CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

21-341

STATISTICAL REVIEW(S)

Statistical Review and Evaluation

NDA 21-341
Name of Drug: Valdecoxib
Proposed Indications: OA, RA and Analgesia
Applicant: GD Searle Corp.
Documents Reviewed: Electronic Submission of the Statistical Section Received by CDER on 01/15/2001
Medical Reviewer: Kent Johnson, M.D.
Statistical Reviewer: Laura Lu, Ph.D.
Date of Review: February-November, 2001

I. Introduction

NDA21-341 has been submitted for approval of valdecoxib for treatment of acute pain in adults, treatment of primary Dysmenorrhea in women, relief of signs and symptoms of osteoarthritis (OA) and rheumatoid arthritis (RA) in adults. This review focuses on the phase III trials of [redacted] analgesia, dysmenorrea, [redacted] OA and RA. [redacted] studies will be reviewed by Dr. Jyoti Zalkikar.

II. [redacted] Analgesia Studies

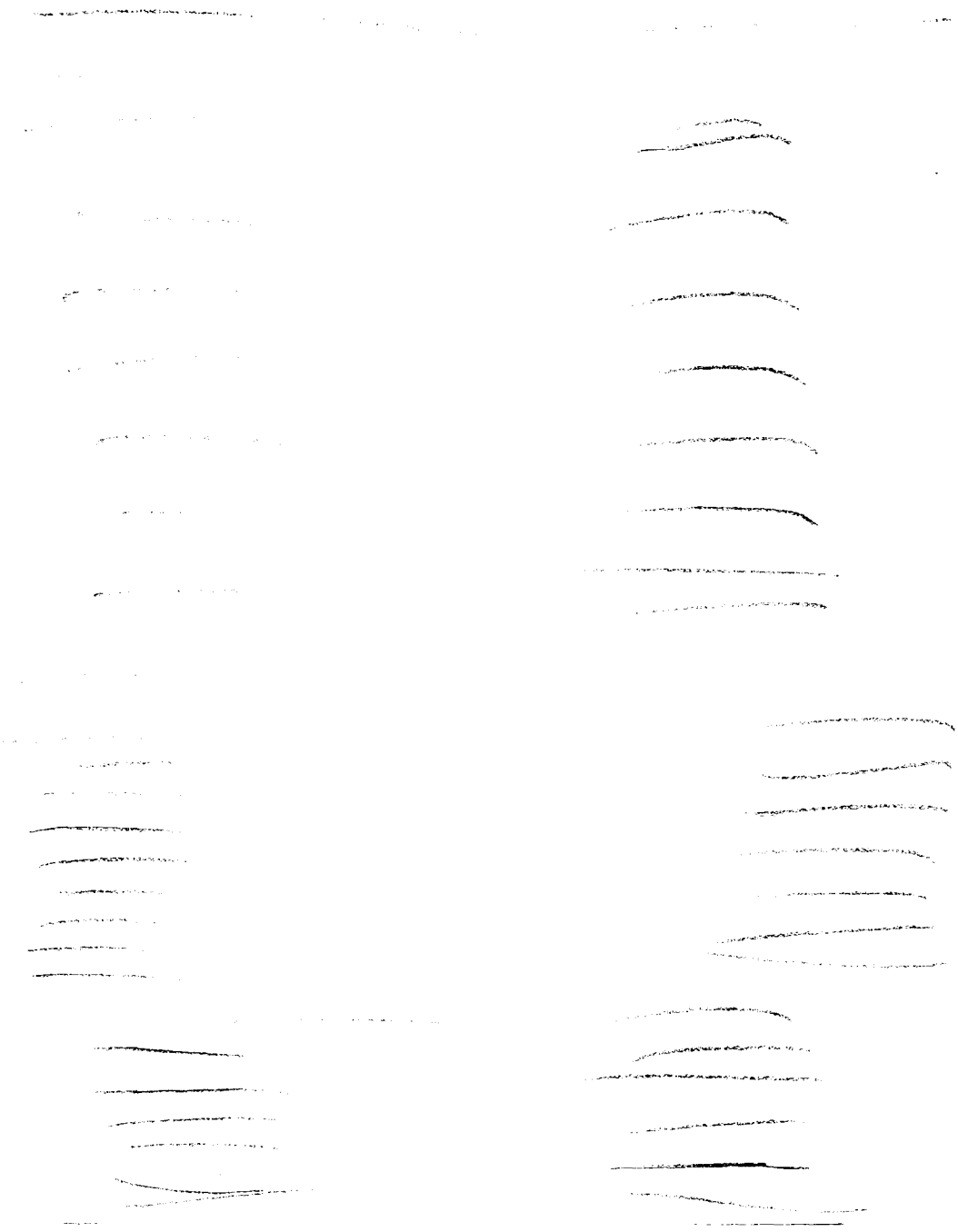
II.1 Study 010

II.1.i Protocol

[redacted]

28 page(s) have been removed because it contains trade secret and/or confidential information that is not disclosable.

2-29



III. Treatment of Dysmenorrhea

III. 1 Study 065

III.1.i Protocol

This is a randomized, placebo- and active-controlled, single center, double-blind, complete block crossover study. This study is designed to demonstrate that valdecoxib 20 mg and valdecoxib 40 mg are efficacious compared to placebo and naproxen sodium 550 mg for the treatment of pain associated with primary dysmenorrhea. The dosing regimen was twice daily as needed for all treatments. The primary study objective is to compare the analgesic efficacy of the initial dose of valdecoxib 20 mg and 40 mg versus placebo in relieving moderate or severe menstrual cramping pain due to primary dysmenorrhea.

Patients who have a history of primary dysmenorrhea that consistently includes menstrual cramping pain of moderate or severe intensity for at least four of the six months immediately preceding study entry and who fulfill all other inclusion/exclusion criteria will be eligible for enrollment into the study. Patients with a history of mild to moderate primary dysmenorrhea will be randomized into one of the following four sequence groups where each sequence as well as each period is a complete randomized block in a 4x4 balanced Latin square in the generalized Youden square design.

SEQUENCE	TREATMENT PERIOD 1	TREATMENT PERIOD 2	TREATMENT PERIOD 3	TREATMENT PERIOD 4
1	Placebo	Valdecoxib 20 mg	Valdecoxib 40 mg	Naproxen 550 mg
2	Valdecoxib 20 mg	Naproxen 550 mg	Placebo	Valdecoxib 40 mg
3	Valdecoxib 40 mg	Placebo	Naproxen 550 mg	Valdecoxib 20 mg
4	Naproxen 550 mg	Valdecoxib 40 mg	Valdecoxib 20 mg	Placebo

The primary measurement of efficacy will consist of the time-weighted sum of pain relief (TOTPAR) and the time weighted sum of pain intensity difference (SPID) over the 8 hour and 12 hour periods after the first dose of study medication. Patients will assess pain intensity (categorical) just prior to taking the first dose of study medication in each cycle (hour 0) and will assess pain intensity (categorical) and pain relief at 0.5, 0.75, 1, 1.5, 2, 3, 4, 5, 6, 7, 8, and 12 hour intervals (or immediately prior to rescue medication or remedication). The secondary measurements of efficacy of the first dose of study medication will include time specific pain relief (PR), time specific pain intensity difference (PID); peak pain relief (PPR); peak pain intensity difference (PPID); time to rescue medication or first remedication, whichever comes first; percentage of patients taking rescue medication before the second dose of study medication; global evaluation of the study medication by the patients recorded at the 12 hour time point (or immediately prior to rescue medications or remedication). The variables for the assessment of dosing regimen include average number of doses of study medication per day, average time between doses, and percentage of patients taking study medication for three days.

For variables pain relief and pain intensity (categorical), if 1or 2 consecutive missing values occur between two time points in which data are obtained, linear interpolation will be used to estimate the missing values. If three or more missing values in a row occur, or there are no evaluations after a certain time point, missing values will be imputed using the last observation carried forward (LOCF).

Time weighted sum of pain relief (TOTPAR), time weighted sum of pain intensity difference (SPID), time specific pain relief (PR), time specific pain intensity difference (PID), peak pain relief (PPR), peak pain intensity difference (PPID), and global evaluation of study mediation prior to rescue or the first remedication on Day 1 will be analyzed by using analyses of variance (ANOVA) with effects for treatment, period, sequence, and patient(sequence). Patient effects are random and all other effects are fixed. Time to rescue medication will be calculated by subtracting the time of administration of first dose of study medication from the time of the rescue medication. Time to first remedication will be calculated by subtracting the time of administration of first dose of study medication from the start time of the first remedication. Time to rescue medication or first remedication will be the minimum of time to rescue medication and first remedication. Time to rescue or first remedication will be analyzed by using survival analysis method. Global evaluation of study medication for the first dose on Day 1, and percentage of patients taking rescue medication will be analyzed by using categorical data analysis method.

The sample size calculation is based on two primary efficacy variables; time weighted sum of pain relief score over the 8 hour period after the first dose study medication (TOTPAR8) and time weighted pain intensity difference over the 8 hour period after the first dose of study medication (SPID8) and the two primary comparisons: each dose of valdecoxib (20 mg and 40 mg) versus placebo. With 90% power and type I error of 0.025 (for a two sided test adjusted for two comparisons), a sample size of 92 patients per treatment group will be needed to detect at least a difference 3 in SPID8 score between valdecoxib 20 mg and placebo, valdecoxib 40 mg and placebo, with estimate of variability at most 5.7 for SPID8.

III.1.ii Sponsor's Main Study Results

a) Patient Disposition

Of the 118 randomized patients, 10 withdrew prior to being dosed with any study medication. An additional 12 patients (5 randomized to sequence 1, 3 randomized to sequence 2, 2 randomized to sequence 3, 2 randomized to sequence 4) were withdrawn from the study during various treatment cycles, resulting in 96 patients who completed all 4 cycles. One patient (0062), who had been marked as having completed the study, had taken rescue medication before 1 hour in cycle 3 and was subsequently considered non-evaluable, resulting in 95 patients evaluable for efficacy. Table 22 below presents detailed information for patient dsiposition.

Table 22. Patient Disposition in Each Cycle

		TREATMENT SEQUENCE 1 (N= 29)	TREATMENT SEQUENCE 2 (N= 30)	TREATMENT SEQUENCE 3 (N= 30)	TREATMENT SEQUENCE 4 (N= 29)	TOTAL (N=118)
CYCLE 1	ENTERED	27	27	29	25	108
	WITHDRAWN FROM STUDY	2	3	1	2	8
	REASON FOR WITHDRAWAL (a)					
	LOST TO FOLLOW-UP	0	0	0	0	0
	PRE-EXISTING VIOLATION	0	0	0	0	0
	PROTOCOL NONCOMPLIANCE	1	3	1	2	7
	ADVERSE SIGN	0	0	0	0	0
	TWO CONSECUTIVE INELIGIBLE CYCLES	1	0	0	0	1
	TOOK RESCUE MEDICATION WITHIN ONE	0	0	0	0	0
CYCLE 2	ENTERED	25	24	28	23	100
	WITHDRAWN FROM STUDY	3	0	1	0	4
	REASON FOR WITHDRAWAL (a)					
	LOST TO FOLLOW-UP	1	0	0	0	1
	PRE-EXISTING VIOLATION	1	0	0	0	1
	PROTOCOL NONCOMPLIANCE	1	0	1	0	2
	ADVERSE SIGN	0	0	0	0	0
	TWO CONSECUTIVE INELIGIBLE CYCLES	0	0	0	0	0
	TOOK RESCUE MEDICATION WITHIN ONE	0	0	0	0	0
CYCLE 3	ENTERED	22	24	27	23	96
	WITHDRAWN FROM STUDY	0	0	0	0	0
	REASON FOR WITHDRAWAL (a)					
	LOST TO FOLLOW-UP	0	0	0	0	0
	PRE-EXISTING VIOLATION	0	0	0	0	0
	PROTOCOL NONCOMPLIANCE	0	0	0	0	0
	ADVERSE SIGN	0	0	0	0	0
	TWO CONSECUTIVE INELIGIBLE CYCLES	0	0	0	0	0
	TOOK RESCUE MEDICATION WITHIN ONE	0	0	0	0	0
CYCLE 4	ENTERED	22	24	27	23	96
	WITHDRAWN FROM STUDY	0	0	0	0	0
	REASON FOR WITHDRAWAL (a)					
	LOST TO FOLLOW-UP	0	0	0	0	0
	PRE-EXISTING VIOLATION	0	0	0	0	0
	PROTOCOL NONCOMPLIANCE	0	0	0	0	0
	ADVERSE SIGN	0	0	0	0	0
	TWO CONSECUTIVE INELIGIBLE CYCLES	0	0	0	0	0
	TOOK RESCUE MEDICATION WITHIN ONE	0	0	0	0	0
COMPLETED STUDY (4 CYCLES)		22	24	27	23	96

Note: Treatment Sequence 1 = A B C D, 2 = B D A C, 3 = C A D B, 4 = D C B A.
Where A = PLACEBO, B = VALDECOXIB 20 mg, C = VALDECOXIB 40 mg, D = NAPROXEN NA 550 mg.
(a) Mutually exclusive and exhaustive categories.

b) Demographics

Demographics and baseline characteristics were generally comparable across treatment groups.

c) Efficacy Evaluation

The 'statistical significance' in the following description was claimed based on Fisher's LSD test for all active treatment groups vs. placebo. Please see the reviewer's comment #3 for the adjustment methods of multiplicity in Section VII.

Primary Endpoints

SPID: Mean SPID scores ranged from 9.77 to 10.87 for the active treatments versus 7.31 for placebo at the 8-hour assessment and 15.16 to 17.39 for the active treatments versus 11.73 for placebo at the 12-hour assessment, respectively. The mean SPID scores of the active treatments were significantly higher than those for placebo for both the 8-hour and 12-hour post first dose assessments. The mean scores of the three active treatments were not statistically significantly different. Detailed results for SPID is presented in Table 23.

TOTPAR: The mean TOTPAR scores of the three active treatments were significantly higher than the score seen for placebo at both 8- and 12-hour post first dose assessments. The three active treatments were not statistically significantly different at the 8- and 12-hour hour assessments. Detailed results for TOTPAR is presented in Table 23 below.

Table 23. Primary Efficacy Parameters (8 and 12 hours post first dose of study medication)

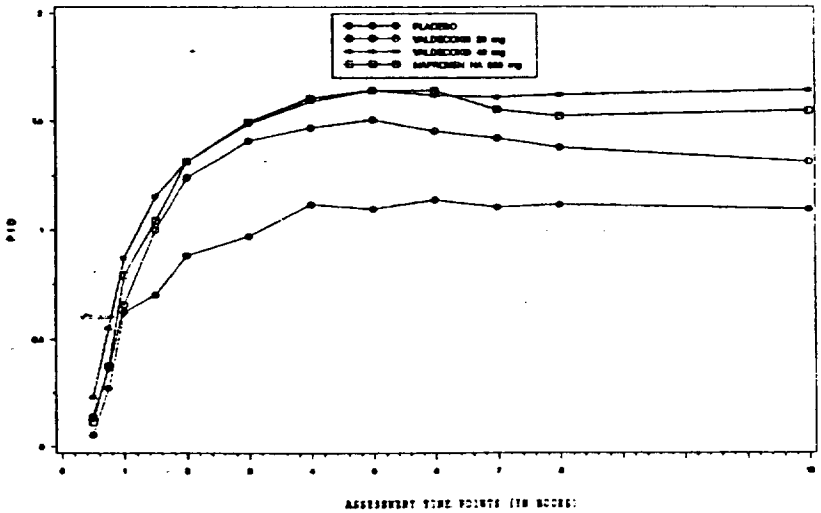
Parameter	Placebo	Valdecoxib 20 mg	Valdecoxib 40 mg	Naproxen Sodium
Sum of Pain Relief (SPID)				
At 8 hours	7.31 (B)	9.77 (A)	10.87 (A)	10.64 (A)
At 12 hours	11.73 (C)	15.16 (B)	17.39 (A)	16.78 (AB)
Total Pain Relief (TOTPAR)				
At 8 hours	15.05 (B)	18.89 (A)	20.80 (A)	20.55 (A)
At 12 hours	23.78 (B)	29.35 (A)	32.90 (A)	32.29 (A)

a: treatments that have the same letter (A or B) were not significantly different in the distribution of the parameter based on the Fisher's Protected Least Significant Difference (LSD) comparisons

Secondary Endpoints

PID: The mean PID scores were numerically in favor of active treatments vs. placebo at all time points. The differences between the valdecoxib 40 mg group and placebo group were statistically significant at all time points except Hour 1. The differences between the valdecoxib 20 mg and placebo groups were statistically significant at all time points except Hours 0.5, 0.75, 1 and 12. The differences between the naproxen 500 mg and placebo groups were statistically significant at all time points except Hours 0.5, 0.75, and 1. There is no clear separation in PID between the active treatment groups. The numerical results for mean PID scores are presented in Figure 13.

