CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER for: 021087

STATISTICAL REVIEW(S)

STATISTICAL REVIEW AND EVALUATION

NDA#:	21-087/S-000
APPLICANT:	Hoffman-La Roche Inc.
NAME OF DRUG:	Oseltamivir
INDICATION:	Treatment of Influenza Infection
TYPE OF REVIEW:	Clinical
DOCUMENTS REVIEWED:	Volumes 3.171, 3.240, 3.252, 3.253, 3.256, 3.271, 3.281, 3.295
MEDICAL INPUT:	Teresa Wu, M.D. (HFD-530)

STATISTICAL REVIEW AND EVALUATION

NDA#: 21-087

1. Background

1.1 Objectives in Trials

The applicant submitted two pivotal randomized, double blind, placebo controlled clinical trials with oseltamivir, trials 15670 and 15671.

The primary objective of these studies was to compare the clinical and antiviral efficacy of oseltamivir at doses of 75 or 150 mg bid to that of placebo in treatment of influenza. The study population in both trials was adults with clinical influenza defined as fever ≥ 38° C plus one respiratory symptom (cough, sore throat, nasal symptoms) plus one constitutional symptom (headache, malaise, myalgia, sweats/chills, fatigue). Subjects were to be recruited within 36 hours of the onset of illness. The ITTI (intent-to-treat infected) population consisted of all subjects randomized who took at least one dose and had laboratory confirmed influenza.

1.2 Summary of Study Designs

1.2.1 Trial 15670

Both studies were double-blind, double-dummy, randomized, three-arm, parallel, placebo-controlled, multi-center trials. Trial 15670 was conducted in Canada (11 centers), Europe (51 centers), and Hong Kong (1 center). Trial 15671 was conducted in the US.

In both trials subjects were randomly assigned in a 1:1:1 ratio to 5 days of 150 mg bid oseltamivir, 75 mg bid oseltamivir, or oseltamivir placebo bid.

In both trials randomization was stratified by current smoking status (yes or no) and by center. In both trials, the details of the randomization are unclear.

1.3 Patient Accounting and Baseline Characteristics

726 patients were enrolled in the trial 15670. Of these, 475 were in the ITTI population. The subjects were enrolled at 63 centers in Canada, Europe, and Hong Kong.

In trial 15670, the ITTI population was 53% male with a mean age of 38 years. They were 94% white and 5% Oriental.

629 patients were enrolled in the trial 15671. Of these, 374 were in the ITTI population. The subjects were enrolled at 60 centers in the US.

In trial 15671, the ITTI population was 52% male with a mean age of 32 years. They were 85% white, 8% black, and 4% Hispanic.

Table 1.3 A summarizes the patient status in trial 15670. Table 1.3 B summarizes the patient status in trial 15671.

TABLE 1.3 A
PATIENT STATUS IN TRIAL 15670

	150 mg	75 mg	Placebo
Randomized	244	243	239
In Treated ITTI	156	158	161
Discontinued	15	8	15
Adverse Event	6	3	6
Refused	4	1	5
Lost to Follow-up	5	4	4

TABLE 1.3 B
PATIENT STATUS IN TRIAL 15671

	150 mg	75 mg	Placebo
Randomized	210	211	208
In Treated ITTI	122	124	128
Discontinued	19	16	11
Adverse Event	6	1	1
Refused	5	4	3
Lost to Follow-up	8	11	7

1.4 Summary of Methods of Assessment

1.4.1 Schedule of Measurements

Patients were given diary cards on which they recorded severity of the 7 of the 8 influenza symptoms used at baseline, (cough, sore throat, nasal symptoms, headache, myalgia, sweats/chills, fatigue), oral temperature, and use of other concomitant medications for influenza. This was done twice daily for eight days. Symptoms were scored as absent, mild, moderate, or severe. Patients returned to the clinic for assessment at day 8. If not all symptoms had been reduced to mild or absent for 3 successive half-day periods, patients were given new diary cards to continue recording symptoms, temperature, and use of other medications until all symptoms were mild or absent for 3 consecutive half-days.

Nasal and throat swabs for viral culture were taken at baseline and every 2 days up to day 8. Subjects with a positive culture at baseline constitute the ITTI (ITT infected) population.

1.4.2 Assessment of Treatment Effects

In both trials, the primary endpoint was time to alleviation of all seven symptoms. Subjects were to record at each diary entry how severe their symptoms were in the previous 12 hours. Thus, time of alleviation was the last time an entry showed any symptom worse than mild provided that entry was followed by two consecutive entries with all symptoms no worse than mild. For subjects who did not alleviate by day 8 (the end of the first diary card), time of alleviation was determined by the answer to the question 'Have your symptoms all been alleviated for the last 24 hours ?' on the second diary card.

1.5 Summary of Statistical Analysis

The primary endpoint was analyzed using Wilcoxon-Gehan scores with stratification by region and smoking status. Regions were defined post enrollment by grouping centers so as to produce 4 regions of approximately equal enrollment. The grouping was done prior to breaking the blind. Differences from placebo were declared statistically significant if they were less

than .05 adjusted for two active arms by the modified Bonferroni procedure.

Missing diary entries were counted as having symptom scores worse than mild. For the primary analysis, subjects who never reported alleviation were considered censored at the time of their last report. For sensitivity analysis, these subjects were considered to alleviate at the longest time actually observed for alleviation.

Unstratified Kaplan-Meier curves and arm medians were used as descriptive statistics for the differences in times to alleviation. 95% confidence intervals for the medians were computed using the bootstrap.

The applicant also conducted two supplemental analyses in which time to healing was defined as the later of the time at which all symptoms were reduced to mild or less for at least 21.5 hours and either the time at which temperature was reduced to 37.2° C or the time after which no medicine for symptom relief was used.

