

**CENTER FOR DRUG EVALUATION AND
RESEARCH**

APPLICATION NUMBER:

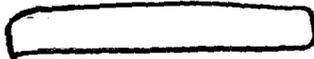
21-110

STATISTICAL REVIEW(S)

STATISTICAL REVIEW AND EVALUATION
CARCINOGENICITY

JUN 22 2000

IND #:



Name of Drug: Rapamune (sirolimus) oral solution

Applicant: Wyeth Ayerst

Documents Reviewed: Volumes 1-8 and electronic data submission

Statistical Reviewer: Nancy Silliman, Ph.D.

Pharmacologist: Steve Kunder, Ph.D.

Project Manager: Matthew Bacho

Key Words: Peto, trend test, pairwise comparisons, adjusted p-values,
adjusted α -levels

Summary of Review

- In this review, a tumor is declared "rare" if the incidence in the control group is $\leq 1\%$ and "common" if the incidence in the control group is $> 1\%$.
- For positive linear trend analyses, "rare" tumors are tested using a significance level of 0.025; "common" tumors are tested using a significance level of 0.005.
- For pairwise comparisons, "rare" tumors are tested using a significance level of 0.05, while "common" tumors are tested using a significance level of 0.01.
- Problems related to the maximum tolerated dose (MTD) being exceeded were seen in the mouse study that was submitted with the original NDA (approved 9/15/99). This submission provides the results from a second study that was conducted in mice. The first study dosed the mice at 0, 12.5, 25, and 50 mg/kg/day, while the current study dosed the mice at 0, 1, 3, and 6 mg/kg/day. When survival in the 6 mg/kg/day group (current study) was compromised due to the occurrence and progression of skin lesions, the surviving mice in this group were sacrificed at week 85 for males and week 97 for females.
- A significant dose-related decreasing trend in survival was observed for both males and females. The trend tests used in this review which look at tumor incidence adjust for differences in intercurrent mortality.
- The percent loss in group mean body weights compared to controls was greater than 10% in the low, medium, and high dose male (14%, 13%, and 19%, respectively) and female mice (11%, 16%, and 17%, respectively). This suggests that these dose groups can be considered appropriate, i.e., close to MTD.
- This reviewer found eleven significant trend tests in the mouse study, all in female mice. The first was for the incidence of adrenal medulla lymphoma ($p=0.007$), a rare tumor. Pairwise comparisons for this tumor were not significant in the low or medium dose groups, but were significant for the high dose group ($p=0.048$). The second was for the incidence of bone marrow histiocytic sarcoma ($p=0.013$), a rare tumor. Pairwise comparisons for this tumor were not significant in the low or medium dose groups, but were significant for the high dose group ($p=0.036$). The remaining nine significant trend tests were all for granulocytic leukemia in the following organs: adrenal cortex ($p=0.017$), gallbladder ($p=0.017$), kidneys ($p=0.017$), mesenteric lymph ($p=0.018$), ovaries ($p=0.017$), salivary gland ($p=0.018$), thymus

($p=0.015$), bone marrow ($p=0.018$), and brain ($p=0.017$). All were considered rare tumors. As the same three mice (one in the medium and two in the high dose group) were found to each have granulocytic leukemia in the 9 organs mentioned above, this reviewer combined granulocytic leukemia across organs to find a significant trend test for the incidence of granulocytic leukemia (multisystemic) in females ($p=0.009$), a rare tumor. Pairwise comparisons for this tumor were not significant in the low, medium, or high dose female groups. (Note: The sponsor also found a significant trend for multisystemic granulocytic leukemia in females, but did not present results for granulocytic leukemia in females by individual organs. In addition, the sponsor performed cross-tabulations of the incidence of the first two tumors mentioned in this paragraph, adrenal medulla lymphoma and bone marrow histiocytic sarcoma, by treatment group in females. However, they did not perform a trend test for either tumor. There was no explanation of why these tests were not performed.)

- There was one tumor that was found to have a significant trend test in the sponsor's analysis that was not found to be significant in this reviewer's analysis. This was for the incidence of hepatocellular adenoma in males ($p=0.003$ sponsor analysis; $p=0.010$ reviewer analysis), a common tumor. In addition, the sponsor combined lymphoma across organs in males and found a significant trend test for the incidence of lymphoma (multisystemic) in males ($p=0.005$), a common tumor.

I. Background

There were problems associated with the MTD being exceeded in the original mouse study that was submitted with NDA 21-083 (approved 9/15/99). As a result, a second study was performed in mice. This IND submission includes the results of that second mouse carcinogenicity study, Study 96047. Rapamune, an immunosuppressant, was administered orally by gavage once daily to evaluate its carcinogenic potential.

Male and female CD-1 VAF mice were dosed at 0, 0, 1, 3, and 6 mg/kg/day for up to 104 weeks. There were 60 mice/sex in the first control group, 75 mice/sex in the second control group, and 75 mice/sex in each of the low, medium, and high dose groups. Control groups were combined for analysis by the sponsor and reviewer. Approximately 55 tissues from each animal were examined macroscopically and microscopically.

A subgroup was sacrificed after 52 weeks to evaluate the chronic toxicity in mice. In addition, when survival in the 6 mg/kg/day group was compromised due to the occurrence and progression of skin lesions, the surviving mice in this group were sacrificed at week 85 for males and week 97 for females.

