CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

208845Orig1s000

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 208845
Supplement #:	SN 0000
Drug Name:	ZILRETTA 40 mg (triamcinolone acetonide extended release injectable suspension) for intra-articular use
Indication(s):	intra-articular injection for the management of osteoarthritis pain of (b) (4)
Applicant:	Flexion Therapeutics, Inc.
Date(s):	Submission Stamp Date: 12/8/16 PDUFA Date: 10/8/172017 Primary Review Completion Due Date: 9/1/2017
Review Priority:	Standard
Biometrics Division:	Division of Biometrics II
Statistical Reviewer:	Kate Meaker, M.S.
Concurring Reviewers:	David Petullo, M.S. - Team Leader
Medical Division:	Division of Division of Anesthesia, Analgesia, and Addiction Products (DAAAP)
Clinical Team:	Pamela Horn, M.D. – Clinical Reviewer Ellen Fields, M.D. – Clinical Team Leader
Project Manager:	Kimberly Compton, R.Ph

Table of Contents

LIST OF TABLES

LIST OF FIGURES

1. EXECUTIVE SUMMARY

This application contains one phase 3 study (FX006-2014-008) and one phase 2b study (FX006-2014-006) to support the efficacy of Zilretta 40 mg for the treatment of pain of OA in the knee. These were randomized, double-blind, parallel group, placebo- or active-controlled clinical studies in patients with osteoarthritis (OA) of the knee,

Study FX006-2014-008 was a placebo-control (saline) and active-control (immediate-release triamcinolone acetonide 40mg) study. Patients were 40 years and older with OA for at least 6 months, and baseline average daily pain (ADP) of 5 - 9 on a 0-10 scale. A total of 486 patients were randomized using a 1:1:1 ratio to the three treatment arms. After the single intra-articular injection of study treatment, patients were followed for 24 weeks. The primary efficacy endpoint was change from baseline to week 12 in ADP. The planned analysis model was a longitudinal mixed effects model (MMRM) with fixed effects for treatment group, study week, treatment-by-week interaction, site and baseline covariate. Patient was included as a random effect. The primary goal was to test superiority of Zilretta 40mg versus placebo, which was successful ($p<0.001$). The second efficacy objective was to test superiority versus the immediate-release triamcinolone 40mg, which was not successful ($p=0.296$). All other comparisons in the planned hierarchical order were considered exploratory.

Study FX006-2014-006 was of similar design to the phase 3 study, enrolled the same patient population, collected the same efficacy assessments, and followed patients for 24 weeks. The key differences were inclusion a lower dose of Zilretta 20mg instead of the active-control arm and a smaller sample size (100 per group). The primary efficacy endpoint and analysis method were the same as in the phase 3 study but Zilretta 40 mg was not statistically significantly superior to placebo. The applicant did further *post hoc* analyses of this study to aid in the planning of the phase 3 study. In my opinion, this study supports efficacy as the results were similar to those observed in the phase 3 study. Lack of a significant difference between Zilretta 40 mg and placebo may be due to a larger placebo effect noted in this study.

2. INTRODUCTION

2.1 Overview

Flexion Therapeutics, Inc. has submitted a new drug application (NDA) for ZILRETTA, an extended release synthetic corticosteroid, for an indication as an intra-articular injection for the management of osteoarthritis pain of (b) (4) It is supplied as a (b) (4) kit containing 40 mg of sterile triamcinolone acetonide (TCA) extended release microsphere powder, 5 mL of sterile diluent, and a sterile vial adapter. The reference product is TCA in an immediate release (IR) formulation, which is approved for intra-articular injection for OA pain.

The applicant discussed the clinical development plan with DAAAP under IND# 111325. The 505(b)(2) pathway was agreed upon at the pre-IND meeting held on June 15, 2011. The Phase 3 protocols, dose, and clinical outcomes were discussed at the End of Phase 2 meeting on September 24, 2013. The statistical analysis plan was reviewed in June, 2015, and pre-NDA questions were addressed in writing in June 2016.

The application includes two clinical efficacy studies, FX006-2014-006 and FX006-2014-008, for consideration. I refer to these as studies 006 and 008, respectively. Study 006 is a Phase 2B study which included two dose arms of Zilretta (20 mg; 40mg) and a placebo arm. Study 008 included Zilretta 40mg, placebo, and an active-control (TCA IR) arm. Placebo control used a saline injection. Study 008 is the pivotal study to confirm efficacy, while Study 006 provides consistent supportive results. In both studies the primary goal was to compare the Zilretta 40mg treatment to placebo to provide evidence of efficacy. In Study 008, a secondary objective was to test for superiority of Zilretta 40 mg against the active-control arm. These two studies are the focus of this review. Although Study 006 did not demonstrate superiority for Zilretta 40mg versus placebo, Dr. Horn requested it be included in the statistical review to confirm direction and size of treatment effect. The efficacy results for the Zilretta 40 mg group were consistent in both studies. The placebo effect size, along with the smaller sample size, contributed to the insufficient evidence to conclude superiority in Study 006.

2.2 Data Sources

All data was supplied by the applicant to the CDER electronic data room (edr) in SAS transport format. The study reports, data, and documentation in the electronic submission are archived under the network path location: \\Cdsesub1\evsprod\NDA208845\0000.

In response to an Information Request, the applicant provided additional analyses of Study 006 to investigate the impact of site on treatment results. This was submitted on March 24, 2017, and is archived at \\Cdsesub1\evsprod\NDA208845\0008.

3. STATISTICAL EVALUATION

3.1 Data and Analysis Quality

The study data was submitted in standard formats, along with all documentation needed to complete my review. I was able to confirm the applicant's efficacy analyses for both studies. The data was clearly organized to conduct my own analyses without difficulty

3.2 Evaluation of Efficacy

3.2.1 Study FX006-2014-008

I will refer to this as study 008.

Study Design and Endpoints

Study 008 was a multicenter, randomized, double-blind, parallel group design. It included three treatment arms: Zilretta 40mg; placebo (saline); and active-control (TCA IR 40mg). It was conducted from January 2015 through January 2016 at 41 centers in the U.S., Canada, Australia, New Zealand, Hong Kong, and the European Union.

Eligible patients were adults, ages 40 and older, with symptoms of OA of the knee for at least 6 months prior to enrollment in the study. Maximum body mass index (BMI) was limited to ≤ 40 kg/m^2. The patient identified the knee with more pain as the index knee. At baseline, pain in the index knee had to occur on >15 days in the previous month, and have a weekly mean average daily pain score of 5 to 9 (0-10 numeric rating scale) during the 7 days prior to enrollment.

A total of 486 patients were randomized at a 1:1:1 ratio to the three treatment arms. Randomization was stratified by mean of the mean average daily pain scores at baseline, with the

following classifications: 5 to < 6, 6 to < 7, and ≥ 7. After screening and randomization, patients were scheduled to receive the intra-articular injection into the index knee. Two patients (1 placebo; 1 TCA IR) were randomized but did not receive the treatment injection. There were 161 patients treated in the Zilretta 40 mg arm, 162 in the placebo arm, and 161 in the TCA IR 40 mg arm.

The injection is prepared by an unblinded pharmacist, and is injected by an unblinded doctor. The contents of the syringe will be screened from the patient. To maintain study blind, the unblinded staff had no further contact with the patient or blinded assessors after the injection had been completed. All interactions with patients including physical exams, and efficacy and safety assessments were carried out by a blinded assessor.