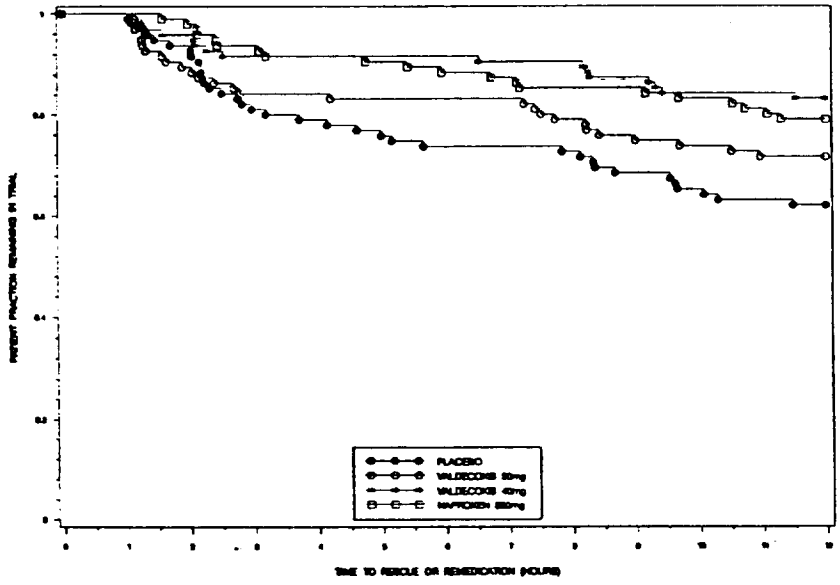
Figure 13. Mean Pain Intensity Difference



PR: The mean PR scores followed the same pattern as that in PID scores.

Time to Rescue Medication or Remedication: The percentages of patients who took rescue medication or remedicated over the median 12-hour time period were 38%, 28%, 17% and 21% in placebo, valdecoxib 20 mg, valdecoxib 40 mg and naproxen treatment groups, respectively. The differences in distribution of time to rescue medication or remedication between the valdecoxib groups and placebo was statistically significant, and so is for naproxen vs. placebo. The differences in distribution of time to rescue medication or remedication between valdecoxib 20 mg and placebo was not statistically significant. The median time to rescue medication or remedication was larger than 12 hours for all treatment groups. The Kaplan-Meier estimators for distribution of time to rescue medication or remedication for the treatment groups are presented in Figure 14 below.

Figure 14. Kaplan-Meier Estimators for Distribution of Time to Rescue Medication or Remedication



III. 2 Study 066

III.2.i Protocol

The protocol of Study 066 is identical to that of Study 065.

III.2.ii Sponsor's Main Study Results

a) Patient Disposition

Of the 120 randomized patients, 19 withdrew prior to being dosed with any study medication. An additional 14 patients (5 randomized to treatment sequence 1, 3 randomized to treatment sequence 2, 2 randomized to treatment sequence 3, and 4 randomized to treatment sequence 4) were withdrawn from the study during various treatment cycles, resulting in 87 patients who completed all four cycles. Two patients (0080 and 0097) who had been marked as having completed the study, were subsequently considered unevaluable: patient 0080 had no times associated with the recorded pain assessments for cycle 2 and patient 0097 had used rescue medication three minutes before one hour had elapsed post-study dose. Therefore, 85 patients were evaluable for efficacy. The detailed patient disposition is presented in Table 24 below.

Table 24. Patient Disposition in Each Cycle

TREATMENT CYCLE	TREATMENT SEQUENCE 1 (N = 30)	TREATMENT SEQUENCE 2 (N = 30)	TREATMENT SEQUENCE 3 (N = 30)	TREATMENT SEQUENCE 4 (N = 30)	TOTAL (N =120)
ENTERED AT CYCLE 1	27	26	22	26	101
WITHDRAWN FROM STUDY	3	2	0	3	8
REASONS FOR WITHDRAWAL (a)					
LOST TO FOLLOW-UP	0	0	0	0	0
PRE-EXISTING VIOLATION	0	0	0	0	0
PROTOCOL NONCOMPLIANCE	1	2	0	1	4
ADVERSE SIGN	0	0	0	1	1
TWO CONSECUTIVE INELIGIBLE CYCLES	2	0	0	1	3
TOOK RESCUE MEDICATION WITHIN ONE HOUR	0	0	0	0	0
ENTERED AT CYCLE 2	24	24	22	23	93
WITHDRAWN FROM STUDY	1	1	1	1	4
REASONS FOR WITHDRAWAL (a)					
LOST TO FOLLOW-UP	0	1	0	0	1
PRE-EXISTING VIOLATION	0	0	0	0	0
PROTOCOL NONCOMPLIANCE	1	0	0	0	1
ADVERSE SIGN	0	0	0	0	0
TWO CONSECUTIVE INELIGIBLE CYCLES	0	0	1	1	2
TOOK RESCUE MEDICATION WITHIN ONE HOUR	0	0	0	0	0
ENTERED AT CYCLE 3	23	23	21	22	89
WITHDRAWN FROM STUDY	1	0	1	0	2
REASONS FOR WITHDRAWAL (a)					
LOST TO FOLLOW-UP	0	0	0	0	0
PRE-EXISTING VIOLATION	0	0	0	0	0
PROTOCOL NONCOMPLIANCE	0	0	0	0	0
ADVERSE SIGN	0	0	0	0	0
TWO CONSECUTIVE INELIGIBLE CYCLES	1	0	1	0	2
TOOK RESCUE MEDICATION WITHIN ONE HOUR	0	0	0	0	0
ENTERED AT CYCLE 4	22	23	20	22	87
WITHDRAWN FROM STUDY	0	0	0	0	0
REASONS FOR WITHDRAWAL (a)					
LOST TO FOLLOW-UP	0	0	0	0	0
PRE-EXISTING VIOLATION	0	0	0	0	0
PROTOCOL NONCOMPLIANCE	0	0	0	0	0
ADVERSE SIGN	0	0	0	0	0
TWO CONSECUTIVE INELIGIBLE CYCLES	0	0	0	0	0
TOOK RESCUE MEDICATION WITHIN ONE HOUR	0	0	0	0	0
COMPLETED STUDY (4 CYCLES)	22	23	20	22	87

Note: Treatment Sequence 1 = A B C D, 2 = B D A C, 3 = C A D B, 4 = D C B A.
Where A = PLACEBO, B = VALDECOXIB 20 MG, C = VALDECOXIB 40 MG, D = NAPROXEN NA 550 MG.
(a) Mutually exclusive and exhaustive categories.

b) Demographics

Demographics and baseline characteristics were generally comparable across treatment groups.

c) Efficacy Evaluation

The 'statistical significance' in the following description was claimed based on Fisher's LSD test for all active treatment groups vs. placebo. Please see the reviewer's comment #3 for the adjustment methods of multiplicity in Section VII.

Primary Endpoints

SPID: Mean SPID scores ranged from 10.32 to 10.76 for the active treatments versus 6.41 for placebo at the 8-hour assessment and 16.14 to 16.54 for the active treatments versus 10.34 for placebo at the 12-hour assessment. The mean SPID scores of the three active treatments were significantly higher than those for placebo at both 8 and 12 hours ($p<0.001$). The mean SPID scores of the three active treatments were not statistically significantly different at both the 8-hour and 12-hour post first dose assessments. Detailed results for SPID is presented in Table 25.

TOTPAR: The mean TOTPAR scores of the three active treatments were significantly higher than the score seen for placebo at both 8-hour and 12-hour first dose assessments ($p<0.001$). The three active treatments were not statisticaly significantly different in mean scores at both assessment timepoints. At the 8-hour assessment, mean scores for the valdecoxib 20 mg, valdecoxib 40 mg, and naproxen sodium treatments were 19.64, 20.94, and 20.71, respectively. At the 12-hour assessment, mean scores for the valdecoxib 20 mg, valdecoxib 40 mg, and naproxen sodium treatments were 30.67, 32.94, and 31.89, respectively. Detailed results for TOTPAR is presented in Table 25 below.

Table 25. Primary Efficacy Parameters (8 and 12 hours post first dose of study medication)

Parameter	Placebo	Valdecoxib 20 mg	Valdecoxib 40 mg	Naproxen Sodium
Sum of Pain Intensity Difference (SPID)				
At 8 hours	6.41 (B[a])	10.32 (A)	10.36 (A)	10.76 (A)
At 12 hours	10.34 (B)	16.14 (A)	16.45 (A)	16.54 (A)
Total Pain Relief (TOTPAR)				
At 8 hours	14.07 (B)	19.64 (A)	20.94 (A)	20.71 (A)
At 12 hours	21.99 (B)	30.67 (A)	32.94 (A)	31.89 (A)

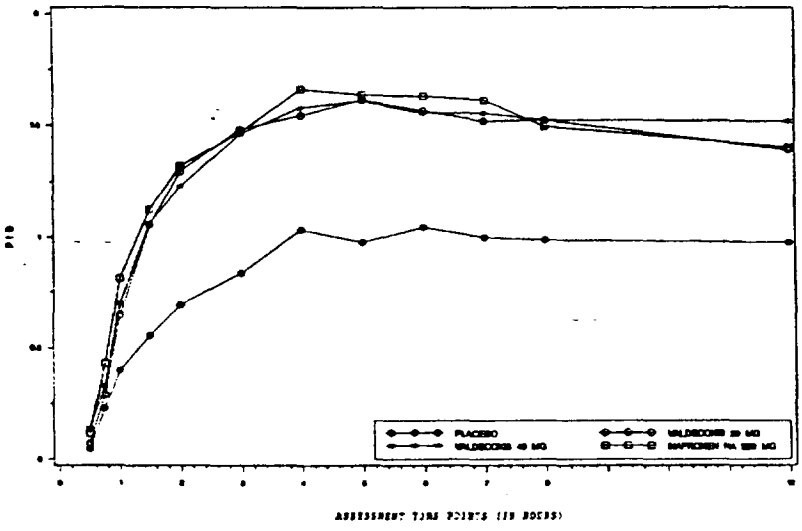
a: treatments that have the same letter (A or B) were not significantly different in the distribution of the parameter based on the Fisher's Protected Least Significant Difference (LSD) comparisons

Secondary Endpoints

PID: The mean PID scores were numerically in favor of active treatments vs. placebo at all time points. The differences between the active treatment groups vs. placebo were statistically significant at all time points except Hours 0.5 and 0.75. The differences between the three active treatment groups were not statistically significant at all time points. The results for mean PID scores are presented in Figure 15.

APPEARS THIS WAY
ON ORIGINAL

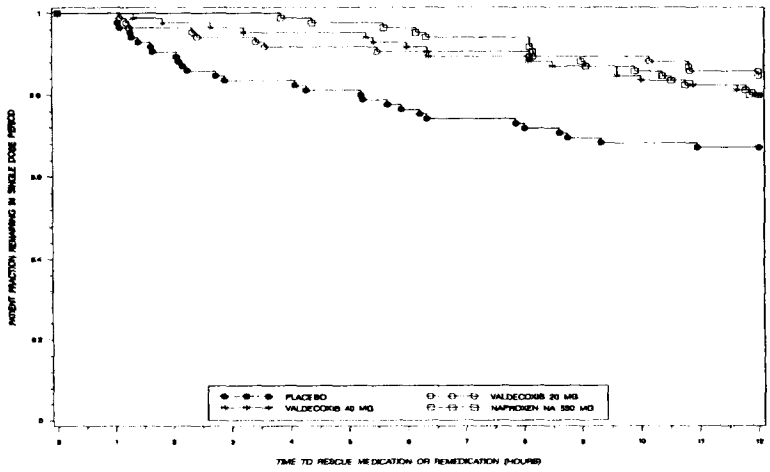
Figure 15. Mean Pain Intensity Difference



PR: The mean PR scores followed the same pattern as that in PID scores.

Time to Rescue Medication or Remedication: The percentages of patients who took rescue medication or remedicated over the median 12-hour time period were 33%, 15%, 20% and 20% in placebo, valdecoxib 20 mg, valdecoxib 40 mg and naproxen treatment groups, respectively. The difference in distribution of time to rescue medication or remedication between active treatments and placebo were statistically significant, There is no clear separation between the active treatments in distribution of time to rescue medication or remedication. The median time to rescue medication or remedication is larger than 12 hours for all treatment groups. The Kaplan-Meier estimators for distribution of time to rescue medication or remedication for the treatment groups are presented in Figure 16 below.

Figure 16. Kaplan-Meier Estimators for Distribution of Time to Rescue Medication or Remedication



IV.1 Study 038

7 page(s) have been removed because it contains trade secret and/or confidential information that is not disclosable.

40-46

V. Osteoarthritis Studies

V.1 Study 049

V.1.i Protocol

This study is designed as a multi-center, double-blind, randomized, placebo-controlled, parallel group trial to evaluate the efficacy and safety of valdecoxib 5 mg QD and 10 mg QD compared with placebo and naproxen 500 mg BID in adult patients with OA. The primary objective of this study is to determine the efficacy of valdecoxib by comparing valdecoxib 5 mg QD and 10 mg QD with placebo in treating the signs and symptoms of OA of the hip.

Patients with symptomatic OA of the hip will be randomly assigned to receive either valdecoxib 5 mg QD, valdecoxib 10 mg QD, naproxen 500 mg BID or placebo. The duration of treatment is 12 weeks, with visits performed at screening, baseline, and Weeks 2, 6 and 12. Primary efficacy endpoints include WOMAC OA Pain Index, Patient's Global Assessment of Arthritis (Index Hip), WOMAC OA Physical Function Index. Secondary efficacy endpoints include Physician's Global Assessment of Arthritis (Index Hip), WOMAC OA Composite Index, WOMAC OA Stiffness Index, Incidence of and Time to Patient Withdrawal Due to Treatment Failure, Patient's Assessment of Arthritis Pain by visual analog score.

All patients who are randomized and take at least one dose of study medication will be included in the intent-to-treat (ITT) cohort. Last-observation-carried-forward approach will be used to impute missing values. Pairwise comparisons for all four treatment groups will be performed by using analysis of covariance (ANCOVA) with treatment and center as factors, and the corresponding baseline score as covariate. Differential effects of gender, age and duration of disease will be examined by ANCOVA models. Changes

from Baseline for the categorical variables will be also analyzed by Cochran-Mantel-Haenszel (CMH) method, stratified by center. Incidence of withdrawal due to treatment failure will be analyzed by Fisher's exact test. Time to withdrawal within each treatment group will be plotted using the Kaplan-Meier product limit estimator. An overall log-rank test for the time to withdrawal due to treatment failure will be performed. In addition, pairwise comparisons will be made between the treatment groups using log-rank test. Pairwise comparisons will be carried out to compare the efficacy of treatments. The results of the pairwise comparisons for the two dose groups (5 mg and 10 mg of valdecoxib) versus placebo will be interpreted using the Hochberg's step down procedure

The sample size calculation is based on the WOMAC Pain Index Scores and the primary comparisons: each dose of valdecoxib versus placebo. A sample size of 120 patients per treatment group will be sufficient to detect differences larger than 1.41 in the mean change from Baseline in the WOMAC Pain Index Scores between treatment groups (valdecoxib 5 mg and 10 mg versus placebo) with at least 80% power and type I error at 0.025 for a two-sided test.