**APPEARS THIS WAY
ON ORIGINAL**

II. Methods

For each species and sex, the sponsor and reviewer analyzed palpable, nonpalpable-lethal, and nonpalpable-nonlethal tumors separately, then combined the results using Peto et al.¹ procedures. For a particular tumor type of interest, the incidence data can be summarized in a $2 \times D$ table, where D is the number of dose groups. The first row contains the number of animals with the tumor of interest, and the second row contains the number of animals without the tumor. However, this summary table can be misleading. If the toxicity of the drug causes animals to die early by some non-cancer related cause, fewer animals will be at risk for tumors in the higher dose groups. Thus, even if the drug also increases the tumor rate, the overall incidence of that tumor in the high dose groups may be smaller than in the control groups. To adjust for the effect that potential differential mortality between the dose groups has on tumor occurrence, the Peto prevalence method breaks up study time into several discrete intervals. The intervals used in this study were weeks 1-52, weeks 53-78, weeks 79-91, weeks 92-104, and terminal sacrifice. The data can thus be represented by several $2 \times D$ tables, one for each time interval. Note that Peto et al. point out that "the effects of differences in longevity on numbers of tumor-bearing animals can be very substantial, and so, whether or not they appear to be they should routinely be corrected for when presenting experimental results."

The dose groups can also be assigned weights in the statistical analysis to test various hypotheses. For example, using weights of $(0, 1, \dots, D)$ gives the trend test which is sensitive to a linear dose effect. Using equal weights $(1, 1, \dots, 1)$ gives a test of association between dose and tumor rate without specifying the form of the relationship. Weight can also be set equal to the actual doses given. Finally, choosing weights close to the actual biological effect of the doses will result in the most sensitive test, but in practice this effect is not known. Linear weights or dose weights are often used. The sponsor used both linear and dose weights in their analysis. As results were similar between the two, this reviewer used dose weights.

For the tumor type of interest, each tumor is classified as "fatal", "non-fatal", or "mortality independent tumors". P-values are calculated for the three classes separately, and then combined to yield a single p-value for the tumor type. Both exact and asymptotic p-values are calculated for most tumor types (statisticians in CDER routinely use the exact method to test for positive linear trend when the number of total tumors across treatment groups is 10 or smaller).

One-sided p-values may be more appropriate than two-sided p-values, since they are more conservative and we are only interested in whether increased doses *increase* tumor incidence. One-sided p-values are reported in this review.

As so many sex/species/organ/tumor type combinations are tested, a simple application of a 0.05 decision rule does not appropriately control the overall false positive rate. It has been suggested by Drs. Lin and Rahman² that if the tumor is "rare" the cutoff should be 0.025 and if the tumor is "common" the cutoff should be 0.005. Tumors are defined as rare or common using historical control data or the control group in the study being analyzed (in this review, the

¹ Peto R, Pike MC, Day NE, Gray RG, Lee PN, Parish S, Peto J, Richards S, and Wahrendorf J (1980). "Guidelines for Simple, Sensitive Significance Tests for Carcinogenic Effects in Long-term Animal Experiments", in Long-term and Short-term Screening Assays for Carcinogens: An Critical Appraisal, World Health Organization.

² Lin KK and Rahman MA (1998), "Overall False Positive Rates in Tests for Linear Trend in Tumor Incidence in Animal Carcinogenicity Studies of New Drugs", *Journal of Biopharmaceutical Statistics* 8(1), pgs. 1-15.

control group is used). The usual practice at FDA is to classify a tumor as common if it occurs in the control group at an incidence of greater than 1%. Using simulation tests on CD-1 rats and CD(BR) mice, Lin and Rahman found that the overall false positive rate resulting from the use of the α -levels 0.025 and 0.005 in the tests for linear trend in a two-species-two-sex study is about 10%. These false-positive rates are judged by the Center for Drug Evaluation and Research as the most appropriate in a regulatory setting.

For pairwise comparisons, the levels of significance that are used are 0.05 for rare tumors and 0.01 for common tumors.

III. Analysis

The percent loss in group mean body weights compared to controls was greater than 10% in the low, medium, and high dose male (14%, 13%, and 19%, respectively) and female mice (11%, 16%, and 17%, respectively). This suggests that these dose groups can be considered appropriate, i.e., close to MTD.

A significant dose-related decreasing trend in survival was observed for both males and females (two-sided $p < 0.001$ using both linear and dose weights for both males and females, separately). For the males, survival rates for the control groups (combined), low, medium, and high dose groups were 37%, 29%, 31%, and 19% (for the latter, this is estimated survival at week 85), respectively. For the females, survival rates for the control groups (combined), low, medium, and high dose groups were 33%, 64%, 46%, and 18% (for the latter, this is estimated survival at week 97), respectively. The trend tests used in this review which look at tumor incidence adjust for differences in intercurrent mortality.

Both the reviewer's and the sponsor's analyses used the Peto et al. procedures described above. Reviewer analyses for male mice (exact and asymptotic trend tests) can be found in Appendix 1. Reviewer analyses for female mice (exact and asymptotic trend tests) are given in Appendix 2.

This reviewer found eleven significant trend tests in the mouse study, all in female mice. The first was for the incidence of adrenal medulla lymphoma ($p=0.007$), a rare tumor. Pairwise comparisons for this tumor were not significant in the low or medium dose groups, but were significant for the high dose group ($p=0.048$). The second was for the incidence of bone marrow histiocytic sarcoma ($p=0.013$), a rare tumor. Pairwise comparisons for this tumor were not significant in the low or medium dose groups, but were significant for the high dose group ($p=0.036$). The remaining nine significant trend tests were all for granulocytic leukemia in the following organs: adrenal cortex ($p=0.017$), gallbladder ($p=0.017$), kidneys ($p=0.017$), mesenteric lymph ($p=0.018$), ovaries ($p=0.017$), salivary gland ($p=0.018$), thymus ($p=0.015$), bone marrow ($p=0.018$), and brain ($p=0.017$). All were considered rare tumors. As the same three mice (one in the medium and two in the high dose group) were found to each have granulocytic leukemia in the 9 organs mentioned above, this reviewer combined granulocytic leukemia across organs to find a significant trend test for the incidence of granulocytic leukemia (multisystemic) in females ($p=0.009$), a rare tumor. Pairwise comparisons for this tumor were not significant in the low, medium, or high dose female groups. (Note: The sponsor also found a significant trend for multisystemic granulocytic leukemia in females, but did not present results for granulocytic leukemia in females by individual organs. In addition, the sponsor performed cross-tabulations of the incidence of the first two tumors mentioned in this paragraph, adrenal medulla lymphoma and bone marrow histiocytic sarcoma, by treatment group in females.

