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:	21-042/SE5
Drug Name:	Vioxx™ (Rofecoxib Tablets)
Indication(s):	Juvenile Rheumatoid Arthritis
Applicant:	Merck and Company Inc. Rahway NJ 07065-0900
Date (s):	Submitted: December 5, 2003 Received: December 5, 2003 Reviewed: May 19, 2004
Review Priority:	Priority
Biometrics Division:	Division of Biometrics III (HFD-725)
Statistical Reviewer:	M. Atiar Rahman, Ph.D. (HFD-725)
Concurring Reviewer:	Stan Lin, Ph.D. (HFD-725)
Medical Division:	Division of Anti-Inflammatory, Analgesic and Ophthalmic Drug Products (HFD-550)
Clinical Team:	Carolyn Yancey, M.D. (HFD-550)
Project Manager:	Barbara Gould (HFD-550)
Keywords:	NDA review, Clinical studies, Juvenile Rheumatoid Arthritis, Non-inferior

Table of Contents

1 EXECUTIVE SUMMARY

1.1 CONCLUSIONS AND RECOMMENDATIONS

In this submission the sponsor included report of a Phase 3 study (Protocol134/135). This was a 12 week, double blind, double dummy, active comparator controlled study in 2 to 17 years old juvenile rheumatoid arthritis patients. There were three treatment groups namely, (1) lower dose rofecoxib- 0.3 mg/kg/day for 2 to 11 year old patients (not to exceed 12.5 mg/day), or 12.5 mg daily for 12 to 17 year old patients; (2) higher dose rofecoxib- 0.6 mg/kg/day for 2 to 11 year old patients (not to exceed 25 mg/day), or 25 mg daily for 12 to 17 year old patients; (3) naproxen- targeted to 15 mg/kg/day. The primary objectives of this study were to examine the therapeutic effects and safety of two doses of rofecoxib and to show the non-inferiority of rofecoxib compared to naproxen.

The primary efficacy endpoint was proportion of patients in the ITT population meeting the JRA 30 criteria. The non-inferiority is claimed if the 95% confidence interval on the relative risk (ratio of percent of patients meeting the JRA 30 criteria) of rofecoxib vs. naproxen was entirely (b) (4)

(b) (4)

However, the Division considered the non-inferiority margin (b) (4) In ITT population the lower limits of the 95% CIs were 0.76 for higher dose rofecoxib vs. naproxen (b) (4)

The medical officer's preference of non-inferiority margin is 0.75 (see medical officer's review). (b) (4)

(b) (4)

1.2 BRIEF OVERVIEW OF CLINICAL STUDIES

In this submission the sponsor included report of a Phase 3 study (Protocol 134/135). This was a 12 week, double blind, double dummy, active comparator controlled study in 2 to 17 year old Juvenile Rheumatoid Arthritis (JRA) patients. Within each study site, allocations were stratified for age: 2 to 11 years and 12 to 17 years, and by pauci articular and poly articular disease. In each age group, patients were allocated to receive 1 of the 3 treatments: (1) lower dose rofecoxib, 0.3 mg/kg/day for 2 to 11 year old patients (not to exceed 12.5 mg/day), or 12.5 mg daily for 12 to 17 year old patients; (2) higher dose rofecoxib, 0.6 mg/kg/day for 2 to 11 year old patients (not to exceed 25 mg/day), or 25 mg daily for 12 to 17 year old patients; (3) naproxen, targeted to 15 mg/kg/day. The 2 to 11 year old patients received suspension formulations, and the 12 to 17 year old patients received tablets. Clinical assessments took place at pre-study screening, randomization, and after 2, 4, 8, and 12 weeks of treatment. In addition, a 14 day post-study follow up was required of all patients after discontinuation. The primary objectives of this study were (1) to examine the therapeutic effects of 2 dose strengths of rofecoxib, taken as oral suspension, in 2 to 11 year old JRA patients (0.3, and 0.6 mg/kg/day); (2) to examine the therapeutic effects of 2 dose strengths of rofecoxib, taken as tablets, in 12 to 17 year old JRA patients (12.5 mg and 25 mg once daily); (3) to demonstrate the safety and tolerability of rofecoxib in children with JRA; (4) to examine the safety and efficacy of naproxen for treatment of JRA, and compare rofecoxib with naproxen (non-inferiority); and (5) To examine treatment effects in patients with pauci articular and poly articular disease, respectively.

1.3 STATISTICAL ISSUES AND FINDINGS

In sponsor's analysis non-inferiority of a dose of rofecoxib to naproxen was claimed if the 95% CI on the risk ratio for that dose vs. naproxen was (b) (4) The Division considers the proposed non-inferiority margin (b) (4) as unacceptably wide. The determination of whether the efficacy of rofecoxib is comparable or not to naproxen in the JRA population, was to be a review issue.

2 INTRODUCTION

2.1 OVERVIEW

In this NDA the sponsor submitted data of a Phase 3 study (Protocol134/135) to support their claim that the use of Vioxx is safe and efficacious for the treatment of Juvenile Rheumatoid Arthritis and that Vioxx is non-inferior in efficacy to Neproxen.

2.2 DATA SOURCES

The submission was both electronic and hard copy. Submitted data were stored in folder \\Cdsesub1\n21042\S_026\2003-12-05\Crt of FDA's Electronic Document Room (EDR). The data quality of the submission was within acceptable limit.

3 STATISTICAL EVALUATION

3.1 EVALUATION OF EFFICACY

3.1.1 STUDY # 143/135 (PHASE-3)

Title: "A 12 Week Active Comparator Controlled Trial to Evaluate the Efficacy and Safety of rofecoxib for Treatment of Juvenile Rheumatoid Arthritis"

3.1.1.1 Design and Objectives

This was a 12 week, double blind, double dummy, active comparator controlled study in 2 to 17 year old JRA patients. Within each study site, allocations were stratified for age: 2 to 11 years and 12 to 17 years, and by pauci articular and poly articular disease. In each age group, patients were allocated to receive 1 of the 3 treatments: (1) lower dose rofecoxib 0.3 mg/kg/day for 2 to 11 year old patients (not to exceed 12.5 mg/day), or 12.5 mg daily for 12 to 17 year old patients; (2) higher dose rofecoxib, 0.6 mg/kg/day for 2 to 11 year old patients (not to exceed 25 mg/day), or 25 mg daily for 12 to 17 year old patients; and (3) naproxen, targeted to 15 mg/kg/day. The 2 to 11 year old patients received suspension formulations, and the 12 to 17 year old patients received tablets. Clinical assessments took place at pre-study screening, randomization, and after 2, 4, 8, and 12 weeks of treatment. In addition, a 14 day post-study follow up was required of all patients after discontinuation. The primary objectives of this study were (1) to examine the therapeutic effects of 2 dose strengths of rofecoxib, taken as oral suspension, in 2 to 11 year old JRA patients (0.3 mg/kg/day, not to exceed 12.5 mg, and 0.6 mg/kg/day, not to exceed 25 mg); (2) to examine the therapeutic effects of 2 dose strengths of rofecoxib, taken as tablets, in 12 to 17 year old JRA patients (12.5 mg and 25 mg once daily); (3) to demonstrate the safety and tolerability of rofecoxib in children with JRA; (4) to examine the safety and efficacy of rofecoxib for treatment of JRA, and compare to naproxen (non-inferiority); and (5) to examine treatment effects in patients with pauci articular and poly articular disease, respectively.

