

**CENTER FOR DRUG EVALUATION AND
RESEARCH**

APPLICATION NUMBER:

21-393 & 21-394

STATISTICAL REVIEW(S)



U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Pharmacoepidemiology and Statistical Science
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION CLINICAL STUDIES

NDA/Serial Number: NDA 21-393
Response to the December 18, 2003 approvable letter

Drug Name: Advil PM (Ibuprofen/Diphenhydramine HCl 200mg/25mg) Liquigels

Indication(s): Pain associated with sleeplessness

Applicant: Wyeth Consumer Healthcare

Date(s): Submitted: June 27, 2005
Received: June 27, 2005
PDUFA: December 26, 2005

Review Priority: Standard

Biometrics Division: Division of Biometrics II (HFD-715)

Statistical Reviewer: Yongman Kim, Ph.D.

Concurring Reviewers: Thomas Permutt, Ph.D.

Medical Division: Division of Anesthesia, Analgesia, and Rheumatology Products
Division of Neuropharmacological Drug Products
Division of Nonprescription Clinical Evaluation

Clinical Team: Elizabeth McNeil, M.D., DNDP

Project Manager: Jane Dean, DAARP

Keywords: ANOVA, CMH test with modified ridit scores

Table of Contents

LIST OF TABLES.....	3
LIST OF FIGURES.....	3
1. EXECUTIVE SUMMARY.....	4
1.1 CONCLUSIONS AND RECOMMENDATIONS	4
1.2 BRIEF OVERVIEW OF CLINICAL STUDY	4
1.3 STATISTICAL ISSUES AND FINDINGS.....	5
2. INTRODUCTION	5
2.1 OVERVIEW	5
2.2 DATA SOURCES	6
3. STATISTICAL EVALUATION.....	6
3.1 EVALUATION OF EFFICACY	6
3.2 EVALUATION OF SAFETY.....	8
4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS	8
5. SUMMARY AND CONCLUSIONS.....	8
5.1 STATISTICAL ISSUES AND COLLECTIVE EVIDENCE.....	8
5.2 CONCLUSIONS AND RECOMMENDATIONS	9
5.3 REVIEW OF CLINICAL STUDIES OF PROPOSED LABEL.....	10
APPENDIX.....	11
SIGNATURES/DISTRIBUTION LIST.....	18

LIST OF TABLES

Table 1. Patient Disposition by Treatment Group	11
Table 2. Patient Demographics and Baseline Pain Variables by Treatment Group (ITT).....	11
Table 3. Analysis of Primary Efficacy Variable: Total Sleep Time derived from Actigraphy (ITT)	12
Table 4. Reviewer Analysis of Primary Efficacy Variable after Logarithmic Transformation: Total Sleep Time derived from Actigraphy (ITT).....	12
Table 5. Reviewer’s Subgroup Analysis of Primary Efficacy Variable: Total Sleep Time derived from Actigraphy (ITT excluding Subjects on Rescue Medication).....	12
Table 6. Analysis of Secondary Efficacy Variables: Subjects’ Assessment of Sleep Duration (ITT).....	13
Table 7. Analysis of Secondary Efficacy Variables: Actigraph Assessment of Sleep Efficiency (ITT).....	13
Table 8. Analysis of Secondary Efficacy Variables: Actigraph Assessment of Wake after Sleep Onset (ITT)..	13
Table 9. Analysis of Secondary Efficacy Variables: Actigraph Assessment of Sleep Latency (ITT)	14
Table 10. Analysis of Secondary Efficacy Variables: Nurse Assessment of Sleep Latency (ITT).....	14
Table 11. Analysis of Secondary Efficacy Variables: Percentage of Taking Rescue Medication (ITT).....	14

LIST OF FIGURES

Figure 1. Schematic of Study Design: Study AE-04-14A.....	15
Figure 2. Box and Whisker Plot of Total Sleep Duration (excerpted from page 79 of the study report)	16
Figure 3. K-M Plot of Actigraph Assessment of Sleep Latency (excerpted from page 50 of the study report)...	17
Figure 4. K-M Plot of Time to Rescue Medication (excerpted from page 51 of the study report)	17

1. EXECUTIVE SUMMARY

1.1 Conclusions and Recommendations

Study AE-04-14A with subjects experiencing sleeplessness associated with nighttime oral surgery pain provided data showing better sleep efficacy of the combination of ibuprofen/diphenhydramine (IBU/DPH) 400/50 mg compared to ibuprofen (IBU) 400 mg alone based on intent-to-treat (ITT) analysis.

The following are primary and secondary efficacy results reported by the sponsor and confirmed by the reviewer:

- IBU/DPH provided a significantly longer duration of sleep derived by actigraph with difference of 1.2 hour ($p=.001$) as primary efficacy variable
- IBU/DPH provided a significantly longer duration of sleep based on subject's assessment ($p=.003$)
- IBU/DPH was significantly higher for the actigraphic measures of sleep efficiency ($p<.001$)
- IBU/DPH was significantly shorter for the actigraphic measures of wake after sleep onset (WASO) ($p<.001$)
- IBU/DPH had significantly fewer subjects taking rescue medications ($p=.031$)
- IBU/DPH had significantly longer time to rescue medication ($p=.020$)
- There were no significant differences between the IBU/DPH combination and IBU alone in sleep latency, as determined both objectively and with an observer

As mentioned in the December 18, 2003 approvable letter, these successful additional study results provide necessary data supporting sleep efficacy indication of IBU/DPH 400/50 mg combination.

1.2 Brief Overview of Clinical Study

In the approvable letter of December 18, 2003, FDA advised that an additional trial was needed to establish adequate evidence of the effect of ibuprofen/diphenhydramine combination on sleep duration and that the trial must be designed to determine the contribution of the components with sleep duration as the primary endpoint with objective measurement and with no forced awakenings. In the letter, FDA recommended using a full factorial design testing the combination against each component and placebo. However, the sponsor designed and conducted a study with two groups of ibuprofen/diphenhydramine combination and ibuprofen alone.

The currently submitted study AE-04-14A was intended to confirm the results of two studies completed as part of the IBU/DPH clinical program: AE-98-01 and AE-98-02 conducted to show sleep efficacy (duration) of the IBU/DPH combination in subjects who had pain due to oral surgery.

Study AE-04-14A was a randomized, stratified (by gender and baseline pain), inpatient, single-dose, double-blind, parallel-group, single-center trial to investigate the safety and effect of IBU/DPH 400mg/50mg combination on sleep duration in subjects experiencing sleeplessness due to oral surgery. The study used the actigraph to measure sleep duration objectively and was designed to have no artificial awakening. Three hundred twenty-nine subjects were randomized to IBU/DPH 400mg/50mg combination (n = 165) and IBU 400mg alone (n = 164). The primary efficacy endpoint was Total sleep duration derived from the actigraph. Secondary efficacy endpoints were Subjective assessment of sleep duration, Actigraphic assessments of sleep efficiency and wake after sleep onset (WASO), Sleep latency based on observation and actigraph, and Time to rescue medication and percentage requiring rescue by wake-up time.

1.3 Statistical Issues and Findings

Missing data issues were not prominent in study AE-04-14A because the dropout rate was very low (about 2%) and, also in algorithms of actigraphic total sleep duration, theoretically 'missing' data of sleep status after rescue medication was defined as 'awake'.

Actigraph-derived total sleep time as primary endpoint was statistically significantly longer for IBU/DPH combination when compared to IBU alone with a difference of 1.2 hour. Subjects' assessment of total sleep duration was also statistically significantly longer for IBU/DPH combination when compared to IBU alone. For other actigraph-derived secondary endpoints such as sleep efficiency, wake after sleep onset, IBU/DPH combination was superior to IBU alone. For other non-actigraph-derived secondary endpoints like subject's assessment of total sleep duration, nurse's assessment of sleep efficiency, IBU/DPH combination was superior to IBU alone. For actigraph-derived sleep latency, however, the difference between the two treatment groups was not statistically significant. IBU/DPH combination had significantly fewer subjects who took rescue medications and had significantly longer time to rescue medication than IBU alone.

2. INTRODUCTION

2.1 Overview

2.1.1 Drug class and regulatory history

The following are quotes from the submission regarding drug development plan and interactions between the sponsor and FDA prior to the current submission:

The sponsor developed ibuprofen/diphenhydramine HCl 200/25mg liquigels for over-the-counter use as an analgesic/sleep-aid combination product. Two partial-factorial, oral surgery

studies (AE-98-01 and AE-98-02) showed that the advantage of the combination is that ibuprofen relieves the pain over several hours, while the diphenhydramine allows for a full night's sleep.

The study AE-04-14A is intended to confirm the results of pivotal studies AE-98-01 and AE-98-02 by evaluating whether IBU/DPH provides better sleep efficacy (duration) compared to IBU alone in subjects who have pain due to oral surgery. Since it is clear that the combination is effective in relieving pain, and in order to address the Agency's concern about the impact of awakening subjects on the assessment of sleep duration, this study will not include any assessment of pain relief. Accordingly, subjects will not be awakened at any time after dosing. The Agency also requested that sleep efficacy be assessed objectively. Accordingly, actigraphic recordings, as well as subjective assessments, will be used to assess sleep efficacy. Total sleep time, as determined by actigraphy, will be the primary efficacy parameter.

This submission is basically the response to the FDA's 2nd approvable letter of December 18, 2003 regarding submission deficiencies. In the letter, FDA recommended that the sponsor conduct an additional study with objective measurement of sleep duration and no artificial awakening. FDA also noted that "clinically significant results from this study supporting your proposed indication, in addition to data presented in the original NDA, would be sufficient for approval."