The average daily pain score, along with other pain assessments and use of rescue medication, were collected daily by IVRS. Patients were permitted to take acetaminophen/paracetamol (up to a maximum of 3 grams per day) as rescue medication on an as needed basis. Patients returned to the study center for follow-up visits every four weeks after the procedure through Week 24. The Western Ontario and McMaster Universities (WOMAC) Osteoarthritis Index pain, stiffness and function domains, and Patient Global Impression of Change (PGIC) outcomes were collected at each visit.

The primary efficacy variable was defined as change from Baseline to Week 12 in the weekly mean of the ADP intensity score. The applicant planned a closed stepdown testing order for the following secondary endpoints:

- Change from Baseline to Week 8 in the weekly mean of the average daily pain intensity score for Zilretta 40mg vs. TCA IR 40mg.
- WOMAC function domain: change from Baseline to Week 12 for Zilretta vs. placebo
- PGIC at Week 12 for Zilretta vs. placebo

The Full Analysis Set (FAS) was defined as all patients who received study medication and have Baseline and at least one post-dose pain evaluation. This was designated as the primary analysis population for efficacy endpoints. The Intent to Treat (ITT) patient population was defined as all randomized patients. The ITT population is generally preferred for efficacy analyses. The applicant provided sensitivity analyses using the ITT population, which only differed by two patients from the FAS set.

The sample size of 150 per arm (450 total) was planned based on two endpoints. The primary comparison was of Zilretta 40 mg to placebo for Change from Baseline to Week 12 in ADP, and the planned secondary comparison of Zilretta 40 mg to active-control group for Change from Baseline to Week 8 in ADP. The applicant assumed a true underlying between-group difference of 1.0 unit at Week 12 versus placebo, 0.9 units at Week 8 versus active-control, with a standard deviation of 2.4; power ≥ 90%, α=0.05, and two-sided tests.

Patient Disposition, Demographic and Baseline Characteristics

A total of 486 patients were randomized, and 484 received study treatment. As shown in Table 1, two did not receive treatment, both for reasons unrelated to treatment. Overall, only 16 patients (3%) discontinued from the study prior to Week 12. Patients were followed for a total of 24 weeks post-treatment.

Table 1: Patient Disposition (Study 008)

	Zilretta 40mg N=161	**Placebo N=162**	**TCA IR 40mg N=161**
Randomized (ITT)	161 (100%)	163 (100%)	162 (100%)
Received Study Treatment (FAS)	161 (100%)	162 (99%)	161 (99%)
Discontinued Prior to Week 12	5 (3%)	8 (5%)	3 (2%)
Adverse Event	0	0	0
Withdrew Consent	2 (1%)	1 (1%)	2 (1%)
Withdrawn by Investigator	0	2 (1%)	0
Lost to Follow-up	0	1 (1%)	0
Lack of Efficacy	2 (1%)	4 (2%)	1 (1%)
Other	1 (1%)	0	0
Discontinued After Week 12 through Week 24	12 (7%)	6 (4%)	9 (6%)
Adverse Event	0	1 (1%)	1 (1%)
Withdrew Consent	3 (2%)	2 (1%)	1 (1%)
Withdrawn by Inv. Or Prot. Violation	1 (1%)	1 (1%)	0
Lost to Follow-up	1 (1%)	2 (1%)	0
Lack of Efficacy	5 (3%)	0	3 (2%)
Other	2 (1%)	0	4 (2%)

Source: CSR Table 10 and ADSL.xpt dataset
All percentages are calculated based on Randomized N per group as denominator.

Baseline Demographics

The three treatment groups were well balanced with respect to relevant demographic and baseline characteristics as shown in Table 2.

Table 2: Demographic and Baseline Pain Characteristics (Study 008)

All Treated (FAS)	**Zilretta 40mg N=161**	**Placebo N=162**	**TCA IR 40mg N=161**
Age (years)			
Mean (SD)	61 (9.5)	62 (8.9)	62 (10.0)
Range	40-83	40-83	40-85
Gender			
Female	103 (64%)	96 (59%)	97 (60%)
Male	58 (36%)	66 (41%)	64 (40%)
Race			
Caucasian	130 (81%)	144 (89%)	131 (81%)
Black	9 (6%)	4 (2%)	12 (7%)
Asian	18 (11%)	13 (8%)	16 (10%)
Other	4 (2%)	1 (1%)	2 (1%)
Region			
United States	68 (42%)	71 (44%)	63 (39%)
Canada	17 (11%)	18 (11%)	17 (11%)
Pacific (AUS; NZ; HK)	41 (25%)	38 (23%)	43 (27%)
Europe	35 (22%)	35 (22%)	37 (23%)
BMI (kg/m^2)			
Mean (SD)	30 (5.0)	30 (4.7)	30 (4.8)
Range	19 – 40	19 – 40	20 – 40
BMI Category:			
Normal (18.5 to 24.9)	28 (17%)	22 (14%)	25 (16%)
Overweight (25.0 to 29.9)	57 (35%)	58 (36%)	53 (33%)
Obesity Class I (30.0 to 34.9)	45 (28%)	52 (32%)	55 (34%)
Obesity Class II (35.0 to 39.9)	31 (19%)	30 (19%)	28 (17%)
Baseline Avg. Daily Pain (0-10 NRS)			
Mean (SD)	6.3 (0.9)	6.3 (1.0)	6.2 (0.9)
Range	5.0 – 8.5	5.0– 8.9	5.0 – 9.0
Baseline Pain Strata			
5 to <6	68 (42%)	68 (42%)	68 (42%)
6 to <7	52 (32%)	52 (32%)	51 (32%)
≥ 7	41 (25%)	42 (26%)	42 (26%)

Sources: ADSL.xpt and ADPAI.xpt datasets

Statistical Methodologies

The statistical analysis planned in the protocol used a longitudinal mixed effects model (MMRM) with fixed effects for treatment group, study week, treatment-by-week interaction, site and baseline covariate. Patient was included as a random effect. Treatment differences were estimated using least squares means along with 95% confidence intervals.

This model assumes missing at random and includes only observed data; missing data are not imputed. This was discussed with the applicant in the statistical review of the protocol. It was decided that, in this particular treatment scenario, where patients cannot discontinue treatment after the injection has be administered, this model would be acceptable. The applicant also agreed to provide sensitivity analyses to assess the impact of the missing data. Those alternate approaches included: 1) imputing baseline-observation carried forward (BOCF) for patients discontinuing due to lack of efficacy and last-observation-carried-forward (LOCF) for patients discontinuing due to AE or other reasons; 2) ITT population with missing data for those who were randomized but not treated imputed using BOCF; and 3) multiple imputation.

The applicant planned a hierarchical testing procedure with a prespecified order for comparisons, each to be tested at two-sided α=0.05. The order of comparisons was:

- Primary: change from baseline to Week 12 in weekly mean of the ADP scores for Zilretta 40mg versus placebo.
- Area under the effect curve (AUE) of change from baseline in weekly mean of the ADP scores from baseline to Week 12 for Zilretta 40mg relative to placebo.
- AUE of change from baseline in weekly mean of the ADP scores from baseline to Week 12 for Zilretta 40mg relative to TCA IR.
- Change from baseline to Week 12 in the weekly mean of the ADP scores for Zilretta 40mg relative to TCA IR
- AUE of change from baseline in weekly mean of the ADP scores from baseline to Week 24 for Zilretta 40mg relative to placebo.

The clinical team preferred the change from baseline to Week 12 for the ADP scores over the area under the effect (AUE) curve measure.

