose of study medication. If a measure is taken both pre-CAE® and post-CAE®, the baseline will be the time point matched value. For changes from pre-CAE® to post-CAE® after the first treatment, the change from pre-CAE® to post-CAE® will be considered the baseline value.

For [redacted] baseline will be measurements taken at Visit 1 (Day -14).

Change from baseline will be calculated as follow-up visit minus baseline visit.

8.4 Data Analysis Conventions

All data analysis will be performed by SDC after the study is completed and the database has been locked and released for unmasking. Statistical programming and analyses will be performed using 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. Minima and maxima will be reported with the same precision as the raw values; means and medians will be presented to one additional decimal place than reported in the raw values. Standard deviations will be presented to one additional decimal place than reported in the raw values. Summaries for discrete variables will include counts and percentages. All percentages will be rounded to one decimal place (i.e., XX.X%). Differences between active treatment groups and Vehicle will be calculated as active minus Vehicle and change from baseline will be calculated as follow-up visit minus baseline.

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 groups will be two-sided at 95% confidence. All p-values will be rounded to 4 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 group and, where appropriate, visit. Listings will be sorted by treatment group, subject number, visit/time point, and parameter as applicable.

For statistical analyses that use site as a fixed effect, any site that enrolls fewer than 10 subjects will be pooled with the site with the next smallest enrollment until each pooled site contains at least 10 subjects.

8.5 Adjustments for Multiplicity

Common factors for multiplicity are multiple groups and multiple endpoints.

1. Multiple groups

The primary treatment comparison will be between the [redacted] K-161 [redacted] treatment group and the vehicle [redacted] treatment group. The other treatment comparisons will be considered secondary/exploratory.

2. Multiple endpoints

The primary and important secondary endpoints will be tested hierarchically to maintain the study-wise Type I error rate of 0.05. First, change from baseline in inferior corneal fluorescein staining will be tested at an alpha level of 0.05. The test for change from baseline scores in the ODS will be conducted at an alpha level of 0.05 only if the previous test demonstrates significance. The test for change from baseline in Schirmer's test will be conducted at an alpha level of 0.05 only if the previous test demonstrates significance. The test for change from baseline in TFBUT will be conducted at an alpha level of 0.05 only if

the previous test demonstrates significance. All other secondary endpoints and treatment comparisons will be considered exploratory.

The hierarchical testing order will be as follows:

Primary comparisons:

1. Comparison of the change from baseline in post-CAE® inferior corneal staining at Visit 5 (Day 29) ([redacted] K-161 [redacted] vs vehicle)
2. Comparison of the change from baseline in post-CAE® ocular discomfort at Visit 5 (Day 29) ([redacted] K-161 [redacted] vs vehicle)

Important Secondary comparisons:

3. Comparison of the change from baseline in pre-CAE® Schirmer's test value at Visit 5 (Day 29) ([redacted] K-161 [redacted] vs vehicle)
4. Comparison of the change from baseline in post-CAE® TFBUT at Visit 5 (Day 29) ([redacted] K-161 [redacted] vs vehicle)

Thus, the study-wise Type I error rate will be maintained at 0.05 by considering only one dose of K-161 as the primary treatment comparison and using a hierarchical testing procedure.

9. Disposition of Subjects

Subject disposition will be presented in terms of the numbers and percentages of subjects who were randomized to blinded study medication with subcategories of dosed with the blinded study medication, did not dose with blinded study medication; who were included in the following analysis populations: ITT, PP, safety, [redacted] and who 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 group and for all subjects. Percentages will be calculated using randomized subjects as the denominator unless otherwise specified.

The total number of enrolled subjects and screen failed subjects will be presented.

The reasons for premature study discontinuation will be summarized by treatment group for all discontinued subjects. Percentages will be calculated using discontinued subjects as the denominator. The reasons for study discontinuation that will be summarized include: AE, Subject Request/Withdrawal, Protocol Violation(s), Administrative Reasons, Sponsor Termination of Study, 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 any deviation, major deviation, and minor deviation will be summarized by treatment group for all randomized subjects. The protocol deviations that will be summarized include the following categories: 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, 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 a masked review.

In addition, subject listings will be provided that include informed consent date, inclusion and exclusion criteria violations, exclusions from the PP population, and screen failures. Details of the study randomization, including randomization date and time, and randomized treatment, will also be included within a subject listing.

10. Demographic and Pretreatment Variables

10.1 Demographic Variables

The demographic variables collected in this study include age, sex, ethnicity, race, whether the subject is Japanese, if the subject's parents or grandparents are of Japanese descent, height, and weight. Subjects who record more than one race will be grouped into a single category denoted as Multiple. 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 = (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, for age category, sex, race, ethnicity, age, sex, ethnicity, race, ethnic Japanese, Japanese descent, and stratification category.

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

10.2 Pretreatment Variables

Baseline disease characteristics will be summarized, overall and by treatment, for the ITT and safety populations using continuous descriptive statistics for inferior corneal staining score; ocular discomfort score; uanesthetized Schirmer's test value; TFBUT; fluorescein staining score (Ora Calibra® scale) in the regions: central, superior, inferior, temporal, corneal sum, conjunctival sum and total eye score; lissamine green staining by region: central, superior, inferior, temporal, nasal, and conjunctival sum as assessed by the Ora Calibra® Corneal and Conjunctival Staining Scale; [redacted] 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

Medical history will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) 22.0.

Ocular and non-ocular medical history will be summarized using discrete summary statistics and presented by treatment and overall at the subject level by System Organ Class (SOC) and Preferred Term (PT) using 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. SOCs are listed in alphabetical order; PTs within a SOC are listed in order of descending frequency across all subjects.

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 the World Health Organization 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 group using frequencies and percentages. Subjects may have more than one medication per ATC text. At each level of subject summarization, a subject will be counted once if he/she reports one or more medications. Percentages will be based on the number of subjects in each treatment group. Listings of prior and 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 actual doses received and comparing that to the number of expected doses as follows:

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

The number of actual doses received will be recorded in the eCRF and determined through the in-office instillations. The number of expected doses that will be used for calculating compliance will be calculated as follows:

- For subjects in [redacted] dosing group: 4 x {[Date of Study Completion/Discontinuation – Date of Visit 2 (Day 1)]} – 1 x [number of visits attended in (Visit 2, Visit 3, Visit 4)]
- For subjects in [redacted] dosing groups: 2 x {[Date of Study Completion/Discontinuation – Date of Visit 2 (Day 1)]}

If a randomized subject discontinues from the study on Day 1, the number of expected doses will be 1.

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 group using the safety population. The compliance category defined above will be summarized with discrete summary statistics.

A subject listing of dosing compliance will also be produced.

A subject listing of run-in, run-in instillation, and run-in replacement will be produced.

A subject listing of study drug assignment, study drug instillation, and study drug replacement will be produced.

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))

If a randomized subject discontinues from the study on Day 1, the extent of exposure will be 1 day.

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 for each subject exposed to study drug will be summarized with continuous descriptive statistics for each treatment group using the safety population. A subject listing of treatment exposure will also be produced.

13. Efficacy Analyses

13.1 Primary Analysis

The primary efficacy endpoints of the study are:

- Mean change from baseline in post-CAE® inferior corneal fluorescein staining score on the Ora Calibra® scale at Visit 5 (Day 29), comparing [redacted] K-161 [redacted] to vehicle [redacted]
- Mean change from baseline in post-CAE® ocular discomfort scale on the Ora Calibra® scale at Visit 5 (Day 29), comparing [redacted] K-161 [redacted] to vehicle [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. Treatment comparisons between [redacted] K-161 [redacted] and vehicle [redacted] will be calculated as [redacted] K-161 [redacted] - vehicle [redacted] All primary endpoints will be described in subject listings.

13.1.1 Inferior Corneal Fluorescein Staining Score At Visit 5 (Day 29) Post-CAE, [redacted] K-161 [redacted] vs Vehicle [redacted]

Fluorescein staining will be conducted at all visits at pre-CAE® and post-CAE®. Grading will be conducted using the Ora Calibra® Scale from 0 to 4 with the use of half grade (0.5) increments, where grade 0 = None and 4 = Severe. The regions to be assessed will be the central, superior, inferior, temporal, nasal, corneal sum, conjunctival sum, and total eye score for study eye only.