2.1.2 Proposed Indication for ADVIL PM

The proposed indication for ADVIL PM in this NDA is for "relief of occasional sleeplessness when associated with _____ minor aches and pains; helps you get to sleep _____"

2.2 Data Sources

The electronic submission on June 27, 2005 can be found on the FDA, CDER electronic document room (EDR).

Final Report:

\\Cdsub1\21393\N_000\2005-06-27\clinstat

Data set:

\\Cdsub1\21393\N_000\2005-06-27\crt\datasets\ac-04-14a

3. STATISTICAL EVALUATION

3.1 Evaluation of Efficacy

3.1.1 Study Design and Endpoints

Study AE-04-14A was a randomized, stratified (by gender and baseline pain), inpatient, single-dose, double-blind, parallel-group, single-center trial to investigate the safety and effect of IBU/DPH 400mg/50mg combination on sleep duration in subjects experiencing sleeplessness associated with nighttime oral surgery pain. IBU 400 mg was the active treatment control. Subjects were randomized to IBU/DPH or IBU in 1:1 ratio.

Figure 1 in Appendix is a schematic of study design.

Two investigators enrolled subjects from a US site and participated in the clinical trial.

The primary efficacy endpoint was the total sleep time derived from the actigraph.

The secondary efficacy endpoints were Subjective assessment of sleep duration, Actigraphic assessment of sleep efficiency, Wake after sleep onset (WASO), Sleep latency based on observation, Sleep latency based on actigraph, Time to rescue medication, and Percentage requiring rescue by wake-up time.

3.1.2 Patient Disposition, Demographic and Baseline Characteristics

As shown in Table 1 in Appendix, a total of seven subjects were discontinued; two in the IBU/DPH group (one withdrew consent, one for a protocol violation), and five in the IBU group (all for protocol violations).

Table 2 in Appendix shows subject demographics by treatment groups. Demographic characteristics were comparable between treatment groups.

The table also shows baseline pain variables by treatment groups. Baseline pain variables were comparable between treatment groups.

3.1.3 Statistical Methodologies

Actigraphic measures of total sleep time, WASO, and Subject's assessment of sleep duration were compared between IBU/DPH and IBU using an ANOVA model with terms for treatment, sex, and baseline pain severity ratings (PSR), as well as via a CMH test using modified riddit scores as a sensitivity analysis due to possible departure from normality assumption. As a sensitivity analysis, I used logarithmic transformation of the total sleep time and fitted the same model as the sponsor's ANOVA model. For comparison of percentages of subjects who took rescue medication, a Mantel-Haenszel chi-square test was used; and for comparison of time to rescue medication, a proportional-hazard regression model with terms for treatment, sex, and baseline PSR was used.

3.1.4 Results and Conclusions

Tables 3–11 and Figures 2 - 4 in Appendix present the statistical analyses done by the sponsor and by me. Following are a summary of the analyses.

Data from the study showed the superiority of IBU/DPH combination to IBU alone in actigraphic measures of total sleep duration, sleep efficiency, wake after sleep onset (WASO) as well as subjective measures of the above parameters in subjects with sleeplessness after oral surgery. There was no statistically significant difference, however, in sleep latency between the two treatment groups. IBU/DPH combination had fewer subjects taking rescue medication and longer time to rescue medication than IBU alone.

The study succeeded in showing that the IBU/DPH combination is superior to IBU alone in sleep duration and sleep efficiency.

3.2 Evaluation of Safety

Safety analyses were done by the clinical reviewer, Elizabeth McNeil, M.D.

No statistical problems or issues were found.

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

Per request from Dr. McNeil, I conducted a subgroup analysis on only the subjects who had not taken rescue medications. The concern by the clinical reviewer was that, because there was a statistically significant difference in taking rescue medications between the treatment groups (28.5% for IBU/DPH vs. 39.6% for IBU, $p=.031$) and the actigraph was manually recorded as 'awake' from the time of taking rescue medication on, the significant significance shown in sleep duration could be driven by this difference. My subgroup analysis also gave a statistically significant difference, however, in the actigraph-derived total sleep duration between the two treatment groups ($p=.044$).

5. SUMMARY AND CONCLUSIONS

5.1 Statistical Issues and Collective Evidence

5.1.1 Statistical Issues

For the efficacy analysis, ITT set was used for the statistical inference on primary and secondary endpoints comparing IBU/DPH combination to IBU alone.

ITT set was defined as all randomized subjects who received at least one dose of study medication.

The sponsor conducted a non-parametric method of CMH with modified ridit scores as a sensitivity analysis with respect to a potential departure from the normality assumption of total sleep duration as well as ANOVA assuming normality. I conducted ANOVA analysis after logarithmic transformation on the total sleep duration. Statistical significance shown in three analyses gave me a reassurance of the sponsor's conclusion on effectiveness with respect to the normal distribution assumption.

Missing data issue was not prominent in the study AE-04-14A because the dropout rate was very low (about 2%), and also in algorithms of actigraphic total sleep duration, theoretically 'missing' data of sleep status after rescue medication was defined as 'awake'.

5.1.2 Collective Evidence

The data from the new study AE-04-14A provided statistically significant evidence of efficacy of IBU/DPH 400/50 mg as sleep aid. The study met the pre-planned objective and agreement between the sponsor and FDA that an additional successful study would be sufficient in addition to data provided previously, especially two studies AE-98-01 and AE-98-02 in the original submission. Furthermore, my sensitivity analysis with respect to departure from the normality assumption and subgroup analysis on subjects who had not taken rescue medication reinforce the efficacy of IBU/DPH combination product.

5.2 Conclusions and Recommendations

Study AE-04-14A with subjects experiencing sleeplessness associated with nighttime oral surgery pain provided data showing better sleep efficacy of the combination of IBU/DPH 400/50 mg compared to IBU 400 mg alone based on ITT analysis.

The following are primary and secondary efficacy results reported by the sponsor and confirmed by the reviewer:

- IBU/DPH provided a significantly longer duration of sleep derived by actigraph with difference of 1.2 hour ($p=.001$) as primary efficacy variable
- IBU/DPH provided a significantly longer duration of sleep based on subject's assessment ($p=.003$)
- IBU/DPH was significantly higher for the actigraphic measures of sleep efficiency ($p<.001$)
- IBU/DPH was significantly shorter for the actigraphic measures of wake after sleep onset (WASO) ($p<.001$)
- IBU/DPH had significantly fewer subjects taking rescue medications ($p=.031$)
- IBU/DPH had significantly longer time to rescue medication ($p=.020$)

- There were no significant differences between the IBU/DPH combination and IBU alone in sleep latency, as determined both objectively and with an observer

As mentioned in the December 18, 2003 approvable letter, these successful additional study results provide necessary data supporting sleep efficacy indication of IBU/DPH 400/50 mg combination.

5.3 Review of Clinical Studies of Proposed Label

Following is the text portion in the _____ section from 'PROPOSED LABELING TEXT':

Data from the study AE-04-14A supports the claim _____ as there was a statistically significant difference in total sleep duration variable, but does not support the claim “helps you get to sleep” as there was no statistically significant difference in the sleep latency variable between IBU/DPH combination and IBU alone.

APPENDIX

Table 1. Patient Disposition by Treatment Group

	IBU/DPH 400/50 MG	IBU 400 MG	TOTAL
RANDOMIZED:	165	164	329
ITT:	165	164	329
COMPLETED, n (%):	163(98.8)	159 (97.0)	322 (97.9)
DISCONTINUED, n (%):	2(1.2)	5 (3.0)	7 (2.1)
Protocol Violation	1	5	6
Consent Withdrawn	1	0	1

Table 2. Patient Demographics and Baseline Pain Variables by Treatment Group (ITT)

	IBU/DPH 400/50 MG (N =165)		IBU 400 MG (N = 164)		p-value
	n	%	n	%	
Age (years)					
Mean ± SD	18.7 ± 2.4		10.0 ± 3.1		0.319
Range	16 - 28		16 - 37		
Gender					
Male	81	49.1	81	49.4	0.961
Female	84	50.9	83	50.6	
Race					
Caucasian	149	90.3	151	92.1	0.584
Black	2	1.2	1	0.6	
Asian	2	1.2	4	2.4	
Hispanic	11	6.7	6	3.7	
Other	1	0.6	2	1.2	
Weight (lbs.)					
Mean ± SD	155.1 ± 33.5		151.4 ± 33.0		0.227
Range	100 – 280		97 – 295		
Baseline Pain					
Pain Intensity (VAS) mm.	76.2 ± 13.2		76.4 ± 14.2		0.873
Pain Severity (Cat)					
Moderate	90	54.5	90	54.9	0.956
Severe	75	45.5	74	45.1	

Table 3. Analysis of Primary Efficacy Variable: Total Sleep Time derived from Actigraphy (ITT)

Total Sleep Time (hour)		
	IBU/DPH 400/50 mg (n=165)	IBU 400 mg (n=164)
LSMean (SE)	9.23 (.26)	8.03 (.26)
Difference (95% CI)	1.20 (.49, 1.91)	
p-value*	.0010	
p-value**	.0007	

*LSMeans and p-values calculated from ANOVA model: $Y = \text{trt} + \text{sex} + \text{baseline PSR}$.

**CMH test using modified ridit scores controlling for sex and baseline PSR

Table 4. Reviewer Analysis of Primary Efficacy Variable after Logarithmic Transformation: Total Sleep Time derived from Actigraphy (ITT)

Total Sleep Time (hour)		
	IBU/DPH 400/50 mg (n=165)	IBU 400 mg (n=164)
Geometric LSMean	8.10	6.88
Ratio (95% CI)	1.18 (1.02, 1.36)	
p-value*	.0269	

*LSMeans and p-values calculated from ANOVA model: $\log Y = \text{trt} + \text{sex} + \text{baseline PSR}$.