The applicant reported results for other secondary outcomes: WOMAC pain; WOMAC function; WOMAC stiffness; Patient Global Impression of Change (PGIC); and proportion of patients experiencing improvement of >30% or >50% in ADP. These were not planned in the hierarchical testing procedure. The clinical reviewer, Dr. Horn was interested in the results for WOMAC function subscale domains and PGIC, as these are relevant to assessing the signs and symptoms of OA of the knee She also requested the WOMAC pain subscale to check consistency with the ADP score results. The PGIC tool is a 7-point Likert scale which asks patients "Since the start of the study, my overall status is:" with 1=very much improved to 7= very much worse. The applicant only planned to test treatment of pain of OA, not the broader indication of signs and symptoms, (b) (4)

Results and Conclusions

Table 3 presents the applicant's results, along with additional endpoints of clinical interest. As there were only a few patients that discontinued, and not much intermittent missing data for the efficacy assessments through Week 12, the various sensitivity analyses conducted to assess the impact of missing data were supportive of the primary analysis.

For the primary efficacy endpoint, the Zilretta 40mg treatment group was statistically significantly better than the placebo treatment group ($p<0.0001$). The second comparison of interest, Zilretta 40mg vs. TCA IR 40mg, did not show sufficient evidence of a treatment difference. Therefore, the testing procedures stopped and all additional comparisons were considered exploratory

Even though the endpoints requested by Dr. Horn, WOMAC function subscale and PGIC, were not part of the applicant's hierarchical testing procedure, I present the results because they are considered clinically relevant. On both scales, low values indicate improvement. As shown in Table 3, the Zilretta 40mg group had more favorable results for both of these outcomes than the placebo group.

Table 3: Efficacy Analysis Results (Study 008)

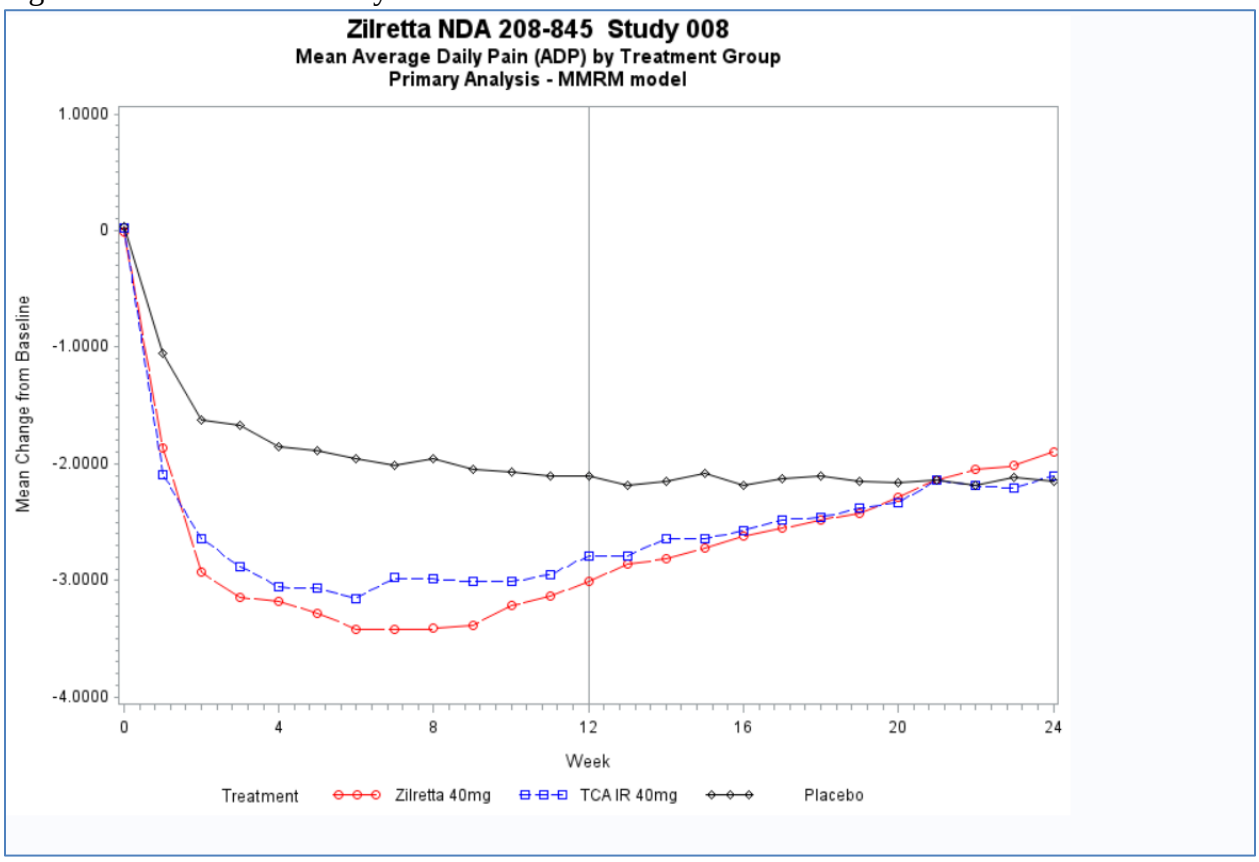
FAS Pt. Popln.		Zilretta 40mg N=161	Placebo N=162	TCA IR 40mg N=161
Change from Baseline to Week 12 in Average Daily Pain	Baseline Mean (SD)	6.3 (0.9)	6.3 (1.0)	6.2 (0.9)
	Chg Wk 12: Mean (SD)	-3.1 (2.4)	-2.2 (2.1)	-2.8 (2.1)
	LS Mean (SE)	-3.1 (0.2)	-2.1 (0.2)	-2.9 (0.2)
	LSM Diff. from Placebo	-0.98		
	2-sided p-value	<0.0001		
	LSM Diff. from TCA IR	-0.26		
	2-sided p-value	0.296		
WOMAC pain subscale {Average of five items; 0=None; 4=Extreme}	Baseline Mean (SD)	2.0 (0.5)	2.0 (0.5)	2.0 (0.5)
	Chg Wk 12: Mean (SD)	-0.9 (0.9)	-0.6 (0.7)	-0.7 (0.8)
	LS Mean (SE)	-0.9 (0.07)	-0.5 (0.07)	-0.7 (0.07)
	LSM Diff. from Placebo	-0.37		
	2-sided p-value	<0.0001		
WOMAC function subscale {Average of 17 items; 0=None; 4=Extreme}	Baseline Mean (SD)	2.1 (0.6)	2.1 (0.5)	2.1 (0.6)
	Chg Wk 12: Mean (SD)	-0.9 (0.8)	-0.6 (0.7)	-0.7 (0.8)
	LS Mean (SE)	-0.9 (0.07)	-0.6 (0.07)	-0.7 (0.07)
	LSM Diff. from Placebo	-0.38		
	2-sided p-value	<0.0001		
Patient Global Impression of Change (PGIC) {7-point scale; 1=very much improved; 7=very much worse}	Week 12: Mean (SD)	2.6 (1.4)	3.1 (1.3)	2.8 (1.4)
	LS Mean (SE)	2.6 (0.1)	3.2 (0.1)	2.8 (0.1)
	LSM Diff. from Placebo	-0.6		
	2-sided p-value	0.0003		
Responders (based on % reduction in pain at Week 12)	>30% Improvement: N (%)	103/161 (64%)	80/162 (49%)	104/161 (65%)
	>50% Improvement: N (%)	80/161 (50%)	56/162 (35%)	75/161 (47%)

Source: Clinical Study Report Tables 14.2.1.1; 14.2.2.1; 14.2.2.4, 14.2.3.1.1; 14.2.6.1, and ADPAI.xpt dataset

[a] The adjusted means and p-values were obtained from MMRM model including effects for treatment, week, treatment-by-week, site, and baseline pain strata.