V.1.ii Sponsor's Main Study Results

a) Patient Disposition

A total of 467 patients at 60 study sites were randomized into this study. Of these patients, 258 (55%) completed the study, and 209 (45%) were prematurely withdrawn. The incidence of withdrawal due to treatment failure was 43% in the placebo group, 27% in the valdecoxib 5 mg QD group, 28% in the valdecoxib 10 mg QD group and 20% in the naproxen 500 mg BID group. The incidence of withdrawal due to adverse events was 6% in the placebo group, 8% in the valdecoxib 5 mg QD group, 10% in the valdecoxib 10 mg QD group and 13% in the naproxen 500 mg BID group. Table 30 shows the detailed information for patient disposition.

Table 30. Patient Disposition

	PLACEBO (N=118)	VALDECOXIB 5 MG QD (N=120)	VALDECOXIB 10 MG QD (N=111)	NAPROXEN 500 MG BID (N=118)	TOTAL (N=467)
COMPLETED STUDY	49 (42%)	73 (61%)	65 (59%)	71 (60%)	258 (55%)
WITHDRAWN	69 (58%)	47 (39%)	46 (41%)	47 (40%)	209 (45%)
REASON FOR WITHDRAWAL(a)					
TREATMENT FAILURE	51 (43%)	32 (27%)	31 (28%)	24 (20%)	138 (30%)
LOST TO FOLLOW-UP	1 (1%)	1 (1%)	0 (0%)	2 (2%)	4 (1%)
PRE-EXISTING VIOLATION	3 (3%)	0 (0%)	1 (1%)	2 (2%)	6 (1%)
NONCOMPLIANCE	7 (6%)	4 (3%)	3 (3%)	4 (3%)	18 (4%)
ADVERSE SIGN OR SYMPTOM	7 (6%)	10 (8%)	11 (10%)	15 (13%)	43 (9%)

(a) Mutually exclusive and exhaustive categories.

b) Demographics

Demographics and baseline characteristics were generally comparable across treatment groups.

c) Efficacy Evaluation

Primary Endpoints

The WOMAC OA Pain Index mean score decreased from baseline to each visit in all treatment groups. The decrease of the LS means was statistically significantly greater in the valdecoxib 5 mg QD, valdecoxib 10 mg QD and naproxen 500 mg BID groups than in the placebo group at all visits ($p<=0.019$). The decrease was not statistically significantly different between the valdecoxib 5 mg QD, valdecoxib 10 mg QD and naproxen 500 mg BID groups at all visits ($p>=0.202$). Detailed results for WOMAC OA Pain score are presented in Table 31.

Table 31. Results on WOMAC OA Pain Index

	Placebo N=117	Valdecoxib 5 mg QD N=120	Valdecoxib 10 mg QD N=111	Naproxen 500 mg BID N=118
Least Squares Mean Change				
Week 2	-0.90	-2.48*	-2.56*	-3.07*
Week 6	-1.09	-2.76*	-3.23*	-3.14*
Week 12	-1.25	-2.54*	-2.83*	-2.94*

*: significantly different from placebo group

The Patient's Global Assessment of Arthritis mean score decreased from Baseline to each visit in all treatment groups. The decrease of the LS means was statistically significantly greater in the valdecoxib 5 mg QD, valdecoxib 10 mg QD and naproxen 500 mg BID groups than in the placebo group at all visits ($p<=0.038$). The decrease was not statistically significantly different between the valdecoxib 5 mg QD, valdecoxib 10 mg QD and naproxen 500 mg BID groups at all visits ($p>=0.086$). Detailed results for Patient's Global Assessment of Arthritis are presented in Table 32.

Table 32. Results on Patient's Global Assessment of Arthritis

	Placebo N=117	Valdecoxib 5 mg QD N=120	Valdecoxib 10 mg QD N=111	Naproxen 500 mg BID N=118
Least Squares Mean Change				
Week 2	-0.90	-2.48*	-2.56*	-3.07*
Week 6	-1.09	-2.76*	-3.23*	-3.14*
Week 12	-1.25	-2.54*	-2.83*	-2.94*

*: significantly different from placebo group

The WOMAC OA Physical Function Index mean score decreased from Baseline to each visit in all treatment groups. The decrease of the LS means was statistically significantly greater in the valdecoxib 5 mg QD, valdecoxib 10 mg QD and naproxen 500 mg BID groups than in the placebo group at all visits ($p<=0.004$). The decrease was not

statistically significantly different between the valdecoxib 5 mg QD, valdecoxib 10 mg QD and naproxen 500 mg BID groups at all visits (p>= 0.084). Detailed results for WOMAC OA Physical Function Index score are presented in Table 33 below.

Table 33. Results on WOMAC OA Physical Function Index

	Placebo N=117	Valdecoxib 5 mg QD N=120	Valdecoxib 10 mg QD N=111	Naproxen 500 mg BID N=118
Least Squares Mean Change				
Week 2	-0.72	-1.10*	-1.26*	-1.31*
Week 6	-0.82	-1.11*	-1.29*	-1.30*
Week 12	-0.87	-1.20*	-1.29*	-1.18*

*: significantly different from placebo group

Secondary Endpoints

Secondary endpoints mirrored the results of the primary endpoints. Active treatment groups showed more improvement than placebo in Physician's Global Assessment of Arthritis (Index Hip), WOMAC OA Composite Index, WOMAC OA Stiffness Index, Patient's Assessment of Arthritis Pain by visual analog score. Active treatment groups also showed lower rates in Patient Withdrawal Due to Treatment Failure than the placebo group. The results for secondary endpoints are presented in Table 34 below.

Table 34. Results for Secondary Endpoints

	Placebo (N=117)	Valdecoxib 5 mg QD (N=120)	Valdecoxib 10 mg QD (N=111)	Naproxen 500 mg BID (N=118)
Physician's Global Assessment of Arthritis				
Baseline Mean	4.1	4.1	4.1	4.1
Least Squares Mean Change				
Week 2	-0.72	-1.10*	-1.22*	-1.32*
Week 6	-0.84	-1.17*	-1.25*	-1.28*
Week 12	-0.88	-1.18*	-1.25*	-1.23*
WOMAC OA Composite Index				
Baseline Mean	52.5	54.7	52.8	51.8
Least Squares Mean Change				
Week 2	-4.31	-10.8*	-12.6*	-14.3*
Week 6	-5.07	-12.3*	-14.7*	-14.7*
Week 12	-5.28	-12.0*	-14.0*	-13.8*
WOMAC OA Joint Stiffness Index				
Baseline Mean	4.6	5.0	4.8	4.7
Least Squares Mean Change				
Week 2	-0.34	-0.82*	-1.14*	-1.25*
Week 6	-0.57	-1.07*	-1.24*	-1.22*
Week 12	-0.60	-1.09*	-1.19*	-1.21*
Patient's Assessment of Arthritis Pain-VAS				
Baseline Mean	71.2	72.3	73.4	69.0
Least Squares Mean Change				
Week 2	-14.4	-21.0*	-24.6*	-27.3*
Week 6	-16.0	-23.3*	-25.8*	-26.1*

Week 12	-15.2	-21.3	-23.2	-22.0
Incidence of Withdrawal due to Treatment Failure				
	43.6%	26.7%	27.9%	20.3%

*: p-value less than 0.05 against placebo

V.2 Study 053

V.2.i Protocol

This study is a multi-center, double-blind, randomized, placebo-controlled, parallel group trial to evaluate the efficacy and safety of valdecoxib 5 mg, 10 mg and 20 mg QD compared with placebo and naproxen 500 mg BID in adult patients with OA. The primary objective of this study is to determine the efficacy of valdecoxib by comparing valdecoxib 5 mg, 10 mg and 20 mg QD with placebo in treating the signs and symptoms of OA of the knee.

This study has similar design with 049 except that
1. Valdecoxib 20 mg group is included.
2. The patients must have knee OA.
3. Secondary endpoints also include Question 2 (How much pain are you having right now?) of American Pain Society (APS) pain measure.

V.2.ii Sponsor's Main Study Results

a) Patient Disposition

A total of 1019 patients with OA of the knee were randomized into this study at 85 sites. Of these patients, 750 (74%) completed the study, and 269 (26%) were prematurely withdrawn. The incidence of withdrawal was 36% in the placebo group, 19% in the valdecoxib 5 mg QD group, 27% in the valdecoxib 10 mg QD group, 22% in the valdecoxib 20 mg QD group, and 27% in the naproxen 500 mg BID group. Table 35 shows the detailed information for patient disposition.

Table 35. Patient Disposition

	PLACEBO (N = 205)	VALDECOXIB 5 MG QD (N = 201)	VALDECOXIB 10 MG QD (N = 206)	VALDECOXIB 20 MG QD (N = 202)	NAPROXEN 500 MG BID (N = 205)	TOTAL (N = 1019)
DURATION IN STUDY						
COMPLETED STUDY	131 (64%)	162 (81%)	150 (73%)	158 (78%)	149 (73%)	750 (74%)
WITHDRAWN	74 (36%)	39 (19%)	56 (27%)	44 (22%)	56 (27%)	269 (26%)
REASON FOR WITHDRAWAL (a)						
LOST TO FOLLOW-UP	4 (2%)	2 (1%)	0 (0%)	2 (1%)	1 (0%)	9 (1%)
PRE-EXISTING VIOLATION	2 (1%)	3 (1%)	5 (2%)	3 (1%)	4 (2%)	17 (2%)
PROTOCOL NONCOMPLIANCE	9 (4%)	6 (3%)	9 (4%)	8 (4%)	12 (6%)	44 (4%)
TREATMENT FAILURE	42 (20%)	16 (8%)	24 (12%)	20 (10%)	13 (6%)	115 (11%)
ADVERSE SIGN OR SYMPTOM	17 (8%)	12 (6%)	18 (9%)	11 (5%)	26 (13%)	84 (8%)

(a) Mutually exclusive and exhaustive categories.

b) Demographics

Except for H. pylori status, demographics and baseline characteristics were generally comparable across treatment groups. H. pylori status was not considered as an influential covariat for efficacy outcomes.

c) Efficacy Evaluation

Primary Endpoints

The WOMAC OA Pain Index mean score decreased from baseline to Weeks 2, 6, and 12 in all treatment groups. Although the changes were numerically greater in all active treatment groups than in the placebo group at all visits, statistical significance was not reached at Week 12 for all active treatment groups vs. placebo by Hochburg adjustment procedure for multiple treatment group comparison (p=0.07 for valdecoxib 5 mg vs. placebo; p=0.14 for valdecoxib 10 mg vs. placebo; p=0.02 for valdecoxib 20 mg vs. placebo) at Week 12. Detailed numerical results for WOMAC OA Pain score are presented in Table 36 below.

Table 36. Results on WOMAC OA Pain Index

	Placebo N=205	Valdecoxib 5 mg QD N=201	Valdecoxib 10 mg QD N=205	Valdecoxib 20 mg QD N=201	Naproxen 500 mg BID N=204
Least Squares Mean Change					
Week 2	-2.21	-2.93*	-3.43*	-3.48*	-3.30**
Week 6	-2.73	-3.44	-3.49	-3.84*	-3.68*
Week 12	-2.99	-3.73	-3.60	-3.92	-3.92

*: significantly different from placebo group

The Patient's Global Assessment of Arthritis mean score decreased from baseline to Weeks 2, 6, and 12 in all treatment groups, and the changes were numerically greater in all active treatment groups than in the placebo group at all visits. Compared to the placebo group, the improvement was significantly greater in the valdecoxib 10 mg and 20 mg groups by Hochburg adjustment procedure for multiple treatment group comparison (p=0.008 for valdecoxib 10 mg vs. placebo; p=0.004 for valdecoxib 20 mg vs. placebo) at Week 12. The improvements in valdecoxib 5 mg group and naproxen group were not statistically significantly different from placebo (p=0.14 for valdecoxib 5 mg vs. placebo; p=0.11 for naproxen vs. placebo) at Week 12. Detailed numerical results for WOMAC OA Pain score are presented in Table 37.

Table 37. Results on Patient's Global Assessment of Arthritis

	Placebo N=205	Valdecoxib 5 mg QD N=201	Valdecoxib 10 mg QD N=205	Valdecoxib 20 mg QD N=201	Naproxen 500 mg BID N=204
Least Squares Mean Change					
Week 2	-1.18	-1.33	-1.36	-1.48*	-1.39*
Week 6	-1.23	-1.41	-1.44	-1.46*	-1.44*
Week 12	-1.24	-1.40	-1.53*	-1.55*	-1.41

*: significantly different from placebo group

The WOMAC OA Physical Function Index mean score decreased from baseline to Weeks 2, 6, and 12 in all treatment groups, and the changes were numerically greater in all active treatment groups than in the placebo group at all visits. The improvements of valdecoxib groups were not statistically significant by Hochburg adjustment procedure for multiple treatment group comparison (p=0.076 for valdecoxib 5 mg vs. placebo; p=0.025 for valdecoxib 10 mg vs. placebo; p=0.041 for valdecoxib 20 mg vs. placebo) at Week 12. Detailed numerical results for WOMAC OA Pain score are presented in Table 38.

Table 38. Results on WOMAC OA Physical Function Index

	Placebo N=205	Valdecoxib 5 mg QD N=201	Valdecoxib 10 mg QD N=205	Valdecoxib 20 mg QD N=201	Naproxen 500 mg BID N=204
Least Squares Mean Change					
Week 2	-7.19	-9.35*	-10.54*	-10.76*	-10.93*
Week 6	-9.24	-10.75	-11.73	-12.12*	-11.90*
Week 12	-9.40	-11.70	-12.29	-12.05	12.57

*: significantly different from placebo group

Secondary Endpoints

Secondary endpoints mirrored the results of the primary endpoints. Active treatment groups showed more improvement than placebo in Physician's Global Assessment of Arthritis (Index Hip), WOMAC OA Composite Index, WOMAC OA Stiffness Index, Patient's Assessment of Arthritis Pain by visual analog score. Active treatment groups also showed lower rates in Patient Withdrawal Due to Treatment Failure than the placebo group. The results for secondary endpoints are presented in Table 39 below.

Table 39. Baseline Means and Changes from Baseline in Secondary Measures of Arthritis Efficacy

	Placebo N=205	Valdecoxib 5 mg QD N=201	Valdecoxib 10 mg QD N=205	Valdecoxib 20 mg QD N=201	Naproxen 500 mg BID N=204
Physician's Global Assessment of Arthritis					
Baseline Mean	4.10	4.07	4.09	4.09	4.10
Least Squares Mean Change					
Week 2	-1.04	-1.31*	-1.37*	-1.42*	-1.35*
Week 6	-1.22	-1.44*	-1.50*	-1.41*	-1.45*
Week 12	-1.22	-1.43*	-1.52*	-1.45*	-1.43*
WOMAC OA Composite Index					
Baseline Mean	53.49	53.03	54.73	53.42	53.67
Least Squares Mean Change					
Week 2	-10.13	-13.26*	-15.05*	-15.44*	-15.47*
Week 6	-12.98	-15.47	-16.74*	-17.33*	-16.99*
Week 12	-13.48	-16.84	-17.34*	-17.22*	-18.04*
WOMAC OA Stiffness Index					
Baseline Mean	4.84	4.87	4.91	4.73	4.94
Least Squares Mean Change					
Week 2	-0.78	-1.03	-1.20*	-1.24*	-1.28*
Week 6	-1.04	-1.25	-1.42*	-1.43*	-1.40*
Week 12	-1.12	-1.33	-1.41	-1.46*	-1.54*
Incidence of Withdrawal due to Treatment Failure					
	42%	16%*	24% *	20% *	13%*
Patient's Assessment of Arthritis Pain-VAS					
Baseline Mean	71.20	71.42	72.41	72.54	72.36
Least Squares Mean Change					
Week 2	-21.19	-28.46*	-30.21*	-32.07*	-31.03*
Week 6	-23.92	-30.81*	-29.85*	-32.28	-31.84*
Week 12	-25.97	-31.33	-30.41	-32.70*	-31.83*

*: p-value less than 0.05 against placebo

IV. Rheumatoid Arthritis Studies

IV.1 Study 060

IV.1.i Protocol

This study is a multi-center, double-blind, randomized, placebo-controlled, parallel group trial to evaluate the efficacy and safety of valdecoxib 10 mg, 20 mg and 40 mg QD compared with placebo and naproxen 500 mg BID in adult patients with RA. The primary objective of this study is to determine the efficacy of valdecoxib by comparing valdecoxib 10 mg QD, 20 mg and 40 mg QD with placebo in treating the signs and symptoms of RA.