The applicant also reported results of a secondary endpoint as being relevant to labelling. This secondary endpoint was the AUC of total symptom score, an alleged measure of symptom severity. This endpoint was computed by coding each symptom as 0, 1, 2, 3 for the categorical values of none, mild, moderate, and severe, respectively. Missing values were computed by linear interpolation. Observed or interpolated values at each time point were added together for all seven symptoms. The standard trapezoidal rule was used to compute AUC for total symptom score. The AUC's for the two active arms were compared to those for the placebo arm using Wilcoxon tests stratified for region and smoking status.

2. Summary of Applicant's Results

2.1 Trial 15670

Results for both the primary endpoint, time to all symptoms sustained at mild or less, and the supplemental endpoint, time to loss of fever, are summarized in table 2.1 A.

TABLE 2.1 A
Times to Healing, Trial 15670

Times to Symptom Alleviation	Placebo	Osel. 75 mg	Osel. 150 mg
N (ITTI)	161	157	155
Median Time (hrs)	116	87	82
95% Confidence Interval	102-138	73-105	68-100
p-value		.017	.007
Times to Afebrile			
Median Time (hrs)	74	44	44
95% Confidence Interval	64-86	36-54	36-47

There was a statistically significant difference of 29-34 hours in time to alleviation of symptoms in both oseltamivir arms compared to placebo. The observed difference in the medians of 5 hours between the 150 mg and the 75 mg doses was not statistically significant (p-value not reported).

There was a practically important heterogeneity in treatment effect between smokers and non-smokers. The treatment effect was larger among smokers than among non-smokers, as shown in table 2.1 B.

TABLE 2.1 B
Times to Alleviation, Trial 15670
Smokers vs Non-smokers

	Placebo	Osel. 75 mg	Osel. 150 mg
Smokers			
N (ITTI)	59	54	54
Median Time (hrs)	135	80	83
95% Confidence Interval	114-181	72-99	68-97
Non-Smokers			
N (ITTI)	102	103	101
Median Time (hrs)	109	88	81
95% Confidence Interval	84-131	73-130	65-114

2.2 Trial 15671

Results are summarized in table 2.2 A.

TABLE 2.2 A
Times to Alleviation, Trial 15671

	Placebo	Osel. 75 mg	Osel. 150 mg
N (ITTI)	129	124	121
Median Time (hrs)	103	72	70
95% Confidence Interval	93-119	60-83	60-88
p-value		.0001	.006
Sensitivity Analysis			
p-value		.0003	.01
Times to Afebrile			
Median Time (hrs)	65	42	42
95% Confidence Interval	59-76	34-48	35-45

There was a statistically significant difference of 31-33 hours in time to alleviation of symptoms in both oseltamivir arms compared to placebo. This difference remained significant even when all subjects lost to follow-up were assumed to have the same time to alleviation as the longest observed time (p-values under sensitivity analysis).

The observed difference in the medians of 2 hours between the 150 mg and the 75 mg doses was not statistically significant (p-value not reported). There was no practically important heterogeneity in treatment effect between smokers and non-

smokers. The treatment effect was slightly smaller among smokers than among non-smokers, as shown in table 2.2 B.

TABLE 2.2 B
Times to Alleviation, Trial 15671
Smokers vs Non-smokers

	Placebo	Osel. 75 mg	Osel. 150 mg
Smokers			
N (ITTI)	28	29	35
Median Time (hrs)	96	61	79
95% Confidence Interval	53-119	35-89	55-117
Non-Smokers			
N (ITTI)	100	92	84
Median Time (hrs)	109	72	70
95% Confidence Interval	93-120	60-86	56-88

2.3 Results with AUC of Symptom Score

The results from both trials with the AUC of total symptom score are summarized in table 2.3 A. There was a statistically significant reduction in cumulative symptom score in both oseltamivir arms, relative to placebo.

TABLE 2.3 A
AUC OF TOTAL SYMPTOM SCORE

Trial	Placebo	75 mg Osel	150 mg Osel
15670			
Mean AUC	1150	900	900
P-value	na	.007	.003
15671			
Mean AUC	1060	760	740
P-value	na	.0001	.0001

3. Summary of Applicant's Conclusions

The applicant concluded that use of either 75 mg or 150 mg bid oseltamivir resulted in significant improvement in compared to placebo when given in the first two days after onset of symptoms. Furthermore, there was no statistically or practically significant difference in the effect of the two doses. Both doses reduced time to symptom alleviation by approximately 1 1/4 days in both trials.

The applicant also concluded that the use of either dose of oseltamivir reduced the severity of symptoms by 25%.

4. Statistical Reviewer's Comments and Analyses

There are a number of issues that should be addressed concerning the applicant's analyses. These include questions about 1) whether the analysis performed properly reflects the randomization used to assign subjects, 2) the exact algorithm used to compute time to healing from the diary card questions, 3) loss to follow-up and relapses, 4) an apparent interaction between smoking status and treatment effect, and 5) the use of AUC of symptom scores to infer that oseltamivir reduces symptom severity beyond the reduction entailed by shorter time to healing.

4.1 Agreement between Randomization and Analysis

First, the randomization used is not fully described. The assignment appears to be stratified randomization with center and smoking status as the stratum variables but it would be preferable that the applicant clarify whether any modification of this scheme, such as dynamic allocation, was used. Second, the applicant's primary analysis uses a post facto grouping of centers into regions which were selected to equalize observed recruitment. P-values based on analyses which stratify on these regions and on smoking status need not concur with those based on the actual stratification. In this particular study, the p-values based on the applicant's technically incorrect stratification are so small that it defies credibility that the p-values based on the true stratification would not also be small enough to reject both null hypotheses.