Inferior corneal fluorescein staining will be summarized at Visit 5 (Day 29) Post-CAE® by treatment groups using continuous descriptive statistics. Primary analysis will use an analysis of covariance (ANCOVA) model adjusted for baseline value, site, and with treatment group as the explanatory variable. Changes from baseline for each treatment group in inferior corneal fluorescein staining will be compared between [redacted] K-161 [redacted] and vehicle [redacted] In addition, treatment by baseline and treatment by study site interactions will be explored in separate models to evaluate how the treatment effect may differ for baseline value and study sites. Analyses will be performed by baseline stratification (≤ 3.0 and >3.0) and/or site to understand how the treatment effect differs for baseline value and study sites. Least squares (LS) means for each treatment group and the LS mean difference between treatment groups will be presented from the models together with standard errors (SE), two-sided p-values, and two-sided 95% CIs.

For the primary analysis, the primary efficacy endpoint of inferior corneal fluorescein staining will have missing data imputed using a control-based pattern mixture model under the assumption of a missing not at random mechanism. The 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 be run. The SAS® code for obtaining multiple pattern mixture model imputation data is:

```
PROC MI DATA = INDATA SEED = 6849548 OUT = MDATA NIMPUTE = 100
      MINIMUM = 0 MAXIMUM = 4 ROUND = 0.5;
  BY TREATMENT;
  MCMC IMPUTE=MONOTONE;
  VAR BASELINE IFS4 IFS5;
RUN;

PROC MI DATA = MDATA SEED = 3541655 OUT = OUTDATA NIMPUTE = 1
```

```
        MINIMUM = . 0 0 0 MAXIMUM = . 4 4 4 ROUND = . 0.5 0.5 0.5;
    BY _Imputation_;
    CLASS TREATMENT;

    MONOTONE REG(IFS4 = BASELINE/ DETAILS);
    MONOTONE REG(IFS5 = BASELINE IFS4/ DETAILS);
    MNAR MODEL(IFS4 IFS5 / MODELOBS=(TREATMENT='Vehicle'));
    VAR BASELINE IFS4 IFS5;
RUN;
```

where

- *INDATA* is the name of the input dataset
- *MDATA* is the name an intermediary dataset with a monotone missing pattern
- *OUTDATA* is the name of the output dataset
- *TREATMENT* is the name of the treatment group variable
- *BASELINE* is the baseline inferior corneal fluorescein staining
- *IFS4-IFS5* are the inferior corneal fluorescein staining at Visits 4 (Day 15) through Visit 5 (Day 29)

If multiple pattern mixture model imputation fails to generate imputations, baseline total sum fluorescein staining score will be added to the `VAR` statement and multiple imputation will be reattempted. If multiple imputation still fails to generate the required imputation, then MINIMUM, MAXIMUM, and ROUND statements will be removed.

After the imputed data sets are obtained, the following code will be used to execute the ANCOVA model on each imputed data set and the results combined from the analyses:

```
PROC MIXED DATA = OUTDATA;
  BY _IMPUTATION_;
  CLASS TREATMENT SITE;
  MODEL CHG = BASELINE TREATMENT SITE/ SOLUTION COVB;
  LSMEANS TREATMENT / CL PDIFF;
  ODS OUTPUT LSMEANS = OUTLS DIFFS = OUTDIFFS;
RUN;
PROC SORT DATA=OUTLS; BY TREATMENT _IMPUTATION_; RUN;
PROC MIANALYZE DATA=OUTLS;
  BY TREATMENT;
  MODELEFFECTS ESTIMATE;
  STDERR STDERR;
RUN;

DATA OUTDIFFS;
  SET OUTDIFFS;
  COMPARISON = TREATMENT||' - '||LEFT(_TREATMENT);
```

```
RUN;
PROC SORT DATA=OUTDIFFS; BY COMPARISON _IMPUTATION_; RUN;
PROC MIANALYZE DATA=OUTDIFFS;
  BY COMPARISON;
  MODELEFFECTS ESTIMATE;
  STDERR STDERR;
RUN;
```

where

- *TREATMENT* is the name of the treatment group variable
- *BASELINE* is the baseline inferior corneal fluorescein staining
- *CHG* is the change from baseline of the inferior corneal fluorescein staining at Visit 5 (Day 29) – *BASELINE*
- *SITE is the site ID*
- *OUTLS* is the name of the output dataset that contains the statistical results for the treatment mean from the ANCOVA model that is run on each of the 100 imputation datasets
- *OUTDIFFS* is the name of the output dataset that contains the statistical results for the difference in treatment mean from the ANCOVA model that is run on each of the 100 imputation datasets

Two sample t-tests will use similar SAS code to execute pattern mixture model multiple imputation analysis.

Sensitivity analyses will also include using ITT population with observed data only and ITT population with last observation carried forward (LOCF).

Markov chain Monte Carlo (MCMC) imputation with the ITT population will be used for sensitivity analysis under the assumption of a missing at random mechanism. The SAS® procedure `PROC MI` will be used for the MCMC imputation of 100 complete data sets. Then, imputation will proceed with The SAS code for obtaining the imputed data is:

```
PROC MI DATA = INDATA SEED = 2568716 OUT = OUTDATA NIMPUTE = 100
      MINIMUM = 0 MAXIMUM = 4 ROUND = 0.5;
  BY TREATMENT;
  MCMC INITIAL = EM;
  VAR BASELINE IFS4 IFS5;
RUN;
```

where

- *INDATA* is the name of the input dataset
- *OUTDATA* is the name of the output dataset
- *TREATMENT* is the name of the treatment group variable

- *BASELINE* is the baseline inferior corneal fluorescein staining
- *IFS4-IFS5* are the inferior corneal fluorescein staining at Visits 4 (Day 15) through 5 (Day 29)

If MCMC multiple imputation fails to generate imputations, baseline total eye fluorescein staining score will be added to the `VAR` statement and multiple imputation will be reattempted. If MCMC multiple imputation still fail to generate the required imputation, then `MINIMUM`, `MAXIMUM`, and `ROUND` statements will be removed.

Tipping point analyses will be included for further sensitivity analyses. The SAS® procedure `PROC MI` will be used to impute non-monotone missing data using MCMC to obtain 100 complete datasets with a monotone missing pattern as described for the control-based pattern mixture model imputation. Then, the monotone datasets will be imputed using a set of shift values for the active treatment group. The set of shift values for the imputation will be from 0 to 4 by 0.1 increments. The SAS® code for obtaining the imputed data is:

```
PROC MI DATA = INDATA SEED = 935769  OUT = MDATA NIMPUTE = 100
      MINIMUM = 0 MAXIMUM = 4 ROUND = 0.5;
      BY TREATMENT;
  MCMC IMPUTE=MONOTONE;
  VAR BASELINE IFS4 IFS5;
RUN;
PROC MI DATA = MDATA SEED = 427586 OUT = OUTDATA NIMPUTE = 1
      MINMAXITER = 1000000 MINIMUM = . 0 0 0 MAXIMUM = . 4 4 4
      ROUND = . 0.5 0.5 0.5;
 BY _Imputation_;
 CLASS TREATMENT;
 MONOTONE REG (IFS5 = BASELINE IFS4/DETAILS);
 MONOTONE REG (IFS4 = BASELINE /DETAILS);
 MNAR ADJUST (IFS5/ SHIFT = X ADJUSTOBS = (TREATMENT = 'ACTIVE
TREATMENT'));
 VAR BASELINE IFS4 IFS5;
RUN;
```

where

- *INDATA* is the name of the input dataset
- *MDATA* is the name an intermediary dataset with a monotone missing pattern
- *OUTDATA* is the name of the output dataset

- *TREATMENT* is the name of the treatment group variable
- *BASELINE* is the baseline inferior corneal fluorescein staining
- *IFS4-IFS5* are the inferior corneal fluorescein staining at Visits 4 (Day 15) through Visit 5 (Day 29)
- X is the shift parameter within the set of shift parameters [0 to 4 by 0.1]

Least squares mean differences, SEs, two-sided CIs and p-values will be reported for each shift value.

Further, primary efficacy analyses will be performed with the PP population with observed data only for sensitivity analysis. The following SAS® code is an example of the ANCOVA model for observed data, LOCF, and PP analyses:

```
PROC MIXED;
      CLASS TREATMENT SITE;
      MODEL CHG = BASELINE TREATMENT SITE / SOLUTION COVB;
      LSMEANS TREATMENT / CL PDIFF;
RUN;
```

Two sample t-tests will be conducted as sensitivity analyses using pattern mixture model imputation with the ITT population, observed data only with the ITT population, LOCF imputation with the ITT population, MCMC imputation with the ITT population, and observed data only with the PP population. Wilcoxon rank sum tests will be conducted as sensitivity analyses using observed data only with the ITT population, LOCF imputation with the ITT population, and observed data only with the PP population.

