

**CENTER FOR DRUG EVALUATION AND
RESEARCH**

APPLICATION NUMBER:

210933Orig1s000

STATISTICAL REVIEW(S)



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

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/BLA #: NDA 210933
Supplement #: 0012
Drug Name: EYSUVIST™ (loteprednol etabonate ophthalmic suspension) 0.25%
Indication(s): For the short-term treatment of the signs and symptoms of dry eye disease
Applicant: Kala Pharmaceuticals, Inc.
Date(s): Submitted: 04/30/2020
Review Completion Goal Date: 08/17/2020
PDUFA Goal date: 10/30/2020
Review Priority: Priority

Biometrics Division: DBIV
Statistical Reviewer: Yunfan Deng, Ph.D.
Concurring Reviewer(s): Greg Soon, Ph.D.

Medical Division: Division of Transplant and Ophthalmology Products
Clinical Team: William Boyd, MD, Reviewer and Team Leader
Project Manager: Dheera Semidey

Keywords: dry eye disease, signs and symptoms, superiority

Table of Contents

1	EXECUTIVE SUMMARY	4
2	INTRODUCTION	7
2.1	OVERVIEW.....	7
2.1.1	<i>Drug Class and Indication.....</i>	7
2.1.2	<i>History of Drug Development.....</i>	8
2.1.3	<i>Studies Reviewed.....</i>	10
2.2	DATA SOURCES	10
3	STATISTICAL EVALUATION	11
3.1	DATA AND ANALYSIS QUALITY	11
3.2	EVALUATION OF EFFICACY	11
3.2.1	<i>Study Design and Endpoints</i>	11
3.2.2	<i>Statistical Methodologies</i>	14
3.2.3	<i>Patient Disposition, Demographic and Baseline Characteristics</i>	15
3.2.3.1	<i>Patient Disposition</i>	15
3.2.4	<i>Results and Conclusions.....</i>	17
3.2.4.1	<i>Primary Symptom Endpoint and Supporting Results</i>	17
3.2.4.2	<i>Primary Sign Endpoint</i>	21
3.2.4.3	<i>Conclusion</i>	21
3.3	EVALUATION OF SAFETY	22
4	FINDINGS IN SPECIAL/SUBGROUP POPULATIONS.....	24
4.1	GENDER, RACE, AND AGE.....	24
4.2	ODS SCORES 1 DAY PRIOR TO DAY 1	24
5	SUMMARY AND CONCLUSIONS	25
5.1	STATISTICAL ISSUES	25
5.2	COLLECTIVE EVIDENCE	25
5.3	CONCLUSIONS AND RECOMMENDATIONS	26
5.4	LABELING RECOMMENDATIONS	26
APPENDIX 1:	(b) (4)	28
APPENDIX 2:	(b) (4)	29
APPENDIX 3:	(b) (4)	31
APPENDIX 4: RESULTS OF THE PRE-DEFINED PRIMARY SIGN AND SYMPTOM ENDPOINTS IN THE PP POPULATION		
		32
APPENDIX 5: SUBGROUP ANALYSES FOR THE PHASE 3 STUDIES.....		
		33

LIST OF TABLES

Table 1: Summary of Primary Efficacy Endpoints.....	5
Table 2: Summary of Efficacy Studies to be assessed in the Statistical Review.....	10
Table 3: Summary of Primary Efficacy Endpoints.....	13
Table 4: Summary of Analysis Population (All Randomized).....	15
Table 5: Subject Disposition (ITT).....	15
Table 6: Demographic and Baseline Characteristics (ITT).....	16
Table 7: Results for the Primary Symptom Endpoint (ITT; Observed).....	18
Table 8: Results for Change from baseline of ODS scores on Day 8 using 3-day mean scores (ITT; Observed).....	18
Table 9: Daily ODS Score (ITT, Observed).....	20
Table 10: Results for the Primary Sign Endpoint (ITT; Observed).....	21
Table 11: Summary of Adverse Events (Pooled Safety Population).....	22
Table 12: Safety Analysis: Adverse Events Associated with ≥ 2 Subjects in Either Treatment Group by System Organ Class and Preferred Term (Studies 002, 006, 007, and 011, Safety Population).....	22
Table 13: Number (%) of Subjects Reporting Common ($\geq 0.3\%$ in Either Treatment Group) Treatment-Related Adverse Events (Pooled Safety Population).....	23
Table 14: Study 002 Procedure.....	29
Table 15: Studies 006, 007, and 011 Procedures.....	30
Table 16: Results for the Primary Symptom Endpoint (PP; Observed).....	32
Table 17: Results for the Primary Sign Endpoint (PP; Observed).....	33
Table 18: Subgroup Analysis of Change from Baseline in Conjunctival Hyperemia on Day 15 by Age, Gender, and Race for All Phase 3 Studies (ITT Observed).....	33
Table 19: Subgroup Analysis of Change from Baseline in ODS on Day 15 Using 3-day Mean by Age, Gender, and Race for All Phase 3 Studies (ITT Observed).....	33

LIST OF FIGURES

Figure 1: Change from Baseline in Conjunctival Hyperemia and ODS using the 3-day mean at Day 15 (ITT; Observed).....	5
Figure 2: Daily ODS Overtime from Days 1 to 15 (ITT; Observed).....	6
Figure 3: Study Schematics.....	12
Figure 4: Change from Baseline in ODS on Day 15 Using 3-day Mean by ODS Score 1 Day Prior to Day 1 (<68 mm and ≥ 68 mm) (ITT Observed).....	25

1 EXECUTIVE SUMMARY

This NDA seeks approval of EYSUVISTM (loteprednol etabonate [LE] ophthalmic suspension) 0.25% dosed four times daily (QID) for the short-term treatment of the signs and symptoms of dry eye disease. This NDA had two major submissions: the original NDA submitted on October 15, 2018 and the current supplement submission on April 30, 2020. To support efficacy and safety, the original submission in 2018 included one Phase 2 (KPI-121-C-002, referred to as Study 002) and two Phase 3 studies (Studies KPI-121-C-006, referred to as Study 006 and KPI-121-C-007, referred to as Study 007); the current NDA supplement included a single Phase 3 study (Studies KPI-121-C-011, referred to as Study 011).

Based on the statistical review of the original submission in 2018 by another statistical reviewer Dr. Qing Zhou, the three studies (002, 006, and 007) demonstrated a consistent treatment effect of LE 0.25% versus vehicle in terms of dry eye sign – conjunctival hyperemia (at Day 29 for Study 002, and at Day 15 for Studies 006 and 007); however, the three studies showed less consistent results for the dry eye symptom – change from baseline in ocular discomfort scores (ODS) at Day 15 using the 3-day mean. A complete response letter was sent to the applicant on 08/07/2019, the Division recommended that to demonstrate substantial evidence of efficacy in the intended patient population, data from at least one additional clinical trial be submitted in which loteprednol etabonate ophthalmic suspension, 0.25% demonstrates superiority to its vehicle in the treatment of signs and symptoms of dry eye disease. Subsequently, the applicant completed Study 011, which focused on the dry eye symptoms endpoint and submitted the results of Study 011 in this supplement. Along with Study 011 submitted in this NDA supplement, for the purpose of across study comparison, this statistical review also presents results for the Phase 2 study (002) and the two Phase 3 studies (006 and 007) previously submitted.

All four studies were similar in design except that the treatment period for the Phase 2 study was 28 days and for the Phase 3 studies was 14 days; while Study 002 and Studies 006 and 007 had both dry eye sign and symptom as the co-primary efficacy endpoints, Study 011 had dry eye symptom as the primary efficacy endpoint and the dry eye sign endpoint of conjunctival hyperemia as the first secondary efficacy endpoint. All four studies were randomized, double-masked, multi-center, vehicle-controlled studies conducted in the United States (US). In the Phase 2 study, 150 subjects were randomized at 1:1 ratio to receive LE 0.25% QID or vehicle QID for 28 days, with the study visits on Days -14 (screening), 1 (baseline/randomization), 8, 15, 22, and 29. For the three Phase 3 studies, each study randomized approximately 900 subjects at 1:1 ratio to receive either LE 0.25% QID or vehicle QID for 14 days with the study visits on Days -14 (screening), 1 (baseline/randomization), 8, and 15.

Table 1 listed the primary sign and symptom endpoints for the four studies. Conjunctival hyperemia scores were 5-level scales from 0 (none) to 4 (severe); corneal staining scores were 4-level scales from 0 (none) to 3 (severe); and ocular discomfort scores (ODS) was a visual analog scale from 0 to 100 mm, where 0 indicated “very mild” and 100 indicated “very severe”.

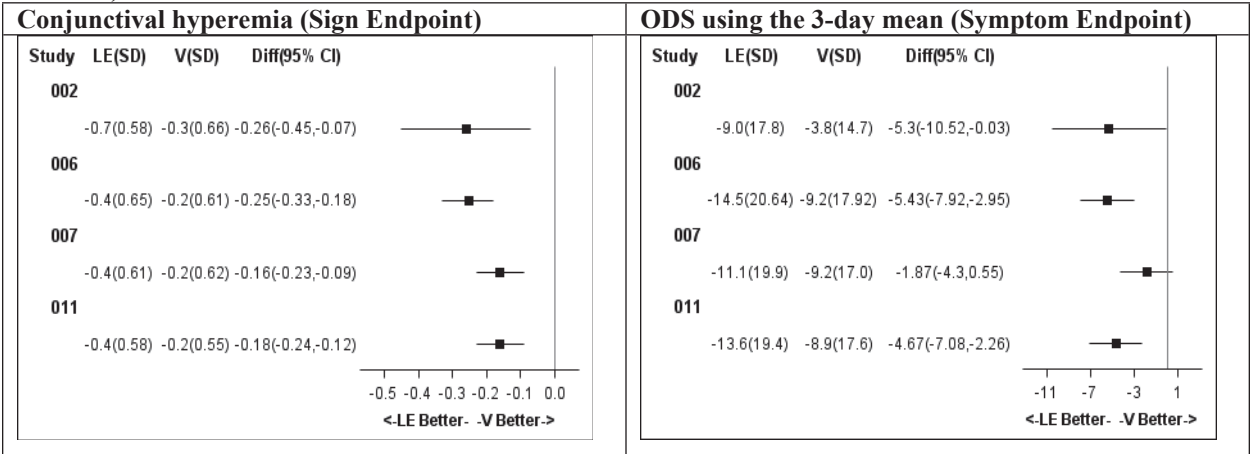
Table 1: Summary of Primary Efficacy Endpoints

	Study 002	Study 006	Study 007	Study 011
Sign	Mean conjunctival hyperemia scores in the worst region of the study eye at Day 29	Change from baseline in conjunctival hyperemia scores in the worst region of the study eye at Day 15 Change from baseline in inferior corneal staining scores in the study eye at Day 15	Change from baseline in conjunctival hyperemia scores in the worst region of the study eye at Day 15	As the first secondary endpoint: Change from baseline in conjunctival hyperemia scores in the worst region of the study eye at Day 15
Symptom	Mean ODS at Day 29 using the 7-day mean	Change from baseline in ODS at Day 15 using the 3-day mean Change from baseline in ODS at Day 15 using the 3-day mean in subgroup of patients with ODS ≥ 68 mm one day prior to Day 1	Change from baseline in ODS at Day 15 using the 3-day mean	Change from baseline in ODS at Day 15 using the 3-day mean Change from baseline in ODS at Day 15 using the 3-day mean in subgroup of patients with ODS ≥ 68 mm one day prior to Day 1

Source: Table 1 of Summary of Clinical Efficacy.

All four studies demonstrated superiority of the LE 0.25% to the vehicle for the primary sign endpoint of change from baseline in conjunctival hyperemia at Day 15 as shown in Figure 1. The Phase 2 study demonstrated superiority of the LE 0.25% to the vehicle for its own primary sign endpoint at Day 29 and for the Phase 3 studies’ primary sign endpoint at Day 15. As presented in Figure 1, two studies (006 and 011) demonstrated superiority of the LE 0.25% to the vehicle for the primary symptom endpoint of change from baseline in ODS at Day 15. However, the other two studies (002 and 007) failed to demonstrate superiority of LE 0.25% for their pre-defined primary symptom endpoint but both showed nominal difference favoring the LE group. Please note that the results achieved nominal significance for the Phase 2 study post-hoc by analyzing the same primary symptom endpoint (change from baseline in ODS at Day 15 using the 3-day mean) as the three Phase 3 studies.

Figure 1: Change from Baseline in Conjunctival Hyperemia and ODS using the 3-day mean at Day 15 (ITT; Observed)



Note: ITT = Intent-to-Treat, LE = LE 0.25%, V = Vehicle, Diff = Difference, CI = Confidence Interval

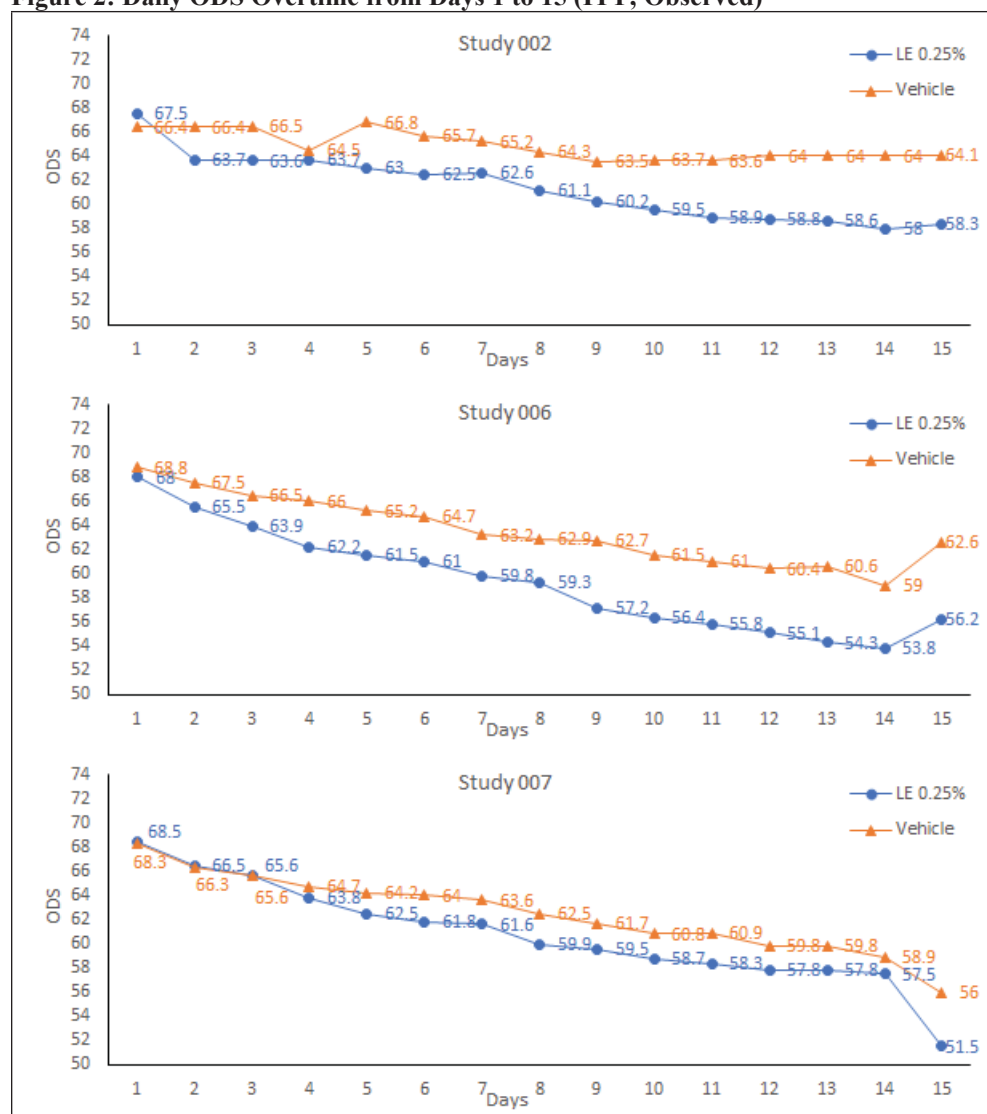
* The estimated difference (95% CI and p-value) was based on an ANCOVA model with treatment, baseline values, and stratification variables.

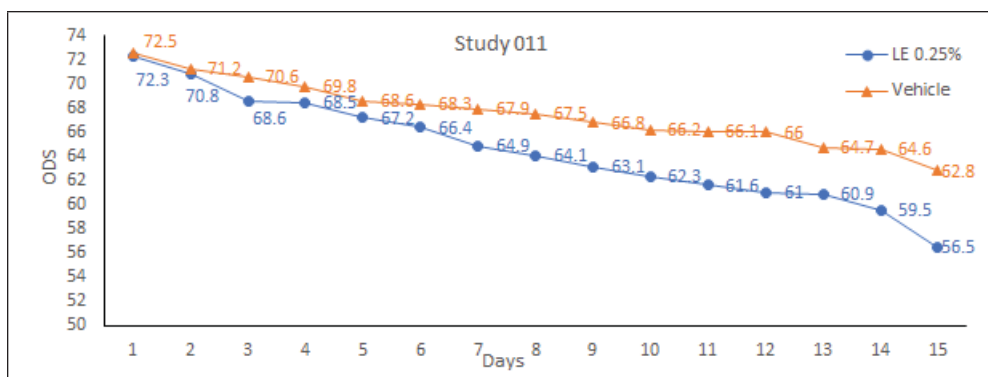
Source: Table 11 of Study 006 CSR, Table 11 of Study 007 CSR, and Table 11 of Study 011 CSR.

Further examining of daily ODS score over time, as presented in Figure 2, showed that there was clear separation in the daily ODS score between the LE group and the vehicle group in all the four studies. Across studies, treatment difference appeared similar among Studies 002, 006 and 011 over time while Study 007 appeared to have less treatment separation.

In conclusion, based on the collective evidence of all four studies, there were treatment effects of LE 0.25% versus vehicle in terms of reducing conjunctival hyperemia (dry eye sign) and lowering ocular discomfort score (dry eye symptom). Therefore, the statistical reviewer recommends the approval of EYSUVIS for the short-term treatment of the signs and symptoms of dry eye disease.

Figure 2: Daily ODS Overtime from Days 1 to 15 (ITT; Observed)





Source: Table 16.1.13.1 of Study 002 add-table, Table 14.2.4.1.1 of Study 006 CSR, Table 14.2.4.1.1 of Study 007 CSR, and Table 14.2.2.4.1 of Study 011 CSR.

2 INTRODUCTION

2.1 Overview

2.1.1 Drug Class and Indication

Dry eye disease is a common condition that occurs when one's tears can't provide adequate lubrication for the eyes. Tears can be inadequate for many reasons. For example, dry eyes may occur if you don't produce enough tears or if you produce poor-quality tears. The medical term for this **condition** is keratoconjunctivitis sicca. Common causes of decreased tear production include:

- Aging
- Certain medical conditions, including diabetes, rheumatoid arthritis, lupus, scleroderma, Sjogren's syndrome, thyroid disorders and vitamin A deficiency
- Certain medications, including antihistamines, decongestants, hormone replacement therapy, antidepressants, and drugs for high blood pressure, acne, birth control and Parkinson's disease
- Laser eye surgery, though symptoms of dry eyes related to this procedure are usually temporary
- Tear gland damage from inflammation or radiation

The disease is characterized by symptoms of ocular discomfort, including but not limited to, eye dryness, foreign body sensation, irritation, burning, tearing, ocular pain, itching, and blurring of vision. Based on epidemiology and market research conducted by the applicant, it is estimated that there are (b) (4) patients with a diagnosis of dry eye disease in the United States. Currently, there are two FDA-approved ophthalmic solutions for the long-term use of dry eye: cyclosporine ophthalmic solution (Restasis; Allergan) and lifitegrast (Xiidra; Shire). The drug product for this NDA submission, loteprednol etabonate (LE) 0.25%, dosed four times daily (QID), has the proposed indication of short-term relief instead of long-term usage for the signs and symptoms of dry eye disease.

LE is an ester-based corticosteroid. LE ophthalmic suspension, 0.5% (Lotemax®; Bausch & Lomb), was approved by the FDA in 1998 (NDA20803) for the treatment of steroid responsive inflammatory conditions of the palpebral and bulbar conjunctiva, cornea and anterior segment of the globe such as allergic conjunctivitis, acne rosacea, superficial punctate keratitis, herpes zoster keratitis, iritis, cyclitis, selected infective conjunctivitides, when the inherent hazard of steroid use is accepted to obtain an advisable diminution in edema and inflammation. In 2012, LOTEMAX® (loteprednol etabonate ophthalmic gel) 0.5% was approved by the Agency for the treatment of post-operative inflammation and pain following ocular surgery. The marketed dosing frequency for LE 0.5% is QID, and three different dosage forms have been approved: suspension (NDA 20583 and 20841; approved in 1998), ointment (NDA 200738; approved in 2011), and gel (NDA 202872; approved in 2012).

LE 1%, dosed twice daily, which was also developed by the applicant, was approved by FDA in August 2018 for the treatment of postoperative pain and inflammation. The applicant also developed the new formulation of LE, KPI-121 0.25%, for the short-term treatment of the signs and symptoms of dry eye disease, utilizing a 2-week course of topical ophthalmic therapy administered four times a day (QID). Both LE 1% and 0.25% are formulated using the applicant-developed proprietary drug-delivery technology known as mucus penetrating particles (MPP). According to the applicant, MPP technology can enhance penetration through the mucous layer of the tear film and thus improve the efficacy of the drug product while maintaining favorable safety profile of LE.

2.1.2 History of Drug Development

In June 2014, under IND 117192, the applicant submitted a protocol for a Phase 2 study (002) to Division of Transplant and Ophthalmology Products (DTOP) to evaluate the safety and efficacy of LE 0.25% for the treatment of dry eye disease. The study included a primary sign endpoint and a primary symptom endpoint. The primary sign endpoint was the mean conjunctival hyperemia score in the worst region of the study eye at Day 29 (see Appendix 1 for the definitions of the study eye and the worst region in the study eye); the primary symptom endpoint was the mean modified symptom assessment in dry eye (SANDE) severity scores using the mean scores for the 7 days prior to Day 29.

On a Type C meeting held with the Division in June 2015, the applicant reported that they had completed Study 002; and statistical significance was reached for the primary sign endpoint but not for the primary symptom endpoint. The following key clinical issues were also discussed with the Division:

- The applicant asked whether DTOP accepted LE 0.25% for a dry eye indication if the duration of the treatment was (b) (4) 14 (b) (4) days. DTOP initially disagreed as dry eye disease is a chronic condition. After discussion, DTOP stated that with the appropriate limitations in container size/volume and labeling, an indication for the temporary relief of the signs and symptoms of dry eye disease could be considered.
- DTOP accepted conjunctival hyperemia as the primary sign endpoint and indicated that they would be open to a responder analysis for total clearing of conjunctival hyperemia.

DTOP also accepted modified SANDE severity scores using the 7-day mean as the primary symptom endpoint. Furthermore, DTOP stated that shorter analysis intervals (e.g., 3 days) for SANDE severity scores might also be acceptable for the statistical analysis.

Subsequently, the applicant submitted protocols for two Phase 3 studies (Studies 006 and 007) with 2 primary sign endpoints: change from baseline in conjunctival hyperemia scores in the worst region of the study eye at Day 15 and change from baseline in inferior corneal staining scores at Day 15; and two primary symptom endpoints: change from baseline in ocular discomfort severity (ODS) scores at Day 15 and change from baseline in ODS scores at Day 15 in subjects with more severe baseline ODS defined as ODS scores ≥ 68 mm 1 day prior to Day 1. It is noted that ODS was defined the same as the modified SANDE severity scores in Study 002, therefore modified SANDE severity scores in Study 002 will be hereinafter referred to as ODS scores for consistency.

In May 2018, after completion of Studies 006 and 007, a Type C meeting was held to discuss the results of Studies 002, 006, and 007. The applicant reported that a statistically significant treatment difference was achieved in all three studies for the primary sign endpoint of conjunctival hyperemia. For the primary symptom endpoint, the significant treatment difference was only achieved in Study 006 but not in the other two studies. The applicant conducted numerous post-hoc analyses. In particular, the applicant analyzed the primary symptom endpoint of the two Phase 3 studies in Study 002 post-hoc, the result was supportive (p-value = 0.0489). The applicant asked if the Division agreed that there was sufficient safety and efficacy data to support an NDA filing based on the results of the three studies. The Division disagreed as Study 007 failed to meet the primary symptom endpoint and recommended an additional study.

In June 2018, the applicant submitted a protocol for a new Phase 3 study (Study 011) for dry eye indication. This new study would randomize approximately 900 subjects to receive either LE or vehicle. This new study modified inclusion criteria according to the exploratory results of the completed studies (Studies 002, 006, and 007) to include potentially more severe patients. In addition to the requirement of a score of > 50 mm on ODS at screening (as in Studies 002, 006, and 007), the inclusion criteria added another requirement of a score of ≥ 45 mm ODS on each of the 3 days prior to Day 1. The two primary endpoints in the new study were both for dry eye symptom: 1) change from baseline in ODS at Day 15 using the 3-day mean in the ITT population and 2) change from baseline in ODS at Day 15 using the 3-day mean in the subgroup of subjects with more severe baseline ODS defined as ODS ≥ 68 mm for the 3 days prior to Day 1.

Nonetheless, without completing the newly designed Study 011, the applicant submitted an NDA application in October 2018 using Studies 002, 006, and 007 as evidence for efficacy and safety. Based on the review of the previous statistical reviewer Dr. Qing Zhou, the submitted 3 studies demonstrated a consistent treatment effect of LE 0.25% versus vehicle in terms of dry eye sign – conjunctival hyperemia (at Day 29 for Study 002, and at Day 15 for Studies 006 and 007); however, the three studies showed less consistent results for the dry eye symptom – change from baseline in ODS at Day 15 using the 3-day mean. Furthermore, the previous statistical reviewer acknowledged that all three studies showed a consistent treatment separation in daily ODS over time, which indicated a treatment activity. Therefore, the previous statistical reviewer deferred to

the approval decision to the clinical review team based on their judgement whether this treatment activity was clinically relevant and sufficient for approval.