4.2 Computation of Times to Healing

The FDA reviewer was unable to reproduce the individual times to symptom alleviation used by the applicant. The FDA reviewer used the following rules to compute times to symptom alleviation. First, symptom alleviation occurs when the subject reports no symptom worse than mild for at least two consecutive diary card records, provided the last such record is at least 21.5 hours after the nearest preceding report of moderate or worse symptoms and provided that there is not a gap of more than 100 hours between the records. Second, symptom alleviation occurs when the subject answers 'yes' to the diary question 'Have all your symptoms alleviated for the last 24 hours?' (This question first appears on the diary cards on day 8.) Time to alleviation is the last time before the earlier of these two events at which the subject reported moderate or worse symptoms during the preceding 12 hours. (This is inferred to mark the beginning of healed period because diary questions all ask 'how did you feel in the previous 12 or 24 hours?' rather than 'how do you feel now?') Subject who do not meet either criterion for alleviation are considered censored at the last time they report symptoms of moderate or worse. Computations are done using reported data only, not interpolations. The applicant reports having used this same algorithm but examination of individual records shows that they have made a number of mistakes.

The FDA reviewer has computed the p-values using Wilcoxon-Gehan tests and Kaplan-Meier curves for the times to healing on all three arms, using both the newly computed times and the times reported by the applicant. Although there were a number of individual discrepancies in the times, there was no difference in the conclusions that both doses of oseltamivir were statistically significantly superior to placebo and that there was no discernible difference between the 75 mg and 150 mg doses. Figure 4.1 shows the Kaplan-Meier curves for the three arms, using the FDA reviewer's computation of times.

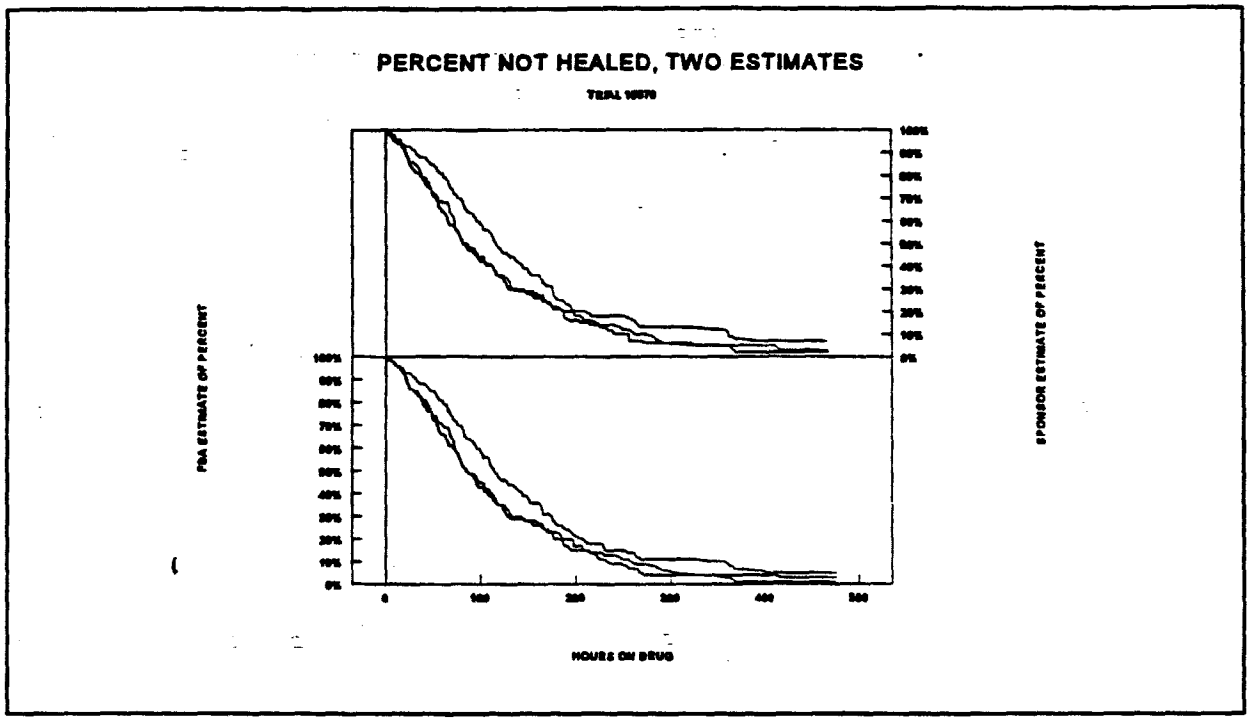

Figure 1

The placebo curve is the one clearly shifted to right of the other two. The two doses of oseltamivir are indistinguishable out to 200 hours on drug.

The p-values for the Wilcoxon-Gehan tests, using the FDA reviewer's calculated time, were .008 for comparing 75 mg to placebo and .004 for comparing 150 mg to placebo.

The magnitude of the difference between the applicant's computed times to alleviation and those of the FDA reviewer can be appreciated from figure 4.2. This graph shows the 95% confidence limits between the percent not healed on placebo and the percent not healed on 75 mg oseltamivir. There are four curves, corresponding to the upper and lower bounds using the FDA's computation of times and using the applicant's computation of times.