Table 5. Reviewer's Subgroup Analysis of Primary Efficacy Variable: Total Sleep Time derived from Actigraphy (ITT excluding Subjects on Rescue Medication)

Total Sleep Time (hour)		
	IBU/DPH 400/50 mg (n=165)	IBU 400 mg (n=164)
LSMean (SE)	10.91 (.15)	10.47 (.16)
Difference (95% CI)	0.44 (.01, .86)	
p-value*	.0010	
p-value**	.0007	

*LSMeans and p-values calculated from ANOVA model: $Y = \text{trt} + \text{sex} + \text{baseline PSR}$.

Table 6. Analysis of Secondary Efficacy Variables: Subjects' Assessment of Sleep Duration (ITT)

Subjects' Assessment of Sleep Duration (hour)		
	IBU/DPH 400/50 mg (n=165)	IBU 400 mg (n=164)
LSMean (SE)	7.90 (.25)	6.86 (.25)
Difference (95% CI)	1.04 (.36, 1.73)	
p-value*	.0030	
p-value**	.0049	

*LSMeans and p-values calculated from ANOVA model: $Y = \text{trt} + \text{sex} + \text{baseline PSR}$.

**CMH test using modified ridit scores controlling for sex and baseline PSR

Table 7. Analysis of Secondary Efficacy Variables: Actigraph Assessment of Sleep Efficiency (ITT)

Sleep Efficiency (% of time spent asleep while in bed)		
	IBU/DPH 400/50 mg (n=165)	IBU 400 mg (n=164)
Mean (STD)	75.9 (24.9)	65.7 (27.5)
Difference (95% CI)	10.2 (4.5, 15.9)	
p-value*	<.0001	

*CMH test using modified ridit scores controlling for sex and baseline PSR

Table 8. Analysis of Secondary Efficacy Variables: Actigraph Assessment of Wake after Sleep Onset (ITT)

WASO (hour)		
	IBU/DPH 400/50 mg (n=165)	IBU 400 mg (n=164)
LSMean (SE)	2.33 (.25)	3.68 (.25)
Difference (95% CI)	-1.35 (-2.05, -.65)	
p-value*	.0002	

*LSMeans and p-values calculated from ANOVA model: $Y = \text{trt} + \text{sex} + \text{baseline PSR}$.

Table 9. Analysis of Secondary Efficacy Variables: Actigraph Assessment of Sleep Latency (ITT)

Sleep Latency (minutes, time until sleep onset)		
	IBU/DPH 400/50 mg (n=165)	IBU 400 mg (n=164)
Median (Range)	23 (4, 144)	23 (3, 199)
Hazard Ratio	1.12	
p-value*	.3317	

*p-values calculated from phreg model: $Y = \text{trt} + \text{sex} + \text{baseline PSR}$.

Table 10. Analysis of Secondary Efficacy Variables: Nurse Assessment of Sleep Latency (ITT)

Sleep Latency (minutes, time until sleep onset)		
	IBU/DPH 400/50 mg (n=165)	IBU 400 mg (n=164)
Median (Range)	20 (10, 180)	20 (10, 180)
Hazard Ratio	1.02	
p-value*	.8333	

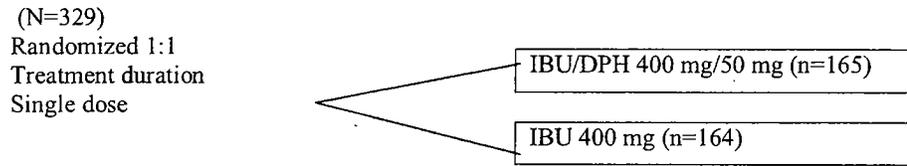
*p-values calculated from phreg model: $Y = \text{trt} + \text{sex} + \text{baseline PSR}$.

Table 11. Analysis of Secondary Efficacy Variables: Percentage of Taking Rescue Medication (ITT)

Percentage of subjects taking rescue medication		
	IBU/DPH 400/50 mg (n=165)	IBU 400 mg (n=164)
Frequency (%)	47 (28.5)	65 (39.6)
p-value*	.0310	

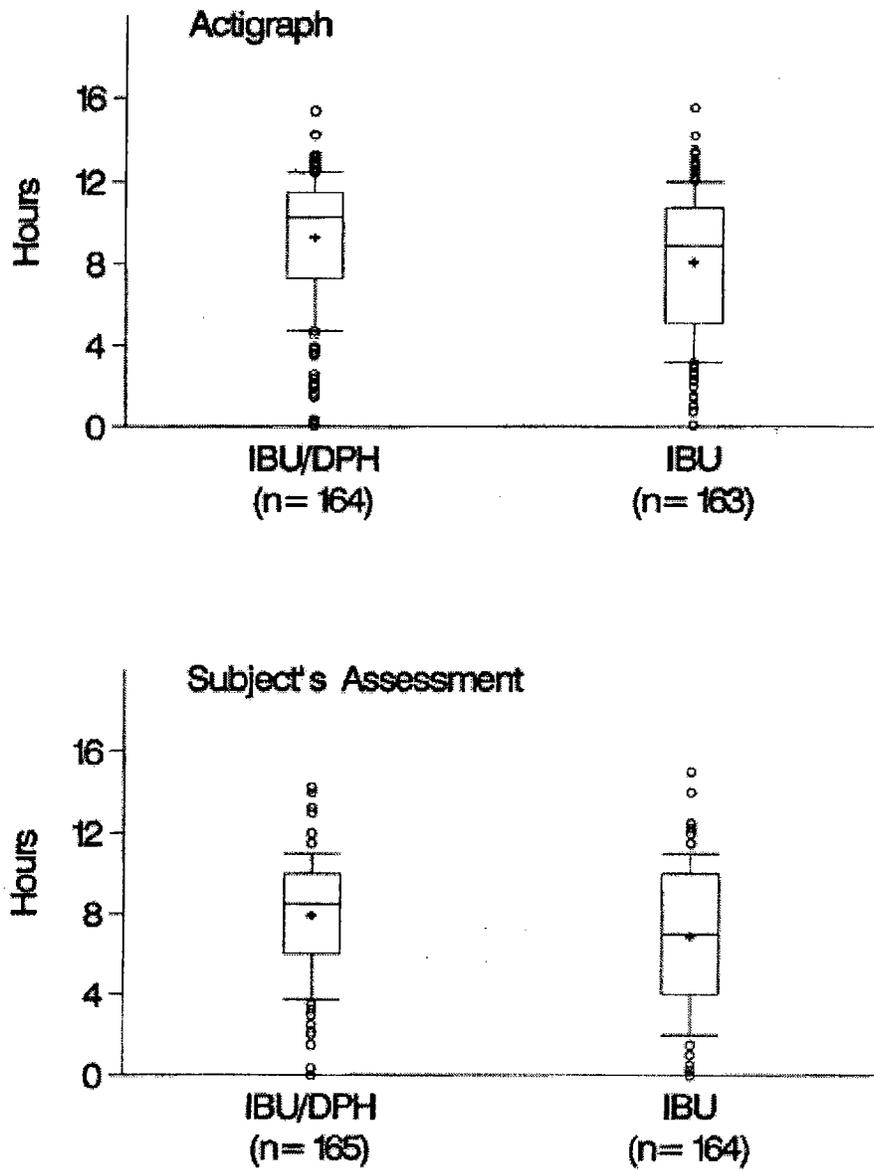
*p-values calculated from Mantel-Haenszel chi-square test.

Figure 1. Schematic of Study Design: Study AE-04-14A



**APPEARS THIS WAY
ON ORIGINAL**

Figure 2. Box and Whisker Plot of Total Sleep Duration (excerpted from page 79 of the study report)



Note: The boxes show the 75th and 25th percentiles. The line within the box is the 50th percentile (median) and the '+' is the mean. Top (bottom) lines represent the 90th (10th) percentiles.

Figure 3. K-M Plot of Actigraph Assessment of Sleep Latency (excerpted from page 50 of the study report)

