CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

204623Orig1s000

STATISTICAL REVIEW(S)

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

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/BLA #:	NDA 204-623
Drug Name:	diclofenac sodium topical solution 2% w/w
Indication(s):	Treatment (b) (4) of osteoarthritis of the knee
Applicant:	Mallinckrodt, Inc.
Date(s):	Letter date: May 4, 2012, PDUFA date: March 4, 2012
Review Priority:	Standard
Biometrics Division:	II
Statistical Reviewer:	Feng Li, Ph.D.
Concurring Reviewers:	Dionne Price, Ph.D.
Medical Division:	Division of Anesthesia, Analgesia, and Addiction Products
Clinical Team:	Jacqueline Spaulding, M.D., Ellen Fields, M.D.
Project Manager:	Swati Patwardhan, Ph.D.
Keywords:	NDA review, Clinical Studies

Table of Contents

LIST OF TABLES

1. EXECUTIVE SUMMARY

Mallinckrodt, Inc. submitted a New Drug Application for diclofenac sodium topical solution 2% seeking an indication for the treatment (b) (4) of osteoarthritis (OA) of the knee. Diclofenac sodium 2% is purported to be a modified formulation of PENNSAID 1.5% topical solution, which is approved for the treatment of signs and symptoms of OA as a four times daily dosage. The desired indication usually requires evidence of efficacy from a 12-week study attaining statistical significance at level 0.05 on pain, physical function and patient's global assessment. A randomized, double-blind, and vehicle-controlled proof-of-concept efficacy study of four weeks duration was submitted to support the efficacy of diclofenac sodium 2% in comparison to vehicle control. There is evidence that diclofenac sodium 2% has an analgesic effect at Week 4. (b) (4)

The clinical development program of diclofenac sodium 2% was discussed at several meetings. Key areas of discussion included the necessity of a 12-week efficacy study and the need for a conservative imputation approach. Study COV05100031 (COV301) was an exploratory study powered to detect a difference in pain between diclofenac sodium 2% and the control with significance level at 0.1. The applicant claimed that the study was in agreement with the FDA guidance for diclofenac topical dosage forms and demonstrated that diclofenac sodium 2% was efficacious.

A total of 260 subjects were randomized equally to apply either diclofenac sodium 2% or vehicle control to the study knee for 4 weeks. Subjects completed the Western Ontario and McMaster Universities (WOMAC) questionnaire for the study knee at the baseline, Week 2 and Week 4 visits. The primary efficacy variable was the change in WOMAC pain subscale score from baseline to Week 4. WOMAC function and patient's global assessments (PGA) were included as secondary efficacy endpoints. The primary analysis was based on an analysis of covariance (ANCOVA) model with baseline score as a covariate and factors for treatment and whether the contralateral knee was treated. Missing WOMAC assessments for subjects who discontinued early due to lack of efficacy or adverse events were imputed using a baseline observation carried forward approach (BOCF). Missing WOMAC assessments due to other reasons were imputed using a last observation carried forward approach (LOCF). Sensitivity analyses were conducted using BOCF and LOCF approaches. For subjects who used rescue medication within 3 days of a study visit, the WOMAC and PGA scores were replaced using the baseline values regardless of discontinuation status or reason.

The diclofenac sodium 2% group had a numerically better response in WOMAC pain, function, and PGA than the control group at Week 4. However, only the difference in WOMAC pain achieved statistical significance at level 0.05. The results were not sensitive to the different imputation methods employed by the applicant as the overall dropout rate was about 8%.

2. INTRODUCTION

2.1 Overview

PENNSAID 1.5% topical diclofenac solution was approved in 2009 for treatment of the signs and symptoms of OA of the knee as a four times daily (QID) dosage. Mallinckrodt, Inc. developed the topical diclofenac 2% solution to improve the product's ease of use. Diclofenac sodium 2% is purported to be a modified formulation of PENNSAID 1.5% topical solution. The applicant seeks to market it as a twice daily (BID) dosage.

The clinical development program of diclofenac sodium 2% was discussed at several meetings under IND 75,045. At the End-of-Phase 2 meeting on September 25, 2006, (b) (4) The division reiterated the necessity of a 12-week study for the proposed indication and the need for a conservative imputation approach. The sponsor was also advised that demonstration of efficacy of diclofenac sodium 2% in pain, physical function, and patient's global assessment are required (b) (4). In 2009, Mallinckrodt, Inc. entered into a licensing agreement with Nuvo to pursue the development of diclofenac sodium 2%.

