

CENTER FOR DRUG EVALUATION AND RESEARCH

**APPLICATION NUMBER:
22511Orig1s000**

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/Serial Number: 22-511 / 00

Drug Name: Vimovo[®] (naproxen/esomeprazole magnesium) Tablets

Indication(s): Signs and symptoms of osteoarthritis, rheumatoid arthritis, and ankylosing spondylitis in patients at risk of developing NSAID associated gastric ulcers

Applicant: AstraZeneca

Date(s): Letter date: 6/30/09
PDUFA date: 4/30/10

Review Priority: Standard

Biometrics Division: Division of Biometrics II

Statistical Reviewer: Kate Meaker, M.S.

Concurring Reviewers: Dionne Price, Ph.D.

Statistical Reviewer for DGP: Freda Cooner, Ph.D. (DB3)

Medical Division: Division of Anesthesia and Analgesia Products (DAAP)
Consult from Division of Gastroenterology Products (DGP)

Clinical Team: Medical Officer: Jin Chen, M.D. (DAAP)
Medical Team Leader: Ellen Fields, M.D. (DAAP)

Project Manager: Anne Simon (DGP)

Keywords: Clinical studies; NDA review; Non-inferiority Comparisons

1. EXECUTIVE SUMMARY OF STATISTICAL FINDINGS	3
1.1 Conclusions and Recommendations	3
1.2 Brief Overview of Clinical Studies	3
1.3 Statistical Issues and Findings	5
2. INTRODUCTION	5
2.1 Overview	5
2.2 Data Sources	6
3. STATISTICAL EVALUATION	6
3.1 Evaluation of Efficacy	6
Study PN400-307 (conducted 4/08 to 12/08)	6
Design	6
Patient Disposition	9
Baseline Demographics	10
Efficacy Results	11
Study PN400-309 (conducted 4/08 to 12/08)	16
Design	16
Patient Disposition	16
Baseline Demographics	17
Efficacy Results	18
3.2 Evaluation of Safety	22
4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS	22
4.1 Gender, Race and Age	22
4.2 Other Special/Subgroup Populations	24
5. SUMMARY AND CONCLUSIONS	24
5.1 Statistical Issues and Collective Evidence	24
5.2 Label Issues	25
5.3 Conclusions and Recommendations	25
Signatures/Distribution List Page	26

1. EXECUTIVE SUMMARY OF STATISTICAL FINDINGS

1.1 Conclusions and Recommendations

The applicant submitted results from four Phase 3 clinical studies intended to assess the efficacy of Vimovo (naproxen/esomeprazole magnesium) Tablets. The applicant is seeking an indication for treatment of signs and symptoms of osteoarthritis, rheumatoid arthritis, and ankylosing spondylitis in patients at risk of developing NSAID-associated gastric ulcers.

This review will only cover two of the clinical studies (307 and 309) which were conducted in patients with osteoarthritis (OA) of the knee and assessed the efficacy of VIMOVO for treatment of the signs and symptoms of OA. Both studies included three double-blind treatment arms: Vimovo 500 mg/20 mg twice daily, Celecoxib 200 mg once daily, and placebo.

The applicant has requested the following language in the Clinical Studies section of the label (Section 14) which constitutes a comparative claim: (b) (4)

[REDACTED]

The results of the two studies were conflicting, (b) (4)
[REDACTED]. In both studies, the results indicate that VIMOVO was not non-inferior to the celecoxib arm, but was statistically significantly superior to placebo. Based on the comparisons to placebo in the two studies, there is sufficient evidence to support the efficacy of VIMOVO for the indication of treatment of OA signs and symptoms. (b) (4)
[REDACTED]

1.2 Brief Overview of Clinical Studies

The naproxen component is currently approved for the treatment of signs and symptoms of osteoarthritis, rheumatoid arthritis, and ankylosing spondylitis. Two of the clinical studies submitted assessed the efficacy of the combination product for this indication. Those studies (307 and 309) will be discussed in this review.

The esomeprazole magnesium component is currently approved for treatment of gastroesophageal reflux disease and risk reduction of NSAID-associated gastric ulcer. Two clinical studies (301 and 302) assessed the efficacy of the combination product for the latter indication. Those studies will be reviewed by Dr. Freda Cooner (Division of Biometrics 3) for the Division of Gastroenterology Products (DGP).

The applicant conducted two prospectively planned, randomized, double-blind, active- and placebo-controlled clinical studies to assess the efficacy of Vimovo for the treatment of signs and symptoms of osteoarthritis (OA). Both studies (307 and 309) had the same design, treatment groups, patient population, efficacy endpoints, planned sample size, and planned analyses.

Patients were adults, ages 50 and older, with a history of at least 6 months of osteoarthritis of the knee. They had to be on a stable dose of NSAIDs, COX-2 inhibitors, or other oral analgesic therapy for at least 6 weeks prior to screening. When the oral analgesic therapy was discontinued, patients who experienced an OA flare, defined as worsening of pain and patient global assessment, were eligible for randomization.

The three double-blind treatment arms were VIMOVO 500 mg/20 mg twice daily (bid), celecoxib 200 mg once daily (qd), and placebo. In each study, eligible patients were randomized using a ratio of 2:2:1 to the three treatment groups.

In both protocols, the applicant stated the primary objective was to demonstrate that VIMOVO was non-inferior to celecoxib 200 mg qd on three primary endpoints: WOMAC pain subscale, WOMAC function subscale, and Patient Global Assessment. All three endpoints are measured on 0-100 mm VAS scales. The applicant planned to show efficacy on all three endpoints and did not plan any statistical adjustment for multiplicity. These have historically been the three efficacy endpoints required by the Agency for the indication of the treatment of signs and symptoms of OA.

Patients were treated for 12 weeks. The timepoint of interest was the change from baseline to Week 12 for the three efficacy measures. These studies did not assess the incidence of gastric ulcers.

The planned primary analysis used an ANCOVA model with terms for treatment and baseline pain as the covariate. The applicant's stated hypothesis was non-inferiority of the VIMOVO treatment group to the celecoxib treatment group on all three endpoints. Non-inferiority was assessed using 95% confidence intervals on the difference between the VIMOVO and celecoxib groups.

The applicant stated in the protocols that a non-inferiority margin of 10mm on the 0-100 mm VAS scales would be used for the comparisons. This was not agreed to by the Agency prior to conducting the studies. Discussion with the Agency on June 10, 2008, described the factors of the analysis which would impact the NI conclusions, including the treatment effect sizes and consistency of treatment response.

1.3 Statistical Issues and Findings

The main issue is the applicant's planned non-inferiority comparisons to celecoxib, with a non-inferiority margin of 10 mm on the three 0-100 mm VAS scales. The VAS scales used in studies 307 and 309 were somewhat different questions, with different outcome scales, than were used to measure efficacy versus placebo in the original application for celecoxib (NDA 20-998; December 29, 1998). The clinical studies submitted to support efficacy of celecoxib used 11-point Likert scales, and the questions were worded differently. The information from the NDA review of celecoxib was not useful toward determining if the proposed 10 unit NI margin was reasonable. AstraZeneca, the applicant for this NDA, did not provide justification to support the 10 unit NI margin.

Since the two studies submitted for this application included both a celecoxib arm and a placebo arm, we evaluated the observed treatment effect sizes for celecoxib versus placebo in these studies as a potential source of information on what an appropriate NI margin would be. In Study 307, the celecoxib treatment effect sizes ranged from 6-7 mm versus placebo. In Study 309, the celecoxib treatment group was not significantly different from placebo, with treatment effect sizes of 1-1.5 mm. These results indicated that the proposed 10 mm NI margin was not reasonable.

2. Introduction

2.1 Overview

VIMOVO consists of two active drug ingredients. Naproxen is a non-steroidal anti-inflammatory drug (NSAID) available in strengths of 250 mg, 375 mg, and 500 mg for oral administration. It is currently approved for treatment of signs and symptoms of osteoarthritis, rheumatoid arthritis, juvenile arthritis, ankylosing spondylitis, tendonitis, bursitis, and acute gout. It is used to reduce swelling and treat pain.

Esomeprazole magnesium (Nexium) is a proton pump inhibitor (PPI) which reduces gastric acid secretion. It is available as an extended-release capsule in strengths of 20 mg or 40 mg for oral administration. The currently approved indications are: treatment of gastroesophageal reflux disease; risk reduction of NSAID-associated gastric ulcer; *H. pylori* eradication to reduce the risk of duodenal ulcer recurrence; and pathological hypersecretory conditions including Zollinger-Ellison Syndrome.

The use of NSAIDs has a recognized risk of gastrointestinal adverse events. The naproxen label (and other NSAID labels) includes the following warning:

NSAIDs cause an increased risk of serious gastrointestinal adverse events including bleeding, ulceration, and perforation of the stomach or intestines, which can be fatal. These events can occur at any time during use and without warning symptoms. Elderly patients are at greater risk for serious gastrointestinal events.

Most osteoarthritis patients are in the higher risk age group, and also commonly require chronic therapy for their symptoms. For this reason, this patient population was selected by the applicant for the development of VIMOVO.

This compound was developed by POZEN Inc. and AstraZeneca under IND 76,301. On June 10, 2008, the sponsor met with the Agency to discuss the Phase 3 clinical development plan. At that meeting the potential for a comparability claim was discussed, along with the factors which would be considered in establishing a non-inferiority margin. At the pre-NDA meeting (March 23, 2009) the sponsor stated they intended to use Last Observation Carried Forward (LOCF) imputation for the primary analysis. The Division of Anesthesia, Analgesia, and Rheumatology Products (DAARP) advised them that the results would be reviewed using a more conservative method of imputing data since most dropouts are nonrandom. Based on this discussion, the applicant included additional sensitivity analyses in the Integrated Summary of Efficacy (ISE) section of the application.

2.2 Data Sources

All data was supplied by the applicant to the CDER electronic data room (edr) in SAS transport format. All necessary documentation, formats, and links were provided as well. The data and final study report for the electronic submission were archived under the network path location <\\CDSESUB1\EVSPROD\NDA022511\022511.ENX>

3. Statistical Evaluation

3.1 Evaluation of Efficacy

Study PN400-307 (conducted 4/08 to 12/08)

Design and Statistical Methods

Study 307 was a randomized, double-blind, parallel arm, multi-center study. The primary objective was to demonstrate that VIMOVO was non-inferior to celecoxib in the treatment of signs and symptoms of osteoarthritis. The applicant's goal was to add comparability claims of VIMOVO to celecoxib in the Clinical Studies section of the label, not to demonstrate efficacy of VIMOVO for this indication. The study included three treatment arms: VIMOVO 500 mg bid, celecoxib 200 mg qd, and placebo. The applicant planned the study based on non-inferiority comparisons between the VIMOVO and celecoxib arms.

Non-inferiority was assessed with three prespecified endpoints for signs and symptoms of osteoarthritis: the Western Ontario McMaster Universities (WOMAC) Osteoarthritis Index Pain Subscale, the WOMAC function subscale, and the Patient Global Assessment of OA (PGA)

visual analogue scale (VAS). Comparisons of the VIMOVO and celecoxib arms to the placebo arm for the endpoints were planned in the protocol as secondary objectives. Improvement on all three endpoints is required for the indication of treatment of signs and symptoms of OA.

The WOMAC instrument consists of 24 items, each measured using a 100 mm VAS scale. The items assess pain, stiffness, and function. The WOMAC Pain subscale is the mean response to 5 questions regarding how much pain a patient has during 5 common daily actions, where 0=no pain and 100=extreme pain. The WOMAC Function subscale is the mean response to 17 items regarding the degree of difficulty in doing daily physical activities such as moving around and looking after oneself, with 0=no difficulty and 100=extreme difficulty. On the WOMAC subscales, lower values indicate the desirable direction.

The patient global assessment (PGA) is a single 0-100 mm VAS question: “Consider all the ways your arthritis affects you, how well are you doing?” with 0=very poor and 100=excellent. On this scale, higher values indicate the desirable outcome.

Patients were age 50 and older, with at least a 6-month history of OA of the knee. They must have been on a stable dose of NSAIDs, COX-2 inhibitors, or other oral analgesic therapy for at least 6 weeks prior to screening, and expected to require continued treatment for at least 12 weeks. After the initial screening visit to determine eligibility, patients discontinued their oral analgesic therapy during a 7-14 day washout period. Patients who experienced an OA flare, defined as worsening of pain on Question 1 of the WOMAC pain scale by at least 15 mm and worsening on a Patient Global Assessment of at least 1 level on a 1-5 Likert scale, during the washout period were eligible for randomization.

Patients were randomized using a 2:2:1 ratio to the three treatment arms. A total of 619 patients were enrolled, with 248 randomized to receive VIMOVO, 247 randomized to receive celecoxib, and 124 randomized to receive placebo. An electronic diary was provided to subjects at the randomization visit to record the WOMAC pain subscale and other assessments daily during the treatment period. The items in the WOMAC function subscale and the PGA (VAS) question were not collected in the e-diary. Those were only recorded at the study visits: baseline (randomization), Week 1, Week 6, and Week 12 (or early discontinuation).

In the protocol, the Intent-to-Treat (ITT) population was defined as all randomized subjects who received at least one dose of study drug and provided at least one post-baseline efficacy evaluation. This is not the preferred definition because subjects who received study drug but discontinued prior to collecting post-baseline efficacy data would be excluded from the analyses. This could result in an artificial inflation of the treatment effect. Instead for my analyses, I used the BOCF imputation for subjects who did not have on-treatment observations so that all randomized were included in the analysis.