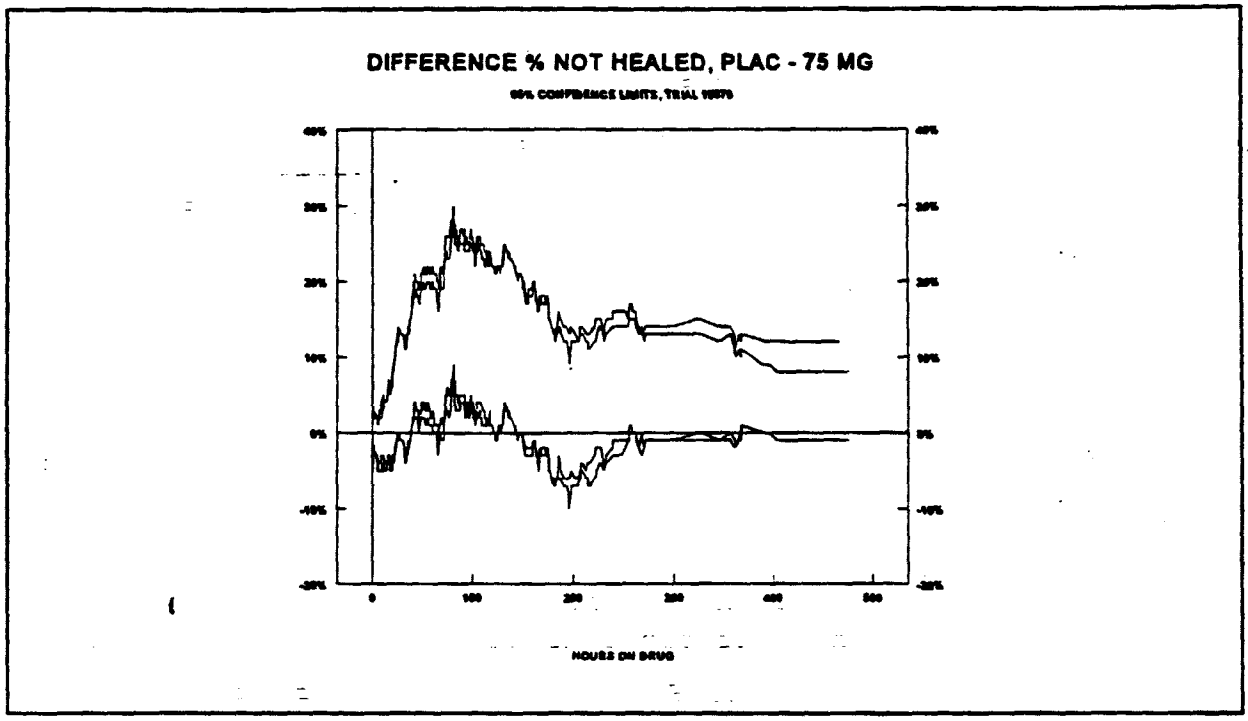

Figure 2

One can see that, at almost all times, the oseltamivir had a statistically significantly higher percent healed and that the differences between the applicant's and the FDA's computations are negligible on practical grounds. As can be inferred from figure 1, the graph for the 95% confidence limits for the difference in percent healed between 150 mg oseltamivir and placebo looks quite similar to figure 2.

Figures 3 and 4 show the same results as figures 1 and 2 for trials 15671. As with trial 15670, the placebo arm is the curve that is shifted to the right in both panels of figure 3. One can see that both oseltamivir arms are essentially the same and that, with 95% confidence, either of them is superior to the placebo arm. Also, there are slight but inconsequential discrepancies between the times to healing as reported by the applicant and as recalculated by the FDA reviewer.