Study COV031 was initially designed as an exploratory study. The study was powered at 80% for the WOMAC pain endpoint with significance level 0.1. The applicant surmised that the study design was consistent with the draft guidance for topical diclofenac sodium and demonstrated that diclofenac sodium 2% was efficacious and well tolerated. The applicant purported that the use of data from Study COV031 to support the desired indication was justified given the new formulation was simply a modified version of the marketed product. The applicant claimed that there was no reason to suspect a non-steroidal anti-inflammatory drug having an effect at Week 4 would subsequently fail to have an effect at later weeks or would have an effect on pain but not on related endpoints.

2.2 Data Sources

The statistical review is based on data submitted for Study COV031.

The raw study data can be found at \\Cdsesub1\evsprod\NDA204623\0000\m5\datasets\cov05100031\tabulations\sdtm. The analyses datasets submitted during the review process per the division's request can be accessed at \\Cdsesub1\evsprod\NDA204623\0005\m5\datasets\cov05100031\analysis\adam\datasets.

3. STATISTICAL EVALUATION

3.1 Data and Analysis Quality

The applicant only submitted the raw datasets in the initial submission. The efficacy datasets and the program code were submitted during the review process per the division's request. Data entry errors were identified by Office of Scientific Investigations at two sites during inspection. These errors were found to be minor and had negligible effects on the final results.

3.2 Evaluation of Efficacy

3.2.1 Study Design and Endpoints

Study COV031 was a Phase 2, randomized, double-blind, multicenter, parallel, and vehicle-controlled study evaluating the efficacy and safety of a diclofenac sodium 2% BID for the treatment of the signs and symptoms of OA of the knees. Eligible subjects were randomized equally to apply either diclofenac sodium 2% or vehicle control to the study knee for 4 weeks. The non-study knee (contralateral knee) was also treated if painful at any point during the study. The first application of the study drug occurred at the study center on Day 1. The study drug was thereafter administered on an outpatient basis.

Subjects reported pain intensity twice daily via an interactive voice response system (IVRS) using an 11-point Numerical Rating Scale (NRS) during the 4-week treatment period. Subjects reported the number of tablets of rescue medication (acetaminophen) taken in each 24-hour period, whether they took rescue medication within 4 hours of an IVRS pain assessment, and whether each dose of study drug was applied. Subjects completed the WOMAC questionnaire at baseline (Day 1 before treatment) and at the Weeks 2 and 4 visits. Each question of the WOMAC index was answered using a 5-point Likert scale. Even if both knees were treated, the WOMAC assessments were done for the study knee only. Subjects were instructed to stop taking rescue medication 3 days prior to the study center visits.

The primary efficacy variable was the change in WOMAC pain subscale score from baseline to Week 4. Secondary efficacy endpoints included the daily pain intensity evaluated using NRS, subject's global assessment (PGA), WOMAC function, WOMAC stiffness, and use of rescue medication.

3.2.2 Statistical Methodologies

The primary efficacy variable was analyzed using an analysis of covariance (ANCOVA) model with baseline score as a covariate and factors for treatment and whether the contralateral knee was treated. The ANCOVA model was determined through a backward elimination process with the initial model including terms for baseline pain score, gender, age (categorized as < 65 or not), Body Mass Index (< 30 or not), and whether the contralateral knee was treated. Through the backward elimination process, terms that were not significant (p-value>0.1) in the initial model were removed. The same ANCOVA model was also used to analyze the continuous secondary efficacy points. The primary analysis was conducted on a modified intention-to-treat (mITT)

population including all subjects who were randomized and applied at least one dose of study medication.

In the primary analysis, missing WOMAC assessments for subjects who discontinued early due to lack of efficacy or adverse events were imputed using a baseline observation carried forward approach (BOCF). Missing WOMAC assessments due to other reasons were imputed using a last observation carried forward approach (LOCF). Sensitivity analyses were conducted based on a BOCF approach and a LOCF approach. For subjects who inappropriately used rescue medication within 3 days of a study visit, the WOMAC and PGA scores were replaced using the baseline values regardless of discontinuation status or reason.