Eligible patients underwent a brief washout of prior NSAID therapy and were assigned to 1 of the 3 treatment groups, in approximately equal proportions. Follow-up clinical assessments were performed at 2, 4, 8, and 12 weeks on study therapy. Acetaminophen was permitted as rescue medication for pain, but use was prohibited within 24 hours of scheduled clinic visits. The primary efficacy timepoint was Week 12.

A second phase of this study was a 12-month open label, active comparator controlled extension of treatment phase (pivotal study). The extension phase had two treatment groups, namely higher dose rofecoxib (0.6 mg/kg to a maximum of 25 mg once daily) and naproxen (15 mg/kg in 2 divided dose). There were 227 patients in the extension phase. In the extension phase some reassignments of patients in the treatment phase were performed following a randomization scheme prepared prior to randomization to treatment. Seventy five percent (75%) of patients in the lower rofecoxib dose group in the treatment phase were reassigned to higher dose rofecoxib and 25% were reassigned to naproxen. Fifty percent (50%) of the naproxen treated patient were reassigned to higher dose rofecoxib and 25% of the higher dose rofecoxib patients were reassigned to naproxen group. A schematic figure of this reassignment is given in Appendix-1.

Sample Size: In the protocol the sponsor stated "The sample size n=75 per dose group has at least 90% power to yield the 95% CI on the ratio of percent of patients improved greater than 0.5, if the true rates are equal for rofecoxib and naproxen and exceed 40%." However, in the actual study the sponsor recruited about 100 patients per treatment group. In their final report the sponsor mentioned "… the sample size of n=100 per dose group had 99% power to yield the 95% CI on the ratio of JRA 30 response rate greater than 0.5, if the true rates are equal for rofecoxib and naproxen and exceed 40%."

3.1.1.2 Efficacy Endpoint

The primary efficacy endpoint was the proportion of patients meeting the JRA 30 criteria[1]. The key Secondary endpoint was the proportion of patients that demonstrated improvement from baseline in parent/patient's assessment of overall well being. Other secondary endpoints in priority order included:

- parent/patient's global assessment of pain (VAS)
- proportion of patients discontinuing due to lack of efficacy JRA 30 Core Set of Variables
- parent/patient's assessment of overall well-being
- investigator's global assessment of disease activity
- patient's assessment of functional ability (CHAQ)
- number of joints with active arthritis
- number of joints with limited range of motion
- ESR

[1] As a result of correspondence with the agency, an amendment was done when the primary efficacy endpoint was changed from improvement in Patient/Parents assessment of overall well being to the JRA 30 (12/06/2001 Written Request and its 5/14/2003 Amendment) The JRA 30 responder criteria were derived from a core set of 6 outcome variables for the assessment of children with JRA Developed for assessment of impact of disease modifying anti rheumatic drug (DMARD) therapy on disease, improvement in patients with JRA was defined as an at-least 30% improvement from baseline in any 3 of the 6 variables in the core set, with not more than 1 of the remaining variables worsened by more than 30% The variables included in the core set were: (1) investigator global assessment of disease activity; (2) parent/patient's global assessment of overall well being; (3) functional ability; (4) number of joints with active arthritis; (5) number of joints with limited range of motion, and (6) ESR In addition to assessment of disease activity, an assessment of the patient's pain was conducted using the Patient's Global Assessment of Pain

3.1.1.3 Patient Analyzed

Modified Intent-to-Treat population: Primary efficacy population was modified intention-to-treat (MITT) population, defined as all patients with a baseline and at least one on treatment-period measurement.

Safety population: MITT population was also used for safety analysis.

Reviewer's comment: For safety analysis all randomized who took one dose of study drug should be more appropriate.

Per Protocol population: The PP analysis population excluded patients or data points with clinically important protocol deviations based on pre-specified criteria.

Patients with any of the following were excluded from the PP analysis:

1. Parent/patient's assessment of overall well being (VAS) was >90 mm at the screening visit.
2. Parent/patient's assessment of overall well being (VAS) was <10 mm at allocation.
3. Patient had fewer than 1 active joint at allocation.
4. Patient had any inflammatory joint disease, other than JRA, that confounded collection of efficacy data.

3.1.1.4 Disposition of Patients, Demography

Disposition of ITT patients and their baseline characteristics are given in Table 1 and Table 2, respectively in Appendix-1. A total of 310 subjects were treated and post-studied, including 109 in rofecoxib 0.3 mg/kg group, 100 in rofecoxib 0.6 mg/kg group, and 101 Naproxen 15 mg/kg group. A total of 46 subjects (15, 22, and 9 in rofecoxib 0.3 mg/kg, rofecoxib 0.6 mg/kg, and Naproxen groups, respectively) were between 2-4 years of age, while a total of 135 subjects (50, 38, and 47 in rofecoxib 0.3 mg/kg, rofecoxib 0.6 mg/kg, and Naproxen groups, respectively) were between 5-11 years of age. Overall about 19% completed the study (24%, 18%, and 14% in rofecoxib 0.3 mg/kg, rofecoxib 0.6 mg/kg, and Naproxen groups, respectively). Most subjects dropped out due to lack of efficacy. The treatment groups were generally comparable with respect to demographic and baseline characteristics. The majority of subjects were female (73.2%) and Caucasian (72.6%) with a mean age of 9.9 years.

