

**CENTER FOR DRUG EVALUATION AND
RESEARCH**

APPLICATION NUMBER:

21-897

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: 21-897
Drug Name: Medisorb Naltrexone
Indication(s): Treatment of alcohol dependence
Applicant: Alkermes, Incorporated
Date(s): Received 03/31/05; user fee (6 months) 09/30/05; extended date (major amendment) 12/30/05;
Review Priority: Priority
Biometrics Division: Division of Biometrics II
Statistical Reviewer: Dionne L. Price, Ph.D.
Concurring Reviewers: Thomas J. Permutt, Ph.D.
Medical Division: Division of Anesthesia, Analgesia, and Rheumatology Products
Clinical Team: Mwango Kashoki, M.D.
Celia Winchell, M.D.
Project Manager: Lisa Basham-Cruz

Keywords: NDA review, clinical studies

Table of Contents

1 EXECUTIVE SUMMARY	3
1.1 CONCLUSIONS AND RECOMMENDATIONS	3
1.2 BRIEF OVERVIEW OF CLINICAL STUDIES	3
1.3 STATISTICAL ISSUES AND FINDINGS	4
2. INTRODUCTION	6
2.1 OVERVIEW.....	6
2.2 DATA SOURCES	6
3. STATISTICAL EVALUATION	7
3.1 EVALUATION OF EFFICACY	7
3.2 EVALUATION OF SAFETY	20
4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS	21
4.1 GENDER, RACE AND AGE	21
4.2 OTHER SPECIAL/SUBGROUP POPULATIONS	21
5. SUMMARY AND CONCLUSIONS	23
5.1 STATISTICAL ISSUES AND COLLECTIVE EVIDENCE	23
5.2 CONCLUSIONS AND RECOMMENDATIONS	24
5.2.1 Labeling.....	24
APPENDICES.....	27
DYNAMIC RANDOMIZATION ALGORITHM.....	28
SUMMARY OF DEMOGRAPHIC AND BASELINE CHARACTERISTICS	31
STATISTICAL FORMULATION OF MODEL	35
RE-RANDOMIZATION ANALYSIS RESULTS	37
EXPLORATION OF PATIENT DISCONTINUATION	38
ANALYSES REPEATED USING ALTERNATE DEFINITION OF A HEAVY DRINKING DAY	40

1 EXECUTIVE SUMMARY

1.1 Conclusions and Recommendations

Alkermes, Incorporated proposes Medisorb Naltrexone for the treatment of alcohol dependence. Based on an evaluation of the event rate of heavy drinking over 24 weeks, the applicant claims that Medisorb Naltrexone 380 mg reduces heavy drinking. My review of the statistical evidence suggests support for the claim. However, I believe that several additional factors warrant consideration when assessing Medisorb Naltrexone. First, protocol violations were identified at two of the three sites inspected by the Division of Scientific Investigations. Alkermes' failure to identify these violations prior to the submission of the NDA diminished my confidence in the overall conduct of the study and resulting data. Furthermore, analyses of the data including and excluding the sites with violations resulted in inconsistent findings further adding to my concern. Since support for Medisorb Naltrexone was derived from a single study, there was no replication of the findings to provide additional assurance. Lastly, multiple safety concerns, such as elevated transaminases and severe allergic reactions, were identified by the review team. While there is statistical evidence that the drug is active, the previously mentioned factors must be assessed collectively by the review team in order to evaluate the risks and benefits of Medisorb Naltrexone. In my opinion, this task is further complicated by the uncertainty surrounding the overall conduct of the study and resulting data.

1.2 Brief Overview of Clinical Studies

Oral naltrexone is approved for the management of alcohol dependence. Alkermes proposes an injectable depot formulation of naltrexone, namely Medisorb Naltrexone. The applicant asserts that Medisorb Naltrexone provides continued exposure for at least a month and may reduce the potential for hepatotoxicity associated with the oral formulation. The drug was introduced to the Division of Anesthesia, Analgesia, and Rheumatology Products via IND 61,138. The clinical development plan, endpoints, and statistical analyses were discussed during several meetings between the applicant and the division.

Prior to submission of the NDA, the applicant sought input from the division regarding the needed number of studies. At that time, the applicant proposed a single study to support the use of the drug. The division stated that two adequate and well-controlled studies were necessary unless the application was submitted under Section 505(b)(2) of the Food, Drug, and Cosmetic Act. On 31 March 2005, Alkermes submitted NDA 21-897 (pursuant to Section 505(b)(2)) in support of Medisorb Naltrexone. The application included a single, double-blind, placebo-controlled, multi-center study and relied on the agency's previous findings of efficacy for oral naltrexone. In the study, patients were randomized to intramuscular injections of Medisorb Naltrexone 190 mg, Medisorb Naltrexone 380 mg, or placebo. Patients randomized to placebo received a matching volume of Medisorb microspheres (i.e. 2 mL or 4 mL) without naltrexone.

Moreover, patients were allocated to treatment for balance on four baseline characteristics using a dynamic randomization scheme. Treatment was administered, along with biopsychosocial support therapy (using the BRENDA approach), during clinic visits occurring every four weeks for the duration of 24 weeks. Patients recorded their alcohol consumption using the timeline follow-back method (TLFB). The primary measure of efficacy was the event rate of heavy drinking over 24 weeks of treatment where a heavy drinking day was defined as a day on which a man consumed at least five drinks or a woman consumed at least four drinks. The applicant defined the event rate as the number of heavy drinking days divided by the number of days at risk for heavy drinking. Additionally, an alcoholic drink was defined as 13.6 grams of absolute ethanol. The applicant employed a stratified Andersen-Gill model for the primary analysis.

1.3 Statistical Issues and Findings

Since the event of interest (i.e. heavy drinking) could potentially occur on multiple days, the applicant employed an Andersen-Gill model to assess the overall effect of treatment. In general, the results produced by the model may be influenced by the non-proportionality of the hazard functions and/or by patient withdrawal that is treatment related. Thus prior to the submission of the NDA, the Division recommended that the applicant consider and propose methodology for use in the event that the proportional hazards (PH) assumption was seriously violated. Moreover, the applicant was urged to conduct a re-randomization test to validate the model inferences. The division additionally suggested the applicant justify and specify how missing data would be handled. To address the former recommendation regarding the PH assumption, the applicant used a stratified Andersen-Gill model. According to the applicant, "A stratified analysis adjusted for different baseline 'hazards' of the prespecified stratification factors. In this way, the treatment effect was not subject to the distortion that a covariate-by-time interaction would induce by inclusion of such a covariate in the model." The applicant additionally proposed a nonparametric Wilcoxon test as an alternative method of analysis if the PH assumption was violated. Alkermes formally tested the assumption by inclusion of an interaction term in the model. To address potential missing data concerns, Alkermes assessed the randomness of the missing data via evaluations of the event rate of heavy drinking by the number of doses received, the Kaplan-Meier curves, and a pattern mixture model.

According to the applicant, there was evidence of a severe violation of the proportional hazards assumption, both overall and for some strata. Additionally, the applicant stated that the re-randomization test based on the stratified Andersen-Gill model produced unstable results because of the small sizes of some of the strata. Based on the evaluation of drop-outs, the applicant concluded that study discontinuations were comparable across treatment groups and were therefore less likely to affect conclusions. I was not convinced that the violation of the proportional hazards could be ignored, nor was I convinced that the missing data occurred randomly. Thus, I focused significant attention on the nonparametric analysis. The nonparametric analysis conducted by the applicant essentially employed a last observation carried forward strategy for missing data. Since I had some concern regarding the possibility that patients withdrew for treatment-related reasons, I performed an additional analysis imputing heavy drinking days for all missing data days. My collective evaluation of the analyses and results suggested the existence of a treatment effect for the 380 mg dose of Medisorb Naltrexone.

The treatment effect was additionally explored via responder analyses. The applicant conducted a series of analyses exploring varying ‘categories’ of responders. Patients were classified into the following response categories: zero heavy drinking days per month, up to one heavy drinking day per month, up to two heavy drinking days per month, up to three heavy drinking days per month, and up to four heavy drinking days per month. The analyses provided some additional evidence of an effect. However, the analyses also raised questions regarding the clinical interpretation and meaningfulness of a reduction in the number of heavy drinking days among the population under study. These issues will be addressed in the medical review of Dr. Mwango Kashoki. To further explore the effects of the treatment, I conducted responder analyses on the subgroups of patients abstinent and non-abstinent at baseline. The response profile among the two subgroups suggested that a response to treatment was more likely to occur among patients abstinent at baseline.

An additional statistical concern was the appropriateness of pooling the placebo groups. The applicant contrasted the analysis based on the pooled placebo groups with the analysis considering separate placebo groups. Additionally, the applicant repeated the primary analysis exploring the treatment differences between the 4 mL placebo and 2 mL placebo groups. The results were consistent for pooled analyses and analyses with separate placebo groups.