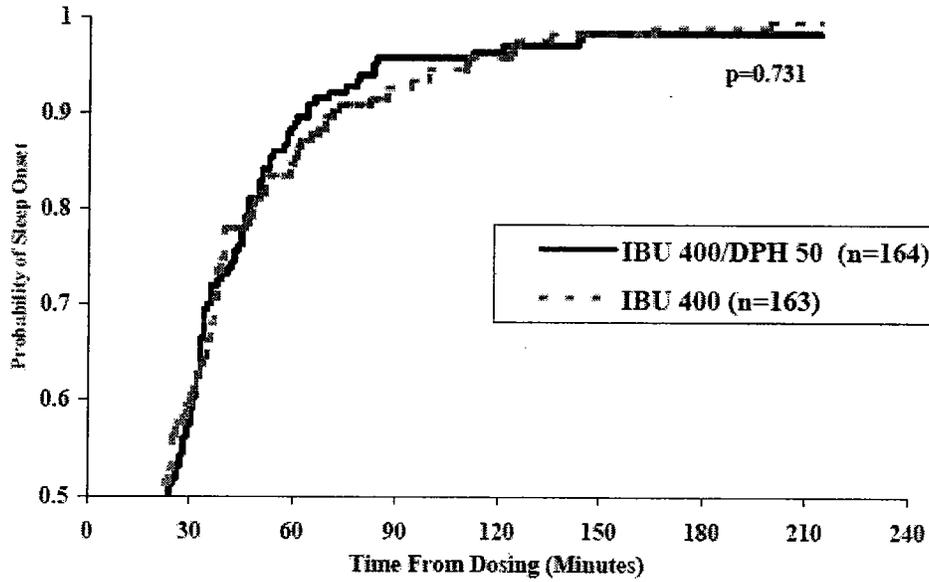
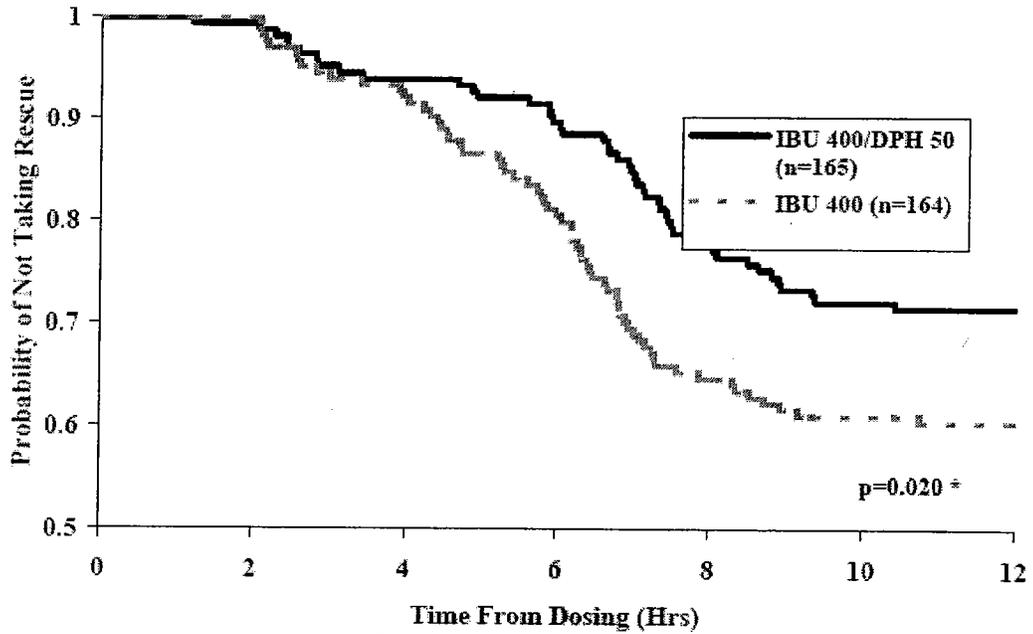


Figure 4. K-M Plot of Time to Rescue Medication (excerpted from page 51 of the study report)



* Significantly better than the IBU 400 group

SIGNATURES/DISTRIBUTION LIST

Primary Statistical Reviewer: Yongman Kim, Ph.D.
Mathematical Statistician

Date: November 30, 2005

Concurring Reviewer: Thomas Permutt, Ph.D.
Statistical Team Leader

cc:

DAARP/Jane Dean
DNDP/Elizabeth McNeil, M.D.
DNDP/John Feeney III, M.D.
HFD-715/Yongman Kim, Ph.D.
HFD-715/Thomas Permutt, Ph.D.
HFD-715/Edward Nevius, Ph.D.
HFD-700/Charles Anello, Ph.D.

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

/s/

Yongman Kim
12/1/2005 08:46:58 AM
BIOMETRICS

Thomas Permutt
12/1/2005 09:25:27 AM
BIOMETRICS
concur

Statistical Review Amendment

NDA: 21393

Name of drug: Advil PM Liqui-Gel (Ibuprofen 400 mg/ diphenhydramine hydrochloride 50 mg)

Applicant: Whitehall-Robins Healthcare

Indication: Relief of occasional sleeplessness when associated with
minor aches and pains

Documents reviewed: Statistical Section of Electronic NDA Submission
(pathway: \\CDSESUB1\N21393\N_000\2003-06-30\clinstat)

Project manager: Jane Dean

Clinical reviewer: Tatiana Oussova, M.D.

Dates: Received 6/30/03

Statistical reviewer: Laura Lu, Ph.D.

Statistics team leader: Stan Lin, Ph.D.

Biometrics division director: Mo Huque, Ph.D.

Table of Contents

1. Introduction..... 2

2. Sponsor’s Response regarding the Influence of awakening Subjects 2

3. Reviewer’s Comments 3

 3.1 About Sponsor’s response regarding awakening subjects..... 3

 3.2 Additional comments..... 4

4. Final conclusion..... 5

1. Introduction

NDA21-393 was originally submitted on 10/16/2001. An approvable letter was issued to the sponsor by the Agency on 8/8/02 regarding this NDA. On 2/5/2003, the sponsor requested for formal dispute resolution at the level of Office of New Drug (OND). Dr. Jenkins, director of OND, wrote a review regarding the sponsor’s request. In his review of administrative records and primary reviews performed by OTC and DAAODP of the NDA 21-393, Dr. Jenkins concurred with the Divisions’ decision to issue an approvable letter. The purpose of the current submission from the sponsor is to address Dr. Jenkins concern regarding the influence of awakening subjects and categorical data for sleep duration. This reviewer will review the sponsor’s response regarding the influence of awakening subjects and provide further comments on efficacy of Advil PM Liqui-Gel. Dr. Kun Jin will provide a separate review regarding the influence of categorical data for sleep duration.

2. Sponsor’s Response regarding the Influence of Awakening Subjects

Two additional analyses of the data from studies AE 98-01 and AE 98-02 were conducted by the sponsor to show that the overall finding that diphenhydramine contributes to sleep duration was a real finding and not an artifact of the design of awakening subjects. First, the sponsor examined the subjects in both studies who had to be awakened at 120 minutes (i.e., those who were asleep at that time point) to assess their pain. These subjects were analyzed to see whether awakening them had an effect on whether they were asleep at 150 minutes, the first observation time point after the scheduled awakening at 120 minutes. As shown in Table 1, the vast majority of subjects in both the IBU and IBU-DPH groups who were awakened at 120 minutes were again asleep at 150 minutes.

Table 1. Subjects awakened at 120 minutes and then asleep at 150 minutes

	Awakened at 120 minutes	Sleep resumed at 150 minutes (%*)
AE 98-01		

IBU-DPH (n=122)	61	57 (93%)
IBU (n=118)	54	48 (89%)
AE 98-02		
IBU-DPH (n=119)	95	87 (92%)
IBU (n=123)	72	64 (89%)

*Based on those awakened at 120 minutes

The sponsor also analyzed the subgroup of subjects in both studies asleep at 150 minutes for sleep duration. In study AE 98-01, 64% (n=76) of the subjects in the IBU-DPH arm and 57% (n=67) of the subjects in the IBU arm were asleep at 150 minutes. In study AE 98-02, 82% (n=97) of the subjects in the IBU-DPH arm and 63% (n=78) of the subjects in the IBU arm were asleep at 150 minutes. Table 2 summarizes the results of this analysis.

Table 2. Duration of sleep for subjects asleep at 150 minutes.

	Sample Size	Mean Duration of Sleep*	p-value
AE 98-01			
IBU-DPH	76	3.62	
IBU	67	2.97	0.03
AE 98-02			
IBU-DPH	97	2.96	
IBU	78	2.37	0.03

* Reported by category (0=<5 hours, 1=>5-6 hours, 2=>6-7 hours, 3=>7-8 hours, 4=>8-9 hours, 5=>9 hours)

For the subjects who were asleep at 150 minutes, those in the IBU-DPH arm had greater total sleep than those in the IBU alone arm in both studies. The sponsor argued that this difference in sleep duration is due to the DPH in the combination therapy.

3. Reviewer's Comments

3.1 About Sponsor's response regarding awakening subjects

To completely answer the question ‘whether awakening subjects leads to bias in comparing sleep duration’, we need data for sleep duration with and without awakening patients up, and the later is not available in the trials conducted by the sponsor. The sponsor tries to answer this question by asking another question: Is awakening subjects at 120 minutes had an effect on whether they were asleep at 150 minutes (see results in Table 1)? The later question can only probe the first question from a narrow angle in terms of sleep latency at 150 minutes and it does not even explore the sleep latency at 120 minutes among the subjects awakened at 90 minutes. In Table 3 below, the reviewer provided information for proportion of subjects asleep at 120 minutes among the subjects awakened at 90 minutes. Tables 1 and 3 show that the proportions of subject who sleep at the next observation time point after being awakened up is consistently higher in the combination group than in the ibuprofen group in both Studies 98-01 and 98-02, especially after being awakened at 90 minutes in Study 98-02 with the difference as high as 18%. So results in Tables 1 and 3 do not provide evidence for the statement that awakening subjects does not lead to bias in comparing sleep duration.

Table 3. Subjects awakened at 90 minutes and then asleep at 120 minutes

	Awakened at 90 minutes	Sleep resumed at 120 minutes (%*)
AE 98-01		
IBU-DPH (n=122)	66	50 (75.8%)
IBU (n=118)	66	47 (71.2%)
AE 98-02		
IBU-DPH (n=119)	82	74 (90.2%)
IBU (n=123)	82	59 (72%)

*Based on those awakened at 90 minutes

Additionally, the results in Table 2 were from a partial dataset (those who were asleep at 150 minutes) from the two studies. Even among the subjects asleep at 150 minutes, the assessment for sleep duration may be influenced by both the awakening effect and drug effect since it is not clear which time point (before or after being awakened up) is counted as the starting point of sleep by patients. These subgroups may also represent a special population due to the selection criteria and the results may not reflect the true treatment effect in the general population. So results in Table 2 do not provide evidence for the statement that awakening subjects does not lead to bias in comparing sleep duration, either.