For each of the three efficacy endpoints, the applicant planned to test the null hypothesis that VIMOVO was inferior to celecoxib. The protocol planned to analyze the efficacy endpoints using an ANCOVA model with terms for treatment and baseline pain. The least square mean changes from baseline and 95% confidence intervals between VIMOVO and celecoxib would be

calculated. The applicant proposed a non-inferiority margin (delta) of 10 mm (out of 0-100 mm VAS scale) for each endpoint. Justification of the 10 point delta value was not discussed in the protocol.

In the Statistical Analysis Plan (SAP) dated January 14, 2009, the applicant planned Last Observation Carried Forward (LOCF) imputation for missing data for the efficacy endpoints at Week 6 or Week 12. At the pre-NDA meeting on March 23, 2009, the sponsor was informed that the Agency preferred a more conservative imputation strategy. The applicant provided the results using the Baseline Observation Carried Forward (BOCF), hybrid LOCF/BOCF, and ITT without imputation (observed data only) in the Integrated Summary of Efficacy (ISE). The hybrid LOCF/BOCF strategy imputed the baseline value for all discontinuations due to adverse events or lack of efficacy and imputed the last observation prior to withdrawal for patients discontinuing due to all other reasons.

The applicant specified in the protocol that no statistical adjustment for multiple comparisons was planned because the non-inferiority comparison had to be demonstrated for all three primary endpoints for the desired comparability claim. The applicant's sample size calculations were:

For the sample size and power calculations, under the null hypothesis it was assumed that Celebrex and PN 400 will have means of 35 and 25, respectively, for WOMAC Pain domain, and for the alternative hypothesis the assumptions were 35 and 33, respectively, with a common standard deviation of 25. With the above assumptions, this study requires 205 subjects each in the celecoxib and PN 400 treatment arms. With an estimated 10% early discontinuation, this study will randomize approximately 228 subjects per active treatment arm and approximately 114 subjects to the placebo arm for a total of approximately 570 subjects, utilizing a 2:2:1 randomization ratio. The sample size is sufficient to reject, using 2.5% onesided test with 90% power, the null hypothesis that PN 400 is inferior to celecoxib, with noninferiority margins of 10 on 100-mm WOMAC Pain and WOMAC Function Domains and 100-mm VAS PGA. Analysis of primary efficacy measures will be based on the intent-to-treat population.

The applicant did not discuss the power for superiority comparisons of VIMOVO to placebo. Using these same values, if placebo has a mean of 25, and VIMOVO has a mean of 33 for WOMAC Pain domain, with a common standard deviation of 25, then the planned sample size of 205 for VIMOVO and 103 for placebo would have power of 75% to detect a statistically significant difference. The applicant did not provide the values used for the sample size calculations for the other two endpoints, but stated that the planned sample size would be sufficient. This indicates the necessary sample sizes for the NI comparisons were no larger for those endpoints, so the power for the superiority comparison to placebo would have been at least 75% for the other two endpoints under the same scenario.

Patient Disposition

Patients were adult males and females with osteoarthritis. A total of 619 patients were randomized to the study, with 248 randomized to VIMOVO, 247 to celecoxib, and 124 to placebo. The dropout rate was consistent across all three groups (14-15%) with no notable differences between the treatment groups in terms of reasons for discontinuation. Table 1 shows the distribution of patient discontinuations by reason.

Table 1: Patient Disposition (Study 307)

	Vimovo 500mg/20mg bid	Celecoxib 200mg qd	Placebo
Randomized	248	247	124
Withdrew Prior to Treatment	1	4	0
Treated	247	243	124
Discontinued without efficacy data	1 (Moved)	1 (Refused to do log pad)	0
Intent-to-Treat (at least one post-baseline efficacy assessment reported)	N=246	N=242	N=124
Discontinued from ITT Population	38 (15%)	34 (14%)	19 (15%)
Adverse event	18 (7%)	16 (7%)	7 (6%)
Lack of efficacy	4 (2%)	3 (1%)	3 (2%)
Withdrew Consent	8 (3%)	9 (4%)	7 (6%)
Lost to Follow-up	0 (0%)	2 (1%)	0 (0%)
Other	8 (3%)	4 (2%)	2 (2%)
Completed	208 (85%)	208 (86%)	105 (85%)

Sources: Clinical Study Report Table 4 and Table 14.1.1.

Baseline Demographics

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

Table 2: Patient Demographics for Intent-to-Treat (ITT) Population (Study 307)

	Vimovo 500mg/20mg bid N=246	Celecoxib 200mg qd N=242	Placebo N=124
Age (years)			
Mean (SD)	62 (8)	62 (8)	62 (8)
Min, Max	50, 84	49, 90	50, 83
Age categories:			
<65 yrs	156 (63%)	164 (68%)	83 (67%)
≥65 yrs	90 (37%)	78 (32%)	41 (33%)
Gender			
Female	161 (65%)	148 (61%)	82 (66%)
Male	85 (35%)	94 (39%)	42 (34%)
Race			
White	194 (79%)	195 (81%)	99 (80%)
Black	43 (17%)	36 (15%)	21 (17%)
Asian	9 (4%)	10 (4%)	3 (2%)
Other	0	1 (<1%)	1 (1%)
Ethnicity			
Hispanic or Latino	12 (5%)	10 (4%)	7 (6%)
Body Mass Index (BMI) kg/m ²			
Mean (SD)	33 (7)	33 (8)	33 (7)
Min, Max	19, 57	19, 62	19, 58

Sources: Clinical Study Report Table 6

Efficacy Results

Although the applicant's primary objective was to demonstrate the non-inferiority of VIMOVO to celecoxib for the purpose of a comparability claim, I initially evaluated the efficacy of the drug compared to placebo. Table 3 presents the results for the analyses of the three primary efficacy endpoints comparing Vimovo to placebo. I have included results from four different imputation approaches. The applicant only collected information on the osteoarthritis endpoints at Baseline, Week 6, and Week 12. Patients who discontinued prior to Week 6 had no on-treatment efficacy measurements to be carried forward. The applicant's ITT-LOCF patient population, as defined in the protocol, did not include patients who did not have on-treatment measurements and was not the desired ITT population.

The results are consistent across the three endpoints and the alternative imputation methods. These results support the efficacy of VIMOVO.

Table 3 - Efficacy Results: Vimovo Compared to Placebo (Study 307)

		Observed data	Applicant's LOCF *	BOCF **	Hybrid *** LOCF/BOCF
	Vimovo N=	187	226	246	246
	Placebo N=	84	108	124	124
Endpoint					
WOMAC Pain	LS Mean: Vimovo	-44.7	-42.0	-34.6	-37.7
	LS Mean: Placebo	-40.5	-35.6	-26.3	-30.4
	Difference	-4.2	-6.4	-8.3	-7.3
	95% CI on Diff.	(-10.2, 1.8)	(-12.0, -0.7)	(-14.4, -2.2)	(-13.2, -1.4)
	p-value	0.17	0.027	0.008	0.015
WOMAC Function	LS Mean: Vimovo	-38.9	-36.4	-30.3	-33.0
	LS Mean: Placebo	-35.5	-30.6	-22.7	-25.9
	Difference	-3.4	-5.8	-7.6	-7.1
	95% CI on Diff.	(-9.4, 2.6)	(-11.3, -0.2)	(-13.3, -1.8)	(-12.7, -1.5)
	p-value	0.26	0.041	0.010	0.013
	Vimovo N=	192	242	246	246
	Placebo N=	87	119	124	124
Pt. Global Assessment	LS Mean: Vimovo	23.3	21.2	19.1	20.9
	LS Mean: Placebo	17.0	14.4	10.5	12.8
	Difference	6.4	6.8	8.7	8.1
	95% CI on Diff.	(-0.1, 12.9)	(1.1, 12.4)	(3.1, 14.2)	(2.6, 13.7)
	p-value	0.055	0.018	0.002	0.004
Source:		SAS datasets	Table 14.2.5.1	Table E2.28	Table E2.29

All comparisons from ANCOVA model with terms for treatment and baseline.

* Applicant's LOCF: If Week 12 missing, carry forward Week 6; If Week 6 also missing, drop from analysis.

** BOCF: Any Week 12 missing data imputed from baseline (Change from baseline = zero)

*** Hybrid LOCF/BOCF: If discontinued for Lack of Efficacy or Adverse events, then baseline carried forward; otherwise last observation (Week 6) carried forward.

During the teleconference (June 10, 2008) to discuss Phase 3 clinical study issues, the applicant asked for advice on the proposed endpoints, non-inferiority margins, statistical testing approach, analysis population, (b) (4)

The response from the Agency (including Dr. Permutt and Dr. Price) was:

(b) (4) however we disagree with your proposed analyses of the two studies planned to assess the efficacy of PN 400 in

patients with OA of the knee using a placebo and an active comparator (celecoxib). The determination of “non-inferiority” will need to be performed in the context of the analyses of all study endpoints. Factors that will contribute to the determination will be the treatment effect sizes versus placebo for both active arms in the studies, the variability in the treatment response, and the degree to which the treatment effects from PN 400 and Celebrex are comparable.

It is therefore not possible to establish an “acceptable” non-inferiority margin at this time. When the results of the study are analyzed, if there is any indication (i.e., any important endpoint) that PN 400 is inferior to Celebrex, there will not be a finding of non-inferiority.

The analyses of endpoints in the proposed trials should be conducted on the intent-to-treat population and the per protocol population.

The sponsor initially stated that the goal of the study was to establish efficacy in the osteoarthritis population. The Agency explained its understanding that the purpose of the study was to investigate the comparability of PN 400 to Celebrex and clarified that a placebo-controlled superiority trial would be required if the goal was to establish efficacy. To assess comparability, the Agency will evaluate the overall profile of PN 400 to Celebrex. The Agency further explained that the concerns regarding multiplicity which motivate the need for pre-specified endpoints and analyses in a trial designed to show efficacy do not apply in the same way to a study specifically designed to evaluate comparability.

Based on the criteria outlined in that discussion, I considered both the comparison of VIMOVO to Celecoxib (see Table 4) and the treatment effect size for Celecoxib vs. placebo (see Table 5).

 (b) (4)
 Table 5 presents the comparisons of Celecoxib to placebo.

Due to missing data in the applicant’s LOCF approach, and based on advice from the Agency to the applicant at the pre-NDA meeting, the BOCF and Hybrid LOCF/BOCF analyses will be the focus of my conclusions. In this study, the conclusions from either of those imputation approaches are consistent.

In Table 4, for all three endpoints, the mean improvement for the VIMOVO group was numerically better than for Celecoxib, so there is no indication that VIMOVO is inferior to Celecoxib. For the confidence intervals shown in Table 4, the direction of interest to determine inferiority is underlined. All are in the range of 3-4 mm.

In Table 5, the treatment effect sizes for celecoxib versus placebo are in the 6-7 mm range. Therefore, the appropriateness of the applicant’s proposed 10 mm difference is questionable since it is larger than the treatment effects observed in the studies.

Table 4 - Efficacy Results: Vimovo Compared to Celecoxib (Study 307)

		Observed data	Applicant's LOCF *	BOCF **	Hybrid *** LOCF/BOCF
	Vimovo N=	187	226	246	246
	Celecoxib N=	179	221	242	242
Endpoint					
WOMAC Pain	LS Mean: Vimovo	-44.7	-42.0	-34.6	-37.7
	LS Mean: celecox.	-43.2	-41.8	-32.6	-37.0
	Difference	-1.5	-0.2	-2.0	-0.6
	95% CI on Diff.	(-6.2, 3.3)	(-4.8, 4.3)	(-7.0, <u>3.0</u>)	(-5.5, <u>4.2</u>)
WOMAC Function	LS Mean: Vimovo	-38.9	-36.4	-30.3	-33.0
	LS Mean: celecox.	-37.9	-36.3	-28.7	-32.2
	Difference	-1.1	-0.1	-1.7	-0.7
	95% CI on Diff.	(-5.8, 3.6)	(-4.6, 4.4)	(-6.4, <u>3.1</u>)	(-5.4, <u>3.9</u>)
	Vimovo N=	192	242	246	246
	Celecoxib N=	180	230	242	242
Pt. Global Assessment	LS Mean: Vimovo	23.4	21.1	19.1	20.9
	LS Mean: celecox.	23.2	21.6	17.5	19.8
	Difference	0.2	-0.5	1.6	1.1
	95% CI on Diff.	(-5.1, 5.4)	(-5.1, 4.1)	(-2.9, 6.1)	(-3.4, 5.7)
Source:		Table 14.2.3	Table 14.2.1	Table E2.25	Table E2.26

* Applicant's LOCF: If Week 12 missing, carry forward Week 6; If Week 6 also missing, drop from analysis.

** BOCF: Any Week 12 missing data imputed to baseline (Change from baseline = zero)

*** Hybrid LOCF/BOCF: If discontinued for Lack of Efficacy or Adverse events, then baseline carried forward; otherwise last observation (Week 6) carried forward.