During the course of the review, the Division of Scientific Investigations identified various protocol violations affecting data collection at two sites. In response, the Division of Anesthesia, Analgesia, and Rheumatology Products subsequently requested that the applicant reanalyze the data excluding the sites. Since a stratified dynamic randomization scheme was used to allocate patients to treatment, I was uncertain about the validity of the model-based inferences when excluding data from the two sites. Thus, I also requested that the applicant use re-randomization tests to verify the results. Alkermes performed the requested analyses and concluded that the supplemental analyses confirmed the efficacy of Medisorb Naltrexone 380 mg. Alkermes maintained that the protocol violations did not affect the study blind. They additionally stated, “It is unlikely that the protocol deviations pertaining to the separation of roles – between the BRENDA therapist and the time line follow back collector – introduced bias into the study.” For these reasons, the applicant strongly believed that the data from the excluded sites should be included in the final analyses of the study. Upon thorough consideration by the review team, the Division was inclined to agree with the applicant’s assessment of the effect of the identified violations. However, I did not agree with the applicant’s conclusions based on the analyses excluding the two sites.

2. INTRODUCTION

2.1 Overview

Alkermes, Incorporated proposes Medisorb Naltrexone, an injectable depot formulation of naltrexone, for the treatment of alcohol dependence. According to the applicant, "Medisorb Naltrexone is a microsphere-based formulation composed of naltrexone incorporated into a biodegradable matrix of polyactide-co-glycolide." Oral naltrexone is currently approved; however, the applicant asserts that the proposed formulation may reduce the potential for hepatotoxicity associated with oral naltrexone. The applicant also claims that the formulation provides continued exposure for at least one month.

Medisorb Naltrexone was introduced to the Division of Anesthesia, Analgesia, and Rheumatology Products (DAARP) via IND 61,138. During the development process, Alkermes submitted several study protocols for division comment. In addition, the product was discussed during a pre-IND meeting, a Type C industry meeting, a pre-NDA meeting, and a CMC meeting. Discussion topics included the clinical development plan, efficacy endpoints of interest, and the statistical analyses. In the pre-IND meeting, the Agency commented that a reduction in heavy drinking was a vague concept and recommended that a responder analysis with respect to absence of heavy drinking be conducted. The Agency further reiterated the recommendation to perform a responder analysis during the pre-NDA meeting. Additionally at the pre-IND meeting, the applicant was encouraged to explore analytical approaches appropriate for multiple failure times. The Agency also agreed that the study population consisting of currently abstinent alcoholics was suitable. During the development process, the study population evolved to include non-abstinent alcoholics. Moreover, the event rate of heavy drinking over a period of time emerged as the primary outcome variable. Methodology appropriate for recurrent event data was proposed and utilized for the primary analysis. The statistical reviewer of the IND, Dr. Milton Fan, expressed several concerns upon review of the draft statistical analysis plan. Dr. Fan's concerns included the need to validate the model-based inference under the dynamic randomization algorithm, the handling of missing data, the appropriateness of pooling the placebo groups, and the validity of the proportional hazards assumption in the primary analysis. Currently, the applicant has submitted NDA 21-897 in support of Medisorb Naltrexone for the treatment of alcohol dependence.

2.2 Data Sources

A single, randomized, placebo-controlled, multi-center, double-blind study was conducted to establish the efficacy of Medisorb Naltrexone. The data and final study reports for the completely electronic submission were archived in the Food and Drug Administration internal document room under the network path location \\Cdsesub1\evsprod\n021897\0000.

3. STATISTICAL EVALUATION

3.1 Evaluation of Efficacy

Study Design

Eligible patients were randomized in a 2:2:1:1 ratio to receive intramuscular injections of Medisorb Naltrexone 190 mg, Medisorb Naltrexone 380 mg, placebo for Medisorb Naltrexone 190 mg, or placebo for Medisorb Naltrexone 380 mg, respectively. Patients randomized to placebo received a matching volume of Medisorb microspheres (i.e. 2 mL or 4 mL) without naltrexone. Treatment was administered during clinic visits occurring at baseline and every 4 weeks thereafter for a 24-week period. During clinic visits, participants also received biopsychosocial support therapy using the BRENDA approach. Alcohol consumption was recorded throughout the study using the timeline follow-back method (TLFB). In the NDA submission the applicant stated, "The BRENDA therapists did not collect the TLFB data reported in the study."

Patients were allocated to treatment for balance on four baseline characteristics using a dynamic randomization procedure. The characteristics were goal of abstinence, gender, abstinence prior to randomization, and investigative site or center. The former three characteristics had two levels while site had 24 levels. The dynamic randomization process was enacted via an interactive voice response system (IVRS). The randomization algorithm (biased coin, $p=0.75$) is provided in the appendix.

The primary measure of efficacy was the event rate of heavy drinking over 24 weeks of treatment. This endpoint was defined as the number of heavy drinking days divided by the number of days at risk for heavy drinking. The applicant's use of the event rate was motivated by the desire to evaluate the drinking events over a defined duration. In addition, a heavy drinking day was defined as a day on which a man consumed at least five drinks or a woman consumed at least four drinks. An alcoholic drink was defined as 13.6 grams of absolute ethanol. Secondary measures of efficacy included days to relapse of heavy drinking, days to relapse of any drinking, number of alcoholic drinks per day, percent of heavy drinking days, percent of days abstinent from alcohol, and the event rate of drinking above the National Institute of Alcohol, Abuse, and Alcoholism derived "safe drinking" level (1 drink/day for women, 2 drinks/day for men).

A sample of size 600 was formulated using log-hazard ratio methods to detect a log event rate ratio of 0.50 to 0.55 with approximately 90% power. In the formulation of the sample size, the applicant assumed, "The proportion of subjects who will be 'abstinent' to heavy drinking will be 0.775 at 24 weeks in 1 of the 2 Medisorb Naltrexone treatment groups as compared with 0.600 in the placebo group and 0.600 in the other Medisorb Naltrexone group."

Patient Disposition, Demographic and Baseline Characteristics

Descriptive demographics and baseline characteristics were summarized for the intent-to-treat (ITT) population composed of all randomized patients who received at least one dose of treatment. The ages of patients ranged from 19 to 79 with a mean age of 45. In the study, 84% of patients were Caucasian, 8% were African-American, and 5% were Hispanic. Sixty-eight percent of the population was male, and the proportion of males to females was approximately 2 to 1 across all treatment groups. Baseline characteristics included weight, height, type of treatment center (i.e. addiction and/or research), patients' treatment goal, lead-in drinking (or abstinent at baseline), employment status, and smoking status. Ninety-two percent of participants consumed alcoholic beverages during the seven days prior to randomization. In addition, 43% of participants had a treatment goal of total abstinence. A detailed table outlining the composition of the study population with respect to demographic and baseline characteristics is presented in the appendix. Demographic and baseline characteristics were similar across the treatment groups.

Of the 627 randomized participants, 209 were randomized to placebo, 210 were randomized to 190 mg, and 208 were randomized to 380 mg. Four-hundred and one participants received all six doses of the treatment. Table 1 was provided in the NDA submission and shows the reasons for incomplete treatment and incomplete data collection. During the review process, the applicant submitted data that further classified discontinuations. The reclassified discontinuations are presented by treatment in Table 2.

Table 1: Reasons Patients Withdrew from Study Treatment and Withdrew from Data Collection
(Source: Reproduced from Final Study Report ALK21-003, Table 14.1.2)

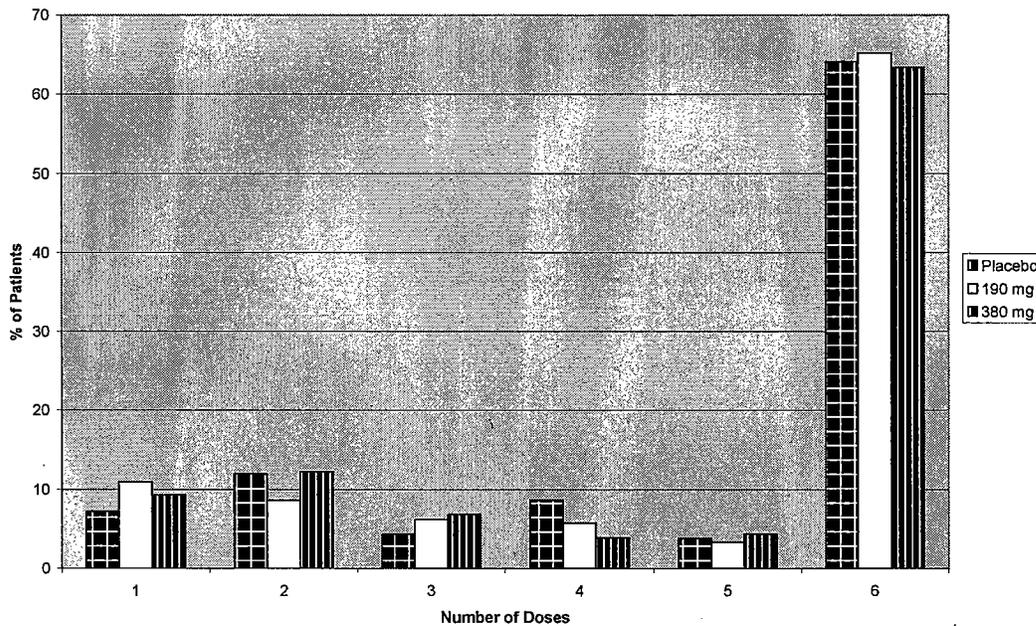
Reason for Incomplete Treatment	Reason for Incomplete Data Collection					Total
	Complete Data Collection	Investigator Judgment	Lost to Follow-up	Withdrew Consent	Other	
Received 6 doses of treatment	388		8	3	2	401
AE	30	2	4	4	17	57
Investigator Judgment		6				6
Lost to Follow-up			83			83
Protocol Violation			1		1	2
Withdrew Consent	7		1	53	1	62
Other	10				6	16
Total	435	8	97	60	27	627

Table 2: Reasons Patients Withdrew from Study Treatment (Reclassified)
 (Source: Adapted from Table 1.1.1 submitted on 29 July 2005)

	190 mg (n=210)	380 mg (n=208)	Placebo (n=209)	Total (n=627)
Completed	137	130	135	402
Adverse events	12	27	13	52
Investigator Judgment	2	3	2	7
Lack of efficacy	9	9	16	34
Lost to follow-up	31	24	28	83
Other	3	0	2	5
Protocol violation	2	0	0	2
Subject withdrew from consent	14	15	13	42

I additionally used a bar graph to depict the percentage of patients within each treatment group that received 1, 2, 3, 4, 5, or 6 doses respectively. A distinguishable pattern of discontinuations was not apparent.