The comparison between diclofenac sodium 2% and the vehicle control was conducted at the 0.10 significance level. The statistical analysis plan stated that this would be an exploratory study, and no adjustment would be made for multiple comparisons.

3.2.3 Patient Disposition, Demographic and Baseline Characteristics

A total of 260 subjects were randomized, 131 to PENNSAID and 129 to the vehicle control. One subject was randomized to PENNSAID in error and did not receive any study drug. All other randomized subjects were included in the mITT population. The subject disposition is shown in Table 1 with percentages based on the number of subjects in the mITT population. Overall, approximately 8% of the subjects discontinued early. The dropout rates of the diclofenac sodium 2% and control groups were 6% and 9% respectively. There were four subjects (3%) from the diclofenac sodium 2% group and five subjects (4%) from the control group who discontinued because of adverse events. Two subjects (2%) from the control group dropped out due to lack of efficacy.

Table 1: Subject Disposition − Number (%) of Patients

	Diclofenac sodium 2%	**Vehicle control**	**Total**
Randomized	131	129	260
Received treatment (mITT)	**130**	**129**	**259**
Terminated prematurely	**8 (6%)**	**12 (9%)**	**20 (8%)**
Reason for discontinuation			
Adverse event	4 (3%)	5 (4%)	9 (3%)
Subject withdrew consent	1 (1%)	2 (2%)	3 (1%)
Lack of efficacy	0	2 (2%)	2 (1%)
Protocol non-compliance	1 (1%)	1 (1%)	2 (1%)
Withdrawal of subject by Sponsor	1 (1%)	0	1
Other	1 (1%)	2 (2%)	3 (1%)

Source: Clinical study report.

The demographic and baseline characteristics were similar between the two treatment groups. A summary of selected demographic and baseline characteristics is provided in Table 2. The majority of the subjects were female (65% in diclofenac sodium arm and 70% in vehicle control). Overall, the mean age was about 60 years. Approximately 85% of the subjects were white.

Table 2: Summary of Demographics and Baseline Characteristics

	Diclofenac sodium 2% N=130	**Vehicle N=129**
Mean age (SD)	60 (9)	62 (9)
Mean weight (SD) (kg)	90 (23)	90 (21)
Mean height (SD) (cm)	170 (10)	168 (10)
Mean BMI (SD) (kg/m^2)	33 (8)	32 (7)
Gender, n (%)		
Male	46 (35%)	39 (30%)
Female	84 (65%)	90 (70%)
Ethnicity, n (%)		
Hispanic or Latino	2 (1%)	0
Not Hispanic or Latino	128 (98%)	129 (100%)
Race, n(%)		
Asian	1 (1%)	0
Black	19 (15%)	21 (16%)
White	110 (85%)	108 (84%)

Source: Clinical study report; SD: standard deviation.

3.2.4 Results and Conclusions

I replicated the applicant's results for the primary efficacy analyses. Table 3 shows the results for WOMAC assessments from an ANCOVA model with terms including treatment, whether the non-study knee was treated, and the baseline value. The ANCOVA model was determined based on the pre-specified backward selection procedure. A negative value of change from baseline indicated an improvement as compared to baseline. The maximum possible values of the WOMAC pain and functions endpoints were 20 and 68 respectively as both were presented as the sum of responses to a number of questionnaires. The diclofenac sodium 2% group had numerically better response in WOMAC pain, function, and patient's global impression than the control group. However, only the difference in WOMAC pain achieved statistical significance at level 0.05.

The protocol stated that a subject could initiate the treatment for the contralateral knee if painful anytime during the study. Thus, I was initially concerned about the inclusion of the treatment

status of the contralateral knee as a factor in the ANCOVA model as the decision to initiate treatment on the contralateral knee could occur after baseline and be affected by the treatment received. My concern was alleviated after I found that only four subjects initiated treatment on the contralateral knee after baseline.