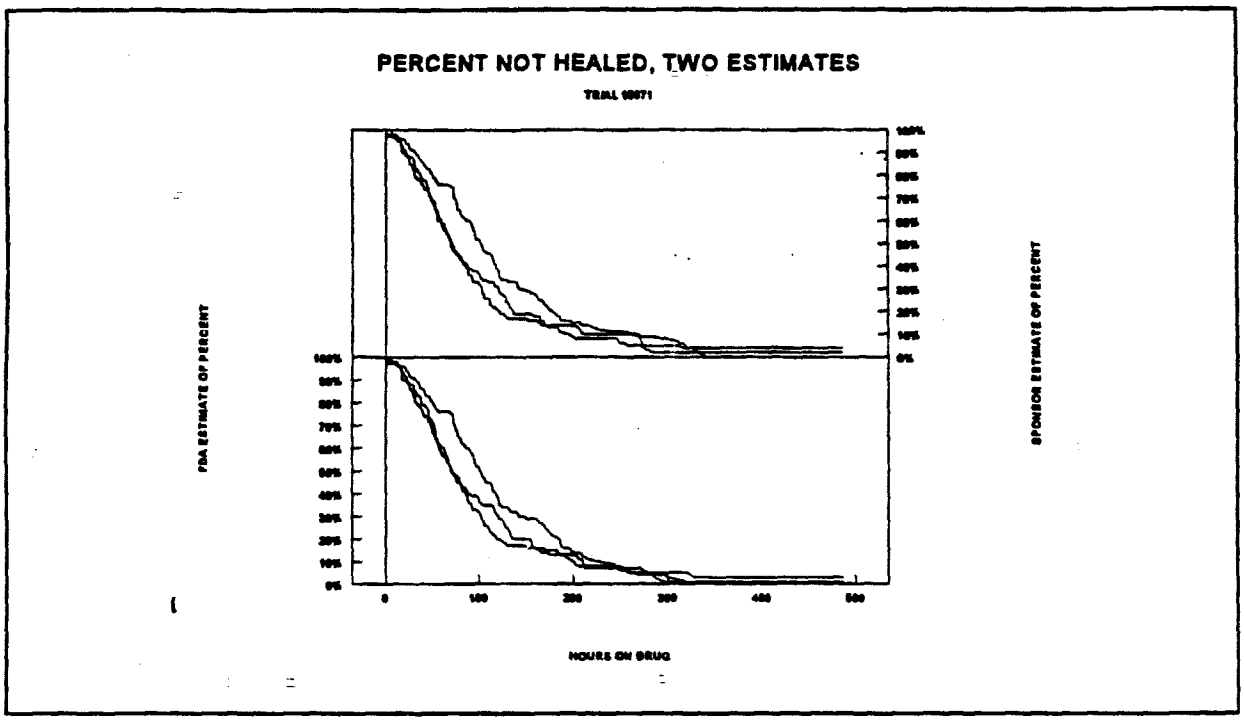

Figure 3

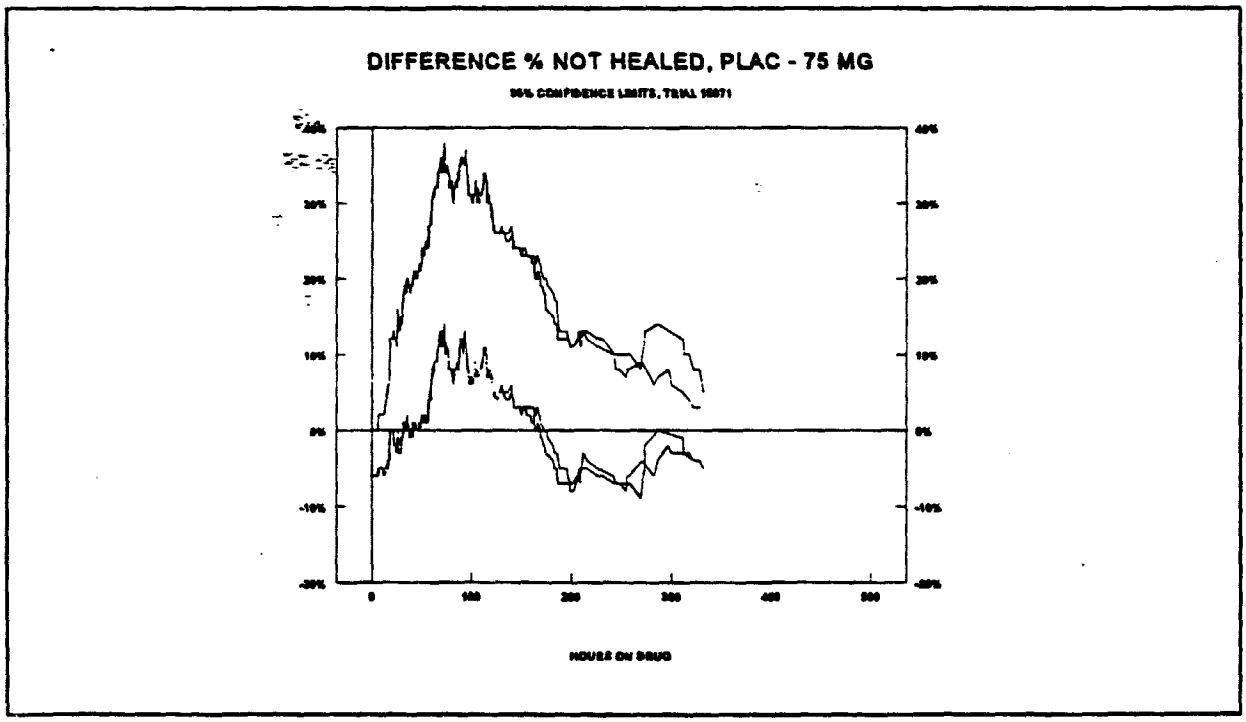

Figure 4

The results for the formal tests for trial 15671 are comparable to those for trial 15670. The p-values for the Wilcoxon-Gehan tests, using the FDA reviewer's calculated time, were .0004 for comparing 75 mg to placebo and .004 for comparing 150 mg to placebo.

4.3 Loss to Follow-up

The applicant performed several sensitivity analyses to assess the effect of loss to follow-up on conclusions. They performed analyses discarding all subjects lost to follow-up before healing and imputing the longest observed healing time to all such subjects. Neither of these analyses substantially altered the conclusions. The FDA reviewer concurs that such analyses are adequate in these trials to conclude that the treatment effect would have remained statistically significant and of essentially the same magnitude had all subject been followed until time to healing.

4.4 Interaction between Smoking Status and Treatment

In trial 15670, there is an observed interaction between treatment and smoking status. The benefit of taking oseltamivir is observed to be much larger in smokers than in non-smokers. This can be seen from the Kaplan-Meier curves for time to symptom alleviation computed for both smoking strata and all three arms. Figure 5 shows these curves using individual times computed by the FDA reviewer; the applicant produced similar curves. One can see that the placebo curve, which is the rightmost in both panels, is much more separated from the oseltamivir curves among smokers.