The clinical review team considered the evidence for dry eye symptom was not sufficient enough for an approval based on the three studies. A complete response letter was sent to the applicant on 08/07/2019, the Division recommended that to demonstrate substantial evidence of efficacy in the intended patient population, data from at least one additional clinical trial be submitted in which loteprednol etabonate ophthalmic suspension, 0.25% demonstrates superiority to its vehicle in the treatment of signs and symptoms of dry eye disease. Subsequently, the applicant completed Study 011, which focused on the dry eye symptoms endpoint and submitted the efficacy and safety results of Study 011 in this supplement.

2.1.3 Studies Reviewed

This review focuses on Study 011 but also includes major design information and results of Studies 002, 006, and 007 for the purpose of cross study comparison. Table 2 below contains a brief summary of Study 011.

Table 2: Summary of Efficacy Studies to be assessed in the Statistical Review

Study No	Design	Objective	Treatment / Sample Size	Study Population	Primary Endpoint
KPI-121-C011	Multi-center, randomized, double-blinded, parallel group, vehicle-controlled	To evaluate the efficacy and safety of LE 0.25% as compared to vehicle	LE 0.25% / 447 Vehicle / 454	Adult patients had a documented clinical diagnosis of dry eye disease in both eyes	Change from baseline in ODS at Day 15 using the 3-day mean Change from baseline in ODS at Day 15 using the 3-day mean in subgroup of patients with ODS $\geq$ 68 mm one day prior to Day 1

Source: Statistical Reviewer's Summary.

2.2 Data Sources

The data sources for this review include clinical study reports, protocols, statistical analysis plan (SAP), and datasets. All data sources were electronic submitted and located at [\\CDSESUB1\evsprod\NDA210933\0001](#), [\\CDSESUB1\evsprod\NDA210933\0003](#), and [\\CDSESUB1\evsprod\NDA210933\0012](#).

3 STATISTICAL EVALUATION

3.1 Data and Analysis Quality

Overall, the submitted data were of good quality with definitions provided for each variable. Results of the primary and secondary efficacy endpoints can be verified by the statistical reviewer with minor data manipulation. The statistical reviewer's analyses were primarily based on the analysis datasets.

3.2 Evaluation of Efficacy

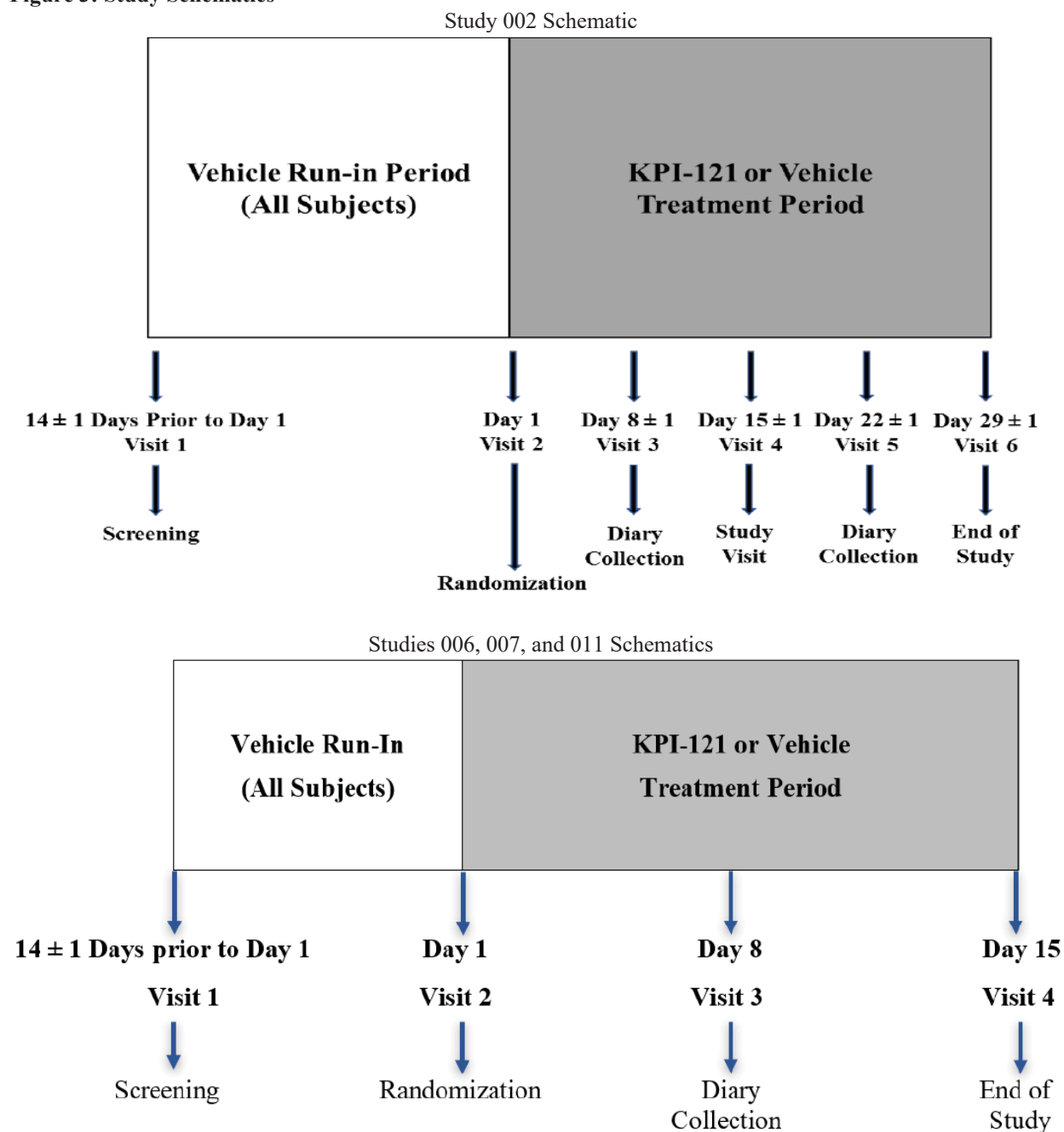
This section evaluates the efficacy results of the four studies together. Sections 3.2.1 and 3.2.2 summarize the study design, endpoints, and statistical methods. Section 3.2.3 summarizes subject disposition as well as demographic and baseline characteristics. Section 3.2.4 discusses the primary analysis and supporting evidence.

3.2.1 Study Design and Endpoints

All four studies were randomized, double-masked, vehicle-controlled studies conducted in the US to evaluate the efficacy and safety of LE 0.25% dosed four times daily (QID) as compared to the vehicle. Study 002 was designed as a Phase 2 study, it had a 14-day vehicle run-in period followed by a 28-day treatment period. There were 6 visits for the study (Days -14 [Screening], 1 [randomization], 8, 15, and 29 [end of study]). Studies 006, 007, and 011 were all Phase 3 studies with a 14-day vehicle run-in period followed by a 14-day treatment period. There were 4 visits for the studies (Days -14 [Screening], 1 [randomization], 8, and 15 [end of study]).

Figure 3 below displays the study schematics for the four studies. Detailed schedules of study assessments are shown in Appendix 2.

Figure 3: Study Schematics



Source: Summary of Clinical Efficacy, Figures 1 and 2.
 Note: KPI-121 is the applicant's internal product name for LE.

The key inclusion criteria in the four studies were:

- Adult patients had a documented clinical diagnosis of dry eye disease in both eyes
- Had ongoing dry eye disease as defined by the following criteria in the same eye or both eyes:

- a. A corneal fluorescein staining score at the screening visit and on Day 1 of ≥ 6 national eye institute (NEI) scale AND
- b. Investigator-rated assessment of bulbar conjunctival hyperemia at the screening visit and on Day 1 of ≥ 2 as assessed using the cornea and contact lens research unit (CCLRU) scale AND
- c. An unanesthetized Schirmer Test score at the screening visit of ≤ 10 mm
- d. A score of > 50 mm on ODS at screening and a score of > 40 mm ODS on the day prior to Day 1 for Studies 002, 006, 007; a score of > 50 mm on ODS at screening and a score of ≥ 45 mm ODS on each of the 3 days prior to Day 1 for Study 011 as the applicant would like to enroll potentially more severe patients.

During the run-in period, subjects in all four studies received a single dose of single-masked vehicle QID in each eye. At Day 1, eligible subjects in these studies were randomized in a 1:1 ratio to receive either LE 0.25% QID or vehicle QID for the entire treatment period (28 days in the Phase 2 study and 14 days in the three Phase 3 studies). In the Phase 2 study, randomization was stratified based on subjects' ODS scores (> 50 mm to 70 mm [later amended to include > 40 mm scores at Day 1], > 70 mm to 100 mm) on the day prior to Day 1 and the investigator-rated conjunctival hyperemia scores ($2, \geq 3$) in the study eye on Day 1. In the Phase 3 studies 006 and 007, randomization was stratified based on subjects' ODS scores (> 40 mm to < 68 mm, ≥ 68 mm to 100 mm) on the day prior to Day 1 and the investigator-rated conjunctival hyperemia scores ($2, \geq 3$) in the study eye on Day 1. In the Phase 3 study 011, randomization was stratified based on subjects' ODS scores (≥ 45 mm to < 68 mm, ≥ 68 mm to 100 mm) on the day prior to Day 1.

The primary efficacy endpoints for the four studies are summarized in the following table. The conjunctival hyperemia scores were the 5 levels grading scales from 0 (none) to 4 (severe) based on photographic assessments conducted at a central reading center. In this review, unless noted otherwise, conjunctival hyperemia scores were assessed in the worst region of the study eye (see Appendix 1 for the definition of the study eye and the worst region). Corneal staining was a four-level grading scales from 0 (none) to 3 (severe) assessed in each of the five sections of corneal (superior, inferior, nasal, temporal, and central) of the study eye. The ODS scores were collected each day (one record per day) in subject's dairy using a visual analog scale from 0 to 100 mm, where 0 indicated "very mild" and 100 indicated "very severe". The primary symptom endpoints defined in the three Phase 3 studies were also analyzed in the Phase 2 study post-hoc. The primary sign endpoint of change from baseline in conjunctival hyperemia in the two Phase 3 studies 006 and 007 were defined as the first secondary efficacy endpoint in Study 011 and as one of secondary efficacy endpoints in the Phase 2 study.

Table 3: Summary of Primary Efficacy Endpoints

	Study 002	Study 006	Study 007	Study 011
Sign	Mean conjunctival hyperemia scores in the worst region of the study eye at Day 29	Change from baseline in conjunctival hyperemia scores in the worst region of the study eye at Day 15	Change from baseline in conjunctival hyperemia scores in the worst region of the study eye at Day 15	First Secondary Endpoint: Change from baseline in conjunctival hyperemia scores in the worst region of the study eye at Day 15

		Change from baseline in inferior corneal staining scores in the study eye at Day 15		
Symptom	Mean ODS at Day 29 using the 7-day mean	Change from baseline in ODS at Day 15 using the 3-day mean Change from baseline in ODS at Day 15 using the 3-day mean in subgroup of patients with ODS $\geq$ 68 mm one day prior to Day 1	Change from baseline in ODS at Day 15 using the 3-day mean	Change from baseline in ODS at Day 15 using the 3-day mean Change from baseline in ODS at Day 15 using the 3-day mean in subgroup of patients with ODS $\geq$ 68 mm one day prior to Day 1

Source: Table 1 of Summary of Clinical Efficacy.

3.2.2 Statistical Methodologies

For the four studies, there were three different analysis populations (also known as analysis sets) defined by the applicant:

- **Intent-to-Treat (ITT) population.** In the Phase 2 study, ITT population was defined as all randomized subjects with at least 1 post baseline evaluation. In the three Phase 3 studies, ITT population was defined as all randomized subjects who received at least 1 dose of study treatment. The primary analysis was performed on the ITT population for all four studies and subjects were analyzed as randomized.
- **Per-Protocol (PP) Population.** In all four studies, PP population was defined as all subjects who completed the randomized study treatment period, fulfilled all study visits, achieved reasonable compliance with the study protocol, and had no major protocol violations.
- **Safety Population.** In all four studies, safety population was defined as all randomized subjects who received at least 1 dose of the study drug and from whom at least 1 safety assessment was obtained after randomization. Subjects were analyzed as treated in the safety population.

All primary endpoints were analyzed using the analysis of covariance (ANCOVA) model with baseline value as covariate and treatment group and the randomization strata as fixed factors. In the Phase 2 study, the two primary endpoints were tested hierarchically (sign first and symptom second). In Study 006, the four primary endpoints were tested using the parallel gatekeeping method with recycling unspent α , details of this method are presented in Appendix 3. In Study 007, the two primary endpoints were tested using the Bonferroni correction. In Study 011, the two primary efficacy endpoints were tested using the Hochberg step-up procedure.

For the missing data, the SAPs for the four studies stated that if more than 5% of data points were missing for a given primary endpoint, multiple imputation would be employed. As the amount of missing was less than 3% for all the studies, the applicant conducted the efficacy analyses based on the available (observed) data without any imputation. Continuous secondary endpoints were

analyzed similarly as the primary endpoints. Binary secondary endpoints were analyzed using the Chi-square test.

The applicant stated in the SAPs of the Phase 3 studies that if all the primary efficacy endpoints achieved statistical significance, all secondary endpoints would be tested hierarchically following the order defined in each protocol.

3.2.3 Patient Disposition, Demographic and Baseline Characteristics

3.2.3.1 Patient Disposition

Studies 002, 006, 007, and 011 each randomized 150, 918, 909, and 901 subjects, respectively. Table 4 is a summary of number of subjects in each analysis population by study and by treatment:

- In Study 006, two subjects were enrolled in the study twice at 2 different sites; one subject was randomized in this study again after this subject completed Study 007. The second participation data from these three subjects were excluded from the analysis populations.
- In Study 007, one subject was enrolled in the study twice at 2 different sites; two subjects were randomized after they had completed Study 006. The second participation data from these subjects were excluded.
- In Study 002, one subject was randomized to LE 0.25% but received vehicle. This subject was included in the ITT population as randomized.
- In Study 011, two subjects received the wrong study treatment; they were assigned vehicle but received LE 0.25%. These subjects were included in the ITT population as randomized and in the safety and PP populations as treated.

Table 4: Summary of Analysis Population (All Randomized)

	Study 002		Study 006		Study 007		Study 011	
	LE 0.25% (N=73)	Vehicle (N=77)	LE 0.25% (N=460)	Vehicle (N=458)	LE 0.25% (N=455)	Vehicle (N=454)	LE 0.25% (N=447)	Vehicle (N=454)
ITT	73 (100%)	77 (100%)	459 (99.8%)	456 (99.6%)	452 (99.3%)	453 (99.8%)	447 (100%)	454 (100%)
PP	67 (91.8%)	66 (85.7%)	410 (89.1%)	409 (89.3%)	416 (91.4%)	419 (92.3%)	377 (84.3%)	405 (89.2)
Safety	73 (100%)	77 (100%)	459 (99.8%)	456 (99.6%)	452 (99.3%)	453 (99.8%)	449 (100.4%)	452 (99.6%)

Source: Table 9 of Study 002 CSR, Table 9 of Study 006 CSR, Table 9 of Study 007 CSR, and Table 9 of Study 011 CSR.

Table 5 shows the subject disposition. Only a small percentage (< 3%) of randomized subjects in the four studies discontinued early from the study. Only one subject discontinued due to adverse events (AEs) in the Phase 2 study and three (3) subjects in the Phase 3 study 011; no subject was discontinued due to AEs in the two Phase 3 studies 006 and 007.

Table 5: Subject Disposition (ITT)

	Study 002		Study 006		Study 007		Study 011	
	LE 0.25% (N=73)	Vehicle (N=77)	LE 0.25% (N=460)	Vehicle (N=458)	LE 0.25% (N=455)	Vehicle (N=454)	LE 0.25% (N=447)	Vehicle (N=454)

Completed Study	73 (100%)	77 (100%)	459 (99.8%)	456 (99.6%)	452 (99.3%)	453 (99.8%)	447 (100%)	454 (100%)
Discontinued	1 (1.4%)	1 (1.3%)	4 (0.9%)	3 (0.7%)	5 (1.1%)	5 (1.1%)	10 (2.2%)	1 (0.2%)
Reasons								
Adverse Event	1 (1.4%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	3 (0.7%)	0 (0%)
Withdrawal by Subject	0 (0%)	0 (0%)	3 (0.7%)	1 (0.2%)	2 (0.4%)	2 (0.4%)	3 (0.7%)	0 (0%)
Lost to Follow-up	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	2 (0.4%)	1 (0.2%)	1 (0.2%)
Other	0 (0%)	1 (1.3%)	1 (0.2%)	2 (0.4%)	3 (0.7%)	1 (0.2%)	3 (0.7%)	0 (0%)

Source: Table 7 of Study 002 CSR; Table 7 of Study 006 CSR; Table 7 of Study 007 CSR; Table 8 of Study 011 CSR.

Demographic and baseline characteristics were generally balanced between treatment groups in the four studies (Table 6). Most study subjects were non-Hispanic (over 70%), female (over 70%), and white (over 70%). The mean age in the four studies was approximately 58 years old (range: 18 to 93). Note that Study 011 had higher mean ODS at baseline since this study recruited subjects with potentially more severe dry eye symptoms.

Table 6: Demographic and Baseline Characteristics (ITT)

	Study 002		Study 006		Study 007		Study 011	
	LE 0.25% (N=73)	Vehicle (N=77)	LE 0.25% (N=460)	Vehicle (N=458)	LE 0.25% (N=455)	Vehicle (N=454)	LE 0.25% (N=447)	Vehicle (N=454)
Age at Randomization (years)								
Mean (SD)	56.3 (17.2)	54.9 (12.3)	58.1 (14.7)	58.3 (14.7)	59.1 (14.5)	59.3 (15.0)	57.6 (15.3)	57.3 (15.5)
Median	58	57	60	60.5	61.5	62.0	59.0	60.0
Min, Max	18, 88	22, 79	18, 90	19, 88	19, 91	18, 93	18, 89	18, 88
Age Group (years)								
< 65	48 (65.8%)	58 (75.3%)	280 (61%)	289 (63.4%)	265 (58.6%)	270 (59.6%)	274 (61.3%)	282 (62.1%)
≥ 65	25 (34.2%)	19 (24.7%)	179 (39%)	167 (36.6%)	187 (41.4%)	183 (40.4%)	173 (38.7%)	172 (37.9%)
Sex								
Female	53 (72.6%)	61 (79.2%)	367 (80%)	359 (78.7%)	332 (73.5%)	354 (78.1%)	339 (75.8%)	330 (72.7%)
Male	20 (27.4%)	16 (20.8%)	92 (20%)	97 (21.3%)	120 (26.5%)	99 (21.9%)	108 (24.2%)	124 (27.3%)
Race								
White	54 (74.0%)	63 (81.8%)	362 (78.9%)	361 (79.2%)	345 (76.3%)	352 (77.7%)	348 (77.9%)	341 (75.1%)
Non-White	19 (26.0%)	14 (18.2%)	97 (21.1%)	95 (20.8%)	107 (23.7%)	101 (22.3%)	99 (22.1%)	113 (24.9%)
Ethnicity								
Hispanic or Latino	15 (20.5%)	16 (20.8%)	65 (14.2%)	76 (16.7%)	118 (26.1%)	118 (26%)	112 (25.1%)	97 (21.4%)
Not Hispanic /Latino	58 (79.5%)	61 (79.2%)	394 (85.8%)	380 (83.3%)	334 (73.9%)	335 (74%)	335 (74.9%)	357 (78.6%)
Baseline conjunctival hyperemia by photographic assessment								
Mean (SD)	2.2 (0.45)	2.1 (0.43)	2.0 (0.6)	2.0 (0.62)	2.0 (0.62)	2.0 (0.57)	2.1 (0.64)	2.1 (0.61)
Median	2	2	2	2	2	2	2	2
Min, Max	1, 4	1, 4	0, 4	0, 4	0, 4	0, 4	1, 4	0, 4

ODS at Day 1								
Mean (SD)	67.5 (15.7)	66.4 (14.8)	68.0 (17.9)	68.8 (16.6)	68.5 (18.2)	68.3 (17.8)	74.3 (13.0)	74.1 (12.7)
Median	69	67.5	70.0	71	71.0	70.0	74.3	74.0
Min, Max	3, 96	20, 97	11, 100	11, 100	12, 100	1, 99	47.3, 100	46, 100
Corneal Staining Inferior at Day 1								
Mean (SD)	2.19 (0.64)	2.38 (0.67)	2.1 (0.70)	2.1 (0.66)	2.1 (0.65)	2.2 (0.65)	2.2 (0.64)	2.2 (0.65)
Median	2	2	2	2	2	2	2	2
Min, Max	0, 3	1, 3	0, 3	0, 3	0, 3	0, 3	0, 3	0, 3
Corneal Staining Nasal at Day 1								
Mean (SD)	1.68 (0.7)	1.77 (0.69)	1.7 (0.70)	1.6 (0.69)	1.6 (0.68)	1.6 (0.67)	1.7 (0.66)	1.76 (0.69)
Median	2	2	2	2	2	2	2	2
Min, Max	0, 3	1, 3	0, 3	0, 3	0, 3	0, 3	0, 3	0, 3

Source: Table 10 of Summary of Clinical Efficacy.

3.2.4 Results and Conclusions

3.2.4.1 Primary Symptom Endpoint and Supporting Results

Table 7 presents the efficacy results for the pre-defined primary symptom endpoint in the four studies. For the pre-defined primary symptom endpoint, both Studies 006 and 011 demonstrated superiority of the LE 0.25% to the vehicle. In Study 006, the changes from baseline in ODS at Day 15 using the 3-day mean in the ITT population were -14.5 and -9.2 for the LE and vehicle groups, respectively; the estimated treatment difference was -5.43 (95% CI: -7.92, -2.95). Analysis of this endpoint in the ITT subgroup of ODS $\geq$ 68 mm 1 day prior to Day 1 yielded similar results (treatment difference = -5.9; 95% CI: -9.35, -2.44). In Study 011, the changes from baseline in ODS at Day 15 using the 3-day mean in the ITT population were -13.6 and -8.9 for the LE and vehicle groups, respectively; the estimated treatment difference was -4.67 (95% CI: -7.08, -2.26). Analysis of this endpoint in the ITT subgroup of ODS $\geq$ 68 mm 1 day prior to Day yielded similar results (treatment difference = -5.44; 95% CI: -8.58, -2.30). The results of this endpoint of both studies were consistent.

Although Study 002 failed to demonstrate superiority of its primary symptom endpoint of mean ODS at Day 29 using 7-day average, post-hoc analysis of the symptom endpoint that was used in the Phase 3 studies (ODS at Day 15 using 3-day average) showed nominal significance favoring the LE 0.25% group. The results of Study 007 also numerically favored the LE group although it failed to demonstrate superiority. When comparing across the three Phase 3 studies, the mean change in ODS at Day 15 in the vehicle group were similar among the three studies, while the mean change in ODS in the LE group in Study 007 was lower than that in the other two Phase 3 studies.

Analyses of the primary endpoints in the PP population yielded similar results (Appendix 4). The applicant did not conduct any sensitivity analyses for missing data because the amount of missing was less than 3%.

Table 7: Results for the Primary Symptom Endpoint (ITT; Observed)

Study	Endpoint	LE 0.25%	Vehicle	LE 0.25% – Vehicle	
		Mean (SD)	Mean (SD)	Difference (95% CI) ¹	P-Value
002	Mean ODS at Day 29 using 7-day mean	n = 73 57.1 (23.9)	n = 77 60.5 (19.8)	-2.64 (-8.51, 3.24)	0.3764
	Post-Hoc Analysis: CFB in ODS at Day 15 using 3-day mean	n = 72 -9.0 (17.8)	n = 77 -3.8 (14.7)	-5.3 (-10.52, -0.03)	0.0489
006	CFB in ODS at Day 15 using 3-day mean	n = 455 -14.5 (20.64)	n = 452 -9.2 (17.92)	-5.43 (-7.92, -2.95)	<0.001
	CFB in ODS at Day 15 using 3-day mean in subgroup of ODS ≥ 68mm 1 day prior to Day 1	n = 257 -16.7 (21.94)	n = 264 -10.8 (18.04)	-5.9 (-9.35, -2.44)	<0.001
007	CFB in ODS at Day 15 using 3-day mean	n = 445 -11.1 (19.9)	n = 444 -9.2 (17.0)	-1.87 (-4.3, 0.55)	0.1298
011	CFB in ODS at Day 15 using 3-day mean	n = 434 -13.6 (19.4)	n = 451 -8.9 (17.6)	-4.67 (-7.08, -2.26)	<0.001
	CFB in ODS at Day 15 using 3-day mean in subgroup of ODS ≥ 68mm 1 day prior to Day 1	n = 296 -15.6 (20.4)	n = 304 -10.2 (18.7)	-5.44 (-8.58, -2.30)	<0.001

CFB: Change from baseline

¹ The estimated difference (95% CI and p-value) was based on an ANCOVA model with treatment, baseline values, and stratification variables.

Source: Table 11 of Study 002 CSR, Table 11 of Study 006 CSR, Table 11 of Study 007 CSR, and Table 11 of Study 011 CSR.

As a secondary efficacy endpoint in Studies 006, 007, and 011, the applicant analyzed the change from baseline of ODS scores on Day 8 using 3-day mean scores as shown in the following table. The change from baseline in ODS at Day 8 were -8.1 and -4.8 for the LE and vehicle group, respectively, in Study 006; were -6.9 and -4.9 for the LE and vehicle group, respectively, in Study 0007; and were -8.1 and -5.8 for the LE and vehicle group, respectively, in Study 011. These results were all supportive of the treatment effect favoring LE.