However, they did not perform a trend test for either tumor. There was no explanation of why these tests were not performed.)

There was one tumor that was found to have a significant trend test in the sponsor's analysis that was not found to be significant in this reviewer's analysis. This was for the incidence of hepatocellular adenoma in males ($p=0.003$ sponsor analysis; $p=0.010$ reviewer analysis), a common tumor. In addition, the sponsor combined lymphoma across organs in males and found a significant trend test for the incidence of lymphoma (multisystemic) in males ($p=0.005$), a common tumor.

IV. Discussion

As is common practice in the Center for Drug Evaluation and Research, in this review a tumor is declared "rare" if the incidence in the control group is $\leq 1\%$ and "common" if the incidence in the control group is $>1\%$. As suggested by Drs. Lin and Rahman, for positive linear trend analyses, "rare" tumors are tested using a significance level of 0.025, while "common" tumors are tested using a significance level of 0.005. For pairwise comparisons, "rare" tumors are tested using a significance level of 0.05, and "common" tumors are tested using a significance level of 0.01.

The data regarding loss of body weight suggests that animals were adequately dosed in this study. A significant dose-related decreasing trend in survival was observed for both males and females. However, the trend tests used in this review which look at tumor incidence adjust for differences in intercurrent mortality.

This reviewer found seven significant trend tests in the mouse study, all in female mice. The first was for the incidence of adrenal medulla lymphoma ($p=0.007$), a rare tumor. Pairwise comparisons for this tumor were not significant in the low or medium dose groups, but were significant for the high dose group ($p=0.048$). The second was for the incidence of bone marrow histiocytic sarcoma ($p=0.013$), a rare tumor. Pairwise comparisons for this tumor were not significant in the low or medium dose groups, but were significant for the high dose group ($p=0.036$). The remaining nine significant trend tests were all for granulocytic leukemia in the following organs: adrenal cortex ($p=0.017$), gallbladder ($p=0.017$), kidneys ($p=0.017$), mesenteric lymph node ($p=0.018$), ovaries ($p=0.017$), salivary gland ($p=0.018$), thymus ($p=0.015$), bone marrow ($p=0.018$), and brain ($p=0.017$). All were considered rare tumors. As the same three mice (one in the medium and two in the high dose group) were found to each have granulocytic leukemia in the 9 organs mentioned above, this reviewer combined granulocytic leukemia across organs to find a significant trend test for the incidence of granulocytic leukemia (multisystemic) in females ($p=0.009$), a rare tumor. Pairwise comparisons for this tumor were not significant in the low, medium, or high dose female groups. (Note: The sponsor also found a significant trend for multisystemic granulocytic leukemia in females, but did not present results for granulocytic leukemia in females by individual organs. In addition, the sponsor performed cross-tabulations of the incidence of the first two tumors mentioned in this paragraph, adrenal medulla lymphoma and bone marrow histiocytic sarcoma, by treatment group in females. However, they did not perform a trend test for either tumor. There was no explanation of why these tests were not performed.)

There was one tumor that was found to have a significant trend test in the sponsor's analysis that was not found to be significant in this reviewer's analysis. This was for the incidence of hepatocellular adenoma in males ($p=0.003$ sponsor analysis; $p=0.010$ reviewer analysis), a common tumor. In addition, the sponsor combined lymphoma across organs in males and found

a significant trend test for the incidence of lymphoma (multisystemic) in males ($p=0.005$), a common tumor.

/S/

6/22/00

Nancy Paul Silliman, Ph.D.
Statistical Reviewer, DB III

/S/

6/22/00

Concur: Karen Higgins, Sc.D.
Acting Team Leader, DB III

cc:

HFD-590

HFD-590/Dr. Albrecht

HFD-590/Dr. Hastings

HFD-590/Dr. Kunder

HFD-590/Dr. Cavaille-Coll

HFD-590/Dr. Tierman

HFD-590/Mr. Eacho

HFD-725/Dr. Fudge

HFD-725/Dr. Higgins

HFD-725/Dr. Silliman

HFD-725/chron.

This review contains 57 pages (including appendices).

NDA #: 21-110

Applicant: Wyeth-Ayerst Laboratories

Name of Drug: Rapamune® (Sirolimus) Tablets

Documents Reviewed: NDA Index and Summary sections (Vol. 1.1-1.3) and Statistical sections (Vols. 1.52-1.59) dated October 29, 1999, SAS datasets, 4-month Safety Update Report dated February 29, 2000, and the electronic regulatory reviewer aids dated October 1999 and April 2000 containing the NDA paper volumes as PDF files and SAS XPT files of the statistical data.

Indication: Prophylaxis of organ rejection in patients receiving renal transplants.

Statistical Reviewer: Cheryl Dixon, Ph.D.

Medical Reviewer: Dr. Rosemary Tieman (HFD-590)

I. Introduction

Rapamune® Oral Solution, approved September 15, 1999, has been shown to be safe and effective in the prophylaxis of organ rejection in patients receiving renal transplants. This NDA seeks approval of the tablet dosage form for the same indication as the approved oral solution. The tablet formulation has several advantages over the oral solution formulation in terms of convenience for the patient. Unlike the oral solution, the tablet can be stored at room temperature and does not require reconstitution. The tablet is not bioequivalent to the oral solution but data has been submitted to demonstrate the clinical comparability of the two dosage forms.

