Study 035 (US)

This study had an identical design and efficacy endpoints than 034. A total of 784 patients were randomized (Appendix 11.1)

Table 19: Study 035. Patient randomization and accounting (6-month analysis)

Patients randomized/ evaluable	Rofecoxib 12.5mg 259 / 251	Rofecoxib 25mg 257 / 251	Diclofenac 50mg TID 268 / 264
Discontinued (% of patients randomized)	29.3	35.4	35.5
Lack of efficacy	13.1	19.5 *	10.1
Adverse events	11.6	10.5	18.3
Patient withdrew consent	1.2	1.6	3.7
Protocol deviation	2.3	2.3	2.6

- $p < 0.05$ vs diclofenac.

The incidence of discontinuation due to lack of efficacy in the diclofenac group was almost a half of the discontinuations due to lack of efficacy in the rofecoxib 25 mg group. This difference was statistically significant, in favor of diclofenac. The incidence of discontinuations due to adverse events was higher (but not statistically significantly different) in the diclofenac group. This difference was mainly due to events related to LFT's elevation.

Efficacy analyses

Table 20. Rofecoxib efficacy in study 035. Analysis of primary endpoints after 12 and 26 weeks.

	WOMAC Pain Walking		Investigator Global DS.		Patient Global Response	
	12 weeks	26 weeks	12 weeks	26 weeks	12 weeks	26 weeks
Rfx 12.5 mg	-31.5	-32.1	-2.3*	-2.3*	-1.4*	-1.5*
Rfx 25 mg	-32.5	-33.1	-2.3*	-2.3*	-1.4*	-1.5*
Diclofenac 150 mg	-35.3	-35.1	-2.5	-2.5	-1.6	-1.6

Statistical comparison to diclofenac $p < 0.05$.

<u>Efficacy analysis</u> over 12 weeks (Appendix 10.1)showed that

- Rofecoxib 12.5 mg/day was not different from 25 mg/d.
- Rofecoxib 25 mg/d was not statistically different from diclofenac in WOMAC Pain Walking on Flat surface but it was statistically different from diclofenac (in favor of diclofenac) for the other two primary endpoints.
- Rofecoxib 12.5 was statistically different from diclofenac in all three primary endpoints (WOMAC Pain Walking on Flat surface, Patient Global of response to therapy, Investigator Global of Disease status).

These differences were numerically small and within the limits of clinical comparability pre-defined by the applicant.

LS Mean changes for secondary endpoints were consistent with the changes in primary endpoints (Appendix 11.1).

Table 21. Study 035. Statistical Comparison of LS Mean changes for efficacy endpoints, from baseline over 12- week period (ITT).

Efficacy end point	12.5 vs 25	12.5 vs diclo	25 vs diclo
WOMAC walking on flat*	No SSD	SSD	No SSD
Pt Global of response to therapy* Investigator Global of disease status* Pt Global of disease status	SSD	"	SSD
WOMAC Physical Function , Pain Stiffness Total Score Average Subscale Average	No SSD	"	No SSD

- Primary endpoints

Analyses of efficacy after 6 months of treatment (secondary analysis) lead to similar results than for 12 weeks (Appendices 10.2 and 11.2).

> *Reviewer's comment: As in study 034, the changes from baseline after 26 weeks are very similar to the ones after 12 weeks. In this particular study, the high number of patients who discontinued due to lack of efficacy in the Rofecoxib 25 mg group (19.5%), might have had some impact in the results of the mean changes averaged over the 6 month period. However, the results are consistent with the results of rofecoxib 12.5 mg QD and the results from study 034.*

> *IN SUMMARY, in STUDY 034 and 035, daily doses of 12.5 and 25 mg QD were clinically COMPARABLE to the effect of diclofenac after twelve weeks and six months of treatment when using the applicant pre-defined criteria of clinical comparability.*

1.3.3. SUPPORTIVE STUDIES

Study 34c (Second six months of studies 034 and 035 and extensions (34-10 and 35-10) up to 86 weeks).

Study 058 and Study 058-10

Studies 044 and 045 (endoscopic studies)

- ## Study 34C : Second six months of 034 and 035 and extensions

1) <u>Design</u>: **034** (multinational) and **035** (US) were multi-center, double-blind, randomized, parallel, active- comparator controlled studies in patients with OA of the knee or hip. These studies were designed as one-year studies, but data from these studies was originally presented separately up to 6 months and pooled for analysis over the 12 month period.

2) <u>Treatment</u>: rofecoxib 12.5 or 25 mg QD, or diclofenac 50 mg TID.

3) <u>Entry criteria</u>: patients who completed 6 months of the base studies 034 and 035 were elegible to continue therapy. Of note, patients were allowed to take concomitant non-NSAID analgesics [opioids, acetaminophen] and systemic corticosteroids [including intraarticular corticosteroids] "without restriction for the supplemental treatment of OA".

4) <u>Demographics</u>: demographics for patients who entered study 034 and 035 are in Appendix A.2.

5) <u>Efficacy end points</u>: During the second six months of these studies, only two primary endpoints were measured: WOMAC Pain Walking on Flat Surface and Investigator Global of Disease Status. Patient Global assessment of Response to Therapy was not measured. For secondary endpoints see Table 3. Efficacy evaluations were done at week 39 and 52.

6) <u>Results</u>

APPEARS THIS WAY
ON ORIGINAL

Table 22. Patient accounting. 52 weeks.Study 034 and 035 pooled data.