Figure 1: Number of Doses (per treatment group)



Statistical Methodologies

The statistical methodologies utilized in the submission resulted from numerous correspondences between the applicant and the agency. As previously stated, concerns expressed by the statistical

reviewer of the IND included the handling of missing data, the appropriateness of pooling the placebo groups, and the validity of the proportional hazards assumption in the primary analysis. The applicant finalized the statistical analysis plan on 19 November 2003, prior to the unblinding of the study. As a result of additional feedback from the agency and statistical issues that arose after unblinding, subsequent analyses were performed by the applicant.

Since the event of interest (i.e. heavy drinking) could occur on multiple days, the statistical methodology used by the applicant accounted for recurrent events across time. Specifically, an Andersen-Gill model was used to assess the overall effect of treatment. The model was stratified by gender, treatment goal of abstinence, and abstinence at baseline (i.e. no drinking seven days prior to the initial treatment administration). Indicator variables representing the treatment effect of the low dose relative to placebo and the high dose relative to placebo were included in the model. The applicant additionally repeated the analysis including a term for baseline percent of heavy drinking. The detailed statistical formulation of the model used in the primary analysis is provided in the appendix. Multiple comparisons, arising from testing each dose of treatment versus placebo, were accounted for via the method of Hochberg. To verify the validity of the model-based inference, statistical significance was evaluated via re-randomization tests. Re-randomization or permutation tests are advantageous in that few, if any, assumptions are required for their application. The following excerpt describes the general implementation of a re-randomization test:

When you analyze an experiment or survey with a parametric test, you compare the observed value of the test statistic with the values in a table of its theoretical distribution. Analyzing the same experiment with a permutation test, you compare the observed value of the test statistic with the set of what-if values you obtain by rearranging and relabeling the data (*excerpt from Permutation Tests by Phillip Good*).

Through the use of the Andersen-Gill model, the applicant sought to provide evidence of a reduction in heavy drinking over time in patients receiving Medisorb Naltrexone. According to the applicant, "The method of analysis estimates the average event rate ratio over time taking into account patient discontinuation." In general, the Andersen-Gill model is formulated by dividing the follow-up time for each patient into intervals defined by actual heavy drinking days. Thus, a patient only contributes data (and belongs to the risk set) for the days having a recorded measurement of the number of drinks consumed. The model assumes that multiple observations per patient are independent, that is, the numbers of events in non-overlapping intervals are independent (also see appendix). Furthermore, another assumption of the model is that of proportional hazards (i.e. the hazard or risk of experiencing a heavy drinking day is constant).

To alleviate concern regarding the appropriateness of the assumption of independent observations, the applicant employed a robust variance estimator approach. Under the approach, the variance estimates were valid even if the dependence structure was modeled incorrectly. The applicant proposed a stratified analysis over covariates for gender, prior drinking, and goal of abstinence to address concerns regarding the proportional hazards assumption. According to the applicant, "A stratified analysis adjusts for different baseline 'hazards' of the prespecified stratification factors. In this way, the treatment effect would not be subject to the distortion that

a covariate-by-time interaction would induce by inclusion of such a covariate in the model.” Moreover, the applicant supplemented the statistical analysis plan (after unblinding) to include an alternate method of analysis if the assumption of proportional hazards was violated. I could not find a pre-specification of the alternate method in the final statistical analysis plan. However, the applicant stated that the approach was “prespecified” in a 21 June 2002 written response to agency comments. The correspondence stated, “...we prefer to provide an alternative to the Andersen-Gill model analysis if the proportional hazards assumption is not met, in the final analysis plan prior to unblinding. However, we will offer our most likely approach in brief here.” Following the proposal of a stratified analysis, the correspondence further stated, “Another approach is simply to collapse event rates over time for each patient such that the marginal event rate for each patient will be incorporated into an analysis of covariance or a non-parametric analysis of event rates (depending on the distribution of event rates over all subjects).”

Since the agency did not fully concur with the proposal to pool placebo groups, the applicant provided a justification. The applicant stated that the low dose injection required a lower volume of microspheres than the high dose injection. Thus, the microsphere volume for placebo injections was matched to active injections for the sole purpose of maintaining the blind. Furthermore the applicant stated, “The undisputed assumption of the study design in the original protocol is that drinking outcomes are independent of whether subjects receive a low volume or high volume placebo injection.” The applicant therefore concluded that pooling of the placebo groups was appropriate. The applicant also contrasted the analysis based on the combined placebo groups with the analysis considering separate placebo groups to alleviate the concern regarding the pooling. Additionally, the applicant repeated the primary analysis exploring the treatment difference between the 4 mL placebo and the 2 mL placebo groups.

Event rates obtained via the Andersen-Gill model were based on available data only. The applicant assumed that uncaptured or missing data occurred randomly and provided no additional insight into the effect of the treatment. To assess the assumption that missing data occurred randomly, the applicant examined the comparability of the treatment groups for subject discontinuation and outcomes via several techniques. The applicant examined the event rate of heavy drinking by the number of doses received, the Kaplan-Meier curves, and a pattern mixture model. In general, a pattern mixture model is a statistical tool designed to model the available or observed data and the missing data mechanism. Using the pattern mixture model approach employed by the applicant, the data was initially stratified by the number of doses (i.e. the missing data pattern). Estimates of the high and low dose treatment effects were then obtained within each stratum. The estimates were subsequently weighted (by 1/variance), and pooled estimates and variances were obtained to formulate conclusions. In the construction of a general pattern mixture model, strata are selected by combining groups with similar missing data patterns. Moreover, an assumption of the approach is that uncaptured data within each stratum is missing randomly.

The applicant conducted a responder analysis whereby patients were classified into the following categories: zero heavy drinking days per month, up to one heavy drinking day per month, up to two heavy drinking days per month, up to three heavy drinking days per month, and up to four

heavy drinking days per month. Heavy drinking days per month were computed via the formula, *Heavy Drinking Days per month = (Percent Heavy Drinking Days*30.4)/100*. Differences between the proportions for patients on active treatment versus placebo were compared via chi-square tests.

Results and Conclusions

The results of the applicant’s primary analyses are shown in Table 3. The applicant concluded that the 380 mg dose of Medisorb Naltrexone significantly reduced the event rate of heavy drinking as compared to placebo. Specifically, patients receiving 380 mg of Medisorb Naltrexone experienced a 25% reduction, as indicated by the hazard ratio of 0.75, in the event rate of heavy drinking compared to the placebo group. An equivalent conclusion was attained when the analysis was adjusted for the percent of heavy drinking at baseline.

Table 3: Event rate of Heavy Drinking^{*†}: Test for Treatment Effect in ALK21-003:
Andersen-Gill (Robust Variance) Stratified Analysis
(Source: Adapted from Final Study Report ALK21-003, Table 8)

	Estimate	Hazard ratio(95% CI)	Unadjusted p-value	Adjusted p-value
190 mg vs. placebo	-0.19	0.83 (0.68,1.02)	0.07	0.07
380 mg vs. placebo	-0.29	0.75 (0.60,0.94)	0.01	0.02

^{*}For each variable (190mg or 380 mg) in the analysis, parameter estimates are obtained for each stratum and pooled by weighting each stratum by 1/var (as described by Wei and Johnson, *Biometrika*, 1985). The hazard ratios are obtained by exponentiating the parameter estimates.

[†]Hochberg method was used to adjust p-value of 190 vs. placebo and 380 mg vs. placebo.

The planned stratified analysis across gender, treatment goal of abstinence, and lead-in drinking (or abstinent at baseline) resulted in eight possible strata. The stratum formed by females, no lead-in drinking, and a treatment goal of abstinence included only five patients, none of whom were in the placebo group. Thus, the applicant performed two sensitivity analyses to support the findings. In the first analysis, the two smallest strata were collapsed resulting in an analysis of seven strata. The second approach used an unstratified analysis. Both analyses supported the conclusions derived from Table 3. According to the applicant, the re-randomization analysis based on the stratified Andersen-Gill model produced unstable results because of the small sizes for some of the strata. A re-randomization based on the unstratified model was conducted and yielded findings consistent with the asymptotic tests (see appendix).