The primary clinical evidence of the safety and efficacy of the sirolimus tablet formulation focuses on a single Phase III study, Study 309. Study 309 was designed to compare the safety and efficacy, and to determine pharmacokinetic profiles, of sirolimus oral solution and tablets in de novo renal transplant recipients. The study is an open-label, multi-center study which examined the safety and efficacy of a tablet formulation of sirolimus (2 mg/d2y) in combination with Neoral®/corticosteroids in recipients of primary or secondary cadaveric or living, non-haploidentical renal allografts. The control group received concomitant immunosuppression consisting of Neoral®/corticosteroids and the oral solution formulation of sirolimus (Rapamune®). Patients were randomly

assigned to treatment group before transplantation. A single dose level of 2 mg/day of sirolimus was selected based upon both the results from phase II studies and the observed overall rate of acute rejection while the oral solution phase III studies were still blinded.

Thirty centers in Australia, Canada and the United States participated in Study 309. Four hundred seventy-seven patients were enrolled in the study; 238 were randomly assigned to sirolimus oral solution treatment and 239 were assigned to sirolimus tablet treatment.

As stated in the protocol, the primary endpoint was efficacy failure in the first 3 months (≤ 104 days) after transplantation. Efficacy failure was defined as the first occurrence of biopsy-confirmed acute rejection, graft loss, or death. Prospectively defined secondary endpoints were patient and graft survival, incidence of biopsy-confirmed acute rejection, graft function (measured by serum creatinine and calculated creatinine clearance), incidence of documented infection or presumptive infection analyzed, and the incidence of histologically confirmed lymphoproliferative disease or other malignancy. The above secondary endpoints were all analyzed at 3, 6 and 12 months after transplant. Efficacy failure at 6 and 12 months after transplant was considered as 'other descriptive analyses' in the protocol.

Reviewer's Comment: The approval of sirolimus oral solution was based upon efficacy failure at 6 months and patient and graft survival at 1 year as co-primary endpoints. Study 309 was designed to utilize efficacy failure at 3 months, along with 12-month patient and graft survival, to assess the activity of the two formulations of sirolimus. The 3-month endpoint was considered to be a valid assessment of efficacy because the highest incidence of acute rejection occurs within the first 3 months after transplantation. At the pre-NDA telecon held on April 7, 1999, the Division stated that the 6-month efficacy failure endpoint would be considered along with the 3-month efficacy failure and 12-month patient and graft survival endpoints.

For the purpose of determining sample size, the rate of efficacy failure at 3 months was estimated to be as high as 16% for the Rapamune[®] oral solution treatment group. With 225 patients randomly assigned to each treatment group, a test of equivalence using the difference in proportions of the primary endpoint with a one-sided 0.025 significance level would have greater than 90% power to reject the null hypothesis. The null hypothesis states that the tablet and oral solution are not equivalent [i.e., the tablet treatment group has a rate of efficacy failure that is at least 15 percentage points higher than that of the oral solution].

There were several issues that led to the acceptability of the open-label nature of this study. There was no preliminary efficacy data with the tablet formulation prior to the start of the study. Thus, there was reluctance on the part of the study investigators to participate in a double-dummy design trial. In the majority of the cases, the same study centers and investigators who participated in the phase III pivotal studies for the oral

solution also conducted study 309. The primary endpoint (acute rejection, graft loss, or death) was composed of well-established, objective, clinically relevant variables, unlikely to be affected by patient or investigator bias. Further, the diagnosis of acute rejection required biopsy confirmation by a trained pathologist who was blinded to patient treatment assignments.

II. Efficacy Evaluation

The primary analysis of efficacy failure consisted of calculating a two-sided 95% confidence interval around the differences in rates for the two formulations (tablet – oral solution). All patients assigned to treatment were included in this analysis. Equivalence of the tablet to the oral solution is demonstrated if the 95% confidence interval crosses zero and remains within a pre-defined upper bound (i.e., if the rate of the endpoint for the tablet formulation did not exceed the oral solution by more than a fixed number of percentage points). The upper bound of the confidence interval was chosen on the basis of the efficacy of the oral solution. The more efficacious the oral solution, the more stringent the definition of equivalence. The prospective definition of equivalence stated by the sponsor is:

Table 1
Definitions of Equivalence

If the rate of the efficacy endpoint for the oral solution is greater than:	But less than or equal to:	The upper bound of the CI* will be no greater than:
0%	10%	10%
10%	20%	15%
20%	30%	20%

*: The 95% confidence interval of the difference in rates of the tablet minus the oral solution.

Reviewer's Comment: Regardless of the rate of the endpoint for the oral solution, the Division would prefer a delta of no more than 10% in assessing the equivalence of transplant products.

Similarity with respect to patient and graft survival incidence rates was assessed with confidence intervals about the difference in rates (tablet – oral solution). The lower bound of the confidence interval is used to assess the maximum decrease in patient and graft survival that one can safely exclude. These rates need to be taken into consideration when assessing the overall efficacy and safety of Rapamune® tablets.

Secondary endpoints defined as binary events and summarized by incidence rates were analyzed using Fisher's exact test. Survival and other time-to-event variables were analyzed by the log-rank test.

- **Patient Demographics**

A total of 477 patients were enrolled in the study and randomized to one of the sirolimus treatment groups: oral solution (n=238) and tablet (n=239). Twenty patients

were randomized into the study but did not receive at least one dose of study medication (9 oral solution and 11 tablet). The most common reason for withdrawal before receiving study medication was protocol violation (n=14). Other reasons include patient request, physician request, compliance issue, and death.