PATIENT ACCOUNTING: Second 6 Months of 1-Year Base Studies (Protocols 034/035)	MK-0966 12.5 mg	MK-0966 25 mg	Diclofenac Sodium 150 mg	Total
ENTERED (second 6 months of 1-year Base Studies):	352	349	347	1048
Male (age range)'	92 (39 to 81)	101 (39 to 80)	83 (40 to 82)	276 (39 to 82)
Female (age range)'	260 (39 to 85)	248 (42 to 83)	264 (39 to 85)	772 (39 to 85)
	N (%)	N (%)	N (%)	N (%)
Completed 1-year Base Studies:	310 (88.1)	300 (85.9)	299 (86.2)	909 (86.7)
And not Extending:	86 (24.4)	82 (23.5)	84 (24.2)	252 (24.0)
And continuing into Extension Studies:	224 (63.6)	218 (62.5)	215 (62.0)	657 (62.7)
Discontinued (after the first 6 months, but before the end of 1 year):	42 (11.9)	49 (14.0)	48 (13.8)	139 (13.3)
Clinical adverse experience	10 (2.8)	16 (4.6)	15 (4.3)	41 (3.9)
Laboratory adverse experience	0 (0.0)	1 (0.3)	0 (0.0)	1 (0.1)
Protocol deviation	8 (22.7)	9 (2.6)	6 (1.7)	23 (2.2)
Patient withdrew consent	9 (2.6)	7 (2.0)	4 (1.2)	20 (1.9)
Lost to follow-up	3 (0.9)	0 (0.0)	1 (0.3)	4 (0.4)
Lack of efficacy	8 (2.3)*	13 (3.7)	19 (5.5)	40 (3.8)
Other reasons²	4 (1.1)	4 (1.5)	3 (0.9)	11 (1.0)

Source: Original NDA. Study 034c, Synopsis. * $p < 0.05$ compared to diclofenac.

Reviewer's comment.

Pooled data (Table 26) suggest that the percentage of patients who discontinued from these studies was between 12 and 14 % and that the discontinuation due to lack of efficacy was numerically higher in the diclofenac group. The Agency did not accept pooled data for efficacy claims.

Upon Agency request, on April 16 1999 the applicant provided data analyses of primary and secondary endpoints for up to 12 months, for studies 034 and 035 separately.

Table 23. Studies 34 and 035. Patient discontinuation (%) at 52 weeks.

	034				035			
	N	Total	LOE	Other	N	Total	LOE	Other
Rfx 12.5 mg	231	35.5	12.1	23.4	259	37.8	13.9	23.9
Rfx 25 mg	232	31.9	11.2	20.7	257	44.7	22.2	22.5
Diclofenac 150 mg	230	33.0	7.0	26.0	268	45.9	16.0	29.9

LOE: Discontinuation due to lack of efficacy. Other: Discontinuation due to clinically and laboratory adverse events, patient withdrew consent, loss to follow up, protocol violation, patient moved and termination by site.

When looking at individual studies the incidence of discontinuation due to lack of efficacy was numerically higher in the rofecoxib groups than in the diclofenac group (except for the 12.5 mg dose in study 035, that was lower than diclofenac).

Table 24. Studies 034 and 035. LS Mean change in Primary endpoints after 12 and 52 weeks of treatment.

| | 034 | | | | 035 | | | |
| | 12 weeks | | 52 weeks | | 12 weeks | | 52 weeks | |
WOMAC Pain Walking	LSM	Diff	LSM	Diff	LSM	Diff	LSM	Diff
Rfx 12.5	-30.6	5.3 *	-31.8	5.5 *	-31.5	3.7*	-34.3	1.8
Rfx 25	-32.5	3.0	-34.0	3.3	-32.5	2.8	-34.1	2.0
Diclo 150	-35.4		-37.3		-35.3		-36.1	
Investigator Global of Disease Status	LSM	Diff	LSM	Diff	LSM	Diff	LSM	Diff
Rfx 12.5	-1.4	0.2*	-1.5	0.1	-1.4	0.2*	-1.4	0.2*
Rfx 25	-1.5	0.1*	-1.5	0.1*	-1.4	0.2*	-1.4	0.2*
Diclo 150	-1.6		-1.6		1-.6		-1.6	

Diff: difference in LS Mean change compared to diclofenac. *$p<0.05$ vs. diclofenac.

When looking at efficacy endpoints, LS Mean changes from baseline were very similar after 12 and 52 weeks but we need to keep in mind that these results are average changes among patients who continued in the study.

Additionally, as mentioned earlier, the pre-established definition of comparability was that three out of three primary endpoints needed to be successful. During the second six months of these studies, only two primary endpoints were measured: WOMAC Pain Walking on Flat Surface and Investigator Global of Disease Status. In a letter to the Agency the applicant stated that Patient Global of Response to Therapy was not measured because unrestricted concomitant non-NSAID analgesic medications were allowed. We share the concern that the use of other analgesics may affect the Patient Global assessment of Response to Therapy, and think that this confounding factor may also interfere with the measurement of other efficacy endpoints during the second six months of the one year studies.

APPEARS THIS WAY
ON ORIGINAL

- **Study 058:** 6-week study of rofecoxib in elderly patients with OA.

1. <u>Design</u>: multi-center US, randomized, double-blind, parallel, placebo and active comparator controlled study in elderly patients with OA of the knee or hip.

2. <u>Treatment</u>: Placebo, rofecoxib 12.5 or 25 mg QD, or nabumetone 500 mg TID

3. <u>Entry criteria</u>: Inclusion/exclusion criteria were similar to other studies except for:

   a) Age: 80 years of age or more
   b) Disease activity criteria: (Non-flare design). No need for flare compared to screening visit :
      <u>For regular NSAIDS users:</u>
      At Visit 1 (prestudy) Pt Global of D Status needed to be <90mm on a 100 mm VAS.
      At Visit 2 ("Flare/Randomization" visit after washout), Pt Global needed to be ≥ 40 mm.

      <u>For Acetaminophen users:</u>
      At both visits 1 and 2 patient gave a global assessment of their OA ≥ 40 and <90 mm on a 100 mm VAS.

   c) Concurrent/previous medications:
      Patients were allowed to take narcotic analgesia (e.g. opioids, tramadol).
      Patients were allowed to take low dose aspirin up to 325 mg/day for cardioprotective/antiplatelet effects. Patients were stratified for low-dose aspirin use.

3. <u>Demographics</u>: Baseline demographics-and endpoint values- prior diseases or medications were similar across groups except that the 25 mg group had higher number of low dose ASA users.

   The baseline values for efficacy endpoints in study 058 were lower than in other studies, for instance Mean Baseline value for Pain Walking on Flat Surface was 55 ± 25 mm compared to 75 ± 15 mm in other studies (Appendix 2). ARA Functional Class was not an entry criterion and data are not available.