Table 8: Results for Change from baseline of ODS scores on Day 8 using 3-day mean scores (ITT; Observed)

Study	LE 0.25%	Vehicle	LE 0.25% – Vehicle	
	Mean (SD)	Mean (SD)	Difference (95% CI) ¹	P-Value
002	n = 73 -4.7 (15.5)	n = 77 -1.9 (11.0)	-2.90 (-7.19, 1.39)	0.18
006	n = 457 -8.1 (16.9)	n = 455 -4.8 (14.9)	-3.38 (-5.4, -1.36)	0.0011
007	n = 449 -6.9 (15.3)	n = 451 -4.9 (14.9)	-2.05 (-4.01, -0.09)	0.041
011	n = 443 -8.1 (15.0)	n = 453 -5.8 (15.2)	-2.18 (-4.13, -0.23)	0.028

CFB: Change from baseline

¹ The estimated difference (95% CI and p-value) was based on an ANCOVA model with treatment, baseline values, and stratification variables.

Source: Table 16.1.13.2 of Study 002 Add-Tables, Table 14.2.1.7 of Study 006 CSR, Table 14.2.1.3 of Study 007 CSR, and Table 14.2.1.1 of Study 011 CSR.

The applicant also analyzed daily ODS over time in the four studies. As demonstrated in the table below, the LE group appeared to have lower daily ODS than the vehicle from Days 4 to 15 in all

these studies consistently. Across studies, treatment difference appeared similar among Studies 002, 006 and 011 over time while Study 007 appeared to have less treatment separation.

Table 9: Daily ODS Score (ITT, Observed)

Days	Study 002			Study 006			Study 007			Study 011		
	LE (N=73)	Vehicle (N=77)	Diff (95% CI)	LE (N=459)	Vehicle (N=456)	Diff (95% CI)	LE 0.25% (N=452)	Vehicle (N=453)	Diff (95% CI)	LE 0.25% (N=447)	Vehicle (N=454)	Diff (95% CI)
1	67.5	66.4	0.9 (-2.4, 4.3)	68.0	68.8	-0.5 (-2.2, 1.1)	68.5	68.3	0.1 (-1.6, 1.8)	72.4	72.5	-0.3 (-1.8, 1.2)
2	63.7	66.4	-2.5 (-6.7, 1.7)	65.5	67.5	-1.7 (-3.4, -0.0)	66.5	66.3	0.1 (-1.6, -1.9)	70.8	71.2	-0.6 (-2.3, 1.0)
3	63.6	66.5	-2.7 (-6.8, 1.4)	63.9	66.5	-2.3 (-4.1, -0.5)	65.6	65.6	-0.1 (-1.9, 1.7)	68.6	70.6	-2.1 (-3.8, -0.4)
4	63.7	64.5	-0.6 (-5.1, 4.0)	62.2	66.3	-3.5 (-5.5, -1.5)	63.8	64.7	-1.0 (-2.9, 0.9)	68.5	69.8	-1.4 (-3.2, 0.4)
5	63.0	66.8	-3.5 (-7.9, 0.9)	61.5	65.2	-3.5 (-5.6, 1.3)	62.5	64.2	-3.5 (-5.6, 1.3)	66.4	68.6	-1.5 (-3.4, 0.4)
6	62.5	65.7	-3.0 (-7.7, 1.6)	61.0	64.7	-3.4 (-5.6, -1.3)	61.8	64.0	-2.3 (-4.4, -0.2)	61.0	68.3	-2.0 (-4.1, 0.0)
7	62.6	65.2	-2.4 (-7.1, 2.3)	59.8	63.2	-3.1 (-5.3, -0.9)	61.6	63.6	-2.2 (-4.3, -0.0)	64.9	67.9	-3.1 (-5.3, -0.9)
8	61.1	64.3	-3.1 (-8.1, 1.8)	59.3	62.9	-3.3 (-5.7, -1.0)	60.0	62.5	-2.6 (-4.8, -0.4)	64.1	67.5	-3.4 (-5.7, -1.2)
9	60.2	63.5	-3.5 (-8.6, 1.6)	57.2	62.8	-5.3 (-7.7, -2.9)	59.6	61.7	-2.2 (-4.5, 0.1)	63.1	66.8	-3.6 (-6.0, -1.3)
10	59.5	63.7	-4.0 (-9.0, 1.1)	56.4	61.5	-4.9 (-7.4, -2.5)	58.7	60.8	-2.1 (-4.4, 0.3)	62.3	66.2	-3.9 (-6.2, -1.5)
11	58.9	63.6	-4.4 (-9.8, 0.9)	55.8	61.0	-4.9 (-7.4, -2.4)	58.3	60.9	-2.6 (-5.1, -0.2)	61.6	66.1	-4.6 (-6.9, -2.2)
12	58.8	64.0	-5.0 (-10.1, 0.2)	55.1	60.4	-5.0 (-7.5, -2.5)	57.8	59.8	-2.1 (-4.6, 0.4)	55.1	60.4	-4.9 (-7.3, -2.5)
13	58.6	64.0	-5.1 (-10.7, 0.5)	54.3	60.6	-6.0 (-8.6, -3.4)	57.8	59.8	-2.2 (-4.7, 0.4)	60.9	64.7	-3.8 (-6.2, -1.3)
14	58.0	64.0	-5.8 (-11.4, -0.1)	53.8	59.0	-5.1 (-7.9, -2.3)	57.5	58.9	-1.3 (-3.9, 1.4)	59.5	64.6	-4.9 (-7.5, -2.3)
15	58.3	64.1	-5.1 (-10.7, 0.5)	56.2	62.6	-8.1 (-13.2, -2.9)	51.5	56.0	-2.8 (-8.9, 3.3)	56.5	62.8	-6.4 (-11.7, -1.7)

¹ The estimated treatment differences and 95% CI were based on the ANOCVA model that included treatment, baseline calculated as the mean of the scores over the three days prior to Day 1, and randomization strata. Diff = treatment difference.

Source: Table 16.1.13.1 of Study 002 add-table, Table 14.2.4.1.1 of Study 006 CSR, Table 14.2.4.1.1 of Study 007 CSR, and Table 14.2.2.4.1 of Study 011 CSR.

3.2.4.2 Primary Sign Endpoint

Table 10 presents the efficacy results for the pre-defined primary sign endpoint in Study 002, the two Phase 3 studies and the pre-defined first secondary endpoint in newly completed Study 011. All four studies demonstrated superiority of the LE 0.25% to the vehicle for the primary sign endpoints of conjunctival hyperemia:

- In Study 002, the mean conjunctival hyperemia scores at Day 29 for the LE 0.25% and vehicle groups were 1.4 and 1.6, respectively; the estimated difference was -0.21 (95% CI: -0.41, -0.01).
- In Study 006, the changes from baseline in conjunctival hyperemia at Day 15 for the LE 0.25% and vehicle groups were -0.4 and -0.2, respectively; the estimated difference was -0.25 (95% CI: -0.33, -0.18). Inferior corneal staining was analyzed as one of the primary sign endpoints only in Study 006. The treatment difference for this endpoint numerically favored the LE group.
- In Study 007, the changes from baseline in conjunctival hyperemia at Day 15 for the LE 0.25% and vehicle groups were -0.4 and -0.2, respectively; the estimated difference was -0.16 (95% CI: -0.23, -0.09).
- In Study 011, the changes from baseline in conjunctival hyperemia at Day 15 for the LE 0.25% and vehicle groups were -0.4 and -0.2, respectively; the estimated difference was -0.18 (95% CI: -0.24, -0.12).

Table 10: Results for the Primary Sign Endpoint (ITT; Observed)

Study	Endpoint	LE 0.25%	Vehicle	LE 0.25% – Vehicle	
		Mean (SD)	Mean (SD)	Difference (95% CI)	P-value
002	Mean Conjunctival hyperemia at Day 29	n = 72 1.4 (0.67)	n = 77 1.6 (0.59)	-0.21 (-0.41, -0.01)	0.0387
006	CFB in conjunctiva hyperemia at Day 15	n = 452 -0.4 (0.65)	n = 451 -0.2 (0.61)	-0.25 (-0.33, -0.18)	<0.001
	CFB in inferior corneal staining at Day 15	n = 455 -0.5 (0.84)	n = 452 -0.4 (0.79)	-0.08 (-0.17, 0.02)	0.1128
007	CFB in conjunctival hyperemia at Day 15	n = 446 -0.4 (0.61)	n = 447 -0.2 (0.62)	-0.16 (-0.23, -0.09)	<0.001
011	CFB in conjunctival hyperemia at Day 15	n = 432 -0.4 (0.58)	n = 449 -0.2 (0.55)	-0.18 (-0.24, -0.12)	<0.001

¹ The estimated difference (95% CI and p-value) was based on an ANCOVA model with treatment, baseline values, and stratification variables. Source: Table 10 of Study 002 CSR, Table 11 of Study 006 CSR, Table 11 of Study 007 CSR, and Table 12 of Study 011 CSR.

Analyses of the primary sign endpoints in the PP population yielded similar results (Appendix 4). The applicant did not conduct any sensitivity analyses for missing data because the amount of missing was less than 3%.

3.2.4.3 Conclusion

All the four studies (002, 006, 007, 011) demonstrated superiority of the LE 0.25% to the vehicle for the pre-defined primary (and the first secondary in Study 011) sign endpoint of conjunctival hyperemia at Day 15.

For the primary symptom endpoint, two studies (006 and 011) among the four studies demonstrated superiority of the LE 0.25% to the vehicle for the pre-defined primary symptom endpoint – change from baseline in ODS at Day 15 using 3-day means; and the other two studies (002 and 007) showed numeric trend in favor of the LE group. Although Study 002 failed to meet the pre-defined primary symptom endpoint, it was able to meet the primary endpoint of the three Phase 3 studies nomically in the post-hoc analysis (p-value = 0. 0489). Therefore, the statistical reviewer concludes that there is enough evidence to support the treatment effect for the symptom endpoint – the change from baseline in ODS scores at Day 15.

3.3 Evaluation of Safety

As presented in the following table, based on the pooled safety analysis, the LE group had slightly more incidents of all AEs, ocular AEs, severe AEs, treatment-related AEs, AEs leading to discontinuation of study treatment, SAEs. There was no death reported in any of the four clinical studies.

Table 11: Summary of Adverse Events (Pooled Safety Population)

Category	LE QID (N=1430)	Vehicle (N=1438)
Any AEs	185 (12.9%)	150 (10.4%)
Ocular AEs	135 (9.4%)	126 (8.8%)
Severe AEs	8 (0.6%)	1 (0.1%)
AEs Related to Treatment	121 (8.5%)	103 (7.2%)
AEs Leading to Discontinuation of Study Treatment	10 (0.7%)	5 (0.3%)
Serious AEs	7 (0.5%)	1 (0.1%)

Source: Table 12 of Summary of Clinical Safety.

Table 12 below presented the adverse events associated with ≥ 2 subjects in either treatment group for each of the four studies.

Table 12: Safety Analysis: Adverse Events Associated with ≥ 2 Subjects in Either Treatment Group by System Organ Class and Preferred Term (Studies 002, 006, 007, and 011, Safety Population)

System Organ Class Preferred Term	Study 002		Study 006		Study 007		Study 011	
	LE (N=72)	Vehicle (N=78)	LE (N=459)	Vehicle (N=456)	LE (N=452)	Vehicle (N=453)	LE (N=449)	Vehicle (N=452)
Eye Disorders	5 (6.9%)	6 (7.7%)	20 (4.4%)	26 (5.7%)	15 (3.3%)	18 (4.0%)	10 (2.2%)	17 (3.8%)
Eye Irritation	1 (1.4%)	1 (1.3%)	5 (1.1%)	7 (1.5%)	2 (0.4%)	3 (0.7%)	2 (0.4%)	1 (0.2%)
Vision Blurred	1 (1.4%)	0	1 (0.2%)	2 (0.4%)	1 (0.2%)	6 (1.3%)	2 (0.4%)	0
Dry Eye	0	2 (2.6%)	n/a	n/a	n/a	n/a	1 (0.2%)	3 (0.7%)
Eye pruritus	1 (1.4%)	1 (1.3%)	1 (0.2%)	4 (0.9%)	3 (0.7%)	4 (0.9%)	1 (0.2%)	3 (0.7%)
Conjunctival hyperemia	n/a	n/a	2 (0.4%)	4 (0.9%)	2 (0.4%)	3 (0.7%)	n/a	n/a
Eye pain	0	1 (1.3%)	2 (0.4%)	1 (0.2%)	2 (0.4%)	3 (0.7%)	n/a	n/a
Conjunctival hemorrhage	n/a	n/a	n/a	n/a	1 (0.2%)	2 (0.4%)	n/a	n/a
Blepharitis	n/a	n/a	3 (0.7%)	0	n/a	n/a	n/a	n/a
Lacrimation increased	n/a	n/a	3 (0.7%)	2 (0.4%)	n/a	n/a	n/a	n/a
Foreign body sensation	n/a	n/a	2 (0.4%)	4 (0.9%)	n/a	n/a	n/a	n/a

Eyelid margin crusting	n/a	n/a	1 (0.2%)	2 (0.4%)	n/a	n/a	n/a	n/a
Eye discharge	0	2 (2.6%)					n/a	n/a
General disorder and administration site conditions	8 (11.1%)	8 (10.3%)	31 (6.8%)	32 (7.0%)	28 (6.2%)	22 (4.9%)	15 (2.3%)	7 (1.5%)
Instillation site pain	7 (9.7%)	8 (10.3%)	28 (6.1%)	28 (6.1%)	26 (5.8%)	20 (4.4%)	13 (2.9%)	7 (1.5%)
Instillation foreign body sensation	n/a	n/a	2 (0.4%)	0	n/a	n/a	n/a	n/a
Instillation site pruritus	1 (1.4%)	1 (1.3%)	1 (0.2%)	3 (0.7%)	n/a	n/a	n/a	n/a
Investigations			3 (0.7%)	1 (0.2%)	n/a	n/a	2 (0.4%)	0
Blood Pressure Increased	n/a	n/a	n/a	n/a	n/a	n/a	2 (0.4%)	0
Intraocular pressure increased	1 (1.4%)	1 (1.3%)	3 (0.7%)	1 (0.2%)	n/a	n/a	n/a	n/a
Nervous system disorders	n/a	n/a	4 (0.9%)	2 (0.4%)	2 (0.4%)	4 (0.9%)	4 (0.9%)	1 (0.2%)
Headache	n/a	n/a	2 (0.4%)	1 (0.2%)	1 (0.2%)	2 (0.4%)	2 (0.4%)	1 (0.2%)
Infections and infestations	n/a	n/a	3 (0.7%)	9 (2.0%)	3 (0.7%)	6 (1.3%)	n/a	n/a
Upper respiratory tract infection	n/a	n/a	1 (0.2%)	2 (0.4%)	0	2 (0.4%)	n/a	n/a
Nasopharyngitis	n/a	n/a	0	3 (0.7%)	n/a	n/a	n/a	n/a
Musculoskeletal and Connective tissue Disorders	n/a	n/a	n/a	n/a	n/a	n/a	2 (0.2%)	0
Neck pain	n/a	n/a	n/a	n/a	n/a	n/a	2 (0.2%)	0
Gastrointestinal disorders	2 (2.8%)	0	n/a	n/a	n/a	n/a	n/a	n/a
Toothache	2 (2.8%)	0	n/a	n/a	n/a	n/a	n/a	n/a

Source: Table 18 of Study 002 CSR, Table 33 of Study 006 CSR, Table 26 of Study 007 CSR, and Table 24 of Study 011 CSR.

The pooled treatment-related adverse events are summarized in the following table. The most frequently reported treatment-related AE was instillation site pain, which was reported by 5.2% of subjects in the LE 0.25% QID group and 4.4% of subjects in the vehicle QID group. The applicant reported that all other treatment-related AEs were $\leq 0.4\%$ of subjects in the LE 0.25% QID group.

Table 13: Number (%) of Subjects Reporting Common ($\geq 0.3\%$ in Either Treatment Group) Treatment-Related Adverse Events (Pooled Safety Population)

Category	LE QID (N=1430)	Vehicle (N=1438)
Number (%) of Subjects Reporting Any Treatment-related AEs	121 (8.5%)	103 (7.2%)
General Disorders and Administration Site Conditions	81 (5.7%)	68 (4.7%)
Instillation Site Pain	74 (5.2%)	63 (4.4%)
Instillation Site Pruritis	3 (0.2%)	5 (0.3%)
Eye Disorders	33 (2.3%)	42 (2.9%)
Eye Irritation	6 (0.4%)	10 (0.7%)

Eye Pruritus	6 (0.4%)	6 (0.4%)
Vision Blurred	5 (0.3%)	8 (0.6%)
Conjunctival Hyperemia	4 (0.3%)	2 (0.1%)
Dry Eye	0	4 (0.3%)
Investigations	4 (0.3%)	2 (0.1%)
Intraocular Pressure Increased	4 (0.3%)	2 (0.1%)

Source: Table 14 of Summary of Clinical Safety.

For a comprehensive review of safety, please refer to Dr. William Boyd's clinical review.

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Race, and Age

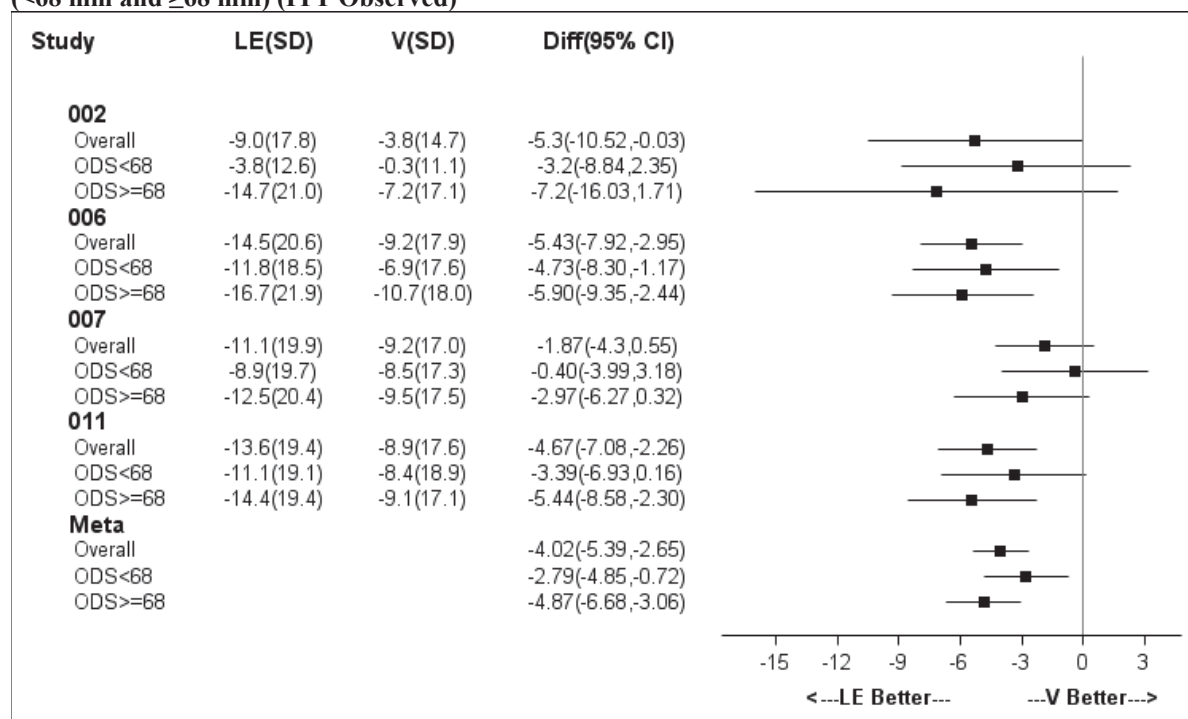
Subgroup analyses based on gender, race, and age for each study were performed (see results in Appendix 5) by the applicant. In all studies, the subgroup analyses results were similar to those seen for the overall population for each demographic subgroup.

4.2 ODS Scores 1 Day Prior to Day 1

As one of the primary efficacy endpoints for Studies 006 and 011, the applicant analyzed and reported change from baseline in ODS at Day 15 using 3-day mean in subgroup of ODS ≥ 68 mm 1 day prior to Day 1. The statistical reviewer further analyzed the change from baseline in ODS at Day 15 using 3-day mean based on ODS scores 1 day prior to Day 1 (<68 mm and ≥ 68 mm) for all the four studies. As presented in the following figure, the analyses results were also consistent with those seen for the overall population for each study.

Since the change from baseline of ODS at Day 15 showed consistently higher treatment effects for the subgroup of subjects who had ODS ≥ 68 mm 1 day prior to Day 1, the statistical reviewer conducted further exploratory meta-analysis using the inverse-variance method to investigate the treatment effects for these two subgroups. As demonstrated in the last three lines of the following figure, although both subgroups demonstrated treatment effects of LE 0.25 over vehicle, it appears that treatment effect for subjects who had lower ODS scores (<68 mm) 1 day prior to Day 1 was slightly lower than those who had higher ODS scores (≥ 68 mm) 1 day prior to Day 1; the difference was about 2 points on the VAS.

Figure 4: Change from Baseline in ODS on Day 15 Using 3-day Mean by ODS Score 1 Day Prior to Day 1 (<68 mm and ≥68 mm) (ITT Observed)



Note: ITT = Intent-to-Treat, LE = LE 0.25%, V = Vehicle, SD = Standard Deviation, Diff = Difference, CI = Confidence Interval; Meta=Meta-Analysis.

The estimated difference (95% CI and p-value) for each study was based on an ANCOVA model with treatment, baseline values, and stratification variables; the meta-analysis estimates were based on inverse-variance method.

Source: Table 11 of Study 006 CSR, and Table 11 of Study 011 CSR, and the statistical reviewer's analysis.

5 SUMMARY AND CONCLUSIONS

5.1 Statistical Issues

There are no major statistical issues identified for this NDA submission.

5.2 Collective Evidence

All the four studies (002, 006, 007, 011) demonstrated superiority of the LE 0.25% to the vehicle for the pre-defined primary (and the first secondary in Study 011) sign endpoint of conjunctival hyperemia.

For the primary symptom endpoint, two studies (006 and 011) among the four studies demonstrated superiority of the LE 0.25% to the vehicle for the pre-defined primary symptom endpoint; and the other two studies (002 and 007) showed numeric trend in favor of the LE group. Although Study 002 failed to meet the pre-defined primary symptom endpoint, it was able to meet the primary endpoint of the three Phase 3 studies in the post-hoc analysis (p-value = 0.0489).

Subgroup analyses showed a consistent treatment effect of the LE 0.25% QID across subgroups of sex, age, and race.

5.3 Conclusions and Recommendations

In conclusion, collectively the four studies provided statistically significant evidence that LE 0.05% is superior to Vehicle in treating both dry eye sign and symptom, the statistical reviewer recommended the approval of LE 0.25% for the short-term treatment of the signs and symptoms of dry eye disease.

5.4 Labeling Recommendations

In the NDA resubmission, the applicant's proposed label had the following text for the clinical studies section.

"The safety and efficacy of EYSUVIS for the treatment of dry eye disease was assessed (b) (4) patients with (b) (4) dry eye disease. Patients received either EYSUVIS or vehicle (1:1 ratio) four times a day for (b) (4) 2 weeks in 4 multi-centered, randomized, double-masked, placebo-controlled trials. The use of artificial tears was not allowed during the trials. (b) (4)

(b) (4)

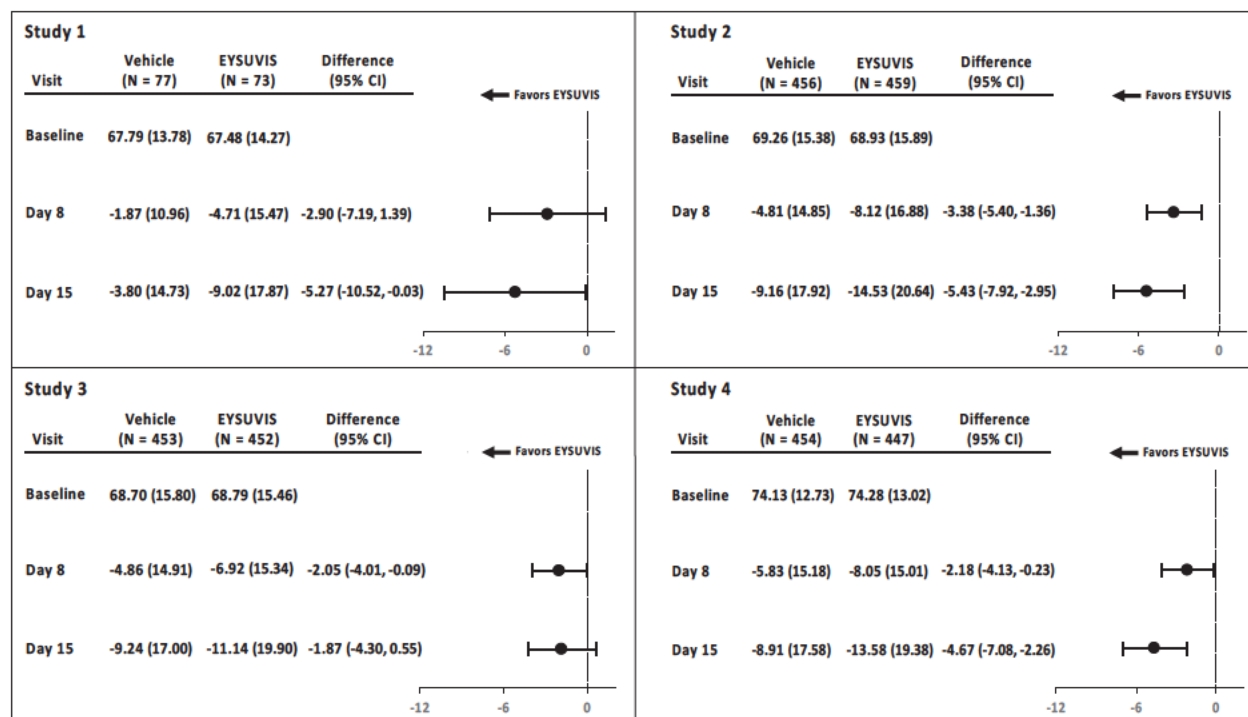
Effects on Symptoms of Dry Eye Disease

Ocular discomfort severity (ODS) was rated by patients daily over the course of the trial using a visual analog scale (0 = very mild, 100 = very severe). (b) (4)

(b) (4) A larger reduction in ocular discomfort severity favoring EYSUVIS was observed at (b) (4) Day 15 in (b) (4) patient population (see Figure 1) (b) (4)

(b) (4)

Figure 1: Mean Change (SD) from Baseline and Treatment Difference (EYSUVIS – Vehicle) in Ocular Discomfort Severity Score in Patients with Dry Eye Disease



Treatment differences between the **EYSUVIS** and vehicle groups are displayed for each study, based on least square means and 2-sided confidence intervals for the change from baseline.