Table 5 - Efficacy Results: Celecoxib Compared to Placebo (Study 307)

		Observed data	Applicant's LOCF *	BOCF **	Hybrid *** LOCF/BOCF
	Celecoxib N=	179	221	242	242
	Placebo N=	84	108	124	124
Endpoint					
WOMAC Pain	LS Mean: celecox.	-43.2	-41.8	-32.6	-37.0
	LS Mean: Placebo	-40.5	-35.6	-26.3	-30.4
	Difference	-2.7	-6.1	-6.3	-6.6
	95% CI on Diff.	(-8.7, 3.2)	(11.8, -0.5)	(-12.4, -0.2)	(-12.5, -0.8)
	p-value	0.36	0.032	0.043	0.026
WOMAC Function	LS Mean: celecox.	-37.9	-36.3	-28.7	-32.2
	LS Mean: Placebo	-35.5	-30.6	-22.7	-25.9
	Difference	-2.3	-5.7	-5.9	-6.4
	95% CI on Diff.	(-8.3, 3.6)	(-11.2, -0.1)	(-11.7, -0.2)	(-12.0, -0.7)
	p-value	0.45	0.045	0.044	0.025
	Celecoxib N=	180	230	242	242
	Placebo N=	87	119	124	124
Pt. Global Assessment	LS Mean: celecox.	23.2	21.6	17.5	19.8
	LS Mean: Placebo	17.0	14.4	10.5	12.8
	Difference	6.2	7.2	7.0	7.0
	95% CI on Diff.	(-0.1, 12.9)	(1.6, 12.9)	(1.5, 12.6)	(1.5, 12.6)
	p-value	0.064	0.013	0.013	0.013
Source:		SAS datasets	Table 14.2.5.1	Table E2.28	Table E2.29

* Applicant's LOCF: If Week 12 missing, carry forward Week 6; If Week 6 also missing, drop from analysis.

** BOCF: Any Week 12 missing data imputed to baseline (Change from baseline = zero)

*** Hybrid LOCF/BOCF: If discontinued for Lack of Efficacy or Adverse events, then baseline carried forward; otherwise last observation (Week 6) carried forward.

Study PN400-309 (conducted 4/08 to 12/08)

Design

All aspects of the study design, patient population, and statistical analyses for Study 309 were identical to Study 307.

Patient Disposition

Patients were adults, ages 50 and over, with a history of OA of the knee. A total of 615 patients were randomized to the study, 244 to the Vimovo treatment group, 247 to the celecoxib treatment group, and 124 to the placebo group. Table 6 shows the distribution of patient discontinuations and reasons. The three groups were somewhat different in terms of their disposition. The celecoxib group had a higher rate of discontinuations (23%) than the Vimovo group (16%) or the placebo group (20%). The celecoxib group had a higher rate that discontinued due to an adverse event, while the placebo group had the higher rate who withdrew consent.

Table 6: Patient Disposition (Study 309)

	Vimovo 500mg/20mg bid	Celecoxib 200mg qd	Placebo
Randomized	244	247	124
Withdrew Prior to Treatment	1	2	2
Treated	243	245	122
Discontinued without efficacy data	2 (Family emerg.; Did not return)	1 (Refused to complete forms)	0
Intent-to-Treat (at least one post-baseline efficacy assessment reported)	N=241	N=244	N=122
Discontinued from ITT Population	38 (16%)	56 (23%)	24 (20%)
Adverse event	15 (6%)	22 (9%)	5 (4%)
Lack of efficacy	1 (<1%)	0 (0%)	2 (2%)
Withdrew Consent	15 (6%)	23 (9%)	14 (11%)
Lost to Follow-up	3 (1%)	3 (1%)	1 (1%)
Other	4 (2%)	8 (3%)	2 (2%)
Completed	208 (84%)	188 (77%)	98 (80%)

Sources: Clinical Study Report Table 4 and Table 14.1.1.

Baseline Demographics

The three treatment groups were well balanced with respect to relevant demographic and baseline characteristics as shown in Table 7.

Table 7: Patient Demographics for Intent-to-Treat (ITT) Population (Study 309)

	Vimovo 500mg/20mg bid N=241	Celecoxib 200mg qd N=244	Placebo N=122
Age (years)			
Mean (SD)	62 (9)	62 (8)	62 (9)
Min, Max	50, 88	50, 89	50, 87
Age categories:			
<65 yrs	157 (65%)	160 (66%)	84 (69%)
≥65 yrs	84 (35%)	84 (34%)	38 (31%)
Gender			
Female	157 (65%)	153 (63%)	77 (63%)
Male	84 (35%)	91 (37%)	45 (37%)
Race			
White	190 (79%)	195 (80%)	100 (82%)
Black	39 (16%)	43 (18%)	18 (15%)
Asian	11 (5%)	2 (1%)	3 (2%)
Other	1 (<1%)	4 (2%)	1 (1%)
Ethnicity			
Hispanic or Latino	19 (8%)	26 (11%)	9 (7%)
Body Mass Index (BMI) kg/m ²			
Mean (SD)	32 (7)	33 (8)	33 (8)
Min, Max	19, 60	21, 62	18, 61

Sources: Clinical Study Report Table 6

Efficacy Results

In the protocol for Study 309, the applicant planned an ANCOVA model with terms for treatment and baseline pain covariate. I provided the same analyses for this study as for Study 307 (see Tables 8-10).

Table 8 presents the results of the superiority comparison of Vimovo to placebo. When the BOCF imputation method was applied, Vimovo was not statistically significantly different from placebo. However, when the Hybrid LOCF/BOCF method was used, Vimovo was statistically significantly different from placebo for all three endpoints. This is due to the different imputation for patients who were coded as discontinuing due to Withdrawn Consent. In the BOCF imputation, all patients who discontinue are imputed as no change from baseline, regardless of reason for discontinuation. In the Hybrid LOCF/BOCF approach, patients whose discontinuation reason is Withdrawn Consent will have a Week 6 observation carried forward, if available. Therefore any improvement of the efficacy assessments achieved by Week 6 is not accounted for in the BOCF imputation. In order to understand the discrepancies between the BOCF and Hybrid LOCF/BOCF imputation results, I investigated the specific descriptions of the reason given for all patients whose discontinuation was coded as Withdrawn Consent to ensure that any subject who mentioned Lack of Efficacy or Adverse Event had the BOCF value imputed. In both the Vimovo and celecoxib groups, most of the patients who were listed as Withdrawn Consent gave reasons unrelated to treatment or OA pain (travel; other medical issues; did not like using log pad) while in the placebo group most (5/7) of the subjects listed as Withdrawn Consent gave reasons related to Lack of efficacy. The Hybrid LOCF/BOCF imputation correctly used that information in carrying forward BOCF values.

Table 8 - Efficacy Results: Vimovo Compared to Placebo (Study 309)

		Observed data	Applicant's LOCF *	BOCF **	Hybrid *** LOCF/BOCF
	Vimovo N=	175	226	240	240
	Placebo N=	89	108	122	122
Endpoint					
WOMAC Pain	LS Mean: Vimovo	-45.5	-42.0	-32.9	-37.6
	LS Mean: Placebo	-40.8	-35.6	-29.9	-31.4
	Difference	-4.7	-6.4	-3.0	-6.2
	95% CI on Diff.	(-10.6, 1.3)	(-12.0, -0.7)	(-9.4, 3.3)	(-12.3, -0.04)
	p-value	0.12	0.027	0.35	0.048
WOMAC Function	LS Mean: Vimovo	-40.5	-36.4	-29.0	-33.2
	LS Mean: Placebo	-34.8	-30.6	-25.4	-26.6
	Difference	-5.7	-5.8	-3.6	-6.5
	95% CI on Diff.	(-11.9, 0.5)	(-11.3, -0.2)	(-9.7, 2.5)	(-12.5, -0.6)
	p-value	0.07	0.041	0.24	0.032
	Vimovo N=	179	242	240	240
	Placebo N=	91	119	122	122
Pt. Global Assessment	LS Mean: Vimovo	30.8	21.2	22.3	27.0
	LS Mean: Placebo	24.1	14.4	19.2	20.4
	Difference	6.7	6.8	3.1	6.5
	95% CI on Diff.	(0.02, 13.3)	(1.1, 12.4)	(-2.8, 9.0)	(0.6, 12.5)
	p-value	0.049	0.018	0.30	0.031
Source:		SAS datasets	Table 14.2.5.1	Table E2.34	Table E2.35

All comparisons from ANCOVA model with terms for treatment and baseline.

* Applicant's LOCF: If Week 12 missing, carry forward Week 6; If Week 6 also missing, drop from analysis.

** BOCF: Any Week 12 missing data imputed to baseline (Change from baseline = zero)

*** Hybrid LOCF/BOCF: If discontinued for Lack of Efficacy or Adverse Events, then baseline carried forward; otherwise last observation (Week 6) carried forward.

Table 9 provides the results of the non-inferiority comparisons of Vimovo to celecoxib. The interpretation of these results depends on the treatment effect sizes of both Vimovo and celecoxib versus placebo. As shown in Table 10 on the following page, the celecoxib treatment group was not statistically significantly different from placebo in this study. Therefore non-inferiority comparisons to celecoxib do not provide valuable information.

Table 9 - Efficacy Results: Vimovo Compared to Celecoxib (Study 309)

		Observed data	Applicant's LOCF *	BOCF **	Hybrid *** LOCF/BOCF
	Vimovo N=	175	213	240	240
	Celecoxib N=	161	220	243	243
Endpoint					
WOMAC Pain	LS Mean: Vimovo	-45.5	-44.2	-32.9	-37.6
	LS Mean: celecox.	-46.0	-42.9	-30.8	-35.7
	Difference	0.5	-1.3	-2.1	-1.9
	95% CI on Diff.	(-4.5, 5.5)	(-5.9, 3.3)	(-7.3, 3.1)	(-6.9, 3.1)
WOMAC Function	LS Mean: Vimovo	-40.5	-38.9	-29.0	-33.2
	LS Mean: celecox.	-39.7	-36.8	-26.8	-30.7
	Difference	-0.8	-2.1	-2.2	-2.4
	95% CI on Diff.	(-6.0, 4.4)	(-6.8, 2.6)	(-7.2, 2.7)	(-7.3, 2.5)
	Vimovo N=	179	235	240	240
	Celecoxib N=	165	234	243	243
Pt. Global Assessment	LS Mean: Vimovo	30.8	29.0	22.3	27.0
	LS Mean: celecox.	29.5	25.6	20.1	22.8
	Difference	1.3	3.5	2.2	4.2
	95% CI on Diff.	(-4.2, 6.9)	(-1.4, 8.3)	(-2.7, 7.0)	(-0.7, 9.0)
Source:		Table 14.2.3	Table 14.2.1	Table E2.31	Table E2.32

* Applicant's LOCF: If Week 12 missing, carry forward Week 6; If Week 6 also missing, drop from analysis.

** BOCF: Any Week 12 missing data imputed to baseline (Change from baseline = zero)

*** Hybrid LOCF/BOCF: If discontinued for Lack of Efficacy or Adverse events, then baseline carried forward; otherwise last observation (Week 6) carried forward.

Table 10 shows the comparisons of celecoxib to placebo. For both the BOCF and Hybrid LOCF/BOCF imputation methods, the celecoxib group was not statistically significantly different from the placebo group. This is due to the imbalance in the percentage of drop-outs and the reasons recorded for discontinuation.

Table 10 - Efficacy Results: Celecoxib Compared to Placebo (Study 309)

		Observed data	Applicant's LOCF *	BOCF **	Hybrid *** LOCF/BOCF
	Celecoxib N=	161	221	243	243
	Placebo N=	89	108	122	122
Endpoint					
WOMAC Pain	LS Mean: Vimovo	-46.0	-41.8	-30.8	-35.7
	LS Mean: Placebo	-40.8	-35.6	-29.9	-31.4
	Difference	-5.2	-6.1	-0.9	-4.3
	95% CI on Diff.	(-11.2, 0.9)	(11.8, -0.5)	(-7.2, 5.4)	(-10.4, 1.9)
	p-value	0.09	0.032	0.78	0.17
WOMAC Function	LS Mean: Vimovo	-39.7	-36.3	-26.8	-30.7
	LS Mean: Placebo	-34.8	-30.6	-25.4	-26.6
	Difference	-4.9	-5.7	-1.4	-4.1
	95% CI on Diff.	(-11.2, 1.3)	(-11.2, -0.1)	(-7.4, 4.7)	(-10.1, 1.9)
	p-value	0.12	0.045	0.65	0.18
	Celecoxib N=	165	230	243	243
	Placebo N=	91	119	122	122
Pt. Global Assessment	LS Mean: Vimovo	29.5	21.6	20.1	22.8
	LS Mean: Placebo	24.1	14.4	19.2	20.4
	Difference	5.3	7.2	1.0	2.4
	95% CI on Diff.	(-1.4, 12.1)	(1.6, 12.9)	(-4.9, 6.9)	(-3.5, 8.3)
	p-value	0.12	0.013	0.75	0.43
Source:		SAS datasets	Table 14.2.5.1	Table E2.34	Table E2.35

* Applicant's LOCF: If Week 12 missing, carry forward Week 6; If Week 6 also missing, drop from analysis.

** BOCF: Any Week 12 missing data imputed to baseline (Change from baseline = zero)

*** Hybrid LOCF/BOCF: If discontinued for Lack of Efficacy or Adverse events, then baseline carried forward; otherwise last observation (Week 6) carried forward.

3.2 Evaluation of Safety

The evaluation of the safety data was conducted by Dr. Jin Chen. The reader is referred to Dr. Chen’s review for information regarding the adverse event profile. No additional safety analyses were requested from me by Dr. Chen.

4. Findings in Special/Subgroup Populations

4.1 Gender, Race and Age

The applicant provided descriptive analyses by age groups, gender, and race, but reported the ITT-LOCF patient population, which did not include all randomized patients. Tables 11a-c show these subgroup analyses for the mean change from baseline to Week 12 for each of the three endpoints using the LOCF/BOCF imputation preferred for the comparison of Vimovo to placebo.