5. <u>Efficacy end points</u>:

   Patient Global Assessment of Disease Status (very well=0 to very poor=100 mm VAS) was specified as <u>the primary efficacy endpoint</u>, but other endpoints were also measured:

   - Investigator Global Assessment of Disease Status (0 to 4 point Likert scale)
   - Discontinuation Due to Lack of Efface or Toxicity
   - Study joint examination
   - Short Form- 346 Health Survey (SF-36)
   - WOMAC VA 3.0 Osteoarthritis Index

The primary analysis of efficacy endpoints planned for this protocol was a comparison of LS mean changes from baseline at 6 weeks. The analysis of LS mean changes averaged OVER 6 weeks was a secondary analysis. Results for these two approaches were similar.

6. Results.

6.2 - Randomization and accounting

Table 25: Study 058 .Patient randomization and accounting:

Randomized/ evaluable for ITT	Placebo 52 / 52	Rofecoxib 12.5 mg/day 118 /118	Rofecoxib 25 mg/day 56 /54	Nabumetone 1500 mg/day 115 / 114
Discontinued (% of randomized)	17.3	14.4	14.3	13.0
Lack of efficacy	11.5	1.7	0.0	1.7
Adverse events	1.9	7.6	8.9	7.0
Patient withdrew consent	3.8	3.4	1.8	2.9

Patient discontinuation due to lack of efficacy was statistically significantly higher in the placebo group than in the active comparator. Discontinuation due to adverse events was evenly distributed among active comparators. Relatively high number of patients withdrew consent in the placebo and rofecoxib 12.5 mg group.

6.2 - Efficacy analysis

The primary analysis for this study (LS Mean changes from baseline *at* six weeks) for the primary endpoint (Patient Global of Disease Activity) showed that rofecoxib at the doses of 12.5 and 25 mg QD was not statistically different from nabumetone 1500 mg/day. There was no statistically significant difference between 12.5 and 25 mg QD. All three treatments were statistically different from placebo.

Table 26. Study 058. Patient Global Assessment of Disease Status. LS Mean changes from baseline at week 6 and between group comparisons.

Placebo	Rfx 12.5 mg QD		Rfx 25 mg QD		Nabumetone 1500 mg/day	
LSM	LSM	Difference	LSM	Difference	LSM	Difference
-14.9	-25.3	-10.5 (-18.7, -2.3)	-25.4	-10.6 (-20.1, -1.0)	-26.0	-11.1 (-19.4, -2.9)

LSM; LS mean changes from baseline. Difference: Difference with placebo (95 % confidence interval).

Reviewer's comment: Although statistically different from placebo, the difference in LS mean changes from baseline for the primary endpoint for the three active treatments was no greater than –11 mm (VAS). In this case a statistically significant difference is not clearly clinically meaningful. Therefore, in this trial, the applicant has not proven that rofecoxib is more effective than placebo in an elderly population. A possible explanation is that

patients were not required to have a flare to enter the study, so there was less "room for improvement".

The analysis of changes from baseline averaged *over* the 6-week period (secondary analysis for this study) showed similar results, although the changes were numerically higher (Appendix 12.).

Most of the other endpoints showed similar results: active treatments were statistically different from placebo but very close or within what would be the limits of clinical comparability to placebo. Neither rofecoxib nor nabumetone showed statistically significant differences to placebo in the SF-36 questionnaire.

In summary, rofecoxib at the doses of 12.5 and 25 mg QD was comparable in efficacy to nabumetone in a six-week trial in an elderly population. However, in this trial, the active treatments did not show a clear clinically meaningful difference from placebo.

058-10 (058 extension)

Entry criteria:

Patients who had completed Visit 6 (last visit) of the Base study were offered to enroll in the First Extension study. Those who met selection criteria and accepted to continue for 24 weeks were enrolled in the extension (Visit 6 base = Visit 0 of extension). The other patients discontinued treatment.

Table 28. Study 058. Patients who completed the base study and entered the extension

		MK-0966		Nabumetone	
	Placebo	12.5 mg	25 mg	1500 mg	Total
Entered Base	N=52	N=118	N=56	N=115	N=341
Completed Base Study	43 (82.7%)	101 (85.6%)	48 (85.7%)	100 (87.0%)	292 (85.6%)
Number (%) of Patients who Entered First Extension[t]	38 (73.1%)	81 (68.6%)	38 (67.9%)	79 (68.7%)	236 (69.2%)

At the time of the cutoff date of 3/31/98, 86.2 %, 74 % and 77.2 % of patients remained in the study. Almost half of the patients who discontinued rofecoxib 12.5 and 25 did so because of lack of efficacy, while only 1/6 of the patients in nabumetone did dropout because of lack of efficacy. Only 3 patients had completed the last visit of this extension, so there is no much to say about this extension.

- **Studies 044, 045:** 6-month endoscopic studies in patients with OA.

   Reviewer's comment: these studies are included here because no other Phase III studies used rofecoxib 50 mg dose for 6 months. There were no study objectives or hypothesis related to OA efficacy in this study, although Patient Global Assessment of Disease Activity Status was measured at each visit using a Likert scale from 0 (very well) to 4 (very poor) to measure OA symptoms. (Of note, all other studies used a VAS scale for Pt global).

1. <u>Design</u>: identical design, multi-center, randomized, double-blind, parallel, placebo and active-comparator controlled study in patients with OA of the knee or hip to determine the incidence of GI ulceration after 12 weeks of treatment with a 12 week extension. **044** was multi-center US study (775 patients). **045** was multinational non-US (742 patients).

2. <u>Treatment</u>: rofecoxib 25 and 50 mg QD and ibuprofen 800 mg TID

3. <u>Entry criteria</u>:
- Patient with OA of knee or hip , ≥ 50 years of age, without gastroduodenal ulceration at baseline.
- No minimun OA disease activity criteria (non-flare study)
- Escape medications: Patients were allowed to take acetaminophen 650 mg 4 times daily in a regular bases and to use local measures (heat, capsaicin, physical therapy). If this was unsuccessful, the patient received a prescription for a non-NSAID pain medication that had successfully controlled their pain in the past (e.g., propoxyphene with acetaminophen; acetaminophen with codeine; tramadol). If this was unsuccessful, the patient was discontinued from the study and underwent endoscopy as soon as possible.