The applicant repeated the primary analyses using the matched placebo groups for the high and low doses, respectively. The analyses supported the initial findings that the 380 mg dose of Medisorb Naltrexone reduced the event rate of heavy drinking. The percent reduction ranged from 30% to 35% for the stratified and unstratified analyses, respectively. Moreover, there was no evidence of a significant treatment difference when comparing the 4 mL and 2 mL placebo groups.

To explore the validity of the applicant's findings from the primary analysis, I evaluated the amount and pattern of uncaptured data as well as the plausibility of the assumptions underlying the Andersen-Gill model. I initially examined the subject discontinuation table provided by the applicant and reproduced in Table 1. During the review of the NDA, the applicant additionally submitted data that further classified discontinuations. The reclassified discontinuations included a category denoting lack of efficacy (previously not included) and are provided in Table 2.

Only 64% of the patients in the ITT population completed the study. However, the percent of completers across treatment groups appeared to be comparable. There were twice as many adverse events among patients in the 380 mg arm as compared to patients in the placebo arm. Additionally, the number of patients lost to follow-up and withdrawing consent was larger than anticipated. The large number of patients in the latter category was concerning since some of the patients could have potentially withdrawn consent because of adverse events. As previously mentioned, the applicant evaluated the comparability of the discontinuations among treatments via Kaplan Meier curves, an examination of the event rate of heavy drinking by number of doses received, and a pattern mixture model. Based on the results shown in the appendix, the applicant states, "Subject discontinuation was unlikely to have had a significant impact on the observed treatment effect." While the Kaplan-Meier curves and the event rate per dose seemed to provide some substantiation of this statement, my concern regarding the effect of treatment related drop-outs remained. Moreover, I had a concern with the appropriateness of the pattern mixture model. The model assumed that uncaptured data within each stratum was missing randomly. Since patients potentially dropped out for treatment related reasons, the appropriateness of the model assumption was questionable given the data.

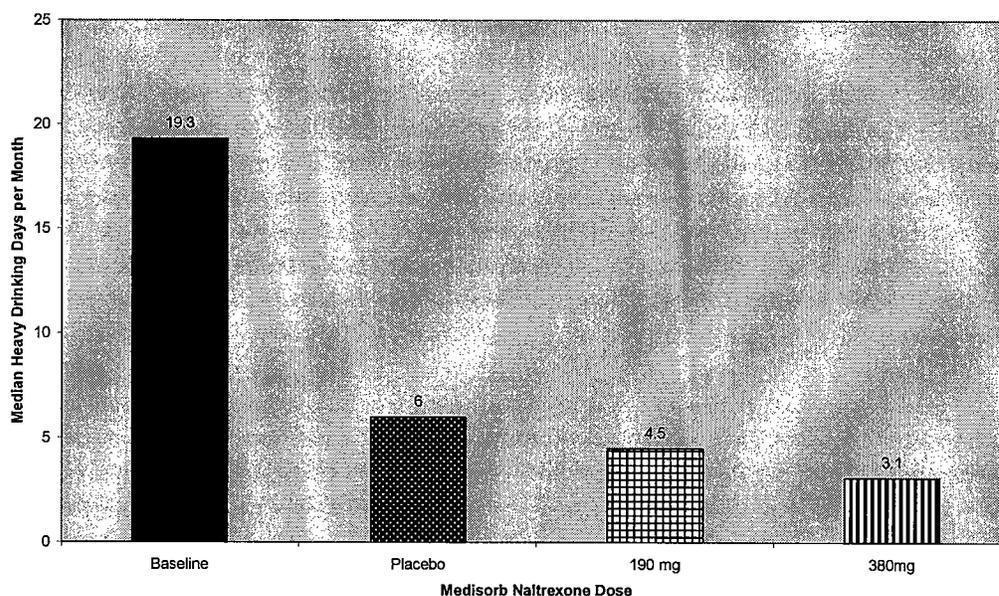
At the request of the agency, the applicant investigated the validity of the assumption of proportional hazards. In general, the assumption of proportional hazards may be relaxed without losing validity of the estimated treatment effect. However, a serious violation of the assumption warrants some concern. The applicant identified a severe violation of the proportional hazards model. As a result of this finding along with the instability of the re-randomization test and my aforementioned concerns regarding missing data, I focused significant attention on the nonparametric analysis of the percent of heavy drinking days. The applicant conducted both an unstratified and stratified analysis with combined and separate placebo groups. My evaluation centered on the unstratified analysis since it was pre-specified. The analysis was conducted on 611 patients and employed a Wilcoxon test, a nonparametric analogue to a t-test. Thirteen patients in the ITT population did not have post-baseline data and were excluded from the analysis. The results are shown in Table 4 and Figure 2. The median numbers of heavy drinking days per month illustrated in Figure 2 were obtained by multiplying the "median event rate of heavy drinking" by 30.4. The applicant used the terminology "median event rate of heavy drinking" to refer to the median percentage of heavy drinking days. Of note, the stratified analysis as well as the analyses with separate placebo groups produced comparable results.

Table 4: Comparison of Median Event rate of Heavy Drinking: Non-Parametric Analyses
(Source: Adapted from final Study Report ALK21-003, Table 18)

Treatment Group	N	Median Event Rate of Heavy Drinking	Percent Difference	p-value Wilcoxon test unstratified
Placebo	204	0.20		
190 mg	206	0.15	25%	0.22
380 mg	201	0.10	48%	<0.01

* p-value compared to placebo

Figure 2 - Median Heavy Drinking Days Per Month
(Source: Reproduced from Final Study Report ALK21-003, Figure 9)



The applicant concluded that there was a 48% reduction in the median percent of heavy drinking for patients receiving 380 mg compared to patients receiving placebo. The applicant further asserted that the large difference between the relative reductions in the event rate of heavy drinking in the nonparametric analysis (i.e. 48%) versus the analysis employing the Andersen-Gill model (i.e. 25%) was because of a minority of subjects who drank very heavily. Approximately 10% of the patients in the 190 mg and 380 mg treatment arms had 77% and 72% of heavy drinking days during the study, respectively. The applicant asserted that this 10% of patients contributed a higher event rate in the Andersen-Gill analysis and therefore diminished the treatment effect. The argument was plausible in my opinion; however, I had concern that the applicant overestimated the percent difference by ignoring the uncaptured data in the nonparametric analysis. The applicant calculated the median percent of heavy drinking days from data at the end of the study and used a last observation carried forward strategy for missing

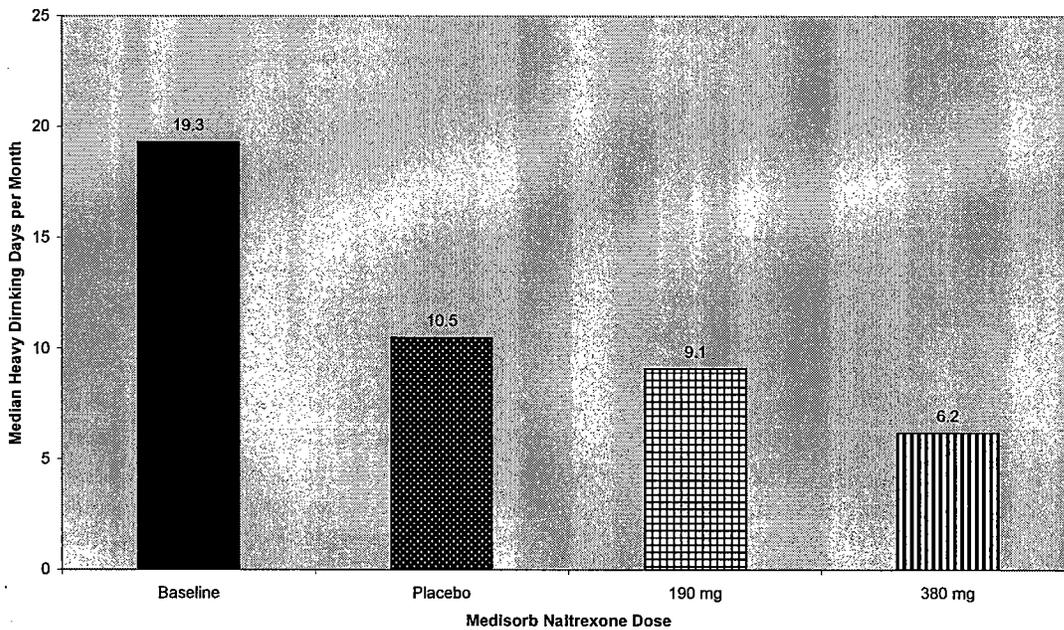
data. I repeated the analysis treating all missing data as heavy drinking days. The results of my reanalysis are found in Table 5 and depicted in Figure 2.

Table 5: Comparison of Median Event rate of Heavy Drinking: Non-Parametric Analyses
Any missing data day is defined as a heavy drinking day

Treatment Group	N	Median Event Rate of Heavy Drinking	Percent Difference	p-value* Wilcoxon test unstratified
Placebo	204	0.35		
190 mg	206	0.30	13%	0.69
380 mg	201	0.20	41%	0.05

*p-value compared to placebo

Figure 3 - Median Heavy Drinking Days Per Month
(All days with missing data are considered as heavy drinking days)



My reanalysis provided evidence of an effect of treatment. Moreover, I performed two additional analyses using varying imputation schemes. In the first analysis, heavy drinking days were imputed for missing data for patients discontinuing due to an adverse event. A last observation carried forward imputation strategy was used for all other discontinuations. The analysis was also repeated with heavy drinking days imputed for missing data for patients discontinuing due to an adverse event or withdrawing consent. These analyses provided additional support for effect of the treatment and alleviated concern regarding the missing data.