The only trend I noticed across any of these subgroups was for the Race groups in Study 307. For all the endpoints and all the treatment arms, the mean improvement in the Caucasian subgroup was more favorable than the mean improvement in the All Other Races subgroup. A reanalysis using an ANCOVA model with terms for treatment, race, the treatment by race interaction, and baseline showed that there was no significant interaction between treatment and race, and that the numerical trend did not change the conclusions from the treatment arm comparisons.

Table 11a: WOMAC Pain: Subgroup Analyses (ITT with LOCF/BOCF Imputation)

Primary Endpoint: Mean Change from Baseline to Week 12: WOMAC Pain	Study 307			Study 309		
	Mean (SE)					
	Vimovo n=246	Celecoxib n=242	Placebo n=124	Vimovo n=241	Celecoxib n=244	Placebo n=122
Age groups						
50-59 years	n=92 -40 (3)	n=111 -37 (3)	n=54 -26 (4)	n=111 -38 (3)	n=110 -37 (3)	n=57 -33 (4)
≥60 years	n=153 -38 (2)	n=127 -36 (3)	n=69 -32 (3)	n=129 -37 (3)	n=133 -35 (3)	n=65 -28 (3)
Gender						
Female	n=160 -38 (2)	n=145 -36 (2)	n=81 -28 (3)	n=156 -38 (2)	n=152 -34 (2)	n=77 -31 (3)
Male	n=85 -41 (3)	n=93 -37 (3)	n=42 -31 (4)	n=84 -37 (3)	n=91 -39 (3)	n=45 -30 (4)
Race						
Caucasian	n=194 -40 (2)	n=192 -37 (2)	n=99 -30 (3)	n=189 -39 (2)	n=194 -35 (2)	n=100 -30 (3)
All others	n=51 -35 (4)	n=46 -33 (5)	n=24 -25 (4)	n=51 -33 (4)	n=49 -39 (4)	n=22 -34 (7)

Sources: SAS datasets

Table 11b: WOMAC Function: Subgroup Analyses (ITT with LOCF/BOCF Imputation)

Primary Endpoint: Mean Change from Baseline to Week 12: WOMAC Function	Study 307			Study 309		
	Mean (SE)					
	Vimovo n=246	Celecoxib n=242	Placebo n=124	Vimovo n=241	Celecoxib n=244	Placebo n=122
Age groups						
50-59 years	n=92 -36 (3)	n=111 -33 (3)	n=54 -21 (3)	n=111 -34 (3)	n=110 -32 (3)	n=57 -29 (4)
≥60 years	n=153 -32 (2)	n=127 -31 (2)	n=69 -28 (3)	n=129 -32 (2)	n=133 -31 (2)	n=65 -23 (3)
Gender						
Female	n=160 -33 (2)	n=145 -33 (2)	n=81 -24 (3)	n=156 -34 (2)	n=152 -31 (2)	n=77 -26 (3)
Male	n=85 -36 (3)	n=93 -30 (3)	n=42 -26 (4)	n=84 -31 (3)	n=91 -32 (3)	n=45 -25 (4)
Race						
Caucasian	n=194 -35 (2)	n=192 -33 (2)	n=99 -26 (3)	n=189 -34 (2)	n=194 -30 (2)	n=100 -25 (3)
All others	n=51 -30 (4)	n=46 -29 (4)	n=24 -21 (5)	n=51 -29 (4)	n=49 -35 (4)	n=22 -29 (7)

Sources: SAS datasets

Table 11c: Patient Global Assessment: Subgroup Analyses (ITT with LOCF/BOCF Imputation)

Primary Endpoint: Mean Change from Baseline to Week 12: Pt Global Assessment	Study 307			Study 309		
	Mean (SE)					
	Vimovo n=246	Celecoxib n=242	Placebo n=124	Vimovo n=241	Celecoxib n=244	Placebo n=122
Age groups						
50-59 years	n=92 20 (4)	n=111 21 (3)	n=54 10 (4)	n=111 27 (3)	n=110 26 (3)	n=57 23 (4)
≥60 years	n=153 22 (3)	n=127 19 (3)	n=69 13 (3)	n=129 24 (3)	n=133 22 (3)	n=65 19 (4)
Gender						
Female	n=160 21 (3)	n=145 20 (2)	n=81 12 (3)	n=156 28 (3)	n=152 24 (3)	n=77 21 (3)
Male	n=85 21 (3)	n=93 20 (3)	n=42 11 (5)	n=84 22 (4)	n=91 23 (3)	n=45 21 (5)
Race						
Caucasian	n=194 22 (2)	n=192 21 (2)	n=99 13 (3)	n=189 27 (3)	n=194 22 (2)	n=100 22 (3)
All others	n=51 17 (5)	n=46 19 (4)	n=24 6 (5)	n=51 22 (5)	n=49 29 (4)	n=22 16 (5)

Sources: SAS datasets

4.2 Other Special/Subgroup Populations

No additional subgroup analyses were requested by Dr. Chen.

5. Summary and Conclusions

5.1 Statistical Issues and Collective Evidence

The applicant provided two prospectively-planned, randomized, double-blind, placebo and active-control, parallel-arm studies to support the efficacy of Vimovo for an indication for treatment of signs and symptoms of osteoarthritis. In both studies, VIMOVO was statistically significantly different from placebo for the three efficacy endpoints. This supports the efficacy of VIMOVO for this indication.

[REDACTED] (b) (4)

During a teleconference to discuss Phase 3 clinical study issues on June 10, 2008, Drs. Thomas Permutt and Dionne Price (DB2) advised the applicant that we did not agree with the delta of 10 units for the non-inferiority criteria and that the totality of the evidence would be used [REDACTED] (b) (4)

[REDACTED] (b) (4) In Study 307, celecoxib was statistically significantly different from placebo, but the treatment effect sizes were 6 to 7 mm for the three endpoints. This is less than the proposed 10 mm non-inferiority criteria planned by the sponsor and suggests that the margin is not reasonable. Moreover since placebo-controlled trials conducted in the original celecoxib application used different scales and questions to assess efficacy, determination of the degree to which the treatment effects should be comparable is difficult to assess. In Study 309, celecoxib was not statistically significantly different from placebo. For that study, comparisons of VIMOVO to celecoxib do not provide meaningful information. [REDACTED] (b) (4)

In both studies, Vimovo was statistically significantly different from placebo for the three efficacy endpoints used to assess the treatment of the signs and symptoms of osteoarthritis.

[REDACTED] (b) (4) Of particular concern was that in study 309 celecoxib was not statistically significantly better than placebo, so non-inferiority comparisons to celecoxib in that study were not reasonable.

5.2 Label Issues

The applicant's proposed label reports the results from the analysis in the Clinical Studies section. The description of the study designs is brief. (b) (4)

I would prefer the following changes in the reporting of the study results:

1. The applicant's first summary statement is "Patients receiving VIMOVO had significantly better results compared to patients receiving placebo, (b) (4) as measured by change from baseline WOMAC scores on domains of pain and physical function as well as on Patient Global Assessment Scores." The conclusions with respect to placebo are acceptable. (b) (4)

2. (b) (4)

5.3 Conclusions and Recommendations

The applicant provided results from two studies planned to demonstrate the efficacy of Vimovo for the treatment of signs and symptoms of osteoarthritis. (b) (4)

Both studies also included a placebo arm, which allowed for comparison of Vimovo to assess efficacy without relying on the non-inferiority comparisons.

Based on the comparisons of Vimovo to placebo for all three endpoints in both studies, there is sufficient evidence to support the efficacy of Vimovo for the treatment of signs and symptoms of osteoarthritis. There was not adequate justification for the planned non-inferiority comparisons, and the results of the studies provide conflicting conclusions on those hypotheses. (b) (4)

Signatures/Distribution List Page

Katherine B. Meaker, M.S.
Mathematical Statistician

Dionne Price, Ph.D.
Team Leader

Application Type/Number	Submission Type/Number	Submitter Name	Product Name
NDA-22511	ORIG-1	POZEN INC	PN 400 NAPROXEN/ESOMEPRAZOLE MAGNESIUM

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
04/28/2010

DIONNE L PRICE
04/28/2010
Concur



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 Number: 22-511 / 000

Drug Name: VIMOVO™ (naproxen/esomeprazole magnesium) Tablets

Indication: Treatment of the signs and symptoms of osteoarthritis, rheumatoid arthritis and ankylosing spondylitis in patients at risk for developing NSAID-associated gastric ulcer

Applicant: POZEN Inc.

Dates: Letter Date: June 30, 2009
Stamp Date: June 30, 2009
PDUFA Date: April 30, 2010

Review Priority: Standard

Biometrics Division: Division of Biometrics 3

Statistical Reviewer: Freda W. Cooner, Ph.D.

Concurring Reviewers: Mike Welch, Ph.D.

Medical Division: Division of Gastroenterology Products (DGP)

Clinical Team: Erica Wynn, M.D., Medical Reviewer
Ruyi He, M.D., Medical Team Leader

Project Manager: Anna Maria Simon

Keywords: Clinical studies, NDA review

Table of Contents

1. EXECUTIVE SUMMARY	3
1.1 CONCLUSIONS AND RECOMMENDATIONS	3
1.2 BRIEF OVERVIEW OF CLINICAL STUDIES	3
1.3 STATISTICAL ISSUES AND FINDINGS	4
2. INTRODUCTION	4
2.1 OVERVIEW.....	4
2.2 DATA SOURCES	8
3. STATISTICAL EVALUATION	9
3.1 STUDY PN400-301 AND STUDY PN400-302	9
3.1.1 EVALUATION OF EFFICACY.....	9
3.1.1.1 STUDY DESIGN AND ENDPOINTS	9
3.1.1.2 PATIENT DISPOSITION, DEMOGRAPHIC AND BASELINE CHARACTERISTICS	11
3.1.1.3 STATISTICAL METHODOLOGIES	13
3.1.1.4 RESULTS AND CONCLUSIONS.....	15
3.1.2 INTEGRATED EFFICACY ANALYSIS	17
3.2 STUDY PN400-303	18
3.3 OVERVIEW OF SAFETY	18
4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS	19
4.1 GENDER, RACE AND AGE	19
4.2 OTHER SPECIAL/SUBGROUP POPULATIONS	21
5. SUMMARY AND CONCLUSIONS	22

1. EXECUTIVE SUMMARY

1.1 Conclusions and Recommendations

Vimovo™ PN 400 Tablet is a fixed dose combination tablet containing 375 mg or 500 mg naproxen in the core and 20 mg esomeprazole in the film coat. Esomeprazole is immediately released from the film coat, whereas the release of naproxen from the enteric coated core is delayed. This NDA is supportive of the oral administration of PN 400 Tablets on a twice daily (bid) regimen for treatment of the signs and symptoms of osteoarthritis, rheumatoid arthritis, and ankylosing spondylitis in patients at risk for developing NSAID-associated gastric ulcers. To demonstrate the efficacy of PN 400 in reducing the risk of gastric ulcer, two identical phase 3 studies were conducted to compare PN 400 with enteric-coated naproxen alone and showed that PN 400 significantly reduced the incidence of gastric ulcers in an at-risk population at the end of six months of treatment. A third trial, intended to study PN 400 in a high-risk population, was terminated early due to insufficient enrollment.

1.2 Brief Overview of Clinical Studies

VIMOVO is a fixed dose combination tablet containing 375 mg or 500 mg naproxen in the core and 20 mg esomeprazole (as the magnesium trihydrate salt) in the film coat. VIMOVO™ Tablets are for oral administration on a twice daily (bid) regimen. The development program and this NDA submission for VIMOVO™ (PN 400) Tablets were of the collaborative efforts of POZEN and AstraZeneca (AZ). Under the licensing agreement between the two companies, this NDA and the supporting IND #76,301 will be transferred to AZ upon approval and AZ will manufacture and market VIMOVO™ Tablets.

The development program to support this 505(b)(2) NDA (RLD's [Reference Listed Drug] Naprosyn® / Nexium®) for PN 400 Tablets included 15 studies in total, 13 of them used PN 400 Tablets, one used PN 200 Tablets (500 mg naproxen and 20 mg omeprazole), and one healthy volunteer pharmacokinetic (PK) study comparing over-encapsulated celecoxib with commercially available Celebrex®. Most of these studies are phase 2 studies and five of them are phase 3 studies. In particular, two phase 3 efficacy and safety Studies PN400-301 and PN400-302 were conducted to demonstrate reductions in gastric ulcer (GU) occurrence with PN 400 compared to EC (enteric-coated) naproxen. Phase 3 efficacy and safety Study PN400-303 in support of high risk population was terminated early due to low and inadequate enrollment. One long term safety Study PN400-304 was also conducted for up to one-year of treatment. Two phase 3 comparative Studies PN400-307 and PN400-309 were additional principal studies to demonstrate non-inferiority of PN 400 to celecoxib in subjects with OA (osteoarthritis) of the knee. This review will focus on Studies PN400-301 and PN400-302. Due to limited additional information from Study PN400-303, it will only be briefly discussed. Studies PN400-307 and PN400-309 will be evaluated by the Division of Anesthesia, Analgesia, and Rheumatology Products (DAARP).