2. <u>Results</u>:   Patient Global assessment

   All active treatments significantly improved the Patient Global Assessment of Disease Activity Status at 12 weeks and over 24 weeks. After 12 weeks and after 24 months rofecoxib 25 and 50 mg QD was comparable to ibuprofen although there was a trend in favor of rofecoxib 50 mg QD. (Appendix 12). Of note, a high number of patients were discontinued from these studies, (particularly from the ibuprofen group) due to finding an endoscopic endpoint or due to adverse events.

In summary Patient Global Assessment of disease Status of rofecoxib 25 and 50 mg QD was comparable to ibuprofen after 12 and 24 weeks. However, a relatively small number of patients were left at the end of the studies and there were confounding factors as the use of analgesics and local treatments. Therefore no efficacy conclusions can be drawn from these trials.

CONCLUSIONS FROM THE OA EFFICACY STUDIES

The following conclusions regarding Rofecoxib and treatment of the signs and symptoms of OA are drawn from the information presented in four pivotal clinical trials (033, 040, 034 and 035):

- Rofecoxib, 12.5 and 25 mg once a day, was efficacious vs. placebo in six-week trials.

- Rofecoxib 12.5 and 25 mg once a day was clinically comparable to the efficacy of Ibuprofen 800 mg TID in 6 week trials and to Diclofenac 50 mg TID up to 6 months, when using the criteria of clinical comparability pre-defined by the applicant (± 10 in a 1-100 mm VAS scale or ± 0.5 in a 0-4 point Likert scale).

Additional information is drawn from four randomized controlled trials (029, 044, 045, and 058):

- Rofecoxib 50 mg once a day was statistically significantly more efficacious than 12.5 and 25 mg once a day, in a six week dose ranging study. These data has not been replicated. No other studies were done to look at the efficacy of the 50 mg dose. Limited data from 6-month studies designed to address safety issues, showed that rofecoxib 50 mg QD was associated with numerically higher number of renal-related and GI adverse events.

- Rofecoxib 12.5 and 25 mg was statistically and clinically comparable to nabumetone in all efficacy endpoints in a six-week study in an elderly population. However, although there was a statistically significant difference between placebo and active treatments, neither nabumetone nor rofecoxib showed a difference from placebo that could be considered clinically meaningful.

APPEARS THIS WAY
ON ORIGINAL

APPENDICES TO OA EFFICACY REVIEW
Appendix 1. - Washout period

NSAID Dose Ranges and Washout Period

NSAID	Total Daily Dose (mg)
3- to 8-Day Washout Period	
Diclofenac	100 to 150 mg
Fenoprofen	900 to 2400 mg
Flurbiprofen	200 to 300 mg
Ibuprofen	1200 to 3200 mg
Indomethacin/INDOCIN SR™	50 to 200 mg
Ketoprofen	150 to 300 mg
Meclofenamate sodium	100 to 200 mg
Mefenamic acid	1000 mg
Tolmetin sodium	1200 to 1800 mg
Etodolac	600 to 1200 mg
Diflunisal	500 to 1000 mg
Aceclofenac	200 mg
4- to 9-Day Washout Period	
Aspirin	1950 to 3250 mg
Choline magnesium trisalicylate	2000 to 3000 mg
Salsalate	2000 to 3000 mg
Sulindac	150 to 400 mg
Naproxen	500 to 1500 mg
Naproxen sodium	550 to 1650 mg
5- to 10-Day Washout Period	
Meloxicam	7.5 to 15 mg
Nabumetone	1000 to 2000 mg
7- to 12-Day Washout Period	
Oxaprozin	600 to 1200 mg
10- to 15-Day Washout Period	
Piroxicam	20 mg

[P034]

Appendix.2 – Baseline characteristics of Phase II and III OA Studies

Study	6-Week Placebo- and Active-Comparator-Controlled Studies				1-Year Studies		All Studies Combined
	Protocol 029	Protocol 033	Protocol 040	Protocol 058	Protocol 034	Protocol 035	
Total Number Randomized	672	736	809	341	693	784	4035
Gender [n (%)]							
Female	477 (71.0)	548 (74.5)	647 (80.0)	217 (63.6)	555 (80.1)	529 (67.5)	2973 (73.7)
Male	195 (29.0)	188 (25.5)	162 (20.0)	124 (36.4)	138 (19.9)	255 (32.5)	1062 (26.3)
Study Joint [n (%)]							
Hip	189 (28.1)	175 (23.8)	180 (22.2)	107 (31.38)	130 (18.8)	190 (24.2)	971 (24.1)
Knee	483 (71.9)	561 (76.2)	629 (77.8)	234 (68.62)	563 (81.2)	594 (75.8)	3064 (75.9)
Treatment (n Per Group)							
Placebo	145	69	74	52	NA	NA	340 (8.43)
MK-0966 5 mg	149	NA	NA	NA	NA	NA	149 (3.69)
MK-0966 12.5 mg	144	219	244	118	231	259	1215 (30.11)
MK-0966 25 mg	137	227	242	56	232	257	1151 (28.53)
MK-0966 50 mg	97	NA	NA	NA	NA	NA	97 (2.40)
Diclofenac 150 mg	NA	NA	NA	NA	230	268	498 (12.34)
Ibuprofen 2400 mg	NA	221	249	NA	NA	NA	470 (11.65)
Nabumetone 1500 mg	NA	NA	NA	115	NA	NA	115 (2.85)
Age							
Mean age (years)	61.7	61.0	63.6	82.7	62.3	62.7	64.0
Age range (years)	39 to 92[1]	39 to 91[1]	40 to 90[1]	79 to 94[1]	38 to 85[1]	39 to 91[1]	38 to 94[1]
Race [n (%)]							
Asian	2 (0.3)	4 (0.5)	2 (0.25)	1 (0.3)	4 (0.6)	1 (0.1)	14 (0.35)
Black	45 (6.7)	53 (7.2)	13 (1.61)	14 (4.1)	9 (1.3)	65 (8.3)	199 (4.93)
Eurasian	NA	1 (0.1)	NA	NA	NA	NA	1 (0.02)
European	NA	1 (0.1)	NA	NA	1 (0.1)	NA	2 (0.05)
Hispanic	19 (2.8)	7 (1.0)	232 (28.68)	4 (1.2)	171 (24.7)	14 (1.8)	447 (11.08)
Indian	NA	2 (0.3)	NA	NA	NA	NA	2 (0.05)
Multiracial	NA	NA	1 (0.12)	NA	15 (2.2)	NA	16 (0.40)
Native American	9 (1.3)	9 (1.2)	2 (0.25)	5 (1.5)	NA	NA	25 (0.62)
Polynesian	NA	NA	1 (0.12)	NA	NA	2 (0.3)	3 (0.07)
White	597 (88.8)	659 (89.5)	558 (68.97)	317 (93.0)	493 (71.1)	702 (89.5)	3326 (82.4)
Weight (kg)							
Mean weight	86.2	88.9	75.77	73.84	77.92	89.4	82.76
Weight range	44.6 to 128.8	49.2 to 176.9	41.7 to 139.0	46.27 to 129.28	40.0 to 135.0	40.4 to 163.3	40.0 to 176.9
Duration of OA							
Mean duration (years)	10.9	10.0	8.65	15.0	8.79	11.3	10.35
ARA Functional Class[*] n [(%)]							
I	106 (15.8)	103 (14.0)	103 (12.7)	‡	119 (17.2)	109 (13.9)	540 (14.62)
II	448 (66.7)	454 (61.7)	479 (59.2)	‡	404 (58.3)	517 (65.9)	2302 (62.32)
III	118 (17.6)	179 (24.3)	227 (28.1)	‡	170 (24.5)	158 (20.2)	852 (23.06)