Table 2 summarizes the demographic and baseline characteristics for all randomized patients. There were no statistically significant differences between the treatment groups. The majority of the patients were male and Caucasian. The source of the donor allograft was primarily cadaveric. The descriptive variables, which include gender, race, and donor source, were evaluated using CMH tests stratified by investigator. Age was evaluated using ANOVA with treatment and investigator as factors.

Table 2
Patient Demographics

	Sirolimus Oral Solution	Sirolimus Tablets	P-value
# Patients	238	239	-
Gender N (%)			0.830
Female	95 (39.9)	93 (38.9)	
Male	143 (60.1)	146 (61.1)	
Age mean (SD)	44.6 (12.8)	46.0 (13.0)	0.216
Min, max	17, 70	16, 74	
Race N (%)			0.921
Caucasian	140 (58.8)	137 (57.3)	
Black	54 (22.7)	59 (24.7)	
Hispanic	24 (10.1)	22 (9.2)	
Oriental (Asian)	9 (3.8)	12 (5.0)	
Other	11 (4.6)	9 (3.8)	
Donor Source N (%)			0.571
Cadaver	158 (66.4)	170 (71.1)	
Living (Related)	55 (23.1)	49 (20.5)	
Living (Unrelated)	25 (10.5)	20 (8.4)	

- **Analysis Results**

Tables 3 and 4 summarize the results of the primary endpoint, efficacy failure, for each treatment group at 3 and 6 months, respectively. The following are included in the tables.

1. The overall rates of efficacy failure for each treatment group and the rates for each component of the composite endpoint.
2. The difference in overall rates of efficacy failure adjusted for investigator and corresponding confidence interval. A difference less than 0 indicates a lower rate of efficacy failure in the sirolimus tablet group than in the oral solution group.

The overall rate of efficacy failure at 3 months in the tablet treatment group (24.7%) is equivalent to that in the oral solution treatment group (23.5%). The upper bound of the 95% confidence interval for the difference in rates is less than 10%.

Table 3
Efficacy Failure at 3 months

	Sirolimus Oral Solution (n=238)	Sirolimus Tablet (n=239)
Overall rate of efficacy failure, n(%)	56 (23.5)	59 (24.7)
Acute rejection	45 (18.9)	42 (17.6)
Graft loss	8 (3.4)	15 (6.3)
Death	3 (1.3)	2 (0.8)
Stratified differences in rates (95% CI)		1.0 (-6.9, 8.9)

Between 3 and 6 months, there were 6 additional efficacy failures in each treatment group. These included 5 acute rejections and 1 death in the oral solution treatment group and 4 acute rejections and 2 deaths in the tablet treatment group. Thus, the overall rates of efficacy failure at 6 months increased to 27.2% in the tablet treatment group and 26.1% in the oral solution treatment group. The upper bound of the 95% confidence interval for the difference in rates is less than 10%, which implies equivalence of the two sirolimus formulations.

Table 4
Efficacy Failure at 6 months

	Sirolimus Oral Solution (n=238)	Sirolimus Tablet (n=239)
Overall rate of efficacy failure, n(%)	62 (26.1)	65 (27.2)
Acute rejection	50 (21.0)	46 (19.2)
Graft loss	8 (3.4)	15 (6.3)
Death	4 (1.7)	4 (1.7)
Stratified differences in rates (95% CI)		1.1 (-7.0, 9.2)

Table 5 includes the results of patient and graft survival 12 months after transplantation for each treatment group. Differences between sirolimus oral solution and tablet were assessed using Fisher's exact test. There was not a statistically significant difference in the rate of patient and graft survival between the two sirolimus formulations. The Rapamune® oral solution treatment group had a slightly better patient and graft survival rate at 12 months than the tablet treatment group. The exact 95% confidence interval about the difference in patient and graft survival rates indicates equivalence at a delta less than 15%. The lower bound of this confidence interval is -10.2. The upper bound of the confidence interval for relative risk implies that the risk of graft loss or death with a functioning graft could be as much as 2 to 3 times greater for a patient on sirolimus tablet compared to the oral solution. Patients who died with a functioning graft accounted for about 33% of the total number of graft losses. There were numerically more pure graft losses in patients who received the tablet formulation than in patients who received the oral solution. According to the sponsor, this numerical difference does

not appear to be due to factors related to drug formulation (see medical reviewer's review for a complete discussion of these graft losses).

Table 5
Patient and Graft Survival at 12 months

	Sirolimus Oral Solution (n=238)	Sirolimus Tablet (n=239)
Patient and Graft survival, n(%)	219 (92.0)	212 (88.7)
Graft loss	11	20
Death w/ functioning graft	8	7
Fisher's exact p-value		0.278
Relative risk (95% CI)		1.42 (0.81, 2.47)
Differences in rates (Exact 95% CI)		-3.3 (-10.2, 2.8)

Table 6 includes the results of patient survival 12 months after transplantation for each treatment group. There was not a statistically significant difference in the rate of patient survival between sirolimus formulations. The exact 95% confidence interval about the difference in survival rates indicates equivalence at a delta less than 5%. The lower bound of this confidence interval is -4.4. The upper bound of the confidence interval for the relative risk implies that the risk of death could be as much as 2 times greater for a patient on sirolimus tablet compared to the oral solution.

Table 6
Patient Survival at 12 months

	Sirolimus Oral Solution (n=238)	Sirolimus Tablet (n=239)
Patient survival, n(%)	228 (95.8)	230 (96.2)
Death	10	9
Fisher's exact p-value		0.271
Relative risk (95% CI)		0.90 (0.37, 2.17)
Differences in rates (Exact 95% CI)		0.4 (-4.4, 5.4)

Reviewer's Comment: *Since there are relatively few additional failures at 6 months compared to 3 months, only the 6 month efficacy failure analyses will be reported in this review for the following supplementary analyses. The conclusions drawn at 6 months are similar to those reported by the sponsor for efficacy failure at 3 months.*

The first acute rejection episode was classified by the criteria of grade I (mild), grade II (moderate), or grade III (severe) acute rejection. Patients not having efficacy failure were categorized as none and patients who had an outcome of graft loss or death were categorized as other. Treatment differences in histological grade of the first acute

rejection episode were assessed through generalized CMH methods (row means score statistic) because of the ordinal nature of the response. The distribution of histological grade of acute rejection is not different between the oral solution and tablet treatment groups.