(b) (4)

(b) (4)

Appendix 1:

(b) (4)

(b) (4)

3 Page(s) have been Withheld in Full as B4 (CCI/TS) immediately following
this page

Appendix 4: Results of the Pre-defined Primary Sign and Symptom Endpoints in the PP Population

Table 16: Results for the Primary Symptom Endpoint (PP; Observed)

Study	Endpoint	LE 0.25%	Vehicle	LE 0.25% vs. Vehicle	
		Mean (SD)	Mean (SD)	Difference	P-value
002	Mean ODS at Day 29	n = 66 55.8 (23.6)	n = 67 60.1 (20.4)	-2.6	0.43
006	CFB in ODS at Day 15	n = 410 -14.67 (20.95)	n = 409 -9.66 (18.09)	-5.12	<0.001
	CFB in ODS at Day 15 in subgroup of ODS $\geq$ 68 mm 1 day prior to Day 1	n = 233 -16.87 (22.11)	n = 238 -11.13 (18.58)	-5.74	0.002
007	CFB in conjunctival hyperemia at Day 15	n = 416 -11.09 (19.76)	n = 418 -9.18 (17.11)	-1.89	0.14
011	CFB in conjunctival hyperemia at Day 15	n = 379 -13.07 (19.15)	n = 403 -8.67 (16.82)	-4.42	<0.001
	CFB in ODS at Day 15 in subgroup of ODS $\geq$ 68 mm 1 day prior to Day 1	n = 255 -15.35 (20.16)	n = 270 -9.69 (17.63)	-5.69	<0.001

Source: Table 14.2.12.2 of Study 002 CSR, Tables 14.2.3.2 and 14.2.3.3 of Study 006 CSR, Table 14.2.3.2 of Study 007 CSR, Tables 14.2.4.2.1 and 14.2.4.2.2 of Study 011 CSR.

Table 17: Results for the Primary Sign Endpoint (PP; Observed)

Study	Endpoint	LE 0.25%	Vehicle	LE 0.25% – Vehicle	
		Mean (SD)	Mean (SD)	Difference (95% CI)	P-value
002	Mean Conjunctival hyperemia at Day 29	n = 65 1.48 (0.64)	n = 67 1.67 (0.53)	-0.24	0.016
006	CFB in conjunctiva hyperemia at Day 15	n = 410 -0.42 (0.66)	n = 409 -0.18 (0.60)	-0.25	<0.001
	CFB in inferior corneal staining at Day 15	n = 410 -0.5 (0.83)	n = 409 -0.4 (0.76)	-0.10	0.06
007	CFB in conjunctival hyperemia at Day 15	n = 446 -0.4 (0.61)	n = 447 -0.2 (0.62)	-0.15	<0.001
011	CFB in conjunctival hyperemia at Day 15	n = 432 -0.4 (0.58)	n = 449 -0.2 (0.55)	-0.18 (-0.24, -0.12)	<0.001

Source: Table 14.2.11.2 of Study 002 CSR, Tables 14.2.3.1 and 14.2.3.4 of Study 006 CSR, Table 14.2.3.1 of Study 007 CSR, Tables 14.2.4.3.1 of Study 011 CSR.

Appendix 5: Subgroup Analyses for the Phase 3 Studies

Table 18: Subgroup Analysis of Change from Baseline in Conjunctival Hyperemia on Day 15 by Age, Gender, and Race for All Phase 3 Studies (ITT Observed)

		Study 006			Study 007			Study 011		
		LE 0.25%	Vehicle	Difference (95% CI) ¹	LE 0.25%	Vehicle	Difference (95% CI) ¹	LE 0.25%	Vehicle	Difference (95% CI) ¹
Age	<65	n = 280 -0.38 (0.68)	n = 284 -0.18 (0.65)	-0.21 (-0.31, -0.11)	n = 263 -0.37 (0.62)	n = 265 -0.26 (0.65)	-0.14 (-0.24, -0.04)	n = 262 -0.37 (0.59)	n = 279 -0.20 (0.55)	-0.18 (-0.26, -0.10)
	≥65	n = 172 -0.45 (0.60)	n = 167 -0.11 (0.53)	-0.32 (-0.43, -0.20)	n = 183 -0.39 (0.59)	n = 182 -0.20 (0.58)	-0.19 (-0.30, -0.08)	n = 170 -0.33 (0.55)	n = 170 -0.15 (0.54)	-0.18 (-0.28, -0.08)
Gender	Male	n = 90 -0.43 (0.74)	n = 96 -0.11 (0.68)	-0.27 (-0.46, -0.08)	n = 117 -0.41 (0.64)	n = 97 -0.15 (0.63)	-0.32 (-0.47, -0.16)	n = 105 -0.41 (0.62)	n = 123 -0.23 (0.60)	-0.21 (-0.35, -0.08)
	Female	n = 362 -0.40 (0.63)	n = 355 -0.17 (0.59)	-0.24 (-0.32, -0.16)	n = 329 -0.37 (0.59)	n = 355 -0.26 (0.62)	-0.12 (-0.21, -0.03)	n = 327 -0.33 (0.56)	n = 326 -0.17 (0.53)	-0.16 (-0.24, -0.09)
Race	White	n = 355 -0.40 (0.66)	n = 356 -0.17 (0.60)	-0.24 (-0.32, -0.16)	n = 339 -0.37 (0.60)	n = 348 -0.24 (0.61)	-0.14 (-0.22, -0.06)	n = 338 -0.38 (0.58)	n = 338 -0.19 (0.56)	-0.20 (-0.27, -0.13)
	Non-White	n = 97 -0.41 (0.61)	n = 95 -0.11 (0.64)	-0.29 (-0.46, -0.12)	n = 107 -0.43 (0.63)	n = 99 -0.23 (0.67)	-0.22 (-0.39, -0.05)	n = 94 -0.27 (0.57)	n = 111 -0.17 (0.50)	-0.11 (-0.24, -0.02)

¹ The estimated difference (95% CI and p-value) was based on an ANCOVA model with treatment, baseline values, and stratification variables.
Source: Tables 14.2.4.6.3, 14.2.4.6.4, 14.2.4.6.5 of Study 006 CSR; Tables 14.2.4.6.1, 14.2.4.6.2, 14.2.4.6.3 of Study 007 CSR; Tables 14.2.2.5.1, 14.2.2.5.2, 14.2.2.5.3, 14.2.2.5.4, 14.2.2.5.5, 14.2.2.5.6 of Study 011 CSR.

Table 19: Subgroup Analysis of Change from Baseline in ODS on Day 15 Using 3-day Mean by Age, Gender, and Race for All Phase 3 Studies (ITT Observed)

		Study 006			Study 007			Study 011		
		LE 0.25%	Vehicle	Difference (95% CI) ¹	LE 0.25%	Vehicle	Difference (95% CI) ¹	LE 0.25%	Vehicle	Difference (95% CI) ¹
Age	<65	n = 280 -15.32 (20.60)	n = 286 -8.20 (17.86)	-7.22 (-10.35, -4.09)	n = 262 -10.08 (20.37)	n = 261 -9.54 (16.91)	-0.26 (-3.46, 2.94)	n = 264 -13.50 (19.45)	n = 281 -8.22 (16.90)	-5.16 (-8.19, -2.12)
	≥65	n = 175	n = 166	-2.59	n = 183	n = 183	-4.12	n = 170	n = 170	-3.54

		-13.3 (20.70)	-10.82 (17.98)	(-6.71, - 1.54)	-12.66 (19.16)	-8.82 (17.17)	(-7.86, - 0.37)	-13.70 (19.32)	-10.04 (18.65)	(-7.52, - 0.44)
Gender	Male	n = 92 -15.53 (20.38)	n = 96 -8.88 (14.07)	-6.24 (-11.28, - 1.20)	n = 118 -10.85 (21.10)	n = 96 -8.32 (15.39)	-2.65 (-7.75, - 2.46)	n = 106 -11.07 (19.05)	n = 123 -8.43 (18.85)	-2.75 (-7.62, - 2.13)
	Female	n = 363 -14.27 (20.72)	n = 356 -9.24 (18.85)	-5.08 (-7.95, - 2.22)	n = 327 -11.24 (19.48)	n = 348 -9.50 (17.43)	-1.76 (-4.54, - 1.02)	n = 328 -14.39 (19.44)	n = 328 -9.09 (17.11)	-5.31 (-8.09, - 2.52)
Race	White	n = 358 -14.00 (19.90)	n = 357 -8.52 (17.79)	-5.60 (-8.33, - 2.87)	n = 339 -11.32 (19.58)	n = 345 -9.03 (17.26)	-2.33 (-5.09, - 0.43)	n = 340 -14.64 (19.90)	n = 339 -10.15 (18.48)	-4.54 (-7.39, - 1.69)
	Non- White	n = 97 -16.46 (23.15)	n = 95 -11.59 (18.32)	-4.70 (-10.63, - 1.23)	n = 106 -10.54 (20.99)	n = 99 -9.98 (16.12)	-0.87 (-6.03, - 4.29)	n = 94 -9.72 (16.92)	n = 112 -5.14 (13.90)	-4.48 (-8.72, - 0.23)

¹ The estimated difference (95% CI and p-value) was based on an ANCOVA model with treatment, baseline values, and stratification variables.
Source: Tables 14.2.4.5.3, 14.2.4.5.4, 14.2.4.5.5 of Study 006 CSR; Tables 14.2.4.5.1, 14.2.4.5.2, 14.2.4.5.3 of Study 007 CSR; Tables 14.2.1.3.1, 14.2.1.3.2, 14.2.1.3.3, 14.2.1.3.4, 14.2.1.3.5, 14.2.1.3.6 of Study 011 CSR.

This is a representation of an electronic record that was signed electronically. Following this are manifestations of any and all electronic signatures for this electronic record.

/s/

YUNFAN DENG
08/14/2020 05:33:11 PM

GUOXING SOON
08/15/2020 02:16:13 PM



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

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/Serial #: NDA 210933/Original

Drug Name: Eysuvis (loteprednol etabonate ophthalmic suspension 0.25%)

Indication(s): Temporary relief of the signs and symptoms of dry eye disease

Applicant: Kala Pharmaceuticals, Inc.

Date(s): Stamp Date: 10/15/2018
PDUFA Date: 08/15/2019
Review Date: 05/7/2019

Review Priority: Standard

Biometrics Division: IV

Statistical Reviewer: Qing Zhou, Ph.D.

Concurring Reviewers: Yan Wang, Ph.D., Team Leader

Medical Division: Division of Transplant and Ophthalmology Products (DTOP)

Clinical Team: William Boyd, MD, Reviewer and Team Leader

Project Manager: Dheera Semidey

Keywords:

Temporary relief of the signs and symptoms of dry eye disease, ocular discomfort severity, conjunctival hyperemia scores

Table of Contents

TABLE OF CONTENTS	2
LIST OF TABLES.....	3
1 EXECUTIVE SUMMARY	5
2 INTRODUCTION	8
2.1 OVERVIEW	8
2.1.1 Background.....	8
2.1.2 History of Drug Development.....	8
2.1.3 Specific Study Reviewed.....	10
2.2 DATA SOURCES	12
3 STATISTICAL EVALUATION	12
3.1 DATA AND ANALYSIS QUALITY.....	12
3.2 EVALUATION OF EFFICACY.....	13
3.2.1 Study Design and Endpoints	13
3.2.2 Statistical Methodologies.....	15
3.2.3 Patient Disposition, Demographic and Baseline Characteristics	16
3.2.4 Results and Conclusion.....	18
3.3 SAFETY EVALUATION	31
4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS.....	32
5 SUMMARY AND CONCLUSIONS	34
5.1 STATISTICAL ISSUES	34
5.2 COLLECTIVE EVIDENCE	34
5.3 CONCLUSIONS AND RECOMMENDATIONS	35
APPENDICES.....	36
APPENDIX 1: (b) (4)	36
APPENDIX 2:	38
APPENDIX 3:	39
APPENDIX 4:	40
APPENDIX 5A: RESULTS FOR THE PRE-DEFINED PRIMARY SIGN ENDPOINTS IN THE PP POPULATION (ITT; OBSERVED).....	44
APPENDIX 5B: RESULTS FOR THE PRE-DEFINED PRIMARY SYMPTOM ENDPOINTS IN THE PP POPULATION (ITT; OBSERVED).....	44
APPENDIX 6: RESULTS FOR INVESTIGATOR-RATED CONJUNCTIVAL HYPEREMIA (ITT; OBSERVED).....	44
APPENDIX 7: DAILY ODS FROM DAYS 1 TO 14 IN ITT SUBGROUP WITH MORE SEVERE BASELINE ODS DEFINED AS ODS $\geq$ 68 MM 1 DAY PRIOR TO DAY 1 IN THE THREE STUDIES (OBSERVED)	45
APPENDIX 8: 3-DAY MEAN ODS FROM DAYS 8 TO 15 IN ITT SUBGROUP WITH MORE SEVERE BASELINE ODS DEFINED AS BASELINE ODS $\leq$ 68 MM (OBSERVED).....	46
APPENDIX 9: CHANGE FROM BASELINE IN ODS USING THE 3-DAY MEAN AT DAY 15 IN MULTIPLE EXPLORATORY MITT POPULATION (OBSERVED)	47
APPENDIX 10: KEY INCLUSION CRITERIA FOR STUDY KPI-121-C-011	48

LIST OF TABLES

Table 1: Primary Endpoints in the Three Studies	5
Table 2: Summary of Studies 002, 006, and 007	11
Table 3: Summary of Analysis Population (All Randomized Subject IDs)	17
Table 4: Subject Disposition (ITT)	17
Table 5: Demographic and Baseline Characteristics (ITT)	17
Table 6: Results for the Primary Sign Endpoints (ITT; Observed)	19
Table 7: Results for the Primary Symptom Endpoints (ITT; Observed)	19
Table 8: Mean Change from Baseline in Conjunctival Hyperemia and ODS (3-Day Mean) at Day 15 in Study 002 (ITT; Observed)	20
Table 9: Change from Baseline in Eye Dryness Scores and Ocular Discomfort Frequency at Days 8 and 15 (ITT; Observed)	29
Table 10: Summary of All AEs (Pooled Safety Population)	32
Table 11: Number of Subjects Reporting Common ($\geq 1\%$ in either treatment group) AEs (Pooled Safety Population)	32

LIST OF FIGURES

Figure 1: Change from Baseline in Conjunctival Hyperemia and ODS using the 3-day mean at Day 15 (ITT; Observed).....	6
Figure 2: Daily ODS using a single day score from Days 1 to 15 (ITT; Observed)	7
Figure 3: Study Design for Studies 002, 006, and 007	13
Figure 4: Forest Plot for Proportion of Subjects with ≥ 1 Grade Improvement from Baseline to Day 15 in Conjunctival Hyperemia, Proportion of Subjects with Grades 0 or 1 on Conjunctival Hyperemia at Day 15 and Proportion of Subjects with Grade 0 on Conjunctival Hyperemia at Day 15 (ITT; Observed)	21
Figure 5: Forest Plot for Change from Baseline in Mean Conjunctival Hyperemia of All Regions at Day 15, Changes from Baseline in Conjunctival Hyperemia at Day 15 in Subjects with ODS of ≥ 68 mm One Day Prior to Day 1 and Subjects with Baseline Conjunctival Hyperemia ≥ 2 , and Changes from Baseline in Inferior and Nasal Corneal Fluorescein Staining at Day 15 (ITT and subgroups of ITT; Observed).....	23
Figure 6: Daily ODS using the 3-Day Mean from Days 5 to 15 (ITT; Observed)	26
Figure 7: Forest Plot for Change from Baseline in ODS at Day 15 using the 7-Day Mean and Change from Baseline ODS at Day 15 using the 3-Day Mean in Subjects with More Severe Baseline Ocular Discomfort (ITT and Subgroups of ITT; Observed)	28
Figure 8: Subgroup Analyses for Primary Sign and Symptom Endpoints (ITT pooled).....	33

1 EXECUTIVE SUMMARY

The applicant seeks approval of Eysuvis (loteprednol etabonate ophthalmic suspension [LE] 0.25%) dosed four times daily (QID) for the temporary relief of the signs and symptoms of dry eye disease. The applicant has conducted one Phase 2 study (Study 002) and two Phase 3 studies (Studies 006 and 007) to support the efficacy and safety of LE 0.25% for this indication.

All three studies were randomized, multicenter, double-masked, vehicle-controlled studies. The Phase 2 study was 29 days in duration and consisted of six visits: Days -14 (screening), 1 (baseline/randomization), 8, 15, 22, and 29. The two Phase 3 studies were 15 days in duration and consisted of four visits: Days -14 (screening), 1 (baseline/randomization), 8, and 15. At Day 1, the Phase 2 study randomized 150 subjects in a 1:1 ratio to receive either LE 0.25% or vehicle for 28 days; the two Phase 3 studies each randomized approximately 910 subjects in a 1:1 ratio to receive either LE 0.25% or vehicle for 14 days.

The primary sign and symptom endpoints for the three studies are listed below in Table 1. Conjunctival hyperemia scores were 5-level scales from 0 (none) to 4 (severe); corneal staining scores were 4-level scales from 0 (none) to 3 (severe); and ocular discomfort scores (ODS) was a visual analog scale from 0 to 100 mm, where 0 indicated “very mild” and 100 indicated “very severe”.

Table 1: Primary Endpoints in the Three Studies

	Study 002	Study 006	Study 007
Sign	Mean conjunctival hyperemia scores in the worst region of the study eye ^a at Day 29	Change from baseline in conjunctival hyperemia scores in the worst region of the study eye ^a at Day 15 Change from baseline in inferior corneal staining scores in the study eye at Day 15	Change from baseline in conjunctival hyperemia scores in the worst region of the study eye ^a at Day 15
Symptom	Mean ODS at Day 29 using the 7-day mean ^b	Change from baseline in ODS at Day 15 using the 3-day mean ^c Change from baseline in ODS at Day 15 using the 3-day mean ^c in subgroup of patients with ODS $\geq$ 68 mm one day prior to Day 1	Change from baseline in ODS at Day 15 using the 3-day mean ^c

^a The definitions of the study eye and the worst region of the study eye are shown in Appendix 1.

^b In Study 002, the primary symptom endpoint was the mean ODS scores over the 7 days prior to Day 29 visit (using the 7-day mean).

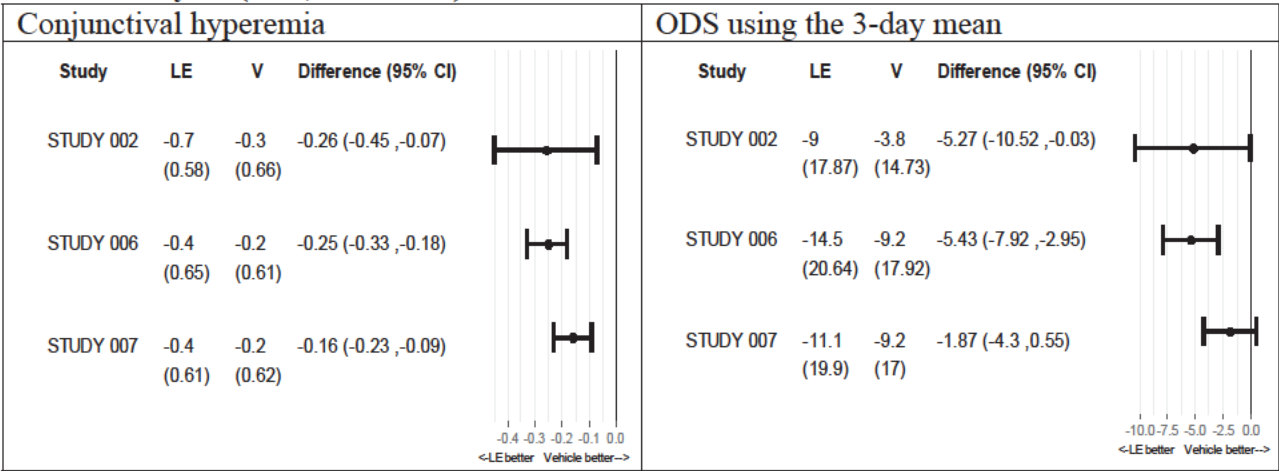
^c In the two Phase 3 studies, the primary symptom endpoints were calculated using the mean scores for the 3 days prior to Day 15 minus the mean of the scores for the 3 days prior to Day 1 (using the 3-day mean). Note that Day 15 was nominal study day instead of the Day 15 study visit.

To control for the overall Type I error (α) of 0.05, the two primary endpoints in the Phase 2 studies were tested sequentially (sign first and symptom second); the four primary endpoints in Study 006 were tested using a parallel gatekeeping method with recycling of unspent α (Appendix 2); and the two primary endpoints in Study 007 were tested using the Bonferroni correction. The analysis method for all of these primary endpoints was analysis of covariance model (ANCOVA) adjusting for baseline and randomization strata.

The three studies showed consistent primary sign results. Both Phase 3 studies demonstrated superiority of the LE 0.25% to the vehicle for the primary sign endpoint of change from baseline in conjunctival hyperemia at Day 15 as shown in Figure 1. The Phase 2 study also demonstrated superiority of the LE 0.25% to the vehicle for its own primary sign endpoint and for the Phase 3 studies' primary sign endpoint.

The three studies, however, showed less consistent primary symptom results. Only Study 006 demonstrated superiority of the LE 0.25% to the vehicle for the primary symptom endpoints. Studies 002 and 007 failed to demonstrate superiority of LE 0.25% for their primary symptom endpoint, although they both showed numerical favorable results (see Table 7 for completed primary symptom results). The applicant analyzed the primary symptom endpoint of the two Phase 3 studies in the Phase 2 study post-hoc. The results achieved nominal significance (Figure 1).

Figure 1: Change from Baseline in Conjunctival Hyperemia and ODS using the 3-day mean at Day 15 (ITT; Observed)

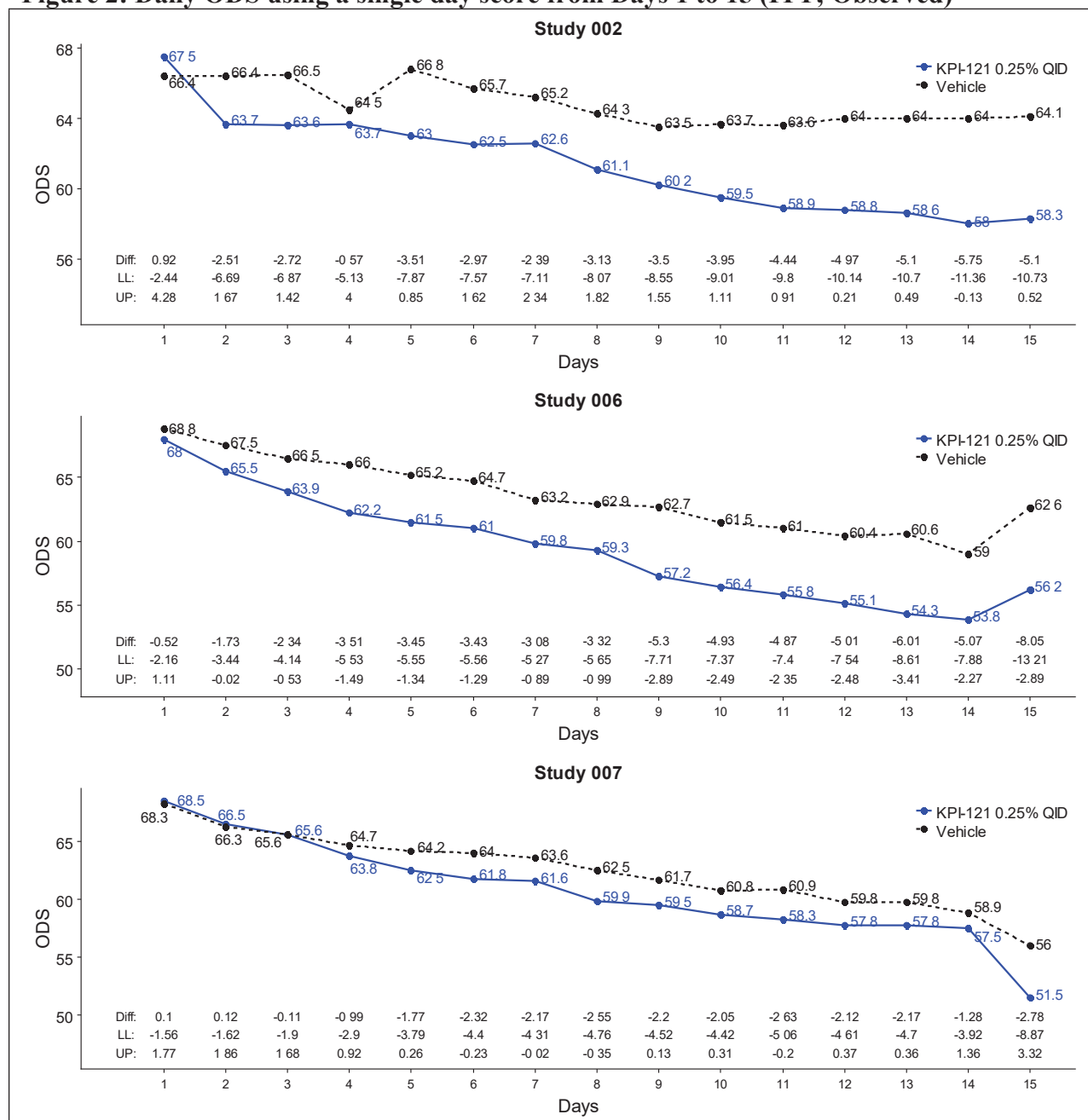


Source: Summary of Clinical Efficacy, Tables 9 and 14. 95% confidence interval (CIs) are calculated by this reviewer.
 Note: ITT = Intent-to-Treat population. In Study 002, ITT population was defined as all randomized subjects with at least 1 post baseline evaluation. In Studies 006 and 007, ITT population was defined as all randomized subjects with at least 1 dose of study treatment. LE = LE 0.25%, V = Vehicle.