LS Means for inferior corneal fluorescein staining changes from baseline pre-CAE® and post-CAE® at Visit 5 (Day 29) will be displayed graphically in a bar chart with 95% CI bars by treatment group based on the ANCOVA analysis of the pattern mixture model imputation with the ITT population.

13.1.2 Ocular Discomfort Scale at Visit 5 (Day 29) Post-CAE, [redacted] K-161 [redacted] vs Vehicle [redacted]

Ocular discomfort scores will be subjectively graded by the subjects using the Ora Calibra® Ocular Discomfort Scale at all visits at pre-CAE® and post-CAE® for each eye separately. The Ocular Discomfort Scale ranges from 0 to 4 where 0 = No Discomfort, 1 = Intermittent Awareness, 2 = Constant Awareness, 3 = Intermittent Discomfort, and 4 = Constant Discomfort. Analyses will only be produced for the study eye.

Ocular discomfort will be summarized at Visit 5 (Day 29) Post-CAE® by treatment groups using continuous descriptive statistics. Primary analysis will use an ANCOVA model adjusted for baseline ocular discomfort, site, and with treatment group as the explanatory variable. Changes from baseline for each treatment group in ocular discomfort will be compared between [redacted] K-161 [redacted] and vehicle [redacted] In addition, treatment by baseline and treatment by study site interactions will be explored in separate models to evaluate how the treatment effect may differ for baseline value and study sites. In the case of a significant interaction at the 0.05 level, analyses will be performed by baseline stratification (≤3.0 and >3.0) and/or site to understand how the treatment effect differs for baseline value and study sites. Least squares means for each treatment

group and the LS mean difference between treatment groups will be presented from the models together with SE, two-sided p-values, and two-sided 95% CIs.

The primary analysis will use control-based pattern mixture model imputation on the ITT population at Visit 5 (Day 29). SAS® code for the multiple imputation analysis will resemble the code in Section 13.1.1 for inferior corneal fluorescein staining. The random number seeds for the monotone imputation and pattern mixture model imputation described in Section 13.1.1 will be `SEED = 354846` and `SEED = 956546`, respectively. In addition, the rounding parameter will be `ROUND = 1`.

Should control-based pattern mixture model imputation fail to produce the required imputations, baseline measurements of overall ocular discomfort, burning, dryness, grittiness, and stinging from the Ora Calibra® Ocular Discomfort & 4-Symptom Questionnaire will added to the `VAR` statement in an iterative fashion until imputation is successful. If multiple imputation still fail to generate the required imputations, then `MINIMUM`, `MAXIMUM`, and `ROUND` statements will be removed.

Sensitivity analyses will be performed using on the ITT population using observed data only and ITT population using LOCF.

Sensitivity analyses under the assumption of a missing at random mechanism will be performed using MCMC imputation on the ITT population as described in section 13.1.1 except the random number seed for `PROC MI` will be `SEED = 146645` and the rounding parameter will be `ROUND = 1`.

Should MCMC imputation fail to produce the required imputations, baseline measurements of overall ocular discomfort, burning, dryness, grittiness, and stinging from the Ora Calibra® Ocular Discomfort & 4-Symptom Questionnaire will added to the `VAR` statement in an iterative fashion until imputation is successful. If MCMC multiple imputation still fail to generate the required imputation, then `MINIMUM`, `MAXIMUM`, and `ROUND` statements will be removed.

Tipping point analysis will be included for further sensitivity analyses in the manner described in section 13.1.1. Then, the monotone datasets will be imputed using a set of shift values for the active treatment group. The random number seeds for the monotone imputation and pattern mixture model imputation described in Section 13.1.1 will be `SEED = 5324652` and `SEED = 236135`, respectively. The set of shift values for the imputation will be from 0 to 4.0 by 0.1 increments. The `MINIMUM`, `MAXIMUM`, and `ROUND` statements will be removed.

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

Two sample t-tests will be conducted as sensitivity analyses using pattern mixture model imputation with the ITT population, observed data only with the ITT population, LOCF imputation with the ITT population, MCMC imputation with the ITT population, and observed data only with the PP population. Wilcoxon rank

sum tests will be conducted as sensitivity analyses using observed data only with the ITT population, LOCF imputation with the ITT population, and observed data only with the PP population.

Least squares means for Ocular Discomfort Score changes from baseline pre-CAE® and post-CAE® at Visit 5 (Day 29) will be displayed graphically in a bar chart with 95% CI bars by treatment group based on the ANCOVA analysis of the multiple pattern mixture model imputation of the ITT population.

13.2 Important Secondary Analyses

13.2.1 Unanesthetized Schirmer's Test at Visit 5 (Day 29), [redacted] K-161 [redacted] vs Vehicle [redacted]

Unanesthetized Schirmer's test will be conducted at all visits at pre-CAE®. The Schirmer's test strip will be placed in the lower temporal lid margin of each eye. After 5 minutes, the test strip will be removed and the length of the moistened area will be recorded in mm for each eye. Lower values indicate less tears produced in the eye. Analyses will only be produced for study eye.

Unanesthetized Schirmer's test will be summarized at Visit 5 (Day 29) Pre-CAE® by treatment groups using continuous descriptive statistics. Primary analysis will use an ANCOVA model adjusted for baseline unanesthetized Schirmer's test, site and with treatment group as the explanatory variable. Changes from baseline for each treatment group in ocular discomfort will be compared between [redacted] K-161 [redacted] and vehicle [redacted] In addition, treatment by baseline and treatment by study site interactions will be explored in separate models to evaluate how the treatment effect may differ for baseline value and study sites. In the case of a significant interaction at the 0.05 level, analyses will be performed by baseline median stratification and/or site to understand how the treatment effect differs for baseline value and study sites. Least squares means for each treatment group and the LS mean difference between treatment groups will be presented from the models together with SE, two-sided p-values, and two-sided 95% CIs.

The primary analysis will use control-based pattern mixture model imputation with the ITT population. The random number seeds for the monotone imputation and pattern mixture model imputation described in Section 13.1.1 will be `SEED = 465485` and `SEED = 849159`, respectively. The `MINIMUM`, `MAXIMUM`, and `ROUND` statements will be removed.

Sensitivity analyses will be performed using on the ITT population using observed data only and ITT population using LOCF.

Sensitivity analyses under the assumption of a missing at random mechanism will be performed using MCMC imputation on the ITT population as described in section 13.1.1 except the random number seed for `PROC MI` will be `SEED = 884564` and the `MINIMUM`, `MAXIMUM`, and `ROUND` statements will be removed.

Tipping point analysis will be included for further sensitivity analyses in the manner described in section 13.1.1 on the ITT population. The random number seeds for the monotone imputation and pattern mixture model imputation described in Section 13.1.1 will be `SEED = 231219` and `SEED = 981982`, respectively.

The set of shift values for the imputation will be from 0 to -5 by an increment of -0.25. The `MAXIMUM`, `MINIMUM` and `ROUND` statements will be removed.

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

Two sample t-tests will be conducted as sensitivity analyses using pattern mixture model imputation with the ITT population, observed data only with the ITT population, LOCF imputation with the ITT population, MCMC imputation with the ITT population, and observed data only with the PP population. Wilcoxon rank sum tests will be conducted as sensitivity analyses using observed data only with the ITT population, LOCF imputation with the ITT population, and observed data only with the PP population.

Sensitivity analyses will be performed on the ITT and PP populations using observed data only, and on the ITT population using LOCF methodology.

Least squares means for unanesthetized Schirmer's test changes from baseline at Visit 5 (Day 29) will be displayed graphically in a bar chart with 95% CI bars by treatment group based on the ANCOVA analysis of the multiple pattern mixture model imputation of the ITT population.

13.2.2 Tear Film Break Up Time

The TFBUT will be recorded at all visits at pre-CAE® and post-CAE®. For each eye, 2 measurements will be taken and averaged unless the 2 measurements are >2 seconds apart and are each <10 seconds, in which case, a third measurement would be taken and the 2 closest of the 3 would be averaged. This average will then be used for analyses. Analyses will be provided for study eye only.