Table 7
Histological Grade of Acute Rejection at 6 Months

	Sirolimus Oral Solution (n=238, 50) ^a	Sirolimus Tablet (n=239, 46)
None	176 (74.0, -) ^b	174 (72.8, -)
Grade I (mild)	29 (12.2, 58.0)	24 (10.0, 52.2)
Grade II (moderate)	18 (7.6, 36.0)	21 (8.8, 45.7)
Grade III (severe)	3 (1.3, 6.0)	1 (0.4, 2.2)
Other	12 (5.0, -)	19 (8.0, -)
p-value		0.357, 0.866 ^c

a: Total number of randomized patients, Number of patients with acute rejections

b: # of patients with event (Percent of all randomized patients, Percent of acute rejections)

c: All randomized patients, Acute rejections only

Rates of efficacy failure were calculated for the following subgroups: recipient race (black, non-black), recipient gender (female, male), donor source (cadaver, living related, living unrelated), and number of HLA mismatches (0 to 2 mismatches, 3 to 6 mismatches). The efficacy failure rates in these subgroups were compared between treatment group using Fisher's exact test. It should be noted, however, that this study was not powered to detect a significant treatment difference in the different subgroups and the total number of patients in some of these subgroups are relatively small. In addition, the effect of the treatment groups on the rate of efficacy failure is assessed by controlling for each of these stratification variables.

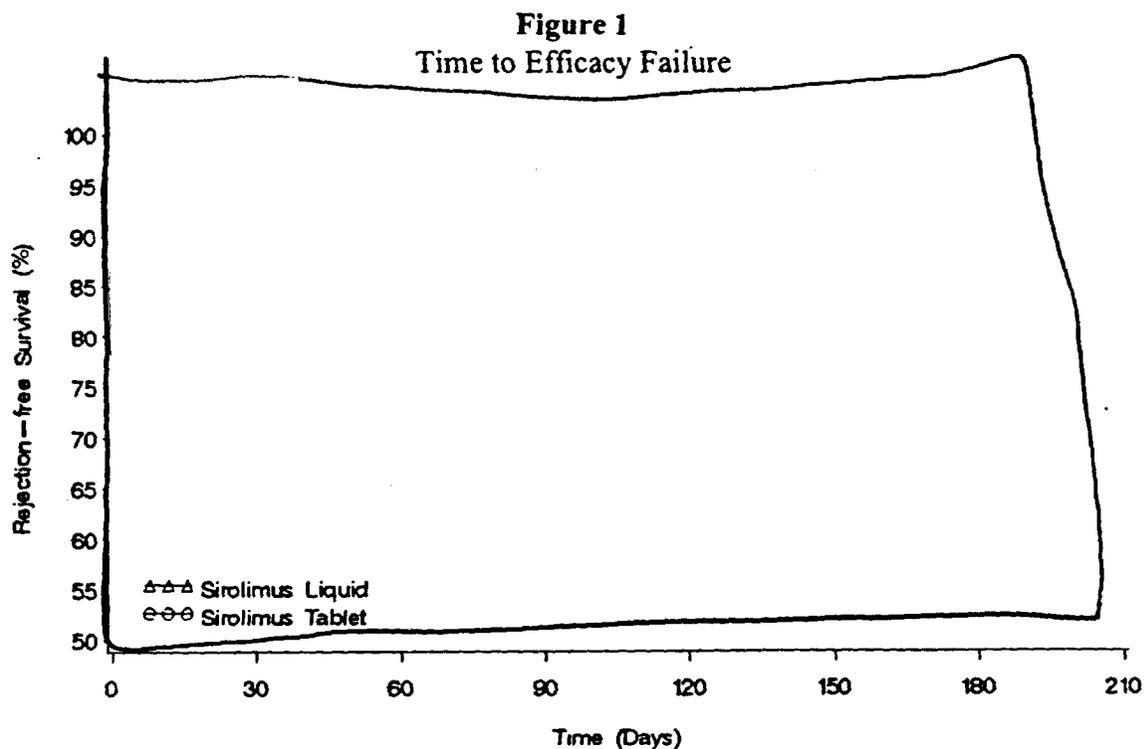
Table 8
Efficacy Failure at 6 months
Selected subgroups

Subgroup	Sirolimus Oral Solution (n=238)	Sirolimus Tablet (n=239)	Stratified Difference (95% CI)
Recipient Race			1.0 (-7.3, 9.3)
Black	18/54 (33.3)	18/59 (30.5)	
Non-black	44/184 (23.9)	47/180 (26.1)	
Recipient Gender			1.2 (-7.1, 9.6)
Female	30/95 (31.6)	27/93 (29.0)	
Male	32/143 (22.4)	38/146 (26.0)	
Donor Source			0.9 (-7.5, 9.3)
Cadaver	47/158 (29.8)	46/170 (27.1)	
Living	15/80 (18.8)	19/69 (27.5)	
Related	8/55 (14.6)	12/49 (24.5)	
Unrelated	7/25 (28.0)	7/20 (35.0)	
Number of HLA mismatches			0.8 (-7.5, 9.1)
0 to 2	8/62 (12.9)	12/55 (21.8)	
3 to 6	54/176 (30.7)	53/184 (28.8)	

The confidence intervals stratified by recipient race, recipient gender, donor source, or number of HLA mismatches are all similar to the confidence interval calculated for the primary analysis of efficacy failure. Thus, the robustness of the results of the primary analysis is supported by these subgroup analyses.