* Patients with ARA Functional Class IV were not permitted to enter the studies.
‡ ARA Functional Class was not an entry criterion in Protocol 058 and data are not available.
[1] Birth date masking to 12/31/XX in database, where XX is the true year of birth, may cause patient's age to appear 1 year younger than the actual age, depending on the date of randomization.
NA=not applicable.

Study	6-Week Placebo- and Active-Comparator-Controlled Studies				1-Year Studies		All Studies Combined
	Protocol 029 (N=672)[‡]	Protocol 033 (N=736)[‡]	Protocol 040 (N=809)[‡]	Protocol 058[*] (N=341)[‡]	Protocol 034 (N=693)[‡]	Protocol 035 (N=784)[‡]	
Pain Walking on a Flat Surface (WOMAC) (0- to 100-mm VAS)							
Mean (SD)	74.1 (16.1)	74.9 (15.3)	72.5 (14.64)	55.1 (24.75)	72.5 (15.41)	76.4 (15.0)	72.5 (16.28)
Range	6 to 100	23 to 100	40 to 100	2.00 to 97.00	40 to 100	29 to 100	6 to 100
Patient Global Assessment of Response to Therapy (No Baseline Values)							
Investigator Global Assessment of Disease Status (0- to 4-Point Likert Scale)							
Mean (SD)	2.9 (0.67)	2.9 (0.60)	3.0 (0.62)	2.25 (0.74)	3.04 (0.64)	2.9 (0.70)	2.89 (0.66)
Range	1 to 4	1 to 4	2 to 4	0 to 4	1 to 4	1 to 4	0 to 4

* Pain Walking on a Flat Surface (WOMAC) and Investigator Global Assessment of Disease Status were other end points for Protocol 058
‡ N represents the total number of patients randomized.

Apppendix 3 Description of efficacy endpoints

A.3.1

Clinical Efficacy End Points (Cont.)

D-6

Correspondence of Osteoarthritis Disease Manifestations to Clinical End Points Measured in MK-0966 Osteoarthritis Program

OA Disease Manifestation	MK-0996 Clinical Program End Points	Designation
Pain	Pain Walking on a Flat Surface (WOMAC)	Primary
	Pain Subscale (WOMAC)	Key Secondary[*]
	Acetaminophen Use (for Rescue)	Other
Global Signs and Symptoms—Patient Perspective	Patient Global Assessment of Response to Therapy	Primary
	Patient Global Assessment of Disease Status	Key Secondary[*]
	Discontinuation Due to Lack of Efficacy	Key Secondary
Global Signs and Symptoms—Investigator Perspective	Investigator Global Assessment of Disease Status	Primary
	Investigator Global Assessment of Response to Therapy	Other
Physical Disability	Physical Function Subscale (WOMAC)	Key secondary
Joint Stiffness	Stiffness Subscale (WOMAC)	Key secondary
Signs (Physical Examination Findings)	Study Joint Tenderness—Knee or Hip	Other
	Study Joint Swelling—Knee only	Other
Health-Related Quality of Life	SF-36 Health Survey — Mental Component Score — Physical Component Score	Other[*]

[*] Primary end point for Protocol 010. Patient Assessment of Arthritic Pain (not shown on table) was a co-primary end point for this protocol.
[*] Primary End Point for Protocol 058, and the Extensions to 029, 034, and 035, and an exploratory end point for Protocols 044 and 045.
[*] Protocols 029 and 058 only.

[PO10; PO29; PO33; PO34; PO34C; PO35; PO40; PO44; PO45; PO58; PO29X; PO29C]

A.3.2 Patient Global assessment of response to Therapy

Patients rated the overall effects of study medication on OA symptoms using a 5-point Likert scale, ranging from "NONE" to "EXCELLENT"

- 0 = None—no good at all, ineffective
- 1 = Poor—some effect, but unsatisfactory
- 2 = Fair—reasonable effect, but could be better
- 3 = Good—satisfactory effect with occasional episodes of pain and/or stiffness
- 4 = Excellent—ideal response, virtually pain free

A.3.3 Physician Global assessment of Disease Activity

Investigators provided global assessments of the patient disease status using a 5-point Likert scale ranging from "VERY POOR" to "VERY WELL"

- 0 = Very well • 1 = Well • 2 = Fair • 3 = Poor • 4 = Very poor

A.3.4 Patient Global assessment of Disease Activity

Patients placed an "X" on a 100-mm VAS ranging from "VERY WELL" (0 mm) to "VERY POOR" (100 mm) representing their assessment of their overall OA disease status.