Figure 1 shows the mean ADP scores by week for the entire 24-week follow-up period. The vertical line at Week 12 represents the primary timepoint for efficacy assessments for treatment of OA pain. The pain relief from the Zilretta 40mg and the TCA-IR 40mg groups was similar across the entire timeframe.

Figure 1: Pain Curves – Study 008



Source: Reviewer

Based on the results, study 008 does provide sufficient evidence to support the efficacy of Zilretta 40mg for the treatment of pain OA of the knee. It does not support any comparative claims to the active-control, or claims based on any secondary endpoints.

3.2.2 Study FX006-2014-006

I will refer to this as study 006.

Study Design and Endpoints

Study 006 was a multicenter, randomized, double-blind, parallel group design. It included three treatment arms: Zilretta 40mg; Zilretta 20mg, and placebo. It was conducted from April 2014 to August 2015, in 48 sites in the U.S. and Canada.

Study 006 was conducted before study 008, and the results of study 006 were used to plan study 008, which is appropriate. The study design, patient population, treatment duration were similar in both studies. However, study 006 included a low-dose Zilretta 20mg treatment arm instead of an active-control arm. The applicant prespecified that efficacy comparisons of the Zilretta 40mg arm to placebo would be completed prior to comparisons of the Zilretta 20 mg arm to placebo. The planned hierarchical order for testing efficacy endpoints was:

- The primary endpoint was the change from Baseline to Week 12 in weekly mean of the average daily pain intensity score s for Zilretta 40mg versus placebo.
- The key secondary efficacy endpoints included the following, listed in order of step-down testing beginning at the top of the list:
    - WOMAC® Osteoarthritis Index C (function): Change from Baseline to Week 12.
    - PGIC: 7-point scale at Week 12.
    - Change from Baseline to Week 16 and then Week 20 and then Week 24 in the weekly mean of the average daily (24-hour) pain intensity scores on a 0 to 10 NRS.

As in Study 008, patients were age 40 and older, BMI ≤40 kg/m^2, with symptoms of OA of the knee for at least 6 moths, and baseline average daily pain of 5 to 9 on the 0-10 NRS scale. Patients were randomized at a 1:1:1 ratio to the three treatment groups. Baseline pain categories were not used as strata for randomization in this study.

The sample size of 92 per arm (276 total) was planned based on the primary comparison of Zilretta 40 mg to placebo for Change from Baseline to Week 12 in ADP. The applicant assumed a true underlying between-group difference of 1.0 unit at Week 12 versus placebo, with a standard deviation of 2.4; power = 80%, α=0.05, and two-sided tests.

The applicant is requesting consideration of the Zilretta 40mg dose in this submission. Study 006 included a lower dose (20mg) Zilretta treatment arm as well as placebo. Only the comparisons of Zilretta versus placebo are of interest in this review, but I included the results for the lower dose group in all tables for completeness.

Patient Disposition, Demographic and Baseline Characteristics

A total of 310 patients were randomized to the three treatment arms. Four patients, all in the placebo arm, discontinued after randomization but before receiving treatment, for reasons unrelated to treatment. As shown in Table 4, fewer patients discontinued from the Zilretta 40mg group prior to Week 12, while more discontinued due to lack of efficacy in the other two groups during that timeframe. Overall, 24 treated patients (8%) discontinued from the study prior to Week 12. Patients were followed for a total of 24 weeks post-treatment

Table 4: Patient Disposition (Study 006)

	Zilretta 20mg N=102	**Zilretta 40mg N=104**	**Placebo N=104**
Randomized (ITT)	102 (100%)	104 (100%)	104 (100%)
Withdrew after Randomization And Before Treatment:			
Patient withdrew consent			2 (1%)
Did not return for IA procedure			1 (1%)
Prohibited Concomitant Medication			1 (1%)
Received Study Treatment (FAS)	102 (100%)	104 (100%)	100 (96%)
Discontinued Prior to Week 12	9 (9%)	3 (3%)	12 (12%)
Adverse Event	1 (1%)	2 (2%)	1 (1%)
Withdrew Consent	1 (1%)	0	4 (4%)
Withdrawn by Investigator	0	0	0
Lost to Follow-up	3 (3%)	0	1 (1%)
Lack of Efficacy	3 (3%)	1 (1%)	3 (3%)
Other	1 (1%)	0	3 (3%)
Discontinued After Week 12 through Week 24	10 (10%)	14 (14%)	7 (7%)
Adverse Event	3 (3%)	2 (2%)	0
Withdrew Consent	0	1 (1%)	1 (1%)
Withdrawn by Inv. Or Prot. Violation	1 (1%)	0	1 (1%)
Lost to Follow-up	1 (1%)	3 (3%)	0
Lack of Efficacy	3 (3%)	5 (5%)	2 (2%)
Other	2 (2%)	3 (3%)	3 (3%)

Source: CSR Table 9 and ADSL.xpt dataset
All percentages are calculated based on Randomized N per group as denominator.

Baseline Demographics

The three treatment groups were well balanced with respect to most demographic and baseline characteristics, as shown in Table 5. Two exceptions were the distribution by gender in the Zilretta 40 mg group, and the distribution by BMI category in the placebo group. Neither of these was a statistically significant imbalance. Both were investigated for subgroup by treatment effects (see Table 9) and did not impact the conclusions.

Table 5: Demographic and Baseline Pain Characteristics (Study 006)

All Randomized	Zilretta 20mg N=102	**Zilretta 40mg N=104**	**Placebo N=104**
Age (years)			
Mean (SD)	58 (8)	59 (8)	60 (8)
Range	43 – 83	41 – 79	40 – 80
Gender			
Female	62 (61%)	51 (49%)	62 (60%)
Male	40 (39%)	53 (51%)	42 (40%)
Race			
Caucasian	81 (79%)	85 (82%)	86 (83%)
Black	17 (17%)	16 (15%)	15 (14%)
Asian	2 (2%)	3 (3%)	2 (2%)
Other	2 (2%)	0	1 (1%)
Country			
United States	84 (82%)	88 (85%)	93 (89%)
Canada	18 (18%)	16 (15%)	11 (11%)
BMI (kg/m^2)			
Mean (SD)	31 (4.9)	31 (4.6)	31 (5.1)
Range	17 – 39	20 – 42	19 – 43
BMI Category:			
Normal (18.5 to 24.9)	14 (14%)	10 (10%)	10 (10%)
Overweight (25.0 to 29.9)	32 (31%)	35 (34%)	35 (34%)
Obesity Class I (30.0 to 34.9)	37 (36%)	39 (38%)	29 (28%)
Obesity Class II (35.0 to 39.9)	19 (19%)	20 (19%)	30 (29%)
Baseline Avg. Daily Pain (0-10 NRS)			
Mean (SD)	6.6 (1.0)	6.5 (1.0)	6.7 (1.1)
Range	5.0 – 8.9	5.0 – 8.4	5.0 – 9.0
Baseline Pain Category			
5 to <6	32 (31%)	34 (33%)	32 (31%)
≥ 6	70 (69%)	70 (67%)	72 (69%)

Sources: ADSL.xpt and ADPAI.xpt datasets

Statistical Methodologies

The statistical analysis used the similar methodology as in Study 008. The only difference was that the term for site was not included in the model for Study 006. They used a longitudinal mixed effects model (MMRM) with fixed effects for treatment group, study week, treatment-by-week interaction and baseline covariate. Patient was included as a random effect. Treatment differences were estimated using least squares means along with 95% confidence intervals.