There were no statistically significant differences between the treatment groups for any of the subgroups. The rate of efficacy failure is numerically greater for black patients than non-black patients. Female patients have numerically greater efficacy failure rates than male patients with both sirolimus formulations. The efficacy failure rates are similar for patients who receive a cadaveric or living donor organ with the exception of patients who receive a living donor organ on sirolimus oral solution. Though, not statistically different, these patients have a numerically lower efficacy failure rate. The low number of failures seen by patients receiving living related donor organs causes this numeric difference. Patients with 3 to 6 HLA mismatches have higher efficacy failure rates than patients with 0 to 2 HLA mismatches. Patients with 0 to 2 HLA mismatches and treated with sirolimus oral solution have a numerically smaller efficacy failure rate than those patients treated with the tablet.

Figure 1 shows the time to efficacy failure in each treatment group during the first 6 months of treatment. There is no statistically significant difference (log-rank, $p=0.6221$) between treatment groups in the time to efficacy failure within in the first 6 months after transplantation.



III. Safety Evaluation

The following is a review of the safety data for the Phase III study submitted by the sponsor. The focus of this section will be treatment emergent adverse events (TEAEs) reported as of the data cut off date (July 30, 1999) used for the 4 month safety update. TEAEs were adverse events not present at baseline or events present at baseline that worsened during treatment. In addition, analyses of the clinical laboratory parameters for renal function at 12 months will be presented. The laboratory parameters used to assess renal function are GFR (glomerular filtration rate) and serum creatinine. For a more complete review of the safety data, please refer to the Medical Reviewer's review.

Of the 477 patients enrolled in the study, 457 received at least one dose of sirolimus and were valid for safety: 229 received the oral solution and 228 received the tablet. One or more TEAEs that were not related to infection or malignancy were reported by 228 (99.6%) sirolimus oral solution patients and 227 (99.6%) sirolimus tablet patients. The most commonly occurring TEAEs (reported in at least 20% of patients in any one treatment group) and the accompanying p-values are summarized by treatment group in Table 9. Acne was the only TEAE reported at a significantly higher ($p=0.035$) rate in the sirolimus oral solution group. No TEAEs were reported at a significantly higher rate in the tablet group.

**APPEARS THIS WAY
ON ORIGINAL**

Table 9Number (%) of Patients Reporting Treatment Emergent Adverse Events ($\geq 20\%$)

Body system Event	Sirolimus Oral Solution (n=229)	Sirolimus Tablet (n=228)	p-value
Any adverse experience (1 or more)	228 (99.6)	227 (99.6)	1.00
Body as a whole			
Abdominal pain	55 (24.0)	60 (26.3)	0.591
Asthenia	60 (26.2)	52 (22.8)	0.447
Back Pain	41 (17.9)	51 (22.4)	0.245
Fever	77 (33.6)	65 (28.5)	0.266
Headache	62 (27.1)	56 (24.6)	0.593
Pain	45 (19.7)	58 (25.4)	0.147
Cardiovascular system			
Hypertension	107 (46.7)	106 (46.5)	1.00
Digestive system			
Constipation	50 (21.8)	64 (28.1)	0.131
Diarrhea	72 (31.4)	78 (34.2)	0.551
Nausea	66 (28.8)	60 (26.3)	0.601
Vomiting	63 (27.5)	56 (24.6)	0.523
Hemic and lymphatic system			
Anemia	69 (30.1)	65 (28.5)	0.758
Metabolic and nutritional			
Creatinine increased	76 (33.2)	73 (32.0)	0.842
Edema	46 (20.1)	44 (19.3)	0.906
Hypercholesteremia	83 (36.2)	82 (36.0)	1.00
Hyperlipemia	95 (41.5)	103 (45.2)	0.451
Peripheral edema	151 (65.9)	149 (65.4)	0.922
Musculoskeletal system			
Arthralgia	48 (21.0)	52 (22.8)	0.552
Nervous system			
Tremor	43 (18.8)	60 (26.3)	0.058
Respiratory system			
Dyspnea	48 (21.0)	48 (21.1)	1.00
Skin and appendages			
Acne	63 (27.5)	43 (18.9)	0.035*
Study event associated with miscellaneous factors			
Local reaction to procedure	88 (38.4)	80 (35.1)	0.497

Table 10 lists the frequency of clinically important TEAE by treatment group. TEAE were identified as clinically important based on incidence rates, the relevance to the renal transplantation population, and/or safety data from previous sirolimus trials.

Table 10
Number (%) of Patients Reporting Clinically Important TEAE
Excluding Infection and Malignancy

Body system Event	Sirolimus Oral Solution (n=229)	Sirolimus Tablet (n=228)	p-value
Body as a whole			
Lymphocele	39 (17.0)	28 (12.3)	0.186
Cardiovascular system			
Myocardial infarction	1 (0.4)	4 (1.8)	0.216
Tachycardia	26 (11.4)	26 (11.4)	1.00
Digestive system			
Liver function tests abnormal	14 (6.1)	24 (10.5)	0.093
Pancreatitis	1 (0.4)	2 (0.9)	0.623
Endocrine system			
Diabetes mellitus	14 (6.1)	25 (11.0)	0.068
Hemic and lymphatic system			
Anemia	69 (30.1)	65 (28.5)	0.758
Leukopenia	18 (7.9)	21 (9.2)	0.620
Thrombocytopenia	32 (14.0)	40 (17.5)	0.307
Thrombotic thrombocytopenia purpura (TTP)	5 (2.2)	2 (0.9)	0.450
Metabolic and nutritional			
Hyperkalemia	43 (18.8)	31 (13.6)	0.162
Hypokalemia	28 (12.2)	22 (14.5)	0.495
Musculoskeletal system			
Arthralgia	48 (21.0)	52 (22.8)	0.652
Nervous system			
Hypertonia	15 (6.6)	5 (2.2)	0.037*
Respiratory system			
Epistaxis	5 (2.2)	10 (4.4)	0.202

The only clinically important TEAE that occurred more frequently in one of the two sirolimus treatment groups was hypertonia. Hypertonia was significantly ($p=0.037$) more frequent in the sirolimus oral solution group.