Patients with adult-onset RA, in a flare state after discontinuing NSAIDs and analgesics, with a Functional Capacity Classification of I-III, will receive either valdecoxib 10 mg QD, valdecoxib 20 mg QD, valdecoxib 40 mg QD, naproxen 500 mg BID or placebo. The duration of treatment was 12 weeks, with visits performed at Screening, Baseline, and Weeks 2, 6 and 12. The primary measures of arthritis efficacy are: 1) ACR-20 Response; 2) Physician's Assessment of Tender/Painful Joint Count; 3) Physician's

Assessment of Swollen Joint Count; 4) Patient's Global Assessment of Disease Activity; 5) Physician's Global Assessment of Disease Activity. An ACR responder is defined as a patient with at least 20% improvement from Baseline in the number of tender/painful joints and in the number of swollen joints as well as at least 20% improvement from Baseline in at least three of the following assessments: 1) Physician's Global; 2) Patient's Global; 3) Patient's Assessment of Pain; 4) C-reactive protein; and 5) mHAQ. The secondary measures of arthritis efficacy are: 1) Tender/Painful Joint score; 2) Swollen Joint score; 3) Patient's Assessment of Arthritis Pain; 4) Patient's Assessment of Physical Function (mHAQ); 5) Acute-phase reactant value (CRP) (EIA method); 6) Duration of morning stiffness; 7) Incidence and time of withdrawal due to treatment failure. Exploratory endpoints will include Response to ACR-50 and ACR-70 Criteria, Severity of Dyspepsia Assessment (SODA) and Patient Satisfaction Questionnaire.

All efficacy analyses will be performed on patients who are randomized and take at least one dose of study medication (Intent-to-Treat Cohort). The efficacy measurements that are missing at Weeks 2, 6, and 12 will be imputed by carrying forward the last efficacy measurement. For ACR-20 response at Weeks 2, 6 and 12, pairwise comparisons will be carried out using the Cochran-Mantel-Haenszel (CMH) test stratified for center. For the mean changes of the efficacy measures from Baseline to each visit, overall comparisons across treatment groups and pairwise comparisons will be carried out by analysis of covariance (ANCOVA), with center and treatment group as factors and Baseline as the covariate. Pairwise comparisons of the categorical changes from Baseline will be carried out by the CMH method, stratified for center. For changes in the Physician's Global Assessment of Disease Activity and Patient's Global Assessment of Disease Activity, categorized as improved, unchanged, or worsened, pairwise comparisons will also be carried out by the CMH method, adjusted for the effects of center. The results of the pairwise comparisons for the two valdecoxib dose groups (20 mg QD and 40 mg QD) versus placebo will be interpreted using Hochberg's step down procedure.

For incidence of withdrawal due to lack of efficacy, overall comparison across treatment groups and pairwise comparisons will be performed by Fisher's exact test. An overall log-rank test on the time to withdrawal was performed. In addition, pairwise comparisons will be made between the treatment groups using log-rank test. The median time to withdrawal for each treatment group will be calculated using the Kaplan-Meier product limit estimator. For ACR-50 and ACR-70 responses, the same analyses will be performed as for ACR-20.

The sample size for this study is based on the expected percent of responders to ACR-20 criteria. It is anticipated that 20% of placebo patients and 35% of patients assigned to receive active treatment would show response. A sample size of 200 per treatment group is sufficient to detect the above difference with α=0.017 and a power of 80%.

IV.2.ii Sponsor's Main Study Results

Although the primary objective of this study was to determine the efficacy of valdecoxib by comparing valdecoxib 10 mg QD, 20 mg and 40 mg QD with placebo in treating the

signs and symptoms of RA, the sponsor only considered 'valdecoxib 20 mg QD vs. placebo' and 'valdecoxib 40 mg QD vs. placebo' as primary comparisons when adjusting for multiple comparisons with Hochburg procedure. When 'valdecoxib 10 mg QD vs. placebo' is included as a primary comparison, the conclusion in terms of statistical significance for the two comparisons 'valdecoxib 20 mg QD vs. placebo' and 'valdecoxib 40 mg QD vs. placebo' are the same. The results for the comparison 'valdecoxib 10 mg QD vs. placebo' were numerically and statistically consistent with that for the comparisons 'valdecoxib 20 mg QD vs. placebo' and 'valdecoxib 40 mg QD vs. placebo'.

a) Patient Disposition

A total of 1090 evaluable patients were enrolled in the study and randomized to receive either valdecoxib 10 mg QD (N=209), valdecoxib 20 mg QD (N=212), valdecoxib 40 mg QD (N=221), naproxen 500 mg BID (N=226) or placebo (N=222) for 12. A total of 466 (42.87%) patients were withdrawn prior to completion of the study including one patient who was randomized to the naproxen 500 mg BID group and withdrew his consent prior to receiving his first dose of study drug. There were 312 (28.7%) patients withdrawn from the study due to treatment failure: 49 (23%) patients in the valdecoxib 10 mg QD, 48 (23%) in the valdecoxib 20 mg QD, 56 (25%) in the valdecoxib 40 mg QD, 57 (25%) in the naproxen 500 mg BID, and 103 (46%) in the placebo treatment group. A total of 66 (6.1%) patients were withdrawn due to adverse events: 11 (5%) patients in the valdecoxib 10 mg QD, 13 (6%) patients in the valdecoxib 20 mg QD, 19 (9%) patients in the valdecoxib 40 mg QD, 13 (6%) patients in the naproxen 500 mg BID, and 10 (5%) patients in the placebo treatment group. Seven (0.64%) patients were lost to follow-up, 43 (4.0%) were withdrawn for pre-existing protocol violations, and 38 (3.5%) patients were withdrawn for protocol non-compliance. Table 40 presents the detailed information for patient disposition.

Table 40. Patient Disposition

	PLACEBO (N=222)	VALDECOXIB 10 MG QD (N=209)	VALDECOXIB 20 MG QD (N=212)	VALDECOXIB 40 MG QD (N=221)	NAPROXEN 500 MG BID (N=226)
COMPLETED STUDY	92(41%)	132(63%)	132(62%)	131(59%)	137(61%)
WITHDRAWN	130(59%)	77(37%)	80(38%)	90(41%)	89(39%)
REASON FOR WITHDRAWAL (a)					
TREATMENT FAILURE	102(46%)	49(23%)	48(23%)	56(25%)	57(25%)
LOST TO FOLLOW-UP	2(<1%)	2(<1%)	2(<1%)	1(<1%)	0(0%)
PRE-EXISTING VIOLATION	10(5%)	7(3%)	7(3%)	11(5%)	8(4%)
PROTOCOL NON-COMPLIANCE	6(3%)	8(4%)	10(5%)	3(1%)	11(5%)
ADVERSE EVENTS	10(5%)	11(5%)	13(6%)	19(9%)	13(6%)

(a) Mutually exclusive and exhaustive categories.

b) Demographics

Demographics and baseline characteristics were generally comparable across treatment groups.

c) Efficacy Evaluation

Although the primary objective of this study was to determine the efficacy of valdecoxib by comparing valdecoxib 10 mg QD, 20 mg and 40 mg QD with placebo in treating the signs and symptoms of RA, the sponsor only considered 'valdecoxib 20 mg QD vs. placebo' and 'valdecoxib 40 mg QD vs. placebo' as primary comparisons when adjusting for multiple comparisons with Hochburg procedure. When 'valdecoxib 10 mg QD vs. placebo' is included as a primary comparison, the conclusion in terms of statistical significance for the two comparisons 'valdecoxib 20 mg QD vs. placebo' and 'valdecoxib 40 mg QD vs. placebo' are the same with Hochburg procedure for three comparisons. The results for the comparison 'valdecoxib 10 mg QD vs. placebo' were numerically and statistically consistent with that for the comparisons 'valdecoxib 20 mg QD vs. placebo' and 'valdecoxib 40 mg QD vs. placebo'.

Primary Endpoints

The **ACR-20 response rates** at Week 12 were 32%, 49%, 48%, 46%, and 44% in the placebo group, valdecoxib 10 mg QD, valdecoxib 20 mg QD, valdecoxib 40 mg QD and naproxen 500 mg BID treatment groups, respectively. Response to the study medication was significantly greater in all active treatment groups compared to the placebo group ($p<=0.003$). No statistically significant difference was found between active treatment groups in ACR-20 response rates. Detailed results for ACR-20 response rate is presented in Table 41.

Table 41. ACR 20 Response Rate (%)

	Placebo N=222	Valdecoxib 10 mg QD N=209	Valdecoxib 20 mg QD N=212	Valdecoxib 40 mg QD N=221	Naproxen 500 mg BID N=225
Week 2	30	47*	46#	51#	51*
Week 6	36	51*	49#	52#	50*
Week 12	32	49*	48#	46#	44*

statistically significant according to the Hochberg procedure.
* $p\leq0.05$ vs. placebo.

The **Physician's Assessment of Tender/Painful Joint Count** mean score decreased from Baseline to Weeks 2, 6, and 12 in all treatment groups, and greater changes were observed with all active treatment groups than in the placebo group at all visits. Decrease in the number of tender/painful joints was significantly greater in the valdecoxib 10 mg QD, 20 mg QD and 40 mg QD treatment groups versus the placebo group at Week 12 ($p<=0.002$). The naproxen 500 mg BID group also showed significant improvement compared with the placebo group ($p<0.001$). There was no statistically significant difference between the valdecoxib 10 mg QD, valdecoxib 20 mg QD, valdecoxib 40 mg QD and naproxen 500 mg BID groups ($p>=0.1$). Detailed results for Physician's Assessment of Tender/Painful Joint Count is presented in Table 42.

Table 42. Mean Changes from Baseline for Physician's Assessment of Tender/Painful Count

Treatment Group	Placebo	Valdecoxib 10 mg QD	Valdecoxib 20 mg QD	Valdecoxib 40 mg QD	Naproxen 500 mg BID
Baseline Mean	27.5	27.3	29.0	29.3	28.9
Least Squares Mean Change					
Week 2	-7.9	-10.7*	-10.5#	-12.1#	-12.3*
Week 6	-8.5	-12.1*	-11.5#	-13.3#	-12.3*
Week 12	-8.1	-11.7*	-11.2#	-12.6#	-11.9*

statistically significant according to the Hochberg procedure.
* p≤0.05 vs. placebo.

The **Physician's Assessment of Swollen Joint Count** mean score decreased from Baseline to Weeks 2, 6, and 12 in all treatment groups, with greater changes observed in the active treatment groups than in the placebo group at all visits. The decrease in the number of swollen joints was not significantly greater in the valdecoxib 20 mg and 40 mg groups than placebo according to Hochburg procedure at Week 12. There was no statistical significant difference between valdecoxib 10 mg QD, valdecoxib 20 mg QD, valdecoxib 40 mg QD and naproxen 500 mg BID groups (p>=0.17). Detailed results for Physician's Assessment of Swollen Joint Count is presented in Table 43.

Table 43. Mean Changes from Baseline for Physician's Assessment of Swollen Joint Count

Treatment Group	Placebo	Valdecoxib 10 mg QD	Valdecoxib 20 mg QD	Valdecoxib 40 mg QD	Naproxen 500 mg BID
Baseline Mean	20.1	20.8	20.2	20.5	21.3
Least Squares Mean Change					
Week 2	-5.8	-7.2	-7.0	-8.0#	-7.6*
Week 6	-6.2	-7.6	-7.3	-7.9*	-7.8*
Week 12	-5.5	-7.9*	-7.0	-7.3*	-7.7*

statistically significant according to the Hochberg procedure.
* p≤0.05 vs. placebo.

The **Patient's Global Assessment of Disease Activity** mean score decreased from Baseline to Weeks 2, 6, and 12 in all treatment groups, with greater changes observed with all active treatment groups versus the placebo group at all visits. When the Patient's Global Assessment of Disease Activity was analyzed as continuous data, the valdecoxib 10 mg QD, 20 mg QD, and 40 mg QD treatment groups showed significant improvement versus the placebo group at Week 12 (p<0.001). There was no statistical significant difference between valdecoxib 10 mg QD, valdecoxib 20 mg QD, valdecoxib 40 mg QD and naproxen 500 mg BID groups (p>=0.281). Detailed results for Patient's Global Assessment of Disease Activity is presented in Table 44.

Table 44. Mean Changes from Baseline for Patient's Global Assessment of Disease Activity

Treatment Group	Placebo	Valdecoxib 10 mg QD	Valdecoxib 20 mg QD	Valdecoxib 40 mg QD	Naproxen 500 mg BID
Baseline Mean	3.7	3.7	3.8	3.8	3.7
Least Squares Mean Change					
Week 2	-0.6	-1.1*	-0.9#	-1.1#	-1.1*
Week 6	-0.5	-1.0*	-1.0#	-1.1#	-1.0*
Week 12	-0.5	-1.0*	-0.9#	-0.9#	-1.0*

statistically significant according to the Hochberg procedure.
* p≤0.05 vs. placebo.

The **Physician's Global Assessment of Disease Activity** mean score decreased from Baseline to Weeks 2, 6, and 12 in all treatment groups, with greater changes observed with all active treatment groups versus the placebo group at all visits. When the Physician's Global Assessment of Disease Activity was analyzed as continuous data, the valdecoxib 10 mg QD, 20 mg QD, and 40 mg QD treatment groups showed significant improvement versus the placebo group at Week 12 ($p<0.001$). There was no statistical significant difference between valdecoxib 10 mg QD, valdecoxib 20 mg QD, valdecoxib 40 mg QD and naproxen 500 mg BID groups ($p>=0.403$). Detailed results for Physician's Global Assessment of Disease Activity is presented in Table 45.

Table 45. Mean Changes from Baseline for Physician's Global Assessment of Disease Activity

Treatment Group	Placebo	Valdecoxib 10 mg QD	Valdecoxib 20 mg QD	Valdecoxib 40 mg QD	Naproxen 500 mg BID
Baseline Mean	3.7	3.6	3.7	3.7	3.6
Least Squares Mean Change					
Week 2	-0.6	-1.0*	-1.0#	-1.1#	-1.0*
Week 6	-0.6	-1.0*	-1.0#	-1.1#	-1.0*
Week 12	-0.5	-1.0*	-1.0#	-1.0#	-1.0*

statistically significant according to the Hochberg procedure.
* p≤0.05 vs. placebo.