Medical colleagues on the review team indicated that the response profile of heavy drinking among participants was of interest. Table 6 illustrates the results from the applicant’s responder analysis. The table shows the number of heavy drinking days per month calculated from the percentage of heavy drinking days. For example, a heavy drinking percentage of 40% would translate to 12 heavy drinking days per month. Similar to the nonparametric analysis, the responder analysis was conducted on the 611 patients having post-baseline data.

Table 6: Responder Analysis
(Source: Reproduced from final Study Report ALK21-003, Table 25)

Post-Baseline ¹ Heavy Drinking Days Per Month ²	Placebo (n=204)	190 mg (n=206)	380 mg (n=201)	190 mg versus placebo*	380 mg versus placebo*
0	23 (11%)	29 (14%)	34 (17%)	0.39	0.10
0-1	44 (22%)	53 (26%)	68 (34%)	0.32	0.01
0-2	56 (27%)	68 (33%)	81 (40%)	0.22	0.01
0-3	68 (33%)	83 (40%)	97 (48%)	0.14	<0.01
0-4	84 (41%)	95 (46%)	110 (55%)	0.31	0.01

¹ Drinking data up to 30 days after the last dose.

² Heavy Drinking Days Per Month = (Percent Heavy Drinking Days * 30.4)/100.

*P-values calculated via Chi-square test.

The applicant concluded that there existed a significant proportion of responders in the Medisorb Naltrexone 380 mg group compared to the placebo group for each response category, with the exception of the zero response category. The medical officer, Dr. Mwango Kashoki, expressed concern with the methodology used to establish the number of heavy drinking days per month. In response, I reanalyzed the data using an actual count of the number of heavy drinking days. The overall conclusions remained unchanged.

In Table 6, participants were classified into categories based on their individual percentages of heavy drinking days reported during participation in the study. In this manner, a patient dropping out prior to the last dose could potentially reflect a positive effect of treatment when in actuality the patient dropped out due to an adverse event. Thirty-six percent of the patients did not receive all six doses; therefore, the applicant’s analysis had the potential to overestimate the treatment effect. I reanalyzed the applicant’s data imputing heavy drinking days for all missing data. The results of my analysis are shown in Table 7.

Table 7: Responder Analysis
(Heavy drinking days are imputed for all missing data.)

Post-Baseline ¹ Heavy Drinking Days Per Month ²	Placebo (n=204)	190 mg (n=206)	380 mg (n=201)	190 mg versus placebo*	380 mg versus placebo*
0	11(5%)	15(7%)	14(7%)	0.43	0.51
0-1	35 (17%)	42 (20%)	46 (23%)	0.40	0.15
0-2	45 (22%)	50 (24%)	62 (31%)	0.60	0.05
0-3	55 (27%)	60 (29%)	74 (37%)	0.63	0.03
0-4	64 (31%)	73 (35%)	83 (41%)	0.38	0.04

¹ Analysis uses 168 days of data for each patient.

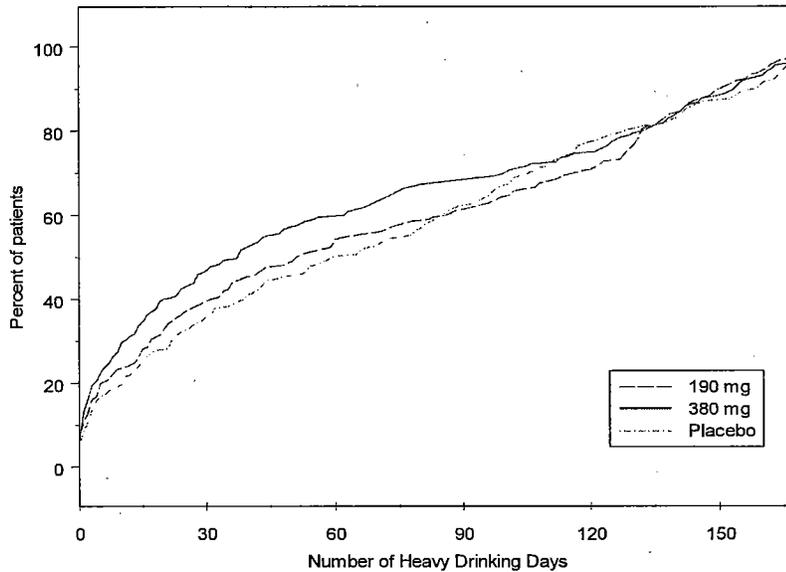
² Heavy Drinking Days Per Month = (Percent Heavy Drinking Days * 30.4)/100.

*P-values calculated via Chi-square test.

From the table, I concluded that Medisorb Naltrexone 380 mg provided some reduction in the percent of heavy drinking days per month. Caution must be exercised when interpreting the results of a supportive responder analysis. In general, a responder analysis is advantageous in that it allows an ease of interpretation and aids in discerning the potential benefit of the treatment; however, the analysis is less powerful to detect a difference among treatments. The results from the analysis indicated that approximately 10% more of the patients receiving the high dose of the treatment had at most two, three, or four heavy drinking days per month (as compared to placebo). This translated into a need for 10 patients to be treated in order to prevent one patient from having up to four heavy drinking days per month. The same interpretation was true of the number needed to treat to prevent up to two or three heavy drinking days per month. In contrast, fifty patients would need to be treated in order for one patient to not experience any heavy drinking days over a 24-week period. The medical review team indicated specific interest in responders who experienced a complete absence from heavy drinking; therefore, I additionally examined the reasons why patients were included in the applicant's responder analysis but excluded from my reanalysis. As suspected, most patients that were excluded from being responders (with zero heavy drinking days) in my analysis dropped out of the study. At the time of withdrawal, the patients had not experienced a heavy drinking day.

To further elucidate the findings, I examined the cumulative distribution curves as well as the response profiles within specific subgroups identified by the clinical review team. The cumulative distribution curves are displayed in Figure 4. The curves clearly separate indicating that the differences detected in the analyses were not purely due to chance. The response profiles within specific subgroups are provided and discussed in Section 4.2 of this review.

Figure 4: Cumulative Distribution Plot



An additional concern of the medical team was the definition of heavy drinking. Based on evolving thoughts on the definition of heavy drinking, the medical team expressed interest in an alternate definition of heavy drinking. Thus, I reanalyzed the data using a definition whereby a heavy drinking day for a male was any day whereby four or more drinks were consumed. Similarly, a heavy drinking day for a female was any day whereby three or more drinks were consumed. The results of my analyses are displayed in the appendix. Overall, the treatment effect did not appear to be evident when the alternate definition of heavy drinking was applied.

The applicant also explored numerous secondary variables. Based on my consultations with the medical officer, I did not conduct additional explorations on these endpoints.

On 23 August 2005, the review team became aware of a series of problems identified during FDA inspections of sites 215 and 217. The problems included protocol violations affecting the collection of data. At site 215, a person who administered the drug (a physician not licensed in the United States) also collected time line follow-back data. In addition, the BRENDA therapists at sites 215 and 217 collected time line follow-back data. The Division of Scientific Investigations recommended that DAARP consider excluding data from sites 215 and 217 in the evaluation. Therefore in an information request to Alkermes dated 25 August 2005, I requested that the applicant reanalyze the data excluding the sites. The applicant was specifically asked to provide results from the primary analysis, the nonparametric analysis, and the responder analyses excluding sites 215 and 217 (with pooled and matched placebo groups). The division also requested that the latter analyses be conducted imputing heavy drinking days for all missing data. To further elucidate findings, the applicant was asked to repeat the requested responder analyses

for several subgroups of patients. Since a stratified dynamic randomization scheme was used to allocate patients to treatment, I was uncertain about the validity of the model-based inferences when excluding data from the two sites. Thus, I requested that the applicant use re-randomization tests to verify the results.

Alkermes provided a response dated 7 September 2005 that included numerous supplemental analyses. In addition to the requested analyses, Alkermes submitted an unstratified Andersen-Gill analysis as well as the Andersen-Gill model adjusted for baseline percent of heavy drinking. I focused attention on the analysis employing the stratified Andersen-Gill model as well as the nonparametric and responder analyses imputing heavy drinking days for all missing data. These analyses were of focus since they best represented the pre-specified analyses originally outlined in the study protocol. The analyses were conducted on 542 patients (excluding the 26 patients from site 215 and 46 patients from site 217). Table 8 and Table 9 depict the results from the applicant's analyses.