For this discussion, serious and clinically important adverse events are limited to patient death, graft loss, malignancy and life-threatening adverse events because of the number and severity of the adverse events that occur in the population of renal transplant patients. The numbers of patients with these events through the data cutoff date (July 30, 1999) used for the 4 month safety update are summarized in Table 11.

Table 11
Summary of Deaths, Graft Loss, Malignancy, and Life-Threatening Adverse Events

Event	Sirolimus Oral Solution (n=238)	Sirolimus Tablet (n=239)
Death	10 (4.2)	9 (3.8)
Graft Loss (pure)	11 (4.6)	20 (8.4)
Malignancy	6 (2.5)	9 (3.8)
Life-Threatening Adverse Event	20 (8.4)	16 (6.7)

Nineteen patients died as of the study data cut-off date. Seventeen patients died after receiving at least one dose of sirolimus and 2 died after they were randomized into the study but before they received and study drug. There was no significant difference in the incidence of death in the oral solution group compared to the tablet group. In the investigators' opinion, one patient death in the oral solution group was thought to be probably related to the study medication. Five patient deaths were thought to possibly be related to the study medication (2 in the oral solution and 3 in the tablet). In both treatment groups, the most common causes of death were infections and cardiovascular events.

There were 31 patients who experienced a pure graft loss as of the study data cut-off date. In addition, there were 12 deaths with a functioning graft (n=6 in each treatment group). The most common causes of graft loss were death with a functioning graft and acute tubular necrosis (oral solution n=2, tablet n=7). Four patients with graft loss received no study medication (n=2 in each treatment group). The majority of graft loss was prior to month 6, with only 4 graft losses after month 6. According to the sponsor, the numerically higher rate of graft loss in the tablet group was due to factors unrelated to the drug formulation.

Fifteen patients had histologically confirmed malignancy as of the study data cut-off date. The distribution of patients with malignancies was similar between treatment groups.

Thirty-six patients had other non-fatal life-threatening adverse events as of the study data cut-off date. The incidence and nature of these events were not atypical for this patient population.

Laboratory parameters regarding renal function are also of interest. The applicant presented an analysis of GFR and serum creatinine at 12 months for only those patients still on therapy. An analysis that includes only patients still on therapy may eliminate a significant proportion of patients who had poor renal function. An exploratory analysis was performed in an attempt to minimize any bias that may result because of this by using all patients who had a study visit at 12 months whether or not they had discontinued study drug. A study visit window of 337 to 393 days post transplant was used to determine a 12 month visit. Even though 22% of the study population (adjusted by excluding patients who died or had a graft loss) was still not included in the FDA analysis population, it was ensured that this population was representative to the overall study population by demonstrating similarity in the rates of rejection.

Table 22 summarizes the results of mean GFR and serum creatinine at 12 months for the FDA analysis population. Mean GFR and mean serum creatinine at 12 months are not significantly different for the sirolimus oral solution compared to the tablet formulation.

Table 22
Mean (sd) GFR and Serum Creatinine at 12 Months

	Sirolimus Oral Solution	Sirolimus Tablet	P-value
GFR (cc/min)	58.3 (21.1) n=166	58.5 (18.3) n=162	0.9102
Serum Creatinine (mg/dL)	1.91 (0.98) n=165	1.86 (0.79) n=163	0.6161

Reviewer's Note: GFR is assessed using the calculated GFR and is reported in cc/min. Serum creatinine is reported in mg/dL. Higher GFR values and lower serum creatinine values indicate better renal function.

**APPEARS THIS WAY
ON ORIGINAL**

**APPEARS THIS WAY
ON ORIGINAL**

Reviewer's Conclusions (which may be conveyed to the sponsor in the action letter)

1. *The results of the Phase III study demonstrate equivalence, within a delta of 10%, of the Rapamune® Oral Solution and tablet formulations with respect to efficacy failure (first occurrence of biopsy-proven acute rejection, graft loss, or death) during the first 3 and 6 months after transplantation.*
2. *The results fail to demonstrate equivalence of the Rapamune® Oral Solution and tablet formulations in 1 year patient and graft survival rates within a delta of 10%. This criterion, however, was only marginally exceeded (lower limit of the 95% CI was -10.2%).*
3. *The results demonstrate equivalence, within a delta of 5%, of the Rapamune® Oral Solution and tablet formulations in 1 year patient survival rates.*
4. *Overall, the Rapamune® Oral Solution and tablet formulations have similar safety profiles. No clinically important adverse events were reported more frequently in the tablet formulation.*

/S/ 8/21/00

Cheryl Dixon, Ph.D.
Biostatistician, DOB III

/S/ 8/21/00

Concur: Karen Higgins, Sc.D.
Team Leader, DOB III

- cc:
- Archival NDA 21-110 Rapamune
 - HFD-590
 - HFD-590/ Dr. Albrecht
 - HFD-590/ Dr. Cavaille Coll
 - HFD-590/ Dr. Tiernan
 - HFD-590/ Mr. Bacho
 - HFD-725/ Dr. Huque
 - HFD-725/ Dr. Higgins
 - HFD-725/ Dr. Dixon
 - HFD-725/ Chron.

This review contains 14 pages.