The applicant also analyzed daily ODS over time in the three studies. As can be seen in Figure 2 below, the LE group appeared to have consistently lower daily ODS than the vehicle from Days 4 to 15 in all three studies. When comparing across studies, Study 007 appeared to have less treatment separation.

In conclusion, in terms of dry eye signs, the consistent primary results in the three studies suggested a treatment effect of LE 0.25% versus the vehicle. In terms of dry eye symptoms, although the three studies showed less consistent results, all studies showed a consistent treatment separation in daily ODS over time. This suggests a treatment activity. As whether this treatment activity is clinically relevant and sufficient for approval, deference is made to the clinical review team.

Figure 2: Daily ODS using a single day score from Days 1 to 15 (ITT; Observed)



Source: KPI-121-002 add-table Table 16.1.13.1; KPI-121-006 CSR, Table 14.2.4.1.1; and KPI-121-007 CSR, Table 14.2.4.1.1

Note: Diff = treatment difference, LL = lower limit of 95% CI, UP = upper limit of 95% CI. The treatment differences and 95% CI were calculated from the ANOCVA that included treatment, baseline calculated as the mean of the scores over the three days prior to Day 1, and randomization strata.

ITT= Intent-to-treat population. In Study 002, ITT population was defined as all randomized subjects with at least 1 post baseline evaluation. In Studies 006 and 007, ITT population was defined as all randomized subjects with at least 1 dose of study treatment.

2 INTRODUCTION

2.1 Overview

2.1.1 Background

Dry eye disease is a chronic disease that affects approximately 33 million people in the United States (US). The disease is characterized by symptoms of ocular discomfort, including but not limited to, eye dryness, foreign body sensation, irritation, burning, tearing, ocular pain, itching, and blurring of vision. Two medications have been approved by the US Food and Drug Administration (FDA) for long-term usage of dry eye: cyclosporine ophthalmic solution (Restasis; Allergan) and lifitegrast (Xiidra; Shire). Cyclosporine ophthalmic solution (Cequa; Sun Ophthalmics) has also been approved to increase tear production in dry eye patients. The proposed drug product, loteprednol etabonate (LE) 0.25%, dosed four times daily (QID), has the proposed indication of temporary relief instead of long-term usage for the signs and symptoms of dry eye disease.

LE as an ester-based corticosteroid is marketed in the US for over 20 years. LE 0.2% has been approved by FDA (NDA 20803; approved in 1998) for the temporary relief of signs and symptoms of seasonal allergic conjunctivitis. LE 0.5% has been approved to treat the post-operative pain and inflammation. The marketed dosing frequency for LE 0.5% is QID, and three different dosage forms have been approved: suspension (NDA 20583 and 20841; approved in 1998), ointment (NDA 200738; approved in 2011), and gel (NDA 202872; approved in 2012).

LE 1%, dosed twice daily, was also approved by FDA in August 2018 for the treatment of post-operative pain and inflammation. LE 1% was developed by the same applicant of the current NDA. Both LE 1% and the proposed drug product, LE 0.25% QID, are formulated using a novel, proprietary drug-delivery technology known as mucus penetrating particles (MPP). According to the applicant, MPP technology can enhance penetration through the mucous layer of the tear film and thus improve the efficacy of the drug product while maintaining favorable safety profile of LE.

2.1.2 History of Drug Development

In June 2014, under IND 117192, the applicant submitted a protocol for a Phase 2 study (KPI-121-C002, also referred to as Study 002) to Division of Transplant and Ophthalmology Products (DTOP) to evaluate the safety and efficacy of LE 0.25% for the treatment of dry eye disease. The study included a primary sign endpoint and a primary symptom endpoint. The primary sign endpoint was the mean conjunctival hyperemia score in the worst region of the study eye at Day 29 (see Appendix 1 for the definitions of the study eye and the worst region in the study eye); the primary symptom endpoint was the mean modified symptom assessment in dry eye (SANDE) severity scores using the mean scores for the 7 days prior to Day 29.

In June 2015, a Type C meeting was held. The applicant reported that Study 002 was completed, and statistical significance was reached for the primary sign endpoint but not for the primary symptom endpoint. The applicant also discussed the following key clinical issues with DTOP:

- The applicant asked whether DTOP accepted LE 0.25% for a dry eye indication if the duration of the treatment was (b) (4) 14 (b) (4) days. DTOP initially disagreed as dry eye disease is a chronic condition. After discussion, DTOP stated that with the appropriate limitations in container size/volume and labeling, an indication for the temporary relief of the signs and symptoms of dry eye disease could be considered.
- DTOP accepted conjunctival hyperemia as the primary sign endpoint and indicated that they would be open to a responder analysis for total clearing of conjunctival hyperemia. DTOP also accepted modified SANDE severity scores using the 7-day mean as the primary symptom endpoint. Furthermore, DTOP stated that shorter analysis intervals (e.g., 3 days) for SANDE severity scores might also be acceptable for the statistical analysis.

In June 2016, the applicant submitted two protocols for two Phase 3 studies (KPI-121-C006 and KPI-121-C007, also referred to as Studies 006 and 007). These two protocols were amended four times. In the final amendment dated July 6, 2017, both studies included two primary sign endpoints and two primary symptom endpoints:

- Two symptom endpoints: Change from baseline in ocular discomfort severity (ODS) scores at Day 15 and change from baseline in ODS scores at Day 15 in subjects with more severe baseline ODS defined as ODS scores ≥ 68 mm 1 day prior to Day 1. The baseline and Day 15 ODS scores were the mean ODS scores over the 3 days prior to Day 1 and Day 15 (hereinafter referred to as using the 3-day mean). Note that ODS was defined identically to the modified SANDE severity scores in Study 002, thus modified SANDE severity scores will be hereinafter referred to as ODS scores.
- Two sign endpoints: Change from baseline in conjunctival hyperemia scores in the worst region of the study eye at Day 15 and change from baseline in inferior corneal staining scores at Day 15

Compared to the primary endpoints in Study 002, the primary endpoints above were evaluated at Day 15 instead of Day 29 and the change from baseline values were analyzed instead of the mean values at a particular study visit/Day. Also, for the primary symptom endpoint of ODS scores, the analysis duration was shortened from 7 days to 3 days. The applicant proposed to

(b) (4)
(see Appendix 2 for details).

In August 2017, a Type B, pre-NDA meeting was held. At the meeting, the applicant asked if the proposed (b) (4) method was acceptable. After receiving clarification of this method, DTOP stated that it understood the statistical theory involved but remained concerned about the strength of the finding regarding the proposal for (b) (4). The applicant also

proposed to submit non-standard analysis datasets. DTOP recommended the applicant also submit SAS codes along with the define documents.

In December 2017, the statistical analysis plan (SAP) for Study 007 was amended following the unmasking and analysis of Study 006 data. (b) (4)

In May 2018, a Type C meeting was held to discuss the results of Studies 002, 006, and 007. The applicant reported that a statistically significant treatment difference was achieved in all three studies for the primary sign endpoint of conjunctival hyperemia. For the primary symptom endpoint, the significant treatment difference was only achieved in Study 006 but not in the other two studies. The applicant conducted numerous post-hoc analyses. In particular, the applicant analyzed the primary symptom endpoint of the two Phase 3 studies in Study 002, the result was supportive (p-value = 0.0489). The applicant asked if DTOP agreed that there was sufficient safety and efficacy data to support an NDA filing based on the results of the three studies. DTOP disagreed as Study 007 failed to meet the primary symptom endpoint of Study 006. DTOP recommended an additional study.

In June 2018, the applicant submitted a protocol for a new Phase 3 study (KPI-121-C-011) for dry eye indication. This new study will randomize approximately 900 subjects to receive either LE or vehicle. This new study also modifies inclusion criteria according to the exploratory results of the completed studies (Studies 002, 006, and 007) to include potentially more severe patients (see reviewer's note in Section 3.2.4.3 for details). The two primary endpoints in the new study are both for dry eye symptoms: 1) change from baseline in ODS at Day 15 using the 3-day mean in the ITT population and 2) change from baseline in ODS at Day 15 using the 3-day mean in the subgroup of subjects with more severe baseline ODS defined as $ODS \geq 68$ mm for the 3 days prior to Day 1.

2.1.3 Specific Study Reviewed

This review focuses on Studies 002, 006, and 007. Table 2 below contains a brief summary of these three studies.

Table 2: Summary of Studies 002, 006, and 007

Trial ID	Design*	Treatment /Sample Size**	Primary Endpoint *** /Analysis*****	Applicant's primary analysis results
Study 002 /Primary objective: To evaluate the efficacy and safety of LE 0.25% as compared to the vehicle	MC, R, DB, VC trial	• LE 0.25% / N=73 • Vehicle / N=77	• Mean conjunctival hyperemia scores at Day 29 • Mean ODS for the 7 days prior to Day 29 • Change from baseline in ODS at Day 15 using the 3-day mean (post-hoc primary symptom endpoint) /ANCOVA	• Superiority of LE 0.25% QID (1.44) to the vehicle (1.61) was demonstrated for the primary sign endpoint (p-value = 0.0387) • Superiority of LE 0.25% QID (57.05) to the vehicle (60.47) was not demonstrated for the primary symptom endpoint (p-values = 0.3764) • Superiority of LE 0.25% QID (-9.02) to the vehicle (-3.80) was demonstrated for the post-hoc primary symptom endpoint (p-value = 0.0489)
Study 006 /Primary objective: To evaluate the efficacy and safety of LE 0.25% as compared to the vehicle	MC, R, DB, VC trial	• LE 0.25% / N=459 • Vehicle / N=456	• Change from baseline in conjunctival hyperemia scores at Day 15 • Change from baseline in inferior corneal staining scores at Day 15 • Change from baseline in ODS at Day 15 using the 3-day mean • Change from baseline in ODS at Day 15 using the 3-day mean in subjects with ODS ≥ 68 mm 1 day prior to Day 1 /ANCOVA	• Superiority of LE 0.25% QID (-0.40) to the vehicle (-0.16) was demonstrated for the primary sign endpoint of conjunctival hyperemia (p-value < 0.001) • Superiority of LE 0.25% QID (-0.5) to the vehicle (-0.4) was not demonstrated for the primary sign endpoint of corneal staining (p-value = 0.1128) • Superiority of LE 0.25% QID (-14.53) to the vehicle (-9.16) was demonstrated for the primary symptom endpoint (p-values < 0.001) • Superiority of LE 0.25% QID (-16.67) to the vehicle (-10.76) was demonstrated for the primary symptom endpoint in subjects with ODS scores ≥ 68 mm 1 day prior to Day 1 (p-values < 0.001)

Study 007 /Primary objective: To evaluate the efficacy and safety of LE 0.25% as compared to the vehicle	MC, R, DB, VC trial	• LE 0.25% / N=452 • Vehicle / N=453	• Change from baseline in conjunctival hyperemia scores at Day 15 • Change from baseline in ODS at Day 15 using the 3-day mean /ANCOVA	• Superiority of LE 0.25% QID (-0.38) to the vehicle (-0.24) was demonstrated for the primary sign endpoint (p-value < 0.001) • Superiority of LE 0.25% QID (-11.14) to the vehicle (-9.24) was not demonstrated for the primary symptom endpoint (p-value = 0.1298)
--	---------------------------	---	---	--

* MC: multi-center, R: randomized, DB: double-blind, VC: vehicle controlled.

** Sample size is the number of the subjects in the Intent-to-Treat (ITT) population. In Study 002, ITT population was defined as all randomized subjects with at least 1 post baseline evaluation. In Studies 006 and 007, ITT population was defined as all randomized subjects with at least 1 dose of study treatment.

*** Conjunctival hyperemia scores were the 5 levels grading scales from 0 (none) to 4 (severe) in the worst region of the study eye based on photographic assessments conducted at a central photographic reading center. The ODS scores were measured using a visual analog scale from 0 to 100 mm derived from data entered into diaries by subjects. In Study 002, this measure was described as modified symptom assessment in dry eye (SANDE) scores and was analyzed based on 7-day mean prior to a visit. In Studies 006 and 007, this measure was analyzed using the 3-day mean prior to a visit.

**** ANCOVA model included treatment, randomized strata, and baseline scores

2.2 Data Sources

The data sources for this review include clinical study reports, protocols, SAP, and datasets. All data sources were electronic submitted and located at [\\CDSESUB1\evsprod\NDA210933\0001](#) and [\\CDSESUB1\evsprod\NDA210933\0003](#).

The last link contains the resubmission of the analysis dataset. As described in Section 3.1 below, the analysis datasets in the original submission were not usable. Upon the information request, the applicant resubmitted a usable analysis dataset.

3 STATISTICAL EVALUATION

3.1 Data and Analysis Quality

The analysis datasets in the original submission were not usable due to the following:

- The datasets did not contain the primary endpoints.
- The define documents did not contain sufficient details.
- The SAS programs were poorly documented and not usable.

Upon our information request, the applicant submitted a usable analysis dataset along with a detailed define file, a reviewer's guide, and the SAS programs in November 2018. The resubmitted analysis dataset was of sufficient quality for this reviewer to perform necessary efficacy analyses.

3.2 Evaluation of Efficacy

This section evaluates the efficacy results of the three studies together. Sections 3.2.1 and 3.2.2 summarize the study design, endpoints, and statistical methods. Section 3.2.3 summarizes subject disposition as well as demographic and baseline characteristics. Section 3.2.4 discusses the primary analysis and various pre-defined and post-hoc exploratory non-primary analyses.

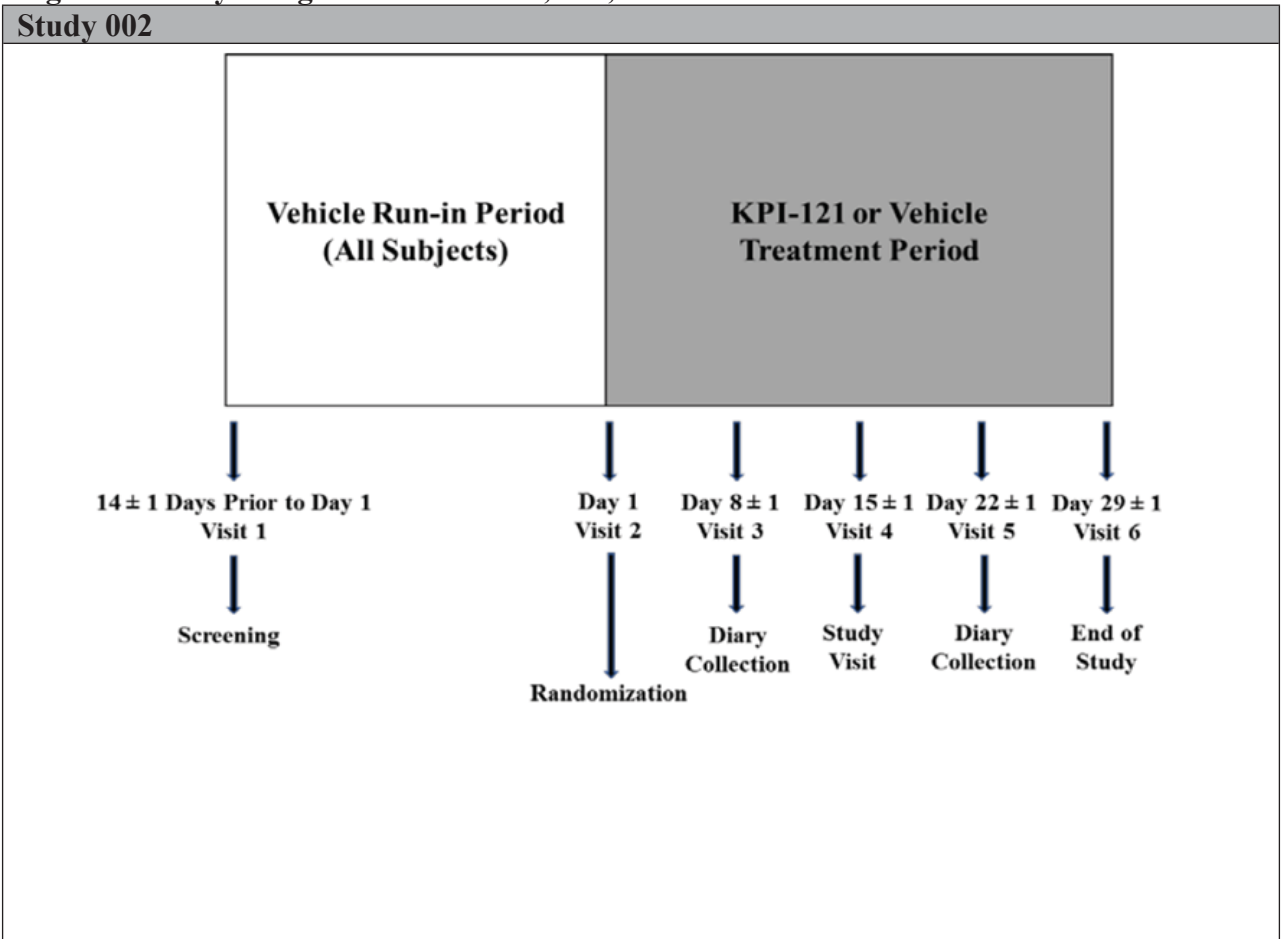
3.2.1 Study Design and Endpoints

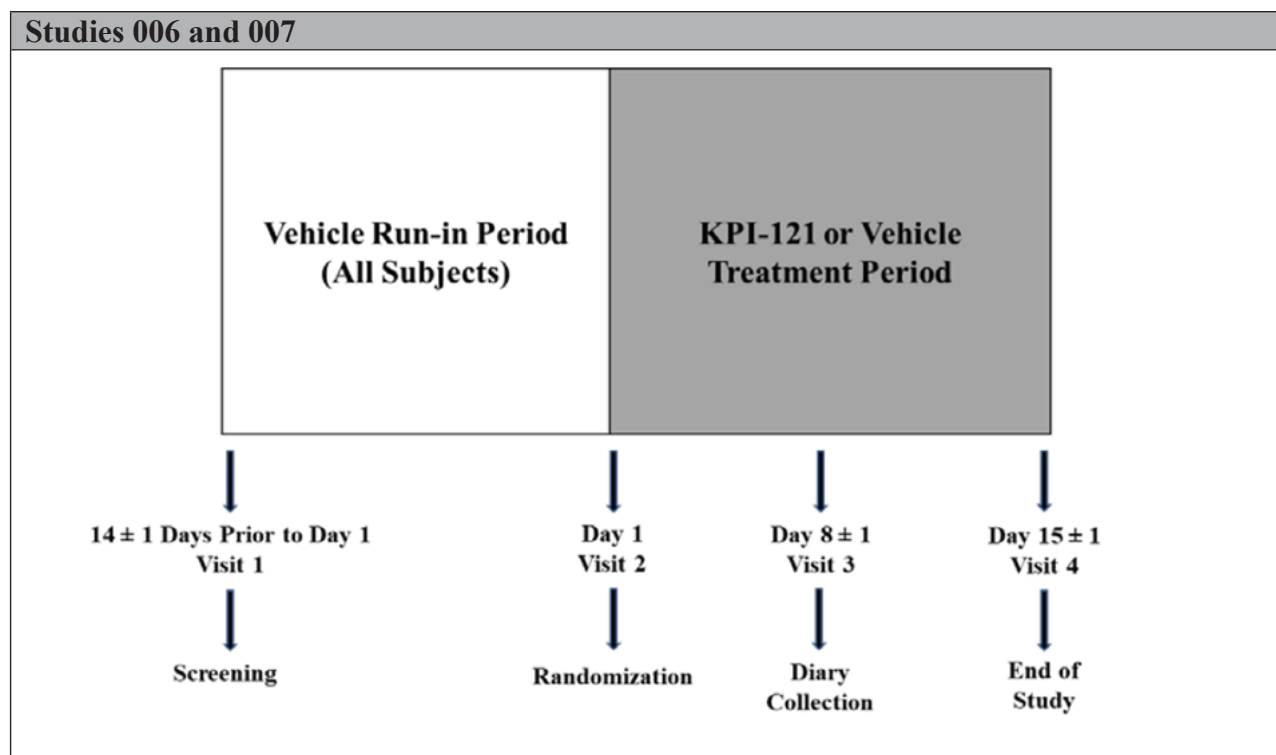
i) Study Design

All three studies were randomized, double-masked, vehicle-controlled studies conducted in the US. Study 002 included a 14-day run-in period followed by a 29-day treatment period. The study visits were at Days -14 (screening), 1 (randomization), 8, 15, 22, and 29 (end of study). Studies 006 and 007 included a 14-day run-in period followed by a 15-day treatment period. The study visits were at Days -14 (screening), 1 (randomization), 8, and 15 (end of study).

Figure 3 below displays the study design for the three studies. Detailed schedules of study assessments are shown in Appendix 3.

Figure 3: Study Design for Studies 002, 006, and 007





Source: Summary of Clinical Efficacy, Figures 1 and 2.
 Note: KPI-121 is the applicant's internal product name for LE.

The key inclusion criteria in the three studies were:

- 1) At least 18 years of age
- 2) Have a documented clinical diagnosis of dry eye disease in both eyes
- 3) Have ongoing dry eye disease as defined by the following criteria in the same eye or both eyes:
 - a. A corneal fluorescein staining score at the screening visit and on Day 1 of ≥ 6 national eye institute (NEI) scale AND
 - b. Investigator-rated assessment of bulbar conjunctival hyperemia at the screening visit and on Day 1 of ≥ 2 as assessed using the cornea and contact lens research unit (CCLRU) scale AND
 - c. A score of 50 mm on ODS at screening and a score of > 40 mm ODS on the day prior to Day 1 AND
 - d. An unanesthetized Schirmer Test score at the screening visit of ≤ 10 mm

During the run-in period, subjects in all three studies received a single dose of single-masked vehicle QID in each eye. At Day 1, eligible subjects in the three studies were randomized in a 1:1 ratio to receive either LE 0.25% QID or vehicle QID for the entire treatment period (28 days in the Phase 2 study and 14 days in the two Phase 3 studies). In the Phase 2 study, randomization was stratified based on subjects' ODS scores (> 50 mm to 70 mm [later amended to include > 40 mm scores at Day 1], > 70 mm to 100 mm) on the day prior to Day 1 and the investigator-rated conjunctival hyperemia scores ($2, \geq 3$) in the study eye on Day 1. In the two Phase 3 studies, randomization was stratified based on subjects' ODS scores (> 40 mm to < 68 mm, ≥ 68 mm to

100 mm) on the day prior to Day 1 and the investigator-rated conjunctival hyperemia scores (2, ≥ 3) in the study eye on Day 1.

Reviewer's note: In all studies, subjects who did not respond adequately to the study treatment could be rescued at the investigator's discretion. The rescued subjects discontinued the study treatment but remained in the study for all efficacy and safety evaluations. Based on this reviewer's analysis, no subject in the ITT population received rescue medication in the three studies.

ii) Study Endpoints

Table 1 in Section 1 shows the pre-defined primary endpoints in the three studies. The conjunctival hyperemia scores were the 5 levels grading scales from 0 (none) to 4 (severe) based on photographic assessments conducted at a central reading center. In this review, unless noted otherwise, conjunctival hyperemia scores were assessed in the worst region of the study eye. Corneal staining was a four-level grading scales from 0 (none) to 3 (severe) assessed in each of the five sections of corneal (superior, inferior, nasal, temporal, and central) of the study eye. The ODS scores were collected each day (one record per day) in subject's dairy using a visual analog scale from 0 to 100 mm, where 0 indicated "very mild" and 100 indicated "very severe". The primary symptom endpoints defined in the two Phase 3 studies were also analyzed in the Phase 2 study in a post-hoc fashion.

The pre-defined secondary sign endpoints in the three studies are listed in Appendix 4. Note that one secondary endpoint in the Phase 2 study was change from baseline in conjunctival hyperemia at Day 15. This secondary endpoint was defined as the primary sign endpoint in the two Phase 3 studies.

Other than the pre-defined endpoints, the applicant explored numerous post-hoc endpoints, and most of these endpoints were for dry eye symptoms, for example, daily ODS based on a single day score, change from baseline in ODS to Days 8 and 15 in subgroup of subjects with more severe baseline ODS defined as ODS scores ≥ 68 mm for the 3 days prior to Day 1. The results for key post-hoc exploratory endpoints will be presented in Sections 3.2.4.2 and 3.2.4.3.

3.2.2 Statistical Methodologies

i) Analysis Population

Three analysis populations were defined in the three studies:

- ITT population. In the Phase 2 study, ITT population was defined as all randomized subjects with at least 1 post baseline evaluation. In the two Phase 3 studies, ITT population was defined as all randomized subjects who received at least 1 dose of study treatment. The primary analysis was performed on the ITT population.