Secondary and Exploratory Endpoints

The **secondary endpoints** demonstrated numerical advantage of the valdecoxib groups over placebo in all visits. Except for swollen joint score and CRP, the nominal p-values for valdecoxib groups vs. placebo were less than 0.01 in all visits for all secondary endpoints. Valdecoxib groups also had higher ACR-50 responder rates over placebo in all visits (response rates were 9%, 16%, 16%, 18%, 20% for placebo, valdecoxib 10 mg, valdecoxib 20 mg, valdecoxib 40 mg and naproxen groups, respectively, at Week 12), but did not demonstrate advantage over placebo in terms of ACR-70 responder rates (response rates were 2%, 4%, 3%, 1%, 5% for placebo, valdecoxib 10 mg, valdecoxib 20 mg, valdecoxib 40 mg and naproxen groups, respectively, at Week 12). Table 46 presents the results for secondary analyses.

Table 46. Results for Secondary Endpoints

Treatment Group	Placebo	Valdecoxib 10 mg QD	Valdecoxib 20 mg QD	Valdecoxib 40 mg QD	Naproxen 500 mg BID
Tender/Painful Joint Score					
Baseline Mean	36.2	34.5	39.0	38.8	37.6
Least Squares Mean Change					
Week 2	-11.5	-15.8**	-16.0**	-17.2***	-17.9***
Week 6	-11.7	-17.0**	-16.4**	-18.8***	-16.8**
Week 12	-11.0	-16.5**	-16.2**	-17.8***	-16.4**
Swollen Joint Score					
Baseline Mean	25.4	25.1	25.9	26.3	26.7
Least Squares Mean Change					
Week 2	-7.4	-9.9*	-9.6*	-11.0***	-10.3**
Week 6	-7.4	-10.1*	-9.3	-10.4**	-10.1*
Week 12	-6.7	-10.5**	-9.1	-9.7*	-10.1**
Patient's Assessment of Arthritis Pain (VAS)					
Baseline Mean	66.6	64.9	68.4	68.9	67.4
Least Squares Mean Change					
Week 2	-10.9	-26.1***	-21.8***	-28.8***	-26.1***
Week 6	-12.0	-25.2***	-24.4***	-29.4***	-25.6***
Week 12	-9.9	-25.1***	-22.8***	-27.6***	-25.5***
Patient's Assessment of Physical Function (mHAQ)					
Baseline Mean	1.4	1.3	1.5	1.4	1.4
Least Squares Mean Change					
Week 2	-0.1	-0.3***	-0.3***	-0.3***	-0.3***
Week 6	-0.1	-0.3***	-0.3***	-0.3***	-0.3***
Week 12	-0.1	-0.3***	-0.3***	-0.3***	-0.3***
Duration of Morning Stiffness					
Baseline Mean	259.1	269.9	269.5	330.2	268.5
Least Squares Mean Change					
Week 2	-56.1	-133.3**	-144.2***	-149.7***	-125.2**
Week 6	-27.9	-128.5***	-124.6**	-148.8***	-104.4*
Week 12	-2.2	-132.1***	96.2**	-136.1***	-100.3**
Acute-Phase Reactant Value (CRP)					
Baseline Mean	23.3	28.9	21.0	31.3	32.4
Least Squares Mean Change					
Week 2	-4.1	-1.3	-2.9	-5.1	10.7
Week 6	0.9	-5.6	2.0	2.1	4.7
Week 12	2.6	-4.5	0.2	1.4	5.2

*p≤0.05 vs. placebo, ** p≤0.01 vs. placebo, *** p≤0.001 vs. placebo.

IV.1 Study 061

IV.1.i Protocol

The protocol of Study 061 is identical to that of Study 060.

IV.2.ii Sponsor's Main Study Results

a) Patient Disposition

A total of 1093 evaluable patients were enrolled in the study and randomized to receive either valdecoxib 10 mg QD (N=226), valdecoxib 20 mg QD (N=219), valdecoxib 40 mg QD (N=209), naproxen 500 mg BID (N=219) or placebo (N=220) for 12 weeks. A total of 442 (40.4%) patients were withdrawn prior to completion of the study. There were 300 patients withdrawn from the study due to treatment failure: 61 (27%) patients in the valdecoxib 10 mg QD, 56 (26%) in the valdecoxib 20 mg QD, 48 (23%) in the valdecoxib 40 mg QD, 43 (20%) in the naproxen 500 mg BID, and 92 (42%) in the placebo treatment group. A total of 65 (5.9%) patients were withdrawn due to adverse events: 10 (4%) patients in the valdecoxib 10 mg QD, 12 (5%) patients in the valdecoxib 20 mg QD, 13 (6%) patients in the valdecoxib 40 mg QD, 21 (10%) patient in the naproxen 500 mg BID, and 9 (4%) patients in the placebo treatment group. Seven (0.64%) patients were lost to follow-up, 44 (4.0%) were withdrawn for pre-existing protocol violations, and 26 (2.4%) patients were withdrawn for protocol non-compliance. Table 47 presents the detailed information for patient disposition.

Table 47. Patient Disposition

	PLACEBO (N=220)	VALDECOXIB 10 MG QD (N=226)	VALDECOXIB 20 MG QD (N=219)	VALDECOXIB 40 MG QD (N=209)	NAPROXEN 500 MG BID (N=219)
COMPLETED STUDY	95(43%)	137(61%)	137(63%)	137(66%)	145(66%)
WITHDRAWN	125(57%)	89(39%)	82(37%)	72(34%)	74(34%)
REASON FOR WITHDRAWAL (a)					
TREATMENT FAILURE	92(42%)	61(27%)	56(26%)	48(23%)	43(20%)
LOST TO FOLLOW-UP	3(1%)	1(<1%)	2(<1%)	0(0%)	1(<1%)
PRE-EXISTING VIOLATION	14(6%)	9(4%)	9(4%)	8(4%)	4(2%)
PROTOCOL NON-COMPLIANCE	7(3%)	8(4%)	3(1%)	3(1%)	5(2%)
ADVERSE EVENTS	9(4%)	10(4%)	12(5%)	13(6%)	21(10%)

(a) Mutually exclusive and exhaustive categories.

b) Demographics

Except for age, demographics and baseline characteristics were generally comparable across treatment groups. The effect of age on efficacy outcome was explored in an ANOVA model and no confounding was found with treatment effect.

c) Efficacy Evaluation

Primary Endpoints

The **ACR-20 response rate** at Week 12 were 32%, 46%, 47%, 50%, and 53% in the placebo group, valdecoxib 10 mg QD, valdecoxib 20 mg QD, valdecoxib 40 mg QD and naproxen 500 mg BID treatment groups, respectively. Response rates were significantly greater in the valdecoxib 10 mg QD, 20 mg QD and 40 mg QD treatment groups versus the placebo group at Week 12 ($p<=0.006$). The naproxen 500 mg BID group showed significantly higher response rate than the placebo group ($p<0.001$) and numerically

higher response rate than the valdecoxib groups at all visits. Detailed results for ACR-20 response rate is presented in Table 48.

Table 48. ACR 20 Response Rate (%)

	Placebo N=220	Valdecoxib 10 mg QD N=226	Valdecoxib 20 mg QD N=219	Valdecoxib 40 mg QD N=209	Naproxen 500 mg BID N=219
Week 2	28	46*	42#	42#	57*
Week 6	31	50*	48#	53#	57*
Week 12	32	46*	47#	60#	53*

statistically significant according to the Hochberg procedure.
* p≤0.05 vs. placebo.

The **Physician's Assessment of Tender/Painful Joint Count** mean score decreased from Baseline to Weeks 2, 6, and 12 in all treatment groups, and greater changes were observed with all active treatment groups than in the placebo group at all visits. Decrease in the number of tender/painful joints was significantly greater in the valdecoxib 10 mg QD, 20 mg QD and 40 mg QD treatment groups versus the placebo group at Week 12 ($p<=0.012$). The naproxen 500 mg BID group also showed significant improvement compared with the placebo group ($p<0.001$). Naproxen 500 mg BID group also had more improvement in Tender/Painful Joint Count than valdecoxib 10 mg and 20 mg QD groups, and the nominal p-values at all visit are less or close to 0.05 ($p<=0.063$). Detailed results for Physician's Assessment of Tender/Painful Joint Count is presented in Table 49.

Table 49. Mean Changes from Baseline for Physician's Assessment of Tender/Painful Count

Treatment Group	Placebo	Valdecoxib 10 mg QD	Valdecoxib 20 mg QD	Valdecoxib 40 mg QD	Naproxen 500 mg BID
Baseline Mean	29.6	28.9	28.6	29.1	29.3
Least Squares Mean Change					
Week 2	-6.9	-10.6*	-11.3#	-13.2#	-13.3*
Week 6	-8.2	-11.7*	-12.5#	-14.7#	-14.8*
Week 12	-8.8	-12.2*	-11.8#	-14.0#	-14.4*

statistically significant according to the Hochberg procedure.
* p≤0.05 vs. placebo.

The **Physician's Assessment of Swollen Joint Count** mean score decreased from Baseline to Weeks 2, 6, and 12 in all treatment groups, with greater changes observed in the active treatment groups than in the placebo group at all visits. No statsitically significant difference were observed between the valdecoxib groups and placebo at Week 12. The nominal p-value was 0.001 for Naproxan 500 mg BID vs. placebo, and the nominal p-values were less than 0.03 for Naproxan 500 mg BID vs. valdecoxib 10 mg QD at Week 12. Detailed results for Physician's Assessment of Swollen Joint Count is presented in Table 50.

Table 50. Mean Changes from Baseline for Physician's Assessment of Swollen Joint Count

Treatment Group	Placebo	Valdecoxib 10 mg QD	Valdecoxib 20 mg QD	Valdecoxib 40 mg QD	Naproxen 500 mg BID
Baseline Mean	20.8	20.5	21.2	20.7	20.5
Least Squares Mean Change					
Week 2	-6.0	-6.7	-7.1	-7.6	-8.5*
Week 6	-6.4	-7.5	-8.3#	-8.7#	-9.4*
Week 12	-6.7	-7.5	-7.9	-8.1	-9.3*

statistically significant according to the Hochberg procedure.
* p≤0.05 vs. placebo.

The **Patient's Global Assessment of Disease Activity** mean score decreased from Baseline to Weeks 2, 6, and 12 in all treatment groups, with greater changes observed with all active treatment groups versus the placebo group at all visits. When the Patient's Global Assessment of Disease Activity was analyzed as continuous data, the valdecoxib 10 mg QD, 20 mg QD, and 40 mg QD treatment groups showed significant improvement versus the placebo group at Week 12 ($p<0.001$). Compared with valdecoxib groups, naproxen 500 mg BID had more improvements in all visits and the nominal p-values for naproxen 500 mg BID vs. valdecoxib 10 mg QD were less than 0.05 at all visits. Detailed results for Physician's Global Assessment of Disease Activity is presented in Table 51.

Table 51. Mean Changes from Baseline for Patient's Global Assessment of Disease Activity

Treatment Group	Placebo	Valdecoxib 10 mg QD	Valdecoxib 20 mg QD	Valdecoxib 40 mg QD	Naproxen 500 mg BID
Baseline Mean	3.6	3.6	3.7	3.6	3.7
Least Squares Mean Change					
Week 2	-0.5	-0.9***	-1.0#	-1.0#	-1.1***
Week 6	-0.4	-0.9***	-1.0#	-1.0#	-1.1***
Week 12	-0.5	-0.9***	-0.9#	-0.9#	-1.1***

statistically significant according to the Hochberg procedure.
* p≤0.05 vs. placebo.

The **Physician's Global Assessment of Disease Activity** mean score decreased from Baseline to Weeks 2, 6, and 12 in all treatment groups, with greater changes observed with all active treatment groups versus the placebo group at all visits. When the Physician's Global Assessment of Disease Activity was analyzed as continuous data, the valdecoxib 10 mg QD, 20 mg QD, and 40 mg QD treatment groups showed significant improvement versus the placebo group at Week 12 ($p<0.001$). Compared with valdecoxib 10 mg and 20 mg QD groups, naproxen 500 mg BID had more improvements in all visits and the nominal p-values for naproxen 500 mg BID vs. valdecoxib 10 mg QD were less than 0.05 at all visits. Detailed results for Physician's Global Assessment of Disease Activity is presented in Table 52.

Table 52. Mean Changes from Baseline for Physician's Global Assessment of Disease Activity

Treatment Group	Placebo	Valdecoxib 10 mg QD	Valdecoxib 20 mg QD	Valdecoxib 40 mg QD	Naproxen 500 mg BID
Baseline Mean	3.6	3.6	3.6	3.6	3.6
Least Squares Mean Change					
Week 2	-0.5	-0.9*	-1.0#	-1.0#	-1.1*
Week 6	-0.5	-0.9*	-1.0#	-1.0#	-1.1*
Week 12	-0.5	-0.9*	-0.9#	-1.0#	-1.0*

statistically significant according to the Hochberg procedure.
* $p \leq 0.05$ vs. placebo.

Secondary and Exploratory Endpoints

The **secondary endpoints** demonstrated numerical advantage of the valdecoxib groups over placebo in all visits. Except for CRP and swollen joint score, the nominal p-values for valdecoxib groups vs. placebo were less than 0.01 in all visits for all secondary endpoints. Valdecoxib groups also had higher ACR-50 responder rates and ACR-70 responder rates over placebo at Week 12, but no statistical significance was found. Naproxen 500 mg BID had higher ACR-50 responder rates and ACR-70 responder rates over placebo at Week 12 with nominal p-values less than 0.05 (0<0.001, 0.018 respectively). The ACR-50 response rates were 12%, 16%, 18%, 17%, 25% for placebo, valdecoxib 10 mg, valdecoxib 20 mg, valdecoxib 40 mg and naproxen groups, respectively, at Week 12. The ACR-70 response rates were <1%, 2%, 3%, 4%, 4% for placebo, valdecoxib 10 mg, valdecoxib 20 mg, valdecoxib 40 mg and naproxen groups, respectively, at Week 12. Table 53 presents the results for secondary analyses.