3.2 Additional comments

The sponsor did not address another concern raised in Dr. Jenkins' letter that the addition of diphenhydramine to ibuprofen decreased ibuprofen's efficacy. The combination drug was statistically significantly worse in a pain endpoint than ibuprofen alone in AE-98-02 and numerically worse in AE-97-01. Also, cumulative percentage of subjects who were asleep at 60 minutes (sleep latency endpoint) was numerically worse for the combination than for ibuprofen in AE-98-01 and AE-98-02. Since SPID and sleep latency were correlated endpoints (e.g., correlation coefficients were 0.43 and 0.49 at 90 minutes and 120 minutes among all patients at study AE 98-02), it is very likely that the combination drug worsened the patient's capability of falling into sleep due to the worsened analgesic effect compared to ibuprofen alone.

4. Final conclusion

The sponsor's response to Dr. Jenkins' letter did not provide evidence for the statement that awakening subjects does not lead to bias in comparing sleep duration. Also, the sponsor did not address Dr. Jenkins' concern that the addition of diphenhydramine to ibuprofen decreased ibuprofen's efficacy in terms of analgesic effect and sleep latency.

Laura Lu, Ph.D.

Mathematical Statistician

Concur: Stan Lin, Ph.D.

Statistics Team Leader

cc:

HFD-550 / Tatian Oussuva
HFD-550 / Jim Witter
HFD-550 / Jane Dean
HFD-725 / Laura Lu
HFD-725 / Stan Lin
HFD-725 / Mo Huque, Ph.D

HFD-700 / Chuck Annelo, Ph.D
HFD-002 / Satya Dubey, Ph.D

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

/s/

Hong Lu
11/4/03 11:34:28 AM
BIOMETRICS

Stan Lin
11/4/03 12:09:14 PM
UNKNOWN

NDA 21-393

**Biometrics
Review**

MEMORANDUM

Date: June 5, 2002
From: Kun Jin, Ph.D.
Statistical Team Leader, DBI/HFD-710
To: File, NDA 21-393
Subject: Comments on the primary analysis of Study 98-02

This memo is an addition to Dr. Paul Andreason's review for this NDA. Dr. Sharon Yan is the statistical reviewer for this NDA. Dr. Yan performed statistical analyses on the sleep parameters of this NDA and confirmed the sponsor's results in the submission. Dr. Yan and I have provided statistical comments to Dr. Andreason during his review. There is no statistical review written for this submission. Much of the subsequent discussions were focused on the sponsor's primary analysis result of sleep duration comparing IB 400-mg/DPH 50-mg arm with IB 400 alone arm in Study 98-02. This issue will be the only focus of this memo. Please see Dr. Andreason's review for more detailed information.

One of the primary sleep endpoints in Study 98-02 was sleep duration. A natural ordinal data for sleep duration should be the number of hours patients slept. The sponsor, instead, collected the following data, 0 for less than 5 hours, 1 for 5 to 6 hours, 2 for 6+ to 7 hours, 3 for 7+ to 8 hours, 4 for 8+ to 9 hours, and 5 for more than 9 hours. The collected patient's sleep duration data is displayed as follows.

**Percentage of Patients in Each Categorical Sleep Duration Group
Study 98-02**

	IBU400/DPH50 N=119	IBU400 N=123
Duration of Sleep		
<5 hours(0)	26 (21.8%)	41 (33.3%)
5 to 6 hours(1)	18 (15.1%)	18 (14.6%)
6+ to 7 hours(2)	12 (10.1%)	15 (12.2%)
7+ to 8 hours(3)	12 (10.1%)	14 (11.4%)
8+ to 9 hours(4)	23 (19.3%)	22 (17.9%)
>9 hours(5)	28 (23.5%)	13 (10.6%)
Mean	2.61	1.98
SD	1.92	1.8

The p-value from the CMH row mean test, controlling for baseline PSR and gender, with modified ridit scores, was 0.009. Dr. Yan calculated the CMH test using general association option; it gave a p-value of 0.099. The simple Chi-square test gave a p-value of 0.098. (In Dr. Andreason's review, this p-value was 0.2424, which I incorrectly reported to him.) The p-values from the simple Chi-square test and the CMH general association test were similar and non-significant. The CMH row mean test was the

protocol specified primary analysis. Notice from the data that the difference of the two groups apparently came from the truncated categories, namely, 0 (< 5 hours) and 5 (> 9 hours). One would like to ask whether such a truncation could probably introduce a bias. More specifically, if the complete sleep duration data were collected, would the CMH row mean test still yield a significant result? In Appendix, I argued that one could not assume this statement is generally true.

The sponsor's collection of data involves two transformations on the natural ordinal data: (a) Truncation of data below 5 or above 10 to 5 and 10, respectively; (b) Linear shift of data downward by 5 units. The CMH row mean test statistic is invariant under a location shift. Therefore, transformation (b) will not introduce a bias to the test result. The truncation above 10 seems to be reasonable since it is not likely to have many data more than 11 hours sleep time. It could be reasonably concluded that the upper end truncation is not materialized. So the real concern is left with the lower end truncation, namely, truncating scores 1,2,3,4, and 5 to the score 5.

In the Appendix, I demonstrated that in some situations, such a truncation could indeed introduce a bias to the population row mean. In the numerical example given in the Appendix, two groups with natural ordinal scores from 1 to 7 had the same population row mean of 4. After the lower-end truncating of 1,2 and 3 to 3, one group had a row mean of 4.551725, the other had a row mean of 4.344828. The truncation introduced a difference between the two group's row means. Since the CMH row mean test is advocated as a sensible test to target such differences, the bias introduced by the truncation is particularly disturbing.

To see the impact of such a truncation bias on the type I error of the CMH row mean test, the following simulation was performed in the Appendix. One thousand trials were simulated with 1000 patients for each drug and placebo groups. The CMH row mean statistic was calculated with the original data as well as the truncated ones. With the significance level 0.05, rejections of the test were recorded. The results are summarized in the following table,

Type I Error of CMH Row Mean Test, $\alpha = 0.05$

	assign 3 to truncated group
with original scores	5.2 %
with truncated scores	86.4%

Here we see that the CMH row mean test with the original scores does not inflate the type I error, but the test with the truncated scores severely inflates the type I error. For the 1000 trials, the mean of p-values for the test with truncated scores was 0.03; the one with the original scores was 0.51. There were 387 (38.6%) trials whose truncated tests had p-values less than **0.001**. The correspondent p-values for the test with original scores had a mean value of **0.41**.

This example is by no means to predict, or estimate, the p-value of the CMH row mean test for the complete sleep duration data. That p-value is unknown. The point I want to make here is that even with a small p-value of 0.009 from the truncated sleep duration

data, one can not assume that the test result will be significant, should the complete sleep duration data be collected.

Kun Jin, Ph.D.
Statistical Team Leader
HFD-710

CC: NDA 21-393
HFD-120/Dr. Katz
HFD-120/Dr. Laughren
HFD-120/Dr. Andreason
HFD-120/Ms. Homonay
HFD-560/Dr. Ganley
HFD-550/Dr. Simon
HFD-120/MS. Dean
HFD-710/Dr. Chi
HFD-710/Dr. Jin
HFD-710/Dr. Yan
HFD-700/Dr. Anello

Appendix

Truncation Bias in Row Mean Analysis of Ordinal Data

Kun Jin, HFD-710

May, 2002

In this submission, a natural ordinal data for sleep duration should be the number of hours patients slept. The sponsor, instead, collected the following data, 0 for less than 5 hours, 1 for 5 to 6 hours, 2 for 6+ to 7 hours, 3 for 7+ to 8 hours, 4 for 8+ to 9 hours, and 5 for more than 9 hours.

Denote a natural sleep scale as i when the sleep duration fall in the interval $[i-1, i)$, $i = 1, 2, \dots$. The sponsor's collection of data involves two transformations on the natural ordinal data: (a) Truncation of data below 5 or above 10 to 5 and 10, respectively; (b) Linear shift of data downward by 5 units.

One would like to ask whether such transformations introduce a bias to the test when a CMH row mean test statistic is used. For a $2 \times r$ table, let a_j be the column score, n_{ij} be counts in cell (i, j) , ($i=1,2, j=1,r$). Denote n_{i+} , and n_{+j} the correspondent margin total counts, and n the total count. Then the CMH row mean test statistic is calculated as follows, (see Stokes et al, 1995)

$$Q_s = n_{1+}(n-1)(f_1 - u_a)^2 / (n-n_{1+}) v_a,$$

where

$$f_1 = \sum_{j=1}^r a_j n_{1j} / n_{1+}, u_a = \sum_{j=1}^r a_j n_{+j} / n, v_a = \sum_{j=1}^r (a_j - u_a)^2 n_{+j} / n.$$

It is easy to see from the formulas that the CMH row mean test statistic is invariant under a location shift, therefore transformation (b) will not introduce a bias to the test result. The truncation above 10 seems to be reasonable since it is not likely to have many data more than 11 hours sleep time. It could be reasonably concluded that the upper end truncation is not materialized. So the real concern is left with the lower end truncation, namely, truncating scores 1,2,3,4, and 5 to the score 5.

Formal statistical justification does not exist for employing such a truncation on ordinal data when a CMH row mean test is used. A naive thinking is that since such a truncation is done for both drug and placebo groups, it is unlikely introducing a bias in favoring one group over the other. This argument could be stated more precisely in statistical term in several ways. I think the following statement would be a relevant one to the current content. Let m_d and m_p be the population mean of original ordinal data of drug and placebo groups, m_{cd} and m_{cp} be the correspondent truncation ones. Under the null hypothesis, $m_d = m_p$, we will also have $m_{cd} = m_{cp}$.

Such a general claim is not true. In the following paragraphs, we will demonstrate cases, in which two populations row means based on original ordinal scales are identical, the truncated row means, however, become different. This illustrates that the low-end truncation could create a treatment mean difference even there is no treatment difference in the original data.