Two identical, randomized, multi-center, double-blind, parallel group, outpatient, active-controlled Studies PN400-301 and PN400-302 were conducted to demonstrate reductions in GU occurrence in subjects who took PN 400 Tablets bid compared with those who took EC naproxen 500 mg bid on a daily basis for six months. Both studies also evaluated differences in NSAID (non-steroidal anti-inflammatory drug)-associated upper gastrointestinal (UGI) adverse events (AEs) and duodenal ulcer (DU) formation in subjects who used PN 400 Tablets bid and those who used naproxen 500 mg bid. The diagnoses for eligibility included subjects with osteoarthritis, ankylosing spondylitis, rheumatoid arthritis (RA) or any other medical condition that would require the use of daily NSAIDs for the study period of six months. The inclusion criteria required that subjects 18 to 49 years of age must have had a documented, uncomplicated GU and/or DU within five years of the study enrollment. Subjects more than 50 years of age, regardless of GU or DU history, were eligible to be randomized. The 1:1 randomization to either PN400 or EC naproxen included stratification on low dose aspirin (LDA) use.

1.3 Statistical Issues and Findings

The primary analyses for the primary efficacy endpoint of gastric ulcer incidence were originally proposed to use a life-table method. Per the FDA's request, the more conventional cumulative rates were used for the treatment comparisons. In one study, missing data were more predominant in the control group, and consequently, the sponsor's imputation method assigning treatment to success was more conservative than the usual treatment-failure approach; although efficacy conclusions were not sensitive to early withdrawals. Protocol modifications were made near study conclusion; of concern were the introduction of new safety-related, key secondary endpoints and a proposed sequential testing order. These secondary endpoints should be considered exploratory and not suitable for efficacy claims.

2. INTRODUCTION

2.1 Overview

Non-steroidal anti-inflammatory drugs (NSAIDs) work by inhibiting the enzyme cyclooxygenase (COX) with two forms (COX-1 and COX-2) discovered in the 1990s. Traditional NSAIDs are considered nonselective because they inhibit both forms of COX. The inhibition of COX-2 by traditional NSAIDs accounts for the anti-inflammatory effect of the drugs, while the inhibition of COX-1 can lead to NSAID toxicity including side effects such as ulcers.

Although NSAIDs remain a key therapy for the management of signs and symptoms of chronic inflammatory conditions such as osteoarthritis (OA), it is well recognized that there is a substantial risk of upper gastrointestinal (UGI) ulcerations and ulcer complications, such as bleeding and perforations, with chronic NSAID therapy. The cumulative incidence of gastroduodenal ulcers with conventional NSAID use has been reported to be as high as 25% to 30% at three months and 45% at six months, while that of placebo is 3% to 7%. Important risk factors for UGI ulcers in NSAID users are advancing age, a history of UGI ulcer or bleeding, and concomitant aspirin use. The sponsor reported that use of NSAID is also associated with

additional UGI events including dyspepsia (in up to 40% of patients), erosive esophagitis (EE; in up to 21% of patients) and an increase in symptoms of gastroesophageal reflux disease (GERD).

In clinical practice gastric anti-secretory agents, especially proton pump inhibitors (PPIs), are frequently used to treat EE and GERD or to manage acid-related dyspepsia. The sponsor stated that these agents have been shown to mitigate the risk from daily NSAID use by significantly reducing the development of gastric ulcers (GUs). While physicians may co-prescribe PPIs with NSAIDs, both inconsistent dosing instructions and lack of subject compliance in “at risk” population may hinder the potential benefit of this approach. Compliance in at risk NSAID users instructed to also take a gastro-protective agent has been reported to be lower than 30% and the likelihood of adherence was further decreased if NSAIDs were prescribed for 90 days or more. The PN Tablets concept was developed as an NSAID product for arthritis that would provide absolute and automatic compliance with the two-drug regimen with every dose since it combines both agents into a single table.

PN 400 is a fixed dose combination tablet containing 375 mg or 500 mg naproxen in the core and 20 mg esomeprazole (present as 22.3 mg esomeprazole magnesium trihydrate) in the film coat. Esomeprazole is immediately released from the film coat, whereas the release of naproxen from the EC core is delayed as it is dependent on elevated pH. Both of the two active components, naproxen (NSAID) and esomeprazole (PPI), are compendia compounds that are currently marketed as single entities. In particular, EC naproxen delayed-release tablets under the trade name EC-NAPROSYN[®] are available in 375 mg and 500 mg bid for the relief of the signs and symptoms of RA (rheumatoid arthritis), OA and AS (ankylosing spondylitis). The current labeling states that for patients who tolerate well, the dose may be increased to 1500 mg/day for periods up to 6 months. Esomeprazole magnesium delayed-release capsules under the trade name NEXIUM[®] are available in 20 mg and 40 mg once daily for up to six months for risk reduction of NSAID-associated GU.

The development program to support this application for PN 400 Tablets included as key elements the need to demonstrate bioequivalence of naproxen in PN 400 Tablets with EC-NAPROSYN[®], as well as the need to demonstrate efficacy of immediate release esomeprazole by improved GI safety relative to EC naproxen alone. The two pivotal studies PN400-301 and PN400-302 were conducted to fulfill the second need above. The sponsor claimed that the design of the pivotal safety and efficacy trials was agreed during discussions with the Agency and through a SPA (special protocol assessment). (b) (4)

On October 3, 2008, after the Agency’s review of the SAPs (statistical analysis plans) for the two pivotal studies, an advice/information request letter was issued recommending the primary analysis be based on analyzing the proportions of cases with GU instead of time-to-GU and recommending multiplicity adjustments for the secondary endpoints. The sponsor issued protocol amendments regarding these issues prior to database lock. The pre-NDA meeting was held on March 23, 2009, the packages did not include the protocols or SAPs, and the aforementioned issues were not discussed during the meeting.

Five phase 3 controlled studies were submitted with the application and are summarized in the table below. Two identical and controlled studies (PN400-301 and PN400-302) were conducted

to demonstrate reductions in GU occurrence in subjects who took PN 400 Tablets bid compared with those who took EC naproxen 500 mg bid on a daily basis for six months. These two studies were designed to study a population of subjects with medical conditions that required daily use of NSAIDs and who were at risk of GI toxicity from the chronic use of NSAIDs. Specifically, the diagnoses for eligibility included subjects with OA, AS, RA or any other medical condition that would require the use of daily NSAIDs for the study period of six months. The inclusion criteria required that subjects 18 to 49 years old must have had a documented, uncomplicated GU and/or DU within five years of the study enrollment. Subjects 50 years or older regardless of GU or DU history were eligible to be randomized. Approximately one quarter (24.0%) of the subjects in Study PN400-301 and 23.1% of the subjects in Study PN400-302 used low-dose aspirin (LDA) and randomization was stratified on LDA use. Study PN400-303, compared PN 400 Tablets bid to ARTHROTEC® (diclofenac 75 mg/ misoprostol 200 mcg) capsules bid, in very high risk subjects, i.e., with a documented history of an ulcer-related serious UGI event such as bleeding, perforation or obstruction. Based on significant difficulty in recruiting subjects, POZEN stopped this study early, and as a result of the inadequate enrollment, reliable data on GU incidence in this population could not be determined. PN 400 Tablets were also evaluated in two controlled studies (PN400-307 and PN400-309) to show non-inferiority to celecoxib in the management of signs and symptoms of OA of the knee. These two studies are to be reviewed separately by DAARP (Division of Anesthesia, Analgesia, and Rheumatology Products).

Table 2.1. Summaries of clinical efficacy and safety studies								
Type of Study	Study Identifier	Objective(s) of the Study	Study Design and Type of Control	Test Product(s) and Dosage Regimen; Route of Administration	Number of Subjects	Healthy Subjects or Diagnosis of Patients	Duration of Treatment	Study Status; Type of Report
Pivotal Safety and Efficacy	PN400-301	Reduction of risk of gastric ulcers in at risk patients	Double-blind, parallel group, randomized, active controlled, multicenter study	Tablets, 500 mg naproxen/20 mg esomeprazole, 500 mg naproxen alone, BID 30-60 mins before food, oral	400 Planned 438 Randomized 434 Treated 333 Completed	Patients with history of OA, RA, ankylosing spondylitis or other medical conditions that require daily NSAID therapy	6 Months	Complete Full
Pivotal Safety and Efficacy	PN400-302	Reduction of risk of gastric ulcers in at risk patients	Double-blind, parallel group, randomized, active controlled, multicenter study	Tablets, 500 mg naproxen/20 mg esomeprazole, 500 mg naproxen alone, BID 30-60 mins before food, oral	400 Planned 423 Randomized 420 Treated 304 Completed	Patients with history of OA, RA, ankylosing spondylitis or other medical conditions that require daily NSAID therapy	6 Months	Complete Full
Safety and Efficacy in High Risk	PN400-303	Incidence of gastric ulcers in high risk population at six months	Double-blind, parallel group, randomized, active controlled, multicenter study	Tablet, 500 mg naproxen/20 mg esomeprazole; Capsule, over-encapsulated ARTHROTEC® 75 (75 mg diclofenac sodium/200 mcg misoprostol), BID Oral	200 Planned 20 Randomized 3 Completed Study terminated	Patients with history of OA, RA, ankylosing spondylitis or other medical conditions that require daily NSAID therapy, with history of documented serious upper gastrointestinal event, such as perforation, obstruction or bleeding	6 Months Planned	Study Terminated Study synopsis complete
Non-inferiority	PN400-307	Non-inferiority of PN 400 and celecoxib in treatment of signs and symptoms of OA	Double-blind, parallel group, randomized, active controlled, multicenter study	Tablet, naproxen 500 mg/20 mg esomeprazole, BID and Capsule, placebo QD; Capsule, 200 mg over-encapsulated CELEBREX® (celecoxib), QD and Tablet, placebo, BID; Tablet, placebo, BID and Capsule, placebo, QD 30-60 mins before meals, oral	570 Planned 619 Randomized 614 Treated 521 Completed	Patients with a history of OA of the knee that requires daily NSAID therapy	3 Months	Complete Full

Table 2.1. (Cont'd) Summaries of clinical efficacy and safety studies								
Type of Study	Study Identifier	Objective(s) of the Study	Study Design and Type of Control	Test Product(s) and Dosage Regimen; Route of Administration	Number of Subjects	Healthy Subjects or Diagnosis of Patients	Duration of Treatment	Study Status; Type of Report
Non-inferiority	PN400-309	Non-inferiority of PN 400 and celecoxib in treatment of signs and symptoms of OA	Double-blind, parallel group, randomized, active controlled, multicenter study	Tablet, naproxen 500 mg/20 mg esomeprazole, BID and Capsule, placebo QD; Capsule, 200 mg over-encapsulated CELEBREX® (celecoxib), QD and Tablet, placebo, BID; Tablet, placebo, BID and Capsule, placebo, QD 30-60 mins before meals, oral	570 Planned 615 Randomized 610 Treated 489 Completed	Patients with a history of OA of the knee that requires daily NSAID therapy	3 Months	Complete Full

Source: Module 5.2 - Tabular Listing of All Clinical Studies

2.2 Data Sources

Materials reviewed include the two phase 3 study reports (PN400-301 and PN400-302), synopsis for the terminated phase 3 study report (PN400-303), and integrated study reports. The original applications were submitted in electronic Common Technical Document (eCTD) format, along with SAS datasets and programs provided, to EDR at <\\Cdsesub1\evsprod\NDA022511\0000>.

3. STATISTICAL EVALUATION

3.1 Study PN400-301 and Study PN400-302

3.1.1 Evaluation of Efficacy

3.1.1.1 Study Design and Endpoints

The primary objective of Studies 301 and 302 were to demonstrate that PN 400 is effective in reducing the risk of GUs in subjects at risk for developing NSAID-associated GUs. The secondary objectives were to determine if PN 400 is effective in reducing the risk of DUs in subjects at risk for developing NSAID-associated ulcers; to compare UGI symptoms in subjects treated with PN 400 versus delayed-release naproxen (hereafter called naproxen) as measured by scores on the Severity of Dyspepsia Assessment (SODA) instrument and the Overall Treatment Evaluation – Dyspepsia (OTE-DP); to compare heartburn symptoms in subjects treated with PN 400 versus naproxen; to evaluate the safety and tolerability of PN 400 and naproxen. The other objective of these two pivotal studies was to assess the effect of concomitant use of LDA on the incidence of gastroduodenal ulcers within each treatment group.

The two pivotal phase 3 studies were identical, randomized, double-blind, parallel-group, controlled, multicenter clinical trials of a six months duration. Approximately 60 U.S. sites were targeted to enroll a total of 400 subjects (200 per treatment group) for each study. Each study had a screening period and a double-blind treatment period. During the study periods, subjects were required not to use disallowed medications. Subjects eligible for inclusion in the studies were male or female with a history of osteoarthritis (OA), rheumatoid arthritis (RA), ankylosing spondylitis (AS), or other medical conditions expected to require daily NSAID therapy for at least six months, who were either

- 18 to 49 years of age and had a history of a documented, uncomplicated gastric or duodenal ulcer (a mucosal break of at least 3 mm in diameter with depth, without any concurrent bleeding, clot, or perforation) within the past five years; or
- 50 years of age and older (did not require a history of a documented, uncomplicated gastric or duodenal ulcer within the past five years).

After all entrance criteria were fulfilled, subjects were randomized to one of two treatment groups, PN 400 or naproxen 500 mg, both taken orally, twice a day (bid) and stratified by LDA use. Subjects also underwent assessments for dyspepsia and related GI symptoms using the SODA instrument and heartburn on the day of randomization. Subjects returned at one and three months for safety assessments, endoscopy, and additional study drug. Also during each visit, subjects were asked about dyspepsia and related GI symptoms using the SODA instrument and had heartburn assessed. If a gastric, duodenal, or esophageal ulcer was detected, study drug was discontinued and the subject was discontinued from the study and placed on appropriate medication for treatment of the ulcer. Subjects completing six months of therapy, discontinued due to GU, or discontinued prematurely returned for a Final Visit for endoscopy (excluding subjects with gastric or duodenal ulcer), SODA and OTE-DP questionnaires and heartburn assessments.