Tear Film Break Up Time will be summarized at Visit 5 (Day 29) Post-CAE® by treatment groups using continuous descriptive statistics. Primary analysis will use an ANCOVA model adjusted for baseline TFBUT, site and with treatment group as the explanatory variable. Changes from baseline for each treatment group in TFBUT will be compared between [redacted] K-161 [redacted] and vehicle [redacted] In addition, treatment by baseline and treatment by study site interactions will be explored in separate models to evaluate how the treatment effect may differ for baseline value and study sites. In the case of a significant interaction at the 0.05 level, analyses will be performed by baseline median stratification and/or site to understand how the treatment effect differs for baseline value and study sites. Least squares means for each treatment group and the LS mean difference between treatment groups will be presented from the models together with standard errors SE, two-sided p-values, and two-sided 95% CIs.

The primary analysis will use control-based pattern mixture model imputation with the ITT population. The random number seeds for the monotone imputation and pattern mixture model imputation described in Section 13.1.1 will be `SEED = 5466765` and `SEED = 903573`, respectively. The `MINIMUM`, `MAXIMUM`, and `ROUND` statements will be removed.

Sensitivity analyses will be performed using on the ITT population using observed data only and ITT population using LOCF.

Tipping point analysis will be included for further sensitivity analyses in the manner described in section 13.1.1. Then, the monotone datasets will be imputed using a set of shift values for the active treatment group. The set of shift values for the imputation will be from 0 to -2 by an increment of -0.1. The random number seeds for the monotone imputation and pattern mixture model imputation described in Section 13.1.1 will be `SEED = 392187` and `SEED = 748451`, respectively. The `MAXIMUM, MINIMUM` and `ROUND` statements will be removed.

Sensitivity analyses under the assumption of a missing at random mechanism will be performed using MCMC imputation on the ITT population as described in section 13.1.1 except the random number seed for `PROC MI` will be `SEED = 234993` and the `MINIMUM`, `MAXIMUM`, and `ROUND` statements will be removed.

Two sample t-tests will be conducted as sensitivity analyses using pattern mixture model imputation with the ITT population, observed data only with the ITT population, LOCF imputation with the ITT population, MCMC imputation with the ITT population, and observed data only with the PP population. Wilcoxon rank sum tests will be conducted as sensitivity analyses using observed data only with the ITT population, LOCF imputation with the ITT population, and observed data only with the PP population.

Least squares means for TFBUT changes from baseline pre-CAE® and post-CAE® at Visit 5 (Day 29) will be displayed graphically in a bar chart with 95% CI bars by treatment group based on the ANCOVA analysis of the multiple pattern mixture model imputation of the ITT population.

13.3 Secondary Analyses

Secondary efficacy variables will be summarized using continuous descriptive statistics (n, mean, SD, median, minimum, and maximum) by visit and treatment group. Two-sample t-tests will be used to analyze the efficacy variables for visit-based data between treatment groups.

Change from baseline in secondary efficacy variables will be summarized by visit and treatment group using continuous descriptive statistics. Change from baseline will be analyzed using an ANCOVA model adjusting for baseline value and site, and with treatment group as the explanatory variable. Two-sample t-tests will be used for sensitivity analysis. Paired t-tests will be used to assess change from baseline within each treatment group.

All secondary efficacy analyses will be performed on the ITT population with control-based multiple pattern mixture model imputation. All secondary endpoints will be described under subject listings.

The following secondary efficacy variables will be tested:

- Fluorescein staining by region: central, superior, inferior, temporal, nasal, and corneal sum as assessed by the Ora Calibra® Corneal and Conjunctival Staining Scale;

- Lissamine green staining by region: central, superior, inferior, temporal, nasal, and conjunctival sum as assessed by the Ora Calibra® Corneal and Conjunctival Staining Scale
- Conjunctival Redness as assessed by the Ora Calibra® Scale
- TFBUT
- Tear Osmolarity
- Unanesthetized Schirmer's Test
- Blink Rate
- Ora Calibra® Ocular Discomfort Scale
- Ora Calibra® Ocular Discomfort & 4-Symptom Questionnaire
- Visual Analog Scale: Burning/Stinging, Itching, Foreign Body Sensation, Blurred Vision, Eye Dryness, Photophobia, and Pain
- Ocular Surface Disease Index

13.3.1 FLUORESCEIN STAINING (ORA CALIBRA® SCALE)

Fluorescein staining will be conducted at all visits at pre-CAE® and post-CAE®. Grading will be conducted using the Ora Calibra® Scale from 0 to 4 with the use of half grade (0.5) increments, where grade 0 = None and 4 = Severe. The regions to be assessed will be the central, superior, inferior, temporal, corneal sum (central, superior, and inferior regions), conjunctival sum (temporal and nasal regions), and total eye score for study eye only.

Visit-based data will be summarized using continuous descriptive statistics for each region, sum score, and time point (pre-CAE®, post-CAE®, and change from pre- to post-CAE®). Change from baseline will also be summarized using continuous descriptive statistics.

Two-sample t-tests will be used to analyze differences between treatment groups 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.

Change from baseline will be analyzed using ANCOVA with adjustment for baseline fluorescein staining score and site, and treatment group as the explanatory variable. 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 ANCOVA model. Two-sample t-tests will be used as sensitivity analysis. Paired t-tests will be used to analyze change from baseline within each treatment group.

Analyses will use ITT population with control-based pattern mixture model imputation as described in section 13.1.1. The random number seeds for the monotone imputation and pattern mixture model imputation will be `SEED = 879921` and `SEED = 562265`, respectively.

A subject level listing of fluorescein staining will also be produced.

13.3.2 LISSAMINE GREEN STAINING (ORA CALIBRA® SCALE)

Subjects will undergo lissamine green staining at all visits at pre-CAE® and post-CAE®. The grading in the inferior, superior, central, temporal, nasal, corneal sum (central, superior, and inferior regions), conjunctival sum (temporal and nasal regions), and total eye sum regions will be measured by the Ora Calibra® Corneal and Conjunctival Staining Scale for lissamine green staining. A standardized grading system of 0 to 4 is used for each of the 5 regions with 0 = No Staining and 4 = Confluent Staining. Half (0.5) grade increments may be used. Analyses will be provided for study eye only.

Visit-based data will be summarized using continuous descriptive statistics by treatment group for each region, sum score, and timepoint (pre-CAE®, post-CAE®, and change from pre- to post-CAE®). Change from baseline will also be summarized with continuous descriptive statistics.

Two-sample t-tests will be used to analyze differences between treatment groups 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.

Change from baseline will be analyzed using ANCOVA adjusted for baseline lissamine green staining score and site, and treatment group as the explanatory variable. 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 ANCOVA model. Two-sample t-tests will be used as sensitivity analysis. Paired t-tests will be used to analyze change from baseline within each treatment group.

Analyses will use ITT population with control-based pattern mixture model imputation as described in section 13.1.1. The random number seeds for the monotone imputation and pattern mixture model imputation will be `SEED = 965464` and `SEED = 331546`, respectively.

A subject level listing of lissamine green staining will also be produced.

13.3.3 CONJUNCTIVAL REDNESS

The Ora Calibra® Conjunctival Redness Scale for Dry Eye will be performed at all visits at pre-CAE® and post-CAE®. The conjunctival redness scale ranges from 0 to 4 (half increments may be used) where 0 = Normal, without Vasodilation; 1 = Trace Ciliary or Conjunctival Vasodilation; 2 = Broad Ciliary Vasodilation, 3 = Broad Ciliary and Slight, Horizontal Conjunctival Vasodilation; and 4 = Broad Ciliary and Prominent, Horizontal Conjunctival Vasodilation. Analyses will be provided for study eye only.

Visit-based data will be summarized using continuous descriptive statistics by treatment group for each time point (pre-CAE®, post-CAE®, and change from pre-CAE® to post-CAE®). Change from baseline will also be summarized with continuous descriptive statistics.

Two-sample t-tests will be used to analyze differences between treatment groups 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.

Change from baseline will be analyzed using ANCOVA adjusted for baseline redness score and site, and treatment group as the explanatory variable. 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 ANCOVA model. Two-sample t-tests will be used as sensitivity analysis. Paired t-tests will be used to analyze change from baseline within each treatment group.

Analyses will use ITT population with control-based pattern mixture model imputation as described in section 13.1.1. The random number seeds for the monotone imputation and pattern mixture model imputation will be `SEED = 465526` and `SEED = 579516`, respectively.

A subject level listing of conjunctival redness will also be produced.

13.3.4 TEAR FILM BREAK-UP TIME

The TFBUT will be recorded at all visits at pre-CAE® and post-CAE®. For each eye, 2 measurements will be taken and averaged unless the 2 measurements are >2 seconds apart and are each <10 seconds, in which case, a third measurement would be taken and the 2 closest of the 3 would be averaged. This average will then be used for analyses. Analyses will be provided for study eye only.