For simplicity, let natural ordinal scales be i , ($i=1,2 \dots 7$), $p(\text{drug patient being } i) = p_i$, $p(\text{placebo patient being } i) = q_i$, ($i = 1, \dots, 7$). Assume $p_i = p_{7-i+1}$, $i=1,2,3$. Let $q_1 = p_1 - 2.5\epsilon$, $q_7 = p_7 - 2.5\epsilon$, $q_i = p_i + \epsilon$, ($i = 2, \dots, 6$), ϵ is positive and small enough to make q_i a probability. Under the assumption, $m_d = m_p = 4$. If we truncate 1,2,3 to 3, the truncated probability become $p_3^c = p_1 + p_2 + p_3$, and $q_3^c = q_1 + q_2 + q_3$, $p_i^c = p_i$, $q_i^c = q_i$. The truncated population row mean of drug group is $m_{cd} = \sum_{i=3}^7 ip_i^c$. For the placebo group, $m_{cp} = \sum_{i=3}^7 iq_i^c = m_{cd} - 4\epsilon < m_{cd}$. This example illustrates that drug and placebo groups could have identical row means with a natural ordinal scale, but the drug group will have a larger row mean than the placebo after the low-end truncation.

Here I will provide a concrete numerical example for non-statistical reviewers. Let $(p_1, \dots, p_7) = (0.2068966, 0.1379310, 0.1034483, 0.1034483, 0.1034483, 0.1379310, 0.2068966)$,

(1)

$(q_1, \dots, q_7) = (0.1034483, 0.1379310, 0.1724138, 0.1724138, 0.1724138, 0.1379310, 0.1034483)$.

(2)

The drug row mean is

$1*0.2068966+2*0.1379310+3*0.1034483+4*0.1034483+5*0.1034483+6*0.1379310+7*0.2068966 = 4$; the placebo row mean is

$1*0.1034483+2*0.1379310+3*0.1724138+4*0.1724138+5*0.1724138+$

$6*0.1379310+7*0.1034483=4$. After the truncation at 3, the probabilities of drug and

placebo groups are $(0.4482759, 0.1034483, 0.1034483, 0.1379310, 0.2068966)$ and

$(0.4137931, 0.1724138, 0.1724138, 0.1379310, 0.1034483)$, respectively. The drug row

mean becomes $3*0.4482759+4*0.1034483+5*0.1034483+6*0.1379310+7*$

$0.2068966=4.551725$; the placebo row mean becomes $3*0.4137931+4*0.1724138$

$+5*0.1724138+6*0.1379310+7*0.1034483=4.344828$.

The CMH row mean test is advocated as a sensible test to target a location difference between groups, (Stokes et al, 1995). The bias introduced by the truncation described above is particularly disturbing when the CMH row mean test is used. To see the impact of such truncation bias on the type I error of the CMH row mean test, the following simulation was performed. Using the probabilities (1) and (2), 1000 patients with natural ordinal scores for each drug and placebo groups were generated. Notice that these samples were generated under a null hypothesis that the row means of two populations are both 4. The CMH row mean test statistic was calculated with the original data as well as the truncated ones. With the significance level 0.05, rejections of the test were recorded. This simulation was repeated 1000 times. The result is in the following table

Type I Error of CMH Row Mean Test, $\alpha = 0.05$

	assign 2 to truncated group	assign 3 to truncated group
--	-----------------------------	-----------------------------

with original scores	5.2%	5.2 %
with truncated scores	50.1%	86.4%

Here we see that the CMH row mean test with original scores does not inflate the type I error, but the test with the truncated scores severely inflates the type I error.

Reference

Stokes, ME, Davis, CS, and Koch GG (1995). *Categorical Data Analysis Using the SAS System*, Cary, SAS Institute Inc.

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

/s/

Kun Jin
6/5/02 10:00:15 AM
BIOMETRICS

1 Executive Summary of Statistical Findings	3
1.1 Overview of Clinical Program and Studies Reviewed	3
1.2 Principal Findings and Conclusions	3
2 Statistical Review and Evaluation of Evidence	3
2.1 Introduction and Background	3
2.2 Data Analyzed and Sources	4
2.3 Statistical Evaluation of Evidence on Efficacy in Study AE9802	4
2.3.1 <i>Protocol</i>	4
2.3.2 <i>Sponsor's Results and Conclusions</i>	6
2.3.2.1 <i>Patient Disposition</i>	6
2.3.2.2 <i>Demographics</i>	6
2.3.2.3 <i>Main Efficacy Results</i>	7
2.3.2.4 <i>Findings in Special Patient Population</i>	11
2.3.3 <i>Reviewer's Comment</i>	11
2.4 Final Conclusion for Study AE9802	11

1 EXECUTIVE SUMMARY OF STATISTICAL FINDINGS

1.1 OVERVIEW OF CLINICAL PROGRAM AND STUDIES REVIEWED

NDA21393 (Advil PM liqui-gel) was submitted for the indication of relief of occasional sleeplessness when associated with — minor aches and pains. Advil PM liqui-gel was a combination of two components: Ibuprofen and diphenhydramine hydrochloride. A total of 4 Phase III efficacy studies were conducted including 3 dental studies (AE9801, AE9802 and AE9803) and 1 headache study (AE9804). The sleep benefit of the combination therapy over ibuprofen 400 mg and placebo are under review by Division of Neuropharmacological Drug and collocated statisticians. Per request of Div of Anti Inflammatory Analgesic, and Ophthalmologic Drug Products, this reviewer provided consult for Study AE9802, the only study showed advantage in ibuprofen 400 mg /diphenhydramine hydrochloride 50 mg over ibuprofen 400 mg in one of the primary endpoints (sleep duration).

1.2 PRINCIPAL FINDINGS AND CONCLUSION

In Study AE9802, both Ibuprofen 400 mg/diphenhydramine hydrochloride 50 mg and ibuprofen 400 mg groups demonstrated statistically significant advantage over placebo in all three primary endpoints: sleep duration, cumulative percentage of subjects asleep at 60 minutes, and sum of pain relief and pain intensity difference (SPRID). Ibuprofen 400 mg/diphenhydramine hydrochloride 50 mg demonstrated statistically significantly longer sleep duration than ibuprofen 400 mg ($p=0.009$ based on Cochran-Mantel-Haenszel analysis with modified ridits), but the results in cumulative percentage of subjects asleep at 60 minutes and SPRID were reversed with p-values 0.112 and 0.05, respectively, for the advantage of ibuprofen 400 mg. Also, ibuprofen 400 mg/diphenhydramine hydrochloride 50 mg did not show additional benefit in the secondary sleep endpoints and was numerically worse in most of the secondary pain endpoints. Therefore, it is questionable whether ibuprofen 400 mg/diphenhydramine hydrochloride 50 mg has any overall benefit compared with ibuprofen 400 mg in terms of sleep and pain.

2 STATISTICAL REVIEW AND EVALUATION OF EVIDENCE

2.1 INTRODUCTION AND BACKGROUND

NDA21393 (Advil PM liqui-gels) was submitted for the indication of relief of occasional sleeplessness when associated with — , minor aches and pains. Since Advil PM liqui-gel is a combination of ibuprofen 400 mg and diphenhydramine hydrochloride 50 mg, a full factorial pilot study (Study AE9701) was conducted to evaluate the efficacy of ibuprofen 400 mg/diphenhydramine hydrochloride 50 mg vs. ibuprofen 400 mg, diphenhydramine hydrochloride 50 mg and placebo. Sponsor also conducted 4 phase III efficacy studies including 3 dental studies and 1 headache study. Only 2 of the dental studies were of partial

factorial design including ibuprofen 400 mg/diphenhydramine hydrochloride 50 mg, ibuprofen 400 mg and placebo groups. The other 2 Phase III studies included only different ibuprofen/diphenhydramine hydrochloride dose groups and placebo group. Therefore, the pain benefit of the combination therapy over diphenhydramine hydrochloride 50 mg can not be assessed in Phase III trials due to the absence of diphenhydramine hydrochloride 50 mg group. The sleep benefit of the combination therapy over ibuprofen 400 mg and placebo are under review by Division of Neuropharmacological Drug and collocated statisticians. Per request of Div of Anti Inflammatory Analgesic, and Ophthalmologic Drug Products, this reviewer provided consult for Study AE9802, the only study showed advantage in ibuprofen 400 mg /diphenhydramine hydrochloride 50 mg over ibuprofen 400 mg in one of the primary endpoints (sleep duration).

2.2 DATA ANALYZED AND SOURCES

The dataset analyzed by this reviewer was `efftran.xpt` submitted by the sponsor in electronic document room with pathway '\\CDSesub1\N21393\N_000\2001-10-16\crt\datasets\ae9802'.

2.3 STATISTICAL EVALUATION OF EVIDENCE ON EFFICACY IN STUDY AE9802

2.3.1 PROTOCOL

This was a randomized, stratified, inpatient, placebo-controlled, partial-factorial, single-dose, double-blind, parallel group, single-center trial. The primary objective of this trial was to evaluate the analgesic and sedative efficacy of Advil PM Liqui-Gels (ibuprofen 400 mg/diphenhydramine hydrochloride 50 mg) compared to ibuprofen liquigel (400 mg) and placebo.