Drug discontinuation was required in circumstances of ulceration, pregnancy, creatinine clearance of < 30 ml/min, or a confirmed > 2.0 g/dL decrease in hemoglobin. A subject was considered to have completed the study if all scheduled assessments at the 6-month visit had been performed, or the primary efficacy endpoint (GU confirmed by endoscopy) had been reached prior to six months. If a DU was detected at any time during study drug treatment, including at the 6-month visit, the subject was withdrawn and was not considered as completing the study. Subjects who were prematurely withdrawn from the study were not replaced.

Subjects were randomized 1:1 to receive one of the following blinded treatments for up to six months: 1) PN 400 (DR naproxen 500 mg / immediate-release esomeprazole 20 mg) tablets; 2) DR naproxen tablet 500 mg, both bid, 30 to 60 minutes before breakfast and 30 to 60 minutes before dinner. Delayed-release (DR) naproxen was used as an active control in order to evaluate the relative effect on GI mucosa of PN 400 compared to naproxen alone. The sponsor claimed that the naproxen active control tablet, which was indistinguishable from the PN 400 tablet with regard to size, shape, and color, was bioequivalent to commercial EC naproxen 500 mg allegedly shown in Study PN400-102. The randomization was stratified by LDA use. Once a subject's eligibility was confirmed and the subject was randomized, he/she was given a 4-digit treatment number obtained from the interactive voice response system (IVRS). Emergency unblinding was allowed only if knowledge of the assigned study drug was deemed necessary to treat the subject and subjects were required to discontinue study drug after.

Endoscopies were performed at Screening Visit 2 prior to randomization and at one, three and six months during the treatment period. The sponsor claimed that every reasonable effort was made to have the same endoscopist perform all endoscopies for a given subject.

The primary efficacy endpoint was the incidence of GUs at any time throughout six months of treatment. An ulcer was defined as a mucosal break of at least 3 mm in diameter (measured by close application of open endoscopic biopsy forceps) with unequivocal crater depth.

The key secondary efficacy/tolerability endpoints of the two pivotal studies were:

- Proportion of subjects with pre-specified NSAID-associated UGI AEs (Adverse Events) or DUs (tolerability endpoint);
- Proportion of subjects discontinuing from the study due to pre-specified NSAID-associated UGI AEs or due to DUs (tolerability endpoint);
- Proportion of subjects developing DUs throughout six months of study treatment.

The sponsor noted that DU was a study endpoint and not collected on the AE Case Report Form (CRF) page, so was stated separately in the two tolerability endpoints. This review will focus on the efficacy endpoints and the two tolerability endpoints will not be discussed in this review. Please refer to the clinical review for more information regarding tolerability endpoints. The secondary efficacy endpoint evaluating DU incidence will be briefly discussed in Section 3.1.4 with recommendation that it should not be considered as a labeling candidate.

The sponsor also investigated some non-key secondary efficacy/tolerability endpoints as follows:

- Proportion of subjects with heartburn ;

- Response on OTE-DP rating;
- Mean change from Baseline for each of the SODA sub-scales;
- Proportion of subjects discontinuing from the study due to any AE or DUs (tolerability endpoint).

Patient reported outcomes (PROs), heartburn assessment, OTE-DP, and SODA were completed at Baseline and at each subsequent study visit; and heartburn assessment and SODA used a 7-day recall period for most of the questions. It should be noted that the sponsor did not propose any multiplicity adjustment for treatment comparisons on any non-key secondary endpoints. So the results on all non-key secondary endpoints listed above should only be considered supportive and exploratory.

The sponsor also proposed two other efficacy endpoints as follows:

- Incidence of gastroduodenal ulcers at any time throughout six months of treatment by low-dose aspirin use (Yes/No) at randomization;
- Incidence of gastroduodenal ulcers at any time throughout six months of treatment.

These endpoints are also exploratory and will not be evaluated in this review.

3.1.1.2 Patient Disposition, Demographic and Baseline Characteristics

The two pivotal studies were conducted concurrently and lasted approximately 12 months. In particular, Study 301 had first subject randomized on September 11, 2007 and last subject completed on September 3, 2008. Study 302 had first subject randomized on September 21, 2007 and last subject completed on September 29, 2008. In total, 861 subjects were randomized in these two efficacy studies with 428 to the PN 400 group and 431 to the naproxen group. Out of these subjects, 854 (more than 99%) were treated with at least one dose of study medication (safety population). Study 301 involved 70 centers in the U.S. and only 59 randomized at least one subject. Study 302 involved 82 U.S. centers and 70 out of them randomized at least one subject. Disposition of the subjects in these two studies is summarized in Table 3.1 below. The sponsor reported that no center enrolled more than 7.5% of the total study population in Study 301 and no center enrolled more than 8.5% of the total study population in Study 302.

The sponsor specified the following analysis populations in the SAPs:

- Intent-to-treat (ITT) population: All randomized subjects who received at least one dose of study drug and had no ulcer detected by endoscopy at the Screening Visit;
- Per-protocol (PP) population: All subjects in the ITT population who did not violate the protocol in any major way that would have impacted the evaluation of efficacy and had at least 70% overall treatment compliance. The sponsor claimed that subjects excluded from the PP population were identified prior to unblinding of the treatment code, and the reason for exclusion was documented.
- Safety population: All randomized subjects who received at least one dose of study drug.

Table 3.1. Accountability and Disposition of Subjects in Studies PN400-301 and PN400-302

	PN 400			EC Naproxen			Total
	301	302	Total	301	302	Total	
Randomized, N	218	212	430	220	211	431	861
Treated (Safety Population)	218 (100%)	210 (99.1%)	428 (99.5%)	216 (98.2%)	210 (99.5%)	426 (98.8%)	854 (99.2%)
ITT Population	218 (100%)	210 (99.1%)	428 (99.5%)	216 (98.2%)	210 (99.5%)	426 (98.8%)	854 (99.2%)
PP Population	203 (93.1%)	180 (84.9%)	383 (89.1%)	201 (91.4%)	180 (85.3%)	381 (88.4%)	764 (88.7%)
Completed study	180 (82.6%)	151 (71.2%)	331 (77.0%)	153 (69.5%)	153 (72.5%)	306 (71.0%)	637 (74.0%)
Completed study without gastric ulcer	171 (78.4%)	136 (64.2%)	307 (71.4%)	103 (46.8%)	102 (48.3%)	205 (47.6%)	512 (59.5%)
Prematurely Discontinued	38 (17.4%)	61 (28.8%)	99 (23.0%)	67 (30.5%)	58 (27.5%)	125 (29.0%)	224 (26.0%)
Adverse event	14 (6.4%)	20 (9.4%)	34 (7.9%)	24 (10.9%)	30 (14.2%)	54 (12.5%)	88 (10.2%)
Withdrew consent	13 (6.0%)	24 (11.3%)	37 (8.6%)	25 (11.4%)	8 (3.8%)	33 (7.7%)	70 (8.1%)
Lost to follow-up	5 (2.3%)	6 (2.8%)	11 (2.6%)	2 (0.9%)	7 (3.3%)	9 (2.1%)	20 (2.3%)
Duodenal ulcer	1 (0.5%)	2 (0.9%)	3 (0.7%)	10 (4.5%)	8 (3.8%)	18 (4.2%)	21 (2.4%)
Other	5 (2.3%)	9 (4.2%)	14 (3.3%)	6 (2.7%)	5 (2.4%)	11 (2.6%)	25 (2.9%)

Source: Reviewer's Table (the results concur with those from the sponsor)

Two studies had noticeable difference in the premature discontinuation patterns. In particular, PN 400 group had higher discontinuation proportion in Study 302 than in Study 301 (28.8% vs. 17.4%) mainly due to withdrawal of consent, while naproxen group had higher discontinuation proportion in Study 301 than in Study 302 (30.5% vs. 27.5%) mainly due to withdrawal of consent as well. The sponsor stated that out of the 24 PN 400 subjects who withdrew their consent in Study 302, five indicated a lack of efficacy and seven not wanting to undergo any additional endoscopies, other reasons included not able to comply with study procedure/scheduling and relocation. The sponsor also stated that of the 25 withdrawal-of-consent subjects under the naproxen treatment in Study 301, five indicated relocation out of the area and three not wanting to undergo any additional endoscopies, other reasons included personal or family emergencies, not able to comply with study procedures/scheduling, lack of efficacy, and needing to start an excluded medication.

For the two studies combined, 65.9% ITT subjects in PN 400 group and 68.3% in naproxen group were female. Racial composition was predominantly White, and the median age was 59 years. The demographic characteristics were generally similar across the treatment arms and the studies except that Study 301 had a larger proportion of elderly patients in comparison to Study 302 and for Study 302 the naproxen group had a larger proportion of white patients compared to the control group. In total, 24.3% of the subjects who were assigned to PN 400 group and 23.6% in naproxen group in the ITT population of Study 301 used LDA, and there were roughly equal histories of documented ulcer (6.0% and 4.6% in PN 400 and naproxen groups, respectively) and

ulcer within the past five years (6.9% and 6.0% in PN 400 and naproxen groups, respectively). Less than one-fourth of the subjects (21.9% in PN 400 group and 24.3% in naproxen group) in the ITT population of Study 302 used LDA. In Study 302, more subjects in naproxen group than in PN 400 group had documented history of ulcer (12.9% vs. 8.6%) and ulcer within previous five years (11.0% vs. 8.6%). Less than one-fourth of the subjects in both treatment groups in the ITT population of the two studies combined used LDA. The percentage of subjects who had a documented ulcer within the past five years was similar between treatment groups in the combined analysis.

3.1.1.3 Statistical Methodologies

For both Studies 301 and 302, the sponsor reported two amendments to the protocol. Amendment 1, dated September 17, 2007, allegedly prior to enrollment of the first subject, updated and clarified some trial conduct specifics. Amendment 2, dated June 2, 2008, added the objective and efficacy variable to assess the effect of concomitant use of LDA on gastroduodenal ulcer incidence within each treatment group. Moreover, Amendment 2 switched the two tolerability endpoints from the safety to the efficacy section. It also excluded subjects who had previously participated in a PN 400 study (although at the time of the amendment, no subject had previously participated in a PN 400 study and enrollment was closed). In addition, Amendment 2 provided the following updates to the statistical analysis section:

- Clarified clinical wording of primary and secondary efficacy endpoint definitions;
- Identified the hierarchical fixed-sequence testing of key secondary and tolerability endpoints as a multiple comparison adjustment, and clarified the step-down procedure for determining statistical significance;
- Specified that the analysis for the effect of concomitant use of LDA on gastroduodenal ulcer incidence would be performed on pooled data from Studies 301 and 302 in the integrated summary of efficacy (ISE);
- Provided the basis for power calculation assumptions for the key secondary endpoints.

The SAP was also amended on September 25, 2008, allegedly prior to database lock, as follows:

- Added a data handling rule for the last assessment date of SODA or heartburn more than ten days after the last dosing date of study drug;
- Added an imputation rule for a partial last dosing date;
- Added a summary of subjects completing the study without a GU;
- Clarified the subgroup of subjects with a history of GU or DU within five years and added a summary of observed GU rate by gender and race;
- Clarified how to summarize the endpoint of gastroduodenal ulcers by LDA use for this study (between and within treatment groups) and how to analyze the endpoint for the ISE;
- Updated SAP Appendix 3 for pre-specified NSAID-associated UGI AE terms.

After the FDA's review of the SAP, a second amendment was made to the SAP prior to database lock (October 15, 2008). Initially, the analysis of the primary efficacy endpoint was planned to be the log-rank test stratified by LDA use (Yes/No) at randomization. However, in a letter dated

October 3, 2008, the FDA recommended using a CMH test stratified by LDA use for the primary analysis of the primary efficacy endpoint. Consequently, an amendment dated October 13, 2008 was made to both the primary analysis of the primary endpoint of incidence of GU and the key secondary endpoints of incidence of DU.

The sponsor also reported that after database lock, it was decided to include additional subgroup analyses of the primary endpoint for ethnicity and smoking status, to add a CMH test for acetaminophen use, and to present additional laboratory shifts tables using the extended laboratory ranges provided by the sponsor.

Most of the above amendments had minimal impact on the primary efficacy assessment. However, the adding and changing order of the key secondary endpoint hierarchy occurred close to the end of the study, very likely after the enrollment concluded. In particular, the original protocol included DU incidence as the first secondary efficacy endpoint without specifying any key secondary endpoints or multiplicity adjustment for the secondary endpoints statistical testing. In Amendment 2, key secondary endpoints section was added to include two tolerability endpoints and the efficacy endpoint of DU incidence with a sequential testing order. Hence, the evaluation of these secondary endpoints should be made with caution.

The determination of sample size for each study was based on the assumption that 15% of subjects treated with naproxen would have a GU over the six-month study, compared to 5% of subjects treated with PN 400. Based on a Fisher's exact test with a two-sided significance level of 5% and 90% power to detect the difference between naproxen and PN 400, it was determined that the sample size in each group was 200. The study also supposedly provided at least 90% power for the three key secondary endpoints using a Fisher's exact test and a two-sided significance level of 5%. However, this determination was part of the analysis plan amendment made late in the study and had no real importance regarding sample size.