Visit-based data will be summarized using continuous descriptive statistics by treatment group at each time point (pre-CAE®, post-CAE®, change from pre- to post-CAE®). Change from baseline will also be summarized with continuous descriptive statistics at each post-baseline visit.

Two-sample t-tests will be used to analyze differences between treatment groups 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.

Change from baseline will be analyzed using ANCOVA adjusted for baseline TFBUT and site, and treatment group as the explanatory variable. 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 ANCOVA model. Two-sample t-tests will be used as sensitivity analysis. Paired t-tests will be used to analyze change from baseline within each treatment group.

Analyses will use ITT population with control-based pattern mixture model imputation as described in section 13.1.1. The random number seeds for the monotone imputation and pattern mixture model imputation will be `SEED = 897985` and `SEED = 381325`, respectively.

A subject level listing of TFBUT will also be produced.

13.3.5 TEAR OSMOLARITY

Tear osmolarity will be measured at Visit 2 (Day 1), Visit 3 (Day 8), Visit 4 (Day 15), and Visit 5 (Day 29) at pre-CAE®. 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.

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

Two-sample t-tests will be used to analyze differences between treatment groups 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.

Change from baseline will be analyzed using ANCOVA adjusted for baseline tear osmolarity and site, and treatment group as the explanatory variable. 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 ANCOVA model. Two-sample t-tests will be used as sensitivity analysis. Paired t-tests will be used to analyze change from baseline within each treatment group.

Analyses will use ITT population with control-based pattern mixture model imputation as described in section 13.1.1. The random number seeds for the monotone imputation and pattern mixture model imputation will be `SEED = 542354` and `SEED = 8976544`, respectively.

A subject level listing of tear osmolarity will also be produced.

13.3.6 UNANESTHETIZED SCHIRMER'S TEST

Unanesthetized Schirmer's test will be conducted at all visits at pre-CAE®. The Schirmer's test strip will be placed in the lower temporal lid margin of each eye. After 5 minutes, the test strip will be removed and the length of the moistened area will be recorded in mm for each eye. Lower values indicate less tears produced in the eye. Analyses will only be produced for study eye.

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

Two-sample t-tests will be used to analyze differences between treatment groups 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.

Change from baseline will be analyzed using an ANCOVA model with adjustment for baseline unanesthetized Schirmer's test reading and site, and treatment group as the explanatory variable. 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 ANCOVA model. Two-sample t-tests will be used as sensitivity analysis. The 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 group.

Analyses will use ITT population with control-based pattern mixture model imputation as described in section 13.1.1. The random number seeds for the monotone imputation and pattern mixture model imputation will be `SEED = 6342576` and `SEED = 765121`, respectively.

A subject level listing of unanesthetized Schirmer's test will also be produced.

13.3.7 BLINK RATE

Blink rate will be recorded at Visit 2 (Day 1), Visit 3 (Day 8), Visit 4 (Day 15), and Visit 5 (Day 29) at pre-CAE®. Analyses will only be produced for study eye.

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

Two-sample t-tests will be used to analyze differences between treatment groups 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.

Change from baseline will be analyzed using an ANCOVA model with adjustment for baseline blink rate and site, and treatment group as the explanatory variable. 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 ANCOVA model. Two-sample t-tests will be used as sensitivity analysis. The 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 group.

Analyses will use ITT population with control-based pattern mixture model imputation as described in section 13.1.1. The random number seeds for the monotone imputation and pattern mixture model imputation will be `SEED = 6754416` and `SEED = 234123`, respectively.

A subject level listing of blink rate will also be produced.

13.3.8 OCULAR DISCOMFORT SCALE (ORA CALIBRA®)

Ocular discomfort scores will be subjectively graded by the subjects using the Ora Calibra® Ocular Discomfort Scale at all visits at pre-CAE® and post-CAE®. The Ocular Discomfort Scale ranges from 0 to 4 where 0 = No Discomfort, 1 = Intermittent Awareness, 2 = Constant Awareness, 3 = Intermittent Discomfort, and 4 = Constant Discomfort. Analyses will only be produced for study eye.

Visit-based data will be summarized using continuous descriptive statistics for each symptom by treatment group at each time point (pre-CAE® and post-CAE®). Change from baseline will also be summarized with continuous descriptive statistics.

Two-sample t-tests will be used to analyze differences between treatment groups 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.

Change from baseline will be analyzed using ANCOVA adjusted for baseline discomfort score and site, and treatment group as the explanatory variable. 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 ANCOVA model. Two-sample t-tests will be used as sensitivity analysis. Paired t-tests will be used to analyze change from baseline within each treatment group.

Analyses will use ITT population with control-based pattern mixture model imputation as described in section 13.1.1. The random number seeds for the monotone imputation and pattern mixture model imputation will be `SEED = 8675478` and `SEED = 523426`, respectively.

A subject-level listing of ocular discomfort scores will also be produced.

13.3.8.1 Ocular Discomfort Scale (ORA Calibra®) During CAE®

Ocular discomfort scores will be assessed every 5 minutes during the CAE® exposure at Visit 1 (Day -14), Visit 2 (Day 1), Visit 4 (Day 15), and Visit 5 (Day 29). Assessments at 0 minutes, 90 minutes, change from 0 to 90 minutes, change from 0 to last timepoint and changes from baseline will be summarized by treatment group, visit, and timepoint using continuous descriptive statistics, including two-sided 95% CIs. Last timepoint is described as last recorded discomfort score within a CAE® exposure.

For the ocular discomfort score at 0 minutes, at 90 minutes, change from 0 to 90 minutes, and change from 0 to last time point, two-sample t-tests will be employed to compare treatment and Vehicle means at each post-baseline visit and time point. The differences in means, two-sided 95% CIs for the difference in means, and p-values will be reported.

Changes from baseline will be compared between treatment groups using ANCOVA models that adjust for baseline and site. The LS means, LS mean differences, SEs, two-sided 95% CIs for the difference in means, and two-sided p-values will be reported. Two sample t-tests will also be conducted. Within each treatment group, paired t-tests will be conducted to compare change from baseline.

LS means from the ANCOVA will be displayed graphically in a line graph with standard error bars by treatment group.

Analyses will use ITT population with control-based pattern mixture model imputation as described in section 13.1.1. The random number seeds for the monotone imputation and pattern mixture model imputation will be `SEED = 728436` and `SEED = 543685`, respectively.

A subject listing of ocular discomfort scores during CAE® will be provided.

13.3.8.2 Ocular Discomfort Scale (ORA Calibra®) During CAE® Mixed Model Repeated Measures

Ocular discomfort scores will be assessed every 5 minutes during the CAE® exposure at Visit 1 (Day -14), Visit 2 (Day 1), Visit 4 (Day 15), and Visit 5 (Day 29). Assessments at 0 minutes, 90 minutes, change from

0 to 90 minutes and changes from baseline will be summarized by treatment group, visit, and timepoint using continuous descriptive statistics, including two-sided 95% CIs.

Change from baseline will be assessed using a mixed-effect model for repeated measures (MMRM) to compare active treatment and vehicle at Visit 4 (Day 15), and Visit 5 (Day 29) while accounting for the correlations among the repeated measurements during the CAE®. The model will include treatment group; time (nominal time will be used); treatment by time interaction, and timepoint matched baseline ocular discomfort score as fixed effects; and subject as the random effect. The following SAS® code will be used:

```
PROC MIXED DATA = INDATA METHOD=REML;
       CLASS SUBJID TREATMENT TIME SITE;
       MODEL CHG_ODS = ODS_BASE SITE TREATMENT | TIME
              / SOLUTION COVB DDFM=KR;
       REPEATED TIME / TYPE = UN SUBJECT = SUBJID;
       LSMEANS TREATMENT TREATMENT*TIME/ CL PDIFF;
       ODS OUTPUT LSMEANS = OUTLS DIFFS = OUTDIFFS;
RUN;
```

where

- *SUBJID* is the subject ID
- *SITE* is the Site ID
- *TREATMENT* is the name of the treatment group variable
- *TIME* is the nominal time of the measurement
- *CHG_ODS* is the change from baseline ocular discomfort score
- *ODS_BASE* is the timepoint matched baseline ocular discomfort score in the study eye
- *OUTLS* is the name of the output dataset that contains the statistical results for the treatment means from the MMRM
- *OUTDIFFS* is the name of the output dataset that contains the statistical results for the differences in treatment means from the MMRM

If the MMRM does not converge with an unstructured (TYPE = UN) covariance matrix, Toeplitz (TYPE = TOEP) and compound symmetry (TYPE = CS) structures will be utilized in order until convergence is achieved. If the MMRM with compound symmetry does not converge, MMRM statistical inferences will represented as not calculable.