Table 8: Event rate of Heavy Drinking, Excluding Sites 215 and 217:
Andersen-Gill (Robust Variance) Stratified Analysis
(Source: Reproduced from Response to Information Request of 8/25/05, Table 1A)

	Hazard ratio(95% CI)	Unadjusted p-value
190 mg vs. placebo	0.94 (0.75,1.18)	0.59
380 mg vs. placebo	0.88 (0.69,1.13)	0.31

Table 9: Nonparametric Analysis, Excluding Sites 215 and 217:
Imputing Missing Data as a Heavy Drinking Day
(Source: Reproduced from Response to Information Request of 8/25/05, Table 4A)

Treatment Group	N	Median Event Rate of Heavy Drinking	Percent Difference	p-value Wilcoxon test unstratified
Placebo	182	0.34		
190 mg	183	0.34	1%	0.72
380 mg	177	0.21	38%	0.12

The applicant's responder analysis excluding data from sites 215 and 217 was conducted on 542 study participants and used data beyond day 168. I also conducted a responder analysis excluding the sites; however, my reanalysis also excluded patients with no post-baseline data and truncated data at day 168 for all patients. For consistency with results displayed previously in my review and for ease of comparison, the results from my responder analysis are displayed in Table 10.

Table 10: My Responder Analysis Excluding Sites 215 and 217:
Imputing Missing Data as a Heavy Drinking Day

	Placebo (n=178)	190 mg (n=180)	380 mg (n=174)	190mg versus placebo	380 mg versus placebo
0	10 (6%)	13 (7%)	14 (8%)	0.54	0.37
0-1	33 (19%)	38 (21%)	41 (24%)	0.54	0.25
0-2	42 (24%)	46 (26%)	56 (32%)	0.67	0.07
0-3	51 (29%)	55 (31%)	66 (38%)	0.69	0.07
0-4	60 (34%)	64 (36%)	75 (43%)	0.71	0.07

The following excerpt summarized the applicant’s conclusions resulting from the re-analyses excluding sites 215 and 217.

The supplemental data analyses requested by the Division confirm the robust findings of efficacy shown in ALK21-003 and contained in the NDA submission. Even with the exclusion of sites 215 and 217, the primary and secondary efficacy analyses continue to show a positive treatment effect for Medisorb Naltrexone 380 mg versus placebo. As expected, exclusion of subjects from sites 215 and 217 (13 % or 82 of 624 total allocated subjects), increases variability, reduces statistical power, and results in nominal p-values that are higher than values for the primary efficacy analysis in the original NDA. In addition, the effect size is slightly narrower compared with the effects shown in the original NDA. Nonetheless, the analyses of the resulting patient subset yield positive treatment effects that are concordant with the primary hypothesis-testing result. Moreover, we found no meaningful differences from the original NDA when carrying out non-parametric and responder analyses with the data from sites 215 and 217 excluded.

I initially examined the applicant’s evaluation of the appropriateness of the proportional hazards assumption. Similar to the original analysis, the assumption was violated. Moreover as in the original analysis, the applicant commented that the re-randomization test could not be executed for the primary analysis because of the small sizes of strata. The inability of the applicant to conduct a re-randomization test added to my initial concerns regarding the validity of the model-based inference when excluding the sites. Thus, I again gave significant attention to the non-parametric analysis. I agreed with the applicant’s assertion that the power to detect a difference was reduced as a result of excluding patients. However, I anticipated that some evidence of the effect of the treatment would be maintained despite the exclusion of patients. In contrast, the treatment effect of the 380 mg dose disappeared when the sites were excluded from the nonparametric analysis (and the primary analysis). No additional information was obtained from the exploration of the response profile excluding sites 215 and 217.

3.2 Evaluation of Safety

The evaluation of the safety data was conducted by Dr. Mwango Kashoki. The reader is referred to Dr. Kashoki’s review for information regarding the adverse event profile.

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Race and Age

For exploratory purposes, the applicant performed a subgroup analysis with respect to gender utilizing an unstratified Andersen-Gill model. I additionally explored possible treatment effects for subgroups defined by age and race. Age was categorized using two subgroups, ages greater than or equal to 50 and ages less than 50. Race was categorized using two subgroups, Caucasian and non-Caucasian.

Medisorb naltrexone 380 mg reduced the event rate of heavy drinking among males. Specifically, males receiving Medisorb Naltrexone 380 mg experienced a 54% reduction in the event rate compared to males in the placebo group. No difference in the event rate of heavy drinking was detected among females. Similarly, there was evidence that the drug reduced the event rate of heavy drinking by 27% among Caucasians. However, no difference in the event rate was detected among non-Caucasians or among the varying age categories.

4.2 Other Special/Subgroup Populations

The medical review team expressed interest in exploring the effects of the treatment in the subgroups of patients that were abstinent seven days prior to randomization and patients non-abstinent at baseline. For clarification, abstinent patients did not consume any alcoholic beverages. Of the randomized patients, only 53 were abstinent at baseline. Since the number of patients abstinent at baseline was very small, I explored the possible effect of treatment on the subgroup via a responder analysis without any attempt to draw a formal statistical inference. The formulation of a responder mimicked that found in Section 3 of this review. The results of my analyses are displayed in Table 13, Table 14, and Figure 5. An examination of the response profile among the two subgroups of patients suggested that a response to treatment was more likely to occur among patients abstinent at baseline. This phenomenon was also apparent in the supplemental analysis excluding sites 215 and 217.

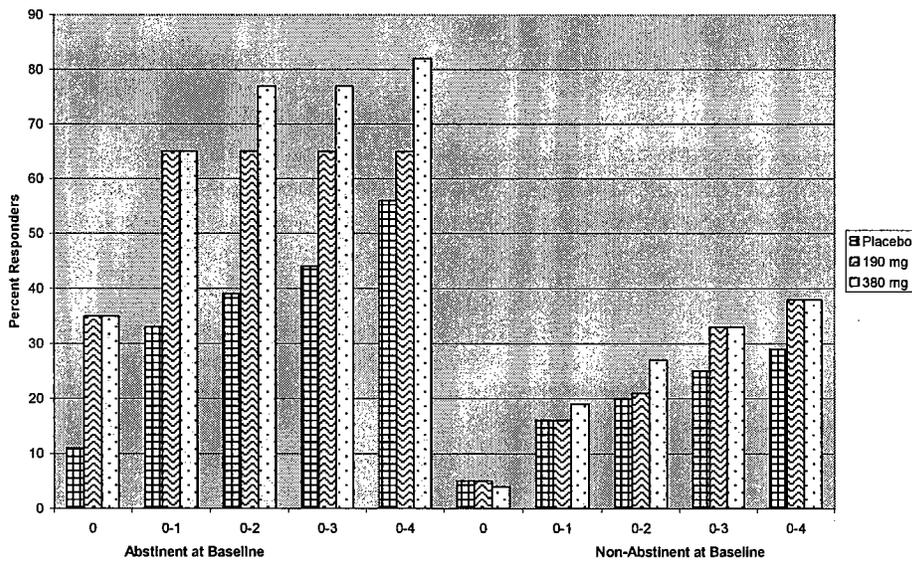
Table 13: Responder Analysis on Patients Abstinent at Baseline
(Heavy drinking days are imputed for all missing data.)

	Placebo (n=18)	190 mg (n=17)	380 mg (n=17)
0	2 (11%)	6 (35%)	6 (35%)
0-1	6 (33%)	11 (65%)	11 (65%)
0-2	7 (39%)	11 (65%)	13 (77%)
0-3	8 (44%)	11 (65%)	13 (77%)
0-4	10 (56%)	11 (65%)	14 (82%)

Table 14: Responder Analysis on Patients Non-abstinent at Baseline
(Heavy drinking days are imputed for all missing data.)

	Placebo (n=186)	190 mg (n=189)	380 mg (n=184)
0	9 (5%)	9 (5%)	8 (4%)
0-1	29 (16%)	31 (16%)	35 (19%)
0-2	38(20%)	39 (21%)	49(27%)
0-3	47(25%)	63 (33%)	61 (33%)
0-4	54 (29%)	72 (38%)	69 (38%)

Figure 5: Reponder Analysis for abstinent and non-abstinent patients



5. SUMMARY AND CONCLUSIONS

5.1 Statistical Issues and Collective Evidence

A primary statistical concern throughout the course of the review was the appropriateness of the Andersen-Gill model for the data. In the statistical literature, the Andersen-Gill model has frequently been the model of choice to analyze recurrent event data where the desire is to assess the overall treatment effect. However in the current setting, the appropriateness of the model was questionable because of the inability of the applicant to conduct a re-randomization test, the severe violation of the proportional hazards assumption, and the significant amount of missing data.

For these reasons, my attention focused on the nonparametric analysis. This analysis was not subject to the model assumptions. Moreover, the results of several of my sensitivity analyses alleviated my concerns regarding the missing data. I additionally evaluated the response profile of heavy drinking. The responder analysis allowed for ease of interpretation and for a more thorough investigation into the clinical meaningfulness of the claimed reduction. However, some caution was advised regarding the interpretation of results. Since the analysis was less powerful to detect a difference among treatments, negative results were not necessarily definitive.

The Division of Scientific Investigations identified various protocol violations affecting data collection at two sites during the review process. The Division of Anesthesia, Analgesia, and Rheumatology Products subsequently requested that the applicant reanalyze the data excluding the sites. In response, Alkermes provided their assessment of the violations and numerous supplemental analyses. Alkermes maintained that the protocol violations did not affect the study blind. They additionally stated, "It is unlikely that the protocol deviations pertaining to the separation of roles – between the BRENDA therapist and the time line follow back collector – introduced bias into the study." Moreover, the applicant concluded that the supplemental analyses confirmed the efficacy of Medisorb Naltrexone 380 mg. I did not concur with this conclusion. However, the Division agreed with the applicant's assessment of the effect of the identified violations, and final conclusions were based on the original submission including all sites.