- Per-Protocol (PP) population. In all three studies, PP population was defined as all subjects who completed the randomized study treatment period, fulfilled all study visits, and achieved reasonable compliance with the study protocol.
- Safety population. In all three studies, safety population was defined as all randomized subjects who received at least 1 dose of the study drug and from whom at least 1 safety assessment was obtained after randomization.

ii) Efficacy Analysis

All primary endpoints were analyzed using the analysis of covariance (ANCOVA) model adjusting for baseline and randomized strata. In the Phase 2 study, the two primary endpoints were tested hierarchically (sign first and symptom second). In Study 006, the four primary endpoints were tested using the parallel gatekeeping method with recycling unspent α . In Study 007, the two primary endpoints were tested using the Bonferroni correction.

Regarding missing data, the SAPs stated that if more than 5% of data points were missing for a given primary endpoint, multiple imputation would be employed. As the amount of missing was less than 2%, the applicant just excluded missing subjects from the efficacy analyses.

Continuous secondary endpoints were analyzed similarly as the primary endpoints. Binary secondary endpoints were analyzed using the Chi-square test. The applicant stated in the SAPs of the two Phase 3 studies that if all the primary efficacy endpoints achieved statistical significance, all secondary endpoint would be tested hierarchically following the order in the Appendix 4. As all three studies failed to achieve statistical significance for all of the primary endpoints, secondary analyses results will be presented based on a nominal $\alpha = 0.05$ in this review. The post-hoc analyses will be presented based on a nominal $\alpha = 0.05$, too.

3.2.3 Patient Disposition, Demographic and Baseline Characteristics

3.2.3.1 Patient Disposition

Studies 002, 006, and 007 each randomized 150, 918, and 909 subjects, respectively. In Study 006, two subjects were randomized twice, and one subject was randomized after the completion of Study 007. The second participation data from these subjects were excluded from the analysis populations. In Study 007, one subject was randomized twice, and two subjects were randomized after the completion of Study 006. The second participation data from these subjects were also excluded. Table 3 is a summary of number of subjects in each analysis population by study and by treatment. Table 4 shows the subject disposition. Only a small percentage of randomized subjects in the three studies discontinued early from the study. Only one subject discontinued due to adverse events (AEs) in the Phase 2 study and no subject was discontinued due to AEs in the two Phase 3 studies.

Table 3: Summary of Analysis Population (All Randomized Subject IDs)

	Study 002		Study 006		Study 007	
	0.25% LE (N=73)	Vehicle (N=77)	0.25% LE (N=460)	Vehicle (N=458)	0.25% LE (N=455)	Vehicle (N=454)
ITT Population	73 (100%)	77 (100%)	459 (99.8%)	456 (99.6%)	452 (99.3%)	453 (99.8%)
PP Population	67 (91.8%)	66 (85.7%)	410 (89.1%)	409 (89.3%)	416 (91.4%)	419 (92.3%)
Safety Population	73 (100%)	77 (100%)	459 (99.8%)	456 (99.6%)	452 (99.3%)	453 (99.8%)

Source: KPI-121-006 CSR, Table 9 and KPI-121-007, Table 9

Note: PP and safety population were summarized as randomized. However, in Study 002, as one subject (b) (6) switched treatment to vehicle, the PP population actually included 66 subjects who received LE and 67 subjects who received vehicle. Similarly, safety population included 72 subjects received LE and 78 subjects received vehicle.

Table 4: Subject Disposition (ITT)

	Study 002		Study 006		Study 007	
	0.25% LE (N=73)	Vehicle (N=77)	0.25% LE (N=459)	Vehicle (N=456)	0.25% LE (N=452)	Vehicle (N=453)
Completed Study	72 (98.6%)	76 (98.7%)	455 (99.1%)	453 (99.3%)	447 (98.9%)	448 (98.9%)
Discontinued early from the study	1 (1.4%)	1 (1.3%)	4 (0.9%)	3 (0.7%)	5 (1.1%)	5 (1.1%)
Adverse Event	1 (1.4%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Withdrawal by Subject	0 (0%)	0 (0%)	3 (0.7%)	1 (0.2%)	2 (0.4%)	2 (0.4%)
Lost to Follow-Up	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	2 (0.4%)
Other	0 (0%)	1 (1.3%)	1 (0.2%)	2 (0.4%)	3 (0.7%)	1 (0.2%)

Source: KPI-121-002 CSR, Table 7; KPI-121-006 CSR, Table 7; KPI-121-007 CSR, Table 7.

3.2.3.2 Demographic and Baseline Characteristics

Demographic and baseline characteristics were generally balanced between treatment groups in the three studies (Table 5). The majority of study subjects were non-Hispanic (over 70%), female (over 70%), and white (over 70%). The mean age in the three studies was approximately 58 years old (range: 18 to 93).

Table 5: Demographic and Baseline Characteristics (ITT)

	Study 002		Study 006		Study 007	
	0.25% LE (N=73)	Vehicle (N=77)	0.25% LE (N=459)	Vehicle (N=456)	0.25% LE (N=452)	Vehicle (N=453)
Age at randomization (years)						
Mean (SD)	56.3(17.17)	54.9(12.29)	58.1(15.41)	58.3(14.66)	59.1(14.48)	59.3(15.01)
Median	58	57	60	60.5	61.5	62.0
Min, Max	18, 88	22, 79	18, 90	19, 88	19, 91	18, 93
Age group (years)						
< 65	48(65.8%)	58(75.3%)	280(61%)	289(63.4%)	265(58.6%)	270(59.6%)
≥ 65	25(34.2%)	19(24.7%)	179(39%)	167(36.6%)	187(41.4%)	183(40.4%)
Sex						
Female	53(72.6%)	61(79.2%)	367(80%)	359(78.7%)	332(73.5%)	354(78.1%)
Male	20(27.4%)	16(20.8%)	92(20%)	97(21.3%)	120(26.5%)	99(21.9%)
Race						
White	54(74.0%)	63(81.8%)	362(78.9%)	361(79.2%)	345(76.3%)	352(77.7%)
Non-White	19(26.0%)	14(18.2%)	97(21.1%)	95(20.8%)	107(23.7%)	101(22.3%)
Ethnicity						
Hispanic or Latino	15(20.5%)	16(20.8%)	65(14.2%)	76(16.7%)	118(26.1%)	118(26%)
NOT Hispanic or	58(79.5%)	61(79.2%)	394(85.8%)	380(83.3%)	334(73.9%)	335(74%)

Latino						
Baseline conjunctival hyperemia by photographic assessment						
Mean (SD)	2.2(0.45)	2.1(0.43)	2.0(0.6)	2.0(0.62)	2.0(0.62)	2.0(0.57)
Median	2	2	2	2	2	2
Min, Max	1, 4	1, 4	0, 4	0, 4	0, 4	0, 4
ODS at Day 1						
Mean (SD)	67.5(15.72)	66.4(14.79)	68.0(17.88)	68.8(16.60)	68.5(18.17)	68.3(17.78)
Median	69	67.5	70.0	71	71.0	70.0
Min, Max	3, 96	20, 97	11, 100	11, 100	12, 100	1, 99
Corneal Staining Inferior at Day 1						
Mean (SD)	2.19 (0.638)	2.38 (0.67)	2.1 (0.70)	2.1 (0.66)	2.1 (0.65)	2.2 (0.65)
Median	2	2	2	2	2	2
Min, Max	0, 3	1, 3	0, 3	0, 3	0, 3	0, 3
Corneal Staining Nasal at Day 1						
Mean (SD)	1.68 (0.705)	1.77 (0.686)	1.7 (0.70)	1.6 (0.69)	1.6 (0.68)	1.6 (0.67)
Median	2	2	2	2	2	2
Min, Max	0, 3	1, 3	0, 3	0, 3	0, 3	0, 3

Source: KPI-121-002 CSR, Tables 9, 10, and 14.2.5.1.1; KPI-121-002 add-table Table 16.1.13.1; KPI-121-006 CSR, Tables 10 14.2.1.12, and 14.2.1.1, and 14.2.4.1.1; and KPI-121-007 CSR, Tables 10, 14.2.1.1, 14.2.1.7, and 14.2.4.1.1

3.2.4 Results and Conclusion

3.2.4.1 Primary Efficacy Results

Table 6 presents the efficacy results for the pre-defined primary endpoints in the three studies. All three studies demonstrated superiority of the LE 0.25% to the vehicle for the primary sign endpoints of conjunctival hyperemia. In Study 002, the mean conjunctival hyperemia scores at Day 29 for the LE 0.25% and vehicle groups were 1.4 and 1.6, respectively; the estimated difference was -0.21 (95% CI: -0.41, -0.01). In Study 006, the changes from baseline in conjunctival hyperemia at Day 15 for the LE 0.25% and vehicle groups were -0.4 and -0.2, respectively; the estimated difference was -0.25 (95% CI: -0.33, -0.18). In Study 007, the changes from baseline in conjunctival hyperemia at Day 15 for the LE 0.25% and vehicle groups were -0.4 and -0.2, respectively; the estimated difference was -0.16 (95% CI: -0.23, -0.09).

Inferior corneal staining was analyzed as one of the primary sign endpoints only in Study 006. The treatment difference for this endpoint numerically favored the LE group.

The results for the primary symptom endpoints were not as consistent as those for the primary sign endpoints of conjunctival hyperemia (Table 7). For the pre-defined primary symptom endpoints, only Study 006 demonstrated superiority of the LE 0.25% to the vehicle. In Study 006, the changes from baseline in ODS at Day 15 using the 3-day mean in the ITT population were -14.5 and -9.2 for the LE and vehicle groups, respectively; the estimated treatment difference was -5.43 (95% CI: -7.92, -2.95). Analysis of this endpoint in the ITT subgroup of ODS $\geq$ 68 mm 1 day prior to Day yielded similar results (treatment difference = -5.9; 95% CI: -9.35, -2.44). Although the other two studies failed to demonstrate superiority, the results numerically favored the LE group. When comparing across the two Phase 3 studies, the mean

change in ODS at Day 15 in the vehicle group were similar between the two studies, while the mean change in ODS in the LE group in Study 007 was lower than that in Study 006.

Analyses of the primary endpoints in the PP population yielded similar results (Appendix 5). The applicant did not conduct any sensitivity analyses for missing data because the amount of missing was less than 5%.

Table 6: Results for the Primary Sign Endpoints (ITT; Observed)

		LE 0.25% QID	Vehicle QID	LE 0.25% QID - vehicle	
Study	Endpoint	Mean (SD)	Mean (SD)	Difference (95% CI)	P-value
Study 002	Mean conjunctival hyperemia at Day 29	n = 72 1.4 (0.67)	n = 77 1.6 (0.59)	-0.21 (-0.41, -0.01)	0.0387
Study 006	CFBL in conjunctiva hyperemia at Day 15	n = 452 -0.4 (0.65)	n = 451 -0.2 (0.61)	-0.25 (-0.33, -0.18)	<0.001
	CFBL in inferior corneal staining at Day 15	n = 455 -0.5 (0.84)	n = 452 -0.4 (0.79)	-0.08 (-0.17, 0.02)	0.1128
Study 007	CFBL in conjunctival hyperemia at Day 15	n = 446 -0.4 (0.61)	n = 447 -0.2 (0.62)	-0.16 (-0.23, -0.09)	<0.001

Source: KPI-121-002 CSR, Table 10; KPI-121-006 CSR, Table 11; KPI-121-007 CSR, Table 11; reviewer's analysis in blue

Note: The estimated treatment difference (95% CI and p-value) was based on an ANCOVA model with treatment, baseline values, and stratification variables.

Table 7: Results for the Primary Symptom Endpoints (ITT; Observed)

		LE 0.25% QID	Vehicle QID	LE 0.25% QID - vehicle	
Study	Endpoint	Mean (SD)	Mean (SD)	Difference (95% CI)	P-value
Study 002	Mean ODS at Day 29	n = 73 57.1 (23.90)	n = 77 60.5 (19.77)	-2.64 (-8.51, 3.24)	0.3764
Study 006	CFBL in ODS at Day 15	n = 455 -14.5 (20.64)	n = 452 -9.2 (17.92)	-5.43 (-7.92, -2.95)	<0.001
	CFBL in ODS at Day 15 in subgroup of ODS $\geq$ 68 mm 1 day prior to Day 1	n = 257 -16.7 (21.94)	n = 264 -10.8 (18.04)	-5.9 (-9.35, -2.44)	<0.001
Study 007	CFBL in ODS at Day 15	n = 445 -11.1 (19.9)	n = 444 -9.2 (17)	-1.87 (-4.3, 0.55)	0.1298

Source: KPI-121-002 CSR, Table 11; KPI-121-006 CSR, Table 11; KPI-121-007 CSR, Table 11; reviewer's analysis in blue

Note: The estimated treatment difference (95% CI and p-value) was based on an ANCOVA model with treatment, baseline values, and stratification variables.

As mentioned previously in Section 3.2.1, the primary sign endpoint of change from baseline in conjunctival hyperemia at Day 15 in the two Phase 3 studies was defined as a secondary endpoint in Study 002. Although the primary symptom endpoint of change from baseline in ODS at Day 15 using the 3-day mean in the two Phase 3 studies was not defined in Study 002, it was analyzed post-hoc in Study 002. Table 8 below shows the results for changes from baseline in conjunctival hyperemia and ODS (using the 3-day mean) at Day 15 for Study 002. The results reached nominal statistical significance level of 0.05 for both endpoints. For the endpoint of change from baseline in ODS at Day 15, the estimated treatment difference of -5.3 was close to that in Study 006 (-5.43).

Table 8: Mean Change from Baseline in Conjunctival Hyperemia and ODS (3-Day Mean) at Day 15 in Study 002 (ITT; Observed)

Endpoint	LE 0.25% QID	Vehicle QID	LE 0.25% QID - vehicle	
	Mean (SD)	Mean (SD)	Difference (95% CI)	P-value
CFBL in conjunctival hyperemia at Day 15	n = 73 -0.7 (0.58)	n = 77 -0.3 (0.66)	-0.26 (-0.45, -0.07)	0.009
CFBL in ODS at Day 15 using the 3-day mean	n = 73 -9.02 (17.87)	n = 77 -3.8 (14.73)	-5.3 (-10.52, -0.03)	0.0489

Source: KPI-121-002 CSR, Table 10; KPI-121-002 add tables, Table 16.1.13.2; reviewer's analysis in blue

Note: The estimated treatment difference (95% CI and p-value) was based on an ANCOVA model with treatment, baseline values, and stratification variables.

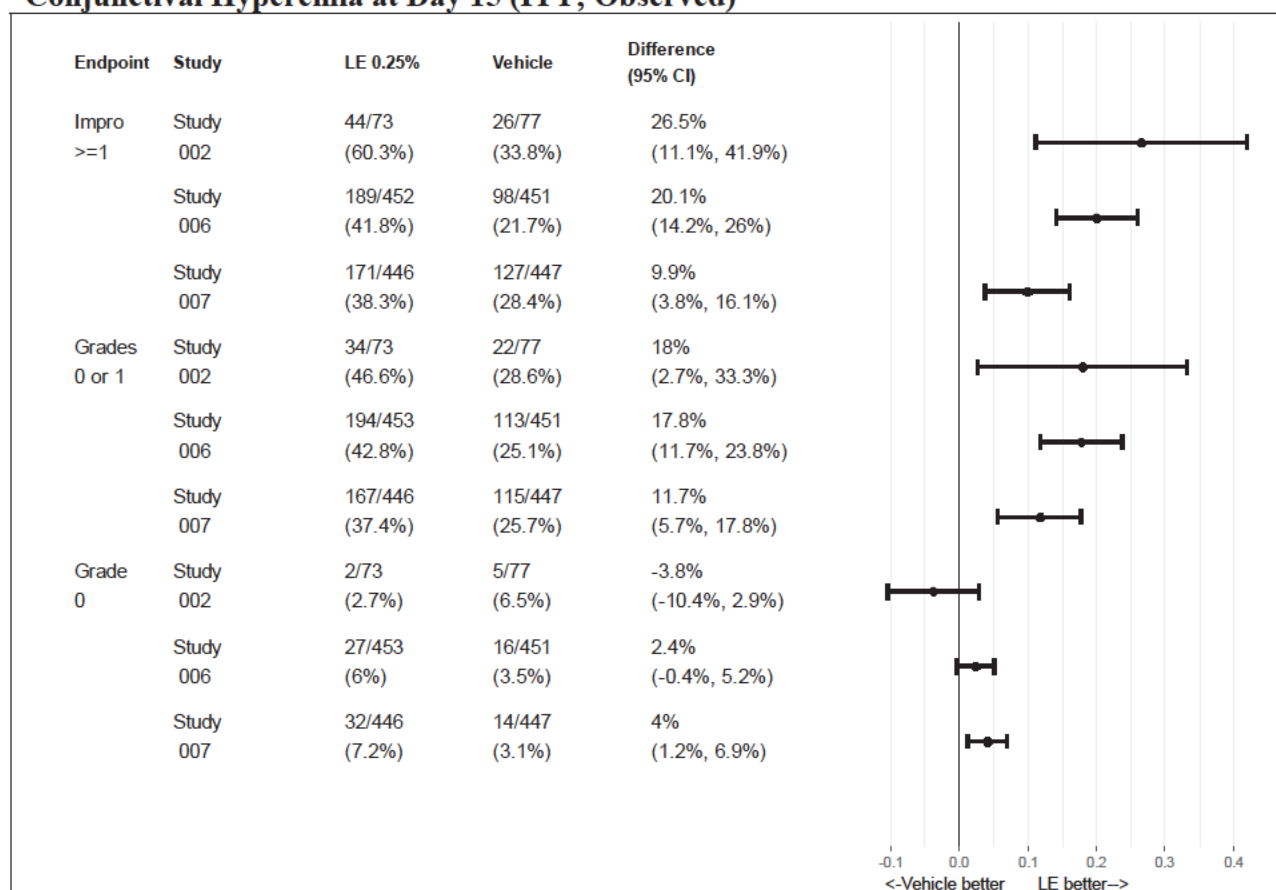
3.2.4.2 Results for Non-Primary Sign Endpoints

Figure 4 presents the efficacy results for the following categorical non-primary sign endpoints:

- Proportion of subjects with ≥ 1 grade improvement from baseline in conjunctival hyperemia at Day 15 (pre-defined secondary endpoints in all three studies)
- Proportion of subjects with grades 0 or 1 on conjunctival hyperemia at Day 15 (post-hoc exploratory endpoints in all three studies)
- Proportion of subjects with grade 0 on conjunctival hyperemia at Day 15 (pre-defined secondary endpoint in Study 006; not defined in the other two studies)

For the first two endpoints, all three studies statistically favored the LE group. For the endpoint of proportion of subjects with grade 0 on conjunctival hyperemia at Day 15, Study 002 did not favor the LE group; Study 006 numerically favored the LE group; and Study 007 statistically favored the LE group.

Figure 4: Forest Plot for Proportion of Subjects with ≥ 1 Grade Improvement from Baseline to Day 15 in Conjunctival Hyperemia, Proportion of Subjects with Grades 0 or 1 on Conjunctival Hyperemia at Day 15 and Proportion of Subjects with Grade 0 on Conjunctival Hyperemia at Day 15 (ITT; Observed)



Source: KPI-121-002 CSR, Table 13; KPI-121-002 CSR addendum, Table 16.1.13.7; KPI-121-006 CSR, Tables 20, 14.2.1.5 and 14.2.4.8; KPI-121-007 CSR, Tables 14.2.1.2 and 14.2.4.8; reviewer's analysis for the proportion of subjects with grade 0 on conjunctival hyperemia at Day 15 in Study 002 and 007.

Note: Impro ≥ 1 is the endpoint of the proportion of subjects with $\geq$ grade improvement from baseline in conjunctival hyperemia at Day 15. Grades 0 or 1 is the endpoint of the proportion of subjects with grades 0 or 1 on conjunctival hyperemia at Day 15. Grade 0 is the endpoint of the proportion of subjects with grade 0 at Day 15 for conjunctival hyperemia.

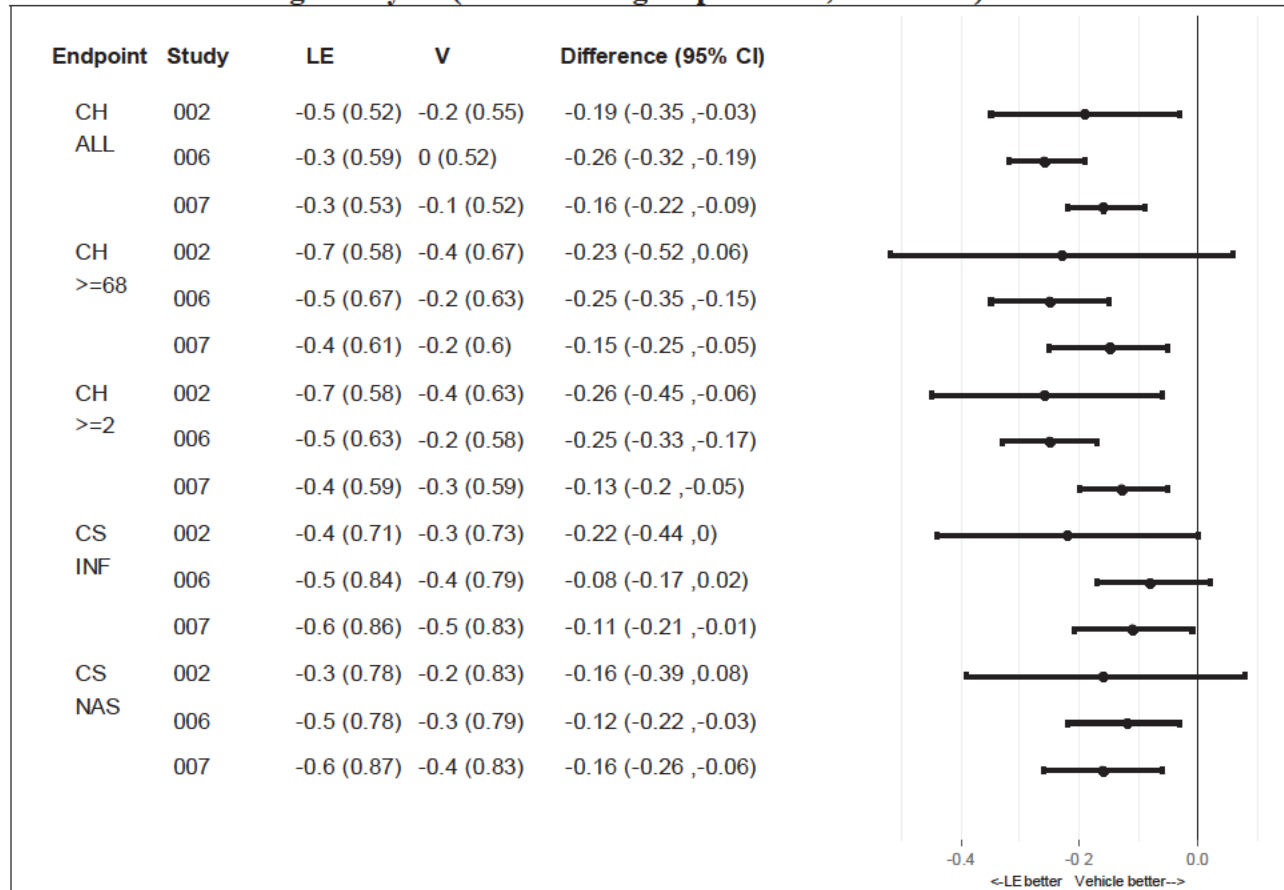
Figure 5 presents the efficacy results for the following continuous non-primary sign endpoints:

- Change from baseline in mean conjunctival hyperemia over all regions (nasal, temporal, and central) in the study eye at Day 15 (pre-defined secondary endpoint in Study 006; not defined in the other studies)
- Change from baseline in conjunctival hyperemia at Day 15 in subjects with ODS ≥ 68 mm one day prior to Day 1 (pre-defined secondary endpoint in Study 006; not defined in the other studies)

- Change from baseline in conjunctival hyperemia at Day 15 in subjects with baseline conjunctival hyperemia ≥ 2 (pre-defined secondary endpoint in Study 006; not defined in the other studies)
- Change from baseline in inferior corneal fluorescein staining at Day 15 (pre-defined primary endpoint in Study 006 and secondary endpoint in the other studies)
- Change from baseline in nasal corneal fluorescein staining at Day 15 (pre-defined secondary endpoint in Studies 002 and 007; not defined in Study 006)

For the above five endpoints, the treatment difference numerically favored the LE group in all three studies. For the endpoint of mean conjunctival hyperemia over all regions of the study eye and for the endpoint of conjunctival hyperemia in subjects with baseline conjunctival hyperemia ≥ 2 , the results statistically favored the LE group in all three studies. For the endpoint of conjunctival hyperemia in subjects with ODS ≥ 68 mm 1 day prior to Day 1, the results statistically favored the LE group in the two Phase 3 studies. For the endpoint of inferior corneal fluorescein staining, the results statistically favored the LE group in Studies 002 and 007. For the endpoint of nasal corneal fluorescein staining, the results statistically favored the LE group in the two Phase 3 studies.

Figure 5: Forest Plot for Change from Baseline in Mean Conjunctival Hyperemia of All Regions at Day 15, Changes from Baseline in Conjunctival Hyperemia at Day 15 in Subjects with ODS of ≥ 68 mm One Day Prior to Day 1 and Subjects with Baseline Conjunctival Hyperemia ≥ 2 , and Changes from Baseline in Inferior and Nasal Corneal Fluorescein Staining at Day 15 (ITT and subgroups of ITT; Observed)



Source: KPI-121-002 CSR, Table 14.2.5.1.1; KPI-121-006 CSR, Tables 14.2.1.3, 14.2.1.2, 14.2.1.6, and 14.2.1.12; KPI-121-007 CSR, Table 14.2.1.7. Reviewer's analyses for endpoints of CH ALL, CH ≥ 68 , and CH ≥ 2 in Studies 002 and 007.

Note: CH ALL is the change from baseline in mean conjunctival hyperemia of all regions in the study eye at Day 15, CH ≥ 68 is change from baseline in conjunctival hyperemia at Day 15 in subjects with ODS of ≥ 68 mm one day prior to Day 1, CH ≥ 2 is the change from baseline in conjunctival hyperemia at Day 15 in subjects with baseline conjunctival hyperemia ≥ 2 , CS INF and CS NAS are mean change from baseline in inferior and nasal corneal fluorescein staining at Day 15.