All efficacy and tolerability endpoint analyses were performed based on the ITT population. In addition, analyses of the primary efficacy endpoint and the key secondary efficacy and tolerability endpoints were performed using the PP population as a supportive analysis.

The primary efficacy endpoint was the proportion of subjects developing GUs throughout six months of study treatment. The cumulative proportion of subjects developing GUs at six months was analyzed using a CMH test stratified by use of LDA (Yes/No) at randomization. In addition, the sponsor also conducted the Kaplan-Meier estimates of the proportion of subjects developing GUs. Time to GU was calculated from the first day of study drug dispensed to the day of confirmed GU or was censored at the last day endoscopic assessment or date of withdrawal (the last date a subject was seen at the investigative site) if no GU developed. A log-rank test stratified by use of LDA (Yes/No) at randomization was used as a secondary analysis to test the difference between treatment groups in the survival curves.

The sponsor performed a sensitivity analysis in which premature withdrawals without a confirmed GU were classified as developing a GU at six months if the subject developed a DU or discontinued due to a pre-specified UGI AE that might have been an indication of a GU. The

cumulative proportion of subjects developing GUs at one month and three months was summarized and analyzed by the sponsor as additional analyses using a CMH test adjusting for use of LDA (Yes/No) at randomization. Moreover, the sponsor conducted an exploratory analysis of the primary endpoint using a conditional logistic regression model. The model included treatment as main effect, use of LDA (Yes/No) as a stratum and history of gastroduodenal ulcer within five years of randomization (Yes/No), and age group (< 60, or ≥ 60 years) as covariates.

The sponsor performed treatment comparisons for the following key secondary efficacy and tolerability endpoints in a sequential order as shown below:

1. The proportion of subjects with pre-specified NSAID-associated UGI AEs (specified in the SAPs allegedly prior to database lock) or DUs;
2. The proportion of subjects discontinuing from the study due to NSAID-associated UGI AEs (specified in the SAPs allegedly prior to database lock) or due to DUs;
3. The proportion of subjects developing DUs throughout six months of study treatment.

The hierarchical fixed-sequence testing approach was used to adjust for multiple comparisons. These endpoints were tested in the specified sequence above with the rule that once a p-value exceeded 0.05, endpoints further down in the sequence were not claimed for statistical significance. The treatment comparisons of the first two key secondary (tolerability) endpoints were performed using a CMH test adjusting for LDA use (Yes/No) at the randomization. The last key secondary (efficacy) endpoint was analyzed in the same manner as the primary efficacy endpoint.

For any treated subject who withdrew prematurely from the study, all available data up to the time of discontinuation were included in the summaries. The sponsor did not specify any missing data handling strategies for the primary endpoint presumably because that the time-to-event type of responses and log-rank test for the treatment comparison were originally intended. However, the sponsor indicated that for the primary endpoint, three categories, including “Gastric Ulcer”, “Maintained Gastric Ulcer Free”, and “Discontinued Gastric Ulcer Free”, were used for the summary and the last two categories were collapsed into one as “Gastric Ulcer Free” for the comparison between treatment groups. In other words, the missing/discontinuation cases without GU reported were treated as “treatment successes” by the sponsor. This is not the usual practice for the primary analysis; however, in this case the results from this imputation method were not in favor of PN 400. More details are in the next section.

3.1.1.4 Results and Conclusions

Observed from Table 3.2 below, significantly fewer subjects who took PN 400 in Studies 301 and 302 had GU than subjects who took EC naproxen at the end of the treatment period (six months). Conventionally, the missing efficacy data would be imputed as treatment failures for the primary analyses. However, in this case, more missing data in the EC naproxen group than in PN 400 for Study 301 and the result was more significantly in favor of PN 400. While for Study 302, missing data was more balanced than those for Study 301, the result was somewhat unchanged. In conclusion, the sponsor’s approach of imputing missing data as treatment successes was the more conservative analysis from a regulatory perspective.

Table 3.2. Proportion of ITT Subjects with Gastric Ulcers by 6 Months

		PN400-301		PN400-302	
		PN 400	EC naproxen	PN 400	EC naproxen
N		218	216	210	210
Gastric Ulcer		9 (4.1%)	50 (23.1%)	15 (7.1%)	51 (24.3%)
Gastric Ulcer Free		209 (95.9%)	166 (76.9%)	195 (92.9%)	159 (75.7%)
	Maintained gastric ulcer free	171 (78.4%)	103 (47.7%)	136 (64.8%)	102 (48.6%)
	Discontinued gastric ulcer free	38 (17.4%)	63 (29.2%)	59 (28.1%)	57 (27.1%)
Difference (95% CI ^a) of (EC naproxen – PN 400)	Missing =		19.0%		17.1%
	Success		(13.0%, 25.6%)		(10.4%, 24.1%)
	Missing =		30.8%		16.2%
	Failure		(21.9%, 39.3%)		(6.6%, 25.5%)
p-value ^b			< 0.001		< 0.001

CI: Confidence Interval

^a Exact CI.

^b CMH test on ulcer occurrence stratified by LDA use (Yes/No).

Source: Reviewer's Table (the results concur with those from the sponsor)

Sensitivity analyses performed by the sponsor and this reviewer did not reveal any inconsistency of the primary analysis results, mainly due to the fact that more DUs or discontinuations occurred in the EC naproxen group than in the PN 400 group. The sponsor also conducted survival analyses with some regrouped data to adjust non-exact GU occurrence time. As expected, the results were consistent with the cumulative analyses with higher estimated GU-free rates for both treatment groups. The sponsor's survival analyses should be considered exploratory due to all the adjustments made.

The three key secondary endpoints proposed by the sponsor contained two tolerability endpoints and one efficacy endpoint (DU incidence) as stated above. The tolerability endpoints were added as key secondary endpoints toward the end of the studies. The results were all in favor of the PN 400 treatment; however, the results for the tolerability endpoints will not be discussed in details within this review. It is not recommended to have these endpoints within the efficacy section of the label due to the safety nature of the tolerability endpoints and the fact that these were redefined late in the study.

The only efficacy endpoint of the three key secondary (efficacy and tolerability) endpoints is DU incidence by six-month. This endpoint was one of the secondary endpoints in the original protocol; however, the sponsor did not indicate any particular statistical testing or multiplicity adjustment at that time. It was brought into the key secondary endpoints list to be tested in Amendment 2; consequently, results regarding DU incidence should be considered exploratory. The DU incidence was very low across treatment arms and studies. The results are presented in the following table.

Table 3.3. Proportion of ITT Subjects with Duodenal Ulcers by 6 Months

	PN400-301		PN400-302	
	PN 400	EC naproxen	PN 400	EC naproxen
N	218	216	210	210
Duodenal Ulcer	1 (0.5%)	11 (5.1%)	2 (1.0%)	12 (5.7%)
Duodenal Ulcer Free	217 (99.5%)	205 (94.9%)	208 (99.0%)	198 (94.3%)
Maintained duodenal ulcer free	171 (78.4%)	103 (47.7%)	136 (64.8%)	102 (48.6%)
Discontinued duodenal ulcer free ^a	46 (21.1%)	102 (47.2%)	72 (34.3%)	96 (45.7%)
Difference (95% CI ^b) of (EC naproxen – PN 400)	Missing =	4.6%		4.7%
	Success	(1.7%, 8.6%)		(1.5%, 9.0%)
	Missing =	30.8%		16.2%
	Failure	(21.9%, 39.3%)		(6.6%, 25.5%)
p-value ^c		0.008		0.016

CI: Confidence Interval

^a Includes subjects with gastric ulcers.^b Exact CI.^c CMH test on ulcer occurrence stratified by LDA use (Yes/No).

Source: Reviewer's Table (the results concur with those from the sponsor except for the p-values)

The sponsor was notified before the submission that Dr. Marker (Site 401) had been the subject of an inspection by the FDA for clinical studies from other sponsors. The three subjects from that site were all in Study 301 and completed the study, none had taken LDA, and none had developed gastric or duodenal ulcers during the treatment period. Two of them were in PN 400 group and the other was in EC naproxen group. The sponsor decided to leave them in for the primary efficacy analysis, and this is acceptable. Excluding these three subjects had no impact on the efficacy results.

3.1.2 Integrated Efficacy Analysis

A combined efficacy analysis was conducted and submitted by the sponsor by pooling the two pivotal studies. Given the highly significant results for each individual study and the much larger combined sample size, the higher significance of the pooled analysis is to be expected; however, these results are only exploratory but useful for subgroup comparisons in Section 4. The primary analysis results on the primary endpoint of GU incidence are presented in the table below.

Table 3.4. Proportion of Pooled ITT Subjects with Gastric Ulcers by 6 Months

	PN 400	EC naproxen
N	428	426
Gastric Ulcer	24 (5.6%)	101 (23.7%)
Gastric Ulcer Free	404 (94.4%)	325 (76.3%)
Maintained gastric ulcer free	307 (71.7%)	205 (48.1%)
Discontinued gastric ulcer free	97 (22.7%)	120 (28.2%)
Difference (95% CI ^a) of (EC naproxen – PN 400)	Missing = Success	18.1% (13.6%, 22.8%)
	Missing = Failure	23.6% (17.1%, 29.9%)
p-value ^b		< 0.001

CI: Confidence Interval

^a Exact CI.^b CMH test on ulcer occurrence stratified by LDA use (Yes/No).

Source: Reviewer's Table (the results concur with those from the sponsor)

3.2 Study PN400-303

A third phase 3 Study PN400-303 was conducted to evaluate the incidence of GU with PN 400 versus diclofenac/misoprostol in subjects who are at high risk for developing NSAID-associated ulcers. A high risk population refers to subjects with a history of a serious UGI event such as perforation, obstruction or bleeding. This study was a randomized, double-blind, parallel-group, controlled, multi-center clinical trial of a six months duration. The active control treatment was ARTHROTEC[®] (75 mg diclofenac/ 200 mcg misoprostol) bid. It was planned to involve approximately 100 U.S. sites to enroll a total of 200 subjects (100 per arm). At least 20% of the subjects enrolled were planned to be age 65 years and older. The study consisted of a Screening period, including washout of disallowed medications of no more than 14 days, a Baseline Endoscopy Visit and three outpatient visits over a six-month period, or until gastric ulcer(s) were confirmed by endoscopy.

Study PN400-303 involved 66 U.S. centers with 16 centers randomized at least one subject. A total of 20 subjects were randomized and treated. The first subject was screened on November 23, 2007, and the last subject completed on June 5, 2008. On May 2, 2008 (prior to any subject completing six months of therapy) POZEN reached an agreement with the FDA to terminate the study due to low and inadequate enrollment. Nine subjects were randomized to the PN 400 group and 11 subjects to the ARTHROTEC[®] group. Although six months of treatment was planned, actually one subject finished four months of treatment, three subjects completed three months, and all others completed less than three months. This study had three subjects developed GUs during the study course. Two of them were in PN 400 group and the other in ARTHROTEC[®] group. The results are only supportive.

3.3 Overview of Safety

This review will only briefly discuss the safety assessments from this application. For more details on the safety of PN 400, please refer to the clinical review.

The sponsor reported that the clinical development plan treated 2337 subjects in six phase 3 studies and 214 normal healthy volunteer (NHV) subjects in seven phase 1 studies. In total, PN 400 Tablets were given to 1326 subjects in the clinical development program; 1166 of these were in phase 3 studies and 1157 were in the studies pooled for safety analysis. Reportedly of these, 491 subjects took PN 400 for six months and 135 took PN 400 for 12 months. The total exposure of subjects to PN 400 in the six phase 3 studies examined in this application was reportedly 458 patient-years.

The study reports indicated that gastritis and diarrhea were reported more frequently in those who took PN 400 compared to those who took EC naproxen. Subjects who took PN 400 did not report increases in frequency of non-GI AEs, or report increases in severity of Adverse Drug Reactions (ADRs) or emergence of new or unexpected ADRs compared to EC naproxen.

There were 58 Serious Adverse Events (SAEs) reported by 53 subjects in the development program all of which occurred in the phase 3 studies. The rates of SAEs in the treatment groups

were reported similar, except fewer were reported in the placebo group (0.4%) than in the PN 400, EC naproxen, or celecoxib groups (2.7%, 3.1%, and 1.6%, respectively).

The sponsor concluded that the overall safety profile of PN 400 was consistent with the profile that has been established for oral naproxen or esomeprazole monotherapy. Moreover, the sponsor reported that continued exposure to PN 400 throughout 12 months did not increase the rates or severity of Treatment Emergent Adverse Event (TEAE) or laboratory changes.

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

The sponsor conducted the primary analyses of the primary efficacy endpoint on the following subgroups in the ITT population when appropriate:

- Use of LDA (Yes/No)
- Age (< 60, or ≥ 60 years)
- History of gastric or duodenal ulcer within the previous five years (Yes/No)
- Race (White, Black, Other)
- Gender
- Ethnicity
- Smoking Status (Yes/No)

All these subgroup analyses were validated by this reviewer but only some of them will be discussed in the sections below per clinical importance. Except as discussed below, these analyses did not appear to show any differences in treatment effect among the sub-populations.

4.1 Gender, Race and Age

This reviewer’s subgroup analyses on age, gender, and race results are presented in Table 4.1 below. The PN 400 appears to have larger treatment effect compared to EC naproxen for patients with more advanced age. However, it should be noted that patients 50 years or older did not need to have history of ulcers to enter the study, while the younger counterparts had that requirement at the baseline. Also it should be noted that male and non-White subgroups only contained small sample sizes. In general, the subgroups did not reveal any internal inconsistencies on the primary endpoint of GU incidence.