3.1.1.5 Sponsor's Analysis of Primary Efficacy Data

The statistical analysis plan, as was described in the protocol, is given in Appendix-2. The study hypotheses was that the proportion of patients that demonstrated improvement, defined as meeting JRA 30 criteria, will be similar between rofecoxib and naproxen treatment groups. This hypothesis was assessed by the 95% CI on the relative risk (ratio of percent of patients meeting the JRA 30 criteria) of rofecoxib versus naproxen, calculated using the Mantel-Haenszel estimate with protocol, joint involvement and age group as stratification factors. (b) (4)

[2] In February 27, 2003 the sponsor submitted the data analysis plan for Study P134/135. In April 20, 2003 the Division sent comments to the sponsor regarding the primary population for analysis and margin of comparability chosen for the study. In those comments the Division considered the proposed non-inferiority margin (b) (4) as unacceptably wide. The determination of whether the efficacy of rofecoxib is comparable or not to naproxen in the JRA population was to be a review issue.

An analysis of covariance (ANCOVA) model, with treatment, protocol, joints involvement, age group, and baseline value as factors was used to analyze all continuous efficacy variables based on their time-weighted average response across Weeks 2, 4, 8, and 12. In addition, the assessment of the treatment response was also done through graphical presentation of the LS mean changes from baseline. The proportion of patients discontinuing test therapy due to lack of efficacy was assessed using the Fisher's exact test.

The therapeutic effect of each dose of rofecoxib was assessed in 2 to 11 year old JRA patients by the ratio of percent of patients meeting the JRA 30 criteria (each dose vs. naproxen) and its associated 95% CI, stratified by protocol and joint involvement. A similar assessment was used for the proportion of patients that demonstrate improvement from baseline in parent/patient's assessment of overall well being. In addition, the comparisons of therapeutic effect between treatments for other efficacy endpoints in 2 to 11 year old JRA patients were assessed by the least squares (LS) mean changes from baseline and their associated 95% CIs, stratified by protocol and joint involvement.

The therapeutic effect of each dose of rofecoxib was assessed in 12 to 17 year old JRA patients by the ratio of percent of patients meeting the JRA 30 criteria (each dose versus naproxen) and its associated 95% CI, stratified by protocol and joint involvement. Similar assessment was used for the proportion of patients that demonstrate improvement from baseline in parent/patient's assessment of overall well being. In addition, the comparisons of therapeutic effect between treatments for other efficacy endpoints in 12 to 17 year old JRA patients were assessed by the LS mean changes from baseline and their associated 95% CIs, stratified by protocol and joint involvement.

Primary efficacy analyses were based on a modified intention-to-treat (MITT) approach (i.e., inclusion of all patients with a baseline and at least one on treatment- period measurement). All measurements (except those from post-study visits) were used; including data collected at discontinuation and unscheduled visits. Dropouts were included in the analysis based on responses obtained up to and including those at the time of discontinuation. Analyses were performed on the time-weighted average response of observed data only, while the last-value-carried-forward method was used only for longitudinal graphs. Since most of the endpoints were analyzed as the time weighted averages over the treatment period, no missing values were imputed except for the longitudinal graphs.

A corroborative per-protocol (PP) analysis was also performed for the primary endpoint. The PP analysis population excluded patients or data points with clinically important protocol deviations based on pre-specified criteria.

A complete list of statistical methods used by the sponsor to analyze the primary and secondary efficacy endpoints are summarized in Table 3 in Appendix 1.

3.1.1.6 Sponsor's Results and Conclusions

<u>Results of sponsor's analysis of primary efficacy endpoints</u>

Results of sponsor's analysis of primary efficacy variable are given in Table 4 in Appendix 1. The proportions of patients in MITT population, meeting the JRA 30 criteria regardless of completion status over the 12 week study were (b) (4) 0.55, and 0.55 in (b) (4) higher dose rofecoxib, and naproxen treatment groups, respectively. The 95% CI on relative risk of higher dose rofecoxib to naproxen was 0.76 to 1.26 (b) (4)

(b) (4)

However, there was a trend to reduced efficacy with lower-dose rofecoxib as suggested by a numerically smaller proportion of patients meeting the JRA 30 criteria than higher dose rofecoxib.

Analysis of proportions of patients meeting the JRA 30 criteria in completers showed similar results. Additionally, the proportion of patients meeting the JRA 30 criteria, based on the time-weighted average up to each time point were similar throughout the base study. Results from the per-protocol population for the primary endpoint corroborated those from the MITT population as shown in Table 5 in Appendix 1.

Results of sponsor's analysis of secondary efficacy endpoints

Sponsor's analysis results of the key secondary endpoint, the proportion of patients with improvement from baseline in parent/patient's assessment of overall well being are given in Table 6 in Appendix 1. For, all 3 treatment groups showed (b) (4) treatment effect of (b) (4) 76.0%, and 73.0%, respectively. Sponsor's analysis results of patient's global assessment of pain are given in Table 7 in Appendix 1. Results showed that both the lower dose and higher dose rofecoxib treatment had numerically greater improvement than naproxen treatment. The mean change from baseline was (b) (4) -13.61 and –9.11 respectively. The discontinuation rates due to lack of efficacy (b) (4) among 3 treatment groups (b) (4) 4.0, and 4.0%, respectively). For the JRA 30 core set of endpoints (b) (4) Naproxen had significantly greater improvement from baseline in number of joints with limited range of motion than rofecoxib treatment groups. However, a similar pattern of difference, although not statistically significant, in the opposite direction was observed for parent/patient global assessment of well being, where it was analyzed as a continuous, rather than a dichotomous variable.

3.1.1.7 Reviewer's analyses and Conclusions

In order to verify sponsor's results, this reviewer recalculated the confidence intervals for the primary efficacy variable. This reviewer's results confirm sponsor's results.

The analysis results from both ITT and completers showed that for both high dose of rofecoxib vs. naproxen and low dose rofecoxib vs. naproxen the entire 95% CIs were above the pre-assigned non-inferiority margin (b) (4) However, as mentioned earlier the Division considered the non-inferiority margin (b) (4) In ITT population the lower limits of the 95% CIs were 0.76 for higher dose rofecoxib vs. naproxen (b) (4) The medical officer's preference of non-inferiority margin is 0.75 (see medical officer's review). Also, in many subgroups the low dose group did not establish non-inferiority even with a non-inferiority margin (b) (4)

(b) (4)