A.3.5 WOMAC (from original NDA 21042)

Several of the end points used in clinical efficacy OA Studies derive from the Western Ontario and McMaster Universities (WOMAC) Osteoarthritis Index questionnaire.

. This is a validated instrument designed to assess the clinical status of patients with OA of the knee or hip. The questionnaire contains 24 individual questions divided among three subscales: Pain, Physical Function (Disability), and Stiffness. The three subscales of the WOMAC questionnaire measure the primary clinical symptoms of lower extremity arthritis from the patient's perspective. Each of the three subscales as well as the first question of the WOMAC questionnaire, Pain Walking on a Flat Surface, were selected as clinical end points.

Pain Subscale (WOMAC)

Patients were asked: "In the past 48 hours, how much pain do you have:
"Walking on a flat surface?" "Going up or down stairs?" "At night while in bed?"
"Sitting or lying?" "Standing upright?"
Possible responses ranged from "NO PAIN" (0 mm) to "EXTREME PAIN" (100 mm) on a VAS. The Pain Subscale (WOMAC) score was the average of the five responses

Physical Function Subscale (WOMAC)

Patients were asked 17 QUESTIONS. Possible responses to each question ranged from "NO DIFFICULTY" (0 mm) to "EXTREME DIFFICULTY" (100 mm) on a VAS for each question. Responses of the 17 questions in the physical function category were averaged

Stiffness Subscale (WOMAC)

Patients were asked: "How severe is your stiffness after first awakening in the morning?" "How severe is your stiffness after sitting, lying, or resting later in the day?" Possible responses To each question ranged from "NO STIFFNESS" (0 mm) to "EXTREME STIFFNESS" (100 mm) on a VAS for each question .The Stiffness Subscale (WOMAC) score was the average of the two individual responses

Total Score Average (WOMAC)

The Total Score Average (WOMAC) was the average of all 24 question scores on the WOMAC. Each individual question was weighted equally.

Average Subscale (WOMAC)

The Subscale Average (WOMAC) was the average of each of the 3 subscale scores (discussed above): Pain, Stiffness, and Physical Function. Each subscale was weighted equally.

A.3.6 Short Form-36 (SF-36)

Each patient's quality of life was measured using the validated *Medical Outcome Study 36-Item Short-Form Health Survey (SF-36)*. The SF-36 is composed of 36 questions grouped into two component scales: the Physical Component Scale (PCS) and the Mental Component Scale (MCS). Scores were computed according to the standard instructions for the SF-36. Each of the component scales was computed and reported as a 0 to 100 score.

The *Physical Component Scale* consists of four domains:
• Physical function • Role physical • Bodily pain • General health

The *Mental Component Scale* consists of four domains:
• Mental health • Role emotional • Social function • Vitality

The two component scales and each of the eight individual domains were analyzed.

Appendix 4. Schedule of observations

<div align="center">

Table 2

Schedule of Clinical Observations and Laboratory Measurements

</div>

Week:	-1 to -2	0	2	4	8	12	19	26		
Visit:	Screening	Flare/ Randomization	Treatment						Discon-	Post-
Visit No:	1	2	3	4	5	6	7	8	tinuation	study
Review of Screening criteria	X									
Informed consent	X									
Medical history	X									
Review of Flare/Randomization criteria		X								
Interim history and monitor for adverse experiences		X	X	X	X	X	X	X	X	X
Vital signs and weight	X	X	X	X	X	X	X	X	X	X
Physical examination and ECG	X								X	
Stool Hemoccult, knee or hip X-ray	X									
CBC, serum chemistry, UA[*]	X	X	X	X	X	X	X	X	X	X
Serum β-HCG[‡]	X									
Urine β-HCG[‡†]		X	X	X	X	X	X	X	X	X
Plasma and urine for archive[ǁ]		X		X				X	X	
Dispense study medication		X		X	X	X	X	X		
Dispense paracetamol[¶]	X	X		X	X	X	X	X	X	
Study medication tablet count			X	X	X	X	X	X	X	
Patient/investigator global assessment of disease status	X	X	X	X	X	X		X	X	
Study joint examination	X	X		X		X		X		
Pain, physical function, stiffness subscales (WOMAC)	X	X	X	X	X	X		X	X	
Patient/investigator global assessment of response to therapy			X	X	X	X		X	X	

[*] Serum chemistry obtained following 8-hour fast at Visits 1 and 4.
[‡] Urine and serum β-HCG samples were obtained from women of childbearing potential only.
[†] Urine β-HCG must have been negative prior to dosing.
[ǁ] At Visit 4, sample was obtained prior to 9:00 AM and taking the morning dose of study medication.
[¶] Only chronic NSAID patients received paracetamol at Visit 1.

Data Source: [3.2]

Appendix.5. Study 010. Results

Table D-10 — Pain Walking on a Flat Surface and Patient and Investigator Global Assessments
Average Over 6 Weeks of Treatment
(Protocol 010)

Protocol	Baseline Mean	Treatment Period Mean	LS Mean' Difference From Baseline (95% CI)	LS Mean' Difference From Placebo (95% CI)
Pain Walking on a Flat Surface (WOMAC) (0- to 100-mm VAS)				
Placebo (n=70)	57.97	51.41	-6.95 (-11.71, -2.20)	N/A
MK-0966 25 mg (n=73)	63.38	34.83	-25.97 (-30.62, -21.31)	-19.01 (-25.71, -12.32)
MK-0966 125 mg (n=74)	58.32	29.57	-28.95 (-33.58, -24.33)	-22.00 (-28.64, -15.36)
Patient Global Assessment of Response to Therapy (0- to 4-Point Likert Scale)				
Placebo (n=70)	N/A	-1.36	-1.33 (-1.56, -1.11)	N/A
MK-0966 25 mg (n=73)	N/A	-2.63	-2.63 (-2.85, -2.41)	-1.29 (-1.61, -0.98)
MK-0966 125 mg (n=70)	N/A	-2.82	-2.81 (-3.04, -2.59)	-1.48 (-1.81, -1.16)
Investigator Global Assessment of Disease Status (0- to 4-Point Likert Scale)				
Placebo (n=69)	2.74	2.21	-0.54 (-0.72, -0.36)	N/A
MK-0966 25 mg (n=70)	2.80	1.28	-1.51 (-1.69, -1.33)	-0.97 (-1.23, -0.71)
MK-0966 125 mg (n=70)	2.77	1.20	-1.57 (-1.75, -1.39)	-1.03 (-1.29, -0.77)