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 MMRM.

Analyses will use ITT population with observed data only.

13.3.9 Ocular Discomfort & 4-Symptom Questionnaire (ORA Calibra® Scale)

Subjects will rate the severity of each of the following symptoms (for both eyes [OU]) with regard to how both their eyes feel in general: overall ocular discomfort, burning, dryness, grittiness, and stinging according to the following 6-point (0 to 5) scale where 0 = None and 5 = Worst. Questionnaires will be recorded at all visits at pre-CAE® and post-CAE®.

Visit-based data will be summarized using continuous descriptive statistics for each symptom by treatment group at each time point pre-CAE® and post-CAE®. Change from baseline will also be summarized with continuous descriptive statistics.

Two-sample t-tests will be used to analyze differences between treatment groups 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.

Change from baseline will be analyzed using ANCOVA adjusted for baseline symptom score and site, and treatment group as the explanatory variable. 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 ANCOVA model. Two-sample t-tests will be used as sensitivity analysis. Paired t-tests will be used to analyze change from baseline within each treatment group.

Analyses will use ITT population with control-based pattern mixture model imputation as described in section 13.1.1. The random number seeds for the monotone imputation and pattern mixture model imputation will be `SEED = 3658465` and `SEED = 9631723`, respectively.

A subject-level listing of symptom scores will also be produced.

13.3.10 Visual Analog Scale

Ocular symptoms of the VAS will recorded at all visits at pre-CAE® and post-CAE®. Subject will be asked to subjectively rate ocular symptoms (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.” The ocular symptoms used for analysis are Burning/Stinging, Itching, Foreign Body Sensation, Blurred Vision, Eye Dryness, Photophobia, and Pain.

Visit-based data will be summarized using continuous descriptive statistics for each symptom by treatment group for each time point (pre-CAE® and post-CAE®). Change from baseline will also be summarized with continuous descriptive statistics.

Two-sample t-tests will be used to analyze differences between treatment groups 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.

Change from baseline will be analyzed using ANCOVA adjusted for baseline ocular symptom score and site, and treatment group as the explanatory variable. 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 ANCOVA. Two-sample t-tests will be used as sensitivity analysis. Paired t-tests will be used to analyze change from baseline within each treatment group.

Analyses will use ITT population with control-based pattern mixture model imputation as described in section 13.1.1. The random number seeds for the monotone imputation and pattern mixture model imputation will be `SEED = 432467` and `SEED = 9434382`, respectively.

A subject-level listing of VAS will also be produced.

13.3.11 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:

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

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 at all visits at pre-CAE®.

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

Two-sample t-tests will be used to analyze differences between treatment groups 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.

Change from baseline will be analyzed using ANCOVA adjusted for baseline OSDI score and site, and treatment group as the explanatory variable. 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 ANCOVA. Two-sample t-tests will be used as sensitivity analysis. Paired t-tests will be used to analyze change from baseline within each treatment group.

Analyses will use ITT population with control-based pattern mixture model imputation as described in section 13.1.1. The random number seeds for the monotone imputation and pattern mixture model imputation will be `SEED = 354715` and `SEED = 863267`, respectively.

A subject-level listing of OSDI will also be produced.

14. [redacted]

[redacted]

[redacted]

14.1.1 [redacted]

[redacted]

14.1.2 [redacted]

[redacted]

[redacted]

14.1.3 [redacted]

[redacted]

14.1.4 EFFICACY ENDPOINTS THAT ARE COMPARED AGAINST [redacted] K-161 FOUR TIMES DAILY ([redacted] (I.E., [redacted] K-161 TWICE DAILY [redacted] K-161 [redacted] AND VEHICLE [redacted]

These analyses will conducted as described in section 13 with secondary endpoints.

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. An AE can arise from any delivery, implantation, or use of a medical device, including medical device failure, subject characteristics that may impact medical device performance (e.g., anatomical limitations), and therapeutic parameters (e.g., energy applied, sizing, dose release, and anatomic fit) associated with medical device use.

All AEs spontaneously reported by the subject and/or in response to an open question from study personnel or revealed by observation, physical examination, or other diagnostic procedures will be recorded in the source document and on the appropriate pages of the CRF. Any clinically relevant deterioration in clinical finding is considered an AE and must be recorded. When possible, signs and symptoms indicating a common underlying pathology will be noted as one comprehensive event.

Documentation regarding the AE will include the nature, date of onset, end date, severity, relationship to IP, action(s) taken, seriousness, and outcome of any sign or symptom observed by the investigator or reported by the subject upon indirect questioning. Exacerbation of conditions related to the signs and symptoms of dry eye disease will not be reported as an AE. All AEs will be coded using MedDRA 22.0.

Treatment-emergent adverse events (TEAE) are defined as any event that occurs or worsens on or after the first dose of study treatment. 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.

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 study drug 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.

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

- Related: A reasonable possibility exists that the IP caused the AE. A related AE can be further defined as follows:
  - Occurs within a reasonable temporal sequence to administration of study drug
  - Cannot be explained by concurrent disease or other drugs or chemicals
  - Improves or disappears on stopping or reducing study drug (de-challenge)
  - Reappears on repeated exposure to study drug (re-challenge)
  - Is an unusual event that is known to be associated with the drug or this class of compound, and cannot be explained by other therapy or the participant's physical condition
  - Unlikely to be attributed to concurrent disease or other drugs or a clinically reasonable response on withdrawal (de-challenge)
- Not Related: A reasonable possibility does not exist that the IP caused the AE. A not related AE can be further defined as follows:
  - Occurs with a temporal relationship to administration of study drug which makes a causal relationship improbable
  - Other drugs, chemicals, or underlying disease provide plausible explanations of causality
  - Is known to be associated with the participant's clinical condition, or with other medication taken by the participant

The expectedness of an AE should be determined based upon existing safety information about the IP using these explanations:

- Unexpected: An AE that is not listed in the Investigator's Brochure (IB) or is not listed at the specificity or severity that has been observed.
- Expected: An AE that is listed in the IB at the specificity and severity that has been observed.
- Not applicable: An AE unrelated to the IP.

An overall summary will be presented that includes the number of events and the number and percentage of subjects who experienced at least one AE, ocular AE, and non-ocular AE by treatment group and over all subjects. This summary will also include breakdowns of TEAEs further categorized as ocular or non-ocular, TEAEs by severity, TEAEs by relationship to study drug, TEAEs causing premature treatment discontinuation, and TEAEs leading to death. TE-SAEs will also be categorized as ocular or non-ocular, TE-SAEs by severity, TE-SAEs by relationship to study drug, TE-SAEs causing premature treatment discontinuation, and TE-SAEs leading to death.

Summaries will be provided for the following categories of AEs:

- Ocular and non-ocular TEAEs by SOC and PT
- Ocular and non-ocular TEAEs by SOC, PT, and maximal severity
- Ocular and non-ocular treatment-related TEAEs by SOC and PT
- Ocular and non-ocular TEAEs by SOC, PT and visit.

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, SOCs will be listed in ascending alphabetical order; PTs will be listed in order of descending frequency for all subjects within each SOC.

All AEs will be presented in a subject listing. In addition, all SAEs will be presented in a separate listing.

15.2 [redacted]

15.3 [redacted]

15.4 [redacted]

15.5 [redacted]

[redacted]

15.6 [redacted]

[redacted]

15.7 [redacted]

[redacted]

- [redacted]
- [redacted]
- [redacted]

[redacted]

15.8 Maximum dose tolerability by comparing the [redacted] K-161 [redacted] arm to the [redacted] and [redacted] K-161 [redacted] arms

Safety analyses as described in Section 15 will include summarizations by treatment group. No additional analyses will be conducted to otherwise examine maximum dose tolerability.

16. Other Measures

16.1 [redacted]

[redacted]

16.2 [redacted]

[redacted]

16.3 [redacted]

[redacted]

16.4 [redacted]

[redacted]

17. [redacted]

18. **Changes from Protocol-Stated Analyses**

In the protocol, it was stated that analyses of exploratory endpoints will use ITT population with pattern mixture model imputation. Analyses for OCT and strip meniscometry will use ITT population with observed data only.

19. **Revision History**

Documentation of revision to the SAP will commence after approval of the final version 1.0.

20. **Tables**

Tables that will be included in the topline delivery are shown in boldface font.