Table 3: Efficacy Results at Week 4

		Least Square Means (Standard Error)			
Endpoint		**Diclofenac sodium 2% (N=130)**	**Vehicle (N=129)**	**Difference (95% CI)**	**P-value**
WOMAC pain	Baseline mean (SD)	12.4 (3.1)	12.6 (3.4)		
	Change from Baseline	-4.4 (0.4)	-3.4 (0.4)	-1.0 (-2.0, 0.0)	0.04
WOMAC function	Baseline mean (SD)				(b) (4)
	Change from Baseline				
Patient's global	Baseline mean (SD)				
	Change from Baseline				

SD: standard deviation; CI: confidence interval.
Source: Clinical Study Report.

The applicant conducted sensitivity analyses using BOCF and LOCF as imputation approaches for dropouts. The results from the sensitivity analyses were similar to those of the primary analysis. All the imputation approaches considered by the applicant were single imputation approaches. The report released by the National Academy of Science (NAS) in 2010 on missing values doesn't recommend single imputation approaches for imputing missing values due to dropouts. In addition, the report emphasizes the necessity to pre-specify the causal estimand in the protocol for a confirmatory trial. The appropriateness of an estimation method depends on the estimand. It is unclear from the protocol what estimand was being estimated by the proposed method. But since the dropout rate was low in both treatment groups, I was not overly concerned about the appropriateness of the imputation approach.

Table 4: Percentage of Subjects Using Any Rescue Medication

Period	**Diclofenac sodium 2% (N=130)**	**Vehicle Control (N=129)**	**P-value [1]**
Week 1 and Week 2	89 (68%)	94 (73%)	0.4
Week 3 and Week 4	68 (52%)	82 (64%)	0.07
Overall	95 (73%)	99 (77%)	0.5

[1] Chi-Square test;
Source: Clinical Study Report

Table 4 presents the percentages of subjects who took rescue during different time periods. Overall, the percentage of subjects who took rescue during the study was similar between the treatment groups. More subjects in the control group than the diclofenac sodium group took

rescue after Week 2. I found that approximately 21% of the patients in the control group and 12% of the patients in the diclofenac sodium 2% took rescue within 3 days of Week 4 visit. For these subjects, the Week 4 WOMAC assessments were replaced with baseline values per the pre-specified primary analysis. I found that if the observed WOMAC assessments were used instead of the baseline values the difference between the groups in the primary endpoint would not be significant, indicating that the control group might have had favorable responses after taking rescue.

3.3 Evaluation of Safety

The evaluation of the safety data was conducted by Dr. Jacqueline Spaulding. The reader is referred to Dr. Spaulding's review for detailed information regarding the adverse event profile. There were no deaths or serious adverse events reported during the study. Generally, the incidence of treatment-related adverse events was similar between treatment groups.

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

The applicant stated that age, sex and BMI were tested as cofactors for the primary efficacy outcome and were found to be non-significant. In addition, there were only about 20 subjects in each group were non-White. Thus, subgroup analyses for these factors were not provided by the applicant. I evaluated subgroups defined by sex, age, BMI and the treatment status of the contralateral knee. The findings in subgroups were generally consistent with the overall population.

4.1 Gender and Age

Table 5 shows the subgroup summaries for gender and age.

Table 5: Subgroup Summaries of Efficacy Results (Part I)

Endpoint	Subgroups	Statistics	Diclofenac sodium 2% (N=130)	Vehicle Control (N=129)
WOMAC pain	Sex			
	Female	n (%)	84 (65%)	90 (70%)
		Mean (SD)	-4 (4)	-3 (4)
	Male	n (%)	46 (35%)	39 (30%)
		Mean (SD)	-5 (5)	-4 (4)
	Age			
	<65	n (%)	95 (73%)	78 (60%)
		Mean (SD)	-5 (5)	-4 (4)
	>=65	n (%)	35 (27%)	51 (40%)
		Mean (SD)	-3 (4)	-3 (4)
WOMAC function	Sex			(b) (4)
	Female	n (%)		
		Mean (SD)		
	Male	n (%)		
		Mean (SD)		
	Age			
	<65	n (%)		
		Mean (SD)		
	>=65	n (%)		
		Mean (SD)		
Patient's global	Sex			
	Female	n (%)		
		Mean (SD)		
	Male	n (%)		
		Mean (SD)		
	Age			
	<65	n (%)		
		Mean (SD)		
	>=65	n (%)		
		Mean (SD)		

4.2 Other Special/Subgroup Populations

Subgroup summaries by BMI and treatment status of the contralateral knee are presented in Table 6. It appears that subjects who had the contralateral knee treated had better responses in both treatment groups.