3.2 EVALUATION OF SAFETY

3.2.1 SPONSOR'S ANALYSIS OF SAFETY DATA

The clinical adverse experience profile is summarized by assigned treatment groups in Table 8 in Appendix 1. Clinical adverse experiences were reported by 196 (63.2%) of 310 patients in the combined base study. One or more clinical adverse experiences occurred in 72 (66.1%), 61 (61.0%),

and 63 (62.4%) patients in the lower-dose rofecoxib, higher dose rofecoxib, and naproxen groups, respectively. Drug-related (determined by the investigator to be possibly, probably, or definitely drug related) clinical adverse experiences occurred in 21 (19.3%), 22 (22.0%), and 28 (27.7%) patients in the lower-dose rofecoxib, higher dose rofecoxib, and naproxen groups, respectively. Serious adverse experiences occurred in 1 (0.9%), 2 (2.0%), and 1 (1.0%) patients in the lower-dose rofecoxib, higher dose rofecoxib, and naproxen groups, respectively. Of these patients, 1 (lower-dose rofecoxib treatment group) discontinued the study. None of the serious adverse experiences was determined to be drug related. No patients died during the study. In total 3 (2.8%), 0 (0.0%), and 2 (2.0%) patients in the lower-dose rofecoxib, higher dose rofecoxib, and naproxen groups, respectively, discontinued study drug due to clinical adverse experiences. Of the patients who discontinued study drug due to adverse experiences, 2 (1.8%), 0 (0.0%), and 2 (2.0%) in the lower-dose rofecoxib, higher dose rofecoxib, and naproxen groups, respectively discontinued due to drug-related adverse experiences. No patient was discontinued due to a serious drug-related adverse experience. The three most commonly reported adverse experiences were abdominal pain, upper abdominal pain, and headache. Drug-related adverse experiences occurred most frequently in the gastrointestinal system. Eighteen (16.5%), 18 (18.0%), and 19 (18.8%), of the patients in the lower-dose rofecoxib, higher dose rofecoxib, and naproxen treatment groups experienced drug-related digestive system adverse experiences. The drug-related adverse experiences most frequently seen in this system were abdominal pain and upper abdominal pain. Drug-related adverse experiences in the nervous system occurred in 3 (2.8%), 3 (3.0%), and 9 (8.9%) patients in the lower-dose rofecoxib, higher dose rofecoxib, and naproxen treatment groups, respectively. The excess of drug-related adverse experiences in the naproxen group was mostly attributable to the incidence of headache which was higher in the naproxen group with 6 (5.9%) of the patients reporting this adverse experience. Three (2.8%) and 1 (1.0%) of the patients in the lower-dose rofecoxib, and higher dose rofecoxib treatment groups, respectively, had drug-related adverse experiences of headache. The adverse experience profile in 2 to 11 year olds and 12 to 17 year olds was similar to the overall population.

Five patients discontinued due to clinical adverse experiences: 3 (3.0%) in the lower-dose rofecoxib treatment group and 2 (2.0%) in the naproxen treatment group. Of the 3 patients in the lower-dose rofecoxib treatment group, 2 patients discontinued due to clinical adverse experiences of abdominal pain, which were determined by the investigator to be study-drug related. The third patient discontinued due to worsening of juvenile rheumatoid arthritis, which was determined by the investigator to be non-study-drug related. Of the 2 patients in the naproxen treatment group, AN 391 discontinued due to a clinical adverse experience of migraine, which was determined by the investigator to be related to study drug. AN 475 discontinued due to a clinical adverse experience of hematochezia, which the investigator determined to be related to study drug (Table 49).

Nonfatal serious clinical adverse experiences occurred in 4 (1.3%) of 310 patients. The incidence of serious clinical adverse experiences was 1, 2, and 1 in the lower-dose rofecoxib, higher dose rofecoxib, and naproxen treatment groups, respectively. None of the serious adverse experiences was determined by the investigator to be drug related. There were no patient deaths in this study.

3.2.2 REVIEWER'S ANALYSIS OF SAFETY DATA

This reviewer did not perform any analysis on the safety data. This reviewer refers to the clinical review for safety analysis.

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

The sponsor performed the following subgroup analysis: Joint involvement (Pauci articular, Poly articular), Age group (2 to 11 year olds, 12 to 17 year olds), Gender (Female, Male), Tanner stage (1, 2, 3, 4, 5), Ethnic group (Black, Caucasian, Hispanic, Multi-racial, Other), Duration of juvenile rheumatoid arthritis (< median years, ≥median years), Erythrocyte sedimentation rate (0 to 20, ≥20), Baseline methotrexate user (Yes, No), Baseline low-dose corticosteroid user (Yes, No), Baseline disease-modifying anti-rheumatic drugs user (Yes, No), Prior naproxen user (Yes, No), and Prior. NSAID users.

The results showed that the therapeutic effect in 2 to 11 year old JRA patients and in 12 to 17 year old JRA patients were similar to those in the combined population of 2 to 17 year old JRA patients except for the following:

- Similar to results in the entire 2 to 17 year old population, for 12 to 17- year olds, the proportion of patients meeting the JRA 30 criteria in the higher dose rofecoxib group was non-inferior to that in the naproxen treatment group. In the lower dose rofecoxib treatment group fewer 12 to 17 year olds responded to treatment, as defined by the JRA 30 criteria. The point estimate of the ratio of response rates, relative to naproxen was 0.63 and the lower limit of the 95% CI was 0.4, less than the pre-specified margin of 0.5.
- For 2 to 11 year olds, the rofecoxib treatment groups had greater improvement from baseline in parent/patient's global assessment of pain than the naproxen treatment group.
- For 12 to 17 year olds, the lower-dose rofecoxib treatment group had a smaller treatment effect than the naproxen treatment group in the number of joints with active arthritis.

(b) (4)

The sponsor also evaluated the treatment effect among patients who completed the 12 week base study. (b) (4)

4.1 REVIEWER'S ANALYSIS OF SUBGROUP POPULATION:

On the advice of the medical officer this reviewer perform subgroup analysis based on the age and bodyweight subgroups. The medical officer selected these subgroups based on some clinical importance. Tables 9 and 10 show this reviewer's analysis results for age and bodyweight subgroups.

Results of this reviewer's analysis showed decreasing treatment effect with increasing age and increasing bodyweight.

5 SUMMARY AND CONCLUSIONS

5.1 STATISTICAL ISSUES AND COLLECTIVE EVIDENCE

In this submission the sponsor included report of a Phase 3 study, namely Study #P134 (Protocol134/135). This was a 12 week, double-blind, double dummy; active comparator controlled study in 2 to 17 years old juvenile rheumatoid arthritis patients. In each age group, patients were allocated to receive 1 of the 3 treatments: (1) lower dose rofecoxib, 0.3 mg/kg/day for 2 to 11 year old patients (not to exceed 12.5 mg/day), or 12.5 mg daily for 12 to 17 year old patients; (2) higher dose rofecoxib, 0.6 mg/kg/day for 2 to 11 year old patients (not to exceed 25 mg/day), or 25 mg daily for 12 to 17 year old patients; (3) naproxen, targeted to 15 mg/kg/day. The 2 to 11 year old patients received suspension formulations, and the 12 to 17 year old patients received tablets. The primary objectives of this study were to examine the therapeutic effects and safety of two doses of rofecoxib and to show the non-inferiority of rofecoxib compared to naproxen.