Following oral surgery, subjects were housed and observed at a clinic site overnight. When subjects experienced at least moderate pain and it was between approximately 6:30 PM and 8:00 PM (at least 3 hours earlier than their usual bedtime), they received ibuprofen 400 mg/diphenhydramine hydrochloride 50 mg, ibuprofen liquigels (400 mg) or placebo in a 3:3:1 ratio and were required to go to bed for the evening. The subject randomization was stratified by pain and gender. Sleep was evaluated by an observer at regular intervals over 3 hours post-dosing. Subjects provided pain assessments at 90 and 120 minutes post-dosing. The nurse observer was scheduled to observe whether the subject was asleep or not at 10, 20, 30, 40, 50, 60, 75, 90, 120, 150 and 180 minutes post dose. Subjects also provided subjective assessments of sleep efficacy as well as global assessments of the study medication as a sleep-aid and as an analgesic the next morning (or at the time rescue medication was used).

In the original protocol, the primary efficacy parameters were cumulative percentage of subjects asleep at 60 minutes post-dose (based on the nurse observer sleep latency assessments), and sum of pain relief combined with pain intensity difference scores (SPRID) over 0-120 minutes. In a later protocol amendment, sleep duration was added as an additional primary endpoint. The secondary efficacy parameters regarding sleep were ease of

falling asleep, global assessment of sleep, sleep latency (based on the nurse observer), actual percentage of subjects asleep at each observed time point and the cumulative percentage of subjects asleep at each observed time point (other than the 60-minute time point), time to rescue medication (for sleeplessness) and the cumulative proportion of subjects taking rescue medication by each pain assessment time point and by the wake-up time. The secondary endpoints regarding pain were PID, relief, and PRID scores at 90 and 120 minutes, SPID and TOTPAR over 0-120 minutes, global assessment of pain, time to rescue medication (for pain) and the cumulative proportion of subjects taking rescue medication by each pain assessment time point and by the wake-up time.

In order to protect the Type I error at 0.05 level, the comparisons will be performed in the following sequential order. Each step must be significant for the following steps to be eligible for significance. However, in order to present the full clinical picture, all pairwise comparisons will be presented.

- 1). IBU/DPH 400/50mg vs placebo: In order to be eligible for being declared significant, both primary sleep parameters and the primary pain parameter should be significant at the 0.05 level.
- 2). IBU/DPH 400/50mg vs IBU 400mg: Duration of sleep will be tested first followed by cumulative percentage of subjects asleep at 60 minutes, each at the 0.05 level. The cumulative percentage of subjects asleep at 60 minutes will be eligible for being declared significant only if the duration of sleep is significant. If duration of sleep is significant, the combination will be considered more effective than ibuprofen alone for sleep.
- 3). IBU 400mg vs placebo: In order to be eligible for being declared significant, the primary pain parameters should be significant at the 0.05 level.

PID, SPID, PRID, SPRID, PRR, TOTPAR, ease of falling asleep, sleep duration, and global scores (sleep and pain) will be analyzed by Analysis of Variance (ANOVA), incorporating effects for treatment, baseline pain severity rating, gender and local anesthetic regimen in the model. Sleep latency and time to rescue medication will be analyzed using the Cox proportional hazards regression model with effects for treatment, baseline PSR, gender, and local anesthetic regimen. Ninety-five percent confidence intervals for the median time to sleep and median time to rescue medication (separately for sleeplessness and pain) will be derived by the method of Simon and Lee. In addition, the actual and cumulative proportion of subjects asleep at each observation time point and the cumulative proportion of subjects who took rescue medication (separately for sleeplessness and pain) by each observation time point will be analyzed by the CMH test controlling for baseline PSR, gender, and local anesthetic regimen.

A total of 280 patients was planned to be included in the study with 120 in each of the active treatment groups and 40 in the placebo group. Assuming that the cumulative percentage of subjects asleep by 60 minutes in the IBU 400mg is 71%, 120 subjects in each active treatment group will provide at least 90% power (at 5% level of significance) to detect a 18% higher percentage of subjects asleep by 60 minutes in the IBU/DPH 400/50mg group.

2.3.2 SPONSOR'S RESULTS AND CONCLUSIONS

2.3.2.1 Patient Disposition

A total of 283 subjects were enrolled, randomized and took study medication. Therefore, 283 subjects were included in the safety analysis. There was one subject without any post-baseline efficacy pain assessments and the remaining 282 subjects were included in the ITT efficacy analysis. One subject was discontinued due to an adverse event (headache) and another subject completed the study, but subsequently, was found to have a protocol violation. An overall summary of subject disposition is provided in Table 1 below.

Table 1. Patient Disposition

	Total		Placebo		IBU400/DPH50		IBU400	
	n	%	n	%	n	%	n	%
Entered	283		40		120		123	
Received Study Medication/ Included for Safety	283		40		120		123	
Completed Study	281	99.3%	39	97.5%	119	99.2%	123	100.0%
Discontinued	2	0.7%	1	2.5%	1	0.8%	0	0.0%
Reason for Discontinuation								
Lost to Follow-up	0	0.0%	0	0.0%	0	0.0%	0	0.0%
Adverse Event	1	0.4%	1	2.5%	0	0.0%	0	0.0%
Rescued <1 HR.	0	0.0%	0	0.0%	0	0.0%	0	0.0%
Uncooperativeness	0	0.0%	0	0.0%	0	0.0%	0	0.0%
Withdrew Voluntarily	1	0.4%	0	0.0%	1	0.8%	0	0.0%
Ineligible	0	0.0%	0	0.0%	0	0.0%	0	0.0%
Protocol Violation	0	0.0%	0	0.0%	0	0.0%	0	0.0%
Administrative/Other	0	0.0%	0	0.0%	0	0.0%	0	0.0%
Intent-to-Treat Subjects	282	99.6%	40	100.0%	119	99.2%	123	100.0%
Evaluable Subjects	280	98.9%	39	97.5%	119	99.2%	122	99.2%

2.3.2.2 Demographics

The demographic characteristics for the 282 subjects from the ITT population are tabulated in Table 2 below. The three treatment groups were comparable for all demographics and medical history. There were 137 (48.6%) males and 145 (51.4%) females; the racial distribution was 95.4% Caucasian, followed by 2.8% Hispanic, 0.7% Black, 0.7% Asian, and 0.4% Other. The mean age was 20.0 years (range: 15-40 years). The mean weight and height were 150.3 pounds (range: 90-310 pounds) and 68.1 inches (range: 59-78 inches) respectively.

Table 2. Patient Demographics

	Total N=282	Placebo N=40	IBU400/DPH50 N=119	IBU400 N=123
GENDER				
MALE	137 (48.6%)	20 (50.0%)	58 (48.7%)	59 (48.0%)
FEMALE	145 (51.4%)	20 (50.0%)	61 (51.3%)	64 (52.0%)
RACE				
CAUCASIAN	269 (95.4%)	37 (92.5%)	114 (95.8%)	118 (95.9%)
BLACK	2 (0.7%)	1 (2.5%)	0 (0%)	1 (0.8%)
ASIAN	2 (0.7%)	0 (0%)	2 (1.7%)	0 (0%)
HISPANIC	8 (2.8%)	1 (2.5%)	3 (2.5%)	4 (3.3%)
OTHER	1 (0.4%)	1 (2.5%)	0 (0%)	0 (0%)
AGE (yrs.)				
MEAN	20.0	20.0	19.7	20.2
STD	4.3	5.0	4.1	4.3
MEDIAN	19.0	18.0	19.0	19.0
RANGE	(15, 40)	(15, 39)	(16, 40)	(16, 39)
WEIGHT (lbs.)				
MEAN	150.3	152.1	148.8	151.2
STD	32.9	42.7	30.3	32.0
MEDIAN	145.0	142.5	145.0	150.0
RANGE	(90, 310)	(105, 310)	(90, 250)	(99, 295)
HEIGHT (ins.)				
MEAN	68.1	67.8	68.3	68.1
STD	3.8	3.9	3.7	3.9
MEDIAN	68.0	67.5	69.0	68.0
RANGE	(59, 78)	(59, 74)	(60, 78)	(59, 78)

2.3.2.3 Main Efficacy Results

Primary Endpoints

Based on a 0 to 5-point scale (0 = *less than 5 hours* to 5 = *greater than 9 hours*), the mean scores for the duration of sleep were 0.05, 2.61 and 1.98 in the placebo, ibuprofen 400 mg/diphenhydramine hydrochloride 50 mg and ibuprofen 400 mg groups, respectively. The cumulative percentage of subjects who had fallen asleep by 60 minutes in the corresponding groups was 27.5%, 66.4% and 75.6% and the 2-hour SPRID scores were 0.3, 7.0 and 7.8 in the placebo, ibuprofen 400 mg/diphenhydramine hydrochloride 50 mg and ibuprofen 400 mg groups, respectively. Both the ibuprofen 400 mg/diphenhydramine hydrochloride 50 mg and the ibuprofen 400 mg dose groups were significantly better than placebo for both sleep parameters as well as the pain parameter. With respect to the comparison of the two actives, ibuprofen 400 mg/diphenhydramine hydrochloride 50 mg subjects experienced a significantly longer duration of sleep than the ibuprofen 400 mg subjects ($p=0.005$). Conversely, ibuprofen 400 mg was numerically better than ibuprofen 400 mg/diphenhydramine hydrochloride 50 mg for the cumulative percentage of subjects asleep by 60 minutes ($p=0.112$) and significantly better for SPRID2 ($p = 0.050$). The results for the primary efficacy variables are presented in Table 3.