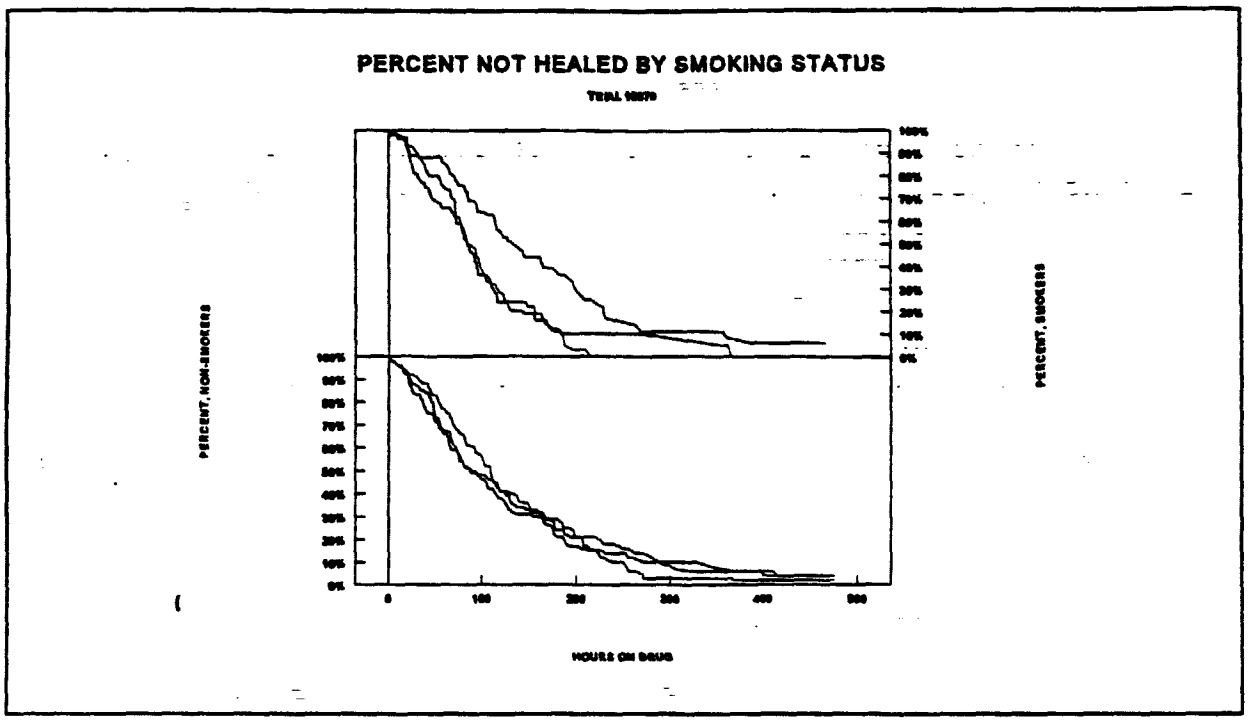

figure 5

Figure 6 shows plots of non-simultaneous 95% confidence limits for the percent still not healed on placebo minus the percent still not healed on 75 mg oseltamivir (positive values show superiority for oseltamivir). The curves for the difference between placebo and 150 mg oseltamivir look similar in each stratum. One can see that even among non-smokers there is an observed superiority which approaches statistical significance for many times. The study was not powered to achieve statistical in each smoking stratum separately. Thus, the most likely conclusion from this study is that the benefit is larger among smokers but not zero among non-smokers.

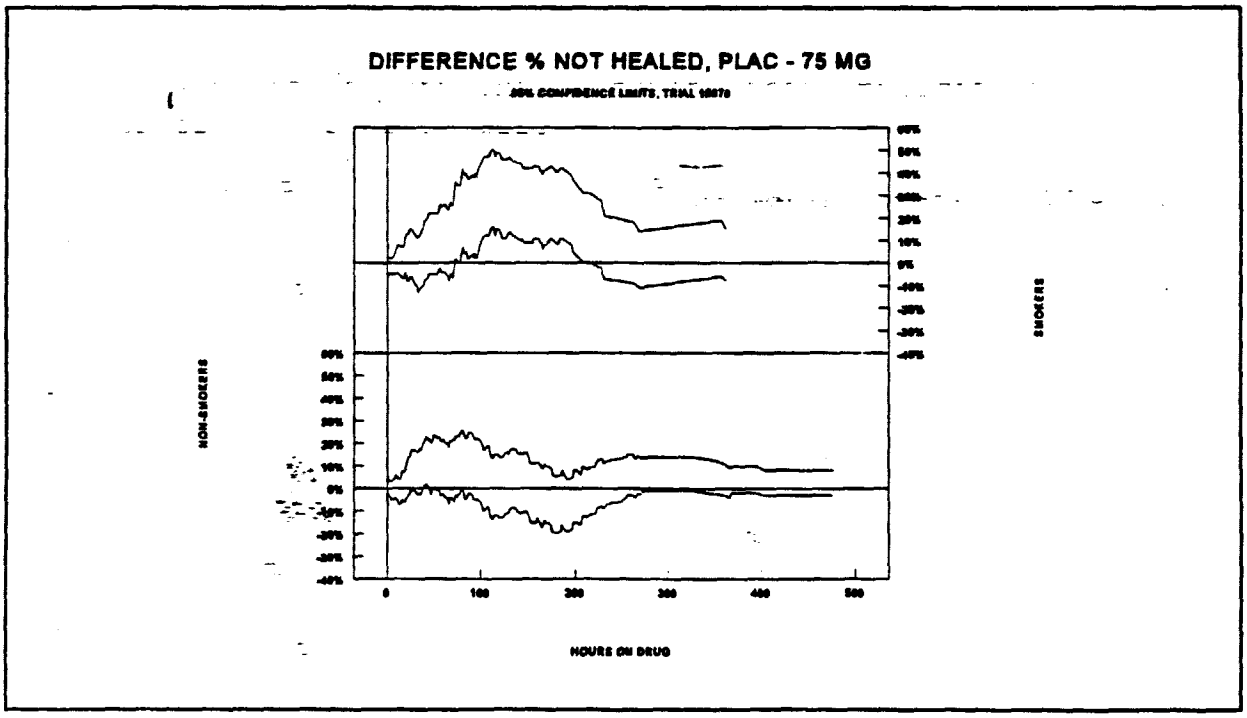

Figure 6

This pattern of interaction was not repeated in the US study. In that study, the non-smokers showed a slightly larger treatment effect than did the smokers. In this study as well, 95% confidence limits for the difference in percent healed between oseltamivir and placebo showed oseltamivir to be statistically significantly superior at many times in both smoking strata. This occurred even thought the study was not powered to detect differences within the individual strata. There does not appear to be enough evidence of an interaction between treatment effect and smoking status to warrant discussion of this issue in the labelling.