This model assumes missing at random and includes only observed data; missing data are not imputed. This was discussed with the applicant in the statistical review of the protocol. It was decided that, in this particular treatment scenario, where patients cannot discontinue treatment after the injection has be administered, this model would be acceptable. The applicant also agreed to provide sensitivity analysis to assess the impact of the missing data by imputing BOCF for patients discontinuing due to lack of efficacy and LOCF for patients discontinuing due to AE or other reasons.

In the study report, testing for secondary endpoints was presented in a different order than planned in the protocol. Since the first comparison failed to show superiority, all other comparisons are exploratory only. I have included the efficacy endpoints which Dr. Horn requested for her own assessment of the Study 006 results.

All comparisons of Zilretta 40mg to placebo would be completed before comparing Zilretta 20mg to placebo on the same ordered list of endpoints.

Results and Conclusions

Table 6 presents the results for the efficacy endpoints. The planned efficacy analyses of Study 006 were not successful for the first comparison of Zilretta 40mg to placebo on mean change from baseline to week 12 for ADP (p=.08). Therefore all other comparisons were for exploratory purposes only.

The results were submitted to this application as supportive evidence. Dr. Horn's main question was if Study 006 was consistent and supportive of the confirmatory results in study 008. The Zilretta 20mg treatment arm is included for completeness, but is not being considered for approval.

Table 6: Efficacy Analysis Results (Study 006)

FAS Patients		Zilretta 20mg N=102	**Zilretta 40mg N=104**	**Placebo N=100**
Change from Baseline to Week 12 in Average Daily Pain	Baseline Mean (SD)	6.6 (1.0)	6.5 (1.0)	6.7 (1.1)
	Chg Wk 12: Mean (SD)	-2.7 (2.3)	-3.2 (2.3)	-2.7 (2.4)
	LS Mean (SE)	-2.6 (0.2)	-3.1 (0.2)	-2.5 (0.2)
	LSM Diff. from Placebo		-0.58	
	2-sided p-value		0.08	
WOMAC pain subscale	Baseline Mean (SD)	2.3 (0.6)	2.1 (0.6)	2.3 (0.7)
	Chg Wk 12: Mean (SD)	-0.9 (0.9)	-0.9 (0.8)	-0.9 (0.9)
	LS Mean (SE)	-0.8 (0.08)	-1.0 (0.08)	-0.8 (0.08)
	LSM Diff. from Placebo		-0.17	
	2-sided p-value		0.15	
WOMAC function subscale	Baseline Mean (SD)	2.3 (0.7)	2.1 (0.6)	2.3 (0.6)
	Chg Wk 12: Mean (SD)	-0.9 (1.0)	-0.9 (0.8)	-0.9 (0.9)
	LS Mean (SE)	-0.8 (0.09)	-1.0 (0.08)	-0.8 (0.09)
	LSM Diff. from Placebo		-0.21	
	2-sided p-value		0.08	
Patient Global Impression of Change (PGIC)	Week 12: Mean (SD)	2.4 (1.3)	2.5 (1.4)	2.6 (1.3)
	LS Mean (SE)	2.5 (0.1)	2.6 (0.1)	2.7 (0.1)
	LSM Diff. from Placebo		-0.1	
	2-sided p-value		0.55	
Responder (based on % reduction in pain at Week 12)	>30% Improvement: N (%)	55/102 (54%)	62/104 (60%)	53/100 (53%)
	>50% Improvement: N (%)	41/102 (40%)	45/104 (43%)	36/100 (36%)

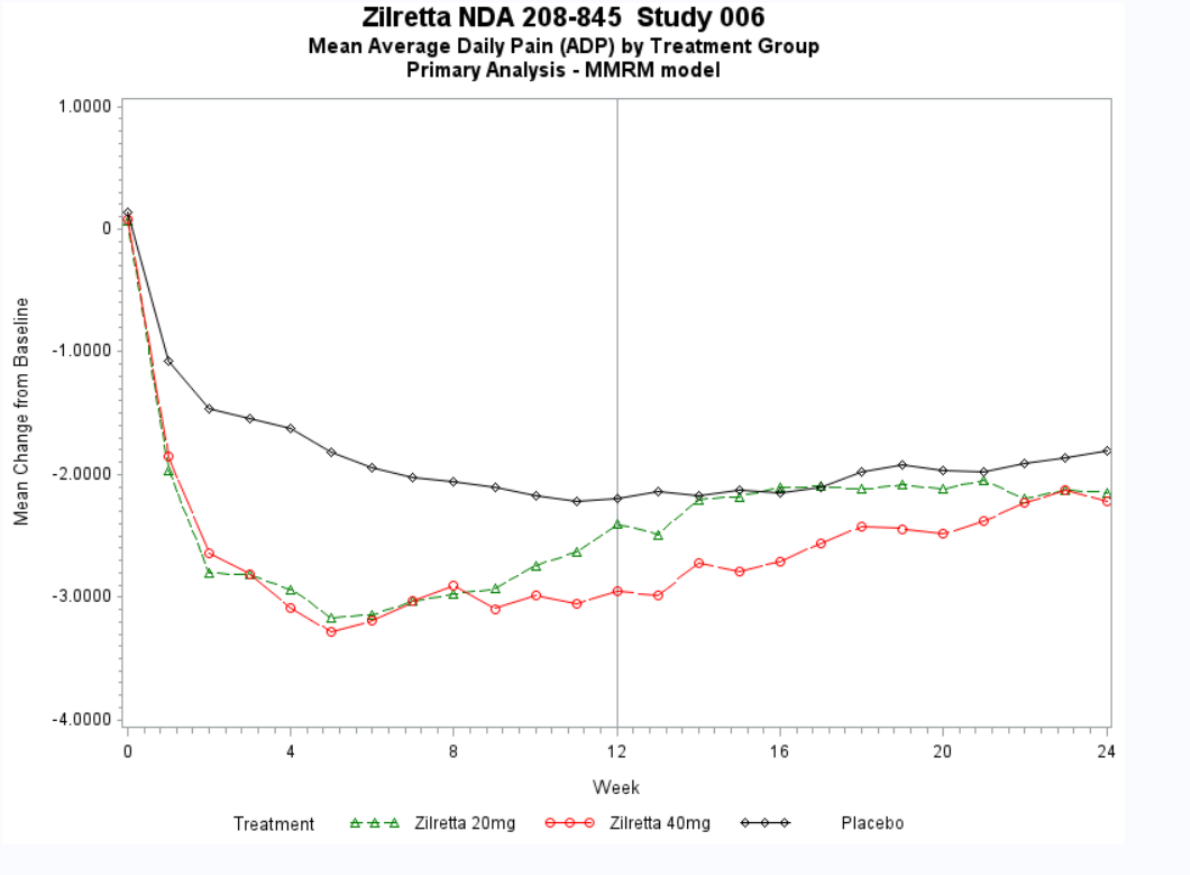
Source: Clinical Study Report Tables

[a] The adjusted means and p-values were obtained from MMRM model including effects for treatment, week, treatment-by-week, and baseline pain.

Dr. Horn asked me to consider potential explanations for Study 006 failing to demonstrate superiority versus placebo. The applicant's MMRM model approach, described above, is not the typical analysis for pain endpoints in a 12-week OA study. Instead I performed an analysis of covariance (ANCOVA) with terms for treatment and baseline pain. This method only includes the Week 12 data, with an appropriate imputation for missing data. Using the BOCF/LOCF imputation generated in the data sets by the applicant, the ANCOVA model results indicated Zilretta 40mg was superior to placebo (p=0.046), with the Zilretta group mean of -3.1 (0.23 SE) and placebo mean of -2.3 (0.23 SE). The estimated treatment effect size in the Zilretta group was consistent (-3.1) across both studies and either model. The estimated placebo response fluctuated slightly higher in Study 006 than in Study 008. Aside from that, I did not identify any other explanation for the failed superiority comparison in Study 006.