Table 3. Results in Primary Endpoints

	Placebo N=40	IBU400/DPH50 N=119	IBU400 N=123
Duration of Sleep			
<5 hours(0)	39 (97.5%)	26 (21.8%)	41 (33.3%)
5 to 6 hours(1)	0 (0%)	18 (15.1%)	18 (14.6%)
6+ to 7 hours(2)	1 (2.5%)	12 (10.1%)	15 (12.2%)
7+ to 8 hours(3)	0 (0%)	12 (10.1%)	14 (11.4%)
8+ to 9 hours(4)	0 (0%)	23 (19.3%)	22 (17.9%)
>9 hours(5)	0 (0%)	28 (23.5%)	13 (10.6%)
P-value* vs. placebo		<0.001	<0.001
P-value* vs. IBU400		0.005	
Cumulative % Asleep at 60 min			
Number (%)	11 (27.5%)	79 (66.4%)	93 (75.6%)
P-value** vs. placebo		<0.001	<0.001
P-value** vs. IBU400		0.112	
SPRID2			
MEAN	0.26	7.03	7.81
STD	2.07	3.47	2.87
MEDIAN	0.00	7.00	8.00
RANGE	(-2, 6)	(-2, 14)	(-2, 14)
P-value* vs. placebo		<0.001	<0.001
P-value* vs. IBU400		0.05	

*: p-values from ANOVA model with treatment, baseline PSR, and gender terms.

** :p-values from the Cochran-Mantel-Haenszel test, controlling for baseline PSR and gender.

Secondary Endpoints

Among sleep-related secondary endpoints (ease of falling asleep, global assessment of sleep, sleep latency (based on the nurse observer), actual percentage of subjects asleep at each observed time point and the cumulative percentage of subjects asleep at each observed time point, and time to rescue medication for sleeplessness) both active treatment groups were significantly better than the placebo group, and the results were very similar in the two active treatment groups. The probability of not taking rescue medication for sleeplessness and sleep latency were plotted in Figures 1 and 2 below to show this general trend in sleep-related secondary endpoints.

Figure 1. Probability of Not Taking Rescue Medication for Sleeplessness

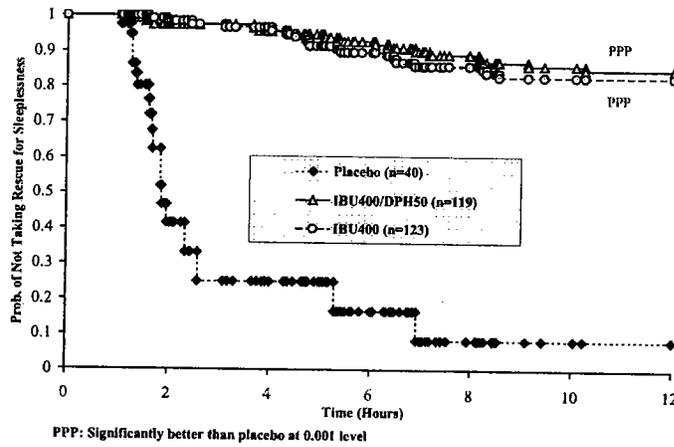
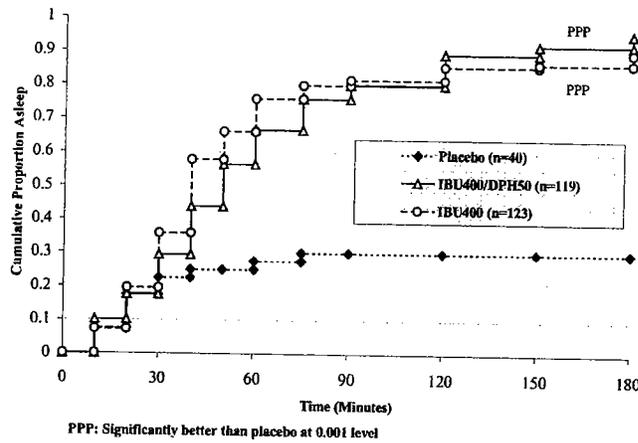
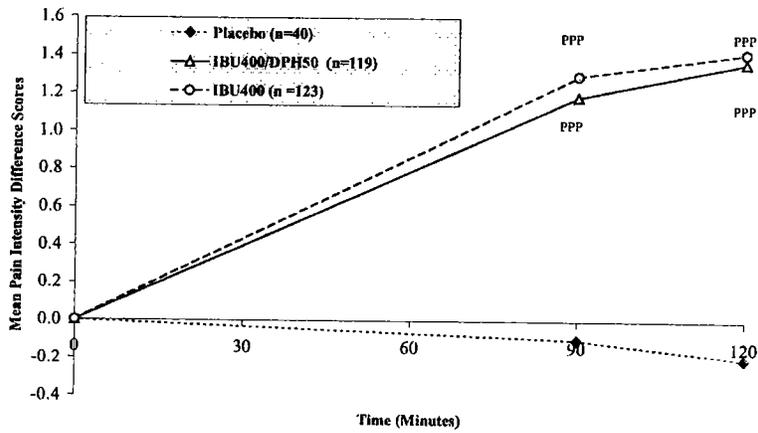


Figure 2. Sleep Latency (Time to Sleep based on Nurse's Observation)



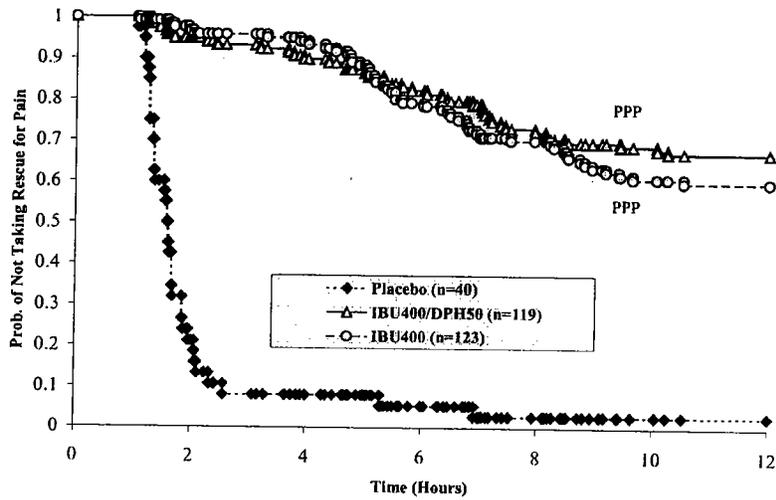
Among the pain-related secondary endpoints (PID, relief, and PRID scores at 90 and 120 minutes, SPID and TOTPAR over 0-120 minutes, global assessment of pain, time to rescue medication for pain), both active treatments were significantly better than placebo. The ibuprofen 400 mg group was numerically better than ibuprofen 400 mg/diphenhydramine hydrochloride 50 mg group in all these endpoints except for time to rescue medication (no separation for this endpoint) and the p-values for the advantage was below 0.05 in TOTPAR over 0-120 minutes. The PID scores at each time points were plotted in Figure 3 below to show this general trend. The probability of not taking rescue medication for pain is also plotted in Figure 4 below.

Figure 3. Pain Intensity Difference at Each Time Point



PPP: Significantly better than placebo at 0.001 level

Figure 4. Probability of Not Taking Rescue Medication for Pain



PPP: Significantly better than placebo at 0.001 level

2.3.2.4 Findings in Special Patient Population

Results in patients with different BPS and gender were compared. There was no evidence for treatment by BPS and treatment by gender interactions.

2.3.3 REVIEWER'S COMMENT

The sleep duration was a ordinal variable with 6 categories: 0 (<5 hours), 1 (5-6 hours), 2 (6-7 hours), 3 (7-8 hours), 4 (8-9 hours), and 5 (>9 hours). Since the categories do not all represent equal time interval and there was substantial number of patients (about 44% in each of the active treatment group) located at the two extreme categories (0 and 5), the ANCOVA analysis proposed by the sponsor which treated the categorical scores as continuous was not appropriate. Based on FDA's suggestion, the sponsor also performed Cochran-Mantel-Haenszel (CMH) test using modified ridit scores and the result shows that both of the active treatment groups had significantly longer sleep duration than the placebo arm and the ibuprofen 400 mg/diphenhydramine hydrochloride 50 mg was significantly better than ibuprofen 400 mg ($p=0.009$). So the ANOVA test and CMH test are consistent in terms of statistical significance.

2. Although ibuprofen 400 mg/diphenhydramine hydrochloride 50 mg was significantly better than ibuprofen 400 mg ($p=0.009$) in sleep duration, the results in cumulative percentage of subjects asleep at 60 minutes post-dose and SPRID were reversed with p-values 0.112 and 0.05, respectively, for the advantage of ibuprofen 400 mg. The secondary sleep endpoints did not provide any support for the benefit of ibuprofen 400 mg/diphenhydramine hydrochloride 50 mg over ibuprofen 400 mg, and most secondary pain endpoints (except for time to rescue medication for pain) showed numerical advantage in ibuprofen 400 mg over ibuprofen 400 mg/diphenhydramine hydrochloride 50 mg. Therefore, it is questionable whether ibuprofen 400 mg/diphenhydramine hydrochloride 50 mg would provide any overall benefit over ibuprofen 400 mg in terms of sleeplessness and pain.

2.4 FINAL CONCLUSION FOR STUDY AE9802

In Study AE9802, ibuprofen 400 mg/diphenhydramine hydrochloride 50 mg was significantly better than ibuprofen 400 mg in sleep duration. However, the results in cumulative percentage of subjects asleep at 60 minutes post-dose and SPRID were reversed with p-value 0.112 and 0.05, respectively, for the advantage of ibuprofen 400 mg. Also, ibuprofen 400 mg/diphenhydramine hydrochloride 50 mg did not show additional benefit in the secondary sleep endpoints and was numerically worse in the secondary pain endpoints. Therefore, it is questionable whether ibuprofen 400 mg/diphenhydramine hydrochloride 50 mg has any overall benefit compared with ibuprofen 400 mg.

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

/s/

Hong Lu
4/29/02 02:00:29 PM
BIOMETRICS

Stan Lin
4/30/02 10:08:09 AM
UNKNOWN