Table 53. Results for Secondary Endpoints

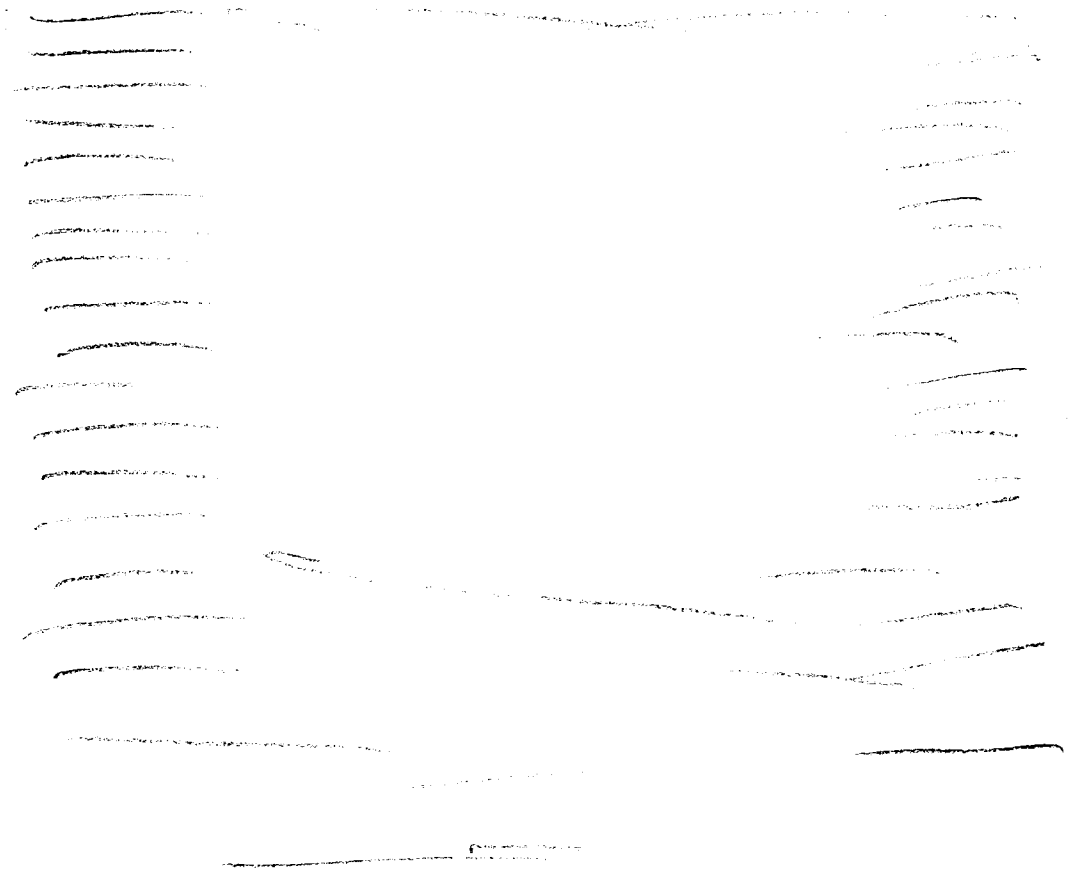
	Placebo N=205	Valdecoxib 5 mg QD N=201	Valdecoxib 10 mg QD N=205	Valdecoxib 20 mg QD N=201	Naproxen 500 mg BID N=204
Tender/Painful Joint Score					
Baseline Mean	38.8	39.6	39.9	38.8	38.5
Least Squares Mean Change					
Week 2	-11.3	-16.2***	-17.1***	-19.0***	-20.0***
Week 6	-12.0	-17.4***	-18.8***	-21.6***	-21.3***
Week 12	-12.7	-17.7**	-17.4**	-20.6***	-21.4***
Swollen Joint Score					
Baseline Mean	26.2	26.5	28.2	26.9	25.6
Least Squares Mean Change					
Week 2	-8.6	-9.5	-10.4	-10.4***	-12.0***
Week 6	-8.7	-10.2	-12.0**	-12.1**	-12.9***
Week 12	-9.2	-9.8	-11.0	-10.9	-12.7**
Patient's Assessment of Arthritis Pain (VAS)					
Baseline Mean	64.6	64.3	65.6	63.9	66.0
Least Squares Mean Change					
Week 2	-9.8	-21.7***	-24.2***	-22.9***	-28.3***
Week 6	-9.6	-20.0***	-20.5***	-23.5***	-28.2***
Week 12	-11.7	-22.1***	-20.9**	-23.8***	-27.1***
Patient's Assessment of Physical Function (mHAQ)					
Baseline Mean	1.3	1.4	1.4	1.3	1.4

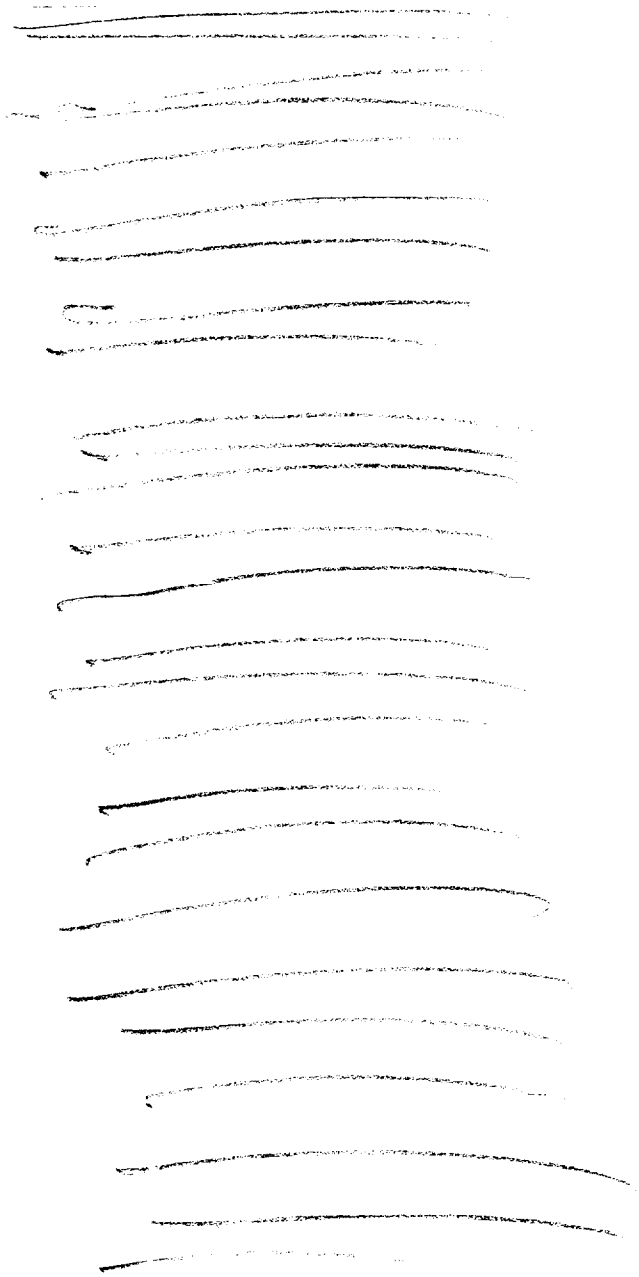
Least Squares Mean Change					
Week 2	-0.1	-0.2***	-0.3***	-0.3***	-0.4***
Week 6	-0.1	-0.3***	-0.3***	-0.3***	-0.4***
Week 12	-0.1	-0.3**	-0.3***	-0.3***	-0.4***
Duration of Morning Stiffness					
Baseline Mean	319.3	335.0	315.4	346.8	390.7
Least Squares Mean Change					
Week 2	-60.8	-173.4***	-197.9***	-206.0***	-231.1***
Week 6	-40.0	-162.0***	-214.1***	-186.9***	-223.1***
Week 12	-61.6	-163.1**	-184.7***	-174.5***	-198.7***
Acute-Phase Reactant Value (CRP)					
Baseline Mean	16.7	21.6	17.7	18.2	14.6
Least Squares Mean Change					
Week 2	-0.4	2.3	-1.2	-1.7	0.8
Week 6	2.5	1.6	-0.5	-0.7	-1.7
Week 12	-0.8	1.6	-0.1	1.7	-1.5

*p≤0.05 vs. placebo, ** p≤0.01 vs. placebo, *** p≤0.001 vs. placebo.

VII. Reviewer's Comments

VII.1. Comments for [redacted] Analgesia Studies





VII.2 Comment for Primary Dysmennorea Studies

A total of 12 (11%) and 14 (14%) of the patients in Study 65 and Study 66, respectively, were excluded from the efficacy analysis due to patients' withdrawal from at least one treatment period. In the observed cohorts (96 patients in Study 65 and 87 patients in Study 66), statsitical significant differences were found between the valdecoxib groups vs. placebo in SPID8 and SPID12 with $p \leq 0.003$. To assess the sensitivity of the results in the observed cohorts, it is of interest to ask the question: 'In order to nullify the statistical significance in the observed group, how bad the results in valdecoxib groups (vs. placebo) need to be in the drop-out cohorts?'. To answer this question, this reviewer did the following analysis:

Step 1. Assume the mean difference in SPID (SPID8 or SPID12) between each valdecoxib group and placebo in the drop-out cohort is Δ, and assume the variation of the SPID differences in the drop-out cohort is the same as that in the observed group.

Step 2. Calculate the threshold Δ values that just nullifies the statistical significance of SPID (SPID8 or SPID12) of the observed group , i.e., the Δ value at which the p-value is 0.025 (by Bonferroni-adjustment method for multiple treatment comparisons) when the observed cohort and drop-out cohort are combined for the comparison between each of the valdecoxib groups vs. placebo. The LS mean differences of SPID are used for the observed cohort based on the originally specified analysis model.

The threshold Δ values for SPID8 and SPID12 for the valdecoxib groups vs. placebo comparisons in Studies 065 and 066 are presented in Table 54 below.

Table 54. Threshold Δ Values That Nullifies Statistical Significance in the Observed Groups

		Valdecoxib 20mg vs. Placebo		Valdecoxib 40mg vs. Placebo	
		LS Mean Difference* (Observed Cohort)	**Threshold Δ Values (Drop-out Cohort)**	LS Mean Difference* (Observed Cohort)	**Threshold Δ Values (Drop-out Cohort)**
Study 065	SPID8	2.39 (N=96)	**-4.9345 (N=12)**	3.57(N=96)	**-14.248 (N=12)**
	SPID12	3.50(N=96)	**-6.2248 (N=12)**	5.80(N=96)	**-24.4441 (N=12)**
Study 066	SPID8	4.33(N=87)	**-17.244 (N=14)**	3.85(N=87)	**-14.2893 (N=14)**
	SPID12	6.17(N=87)	**-23.3737 (N=14)**	5.66(N=87)	**-20.2467 (N=14)**

*: Least-square mean difference based on ANOVA model with treatment, period, sequence, and patient(sequence) as factors.

Since the Δ values are all negative, with the equal variation assumption in Step 1, the statistical significance for valdecoxib groups vs. placebo will be retained unless placebo show certain advantage (see Δ values in Table 54) over valdecoxib groups in the drop-out cohorts.

VII.3. Comment for Multiplicity Adjustment Methods in Analgesic Studies

In the analgesic studies included in this review [redacted] Dysmenorrhea and [redacted] studies), Fisher's LDS methods was used to protect type I error rate for multiple between group comparisons. Since Fisher's LSD method only controls type I error rate when there are <=3 treatment groups are included in a study, the 'statistical significance' claim in studies with more than 3 treatment groups (Study 010, Study 011, Study 052, Study 065, Study 066) may not be valid. To check the validity of the 'significance' claim, this reviewer used Bonferroni method to adjust the multiple pairwise comparisons between the valdecoxib groups and placebo. The results by Bonferroni method is consistent with that by Fisher's LSD method in terms of 'statistical significance' claim for valdecoxib groups vs. placebo.

VII.4 ITT Analyses vs. 'All Randomized' Analyses in [redacted] Analgesia Studies and [redacted] Studies

In [redacted] analgesia studies and [redacted] studies, the ITT populations have excluded patients with requirements specified in the protocol. In [redacted] analgesia studies, the reasons for exclusion were 'incomplete measurements within 1 hour' and 'patient vomited within 30 minute after first dose of study medication'. In [redacted] studies, the reasons for exclusion were 'patient required analgesia within 30 minutes of the [redacted] and 'patient vomited within 30 minute after first dose of study medication'. The number of patient included in the ITT populations were 78%-89% of that of all randomized patient. Upon this reviewer's request, the sponsor conducted analyses for all randomized patients for [redacted] analgesia studies and [redacted] studies, the results of the 'all randomized' analyses were consistent with that of the ITT analyses.

VII.4 Comment for RA Studies

In Studies 60 and 61, ACR-20 response rate was one of the primary endpoints. The method in dealing with early withdrawals for this endpoint is LOCF, i.e., a patient was counted as a responder as long as the patient satisfied the ACR-20 criteria at the last visit before the patient left study. Since patients' withdrawal due to lack of efficacy or adverse events are usually considered as treatment failures, a sensitivity analysis is conducted for ACR-20 response by counting all patients who withdrew due to treatment failures as non-responders. The ACR-20 rates by the new analysis are lower in each treatment groups than that in the original analysis. However, the result in terms of statistical significance by the new analysis is generally consistent with that of the original analysis. The detailed results are in Table a.1 and a.2 in Appendix A.

VIII. Final Conclusions

VIII.1 Conclusions Based on [redacted] Analgesia Studies

[redacted]

1 page(s) have been removed because it contains trade secret and/or confidential information that is not disclosable.

69

[redacted]

VIII.2 Conclusions Based on Primary Dysmenorrhea Studies

In both Study 065 and Study 066, valdecoxib 20 mg (twice daily as needed) and valdecoxib 40 mg (twice daily as needed) demonstrated statistically significant advantage over placebo in terms of SPID, TOTPAR, time specific pain intensity difference and pain relief during the first dosing period. The median time to rescue medication or remedication was larger than 12 hours for all treatment groups in both studies. Valdecoxib 20 mg and valdecoxib 40 mg were not clearly separated from each other in terms of SPID, TOTPAR, time specific pain intensity difference and pain relief.

VIII.3. Conclusions Based on [redacted] Studies

[redacted]

VIII.4 Conclusions Based on OA Studies

Valdecoxib 10 mg BID has demonstrated statistically significant advantage over placebo in all three primary endpoints: WOMAC OA Pain Index, Patient's Global Assessment of Arthritis (Index Hip), WOMAC OA Physical Function Index in Study 049, but only in Patient's Global Assessment of Arthritis (Index Knee) in Study 053. Valdecoxib 5 mg BID also demonstrated statistically significant advantage over placebo in the three primary endpoints in Study 049. Despite the numerical advantage, Valdecoxib 5 mg failed to demonstrate statistically significant advantage over placebo in any primary

endpoint in Study 053. Compared with Valdecoxib 10 mg BID, there is no additional benefit from valdecoxib 20 mg BID.

VIII.5 Conclusions Based on RA Studies

In both Studies 060 and 061, valdecoxib 10 mg BID, 20 mg BID and 40 mg BID all demonstrated statistically significant advantage over placebo in 4 out of the 5 primary endpoints: ACR-20 Response, Physician's Assessment of Tender/Painful Joint Count, Patient's Global Assessment of Disease Activity and Physician's Global Assessment of Disease Activity. However, despite the numerical advantage, all the three valdecoxib dose levels failed to demonstrate statistically significant advantage over placebo in Physician's Assessment of Swollen Joint Count. Compared with valdecoxib 10 mg BID, valdecoxib 20 mg BID and 40 mg BID did not demonstrate additional benefit.

Laura Lu, Ph.D.

Mathematical Statistician

Concur:

Stan Lin, Ph.D.
Team Leader

CC:
NDA21341
HFD-550/MO/Johnson/Goldkind/Bull
HFD-550/PM/Schmidt
HFD-550/Div. File
HFD-725/Lu/Lin ST./Huque/Anello
HFD-725/Div. File

Page Number

72

is a blank page.