Table Number	Title	Population
Table 14.1.1	**Subject Disposition**	**All Screened Subjects**
Table 14.1.2.1	**Demographic Characteristics**	**ITT Population**
Table 14.1.2.2	Demographic Characteristics	Safety Population
Table 14.1.3.1	**Baseline Disease Characteristics**	**ITT Population**
Table 14.1.3.2	Baseline Disease Characteristics	Safety Population
Table 14.1.4.1	Ocular Medical History	Safety Population
Table 14.1.4.2	Non-Ocular Medical History	Safety Population
Table 14.1.5.1	Ocular Concomitant Medications	ITT Population
Table 14.1.5.2	Non-Ocular Concomitant Medications	ITT Population
Table 14.2.1.1.1	**Inferior Corneal Fluorescein Staining (Ora Calibra Scale) [redacted] K-161 [redacted] vs Vehicle**	**ITT Population with PMM**
Table 14.2.1.1.2	Inferior Corneal Fluorescein Staining (Ora Calibra Scale) by Baseline Stratification and Site [redacted] K-161 [redacted] vs Vehicle	ITT Population PMM

Table 14.2.1.1.3	Inferior Corneal Fluorescein Staining (Ora Calibra Scale) by Baseline Stratification [redacted] K-161 [redacted] vs Vehicle	ITT Population with PMM
Table 14.2.1.1.4	Inferior Corneal Fluorescein Staining (Ora Calibra Scale) by Site [redacted] K-161 [redacted] vs Vehicle	ITT Population with PMM
Table 14.2.1.1.5	Inferior Corneal Fluorescein Staining (Ora Calibra Scale) [redacted] K-161 [redacted] vs Vehicle	ITT Population with Observed Data Only
Table 14.2.1.1.6	Inferior Corneal Fluorescein Staining (Ora Calibra Scale) [redacted] K-161 [redacted] vs Vehicle	ITT Population with LOCF
Table 14.2.1.1.7	Inferior Corneal Fluorescein Staining (Ora Calibra Scale) [redacted] K-161 [redacted] vs Vehicle	ITT Population with MCMC
Table 14.2.1.1.8	Tipping Point Analysis: Inferior Corneal Fluorescein Staining (Ora Calibra Scale) [redacted] K-161 [redacted] vs Vehicle	ITT Population with PMM
Table 14.2.1.1.9	Inferior Corneal Fluorescein Staining (Ora Calibra Scale) [redacted] K-161 [redacted] vs Vehicle	PP Population with Observed Data Only
Table 14.2.1.2.1	**Ocular Discomfort Scale (Ora Calibra Scale) [redacted] K-161 [redacted] vs Vehicle**	**ITT Population with PMM**
Table 14.2.1.2.2	Ocular Discomfort Scale (Ora Calibra Scale) by Baseline Median and Site [redacted] K-161 [redacted] vs Vehicle	ITT Population with PMM
Table 14.2.1.2.3	Ocular Discomfort Scale (Ora Calibra Scale) by Baseline Median [redacted] K-161 [redacted] vs Vehicle	ITT Population with PMM
Table 14.2.1.2.4	Ocular Discomfort Scale (Ora Calibra Scale) by Site [redacted] K-161 [redacted] vs Vehicle	ITT Population with PMM
Table 14.2.1.2.5	Ocular Discomfort Scale (Ora Calibra Scale) [redacted] K-161 [redacted] vs Vehicle	ITT Population with Observed Data Only
Table 14.2.1.2.6	Ocular Discomfort Scale (Ora Calibra Scale) [redacted] K-161 [redacted] vs Vehicle	ITT Population with LOCF
Table 14.2.1.2.7	Ocular Discomfort Scale (Ora Calibra Scale) [redacted] K-161 [redacted] vs Vehicle	ITT Population with MCMC
Table 14.2.1.2.8	Tipping Point Analysis: Ocular Discomfort Scale (Ora Calibra Scale) [redacted] K-161 [redacted] vs Vehicle	ITT Population with PMM
Table 14.2.1.2.9	Ocular Discomfort Scale (Ora Calibra Scale) [redacted] K-161 [redacted] vs Vehicle	PP Population with Observed Data Only
Table 14.2.2.1.1	**Unanesthetized Schirmer's Test (mm) [redacted] K-161 [redacted] vs Vehicle**	**ITT Population with PMM**
Table 14.2.2.1.2	Unanesthetized Schirmer's Test (mm) by Baseline Median and Site [redacted] K-161 [redacted] vs Vehicle	ITT Population with PMM
Table 14.2.2.1.3	Unanesthetized Schirmer's Test (mm) by Baseline Median [redacted] K-161 [redacted] vs Vehicle	ITT Population with PMM
Table 14.2.2.1.4	Unanesthetized Schirmer's Test (mm) by Site [redacted] K-161 [redacted] vs Vehicle	ITT Population with PMM
Table 14.2.2.1.5	Unanesthetized Schirmer's Test (mm) [redacted] K-161 [redacted] vs Vehicle	ITT Population with Observed Data Only
Table 14.2.2.1.6	Unanesthetized Schirmer's Test (mm) [redacted] K-161 [redacted] vs Vehicle	ITT Population with LOCF
Table 14.2.2.1.7	Unanesthetized Schirmer's Test (mm) [redacted] K-161 [redacted] vs Vehicle	ITT Population with MCMC
Table 14.2.2.1.8	Tipping Point Analysis: Unanesthetized Schirmer's Test (mm) [redacted] K-161 [redacted] vs Vehicle	ITT Population with PMM

Table 14.2.2.1.9	Unanesthetized Schirmer's Test (mm) [redacted] K-161 [redacted] vs Vehicle	PP Population with Observed Data Only
Table 14.2.2.2.1	**Tear Film Break-Up Time (seconds) [redacted] K-161 [redacted] vs Vehicle**	**ITT Population with PMM**
Table 14.2.2.2.2	Tear Film Break-Up Time (seconds) by Baseline Median and Site [redacted] K-161 [redacted] vs Vehicle	ITT Population with PMM
Table 14.2.2.2.3	Tear Film Break-Up Time (seconds) by Baseline Median [redacted] K-161 [redacted] vs Vehicle	ITT Population with PMM
Table 14.2.2.2.4	Tear Film Break-Up Time(seconds) by Site [redacted] K-161 [redacted] vs Vehicle	ITT Population with PMM
Table 14.2.2.2.5	Tear Film Break-Up Time (seconds) [redacted] K-161 [redacted] vs Vehicle	ITT Population with Observed Data Only
Table 14.2.2.2.6	Tear Film Break-Up Time (seconds) [redacted] K-161 [redacted] vs Vehicle	ITT Population with LOCF
Table 14.2.2.2.7	Tear Film Break-Up Time (seconds) [redacted] K-161 [redacted] vs Vehicle	ITT Population with MCMC
Table 14.2.2.2.8	Tipping Point Analysis: Tear Film Break-Up Time (seconds) [redacted] K-161 [redacted] vs Vehicle	ITT Population with PMM
Table 14.2.2.2.9	Tear Film Break-Up Time (seconds) [redacted] K-161 [redacted] vs Vehicle	PP Population with Observed Data Only
Table 14.2.3.1	**Fluorescein Staining (Ora Calibra Scale)**	**ITT Population with PMM**
Table 14.2.3.2	Lissamine Green Staining (Ora Calibra Scale)	ITT Population with PMM
Table 14.2.3.3	Conjunctival Redness (Ora Calibra Scale)	ITT Population with PMM
Table 14.2.3.4	**Tear Film Break-Up Time (seconds)**	**ITT Population with PMM**
Table 14.2.3.5	Tear Osmolarity (mOsm/L)	ITT Population with PMM
Table 14.2.3.6	Unanesthetized Schirmer's Test (mm)	ITT Population with PMM
Table 14.2.3.7	Blink Rate (Blinks per 60 seconds)	ITT Population with PMM
Table 14.2.3.8	**Ocular Discomfort Scale (Ora Calibra)**	**ITT Population with PMM**
Table 14.2.3.9.1	Ocular Discomfort Scale during CAE (Ora Calibra)	ITT Population with PMM

Table 14.2.3.9.2	Ocular Discomfort Scale during CAE (Ora Calibra) MMRM Analysis	ITT Population with Observed Data Only
Table 14.2.3.10	Ocular Discomfort & 4-Symptom Questionnaire (Ora Calibra Scale)	ITT Population with PMM
Table 14.2.3.11	Visual Analogue Scale	ITT Population with PMM
Table 14.2.3.12	Ocular Surface Disease Index	ITT Population with PMM
[redacted]		
Table 14.2.4	Summary of Efficacy Analyses	
Table 14.3.1	**Overall Summary of Adverse Events (AE) by Treatment Group**	**Safety Population**