Table 6: Subgroup Summaries of Efficacy Results (Part II)

Endpoint	Subgroups	Statistics	Diclofenac sodium 2% (N=130)	Vehicle Control (N=129)
WOMAC pain	BMI			
	<30	n (%)	56 (43%)	57 (44%)
		Mean (SD)	-5 (4)	-3 (4)
	>=30	n (%)	74 (57%)	72 (56%)
		Mean (SD)	-5 (5)	-4 (4)
	Contralateral Knee			
	Not treated	n (%)	48 (37%)	39 (30%)
		Mean (SD)	-3 (4)	-3 (3)
	Treated	n (%)	82 (63%)	90 (70%)
		Mean (SD)	-5 (5)	-4 (5)
WOMAC function	BMI			
	<30	n (%)		(b) (4)
		Mean (SD)		
	>=30	n (%)		
		Mean (SD)		
	Contralateral Knee			
	Not treated	n (%)		
		Mean (SD)		
	Treated	n (%)		
		Mean (SD)		
Patient's global	BMI			
	<30	n (%)		
		Mean (SD)		
	>=30	n (%)		
		Mean (SD)		
	Contralateral Knee			
	Not treated	n (%)		
		Mean (SD)		
	Treated	n (%)		
		Mean (SD)		

5. SUMMARY AND CONCLUSIONS

5.1 Statistical Issues

Statistical issues identified included the use of a potential post-baseline factor in the analysis model, the lack of an adjustment for multiplicity and the use of a single imputation method. The applicant included whether the contralateral knee was treated as a factor in the final ANCOVA model. I was initially concerned as the decision to initiate treatment on the contralateral knee could occur after baseline and be affected by the treatment received. My concern was alleviated after I found that the majority of subjects initiated treatment on the contralateral knee immediately after treatment on the study knee on the first day. In addition, more subjects in the control group had the contralateral knee treated. There was no multiplicity adjustment for comparisons on WOMAC pain, WOMAC function and patient's global. However, I was not

concerned given that statistical significance on the pain endpoint is always required for the pain indication, and it was achieved at level 0.05. Single imputation was used in the primary analysis. This is a concern, in general, especially when the causal estimand is not explicitly specified. I was not overly concerned as the dropout rate in the study was very low.

5.2 Collective Evidence

An indication (b) (4) of OA usually requires evidence of efficacy from a 12-week study attaining statistical significance at level 0.05 on pain, physical function and patient's global assessment as conveyed by the division during the development of the product. The diclofenac sodium 2% group had a numerically better response in WOMAC pain, physical function, and patient's global assessment than the control group at Week 4. However, only the difference in WOMAC pain achieved statistical significance at level 0.05. (b) (4)

There was some evidence of efficacy in treatment of pain up to Week 4. However, there was no data from the study to demonstrate that diclofenac sodium 2% was superior to the vehicle control in pain management beyond Week 4.

5.3 Conclusions and Recommendations

The proof-of-concept study has provided some evidence of analgesic efficacy at Week 4. (b) (4) The applicant has applied the FDA guidance for topical dosage forms to their development plan and uses this as justification for pursuing a different path than that previously recommended by the division. The review team will consequently need to determine if this guidance is applicable to diclofenac sodium 2%.

5.4 Labeling Recommendations

The applicant submitted the following wording to add to the clinical study section of the label for review:

(b) (4)

Table 4: Change in Treatment Outcomes after 4 Weeks of Treatment with PENNSAID (b) (4)

Efficacy Variable	Treatment	
	PENNSAID PUMP N=130	Vehicle Control N=129
WOMAC Pain Subscale		
Baseline	12.4	12.6
Mean Change from Baseline	-4.5	-3.6
(b) (4)		

I recommend the applicant reword the third sentences to read as follows: "In this trial, PENNSAID (b) (4) treatment resulted in improvement compared to vehicle in WOMAC pain. (b) (4) I also recommend the applicant delete (b) (4) Lastly, I recommend the addition of a footnote to the table as follows: WOMAC pain subscale is based on the sum of pain scores for five items using a 5-point Likert scale.

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

/s/

FENG LI
01/28/2013

DIONNE L PRICE
01/28/2013
Concur