The primary efficacy endpoint was proportion of patients in the ITT population meeting the JRA 30 criteria. The non-inferiority is claimed if the 95% confidence interval on the relative risk (ratio of percent of patients meeting the JRA 30 criteria) of rofecoxib vs. naproxen was entirely (b) (4) However, the Division considered the non-inferiority margin (b) (4)

Since there was only one study, the collective evidence is based on only one study. The results showed that in ITT population the lower limits of the 95% CIs were 0.76 for higher dose rofecoxib vs. naproxen (b) (4)

5.2 CONCLUSIONS AND RECOMMENDATIONS

The primary efficacy endpoint was proportion of patients in the ITT population meeting the JRA 30 criteria. The non-inferiority is claimed if the 95% confidence interval on the relative risk (ratio of percent of patients meeting the JRA 30 criteria) of rofecoxib vs. naproxen was entirely (b) (4)

(b) (4)

However, the Division considered the non-inferiority margin (b) (4) In ITT population the lower limits of the 95% CIs were 0.76 for higher dose rofecoxib vs. naproxen (b) (4)

The medical officer's preference of non-inferiority margin is 0.75 (see medical officer's review). Also, in many subgroups the low dose group did not establish non-inferiority even with a non-inferiority margin (b) (4)

(b) (4)

M. Atiar Rahman, Ph.D.
Mathematical Statistician

Concur: Stan Lin, Ph.D.
Team Leader

cc:
Archival NDA 21-565
HFD-550/Division File
HFD-550/Dr. Harvey
HFD-550/Dr. Hertz
HFD-550/ Dr. Chambers
HFD-550/ Dr. Yancey
HFD-550/ Mr. Rodrigues
HFD-725/ Chron
HFD-725/ Dr. Huque
HFD-725/ Dr. Lin
HFD-725/ Dr. Rahman
HFD-700/Dr. Anello

6 APPENDIX-1

6.1 TABLE 1: PATIENT DISPOSITION

Overall Disposition of Patients

Time Frame	Rofecoxib 0.3 mg/kg or Rofecoxib 12.5 mg (N=109)	Rofecoxib 0.6 mg/kg or Rofecoxib 25 mg (N=100)	Naproxen 15 mg/kg (N=101)
Treatment and Post-study	n=109	n=100	n=101
patient completed	26	18	14
patient discontinued	10	5	10
clinical AE	3	0	3
laboratory AE	3	1	0
lack efficacy	3	4	4
lost to follow-up	0	0	3
patient discontinued for other	1	0	0
patient extended	73	77	77

Although patients are counted only once within a Time Frame, patients may be counted in more than one Time Frame.
Source Table 4.31.7 of sponsor's analysis

6.2 TABLE 2: BASELINE DEMOGRAPHIC CHARACTERISTICS

	Rofecoxib 0.3 mg/kg or Rofecoxib 12.5 mg (N = 109)		Rofecoxib 0.6 mg/kg or Rofecoxib 25 mg (N = 100)		Naproxen 15 mg/kg (N = 101)		Total (N = 310)	
	n	(%)	n	(%)	n	(%)	n	(%)
Gender								
Female	83	(76.1)	70	(70.0)	74	(73.3)	227	(73.2)
Male	26	(23.9)	30	(30.0)	27	(26.7)	83	(26.8)
Age								
1 and Under	0	(0.0)	0	(0.0)	0	(0.0)	0	(0.0)
2 to 4	15	(13.8)	22	(22.0)	9	(8.9)	46	(14.8)
5 to 11	50	(45.9)	38	(38.0)	47	(46.5)	135	(43.5)
12 to 17	44	(40.4)	40	(40.0)	45	(44.6)	129	(41.6)
Over 12 to 17	0	(0.0)	0	(0.0)	0	(0.0)	0	(0.0)
Mean		9.7		9.4		10.7		9.9
SD		4.26		4.27		3.99		4.20
Median		10.0		10.0		11.0		10.0
Range		2-17		2-16		2-17		2-17
Race								
Asian	0	(0.0)	1	(1.0)	0	(0.0)	1	(0.3)
Black	1	(0.9)	4	(4.0)	9	(8.9)	14	(4.5)
Eurasian	1	(0.9)	0	(0.0)	0	(0.0)	1	(0.3)
European	0	(0.0)	1	(1.0)	0	(0.0)	1	(0.3)
Hispanic American	6	(5.5)	4	(4.0)	5	(5.0)	15	(4.8)
Indian	1	(0.9)	0	(0.0)	0	(0.0)	1	(0.3)
Multi-Racial	15	(13.8)	20	(20.0)	16	(15.8)	51	(16.5)
Polynesian	0	(0.0)	1	(1.0)	0	(0.0)	1	(0.3)
White	85	(78.0)	69	(69.0)	71	(70.3)	225	(72.6)

Source Table 4.31.10 of sponsor's analysis

6.3 TABLE 3: ENDPOINTS AND THEIR STATISTICAL ANALYSES

Endpoint	Statistical Method	Analysis Approaches
Primary		
Proportion of Patient Meeting the JRA 30 Criteria	Mantel-Haenszel method	MITT and PP
Key Secondary		
Proportion of Patient with Improvement from Baseline in Parent/Patient's Assessment of Overall Well-being	Mantel-Haenszel method	MITT
Other Secondary		
Patient's Global Assessment of Pain	ANCOVA	MITT
Discontinuation Due to Lack of Efficacy	Fisher's exact test	MITT
JRA 30 Core Set of Variables:		
Parent/Patient's Assessment of Overall Well-Being	ANCOVA	MITT
Investigator's Global Assessment of Disease Activity	ANCOVA	MITT
Functional Ability (CHAQ)	ANCOVA	MITT
Number of Joints With Active Arthritis	ANCOVA	MITT
Number of Joints With Limited Range of Motion	ANCOVA	MITT
Erythrocyte Sedimentation Rate	ANCOVA (log scale)	MITT