Table 14.3.2.1	**Ocular Treatment-Emergent Adverse Events (TEAE) by System Organ Class and Preferred Term**	**Safety Population**
Table 14.3.2.2	Non-Ocular Treatment-Emergent Adverse Events (TEAE) by System Organ Class and Preferred Term	Safety Population
Table 14.3.3.1	Ocular Treatment-Emergent Adverse Events (TEAE) by System Organ Class, Preferred Term, and Maximal Severity	Safety Population
Table 14.3.3.2	Non-Ocular Treatment-Emergent Adverse Events (TEAE) by System Organ Class, Preferred Term, and Maximal Severity	Safety Population
Table 14.3.4.1	Ocular Treatment-Related Treatment-Emergent Adverse Events (TEAE) by System Organ Class and Preferred Term	Safety Population
Table 14.3.4.2	Non-Ocular Treatment-Related Treatment-Emergent Adverse Events (TEAE) by System Organ Class and Preferred Term	Safety Population
Table 14.3.5.1	All Ocular Treatment-Emergent Adverse Events (TEAE) by System Organ Class, Preferred Term, and Study Day of Onset	Safety Population
Table 14.3.5.2	All Non-Ocular Treatment-Emergent Adverse Events (TEAE) by System Organ Class, Preferred Term, and Study Day of Onset	Safety Population
[redacted]		

Table 14.3.15	Exposure to Study Drug	Safety Population

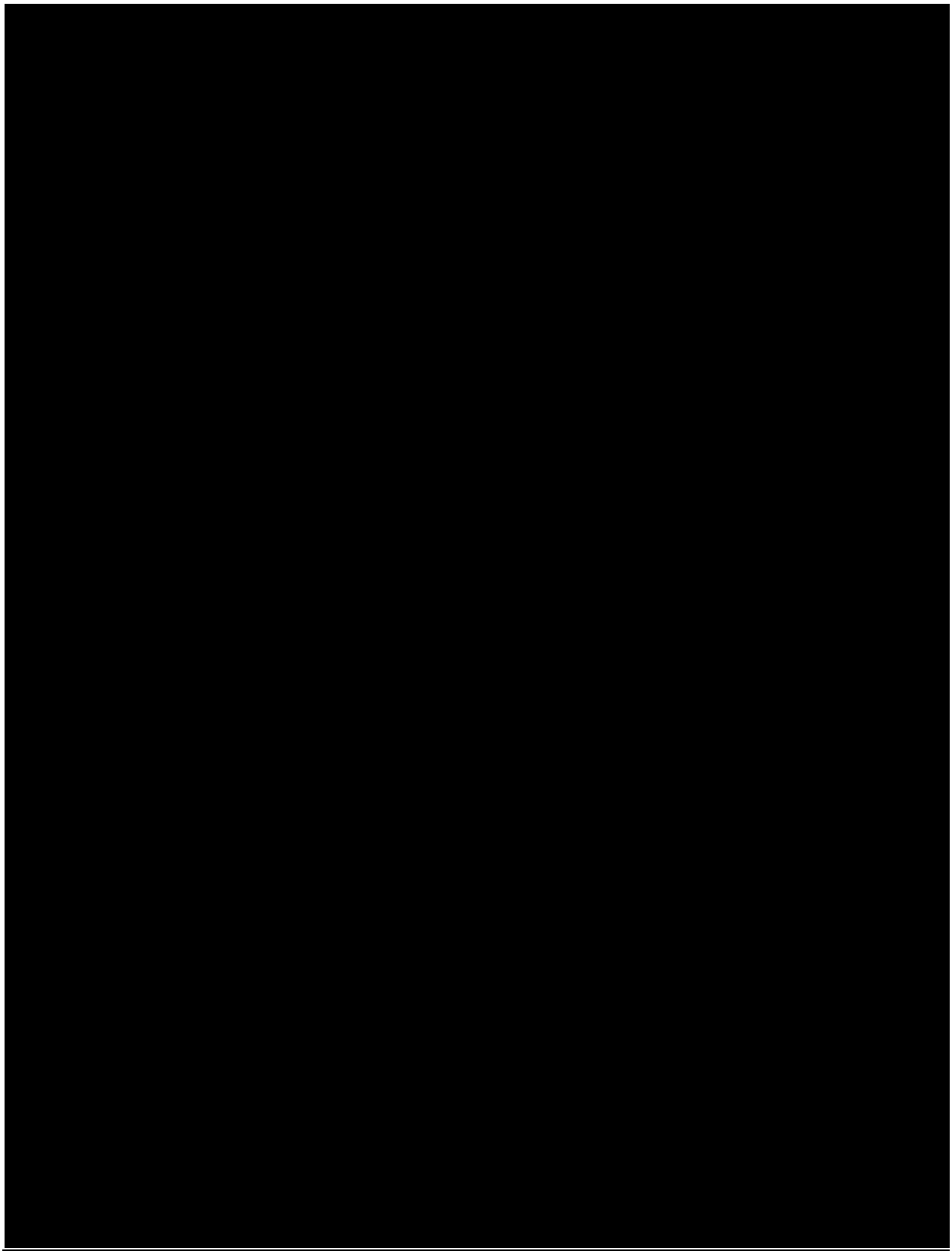
21. Listings

Listing Number	Title
Listing 16.1.7	Randomization Schedule
Listing 16.2.1.1	Subject Disposition
Listing 16.2.1.2	Inclusion/Exclusion Criteria and Screen Failure
Listing 16.2.2	Protocol Deviations
Listing 16.2.3	Study Population Inclusion
Listing 16.2.4.1	Demographics
Listing 16.2.4.2	Ocular Medical History
Listing 16.2.4.3	Non-Ocular Medical History
Listing 16.2.4.4	Ocular Prior and Concomitant Medications
Listing 16.2.4.5	Non-Ocular Prior and Concomitant Medications
Listing 16.2.5.1	Run-In, Run-In Instillation, and Run-In Replacement
Listing 16.2.5.2	Study Drug Assignment, Instillation, and Replacement
Listing 16.2.5.3	Study Drug Exposure and Dosing Compliance
Listing 16.2.5.4	Study Drug Accountability
Listing 16.2.6.1	Fluorescein Corneal and Conjunctival Staining (Ora Calibra)
Listing 16.2.6.2	Ocular Discomfort Scale
Listing 16.2.6.3	Unanesthetized Schirmer's Test
Listing 16.2.6.4	Tear Film Break-Up Time (TFBUT)
Listing 16.2.6.5	Lissamine Green Corneal and Conjunctival Staining (Ora Calibra)
Listing 16.2.6.6	Conjunctival Redness
Listing 16.2.6.7	Tear Osmolarity
Listing 16.2.6.8	Blink Rate (Blinks per 60 seconds)
Listing 16.2.6.9	Ocular Discomfort Scale (Ora Calibra®) during CAE
Listing 16.2.6.10	Ocular Discomfort & 4-Symptom Questionnaire (Ora Calibra)
Listing 16.2.6.11	Visual Analog Scale (VAS)
Listing 16.2.6.12	Ocular Surface Disease Index (OSDI)
[redacted]	

[redacted]	
Listing 16.2.7.1	All Adverse Events
Listing 16.2.7.2	Ocular Adverse Events
Listing 16.2.7.3	Non-Ocular Adverse Events
Listing 16.2.7.4	Serious Adverse Events
[redacted]	

22. Figures

Figure Number	Title	Population
Figure 14.2.1.1.1	**Inferior Corneal Fluorescein Staining (Ora Calibra Scale) [redacted] K-161 [redacted] vs Vehicle**	**ITT Population with PMM**
Figure 14.2.1.2.1	**Ocular Discomfort Scale (Ora Calibra Scale) [redacted] K-161 [redacted] vs Vehicle**	**ITT Population with PMM**
Figure 14.2.2.1.1	**Unanesthetized Schirmer's Test (mm) [redacted] K-161 [redacted] vs Vehicle**	**ITT Population with PMM**
Figure 14.2.2.2.1	**Tear Film Break-Up Time (seconds) [redacted] K-161 [redacted] vs Vehicle**	**ITT Population with PMM**
Figure 14.2.3.1	Ocular Discomfort Scale during CAE (Ora Calibra)	ITT Population with PMM



[redacted]