4.5 Relapses

The issue of post-healing relapse has been raised in other influenza treatment applications. The nature of data collection in these trials limits the extent to which that issue can be pursued here. The FDA reviewer has computed the time from calculated healing until either time that symptom scores were last reported or until one or more symptom scores were again reported to be moderate or worse. The time from healing until the earlier of these times was considered to be the time to relapse or until censored for relapse. The counts of subjects observed to relapse or censored for relapse are given for each arm for each day post healing in figures 7 and 8, for trials 15670 and 15671, respectively. Subjects with no post healing data are counted as NA under those censored. Subjects not observed to heal at all are excluded from the graphs.

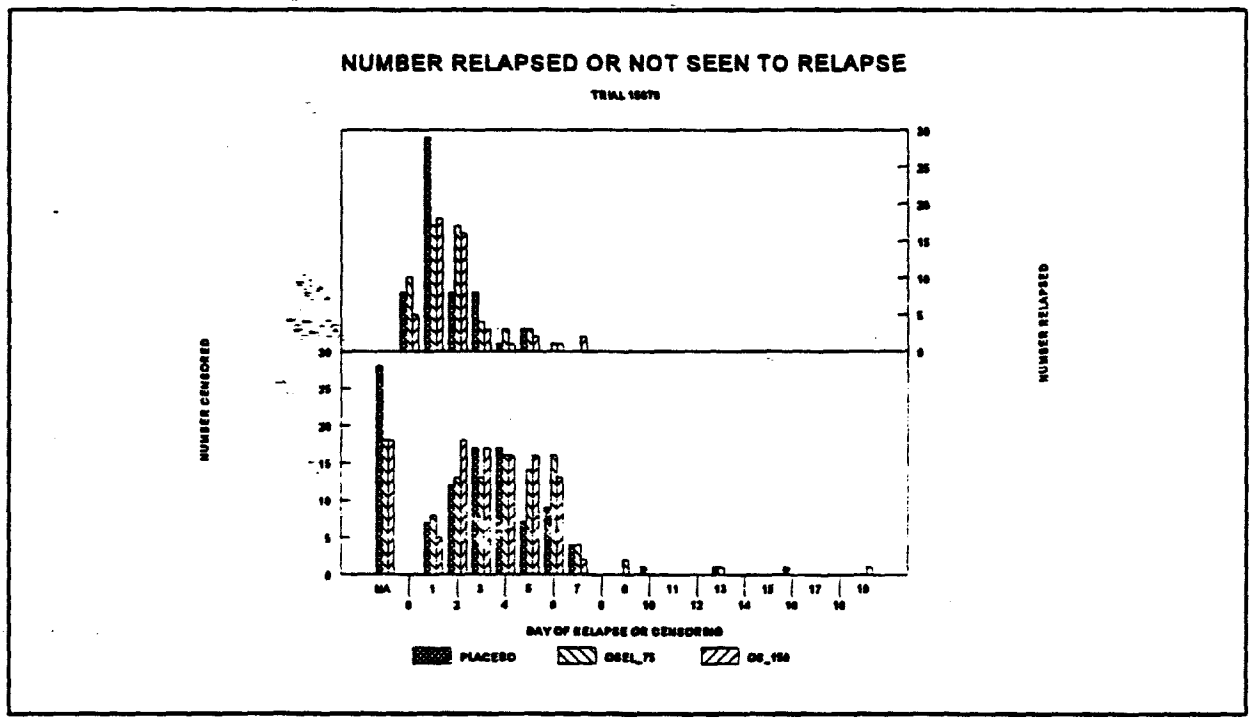

Figure 7

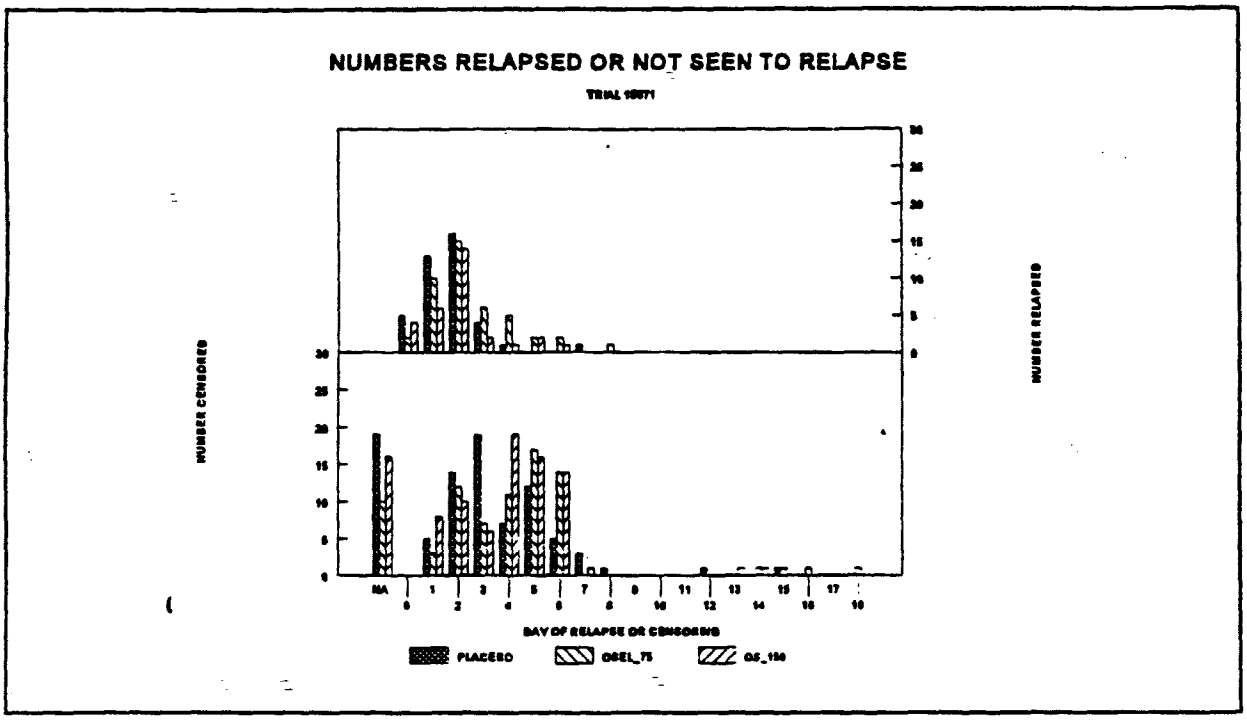

Figure 8

One can see for both trials that somewhat more subjects relapse on days 1-3 post-healing in placebo arm than in the two oseltamivir arms. Because post-healing data for subjects is lost in a potentially biased fashion, one cannot conclude from this that the oseltamivir provides post-healing protection. However, to the limited extent that these data permit, one can see no clear evidence that the earlier healing time for oseltamivir is vitiated by earlier relapses.

4.6 Problems with Center 19167

FDA inspection of sites has revealed possible defects in the recording of the data at center 19167 in trial 15670. Therefore, the FDA reviewer has conducted analyses of this trial with all subjects from center 19167 deleted. The results of this analysis are given in table 4.6 below.

TABLE 4.6

Arm	N	Wilcoxon-Gehan p-value	Median - Placebo Median
75 mg, All subjects	160	.008	28 Hours
75 mg, wo Center 19167	149	.013	28 Hours
150 mg, All subjects	154	.004	30 Hours
150 mg, wo Center 19167	146	.013	22 Hours
Placebo, All subjects	161		
Placebo, wo Center 19167	150		

One can see that even though power diminishes with the loss of 8-11 subjects per arm, the effect remains statistically significant. Furthermore, the difference in median time to symptom alleviation is not changed in the 75 mg arm. It becomes somewhat shorter in the 150 mg arm. However, a conclusion that oseltamivir reduces time to healing by one to one and one quarter days continues to be reasonable even after omission of the data from center 19167.

4.7 Reduction of Symptom Severity