ANCOVA = Analysis of Covariance.
CHAQ = Child Health Assessment Questionnaire.
MITT = Modified Intention To Treat.
PP = Per Protocol.
Data Source: [3.4]
Source Table 10 of sponsor's analysis

6.4 TABLE 4: SPONSOR'S ANALYSES OF PRIMARY EFFICACY VARIABLES

(Modified Intention-to-Treat Approach)

JRA 30 Responder During the 12 week Base Study: Regardless of Completion Status (Primary)†

Treatment	Frequency (%)	(b) (4)
Higher dose Rofecoxib	54 /99	(54.5)
Naproxen	54 /98	(55.1)
Between-Group Comparison	Relative Risk‡ (95% CI)	Difference§ (95% CI)
Higher Dose Rofecoxib vs. Naproxen	0.98 (0.76, 1.26)	-1.3 (-15.1, 12.5)

(b) (4)

JRA 30 Responder and Completer (Secondary)

Treatment	Frequency (%)	(b) (4)
Higher dose Rofecoxib	54 /99	(54.5)
Naproxen	53 /99	(53.5)
Between-Group Comparison	Relative Risk‡ (95% CI)	Difference§ (95% CI)
Higher Dose Rofecoxib vs. Naproxen	1.00 (0.78, 1.29)	0.1 (-13.7, 13.8)

(b) (4)

† The numerator is number of patients who met the JRA 30 criteria; the denominator is the number of patients with evaluable JRA 30 criteria.

‡ From Mantel-Haenszel estimate with protocol, age group, and joint involvement as stratification factors.

§ From the normal approximation for a Cochran-Mantel-Haenszel weighted average of the differences over all strata.

In order to be a responder, the patient had to complete the 12 week study and meet the JRA 30 criteria; but to be a non-responder; the patient either did not complete the 12 week study or did not meet the JRA 30 criteria. AN 10 in the naproxen treatment group discontinued and his JRA 30 response criteria could not be evaluated because his efficacy measurements were not collected during on-treatment period. This patient was counted as a non-responder in the secondary analysis because the patient had discontinued. However, since his JRA 30 criteria could not be evaluated, this patient could not be included in the primary analysis.

JRA = Juvenile Rheumatoid Arthritis.

Source Table 21 of sponsor's analysis

6.5 TABLE 5: SPONSOR'S ANALYSES OF PRIMARY EFFICACY VARIABLES

(Per-Protocol)

JRA 30 Responder: Regardless of Completion Status (Primary)[†]

Treatment	Frequency (%)	(b) (4)
Higher dose Rofecoxib	52 /90	(57.8)
Naproxen	48 /87	(55.2)
Between-Group Comparison	Relative Risk[‡] (95% CI)	Difference[§] (95% CI)
Higher Dose Rofecoxib vs. Naproxen	1.04 (0.80, 1.35)	2.3 (-12.0, 16.6)

(b) (4)

JRA 30 Responder and Completer (Secondary)

Treatment	Frequency (%)	(b) (4)
Higher dose Rofecoxib	51 /92	(55.4)
Naproxen	47 /94	(50.0)
Between-Group Comparison	Relative Risk[‡] (95% CI)	Difference[§] (95% CI)
Higher Dose Rofecoxib vs. Naproxen	1.08 (0.83, 1.42)	4.2 (-9.9, 18.2)

(b) (4)

† The numerator is number of patients who met the JRA 30 criteria; the denominator is the number of patients with evaluable JRA 30 criteria.

‡ From Mantel-Haenszel estimate with protocol, age group, and joint involvement as stratification factors.

§ From the normal approximation for a Cochran-Mantel-Haenszel weighted average of the differences over all strata.

In order to be a responder, the patient had to complete the 12 week study and meet the JRA 30 criteria; but to be a non-responder; the patient either did not complete the 12 week study or did not meet the JRA 30 criteria. Four patients (ANs 237, 279, 552, and 636) on lower-dose Rofecoxib, 2 patients (ANs 4 and 128) on the higher dose Rofecoxib, and 7 patients (ANs 1, 10, 116, 293, 312, 391, and 475) on naproxen discontinued and their JRA 30 response criteria could not be evaluated because insufficient efficacy measurements were collected. These patients were counted as non-responders in the secondary analysis because they had discontinued. However, since these patients' JRA 30 criteria could not be evaluated, they could not be included in the primary analysis.

JRA = Juvenile Rheumatoid Arthritis.

Data Source: [4.3]

6.6 TABLE 6: PARENT/PATIENT'S ASSESSMENT OF OVERALL WELL-BEING

(Pop: Modified Intent to Treat)

Treatment	Frequency†	(%)
		(b) (4)
Higher dose Rofecoxib	76 /100	(76.0)
Naproxen	73 /100	(73.0)
Between-Group Comparison	Relative Risk‡ (95% CI)	Difference§ (95% CI)
Higher dose Rofecoxib vs. Naproxen	1.04 (0.89, 1.22)	3.1 (-8.8, 15.0)
		(b) (4)

† Frequency = m/n, where n is the total number of patients with nonmissing values, and m is the number of patients with improvement from baseline in patient/parent's assessment of overall well-being.
‡ From Mantel-Haenszel estimate with protocol, age group, and joint involvement as stratification factors.
§ From the normal approximation for a Cochran-Mantel-Haenszel (CMH) weighted average of the differences over all strata.

Data Source: [4.3]

6.7 TABLE 7: PATIENT'S GLOBAL ASSESSMENT OF PAIN

(Modified Intention-to-Treat Approach)

Treatment Group	N	Baseline Mean	Treatment Period Mean	Mean Change	SD of Change	LS Mean† Change	95% CI for LS Mean† Change
							(b) (4)
Higher dose Rofecoxib	100	41.85	28.24	-13.61	24.51	-13.12	(-16.75, -9.48)
Naproxen	100	42.71	33.60	-9.11	22.49	-8.43	(-11.98, -4.88)

Comparisons Between Treatment Groups	Difference in LS Mean	95% CI for Diff.	p-Value
Between Active Treatments			
Higher dose Rofecoxib vs. Naproxen	-4.69	(-9.68, 0.30)	0.065
			(b) (4)

Effect:	p-Value	Pooled SD
Baseline Covariate	<0.001	17.83
Protocol	0.140	
Age Group	0.647	
Joint Involvement	0.937	
Treatment	0.132	

† Least-squares mean.