Figure 2 shows the mean average daily pain scores by week for the entire 24-week follow-up period. The vertical line at Week 12 represents the primary timepoint for efficacy assessments for treatment of OA pain. The pain relief from the Zilretta 40mg and the placebo groups remain separate through the 12 weeks after treatment is administered. This is an exploratory result only and does not imply statistical significance at any timepoint.

Figure 2: Change in Average Daily Pain (Study 006)



Source: ADPAI.xpt dataset

Based on the efficacy results, study 006 does provide evidence to support the Zilretta 40mg for the treatment of pain of OA of the knee. The treatment effect for the Zilretta 40 mg arm are of consistent magnitude and direction as in Study 008. The notable difference in the results across the two studies is the larger treatment response in the placebo group in Study 006.

3.3 Evaluation of Safety

Dr. Horn completed the safety review for this study. She did not request any additional safety analyses.

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Age, Gender, Race, Region, and Body Mass Index (BMI)

I produced exploratory analyses for the primary efficacy endpoint by age group, gender, race, region, and Body Mass Index (BMI) (See Tables 7 and 8).

Table 7: Subgroup Analyses: (Study 008) – Reviewer's Results; All Treated

Change from Baseline to Week 12 Average Daily Pain N Mean (SD)	**Study 008**		
	Zilretta 40mg N=161	**Placebo N=162**	**TCA IR 40mg N=161**
Age group			
≤ 65 years	N=107 -3.0 (2.6)	N=99 -2.2 (2.1)	N=99 -2.7 (2.1)
> 65 years	N=54 -3.0 (2.3)	N=63 -2.0 (2.1)	N=62 -3.0 (2.2)
Sex			
Female	N=103 -3.3 (2.4)	N=96 -2.1 (2.1)	N=97 -3.0 (2.1)
Male	N=58 -2.6 (2.5)	N=66 -2.2 (2.1)	N=64 -2.5 (2.1)
Race			
Caucasian	N=130 -3.2 (2.4)	N=144 -2.1 (2.1)	N=131 -2.8 (2.2)
Non-Caucasian	N=31 -2.2 (2.9)	N=18 -2.0 (1.8)	N=30 -3.1 (1.9)

Table 7 (cont.): Subgroup Analyses: (Study 008) – Reviewer's Results; All Treated

Change from Baseline to Week 12 Average Daily Pain N Mean (SD)	**Zilretta 40mg N=161**	**Placebo N=162**	**TCA IR 40mg N=161**
Region			
United States	N=68 -3.5 (2.7)	N=71 -1.9 (2.2)	N=63 -2.9 (2.2)
Canada	N=17 -3.6 (2.3)	N=18 -1.7 (1.8)	N=17 -4.1 (1.9)
Pacific	N=34 -2.1 (2.5)	N=36 -2.2 (2.0)	N=40 -2.2 (2.1)
Europe	N=35 -2.8 (1.8)	N=35 -2.6 (2.2)	N=38 -2.6 (2.0)
Body Mass Index (kg/m^2)			
≤ 24.9	N=28 -3.6 (2.4)	N=22 -2.2 (2.1)	N=25 -2.0 (1.8)
25.0 to 29.9	N=57 -2.7 (2.3)	N=58 -2.1 (2.1)	N=53 -3.0 (2.0)
30.0 to 34.9	N=45 -3.1 (2.5)	N=52 -2.2 (2.1)	N=55 -2.8 (2.4)
≥ 35.0	N=31 -3.1 (2.8)	N=30 -2.1 (2.3)	N=28 -3.3 (2.0)

Source: Reviewer

In Study 008, the only notable difference among the subgroups is that there was very little difference in the mean change in ADP across the treatment groups in the Pacific (Australia; New Zealand; Hong Kong) and Europe (Denmark, Estonia, Lithuania, Romania) regions. There was a statistically significant treatment by region interaction (p=0.03) but we have no further information to investigate this unforeseen association. The treatment by race and treatment by BMI interactions were not significant. Study 006 was conducted in the U.S. and Canada so it did not provide any insight into the Pacific or Europe regions.

Table 8: Subgroup Analyses: (Study 006) – Reviewer's Results; All Treated

Change from Baseline to Week 12 Average Daily Pain	**Study 006**		
N Mean (SD)	Zilretta 20mg N=102	**Zilretta 40mg** **N=104**	**Placebo** **N=100**
Age group			
≤ 65 years	N=83 -2.3 (2.3)	N=78 -3.1 (2.4)	N=78 -2.5 (2.4)
> 65 years	N=19 -2.8 (2.4)	N=26 -2.8 (2.4)	N=22 -2.3 (2.1)
Sex			
Female	N=62 -2.9 (2.5)	N=51 -3.0 (2.6)	N=61 -2.5 (2.2)
Male	N=40 -1.7 (1.6)	N=53 -3.0 (2.1)	N=39 -2.3 (2.7)
Race			
Caucasian	N=81 -2.5 (2.3)	N=85 -3.2 (2.5)	N=82 -2.2 (2.1)
Non-Caucasian	N=21 -2.2 (2.2)	N=19 -2.0 (1.6)	N=18 -3.4 (3.3)
Region			
United States	N=84 -2.5 (2.3)	N=88 -3.1 (2.3)	N=89 -2.4 (2.4)
Canada	N=18 -2.0 (2.3)	N=16 -2.7 (2.6)	N=11 -2.3 (2.4)
Body Mass Index (kg/m^2)			
≤ 24.9	N=14 -2.4 (2.1)	N=10 -2.4 (2.4)	N=10 -1.2 (1.6)
25.0 to 29.9	N=32 -2.8 (2.4)	N=35 -2.6 (2.2)	N=33 -2.3 (2.4)
30.0 to 34.9	N=37 -1.9 (2.0)	N=39 -2.7 (2.4)	N=29 -2.9 (2.5)
≥ 35.0	N=19 -2.8 (2.8)	N=20 -4.5 (2.1)	N=28 -2.5 (2.4)

Source: Reviewer

In Study 006, the mean change in ADP at Week 12 was consistently higher in the Zilretta 40 mg arm than the placebo arm for all subgroups except Non-Caucasian. The treatment by race interaction was not statistically significant. These are descriptive analyses only and are not

intended for inferential purposes. None of the results from the subgroup analyses give reason to question the overall efficacy results.

5. SUMMARY AND CONCLUSIONS

5.1 Statistical Issues and Collective Evidence

This application contained two prospectively planned, multicenter, randomized, double-blind, active- and placebo- controlled parallel arm clinical studies to support efficacy for Zilretta for intra-articular injection for the treatment of pain due to OA in the knee. Both were appropriately designed with the desired patient population. The primary efficacy endpoint was the change from baseline to Week 12 in mean ADP.

Study 006 was a Phase 2B study, and the planned MMRM model included terms for treatment, week, treatment-by-week, and baseline pain. The results did not show statistical significance for the Zilretta 40mg arm versus placebo ($p=0.08$). The applicant performed additional *post hoc* analyses and found that when a term for SITE was added to the model, the treatment comparison did show significance ($p=0.034$). Investigations into the site effect indicated that no single site or country (US / Canada) was driving the results.