$p<0.001$ LS mean difference from placebo for both MK-0966 treatment groups. Decreasing values represent improvement.
[P010]

APPEARS THIS WAY ON ORIGINAL

Appendix 6. 1 to 6.5

A. 6.1 . Study 029. Patient accounting

	Placebo	MK-0966				Total
		5 mg	12.5 mg	25 mg	50 mg[a]	
ENTERED:	145	149	144	137	97	672
Male (age range)[b]	46 (44 to 89)	42 (46 to 78)	41 (45 to 92)	33 (45 to 78)	33 (40 to 86)	195 (40 to 92)
Female (age range)[b]	99 (39 to 83)	107 (39 to 80)	103 (38 to 85)	104 (39 to 83)	64 (42 to 82)	477 (38 to 85)
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
COMPLETED:	111 (76.6)	124 (83.2)	122 (84.7)	123 (89.8)	85 (87.6)	565 (84.1)
DISCONTINUED:	34 (23.4)	25 (16.8)	22 (15.3)	14 (10.2)	12 (12.4)	107 (15.9)
Clinical adverse experience	2 (1.4)	6 (4.0)	4 (2.8)	6 (4.4)	5 (5.2)	23 (3.4)
Laboratory adverse experience	0 (0.0)	0 (0.0)	1 (0.7)	1 (0.7)	0 (0.0)	2 (0.3)
Deviation from protocol	0 (0.0)	1 (0.7)	1 (0.7)	0 (0.0)	3 (3.1)	5 (0.7)
Patient withdrew consent	0 (0.0)	3 (2.0)	1 (0.7)	1 (0.7)	0 (0.0)	5 (0.7)
Lost to follow-up	3 (2.1)	0 (0.0)	1 (0.7)	0 (0.0)	0 (0.0)	4 (0.6)
Lack of efficacy	28 (19.3)	15 (10.1)	12 (8.3)	6 (4.4)	3 (3.1)	64 (9.5)
Other reasons[c]	1 (0.7)	0 (0.0)	2 (1.4)	0 (0.0)	1 (1.0)	4 (0.6)

[a] Patients with OA of the hip did not receive 50 mg.

[b] Birth date masking to 31DECXX in database, where XX is the true year of birth, may cause patient's age to appear 1 year younger than the actual age, depending on the date of randomization. All patients were ≥39 years of age when randomized.

[c] Includes reasons other than those listed.

Data Source: 14.22 to 4.27.201

APPEARS THIS WAY
ON ORIGINAL

A.6.2. Study 029. Primary endpoint results

Table D-11

Primary End Points
Average Over 6-Week Treatment Period
(Protocol 029)

Protocol	Baseline Mean	Treatment Period Mean	LS Mean' Difference From Baseline (95% CI)	LS Mean' Difference From Placebo (95% CI)
Pain Walking on a Flat Surface (WOMAC) (0- to 100-mm VAS)				
Placebo (n=139)	74.49	60.09	-17.50 (-21.26, -13.74)	N/A
MK-0966 5 mg (n=147)	74.35	44.66	-32.54 (-36.20, -28.88)	-15.04* (-20.29, -9.80)
MK-0966 12.5 mg (n=143)	71.38	43.85	-31.78 (-35.49, -28.07)	-14.28* (-19.57, -9.00)
MK-0966 25 mg (n=135)	76.05	45.43	-33.00 (-36.82, -29.18)	-15.50* (-20.86, -10.14)
MK-0966 50 mg (n=97)	74.62	35.53	-41.09 (-45.60, -36.59)	-23.60** (-29.47, -17.73)
Patient Global Assessment of Response to Therapy (0- to 4-point Likert Scale)				
Placebo (n=140)	N/A	-1.17	-1.22 (-1.38, -1.05)	N/A
MK-0966 5 mg (n=145)	N/A	-1.91	-1.98 (-2.14, -1.82)	-0.76* (-0.99, -0.53)
MK-0966 12.5 mg (n=143)	N/A	-2.20	-2.24 (-2.41, -2.08)	-1.02*,** (-1.25, -0.79)
MK-0966 25 mg (n=135)	N/A	-2.19	-2.27 (-2.44, -2.10)	-1.05** (-1.29, -0.82)
MK-0966 50 mg (n=97)	N/A	-2.53	-2.55 (-2.75, -2.35)	-1.33** (-1.59, -1.08)
Investigator Global Assessment of Disease Status (0- to 4-point Likert Scale)				
Placebo (n=139)	3.01	2.27	-0.71 (-0.85, -0.56)	N/A
MK-0966 5 mg (n=146)	2.88	1.74	-1.19 (-1.33, -1.05)	-0.48* (-0.68, -0.28)
MK-0966 12.5 mg (n=143)	2.94	1.59	-1.37 (-1.51, -1.22)	-0.66* (-0.86, -0.46)
MK-0966 25 mg (n=135)	2.93	1.59	-1.36 (-1.51, -1.22)	-0.66* (-0.86, -0.45)
MK-0966 50 mg (n=96)	2.86	1.21	-1.68 (-1.86, -1.51)	-0.98** (-1.20, -0.76)

Decreasing values represent improvement.
* $p < 0.05$ versus 50 mg.
** $p < 0.05$ versus 5 mg.
$p < 0.001$ LS mean difference from placebo for all active-treatment groups.

[P029]

A.6.3 . Study 029. Secondary endpoints results.