The applicant also reported the analysis results for the pre-defined non-primary endpoint of investigator-rated conjunctival hyperemia. In Study 002, the treatment difference for the mean investigator-rated conjunctival hyperemia was not significant at Days 15 and 29 (p-value = 0.7588 and 0.1402, respectively; Appendix 6). In the two Phase 3 studies, the treatment difference for the change from baseline in investigator-rated conjunctival hyperemia was statistically significant in nasal, temporal, and frontal regions of the study eye at Day 15 (all p-values < 0.001).

3.2.4.3 Results for Non-Primary Symptom Endpoints

This section presents the efficacy results for the following pre-defined and post-hoc non-primary symptom endpoints:

- Daily ODS using a single day score from Days 1 to 15 (post-hoc exploratory endpoints in all three studies)
- Change from baseline in ODS at Days 2, 3, and 4, calculated as the single day score at Days 2, 3, and 4 minus the mean of the scores for the 3 days prior to Day 1 (Day 4 was pre-defined in the Phase 3 studies as a secondary endpoint, while Days 2 and 3 were pre-defined in Study 007 as secondary endpoints)
- Daily ODS using the 3-day mean from Days 5 to 15 (post-hoc exploratory endpoints in all three studies)
- Change from baseline in ODS at Day 8 using the 3-day mean (not defined in Study 002; pre-defined secondary endpoint in the two Phase 3 studies)
- Change from baseline in ODS at Day 15 using the 7-day mean (post-hoc exploratory endpoint in all three studies)
- Change from baseline in ODS at Day 15 using the 3-day mean in subjects with more severe baseline ODS defined as $\text{ODS} \geq 68$ mm 1 day prior to Day 1 (not defined in Study 002; pre-defined primary endpoint in Study 006 and secondary endpoint in Study 007)
- Change from baseline in ODS at Day 15 using the 3-day mean in subjects with more severe baseline ODS defined as $\text{ODS} \geq 68$ mm for the 3 days prior to Day 1 (post-hoc exploratory endpoints in all three studies)
- Change from baseline in eye dryness scores at Days 8 and 15 (not defined in Study 002; pre-defined other endpoint in Study 006 and secondary endpoint in Study 007)
- Change from baseline in ocular discomfort frequency at Days 8 and 15 using the 3-day mean (not defined in Study 002; pre-defined other endpoint in Study 006 and secondary endpoint in Study 007)

This section also briefly summarizes the results for the primary symptom endpoint of change from baseline in ODS at Day 15 (3-day mean) in the three post-hoc defined modified ITT (mITT) populations.

This section does not discuss the following endpoints in the Phase 2 study: ocular discomfort frequency based on the 7-day mean at Days 15 and 29, subject-rated ocular discomfort severity and frequency scores at Days 15 and 29. Note that for all these endpoints, the results were not statistically significant.

Daily ODS using a single day score from Days 1 to 15

Figure 2 displays daily ODS using a single day score from Days 1 to 15. In Study 002, the LE group had consistently lower ODS from Days 2 to 15 as compared to the vehicle. In Study 006, the LE group had consistently lower ODS throughout the study. In Study 007, the LE group had consistently lower ODS from Days 4 to 15. When comparing across the studies, the treatment differences in Studies 002 and 006 appeared to be constantly greater than those in Study 007 from Days 2 to 15.

Figure 2 also presents the treatment differences and the corresponding 95% CI estimated from an ANCOVA that included treatment, baseline, and randomization strata. Based on this ANCOVA, the treatment difference was most evident in Study 006. In Study 006, the treatment difference was significant from Days 2 to 15. In Study 002, although the treatment difference was insignificant except at Day 14, the point estimates were relatively large and at many time points, close to those in Study 006. Thus, the insignificant results in Study 002 were likely due to the relatively small sample size. In Study 007, the treatment difference was insignificant at most of the time points except at Days 6, 7, 8, and 11. The point estimates were relatively small as compared to the other two studies.

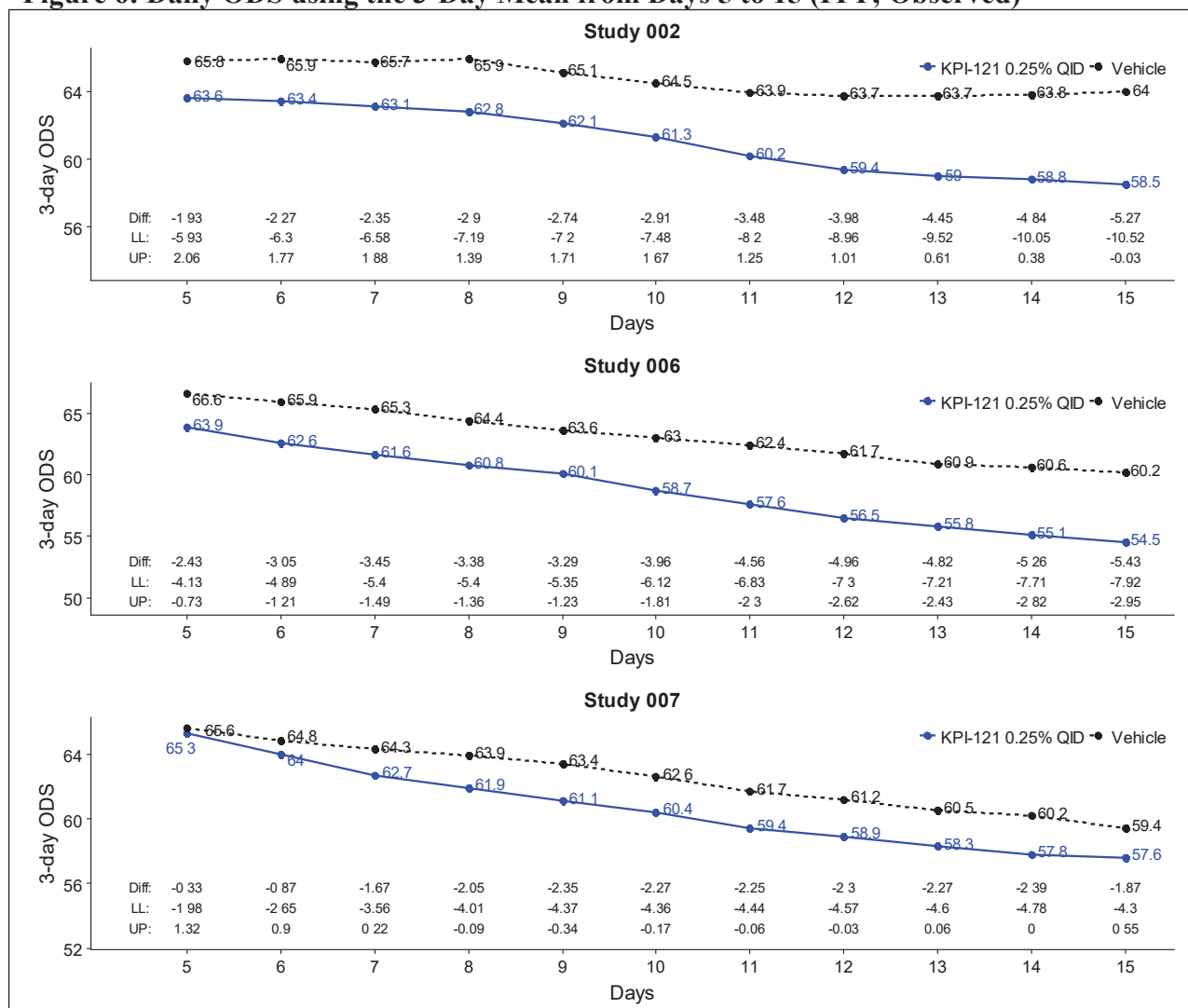
The applicant plotted daily ODS using a single day score from Days 1 to 14 in subgroup of subjects with more severe baseline ODS defined as $\text{ODS} \geq 68$ mm 1 day prior to Day 1. The results were similar to those in the ITT population (Appendix 7).

Daily ODS using the 3-day mean from Days 5 to 15

The results in Figure 2 were also supported by the analysis of daily ODS calculated as the means over the 3 days prior to Days 5 to 15 (Figure 6). Similar to daily ODS based on a single day score, a consistent treatment separation can be observed in all three studies, and Study 007 had a smaller separation as compared to the other two studies. At Days 8 through 12 and at Day 14, the treatment difference was significant for both Phase 3 studies.

The applicant also analyzed the daily ODS using the 3-day mean from Days 8 to 15 in subgroup of subjects with more severe baseline ODS defined as $\text{ODS} \geq 68$ mm 1 day prior to Day 1. The results were similar (Appendix 8).

Figure 6: Daily ODS using the 3-Day Mean from Days 5 to 15 (ITT; Observed)



Source: KPI-121-002 add-table Table 16.1.13.2; KPI-121-006 CSR, Table 14.2.4.2.1; and KPI-121-007 CSR, Table 14.2.4.2.1

Note: Diff = treatment difference, LL = lower limit of 95% CI, UP = upper limit of 95% CI.

Change from baseline in ODS at Days 2, 3, and 4

Change from baseline in ODS at Day 8 using the 3-day mean

The results in Figures 2 and 6 were supported by the applicant's analyses of change from baseline in ODS at Days 2, 3, 4, and 8. For the change from baseline in ODS at Day 4, statistical significance was achieved in Study 006 (treatment difference = -3.51; p-value = 0.0007) but not in Study 007 (treatment difference = -0.99; p-value = 0.3104). For the changes from baseline in ODS at Days 2 and 3, statistical significance was not achieved in Study 007 (difference = 0.12 and -0.11, respectively; p-value = 0.8939 and 0.9034, respectively).

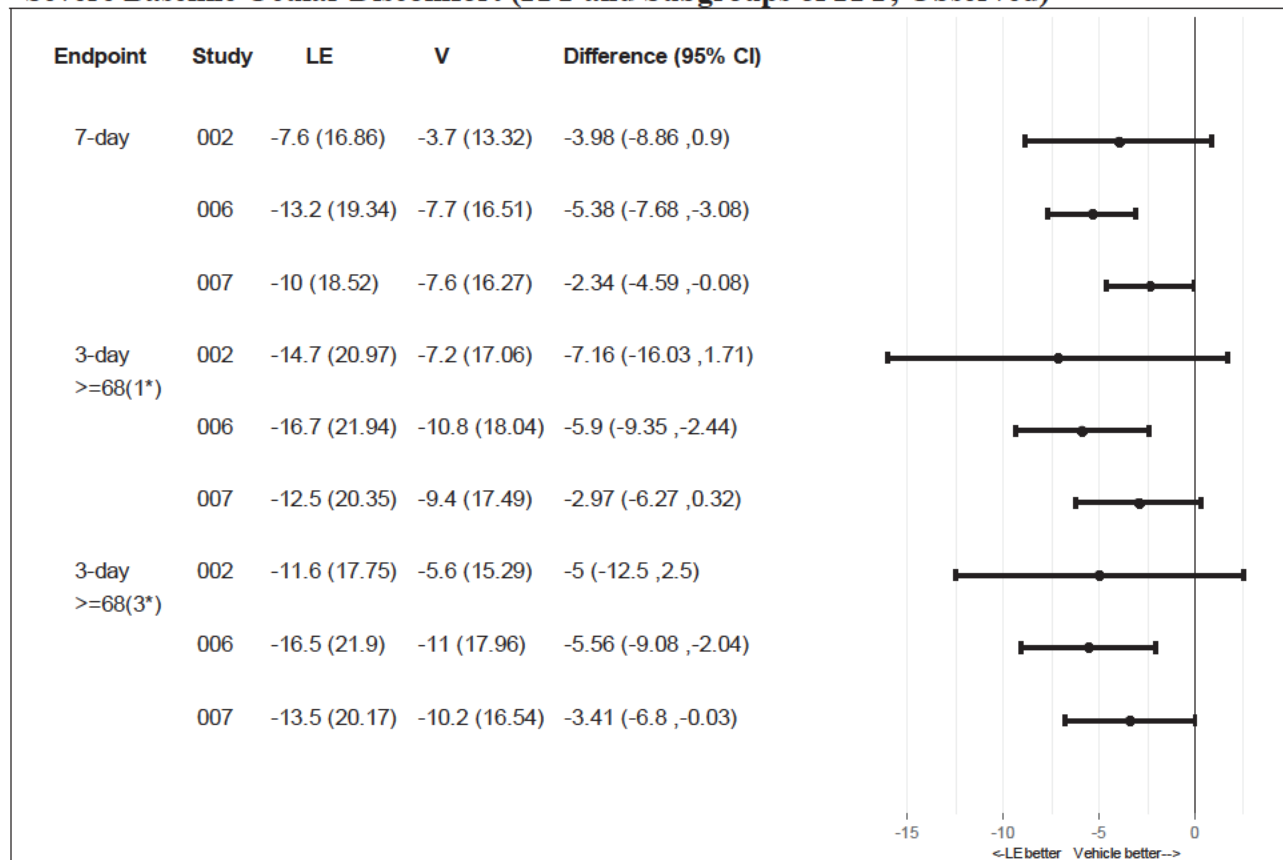
For the change from baseline in ODS at Day 8, the significant results were shown in both Phase 3 studies. The change from baseline in ODS at Day 8 were -8.1 and -4.8 for the LE and vehicle group, respectively, in Study 006 and were -6.9 and -4.9 for the LE and vehicle group, respectively, in Study 007. The estimated treatment differences and 95% CI can be seen in Figure 6 (analyses of change from baseline in ODS at Day 8 and mean scores at Day 8 yielded the same treatment difference and 95% CI).

Change from baseline in ODS at Day 15 using the 7-day mean

Change from baseline in ODS at Day 15 using the 3-day mean in subgroup of subjects with more severe baseline ODS

As shown in Figure 7 below, for the endpoint of change from baseline in ODS at Day 15 based on the 7-day mean, the treatment differences were statistically significant in the two Phase 3 studies but not in the Phase 2 study. For the endpoint of change from baseline in ODS at Day 15 using the 3-day mean in subgroup of subjects with more severe baseline ODS defined as $\text{ODS} \geq 68$ mm 1 days prior to Day 1, the significant treatment difference was only achieved in Study 006 but not in the other two studies. When more severe baseline ODS was redefined post-hoc as $\text{ODS} \geq 68$ mm for the 3 days prior to Day 1, the treatment difference was significant in both Phase 3 studies.

Figure 7: Forest Plot for Change from Baseline in ODS at Day 15 using the 7-Day Mean and Change from Baseline ODS at Day 15 using the 3-Day Mean in Subjects with More Severe Baseline Ocular Discomfort (ITT and Subgroups of ITT; Observed)



Source: KPI-121-002 add-table Tables 16.1.13.4, 16.1.13.5, and 16.1.13.6; KPI-121-006 CSR, Tables 14.2.4.3, 14.2.1.9, and 14.2.4.4; and KPI-121-007 CSR, Tables 14.2.4.3, 14.2.1.5, and 14.2.4.4

Note: 7-day = change from baseline in ODS at Day 15 using the 7-day mean; 3-day >= 68 (1*) = change from baseline in ODS at Day 15 using the 3-day mean in subgroup of subjects with ODS ≥ 68 mm 1 days prior to Day 1; and 3-day >= 68 (3*) = change from baseline in ODS at Day 15 using the 3-day mean in subgroup of subjects with ODS ≥ 68 mm 3 days prior to Day 1

Eye dryness scores

Ocular discomfort frequency

ODS in mITTs

The applicant analyzed the change from baseline in eye dryness scores and ocular discomfort frequency (using the 3-day mean) at Days 8 and 15 in two Phase 3 studies (Table 9). The results statistically favored the LE group at Days 8 and 15 in Study 006; however, these significant results were not shown in Study 007.

Table 9: Change from Baseline in Eye Dryness Scores and Ocular Discomfort Frequency at Days 8 and 15 (ITT; Observed)

			LE 0.25% QID	Vehicle QID	LE 0.25% QID - vehicle	
Endpoint	Study	Time	Mean (SD)	Mean (SD)	Difference (95% CI)	P-value
Eye Dryness Scores	Study 006	Day 8	-11.6 (20.61)	-8.5 (18.82)	-3.84 (-6.25, -1.42)	0.0019
		Day 15	-17.4 (22.91)	-14.3 (21.11)	-3.68 (-6.46, -0.89)	0.0098
	Study 007	Day 8	-9.6 (18.66)	-8 (18.94)	-1.51 (-3.85, 0.84)	0.2084
		Day 15	-14 (22.78)	-11.5 (20.53)	-2.47 (-5.21, 0.27)	0.0776
Ocular Discomfort Frequency	Study 006	Day 8	-7.7 (17.32)	-4.7 (14.96)	-3.04 (-5.07, -1.01)	0.0034
		Day 15	-13.7 (21.08)	-9.1 (18.21)	-4.52 (-7.04, -2.01)	< 0.001
	Study 007	Day 8	-6 (16.13)	-4.4 (16.13)	-1.65 (-3.7, 0.39)	0.1131
		Day 15	-10.2 (20.72)	-8.4 (17.97)	-1.92 (-4.41, 0.57)	0.1309

Source: KPI-121-006 CSR, Tables 14.2.1.13 and 14.2.1.11; KPI-121-007 CSR, Tables 14.2.1.8 and 14.2.1.6

Furthermore, the applicant conducted exploratory analyses for the change from baseline in ODS at Day 15 using the 3-day mean on three mITT populations:

- 1) ITT population excluding subjects with ODS improvement from screening to baseline ≥ 25 mm
- 2) ITT population excluding subjects had total corneal fluorescein staining scores of less than 6 in either eye at either the screening or baseline visit
- 3) ITT population excluding subjects met either one of the two above criteria

These exploratory analyses yielded statistically significant results favoring the LE group in both Phase 3 studies. In the Phase 2 study, the results statistically favored the LE group in the second mITT population (Appendix 9).

Reviewer's note: The above mITT results appear to influence the inclusion criteria in the ongoing Study KPI-121-C-011. In Study KPI-121-C-011, the inclusion criteria require both eyes to have total corneal fluorescein staining score less than 6 at either screening and Day 1 and subjects must have less than 25 mm improvement in ODS during the run-in period before Day 1. Appendix 10 shows the detailed inclusion criteria of Study KPI-121-C-011

3.2.4.4 Efficacy Conclusion

The three studies showed consistent primary sign results: all three studies demonstrated superiority of the LE group to the vehicle for the primary sign endpoints of conjunctival hyperemia. These significant primary results were supported by the analyses of the primary sign endpoints in the PP population and by numerous non-primary analyses. For example, the following non-primary sign results were statistically significant in favor of the LE group:

- Results for the endpoint of change from baseline in conjunctival hyperemia at Day 15 in Study 002 statistically favored the LE group.

- Results for the endpoint of proportion of subjects with ≥ 1 grade improvement in conjunctival hyperemia from baseline at Day 15 statistically favored the LE group in all three studies.
- Results for the endpoint of proportion of subjects with grades 0 or 1 on conjunctival hyperemia at Day 15 statistically favored the LE group in all three studies.
- Results for the endpoint of proportion of subjects with grade 0 on conjunctival hyperemia at Day 15 statistically favored the LE group in Study 007
- Results for the endpoint of change from baseline in mean conjunctival hyperemia over all regions of the study eye at Day 15 statistically favored the LE group in all three studies.
- Results for the endpoint of change from baseline in conjunctival hyperemia at Day 15 in subjects with ODS of ≥ 68 mm one day prior to Day 1 statistically favored the LE group in the two Phase 3 studies.
- Results for the endpoint of change from baseline in conjunctival hyperemia at Day 15 in subjects with baseline conjunctival hyperemia ≥ 2 statistically favored the LE group in all three studies.
- Results for the endpoint of change from baseline in inferior corneal fluorescein staining at Day 15 statistically favored the LE group in Studies 002 and 007
- Results for the endpoint of change from baseline in nasal corneal fluorescein staining at Day 15 statistically favored the LE group in the two Phase 3 studies.
- Results for the endpoint of change from baseline in investigator-rated conjunctival hyperemia at Day 15 in nasal, temporal, and frontal regions statistically favored the LE group in the two Phase 3 studies.

The three studies, however, showed less consistent primary symptom results: only Study 006 demonstrated superiority of the LE group to the vehicle for the pre-defined primary symptom endpoints. In Study 006, the significant primary results were also supported by the PP analyses and by the analyses for numerous non-primary endpoints. For example, statistical significance was achieved for the following pre-defined non-primary symptom endpoints in Study 006: change from baseline in ODS at Days 4 and 8, change from baseline in eye dryness scores and ocular discomfort frequency at Days 8 and 15. Statistical significance was also achieved for the following post-hoc defined non-primary symptom endpoints: daily ODS using a single day scores from Days 2 to 15, daily ODS using the 3-day mean from Days 5 to 14, change from baseline in ODS at Day 15 using the 7-day mean, and change from baseline in ODS using the 3-day mean in subgroups of subjects with more severe baseline ODS. In addition, the applicant showed that statistical significance was achieved when analyzing the primary symptom endpoint in post-hoc defined mITT populations.

Study 007 failed to meet the primary symptom endpoint. Study 007 also failed to meet most of the pre-defined non-primary symptom endpoints even though the results numerically favored the LE group. The applicant conducted numerous post-hoc exploratory analyses. Analysis of daily ODS using a single day score showed a small but consistent treatment separation between the LE and vehicle groups from Days 4 to 15 (significant difference at Days 6, 7, 8, and 11). Analysis of daily ODS using the 3-day mean showed statistically significant treatment differences at Days 8, 9, 10, 11, 12, and 14. Other significant exploratory results favoring the LE group included results for the change from baseline in ODS at Day 15 using the 7-day mean, results for the change from baseline in ODS at Day 15 using the 3-day mean in subgroup of subjects with more severe baseline ODS defined as $\text{ODS} \geq 68$ mm 3 days prior to Day 1, and results from analyzing the primary symptom endpoint in the post-hoc defined mITT populations.

Study 002 failed to meet the pre-defined primary and non-primary symptom endpoints. A post-hoc analysis of the primary symptom endpoint defined in the two Phase 3 studies (i.e., change from baseline in ODS at Day 15 using the 3-day mean) favored the LE group ($p\text{-value} = 0.0489$). Exploratory analyses of daily ODS score using a single day mean and 3-day mean both showed a consistent treatment separation throughout the study. Analysis of change from baseline in ODS at Day 15 using the 3-day mean in the post-hoc modified ITT population removing all subjects with < 6 total corneal fluorescein staining scores in either eye at screening or baseline yielded significant results.

3.3 Safety Evaluation

This section briefly summarizes the safety in the three studies. For a comprehensive review of safety, please refer to Dr. Martin Nevitt's clinical review.

As shown in Table 10, in the pooled safety population, the LE group had slightly more incidences of AEs, ocular AEs, Severe AEs, AEs related to treatment, AEs leading to discontinuation of study treatment, and serious AEs than the vehicle group. No death was reported in the three studies. The most frequently reported AE was instillation site pain (6.2% in the LE group and 5.7% in the vehicle group; Table 11). The applicant reported that the serious AEs in the LE group were diverticulum intestinal, cholelithiasis, hip fracture, and intervertebral disc protrusion (1 subject each).

The applicant also stated that no clinically important differences were observed between the LE and vehicle groups for BCVA, slit-lamp, and dilated ophthalmoscopy findings. In terms of IOP, the applicant stated that the mean IOP values were similar in the LE and vehicle groups at scheduled study visits (Days 15 and 29). The percentages of subjects with > 5 mmHg increase from baseline, > 5 mmHg increase from baseline and ≥ 21 mmHg measurement at any postbaseline visit, ≥ 10 mmHg increase from baseline, and ≥ 10 mmHg increase from baseline and ≥ 21 mmHg measurement at any postbaseline visit were numerically higher in the LE group as compared to the vehicle.