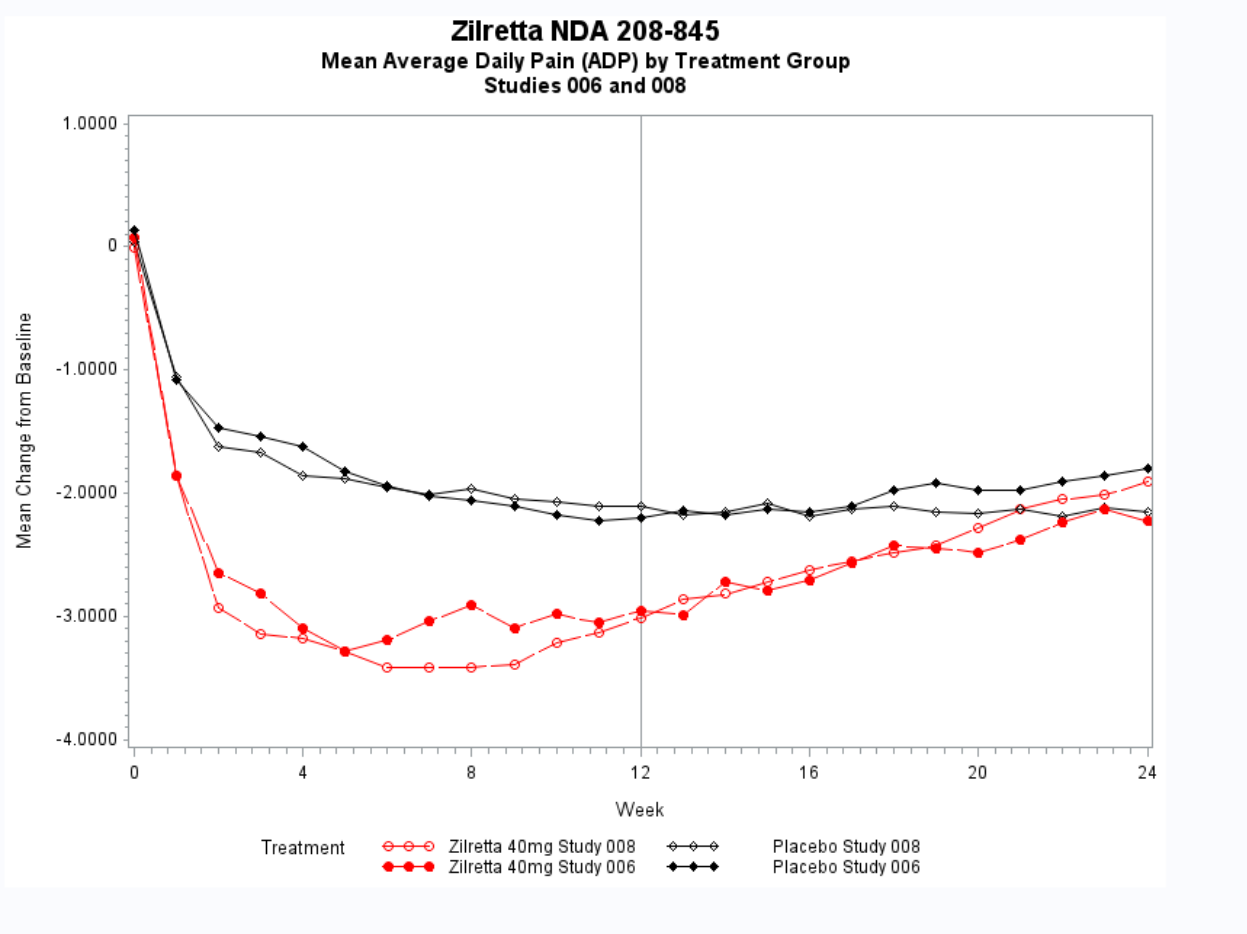
Based on the results of Study 006, the planned MMRM model in Study 008 included site, along with terms for treatment, week, treatment-by-week, and baseline pain. The results showed statistical significance ($p<0.001$) for the comparison of Zilretta 40 mg to placebo. The applicant also provided results for the MMRM model without the site term, and the results were consistent.

The only notable issue in my review was the model used to analyze the weekly pain data. The applicant planned in the protocols to apply a MMRM model. This is not typically our preferred approach because it does not account for missing data. The applicant provided alternative analyses applying either multiple imputation, or a BOCF/LOCF approach depending on reason for discontinuation. The results and conclusions from all three models were the same.

In this treatment scenario, a patient receives a single injection and then reports pain over 12 weeks. Patients may discontinue reporting to the study, but cannot choose to discontinue treatment after the injection has been completed. Therefore, concerns about the assumptions regarding Missing At Random (MAR) for the MMRM approach are minimized in this setting.

Figure 3 presents the pain curves for the Zilretta 40mg groups (red lines) and the placebo groups (black lines) for each of the two studies discussed in my review. I provided this to confirm that, in spite of Study 006 failing to show superiority, it did not contradict the results from Study 008. The vertical line represents Week 12, the primary timepoint for efficacy determination. The results are very similar visually, which suggests the placebo treatment effect in Study 006, along with the smaller sample size, affect the insufficient evidence to show superiority. The study 006 results provide support for the conclusion of efficacy for Zilretta 40mg.

Figure 3: Mean Change in Average Daily Pain Scores - Zilretta 40mg versus Placebo
Studies 006 and 008



Source: Reviewer

5.2 Conclusions and Recommendations

This application contains two prospectively planned, multicenter, randomized, double-blind, active- and placebo- controlled clinical studies to provide evidence of efficacy for Zilretta for intra-articular injection for the treatment of pain due to OA in the knee. Both were appropriately

designed with the desired patient population. Study 006 was a Phase 2B study, which did not achieve statistical significance, but did show consistent and supportive results for the Zilretta 40mg dose. Study 008 was the confirmatory Phase 3 study, which showed statistical significance in the reduction of ADP from baseline to Week 12, the primary efficacy endpoint.

The only notable issue in my review was the model used to analyze the weekly pain data. The applicant planned in the protocols to apply a MMRM model. This is not typically our preferred approach because it does not account for missing data. The applicant provided alternative analyses applying either multiple imputation, or a BOCF/LOCF approach depending on reason for discontinuation. The results and conclusions from all three models were the same.

Based on the consistent results from Studies 008 and 006, there is sufficient evidence to support Zilretta 40 mg for intra-articular injection for the treatment of pain due to osteoarthritis in the knee.

5.3 Labeling Review

.The applicant proposed the following indication statement:

(b) (4)

The following is the current proposed language for the Clinical Studies section:

The efficacy of ZILRETTA was demonstrated in a multi-center, international, randomized, double-blind, parallel-arm, placebo and active-controlled study in patients with osteoarthritis (OA) pain of the knee. A total of 484 patients (ZILRETTA (b) (4) mg, N=161; placebo [saline], N=162; active control [a crystalline suspension, immediate release formulation of triamcinolone acetonide 40 mg], N=161) were treated and followed for up to 24 weeks. Patients had a mean age of 62 (range 40 to 85 years); baseline demographics and disease characteristics were balanced across treatment arms. Twenty-five percent (25%) of patients had received at least one prior corticosteroid intra-articular injection more than 3 months prior to treatment. A total of 470 patients (97%) completed follow-up to Week 12, the time point for primary efficacy determination, and 443 (91.5%) completed to Week 24.

*The primary efficacy endpoint comparing ZILRETTA to placebo was change from baseline at Week 12 in the weekly mean of the Average Daily Pain intensity scores (ADP) as assessed by a 0-10 Numeric Rating Scale (NRS). ZILRETTA demonstrated a significant (b) (4) reduction in pain intensity at the primary endpoint (b) (4) ZILRETTA also demonstrated a (b) (4) reduction in pain intensity scores each week from Weeks 1 – 12 (**Figure***

(b) (4)


(b) (4)

Figure (b) (4): Weekly Change from Baseline to Week 12 in Average Daily Pain

ADP = Average Daily Pain; LS = Least Square; (b) (4)

The first paragraph, describing the Phase 3 study design and patient population is appropriate. (b) (4)

.