Data Source: [4.3]

6.8 TABLE 8: SPONSOR'S ANALYSES OF SAFETY DATA

	Rofecoxib 0.3 mg/kg or Rofecoxib 12.5 mg (N=109)		Rofecoxib 0.6 mg/kg or Rofecoxib 25 mg (N=100)		Naproxen 15 mg/kg (N=101)	
	n	(%)	n	(%)	n	(%)
Number (%) of patients:						
with one or more adverse experiences	72	(66.1)	61	(61.0)	63	(62.4)
with no adverse experience	37	(33.9)	39	(39.0)	38	(37.6)
with drug-related adverse experiences†	21	(19.3)	22	(22.0)	28	(27.7)
with serious adverse experiences	1	(0.9)	2	(2.0)	1	(1.0)
with serious drug-related adverse experiences	0	(0.0)	0	(0.0)	0	(0.0)
who died	0	(0.0)	0	(0.0)	0	(0.0)
discontinued due to adverse experiences	3	(2.8)	0	(0.0)	2	(2.0)
discontinued due to drug-related adverse experiences	2	(1.8)	0	(0.0)	2	(2.0)
discontinued due to serious adverse experiences	1	(0.9)	0	(0.0)	0	(0.0)
discontinued due to serious drug-related adverse experiences	0	(0.0)	0	(0.0)	0	(0.0)

† Determined by the investigator to be possibly, probably, or definitely drug related.

Data Source: [4.1; 4.20; 4.32]

6.9 TABLE 9: SUBGROUP ANALYSIS OF JRA 30 RESPONDER RATES BY AGE

Assuming missing value as missing			
Dose Groups	Age Sub-Group (Years)	Number of Patients	Relative Risk (95% CI)
Higher Dose Rofecoxib vs. Naproxen	All Age Group	54/99, 54/98	0.99 (0.76, 1.28)
Higher Dose Rofecoxib vs. Naproxen	2≤ Age ≤5	15/24, 6/11	1.15 (0.64, 2.87)
Higher Dose Rofecoxib vs. Naproxen	6≤ Age ≤11	18/35, 22/42	0.98 (0.62, 1.53)
Higher Dose Rofecoxib vs. Naproxen	12≤ Age ≤17	21/40, 26/45	0.91 (0.61, 1.34)
[redacted]			(b) (4)
Assuming missing value as failure			
Dose Groups	Age Sub-Group (Years)		Relative Risk (95% CI)
Higher Dose Rofecoxib vs. Naproxen	All Age Group	54/100, 54/101	1.01 (0.78, 1.31)
Higher Dose Rofecoxib vs. Naproxen	2≤ Age ≤5	15/25, 6/12	1.20 (0.65, 3.06)
Higher Dose Rofecoxib vs. Naproxen	6≤ Age ≤11	18/35, 22/44	1.03 (0.65, 1.61)
Higher Dose Rofecoxib vs. Naproxen	12≤ Age ≤17	21/40, 26/45	0.91 (0.60, 1.35)
[redacted]			(b) (4)

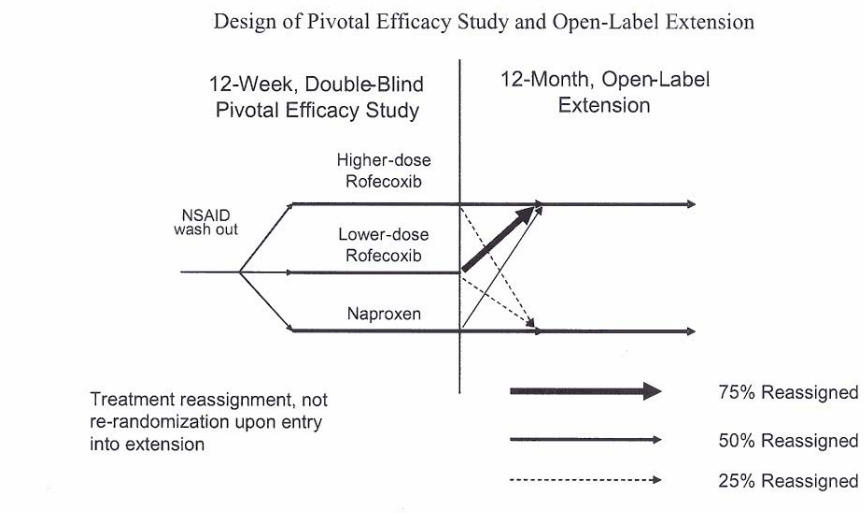
Reviewer's table

6.10 TABLE 10: SUBGROUP ANALYSIS OF JRA 30 RESPONDER RATES BY BODYWEIGHT

Dose Groups	Bodyweight Sub-Group (Kg)	Age (Years) n, Mean (Min, Max)	Number of Patients	Relative Risk (95% CI)
Assuming missing value as missing				
Higher Dose Rofecoxib vs. Naproxen	All Bodyweight Group	197, 10.15 (2.00, 17.00)	54/99, 54/98	0.99 (0.76, 1.28)
Higher Dose Rofecoxib vs. Naproxen	10≤ Bodyweight <20	41, 4.05 (2.00, 7.00)	15/25, 9/16	1.07 (0.62, 2.10)
Higher Dose Rofecoxib vs. Naproxen	20≤ Bodyweight <40	75, 9.77 (6.00, 15.00)	20/38, 19/37	1.03 (0.64, 1.63)
Higher Dose Rofecoxib vs. Naproxen	Bodyweight<=40	81, 13.58 (3.00, 17.00)	19/36, 26/45	0.91 (0.59, 1.36)
[redacted]				(b) (4)
Assuming missing value as failure				
Higher Dose Rofecoxib vs. Naproxen	All Bodyweight Group	201, 10.07 (2.00, 17.00)	54/100, 54/101	1.01 (0.78, 1.31)
Higher Dose Rofecoxib vs. Naproxen	10≤ Bodyweight <20	43, 4.00 (2.00, 7.00)	15/26, 9/17	1.09 (0.62, 2.16)
Higher Dose Rofecoxib vs. Naproxen	20≤ Bodyweight <40	76, 9.76 (6.00, 15.00)	20/38, 19/38	1.05 (0.65, 1.67)
Higher Dose Rofecoxib vs. Naproxen	Bodyweight >=40	82, 13.55 (3.00, 17.00)	19/36, 26/46	0.93 (0.60, 1.40)
[redacted]				(b) (4)

Reviewer's table

Figure 1



Source: Figure 2.7.3:1 of sponsor's submission

7 APPENDIX- 2

7.1 STATISTICAL ANALYSIS PLAN

I. DATA ANALYSIS

Statistical analysis of the data from this study will be the responsibility of the Clinical Biostatistics department of Merck Research Laboratories.