Table D-12 Key Secondary End Points
Average Over 6-Week Treatment Period
(Protocol 029)

Group	Baseline Mean	Treatment Period Mean	LS Mean Difference From Baseline (95% CI)	LS Mean[†] Difference From Placebo (95% CI)
Physical Function Subscale (WOMAC) (0- to 100-mm VAS)				
Placebo (n=139)	68.36	61.45	-8.78 (-12.12, -5.45)	N/A
MK-0966 5 mg (n=147)	67.64	49.30	-20.23 (-23.47, -16.99)	-11.45* (-16.10, -6.80)
MK-0966 12.5 mg (n=143)	66.62	47.54	-21.30 (-24.59, -18.01)	-12.52* (-17.21, -7.84)
MK-0966 25 mg (n=135)	69.13	47.65	-22.92 (-26.31, -19.54)	-14.14* (-18.89, -9.39)
MK-0966 50 mg (n=97)	67.44	39.28	-30.05 (-34.05, -26.06)	-21.27** (-26.48, -16.07)
Stiffness Subscale (WOMAC) (0- to 100-mm VAS)				
Placebo (n=139)	72.40	64.19	-10.64 (-14.18, -7.10)	N/A
MK-0966 5 mg (n=147)	72.40	51.50	-23.04 (-26.49, -19.60)	-12.40* (-17.34, -7.46)
MK-0966 12.5 mg (n=142)	70.77	50.23	-23.34 (-26.85, -19.84)	-12.70* (-17.68, -7.72)
MK-0966 25 mg (n=135)	74.18	49.46	-26.26 (-29.86, -22.67)	-15.62* (-20.67, -10.58)
MK-0966 50 mg (n=96)	74.51	41.74	-33.92 (-38.19, -29.66)	-23.28** (-28.82, -17.74)
Patient Global Assessment of Disease Status (0- to 100-mm VAS)				
Placebo (n=139)	69.19	63.51	-7.85 (-11.51, -4.20)	N/A
MK-0966 5 mg (n=147)	70.02	47.07	-24.41 (-27.97, -20.86)	-16.56* (-21.66, -11.47)
MK-0966 12.5 mg (n=143)	67.56	45.40	-25.11 (-28.71, -21.51)	-17.26* (-22.39, -12.13)
MK-0966 25 mg (n=134)	70.23	43.45	-28.09 (-31.81, -24.37)	-20.24 (-25.45, -15.02)
MK-0966 50 mg (n=97)	68.52	36.86	-33.28 (-37.65, -28.91)	-25.43** (-31.13, -19.73)

Discontinuation Due to Lack of Efficacy (%)

Group	% Frequency at Week 6	Difference in Percent From Placebo (95% CI)
Placebo (n=139)	19.31	N/A
MK-0966 5 mg (n=147)	10.07	-9.24* (-17.38, -1.20)
MK-0966 12.5 mg (n=143)	8.33	-10.98 (-18.83, -3.12)
MK-0966 25 mg (n=135)	4.38	-14.93 (-22.21, -7.65)
MK-0966 50 mg (n=97)	3.09	-16.22** (-23.51, -8.93)

† Decreasing values represent improvement.
* $p<0.05$ versus 50 mg.
** $p<0.05$ versus 5 mg.
$p<0.001$ LS mean difference from placebo for all active-treatment groups.

[P029]

Appendix7.1. 010 and 029 "Integrated" analysis and "best fitted" curve.

Figure D-4

Pain Walking on a Flat Surface (WOMAC) (0- to 100-mm VAS)

Least Squares Mean Difference From Placebo in Average Change From Baseline Over the 6-Week Treatment Period (Protocols 010 and 029)

_____ = BEST FIT CURVE   ● =Protocol 010,   O =Protocol 029
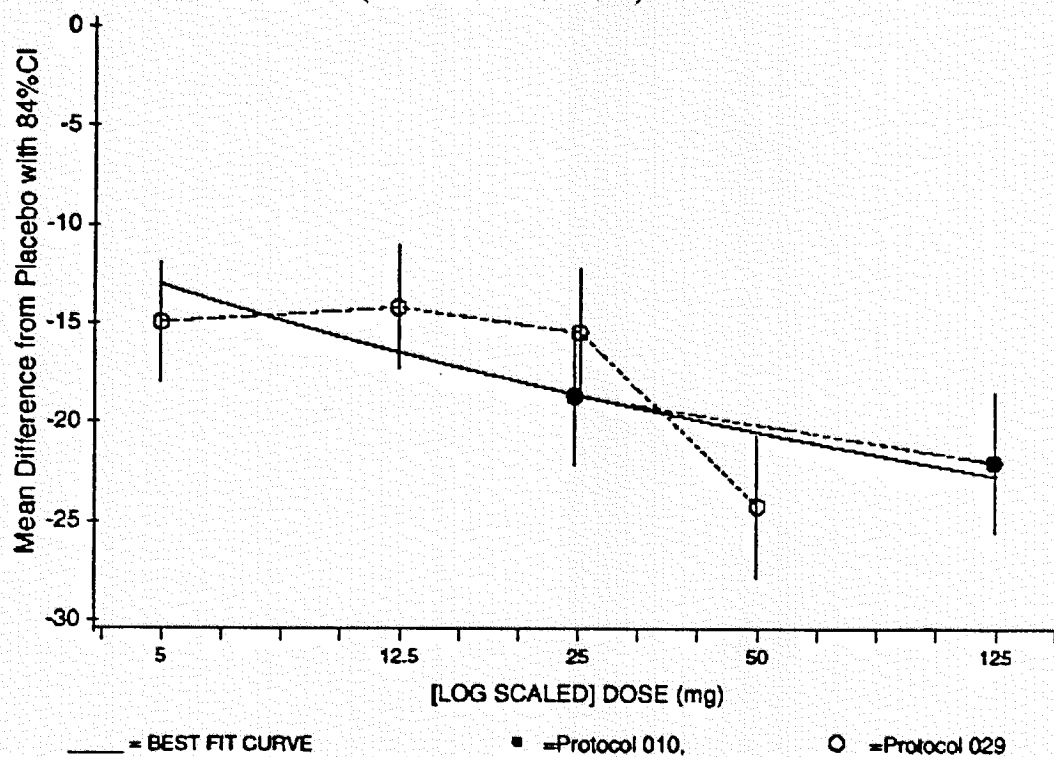
[304]

Similar curves were generated for Patient Global of Response to Therapy and Investigator Global of Disease Status.