Figure (b) (4) shows the results for pain over the 12 week treatment period. The applicant proposed only showing the Zilretta and Placebo groups. Dr. Horn requested the applicant revise the figure to also show the active-control (TCA IR) group. (b) (4)

My suggested revisions for the second paragraph of the clinical studies section is:

> The primary efficacy endpoint comparing ZILRETTA to placebo was change from baseline at Week 12 in the weekly mean of the Average Daily Pain intensity scores (ADP) as assessed by a 0-10 Numeric Rating Scale (NRS). ZILRETTA demonstrated a significant (b) (4) reduction in pain intensity at the primary endpoint versus placebo. (b) (4)


(b) (4)

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

/s/

KATHERINE B MEAKER
09/01/2017

DAVID M PETULLO
09/01/2017
I concur.

STATISTICAL REVIEW AND EVALUATION
FILING REVIEW OF AN NDA/BLA

NDA #:	208-845/ 000
Related IND #:	111,325
Product Name:	Zilretta (extended release corticosteroid triamcinolone acetonide for injection)
Indication(s):	Intra-articular (IA) injection for the management of osteoarthritis (OA) (b) (4)
Applicant:	Flexion
Dates:	Received: December 8, 2016 PDUFA: October 6, 2017 Primary Review Due Date: September 1, 2017
Review Priority:	Standard
Biometrics Division:	Division of Biometrics II
Statistical Reviewer:	Kate Meaker, M.S.
Concurring Reviewers:	David Petullo, M.S.
Medical Division:	Division of Division of Anesthesia, Analgesia, and Addiction Products (DAAAP)
Clinical Team:	Pamela Horn, M.D.; Ellen Fields, M.D.
Project Manager:	Kimberly Compton, RPh

1. Summary of Efficacy/Safety Clinical Trials to be Reviewed

This submission includes two clinical studies to support the efficacy of Zilretta for the indication of treatment of pain of OA (b) (4) (See Table 1 for details). Study FX006-2014-008 is a Phase 3, randomized, double-blind, active- and placebo-control, parallel arm study. Study FX006-2014-006 is a Phase 2b, randomized, double-blind, placebo-control, parallel arm study.

Table 1: Summary of trials to be reviewed

Trial ID	Design*	Treatment/ Sample Size	Endpoint/Analysis	Preliminary Findings
FX-2014-008	MC, R, DB, PG, AC and PC trial ; Knee OA; (24 weeks)	FX006 40mg N161 Placebo N=162 TCA-IR N=162	Primary: Change from Baseline to Week 12 in Avg. Daily Pain (0-10 NRS)	1°: Superiority vs. Placebo (p<0.001)
FX-2014-006	MC, R, DB, PG, PC trial ; Knee OA; (24 Weeks)	FX006 40mg N=104 FX006 20mg N=102 PlaceboN=104	Primary: Change from Baseline to Week 12 in Avg. Daily Pain (0-10 NRS)	1°: Superiority vs. Placebo (p=0.082)

* MC: multi-center, R: randomized, DB: double-blind, PG: parallel group, PC: placebo-controlled, AC: active-controlled

Additional Notes:

All the clinical studies were conducted in patients with OA of the knee. (b) (4) (b) (4) (b) (4) The clinical team will decide the final wording for the label.

The active comparator product, TCA IR 40mg has been approved since 1965. The indication for intra-articular injection for the reference product (Kenalog 40) is: "The intra-articular or soft tissue administration of Kenalog-40 Injection is indicated as adjunctive therapy for short-term administration (to tide the patient over an acute episode or exacerbation) in acute gouty arthritis, acute and subacute bursitis, acute nonspecific tenosynovitis, epicondylitis, rheumatoid arthritis, synovitis, or osteoarthritis." The label does not list clinical studies, and does not designate specific joints for the treatment of OA.

Study 008:

- Conducted Jan. 2015 – Jan. 2016
- US, Canada, Australia, NZ, Hong Kong, EU
- Subjects were ≥40 years old with OA of the knee
- Baseline average daily pain (ADP) intensity score between 5 and 9 on 0-10 VAS scale

Study 006:

- Conducted April 2014 – August 2015
- Sites were in US and Canada
- Subjects were ≥40 years old with OA of the knee
- Baseline average daily pain (ADP) intensity score between 5 and 9 on 0-10 VAS scale

2. Assessment of Protocols and Study Reports

Table 2: Summary of Information Based Upon Review of the Protocol(s) and the Study Report(s)

Content Parameter	Response/Comments
Designs utilized are appropriate for the indications requested.	**yes**
Endpoints and methods of analysis are specified in the protocols/statistical analysis plans.	**yes**
Interim analyses (if present) were pre-specified in the protocol with appropriate adjustments in significance level. DSMB meeting minutes and data are available.	**NA**
Appropriate details and/or references for novel statistical methodology (if present) are included (e.g., codes for simulations).	**NA**
Investigation of effect of missing data and discontinued follow-up on statistical analyses appears to be adequate.	**Yes;** **Section 9.7.1.10.2.1 in CSR-006;** **Section 9.7.1.9.1.1 in CSR-008**

3. Electronic Data Assessment

Table 3: Information Regarding the Data

Content Parameter	Response/Comments
Dataset location	\\cdsesub1:NDA208845/000/m5/datasets
Were analysis datasets provided?	Yes
Dataset structure (e.g., SDTM or ADaM)	ADAM and SDTM for studies 008 and 006; SDTM only for study 001
Are the define files sufficiently detailed?	yes
List the dataset(s) that contains the primary endpoint(s)	ADPAI.xpt
Are the *analysis datasets* sufficiently structured and defined to permit analysis of the primary endpoint(s) without excess data manipulation?	Yes
Are there any initial concerns about site(s) that could lead to inspection? If so, list the site(s) that you request to be inspected and the rationale.	No.
Safety data are organized to permit analyses across clinical trials in the NDA/BLA.	Yes;

* This might lead to the need for an information request or be a refuse to file issue depending on the ability to review the data.

4. Filing Issues

Table 4: Initial Overview of the NDA/BLA for Refuse-to-file (RTF):

Content Parameter	Yes	No	NA	Comments
Index is sufficient to locate necessary reports, tables, data, etc.	√			
ISS, ISE, and complete study reports are available (including original protocols, subsequent amendments, etc.)	√			
Safety and efficacy were investigated for gender, racial, and geriatric subgroups investigated.	√			
Data sets are accessible, sufficiently documented, and of sufficient quality (e.g., no meaningful data errors).	√			
Application is free from any other deficiency that render the application unreviewable, administratively incomplete, or inconsistent with regulatory requirements	√			

IS THE APPLICATION FILEABLE FROM A STATISTICAL PERSPECTIVE? Yes

5. Comments to be Conveyed to the Applicant

5.1. Refuse-to-File Issues

None

5.2. Information Requests/Review Issues

The following should be requested from the applicant:

1. Provide efficacy results BY SITE for both Study 006 and Study 008.
2. Discuss the sites which impacted the efficacy results in the *post hoc* analysis of Study 006.
3. Provide subgroup analyses (Gender; Age, Race, Region) for each efficacy study separately. The Complete Study Reports submitted do not include these. The ISE presents these results on pooled data only.

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

/s/

KATHERINE B MEAKER
02/17/2017

DAVID M PETULLO
02/23/2017
I concur.