Table 4.1. Proportion of ITT Subjects with Gastric Ulcers by 6 Months by Subgroups

	PN400-301		PN400-302		Combined	
	PN 400	Naproxen	PN 400	Naproxen	PN 400	Naproxen
Age (< 60 years)						
N	105	97	111	120	216	217
Gastric Ulcer	8 (7.6%)	20 (20.6%)	10 (9.0%)	26 (21.7%)	18 (8.3%)	46 (21.2%)
Gastric Ulcer Free	97 (92.4%)	77 (79.4%)	101 (91.0%)	94 (78.3%)	198 (91.7%)	171 (78.8%)
Maintained gastric ulcer free	76 (72.4%)	46 (47.4%)	74 (66.7%)	63 (52.5%)	150 (69.4%)	109 (50.2%)

Table 4.1. (Cont'd) Proportion of ITT Subjects with Gastric Ulcers by 6 Months by Subgroups

	PN400-301		PN400-302		Combined	
	PN 400	Naproxen	PN 400	Naproxen	PN 400	Naproxen
Age (< 60 years)						
Discontinued gastric ulcer free	21 (20.0%)	31 (32.0%)	27 (24.3%)	31 (25.8%)	48 (22.2%)	62 (28.6%)
Difference (95% CI ^a) of (EC naproxen – PN 400)	13.0% (3.2%, 23.2%)		12.7% (3.4%, 22.1%)		12.9% (6.3%, 19.7%)	
p-value ^b	0.014		0.013		< 0.001	
Age (≥ 60 years)						
N	113	119	99	90	212	209
Gastric Ulcer	1 (0.9%)	30 (25.2%)	5 (5.1%)	25 (27.8%)	6 (2.8%)	55 (26.3%)
Gastric Ulcer Free	112 (99.1%)	89 (74.8%)	94 (94.9%)	65 (72.2%)	206 (97.2%)	154 (73.7%)
Maintained gastric ulcer free	95 (84.1%)	57 (47.9%)	62 (62.6%)	39 (43.3%)	157 (74.1%)	96 (45.9%)
Discontinued gastric ulcer free	17 (15.0%)	32 (26.9%)	32 (32.3%)	26 (28.9%)	49 (23.1%)	58 (27.8%)
Difference (95% CI ^a) of (EC naproxen – PN 400)	24.3% (16.9%, 33.3%)		22.7% (12.6%, 33.6%)		23.5% (17.2%, 30.2%)	
p-value ^b	< 0.001		< 0.001		< 0.001	
Gender (Female)						
N	150	149	132	142	282	291
Gastric Ulcer	4 (2.7%)	35 (23.5%)	11 (8.3%)	36 (25.4%)	15 (5.3%)	71 (24.4%)
Gastric Ulcer Free	146 (97.3%)	114 (76.5%)	121 (91.7%)	106 (74.6%)	267 (94.7%)	220 (75.6%)
Maintained gastric ulcer free	115 (76.7%)	72 (48.3%)	84 (63.6%)	73 (51.4%)	199 (70.6%)	145 (49.8%)
Discontinued gastric ulcer free	31 (20.7%)	42 (28.2%)	37 (28.0%)	33 (23.2%)	68 (24.1%)	75 (25.8%)
Difference (95% CI ^a) of (EC naproxen – PN 400)	20.8% (13.7%, 28.7%)		17.0% (8.4%, 25.9%)		19.1% (13.6%, 24.9%)	
p-value ^b	< 0.001		< 0.001		< 0.001	
Gender (Male)						
N	68	67	78	68	146	135
Gastric Ulcer	5 (7.4%)	15 (22.4%)	4 (5.1%)	15 (22.1%)	9 (6.2%)	30 (22.2%)
Gastric Ulcer Free	63 (92.6%)	52 (77.6%)	74 (94.9%)	53 (77.9%)	137 (93.8%)	105 (77.8%)
Maintained gastric ulcer free	56 (82.4%)	31 (46.3%)	52 (66.7%)	29 (42.6%)	108 (74.0%)	60 (44.4%)
Discontinued gastric ulcer free	7 (10.3%)	21 (31.3%)	22 (28.2%)	24 (35.3%)	29 (19.9%)	45 (33.3%)
Difference (95% CI ^a) of (EC naproxen – PN 400)	15.0% (3.0%, 27.8%)		16.9% (5.8%, 28.9%)		16.1% (8.2%, 24.6%)	
p-value ^b	0.034		0.006		< 0.001	

CI: Confidence Interval

^a Exact CI.

^b CMH test on ulcer occurrence stratified by LDA use (Yes/No).

Table 4.1. (Cont'd) Proportion of ITT Subjects with Gastric Ulcers by 6 Months by Subgroups

	PN400-301		PN400-302		Combined	
	PN 400	Naproxen	PN 400	Naproxen	PN 400	Naproxen
Race (White)						
N	184	181	183	190	367	371
Gastric Ulcer	8 (4.3%)	42 (23.2%)	11 (6.0%)	46 (24.2%)	19 (5.2%)	88 (23.7%)
Gastric Ulcer Free	176 (95.7%)	139 (76.8%)	172 (94.0%)	144 (75.8%)	348 (94.8%)	283 (76.3%)
Maintained gastric ulcer free	141 (76.6%)	88 (48.6%)	117 (63.9%)	91 (47.9%)	258 (70.3%)	179 (48.2%)
Discontinued gastric ulcer free	35 (19.0%)	51 (28.2%)	55 (30.1%)	53 (27.9%)	90 (24.5%)	104 (28.0%)
Difference (95% CI ^a) of (EC naproxen – PN 400)	18.9% (12.2%, 26.1%)		18.2% (11.2%, 25.5%)		18.5% (13.7%, 23.6%)	
p-value ^b	< 0.001		< 0.001		< 0.001	
Race (non-White)						
N	34	35	27	20	61	55
Gastric Ulcer	1 (2.9%)	8 (22.9%)	4 (14.8%)	5 (25.0%)	5 (8.2%)	13 (23.6%)
Gastric Ulcer Free	33 (97.1%)	27 (77.1%)	23 (85.2%)	15 (75.0%)	56 (91.8%)	42 (76.4%)
Maintained gastric ulcer free	30 (88.2%)	15 (42.9%)	19 (70.4%)	11 (55.0%)	49 (80.3%)	26 (47.3%)
Discontinued gastric ulcer free	3 (8.8%)	12 (34.3%)	4 (14.8%)	4 (20.0%)	7 (11.5%)	16 (29.1%)
Difference (95% CI ^a) of (EC naproxen – PN 400)	19.9% (4.3%, 37.5%)		10.2% (13.7%, 35.8%)		15.4% (2.2%, 29.7%)	
p-value ^b	0.049		0.664		0.046	

CI: Confidence Interval

^a Exact CI.

^b CMH test on ulcer occurrence stratified by LDA use (Yes/No).

Source: Reviewer's Table (the results concur with those from the sponsor)

4.2 Other Special/Subgroup Populations

Treatment group comparisons within LDA use subgroups are summarized in the table below. Comparisons are based on the chi-square test, as CMH testing stratified by LDA use would not be applicable. It should be noted that these results should only be considered exploratory.

Table 4.2. Proportion of ITT Subjects with Gastric Ulcers by 6 Months by LDA Uses

	PN400-301		PN400-302		Combined	
	PN 400	Naproxen	PN 400	Naproxen	PN 400	Naproxen
LDA use (Yes)						
N	53	51	46	51	99	102
Gastric Ulcer	1 (1.9%)	12 (23.5%)	2 (4.3%)	17 (33.3%)	3 (3.0%)	29 (28.4%)
Gastric Ulcer Free	52 (98.1%)	39 (76.5%)	44 (95.7%)	34 (66.7%)	96 (97.0%)	73 (71.6%)
Maintained gastric ulcer free	45 (84.9%)	25 (49.0%)	31 (67.4%)	16 (31.4%)	76 (76.8%)	41 (40.2%)

Table 4.2. (Cont'd) Proportion of ITT Subjects with Gastric Ulcers by 6 Months by LDA Uses

	PN400-301		PN400-302		Combined	
	PN 400	Naproxen	PN 400	Naproxen	PN 400	Naproxen
LDA use (Yes)						
Discontinued gastric ulcer free	7 (13.2%)	14 (27.5%)	13 (28.3%)	18 (35.3%)	20 (20.2%)	32 (31.4%)
Difference (95% CI ^a) of (EC naproxen – PN 400)	21.6% (9.9%, 35.6%)		29.0% (14.4%, 43.6%)		25.4% (16.1%, 35.3%)	
p-value ^b	0.002		< 0.001		< 0.001	
LDA use (No)						
N	165	165	164	159	329	324
Gastric Ulcer	8 (4.8%)	38 (23.0%)	13 (7.9%)	34 (21.4%)	21 (6.4%)	72 (22.2%)
Gastric Ulcer Free	157 (95.2%)	127 (77.0%)	151 (92.1%)	125 (78.6%)	308 (93.6%)	252 (77.8%)
Maintained gastric ulcer free	126 (76.4%)	78 (47.3%)	105 (64.0%)	86 (54.1%)	231 (70.2%)	164 (50.6%)
Discontinued gastric ulcer free	31 (18.8%)	49 (29.7%)	46 (28.0%)	39 (24.5%)	77 (23.4%)	88 (27.2%)
Difference (95% CI ^a) of (EC naproxen – PN 400)	18.2% (11.1%, 25.8%)		13.5% (5.9%, 21.4%)		15.8% (10.7%, 21.2%)	
p-value ^b	< 0.001		0.001		< 0.001	

CI: Confidence Interval

^a Exact CI.

^b Chi-Square test on ulcer occurrence.

Source: Reviewer's Table (the results concur with those from the sponsor)

5. SUMMARY AND CONCLUSIONS

The efficacy of PN 400 Tablets, delayed-released 500 mg naproxen in combination with immediate-release 20 mg esomeprazole twice daily was demonstrated in two identical phase 3 controlled clinical studies designed to show reduction in occurrence of gastric ulcers at the end of a six-month treatment period as compared to enteric-coated 500 mg naproxen twice daily. The efficacy appears to be consistent among age, gender, and racial subgroups. The secondary endpoints based on duodenal ulcer occurrence and NSAID-associated upper gastrointestinal adverse events were added as key secondary variables after the enrollment concluded and should be considered exploratory and not candidates for efficacy labeling.

Application Type/Number	Submission Type/Number	Submitter Name	Product Name
NDA-22511	ORIG-1	POZEN INC	PN 400 NAPROXEN/ESOMEPRAZOLE MAGNESIUM

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

/s/

FREDA COONER
04/17/2010

MICHAEL E WELCH
04/19/2010
Concur with review.

**Screening of New NDA
Division of Biometrics II**

Date: 8/19/09

NDA #: 22-511

Priority Classification: S

Trade Name: VIMOVO

Applicant: Pozen, Inc.

Generic Name: Naproxen/esomeprazole

Date of Submission: 6/30/09

Indication: treatment of signs and symptoms of osteoarthritis, rheumatoid arthritis, and ankylosing spondylitis in patients at risk for developing NSAID-associated gastric ulcers

No. of Controlled Studies: 2

User Fee Goal Date: 4/30/10

Date of 45-Day Meeting: 8/19/09

Medical Officer: Jin Chen, M.D. (DAARP)

Project Manager: Chantal Phillips (DGP)

Screened by: Kate Meaker, M.S.

Statistical sections: Sections 2.5, 2.7, and 5.3.5

Anticipated Review Completion Date: 1/15/10

Comments:

1. This NDA is in the Division of Gastroenterology Products (DGP). A formal consult was requested of DAARP to review two non-inferiority studies of the efficacy in reduction of pain in patients with osteoarthritis.
2. It is fileable.

CHECKLIST

Item	Check (NA if not applicable)
Index sufficient to locate necessary reports, tables, etc.	Yes
Original protocols & subsequent amendments available in the NDA	Yes
Designs utilized appropriate for the indications requested	Yes
Endpoints and methods of analysis spelled out in the protocols	Yes
Interim analyses (if present) planned in the protocol and appropriate adjustments in significance level made	NA
Appropriate references included for novel statistical methodology (if present)	NA
Sufficient data listings and intermediate analysis tables to permit statistical review	Yes
Data from primary studies in electronic data room	Yes
Intent-to-treat analysis	Yes
Effects of dropouts on primary analyses investigated	Yes
Safety and efficacy for gender, racial, and geriatric subgroups investigated	Yes

BRIEF SUMMARY OF CONTROLLED CLINICAL TRIALS

Study Number (Dates Conducted)	Number of Centers (Locations)	Sample Size	Type of Control	Design	Duration of Treatment
PN400-307 (4/08 – 12/08)	75 centers (All US)	PN400 n=248 Celecoxib n=247 Placebo n=124	Active control (celecoxib 200 mg qd) Placebo	Randomized, Double-blind, Parallel Arm, Active-controlled and Placebo-controlled, Multicenter	12 weeks
PN400-309 (4/08 – 12/08)	82 centers (All US)	PN400 n=244 Celecoxib n=247 Placebo n=124	Active control (celecoxib 200 mg qd) Placebo	Randomized, Double-blind, Parallel Arm, Active-controlled and Placebo-controlled, Multicenter	12 weeks

Katherine B. Meaker
Mathematical Statistician

Concur: Dionne Price Ph.D.
Team Leader

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
10/01/2009

DIONNE L PRICE
10/01/2009
Concur