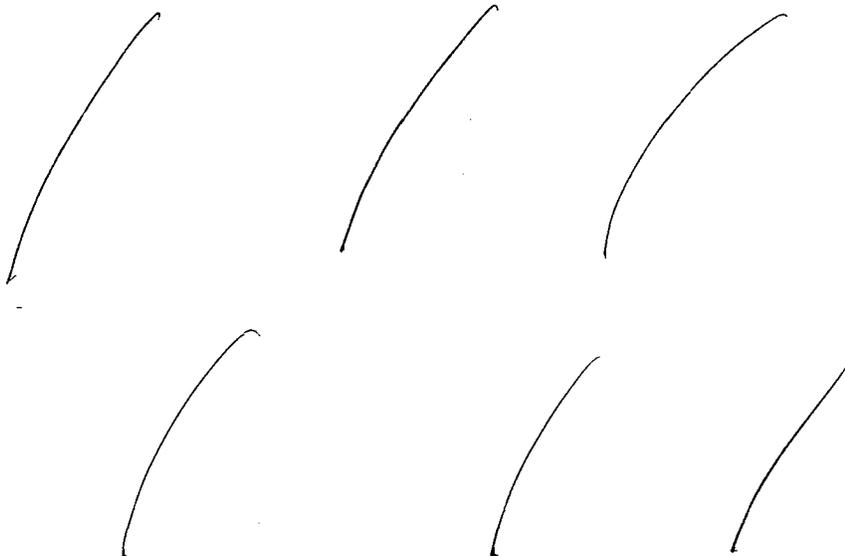
My collective evaluation of the analyses and results found that Medisorb Naltrexone 380 mg reduced the number of heavy drinking days over the 24-week period. Additional post-hoc analyses suggested that a response to treatment was more likely to occur among patients abstinent at baseline.

5.2 Conclusions and Recommendations

The applicant submitted NDA 21-897 to provide evidence of the efficacy and safety of Medisorb Naltrexone for the treatment of alcohol dependence. The applicant claims that the Medisorb Naltrexone 380 mg reduces heavy drinking based on an evaluation of the event rate of heavy drinking over 24 weeks.

I have reviewed the data and conclude that Medisorb Naltrexone has shown evidence of activity. Specifically, there is support for the claim that heavy drinking is reduced. However, the demonstrated reduction must be weighed with the totality of the findings in order to establish the merits of the treatment. Specifically, several factors should be considered when evaluating Medisorb Naltrexone. Alkermes' failure to identify protocol violations prior to the submission of the NDA diminishes my confidence in the overall conduct of the study and subsequent data. Support for Medisorb Naltrexone is derived from a single study; therefore, there is no replication of the findings to enhance my overall confidence in the study, data, and resulting conclusions. In addition, multiple safety concerns have been identified by the review team including elevated transaminases and numerous severe allergic reactions possibly associated with treatment. In conclusion, there is statistical support for the drug. However, the previously mentioned factors must be assessed collectively by the review team in order to evaluate the risks and benefits of Medisorb Naltrexone.

5.2.1 Labeling



2 Page(s) Withheld

 Trade Secret / Confidential

 / Draft Labeling

 Deliberative Process

APPENDICES

**APPEARS THIS WAY
ON ORIGINAL**

**APPEARS THIS WAY
ON ORIGINAL**

3 Page(s) Withheld

 Trade Secret / Confidential

 Draft Labeling

 Deliberative Process

Summary of Demographic and Baseline Characteristics

The following table was provided in the applicant's Clinical Study Report (Table 6).

	Treatment Group			
	All Subjects	Placebo	180mg	380mg
No. in the ITT Population	624	209	210	205
Sex (N, %) ¹				
Male	423 (68)	143 (68)	142 (68)	138 (67)
Female	201 (32)	66 (32)	68 (32)	67 (33)
Age (years)				
Mean	45	45	45	45
Std. Deviation	11	11	11	10
Median	45	44	44	45
Min-Max	19-79	21-79	19-72	21-72
Race/Ethnicity (N, %) ¹				
Caucasian	521 (84)	180 (86)	169 (81)	172 (84)
African American	50 (8)	17 (8)	17 (8)	16 (8)
Hispanic	32 (5)	7 (3)	15 (7)	10 (5)
Other	15 (2)	3 (1)	7 (3)	5 (2)
Asian	3 (.5)	1 (.5)	1 (.5)	1 (.5)
Native American	3 (.5)	1 (.5)	1 (.5)	1 (.5)
Male's Weight (kg)				
N	423	143	142	138
Mean	89	86	89	91
Std. Deviation	18	16	19	19
Median	85	82	85	89
Min-Max	50-159	5-137	51-159	50-156
Female's Weight (kg)				
N	200	66	68	66
Mean	71	72	71	71
Std. Deviation	16	16	15	17
Median	67	68	67	66
Min-Max	46-139	46-113	50-120	46-139
Male's Height (cm)				
N	422	143	141	138
Mean	178	178	178	179
Std. Deviation	7	7	8	7
Median	178	178	178	180
Min-Max	155-205	157-195	155-205	165-198

(table continues on next page)

	Treatment Group			
	All Subjects	Placebo	180mg	380mg
Female's Height				
(cm)				
N	200	66	68	66
Mean	165	165	166	164
Std. Deviation	7	6	7	7
Median	165	165	166	165
Min-Max	150–185	155–185	150–180	150–183
Site (N,%)¹				
217	46 (7)	15 (7)	15 (7)	16 (8)
225	40 (6)	14 (7)	12 (6)	14 (7)
209	39 (6)	14 (7)	12 (6)	13 (6)
210	38 (6)	14 (7)	12 (6)	12 (6)
215	36 (6)	12 (6)	12 (6)	12 (6)
214	35 (6)	11 (5)	12 (6)	12 (6)
213	33 (5)	11 (5)	11 (6)	11 (5)
208	34 (5)	10 (5)	13 (6)	9 (4)
224	31 (5)	10 (5)	10 (5)	11 (5)
218	31 (5)	11 (5)	10 (5)	10 (5)
216	30 (5)	11 (5)	10 (5)	9 (4)
212	30 (5)	10 (5)	10 (5)	10 (5)
202	27 (4)	10 (5)	9 (4)	8 (4)
230	27 (4)	8 (4)	10 (5)	9 (4)
221	26 (4)	8 (4)	9 (4)	9 (4)
211	25 (4)	8 (4)	9 (4)	8 (4)
227	20 (3)	7 (3)	7 (3)	6 (3)
229	17 (3)	5 (2)	6 (3)	6 (3)
228	17 (3)	5 (2)	6 (3)	6 (3)
220	13 (2)	4 (2)	4 (2)	5 (2)
207	12 (2)	4 (2)	4 (2)	4 (2)
226	8 (1)	3 (1)	3 (1)	2 (1)
219	6 (1)	2 (1)	2 (1)	2 (1)
223	5 (1)	2 (1)	2 (1)	1 (.5)
Treatment Centers				
(N,%) ¹				
Addiction	303 (49)	104 (50)	102 (49)	97 (47)
Both	109 (18)	36 (17)	36 (17)	37 (18)
Addiction/Research				
Research	212 (34)	69 (33)	72 (34)	71 (35)
Subjects' Treatment Goal¹				
Total Abstinence	270 (43)	90 (43)	90 (43)	90 (44)
Total Abstinence, but a lapse is possible	64 (10)	19 (9)	24 (11)	21 (10)
Occasional use	191 (31)	68 (33)	61 (29)	62 (30)
Temporary Abstinence	9 (1)	4 (2)	3 (1)	2 (1)
Regular Use but quantity controlled	75 (12)	23 (11)	29 (14)	23 (11)
No goal	15 (2)	5 (2)	3 (1)	7 (3)

(table continues on next page)

	Treatment Group			
	All Subjects	Placebo	180mg	380mg
No of subjects with lead in Drinking (N,%) ¹	571 (92)	190 (91)	193 (92)	188 (92)
% Heavy Drinking Days Pre First Dose				
N	624	209	210	205
Mean	65	65	67	64
Std. Dev.	26	25	26	26
Median	63	67	63	63
Min-Max	0-100	0-100	0-100	0-100
No. of Heavy Drinking Days 30 Days Pre First Dose				
N	624	209	210	205
Mean	20	20	20	19
Std. Dev.	8	8	8	8
Median	19	20	19	19
Min-Max	0-30	0-30	0-30	0-30
% Drinking Days 30 Days Pre First Dose				
N	624	209	210	205
Mean	76	76	77	76
Std. Dev.	23	23	23	23
Median	83	81	83	83
Min-Max	0-100	0-100	0-100	0-100
No. of Drinking Days 30 Days Pre First Dose				
N	624	209	210	205
Mean	23	23	23	23
Std. Dev.	7	7	7	7
Median	25	24	25	25
Min-Max	0-30	0-30	0-30	0-30
Alcohol Dependence Scale Score*				
N	306	100	103	103
Mean	17	17	18	17
Std. Dev.	7	7	7	8
Median	17	16	17	16
Min-Max	1-42	2-42	4-40	1-39
Unemployed at Baseline				
No	533 (85)	177 (85)	178 (85)	178 (87)
Yes	89 (14)	31 (15)	31 (15)	27 (13)

(table continues on next page)

	Treatment Group			
	All Subjects	Placebo	180mg	380mg
Attended Any Self Help Groups at Baseline? ¹				
No	553 (89)	185 (89)	187 (89)	181 (88)
Yes	69 (11)	23 (11)	22 (11)	24 (12)
Smoking Status at Baseline ¹				
No	328 (53)	120 (57)	103 (49)	105 (51)
Yes	293 (47)	88 (42)	106 (51)	99 (48)
Unknown	3 (.5)	1 (.5)	1 (.5)	1 (.5)

¹ Percentages are out of the number of subjects in the ITT population

*The ADS was added to the protocol in April 2002, after enrollment had begun. Subjects enrolled prior to that date did not complete the questionnaire.