Hypotheses

The study hypotheses are as follows:

1. The proportion of patients that demonstrate improvement will be similar between rofecoxib and naproxen treatment groups. The primary endpoint for evaluating efficacy at the onset was the parent/patient's assessment of overall well being; however, as explained in the background section, prior to database unblinding, the primary endpoint was changed to JRA 30.

   Criteria for evaluation: The 95% CI for ratio of percent of patients demonstrating improvement from baseline (rofecoxib versus naproxen) will lie entirely (b) (4)

2. Administration of rofecoxib to children with JRA will be safe and well tolerated.

   Criteria for evaluation: Differences between rofecoxib doses and naproxen in proportions of patients with adverse experiences or exceeding predefined limits of change in laboratory or vital sign variables will be assessed in the context of the magnitude of the proportions and differences observed and their clinical relevance.

Addressed objectives:

1. Mean change from baseline (averaged over all times of observation) will be compared between the 2 doses of rofecoxib in the 2- through 11 year old JRA patients, stratified by joint involvement (pauci versus poly).
2. Mean change from baseline (averaged over all times of observation) will be compared between the 2 doses of rofecoxib in the 12- through 17 year old JRA patients, stratified by joint involvement (pauci versus poly).
3. The safety and tolerability of rofecoxib in children with JRA will be assessed as described for the second hypothesis above.
4. The safety and efficacy profile of naproxen for treatment of JRA will be assessed and compared with that of rofecoxib as described for the hypotheses above.
5. The effects of treatment in patients with pauci articular and poly articular disease will be assessed by stratification by this factor and assessing treatment-by-joint status (pauci versus poly) interaction.

Variables/Time Points of Interest (Metric, Parameter)

The efficacy variables are (in priority order):

1. JRA 30
2. Parent/Patient's Assessment of Overall Well Being
3. Investigator's Global Assessment of Disease Activity
4. Functional Ability (CHAQ)
5. The Parent/Patient's Global Assessment of Pain
6. Number of Joints With Active Arthritis
7. Number of Joints With Limited Range of Motion
8. Proportion of patients discontinuing due to lack of efficacy
9. ESR

Initially, the parent/patient's assessment of overall well being was the primary endpoint; others were considered secondary. The primary endpoint has been replaced by JRA 30 after completion of the clinical portion of the trial, but prior to unblinding the database-refer to Background section for more detail. Parent/Patient's Assessment of Overall Well Being will also be assessed similarly to JRA 30, and inserted in the above priority list immediately after JRA 30, which will be first in the priority list. Note that power for JRA 30 is the same as for Parent/Patient's Assessment of Overall Well Being since both are binary endpoints; thus, the power section of this data analysis section is not revised; however, it is understood to apply to JRA 30.

The primary measure of improvement for each endpoint will be time-weighted average change from baseline across all treatment visits (Visit 3.0 through Visit 6.0 and any unscheduled visits between 2.0 and 6.0). Visit 2.0 is considered baseline. In addition, mean change from baseline (±SE) by treatment group will be summarized at each observation week in single variable plots; for these plots only, missing values will be imputed via the last value carried forward technique. In addition to the between-treatment group comparisons of proportions of patients with AEs and exceeding predefined limits of change in safety parameters as described above, means ±SE will be plotted over time for each laboratory and vital signs parameter.
For safety, the following list is of primary interest: discontinuations due to digestive adverse experiences or abdominal pain, clinical adverse experiences of hypertension and blood pressure increased, clinical adverse experiences of fluid retention and edema, laboratory adverse experiences of increased serum creatinine and increased serum hepatic transaminases (ALT and AST). Statistical significance testing for between-treatment group differences will be carried out for these endpoints. No significance testing will be carried out for other safety endpoints; their clinical relevance will be assessed on the basis of magnitude of effects using confidence intervals.

Approaches to Analyses

The primary analysis will be a modified intent-to-treat approach. All patients who take at least 1 dose of study drug and provide baseline and at least 1 postbaseline response will be included in the analysis of efficacy. Dropouts for various reasons are not unexpected. Dropouts will be included in the primary analysis based on their responses obtained up to and including the time of discontinuation. The primary analysis of clinical efficacy will be based on a stepdown procedure (high-dose rofecoxib versus naproxen first, and if the CI >0.5, followed by low-dose rofecoxib versus naproxen).

If the number of major protocol violations in Part I is not negligible, then a secondary analysis with protocol violations removed will be carried out. All protocol violations will be identified, and a decision about the need for a perprotocol

analysis will be made prior to the unblinding of the data. The list of major protocol violations will be documented prior to unblinding the data.

Statistical Methods

The ratio of proportions of patients meeting the JRA30 criteria and the ratio of patients with improvement from baseline in parent/patient's assessment of overall well-being assessed using a Mantel-Haenszel ratio of rates. Discontinuations due to lack of efficacy, and for each primary safety endpoint will be assessed using Fisher's exact test since their expected rates are lower than those for the dichotomized efficacy endpoints. All individual efficacy variables except proportions of patients will be assessed by ANOVA (model to include terms for joint involvement stratum, age stratum, baseline covariate, and treatment group). The interactions with treatment will be evaluated; if significant at the 0.05 level, their qualitative nature will be assessed using exploratory data analytic techniques. Least-squares mean differences between rofecoxib doses and naproxen will be compared via t-tests derived from the ANOVA. Assumptions of normality and homogeneity will be assessed by the Shapiro-Wilk statistic and Levene's test.

Multiplicity

No adjustment for multiplicity is needed because there is only 1 primary hypothesis for efficacy. Making no multiplicity adjustment for safety is conservative, and enhances power to find untoward effects if present. Hence, no multiplicity adjustment will be made.

Sample Size and Power Calculations

The sample size N=75 per dose group has at least 90% power to yield the 95% CI for the ratio of percent of patients improved greater than 0.5, if the true rates are equal for rofecoxib and naproxen and exceed 40%. This was computed using the log transformation of the ratio of 2 binomial rates and the normal approximation to the binomial distributions.

Interim Analyses

The sample size may be adjusted during the trial based on blinded assessment of the overall study response rate. This is because the variance of the rate ratio depends on the rates. Since this type of adjustment is made blinded to treatment, and there will be only 1 final unblinded analysis, there will be no adjustment to the alpha level of 0.05. There will be no unblinded interim analysis.

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

/s/

Atiar Rahman
5/21/04 07:09:26 PM
BIOMETRICS

Stan Lin
5/28/04 01:49:10 PM
UNKNOWN