Table A. Appendix A

Table a1. Categorical Status Based on the ACR Responder's Index (20%)
With Treatment Related Dropouts (Due to Lack of Efficacy or AE) as Non-Responders
Number of Patients (%) (a)

Intent-to-Treat Cohort (ITT)

	PLACEBO (N=222)	VALDECOXIB 10 MG QD (N=209)	VALDECOXIB 20 MG QD (N=212)	VALDECOXIB 40 MG QD (N=221)	NAPROXEN 500 MG BID (N=225)	LINEAR TREND P-VALUE (c)
WEEK2						<0.001
RESPONDER (b)	50(23%)	82(39%)	80(38%)	90(41%)	95(42%)	
NON-RESPONDER	172(77%)	127(61%)	132(62%)	131(59%)	130(58%)	
TOTAL	222(100%)	209(100 %)	212(100%)	221(100%)	225(100%)	
WEEK6						<0.001
RESPONDER (b)	65(29%)	96(46%)	88(42%)	100(45%)	106(47%)	
NON-RESPONDER	157(71%)	113(54%)	124(58%)	121(55%)	119(53%)	
TOTAL	222(100%)	209(100 %)	212(100%)	221(100%)	225(100%)	
WEEK12						0.001
RESPONDER (b)	59(27%)	94(45%)	91(43%)	91(41%)	96(43%)	
NON-RESPONDER	163(73%)	115(55%)	121(57%)	130(59%)	129(57%)	
TOTAL	222(100%)	209(100 %)	212(100%)	221(100%)	225(100%)	

P-VALUE FOR TREATMENT COMPARISONS (d):

	PRIMARY		SECONDARY							
	20 MG QD VS. PLACEBO	40 MG QD VS. PLACEBO	10 MG QD VS. PLACEBO	20 MG QD VS. 10 MG QD	40 MG QD VS. 10 MG QD	40 MG QD VS. 20 MG QD	NAPROXEN VS. PLACEBO	NAPROXEN VS. 10 MG QD	NAPROXEN VS. 20 MG QD	NAPROXEN VS. 40 MG QD
WEEK2 :	<0.001#	<0.001#	<0.001	0.701	0.628	0.334	<0.001	0.594	0.472	0.924
WEEK6 :	0.004#	<0.001#	<0.001	0.279	0.945	0.281	<0.001	0.882	0.305	0.992
WEEK12 :	<0.001#	<0.001#	<0.001	0.583	0.529	0.880	<0.001	0.605	0.821	0.952

Note: The ITT cohort includes only patients who had at least one dose of study medication
(a) This table is based on the last observation carried forward approach
(b) Responder: At least 20% improvement from baseline in the number of tender/painful joints and in the number of swollen joints as well as at least 20% improvement from baseline in at least three of the following assessments: 1) Physician's Global 2) Patient's Global 3) Patient's Assessment of Pain 4) C-Reactive Protein 5) mHAQ.
(c) Cochran-Mantel-Haenszel test of linear dose trend stratified by center, p-value for Nonzero Correlation
(d) Cochran-Mantel-Haenszel test of treatment comparison stratified by center, p-value for Row MEAN Scores Differ
Statistically significant according to the Hochberg procedure(primary pairwise comparisons only)

Table a2.Categorical Status Based on the ACR Responder's Index (20%)
With Treatment Related Dropouts (Due to Lack of Efficacy or AE) as Non-Responders
Number of Patients (%) (a)

Intent-to-Treat Cohort (ITT)

	PLACEBO (N=220)	VALDECOXIB 10 MG QD (N=226)	VALDECOXIB 20 MG QD (N=219)	VALDECOXIB 40 MG QD (N=209)	NAPROXEN 500 MG BID (N=219)	LINEAR TREND P-VALUE (c)
WEEK2						0.130
RESPONDER (b)	54(25%)	84(37%)	77(35%)	68(33%)	96(44%)	
NON-RESPONDER	166(75%)	142(63%)	142(65%)	141(67%)	123(56%)	
TOTAL	220(100%)	226(100 %)	219(100%)	209(100%)	219(100%)	
WEEK6						<0.001
RESPONDER (b)	57(26%)	102(45%)	89(41%)	98(47%)	102(47%)	
NON-RESPONDER	163(74%)	124(55%)	130(59%)	111(53%)	117(53%)	
TOTAL	220(100%)	226(100 %)	219(100%)	209(100%)	219(100%)	
WEEK12						<0.001
RESPONDER (b)	60(27%)	94(42%)	89(41%)	94(45%)	98(45%)	
NON-RESPONDER	160(73%)	132(58%)	130(59%)	115(55%)	121(55%)	
TOTAL	220(100%)	226(100 %)	219(100%)	209(100%)	219(100%)	

P-VALUE FOR TREATMENT COMPARISONS (d):

	PRIMARY		SECONDARY							
	20 MG QD VS. PLACEBO	40 MG QD VS. PLACEBO	10 MG QD VS. PLACEBO	20 MG QD VS. 10 MG QD	40 MG QD VS. 10 MG QD	40 MG QD VS. 20 MG QD	NAPROXEN VS. PLACEBO	NAPROXEN VS. 10 MG QD	NAPROXEN VS. 20 MG QD	NAPROXEN VS. 40 MG QD
WEEK2 :	0.049	0.082	0.006	0.615	0.397	0.605	<0.001	0.152	0.045	0.023
WEEK6 :	0.002#	<0.001#	<0.001	0.309	0.648	0.173	<0.001	0.705	0.140	0.976
WEEK12 :	0.011#	<0.001#	0.002	0.674	0.423	0.277	<0.001	0.528	0.292	0.952

Note: The ITT cohort includes only patients who had at least one dose of study medication
(a) This table is based on the last observation carried forward approach
(b) Responder: At least 20% improvement from baseline in the number of tender/painful joints and in the number of swollen joints as well as at least 20% improvement from baseline in at least three of the following assessments: 1) Physician's Global 2) Patient's Global 3) Patient's Assessment of Pain 4) C-Reactive Protein 5) mHAQ.
(c) Cochran-Mantel-Haenszel test of linear dose trend stratified by center, p-value for Nonzero Correlation
(d) Cochran-Mantel-Haenszel test of treatment comparison stratified by center, p-value for Row MEAN Scores Differ
Statistically significant according to the Hochberg procedure(primary pairwise comparisons only)

Statistical Review and Evaluation
(Carcinogenicity Review)

NDA #: 21-341

Drug Name: Valdecoxib Tablets

Sponsor: G. D. Searle LLC Subsidiary of Pharmacia Corp.

Date Submission: January 15, 2001

Documents Reviewed: Carcinogenicity Portion from the Electric Submission.

Reviewing Pharmacologist: Josie Yang, Ph.D.

1. Background and Introduction

In this submission, total of 2 animal carcinogenicity studies are included:

Study No. SA4630/MSE-N 97095: Rat Carcinogenicity Study
Study No. SA4627/MSE-N 97091: Mouse Carcinogenicity Study

2. Reviewer's analyses of Rat Study (SA4630/MSE-N 97095)

2.1. Study Design

SC-65872, an anti-inflammatory compound under development for use in the treatment of osteoarthritis and rheumatoid arthritis and the management of pain, was administered once daily by oral gavage for least 104 weeks to male and female rats (98 weeks for high dose females). The objective of the study was to evaluate the oncogenic potential of SC-65872 when administered orally to rats.

The high dosages of 12.5 mg/kg/day in males and 5 mg/kg/day in females were projected to produce systemic exposures (AUC) that were comparable to the exposures at which toxicity was observed in the 13 week rat study. The above dose levels of SC-65872 were decreased once for males, low and mid-dose females and twice for high dose females during the study as a result of excessive mortality due to intestinal toxicity. The dose levels of the male groups from Day 1 through 158 (23 week) were 0, 2.5, 5.0 and 12.5 mg/kg/day; and from Day 159 to termination, 0, 2.5, 5.0 and 7.5 mg/kg/day for control, low, mid and high Toxicology and Pharmacokinetic groups, respectively. The female dosages from Day through 88 (13 week) were 0, 1.25, 2.5 and 5.0 mg/kg/day; from Day 89 through 158 the dosages were 0, 1.25, 2.5 and 3.75 mg/kg/day; and from Day 159 to termination the dosages were 0, 0.5, 1.0, and 1.5 mg/kg/day for control, low, mid and high Toxicology and Pharmacokinetic groups, respectively.

Due to high mortality in the test groups, all surviving animals in the Pharmacokinetic groups were reassigned to the Toxicology groups after the Week 52 pharmacokinetic bleeds. After reassignment, these animals were treated the same as the Toxicology animals.

The experimental design is summarized in the following table:

Table 2.1 Overview of study design / Animal type: Rat

Group Designation	Dosage mg/kg/day Males	Dosage mg/kg/day Females	No./ Sex	Females Sacrificed at week 99	Females Sacrificed at week 105	Males Sacrificed at week 105
Toxicology Animals						
V-T (Control-Toxicology)	0 (a)	0 (a)	100	10 Females	All Surviving Females	All Surviving Males
1	2.5	0.5	100		All Surviving Females	All Surviving Males
2	5.0	1.0	100		All Surviving Females	All Surviving Males
3	7.5	1.5	100	All Surviving Females		All Surviving Males
Pharmacokinetic Animals						
4 (V-P. Control Pharmacokinetic)	0 (a)	0 (a)	10			
5	2.5	0.5	25			
6	5.0	1.0	25			
7	7.5	1.5	25			

2.2. Statistical Methodology

2.2.1. General

All the analyses are performed by gender – male and female.

2.2.2. Evaluation of Validity of the Design

The evaluation of the validity of the study design can be performed by checking two questions:

1. If there were sufficient numbers of animals living long enough to get adequate exposure to the chemical and to be at risk of forming late-developing tumors.
2. If the doses used were high enough to present a reasonable tumor challenge to the tested animals.

The first question can be checked using criteria proposed in Lin and Ali (1994) as follow:
As a rule of thumb, a 50% survival rate of the initial animals in any treatment group between weeks 80-90 of a two-year study may be considered as a sufficient number and adequate exposure.

For the second question, it is generally accepted that the high dose should be close to the MTD (maximum tolerated dose). In Chu, Cueto and Ward (1981), the following criteria are mentioned for dose adequacy.

(i) "A dose is considered adequate if there is a detectable loss in weight gain of up to 10% in a dosed group relative to the controls."

(ii) "The administered dose is also considered an MTD if dosed animals exhibit clinical signs or severe histopathologic toxic effects attributed to the chemical."

(iii) "In addition, doses are considered adequate if the dosed animals show a slight increased mortality compared to the controls."

If one of the above applies, then the doses are considered to be properly selected.

2.2.3. Analyses of Mortality

The intercurrent data are tested first to see if the survival distributions of the treatment groups are significantly different and if the linear trend in mortality is significant. Cox test and Kruskal-Wallis test are used.

2.2.4. Analyses of Tumor Incidence

A. Tests of positive linear trend in incidence rates for individual tumor and tissue are performed. The prevalence method, the death-rate method, and the onset-rate method described in Peto et al. (1980) are used to analyze tumor data observed in incidental, fatal, and mortality-independent contexts of observation, respectively. For the prevalence method, the time intervals used are 0-50, 51-80, 81-104, 105-108.

B. For the linear trend test in incidence rates for individual tumor and tissue, either Exact Test or Asymptotic Test is used by following rule:

- Exact test: The statistical interpretation of significance is based on the exact test, if one of the two following situation applies.
  1. The tumor is found either fatal to all the animals or non-fatal to all the animals
  2. The tumor is fatal only to some but not to all animals, and time-intervals for both situations of lethality do not overlap

  The exact test is done using the Permutation test with general scores, which are the actual dose values. When the scores are set to be equally spaced, the above test is known as the Cochran-Armitage test.

- Asymptotic test: The statistical interpretation of significance is based on the asymptotic test, if none of the above situations applies. The asymptotic test uses the Z-statistic, following the standard normal distribution.

C. For the dual controls, all comparisons between one or pooled control groups with the test article are performed.

D. Since linear trend tests are performed on all the tumors and tissues, the overall false positive rate would be very large if each tumor and tissue was tested at 0.05 level of significance. Haseman (1983) proposed a rule to adjust for the effect of multiple testing. A modified rule, proposed by the Divisions of

Biometrics, CDER/FDA is applied to the trend tests in this review. In order to keep the overall type-I error at the level of about 10%, this rule states:

- Tumors with a spontaneous tumor rate of 1% or less are tested at the 0.025 significance level.
- Other wise, the 0.005 significance level is used

E. Combined analyses of tumors and Organs can be performed by pharmocological reviewer's request. When the combined analyses are performed, if more than one tumor incidences are obtained from one animal, it is counted as one incidence.

2.3. Study results by reviewer's analyses

2.3.1. Evaluation of Validity of the Design

For male and female rats, about 50% survived between 80-90 weeks for each treatment group as shown in Table 2.2, Figure 2.1, and Figure 2.2 below. The reviewer's evaluation of the validity of the design showed that, based on the survival data, the selected high doses were close to MTD. However, clinical signs or severe histopathologic toxic effects exhibited in the treated animals should also be considered in the final decision of the appropriateness of the selected high dose level.

2.3.2. Analyses of Mortality

Highly significant increases in mortality are observed.

Table 2.2 Mortality incidents for rat

WEEK	Control	Low	Med	High	Total
Male					
0-50	11	16	30	38	95
51-80	26	32	37	37	132
81-104	45	40	38	26	149
105-105	26	32	14	12	84
Total	108	120	119	113	460
Female					
0-50	2	4	33	66	105
51-80	31	36	30	19	116
81-104	47	51	29	13	140
105-105	19	32	19	.	70
Interim	10	.	.	12	22
Total	109	123	111	110	453

Figure 2.1 Kaplan-Meier Survival function for Rat/Male

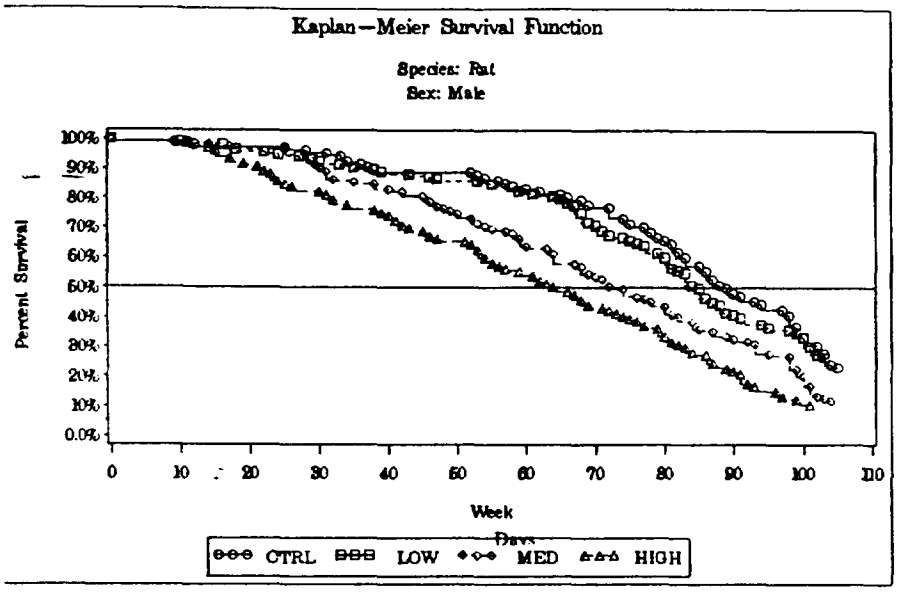
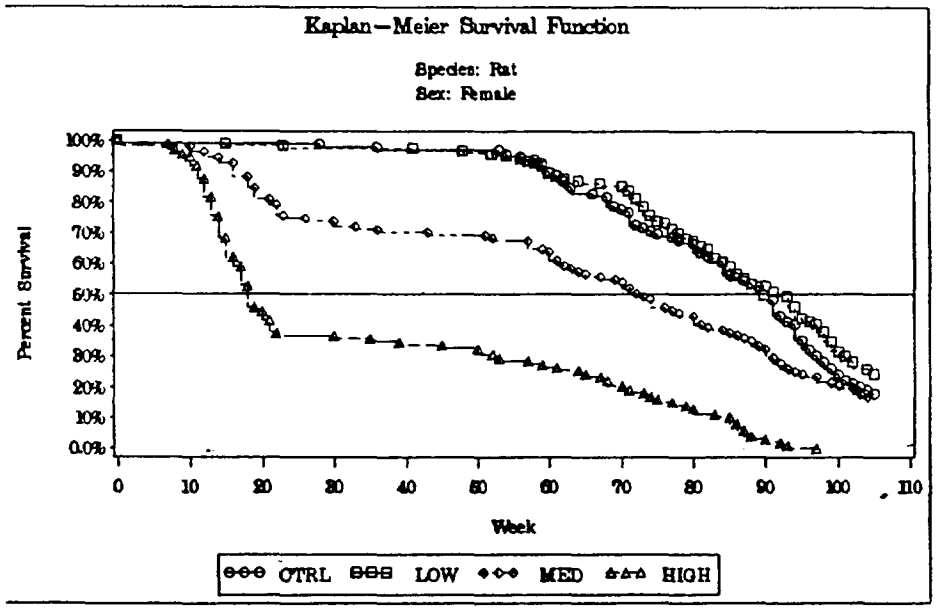


Figure 2.2 Kaplan-Meier Survival function for Rat/Female



As shown in the Table 2.2 and Figure 2.1, 2.2, positive trend of mortality is detected as doses are increased. For both male and female rat, except the Low dose group, higher mortality is observed on the Med and High dose group.

The following table displays the p-values of the test of homogeneity and of positive linear trends for males and females using the Cox test and the Kruskal-Wallis test.