The applicant makes a label claim of reduced severity of symptoms. This claim is based on differences between the arms in AUC of total symptom scores. This analysis is problematic. It is impossible to compute any AUC even for individual symptom scores because symptom scores are not numeric. They are only ordinal categorical data: severe, moderate, mild, and none. The assignment of numeric values 3, 2, 1, 0 to these categories is only a device to facilitate computation of time to alleviation and is not meant to reflect the actual intensity of discomfort. Use of the AUC with these numeric values implies, for instance, that every subject would agree that two days of moderate symptoms are exactly equal in discomfort to one day of severe symptoms and one day of mild symptoms. This assumption is problematic at best. Using total symptom scores requires one to make the even more dubious assumption that one day of severe nasal congestion and mild sore throat is equal in discomfort to one day of moderate cough and moderate fatigue. The reduction of all symptoms to mild level more quickly on oseltamivir tacitly implies that symptoms will be reduced in severity more quickly.

The AUC analysis does not support any stronger conclusion than this.

The FDA reviewer performed an analysis that is designed to identify reductions in symptom scores that are greater than might be expected from the reduction in time to alleviation of all symptoms. For each symptom, the reviewer computed the percent of subjects who had severe levels of that symptom and the percent of subjects who had moderate or severe levels of that symptom on the day prior to healing, on the day two days prior to healing, on the day three days prior to healing, ..., on the day one week prior to healing. When these 14 curves (7 symptoms, 2 levels of severity) for each arm, the 14 placebo curves were indistinguishable from the corresponding oseltamivir curves. (This plot is not reproduced here.) In other words, the entire difference in severity of symptoms was due to faster healing times for oseltamivir.

5. Statistical Reviewer's Summary

The applicant has demonstrated that there is a statistically significantly shorter time to symptom alleviation and loss of fever with either 75 mg or 150 mg oseltamivir bid than with placebo in subjects with confirmed influenza. The improvement in time to symptom alleviation is about 1 1/4 days with either oseltamivir dose. These conclusions are robust to a reasonable variety of methods of handling loss to follow-up. There are inadequate data to draw conclusions about differences among the arms during the period following first reported symptom alleviation. There is also no convincing evidence of a difference in treatment effects between smokers and non-smokers. Finally, there is no convincing evidence that oseltamivir reduces symptom severity to any extent beyond that implicit in the shorter time to symptom alleviation.

/S/

Thomas Hammerstrom, Ph.D.
Mathematical Statistician

/S/

Concur: Dr. Aras

cc:
Archival NDA #21-087

HFD-530
HFD-530/Dr. Jolson
HFD-530/Dr. Birnkrant
HFD-530/Dr. Dempsey
HFD-530/Ms. Carmouze
HFD-530/Dr. Murray
HFD-530/Dr. Wu
HFD-725/Dr. Hammerstrom
HFD-725/Dr. Huque
HFD-725/Ms. Shores

APPEARS THIS WAY
ON ORIGINAL