**APPEARS THIS WAY
ON ORIGINAL**

2 Page(s) Withheld

Trade Secret / Confidential

Draft Labeling

Deliberative Process

Re-randomization Analysis Results

The following table was provided as a response to an Information Request dated 05 August 2005.

Re-randomization Analysis¹ Test Results for Treatment Effects with Andersen-Gill Model (Unstratified Analysis for Event Rate of Heavy Drinking)

Effects in the Model	P-value (2-sided)			
	Medisorb Naltrexone 380 mg vs. Placebo		Medisorb Naltrexone 190 mg vs. Placebo	
	Asymptotic Test	Re-randomization Test	Asymptotic Test	Re-randomization Test
380 mg vs. Placebo	0.02	0.03	0.32	0.35
190 mg vs. Placebo				

¹ Based on simulation with 10,000 realizations.

The following excerpt (dated 07 September 2005) from the applicant provides additional detail regarding the inability to complete the re-randomization test based on the stratified Andersen-Gill model:

For the unstratified analysis, asymptotic p-values correlated well with p-values from the re-randomization analyses. A re-randomization analysis was attempted for the stratified Andersen-Gill analyses. However, the analysis could not be completed because of small sample sizes in certain strata. For example, realizations occur such that all of the subjects in one of the treatment groups within a stratum have no drinking events. When these realizations occur, the estimate of the log hazard ratio for the strata in question is undefined ($\pm\infty$). This appears to have caused the re-randomization program to discontinue or "crash". Other factors may have also affected the attempted re-randomization analysis. A re-randomization analysis for the stratified analysis can not be performed without introducing certain assumptions into the program code.

Exploration of Patient Discontinuation

The applicant explored patient discontinuation via Kaplan-Meier curves (Figure 14.1.1 of the Clinical Study Report), an examination of the event rate by dose received (Table 5 in the Statistical Supplement), and a pattern mixture model (Table 7 in the Statistical Supplement). Each is replicated below.

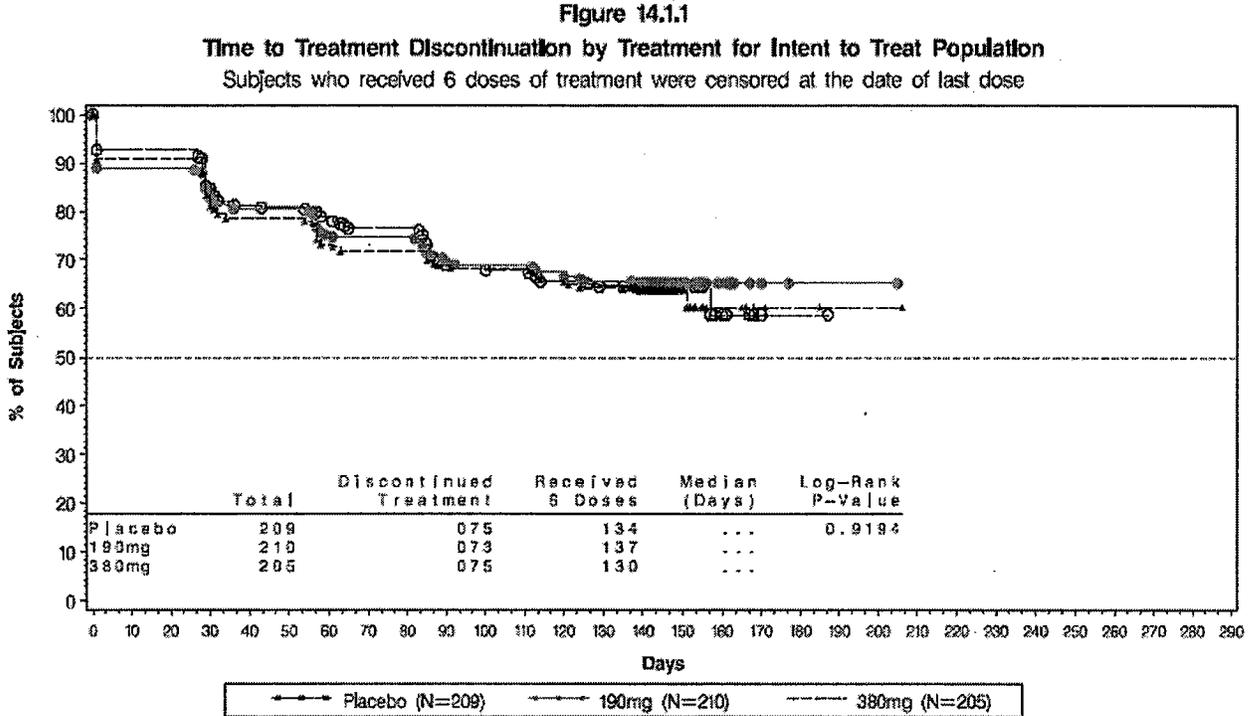


Table5: Event rate of Heavy Drinking: Test for Treatment Effect by Dose Received – 8 Strata Analysis

Time Period	Treatment	N	Hazard Ratio	p-value
30 Days After				
1 st Dose	Placebo	209		
	190 mg	210	0.52	<0.01
	380 mg	205	0.41	<0.01
2 nd Dose	Placebo	194		
	190 mg	187	0.88	0.16
	380 mg	186	0.75	<0.01
3 rd Dose	Placebo	169		
	190 mg	169	0.86	0.11
	380 mg	161	0.73	<0.01
4 th Dose	Placebo	160		
	190 mg	156	0.86	0.12
	380 mg	147	0.71	<0.01
5 th Dose	Placebo	142		
	190 mg	144	0.84	0.08
	380 mg	139	0.73	<0.01
6 th Dose	Placebo	134		
	190 mg	137	0.83	0.07
	380 mg	130	0.75	0.01

All subject drinking days are analyzed including subjects that discontinued. N is the number of subjects who received at least the number of doses shown in that row. P-values are compared with placebo.

Table 7: Assessment of the Impact of Subject Discontinuation From Treatment: Pattern Mixture Analysis ¹

Doses Received	Treatment	N	Hazard Ratio	p-value
1	Placebo	15		
	380 mg	19	0.15	<0.01
2	Placebo	25		
	380 mg	25	0.95	0.87
3	Placebo	9		
	380 mg	14	1.18	0.63
4	Placebo	18		
	380 mg	8	0.63	0.45
5	Placebo	8		
	380 mg	9	0.51	0.15
6	Placebo	134		
	380 mg	130	0.75	0.05
Pattern Mixture Analysis	Placebo	209		
	380 mg	205	0.70	<0.01

¹ Stratified by N doses subjects received.

Note: N is the number of subjects who received the exact number of doses shown in that row.

Analyses Repeated Using Alternate Definition of a Heavy Drinking Day

The table below corresponds to Table 3 in the review. However, an alternate definition of heavy drinking day was used in the calculation. A heavy drinking day was defined as a day whereby a man had four or more drinks or a woman had three or more drinks.

	Estimate	Hazard ratio	Unadjusted p-value	Adjusted p-value
190 mg vs. placebo	-0.10	0.90	0.25	0.25
380 mg vs. placebo	-0.20	0.82	0.04	0.07

The table below corresponds to Table 5 in the review. However, an alternate definition of heavy drinking day was used in the calculation. A heavy drinking day was defined as a day whereby a man had four or more drinks or a woman had three or more drinks.

Treatment Group	N	Median Event Rate of Heavy Drinking	Percent Difference	p-value Wilcoxon test unstratified
Placebo	204	0.46		
190 mg	206	0.41	12%	0.85
380 mg	201	0.30	34%	0.09

The table below corresponds to Table 7 in the review. However, an alternate definition of heavy drinking day was used in the calculation. A heavy drinking day was defined as a day whereby a man had four or more drinks or a woman had three or more drinks.

	Placebo (n=204)	190 mg (n=206)	380 mg (n=201)	190 mg versus placebo	380 mg versus placebo
0	8 (4%)	13 (6%)	13(7%)	0.27	0.25
0-1	26 (13%)	31 (15%)	34 (17%)	0.50	0.24
0-2	34 (17%)	36 (18%)	44 (22%)	0.83	0.18
0-3	42 (21%)	42 (20%)	53 (26%)	0.96	0.17
0-4	53 (26%)	55 (27%)	63 (31%)	0.87	0.23

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

/s/

Dionne Price
12/16/2005 05:18:45 PM
BIOMETRICS

Thomas Permutt
12/16/2005 05:20:35 PM
BIOMETRICS
concur

S. Edward Nevius
12/16/2005 05:29:20 PM
BIOMETRICS
Concur with review.