Table 3.4 Results of homogeneity and positive linear trends test

SEX	METHOD	TIME-ADJUSTED	P-VALUE
Male	Cox	Dose-Mortality Trend	< 0.001*
		Homogeneity	< 0.001*
	Kruskal-Wallis	Dose-Mortality Trend	< 0.001*
		Homogeneity	< 0.001*
Female	Cox	Dose-Mortality Trend	< 0.001*
		Homogeneity	< 0.001*
	Kruskal-Wallis	Dose-Mortality Trend	< 0.001*
		Homogeneity	< 0.001*

* : statistically significant with significant level 0.05

The test of homogeneity and the test of linear trend yield highly significant results consistently for both male and female rats. So, the increment of mortality has been detected as dose increases, especially on Med and High dose group for both male and female rat.

2.3.3. Analyses of Tumor Incidence

Because of the early termination and euthanasia for the female rat, tumor incedences were analyzed for both including and excluding High dose group for female. For both male and female, and for both including and excluding High dose group for female, no significant trend test is observed in any organ and tumor based on the significant level adjusted for multiple comparison (See 3.2.4.D). All the analysis results of the organs and tumors are provided in the Table 7, Table 8a, Table 8b in the Appendix.

3. Reviewer's analyses of Mouse Study (SA4627/MSE-N 97091)

3.1. Study Design

SC-65872, a selective inhibitor of cyclooxyase-2 (COX-2), with potential therapeutic applications in the treatment of osteo- and rheumatoid arthritis and management of pain was administered by dietary admix for up to 105 weeks to Charles River CD-1 mice. The dietary route of exposure to SC-65872 was established in a 2-week feasibility study in CD-1 mice. The doses for this carcinogenicity study were selected based on the results of a 13-week study in CD-1 mice.

Initially, animals were administered diets intended to deliver 0, 12.5, 25, or 50 mg/kg/day (males) and 0, 25, 50, or 100 mg/kg/day (females). However, because of excessive mortality that occurred during the first 27 weeks of the study, intended doses were reduced by 50%. From Week 28 until the end of the study, male mice were administered diets intended to deliver 0, 6.25, 12.5 or 25 mg/kg/day and female mice were administered diets intended to deliver 0, 12.5, 25 or 50 mg/kg/day. In addition, interim sacrifices that were planned for Week 27 and Week 53 of the study were canceled in order to maximize the number of animals evaluated at termination.

The experimental design is summarized in the following table:

Table 3.1 Overview of study design / Animal type: Mice

Group Designation	Male Dosage (mg/kg/day)[a]	Female Dosage (mg/kg/day)[a]	No./ Sex	Animals/Sex Sacrificed at Termination
N	0[b]	0[b]	100	All Survivors
1	6.25	12.5	100	All Survivors
2	12.5	25	100	All Survivors
3	25	50	100	All Survivors

[a]Because of high mortality, all dose levels were decreased by 50% beginning at Week 28. The dose levels for Weeks 1 to 27 were 0,12.5,25, and 50 mg/kg/day males and 0, 25, 50 and 100 mg/kg/day (females).
[b]Basal diet only

3.2. Statistical Methodology

Same statistical methods as in rat study (SA4627/MSE-N 97091) are performed in the analyses of mouse study, excluding the time intervals used for the tumor incidence analyses of incidental tumor. Since all the female mouse were sacrificed at week 103, 0-50, intervals of 51-80, 81-101, 101-103 were used.

3.3. Study Results

3.3.1. Evaluation of Validity of the Design

For each male and female mouse survived about half between 80-90 weeks for each treatment group as shown in Table 3.2, Figure 3.1, and Figure 3.2. The reviewer's evaluation of the validity of the design showed that, based on the survival data, the selected high doses were close to MTD. However, clinical signs or severe histopathologic toxic effects exhibited in the treated animals should also be considered in the final decision of the appropriateness of the selected high dose level.

3.3.2. Analyses of Mortality

Table 3.2 Mortality incidents for mouse

WEEK	Control	Low	Med	High	Total
Male					
0-50	6	7	17	30	60
51-80	33	19	33	37	122
81-104	31	45	28	17	121
105-105	30	29	22	16	97
Total	100	100	100	100	400
Female					
0-50	3	6	15	47	71
51-80	17	24	33	24	98
81-101	34	40	29	16	119
102-103	46	30	23	13	112
Total	100	100	100	100	400

Figure 3.1 Kaplan-Meier Survival function for Mice/Male

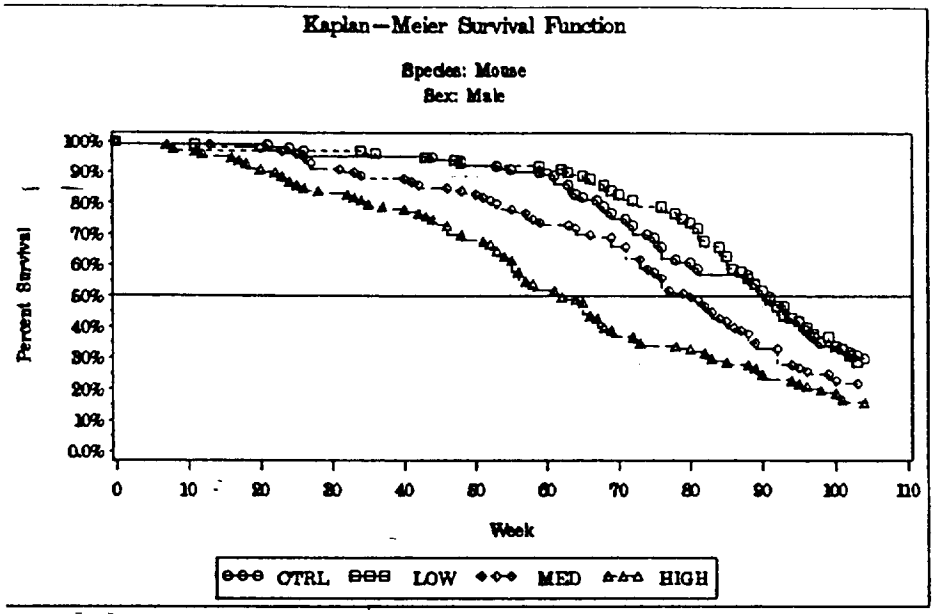
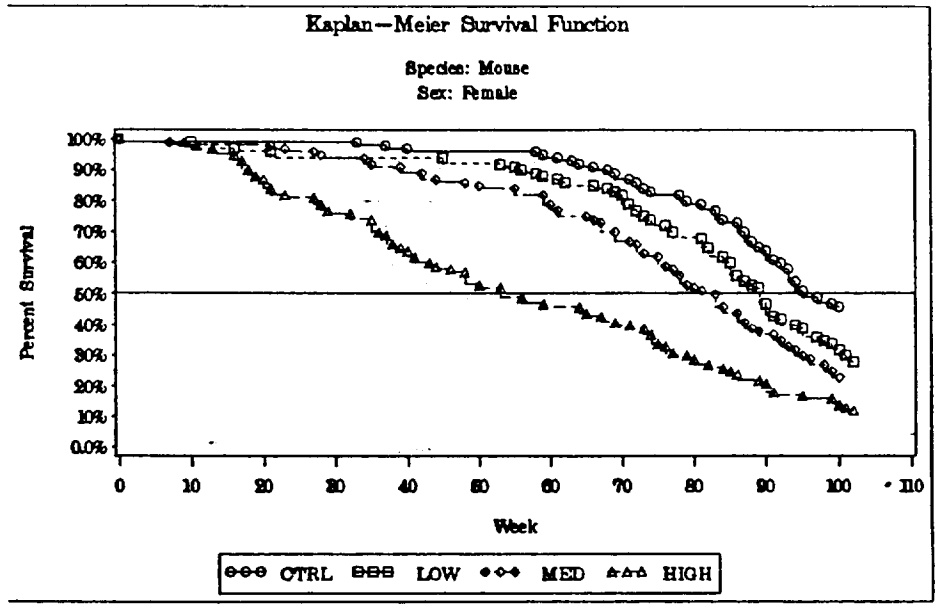


Figure 3.2 Kaplan-Meier Survival function for Mice/Female



As shown in the above table and figures, The High and Med dose group shows a higher mortality than controlled group for both male and female mice.

The following table displays the p-values of the test of homogeneity and of positive linear trends for males and females using the Cox test and the Kruskal-Wallis test.

Table 3.3 Results of homogeneity and positive linear trends test

SEX	METHOD	TIME-ADJUSTED	P-VALUE
Male	Cox	Dose-Mortality Trend	<0.001*
		Homogeneity	<0.001*
	Kruskal-Wallis	Dose-Mortality Trend	<0.001*
		Homogeneity	<0.001*
Female	Cox	Dose-Mortality Trend	<0.001*
		Homogeneity	<0.001*
	Kruskal-Wallis	Dose-Mortality Trend	<0.001*
		Homogeneity	<0.001*

*: statistically significant with significant level 0.05

The test of homogeneity and the test of linear trend yield highly significant results consistently for both male and female mouse. So, the increment of mortality has been detected as dose increases, especially on Med and High dose group for both male and female mice.

3.3.3. Analyses of Tumor Incidence

For both male and female, no significant trend test is observed in any organ and tumor based on the significant level adjusted for multiple comparison (See 3.2.4.D). All the analysis results of the organs and tumors are provided in the Table 15, and Table 16 of the Appendix.

4. Additional analyses for combined organs and tumors

Tumor incidences for combined organs and tumors were analyzed in addition for each rat and mice for both male and female. Methods of combinations were discussed with the pharmacologist, and the final decisions are summarized in Table 4.1, which was also reviewed by the pharmacologist.

However, none of the tested tumor type combinations showed statistically significant positive linear trend or increased incidence in the treated groups when compared with the control. Details of the combined analysis results are shown in Table 17 through Table 21 of Appendix.

Table 4.1 Combination list of Organs and Tumors

Organ	Tumor Combination	Rat		Mice	
		Male	Female	Male	Female
Skin	basal cell: all types of neoplasm	N/A	N/A	N/A	N/A
	squamous cell: papilloma + carcinoma	**YES***	N/A	N/A	N/A
	sebaceouscell: adenoma + carcinoma	**YES**	N/A	N/A	N/A
	leiomyoma + leiomyosarcoma	N/A	N/A	N/A	N/A
	fibroma + fibrosarcoma	**YES**	**YES**	N/A	N/A
Subcutis	basal cell: all types of neoplasm	N/A	N/A	N/A	N/A
	squamous cell: papilloma + carcinoma	N/A	N/A	N/A	N/A
	sebaceouscell: adenoma + carcinoma	N/A	N/A	N/A	N/A
	leiomyoma + leiomyosarcoma	N/A	N/A	N/A	N/A
	fibroma + fibrosarcoma	N/A	N/A	**YES**	N/A
Mammary Gland	adenoma + fibroma + fibroadenoma	**YES**	N/A	N/A	N/A
	carcinoma + adenocarcinoma	N/A	N/A	N/A	N/A
Lung	bronchioaveolar adenoma + carcinoma	N/A	N/A	N/A	N/A
All Organs	hemagioma	**YES**	N/A	**YES**	**YES**
	hemangiosarcoma	**YES**	N/A	N/A	N/A
	hemagioma + hemangiosarcoma	**YES**	N/A	N/A	N/A
	lymphoma + lymphocytic + lymphoblastic + histiocytic	N/A	N/A	N/A	N/A
Oral Cavity	squamous cell: papillma + carcinoma	N/A	N/A	N/A	N/A
	adenoma + adenosarcomas + adenomatous polyps	N/A	N/A	N/A	N/A
esophagus	squamous cell: papillma + carcinoma	N/A	N/A	N/A	N/A
	adenoma + adenosarcomas + adenomatous polyps	N/A	N/A	N/A	N/A
fore stomach	squamous cell: papillma + carcinoma	N/A	N/A	N/A	N/A
	adenoma + adenosarcomas + adenomatous polyps	N/A	N/A	N/A	N/A
glandular stomach	squamous cell: papillma + carcinoma	N/A	N/A	N/A	N/A
	adenoma + adenosarcomas + adenomatous polyps	N/A	N/A	N/A	N/A
small intenstine	squamous cell: papillma + carcinoma	N/A	N/A	N/A	N/A
	adenoma + adenosarcomas + adenomatous polyps	N/A	N/A	N/A	N/A
large intestine	squamous cell: papillma + carcinoma	N/A	N/A	N/A	N/A
	adenoma + adenosarcomas + adenomatous polyps	N/A	N/A	N/A	N/A
Liver	hepatocellular adenoma + carcinoma	**YES**	N/A	**YES**	**YES**
Kidney	tubular cell: adenoma + carcinoma	**YES**	**YES**	N/A	N/A
Urinary Bladder	transitional cell: adenoma + carcinoma	N/A	N/A	N/A	N/A
Pituitary (Gland)	ParsDistailis: adenoma + carcinoma	**YES**	**YES**	**YES**	**YES**
	adenoma + carcinoma	N/A	N/A	N/A	N/A
Thyroid (Gland)	adenoma + carcinoma	N/A	N/A	N/A	N/A
	follicular cell:adenoma + carcinoma	N/A	**Yes**	N/A	N/A
	"C"-cell: adenoma + carcinoma	**YES**	**YES**	N/A	N/A
Pancreas (Islets)	acinar cells: adenoma + carcinoma	**YES**	N/A	N/A	N/A
	islet cell: adenoma + carcinoma	**YES**	N/A	N/A	N/A
Adrenal (Gland)	cortical adenoma + carcinoma	N/A	**YES**	**YES**	N/A
	pheochromocytoma + malignant pheochromocytoma	**YES**	**YES**	N/A	N/A
Ovary / Testis	germ cell neoplasms: all types	N/A	N/A	N/A	N/A
	stromal cell neoplasms: all types	N/A	N/A	N/A	N/A
Uterus and Cervix	glandular carcinoma + glandular adenoma	N/A	N/A	N/A	N/A
	stromal polyps + stromal sarcoma	N/A	N/A	N/A	N/A
Prostate	adenoma + carcinoma	**YES**	N/A	N/A	N/A
Nervous System	glioma + oligodendroglioma + astrocytoma	N/A	N/A	N/A	N/A

* YES: Organs and Tumors are combined and analyzed in this review.

5. Summary

5.1. Rat Study

Statistically significant positive linear trend and increased mortality were detected in the treated groups when compared to the control in both male and female rats. About 50% of the rats survived between 80-90 weeks for each of the other treatment groups, so this study showed adequate exposure to the chemical for the study animals.

In tumor incidence analyses, differences in mortality among treatment groups was adjusted. Because of the early termination and euthanasia, the analyses of tumor incidence for female rats are performed for both including and excluding High dose group. None of the tested tumor types showed statistically significant positive linear trend or increased incidence in the treated groups when compared with the control.

Additional analyses (as requested by the Reviewing Pharmacologist) for the male as well as female rats showed no statistically significant dose-tumor positive linear trend in the treated groups when compared with the control.

5.2. Mouse Study

Statistically significant positive linear trend and increased mortality were detected in the treated groups when compared to the control in both male and female mouse. Again, about 50% of the mouse survived between 80-90 weeks for each of the other treatment groups, so this study had adequate chemical exposure to the animals.

In tumor incidence analyses, differences in mortality among treatment groups was adjusted. None of the tested tumor types showed statistically significant positive linear trend or increased incidence in the treated groups when compared with the control.

Additional analyses (as requested by the Reviewing Pharmacologist) for the male as well as female mouse showed no statistically significant dose-tumor positive linear trend in the treated groups when compared with the control.

Suktae Choi, Ph.D.
Mathematical Statistician

Concur: Stan Lin, Ph.D.
Team Leader

Cc: — Archival NDA 21-341
HFD-550/Schmidt/Osterberg/Yang
HFD-725/Choi/S.Lin/Huque/Anello
HFD-725/Division File/Chron