Table 10: Summary of All AEs (Pooled Safety Population)

Category	LE 0.25% QID (N = 983)	Vehicle QID (N = 987)
Any AEs	143 (14.5%)	123 (12.5%)
Ocular AEs	111 (11.3%)	102 (10.3%)
Severe AEs	6 (0.6%)	1 (0.1%)
AEs Related to Treatment	97 (9.9%)	87 (8.8%)
AEs Leading to Discontinuation of Study Treatment	5 (0.5%)	4 (0.4%)
Serious AEs	4 (0.4%)	0
Deaths	0	0

Source: Summary of clinical safety, Table 11

Table 11: Number of Subjects Reporting Common ($\geq 1\%$ in either treatment group) AEs (Pooled Safety Population)

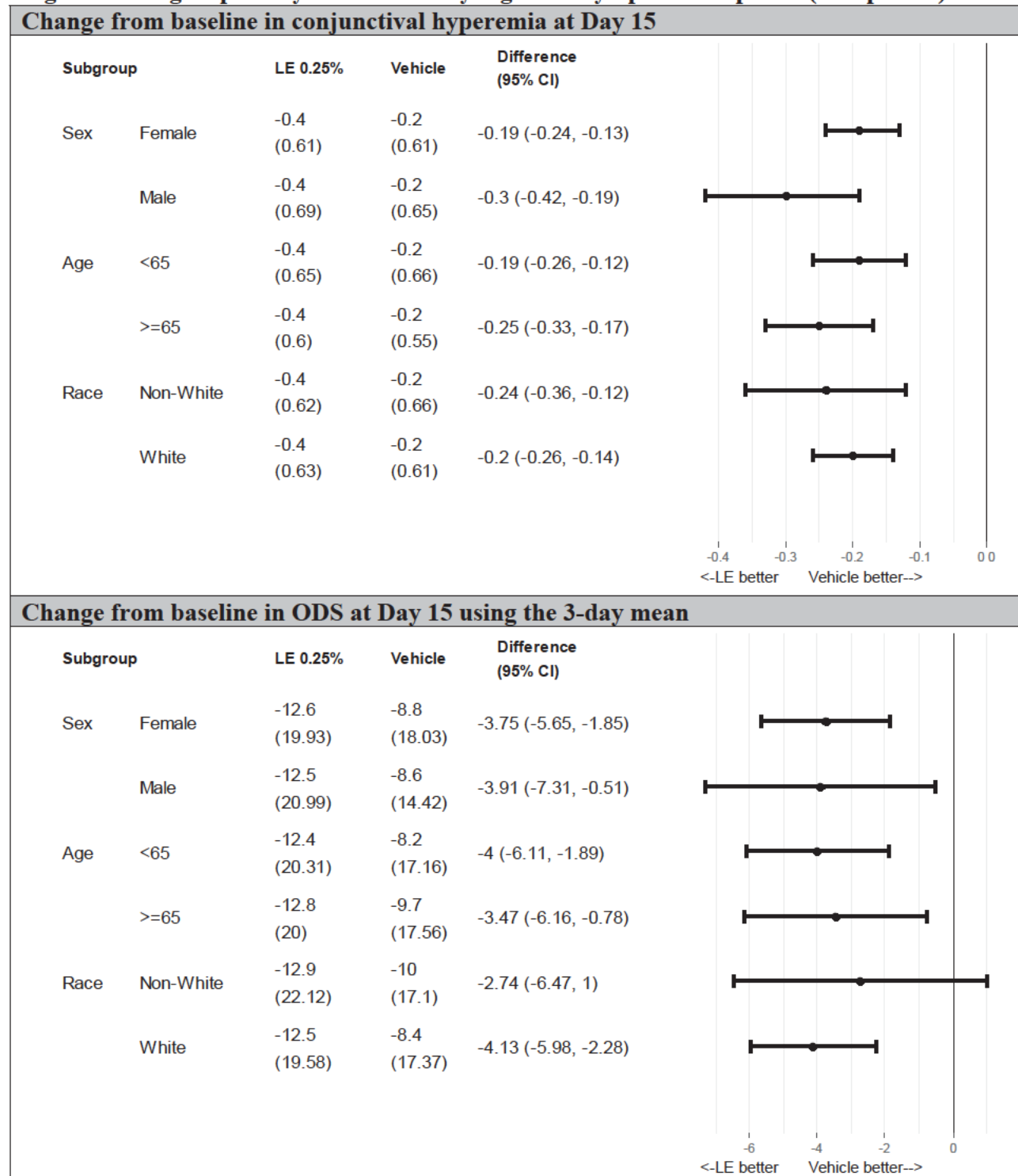
System Organ Class Preferred Term	KPI-121 0.25% QID (N = 983)	Vehicle QID (N = 987)
Number (%) of Subjects Reporting Any AEs	143 (14.5%)	123 (12.5%)
General Disorders and Administration Site Conditions	67 (6.8%)	62 (6.3%)
Instillation Site Pain	61 (6.2%)	56 (5.7%)
Eye Disorders	40 (4.1%)	50 (5.1%)
Eye Irritation	8 (0.8%)	12 (1.2%)

Source: Summary of clinical safety, Table 12

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

Figure 8 displays the results for the primary endpoints of change from baseline in conjunctival hyperemia and ODS (using the 3-day mean) at Day 15 in demographic subgroups of the pooled ITT population from the three studies. For the change from baseline in conjunctival hyperemia at Day 15, the results statistically favored the LE group in all subgroups. For the change from baseline in ODS at Day 15 using the 3-day mean, the results statistically favored the LE group in most of the subgroups, except for the subgroup of subjects who were non-white. In subgroup of non-white subjects, the treatment difference numerically favored the LE group.

Figure 8: Subgroup Analyses for Primary Sign and Symptom Endpoints (ITT pooled)



Source: Reviewer's analysis.

5 SUMMARY AND CONCLUSIONS

5.1 Statistical Issues

No major statistical issues were identified.

5.2 Collective Evidence

Regarding the dry eye signs, all three studies demonstrated superiority of the LE 0.25% to the vehicle for the pre-defined primary sign endpoints of conjunctival hyperemia. The analyses of non-primary sign endpoints provided further assurance that LE 0.25% was effective in improving dry eye sign. For example, the treatment difference was statistically significant for the following non-primary endpoints:

- Change from baseline in conjunctival hyperemia at Day 15 in Study 002
- Proportion of subjects with ≥ 1 grade improvement in conjunctival hyperemia and proportion of subjects with grades 0 or 1 on conjunctival hyperemia at Day 15 in the three studies
- Proportion of subjects with grade 0 on conjunctival hyperemia at Day 15 in Study 007
- Change from baseline in mean conjunctival hyperemia over all regions in the study eye at Day 15 in the three studies
- Change from baseline in conjunctival hyperemia at Day 15 in subjects with ODS of ≥ 68 mm one day prior to Day 1 in the two Phase 3 studies
- Change from baseline in conjunctival hyperemia in subjects with baseline conjunctival hyperemia ≥ 2 in the three studies
- Change from baseline in inferior corneal fluorescein staining at Day 15 in Studies 002 and 007
- Change from baseline in nasal corneal fluorescein staining at Day 15 in the two Phase 3 studies
- Change from baseline in investigator-rated conjunctival hyperemia in nasal, temporal, and frontal regions at Day 15 in the two Phase 3 studies.

Regarding the dry eye symptoms, despite that the numeric trend of the pre-defined primary symptom endpoint favored the LE group in the three studies; only Study 006 demonstrated superiority of the LE group to the vehicle. Although Study 002 failed to meet the pre-defined primary symptom endpoint, it was able to meet the primary endpoint of the two Phase 3 studies in the post-hoc analysis (p-value = 0.0489). The following pre-defined and post-hoc non-primary results may suggest a possible treatment activity of LE 0.25% for the dry eye symptoms:

- Analyses of pre-defined secondary endpoint of change from baseline at Day 8 using the 3-day mean statistically favored the LE group in the two Phase 3 studies.
- Post-hoc analyses of daily ODS using a single day score from Days 1 to 15 showed a consistent treatment separation over time in the three studies. The results statistically

favored the LE group at Day 14 in Study 002, from Days 2 to 15 in Study 006 and at Days 6, 7, 8, and 11 in Study 007.

- Post-hoc analyses of daily ODS using the 3-day mean from Days 5 to 15 showed a consistent treatment separation over time in the three studies, too. The results statistically favored the LE group from Days 5 to 15 in Study 006 and from Days 8 to 12 and at Day 14 in Study 007.
- Post-hoc analyses of change from baseline in ODS at Day 15 using the 7-day mean statistically favored the LE group in the two Phase 3 studies.
- Post-hoc analyses of change from baseline in ODS at Day 15 using the 3-day mean in subgroup of subjects with ODS ≥ 68 mm for the 3 days prior to Day 1 statistically favored the LE group in the two Phase 3 studies.
- Analyses of pre-defined endpoints of change from baseline in ODS at Day 4 and changes from baseline in eye dryness scores and ocular discomfort frequency at Days 8 and 15 statistically favored the LE group in Study 006.

Subgroup analyses showed a consistent treatment effect of the LE 0.25% QID across subgroups of sex, age, and race.

5.3 Conclusions and Recommendations

In conclusion, the three studies provide statistical evidence that the LE 0.25% is effective for the treatment of dry eye signs. For dry eye symptoms, as the two out of the three studies failed to meet the pre-defined primary symptom endpoints, the evidence was less substantial. Based on the collective evidence summarized above, a possible treatment activity of LE 0.25% is shown. As whether this treatment activity is clinically relevant and sufficient to confirm the efficacy of LE 0.25% for the treatment of dry eye symptoms, we defer to the clinical review team. In addition, the applicant is conducting another Phase 3 study (KPI-121-C-011) focusing on dry eye symptoms (i.e., both primary endpoints were symptom endpoints). The results of this ongoing study may provide further information regarding the efficacy of LE 0.25% for the treatment of dry eye symptoms.

8 Page(s) have been Withheld in Full as B4 (CCI/TS) immediately following this page

Appendix 5A: Results for the Pre-defined Primary Sign Endpoints in the PP population (ITT; Observed)

		LE 0.25% QID	Vehicle QID	LE 0.25% QID - vehicle
Study	Endpoint	Mean (SD)	Mean (SD)	Difference, P-value
Study 002	Mean conjunctival hyperemia at Day 29	1.48 (0.64)	1.67 (0.533)	-0.24, 0.0161
Study 006	CFBL in conjunctiva hyperemia at Day 15	-0.42 (0.663)	-0.18 (0.601)	-0.25, < 0.001
	CFBL in inferior corneal staining at Day 15	-0.5 (0.83)	-0.4 (0.76)	-0.1, 0.0581
Study 007	CFBL in conjunctival hyperemia at Day 15	-0.38 (0.597)	-0.24 (0.632)	-0.15, < 0.001

Source: KPI-121-002 CSR, Table 14.2.11.2; KPI-121-006 CSR, Tables 14.2.3.1 and 14.2.3.4; KPI-121-007 CSR, Table 14.2.3.1.

Note: The estimated treatment difference (95% CI and p-value) was based on an ANCOVA model with treatment, baseline values, and stratification variables.

Appendix 5B: Results for the Pre-defined Primary Symptom Endpoints in the PP population (ITT; Observed)

		LE 0.25% QID	Vehicle QID	LE 0.25% QID - vehicle
Study	Endpoint	Mean (SD)	Mean (SD)	Difference, P-value
Study 002	Mean ODS at Day 29	55.76 (23.610)	60.06 (20.384)	-2.58, 0.4259
Study 006	CFBL in ODS at Day 15	-14.67 (20.947)	-9.66 (18.094)	-5.12, < 0.001
	CFBL in ODS at Day 15 in subgroup of ODS $\geq$ 68 mm 1 day prior to Day 1	-16.87 (22.112)	-11.13 (18.580)	-5.74, 0.0025
Study 007	CFBL in ODS at Day 15	-11.09 (19.764)	-9.18 (17.110)	-1.89, 0.1395

Source: KPI-121-002 CSR, Table 14.2.12.2; KPI-121-006 CSR, Tables 14.2.3.3 and 14.2.3.2; KPI-121-007 CSR, Table 14.2.3.2.

Note: The estimated treatment difference (95% CI and p-value) was based on an ANCOVA model with treatment, baseline values, and stratification variables.

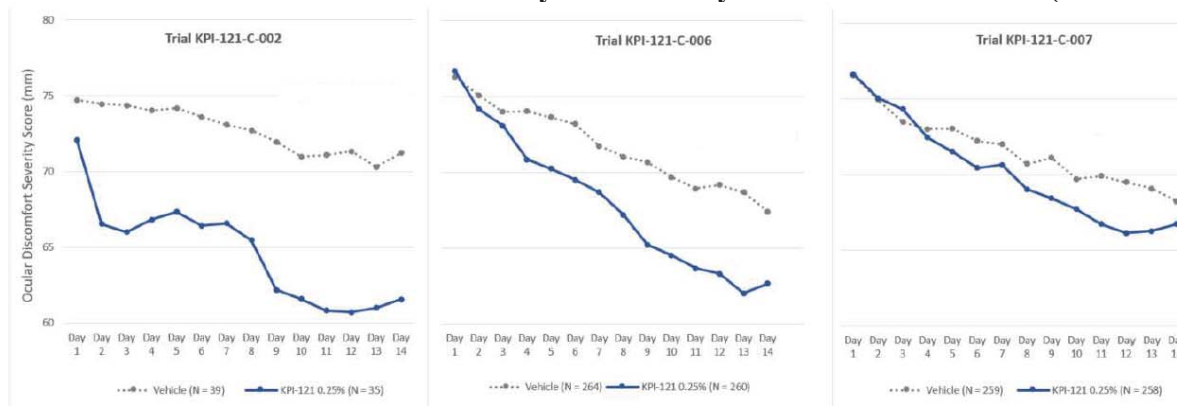
Appendix 6: Results for Investigator-rated Conjunctival Hyperemia (ITT; Observed)

		LE 0.25% QID	Vehicle QID	LE 0.25% QID - vehicle
Study	Day/Region	Mean (SD)	Mean (SD)	Difference, P-value
Study 002	Day 15	2.21 (0.529)	2.19 (0.629)	0.03, 0.7588
	Day 29	2.01 (0.514)	2.14 (0.555)	-0.13, 0.1402
Study 006	Nasal	-0.46 (0.698)	-0.30 (0.683)	-0.17, < 0.0001
	Temporal	-0.49 (0.721)	-0.30 (0.708)	-0.19, < 0.0001
	Frontal	-0.47 (0.765)	-0.29 (0.645)	-0.15, 0.0005
Study 007	Nasal	-0.52 (0.672)	-0.34 (0.692)	-0.17, < 0.0001
	Temporal	-0.51 (0.682)	-0.33 (0.672)	-0.17, < 0.0001
	Frontal	-0.40 (0.672)	-0.27 (0.649)	-0.13, 0.0011

Source: KPI-121-002 CSR, Table 14.2.7.1.2; KPI-121-006 CSR, Table 23; KPI-121-007 CSR, Table 16

Note: In Study 002, mean investigator-rated conjunctival hyperemia scores at Days 15 and 29 were reported. In the two Phase 3 studies, the mean changes from baseline at Day 15 were reported

Appendix 7: Daily ODS from Days 1 to 14 in ITT Subgroup with More Severe Baseline ODS defined as ODS ≥ 68 mm 1 Day Prior to Day 1 in the Three Studies (Observed)



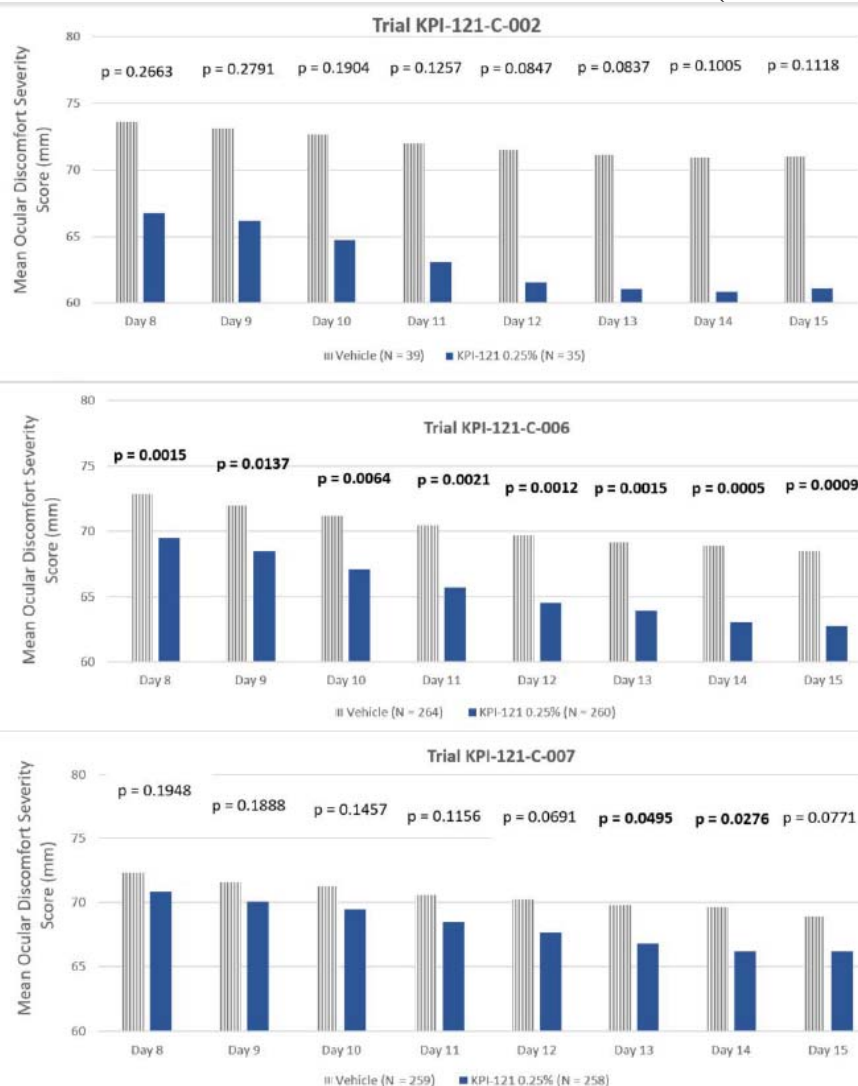
ITT = intent to treat; ODS = ocular discomfort severity

Note: In all trials, KPI-121 0.25% or vehicle was applied as 1 or 2 drops in each eye 4 times a day. In Trial KPI-121-C-002, the treatment period was 28 days. In Trials KPI-121-C-006 and KPI-121-C-007, the treatment period was 14 days. ODS scores (from 0 to 100 mm) were derived from data entered into diaries by subjects. Scores were general evaluations across both eyes; there were not separate questions for each eye. Higher scores indicate more severe ocular discomfort. Daily means are calculated as mean single-day ODS scores for a given day. Only available data are used without imputation. Ns by treatment group reflect the ITT subgroup defined as subjects with ODS scores ≥ 68 mm 1 day prior to Day 1. Ns at each study day are provided in the source tables.

Source: Module 5.3.5.1, CSR KPI-121-C-002, [Tables 16.1.13.1.1](#), CSR KPI-121-C-006, [Tables 14.2.4.1.2](#), CSR KPI-121-C-007, [Table 14.2.4.1.2](#)

Source: Summary of Clinical Efficacy; Figure 9

Appendix 8: 3-Day mean ODS from Days 8 to 15 in ITT Subgroup with More Severe Baseline ODS Defined as Baseline ODS ≤ 68 mm (Observed)



ITT = intent to treat; ODS = ocular discomfort severity

Note: In all trials, KPI-121 0.25% or vehicle was applied as 1 or 2 drops in each eye 4 times a day. In Trial KPI-121-C-002, the treatment period was 28 days. In Trials KPI-121-C-006 and KPI-121-C-007, the treatment period was 14 days. ODS scores (from 0 to 100 mm) were derived from data entered into diaries by subjects. Scores were general evaluations across both eyes;

there were not separate questions for each eye. Higher scores indicate more severe ocular discomfort. Three-day means are calculated as mean scores for the 3 days prior to a given study day. Only available data are used without imputation; if the target number of prior days of diary data was not available, the number of days contributing to the average could be less than the target. Ns reflect the overall ITT population by treatment group. Ns at each study day are provided in the source tables. P-values are on the difference between KPI-121 and vehicle by t-test on difference of adjusted least square means on change from baseline from an analysis of covariance (ANCOVA) with treatment and randomized strata as fixed factors and baseline score as a covariate in the model. P-values < 0.05 are shown in bold.

Source: Module 5.3.5.1, CSR KPI-121-C-002, Table 16.1.13.2.1, CSR KPI-121-C-006, Table 14.2.4.2.2, CSR KPI-121-C-007, Table 14.2.4.2.2

Source: Summary of Clinical Efficacy; Figure 11

Appendix 9: Change from Baseline in ODS using the 3-Day Mean at Day 15 in Multiple Exploratory mITT Population (Observed)

	Phase 2 Trial KPI-121-C-002			Phase 3 Trial KPI-121-C-006			Phase 3 Trial KPI-121-C-007		
	KPI-121 0.25%	Vehicle	Trt Diff P-value ^a	KPI-121 0.25%	Vehicle	Trt Diff P-value ^a	KPI-121 0.25%	Vehicle	Trt Diff P-value ^a
ITT Populations (Prespecified Phase 3 Analysis)									
Day 1	n = 73 67.48 (14.269)	n = 77 67.79 (13.784)	-5.27 0.0489	n = 459 68.93 (15.888)	n = 456 69.26 (15.378)	-5.43 < 0.0001	n = 452 68.79 (15.459)	n = 453 68.70 (15.799)	-1.87 0.1298
Change at	n = 73	n = 77		n = 455	n = 452		n = 445	n = 444	
Day 15	-9.02 (17.872)	-3.80 (14.727)		-14.53 (20.636)	-9.16 (17.924)		-11.14 (19.901)	-9.24 (17.002)	
mITT Removing Subjects with ≥ 25 mm Improvement in ODS Scores During Vehicle Run-In ^b									
Day 1	n = 68 68.03 (14.288)	n = 71 68.77 (13.845)	-5.19 0.0662	n = 416 70.50 (15.422)	n = 413 70.78 (14.983)	-5.35 < 0.0001	n = 408 70.43 (15.032)	n = 415 70.07 (15.349)	-2.94 0.0200
Change at	n = 68	n = 71		n = 412	n = 409		n = 403	n = 407	
Day 15	-9.74 (18.065)	-4.67 (14.837)		-14.64 (20.377)	-9.35 (17.374)		-11.66 (19.109)	-8.70 (16.917)	
mITT Removing Subjects with < 6 Total Corneal Fluorescein Staining Score ^c in Either Eye at Screening or Baseline									
Day 1	n = 53 66.76 (15.217)	n = 59 68.61 (14.173)	-6.32 0.0334	n = 377 69.24 (16.091)	n = 368 70.34 (15.277)	-5.19 0.0004	n = 344 69.62 (15.542)	n = 349 69.06 (15.445)	-3.33 0.0174
Change at	n = 53	n = 59		n = 375	n = 364		n = 341	n = 343	
Day 15	-8.95 (17.271)	-2.84 (13.429)		-14.82 (21.046)	-9.91 (18.617)		-11.86 (19.681)	-8.48 (16.941)	
mITT Removing Subjects with ≥ 25 mm Improvement in ODS Scores During Vehicle Run-In ^b or with < 6 Total Corneal Fluorescein Staining Score ^c in Either Eye at Screening or Baseline									
Day 1	n = 49 67.46 (15.212)	n = 54 69.85 (14.118)	-6.13 0.0545	n = 341 70.99 (15.482)	n = 332 71.94 (14.788)	-4.82 0.0014	n = 315 70.98 (15.248)	n = 320 70.59 (14.819)	-4.37 0.0022
Change at	n = 49	n = 54		n = 339	n = 328		n = 312	n = 314	
Day 15	-9.40 (17.737)	-3.48 (13.731)		-14.93 (21.031)	-10.33 (18.024)		-12.35 (18.869)	-7.98 (16.799)	

ITT = intent to treat; mITT = modified intent-to-treat; ODS = ocular discomfort severity; SD = standard deviation; Trt Diff = between-group treatment difference (KPI-121 minus vehicle) using least-square means

Note: In both trials, KPI-121 0.25% or vehicle was applied as 1 or 2 drops in each eye 4 times a day for 14 days. ODS scores (from 0 to 100 mm) were derived from data entered into diaries by subjects. Scores were general evaluations across both eyes; there were not separate questions for each eye. Higher scores indicate more severe ocular discomfort. Three-day means were calculated using available data from the 3 days prior to Days 1 and 15. Only available data are used without imputation; if the target number of prior days of diary data was not available, the number of days contributing to the average could be less than the target.

^a P-values are on the difference between KPI-121 and vehicle by t-test on difference of adjusted least square means on change from baseline from an analysis of covariance (ANCOVA) with treatment and randomized strata as fixed factors and baseline score as a covariate in the model.

^b Subjects were removed from the analysis if their change in ODS scores during the run-in period (score 1 day prior to Day 1 minus the first score recorded in the subject diary) was -25 mm or less (ie, demonstrating an improvement of 25 mm or more).

^c Corneal staining scores (using the NEI scale, from 0 to 3) were based on evaluations by the investigator; the total score was the sum of scores for the 5 sections of the cornea (superior, inferior, nasal, temporal, central).

Source: Module 5.3.5.1, CSR KPI-121-C-002, Tables 16.1.13.3, 16.1.13.3.1, 16.1.13.3.2, and 16.1.13.3.3, CSR KPI-121-C-006, Tables 14.2.1.7, 14.2.4.7.1, 14.2.4.7.2, and 14.2.4.7.3; CSR KPI-121-C-007, Tables 14.2.1.3, 14.2.4.7.1, 14.2.4.7.2, and 14.2.4.7.3

Source: Summary of Clinical Efficacy, Table 18

Appendix 10: Key Inclusion Criteria for Study KPI-121-C-011

Visit 1 (Screening) and Visit 2 (Randomization), individuals of either gender or any race will be eligible for study participation if they:

1. Provide written informed consent and HIPAA authorization prior to any study-related procedures.
2. Are 18 years of age or older.
3. Are willing and able to follow instructions and can be present for the required study visits for the duration of the study, including:
 - a. Single-masked investigational product use compliance of at least 80% during the final week of Stage 1 AND
 - b. Diary completion compliance for ocular discomfort scoring of at least 80% in the final week of Stage 1.
4. Have a documented clinical diagnosis of dry eye disease in both eyes.
5. Have ongoing dry eye disease as defined by the following criteria in the same eye or both eyes:
 - a. A corneal fluorescein staining score at Visit 1 and Visit 2 of ≥ 6 [National Eye Institute (NEI) scale] in both eyes AND
 - b. Investigator-rated assessment of bulbar conjunctival hyperemia at Visit 1 and Visit 2 of ≥ 2 as assessed using the Cornea and Contact Lens Research Unit (CCLRU) scale in one or both eyes AND
 - c. A score of > 50 mm as graded by the subject utilizing the modified SANDE visual analog scale for ocular discomfort severity (ODS) at Visit 1 AND
 - d. A score of ≥ 45 mm on the 3 days prior to Visit 2 (Day 1) as graded by the subject utilizing the modified SANDE visual analog scale for ODS AND
 - e. An unanesthetized Schirmer Test score at Visit 1 of ≤ 10 mm (in the same eye as qualified for 5b criterion above) AND
 - e. No greater than 25 mm decrease in ODS score during Stage 1 as assessed by comparing the first ODS score recorded in the subject diary to the ODS score on the last diary day before Visit 2 (Day 1).

Source: Page 29 of the Study KPI-121-C-011 protocol (version 1.1; 06/19/2018)

This is a representation of an electronic record that was signed electronically. Following this are manifestations of any and all electronic signatures for this electronic record.

/s/

QING N ZHOU
05/13/2019 12:02:36 PM

YAN WANG
05/13/2019 12:13:29 PM
I concur.