CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

50-708/S027
50-709/S021

STATISTICAL REVIEW(S)

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/Serial Number:	50708/S027, 50709/S021
Drug Name:	Prograf (Tacrolimus)
Indication(s):	Prevention of Renal Allograft Rejection
Applicant:	Astellas Pharma US, Inc
Date(s):	Application Dates: February 13, 2006 and March 31, 2008
	Mid-cycle review: January 29, 2009
Review Priority:	Standard
Biometrics Division:	IV
Statistical Reviewer:	Hongling Zhou, Ph.D.
Concurring Reviewers:	Karen Higgins, Sc.D.
Medical Division:	Division of Special Pathogen and Transplant Products
Clinical Team:	Patrick Archdeacon, M.D., Medical Reviewer
	Marc Cavaille Coll, M.D., Medical Team Leader
Project Manager:	June Germain

Keywords:
clinical studies, missing data, imputation/data imputation, open-label, multiple comparisons.

Table of Contents

1. EXECUTIVE SUMMARY

1.1 Conclusions and Recommendations

In the ELiTE-Symphony study, four immunosuppression regimens with mycophenolate mofetil, and corticosteroids were compared for efficacy and safety. A standard regimen of mycophenolate mofetil, cyclosporine and corticosteroids (group A) was compared to regimens with mycophenolate mofetil, daclizumab and corticosteroids as mainstay immunosuppression in combination with low-dose cyclosporine (group B), low-dose tacrolimus (group C) or low-dose sirolimus (group D). In the assessment of this reviewer, group C was shown to be superior to groups A, B and D on the primary efficacy endpoint defined by the sponsor, the glomerular filtration rate (GFR) 12 months after transplantation which was determined from serum creatinine using the Cockcroft-Gault formula. Group C was also shown to be superior compared with other three groups on the endpoint of efficacy failure defined by rate of BPAR (biopsy proven acute rejection), graft loss, death or loss to follow-up. The results seem to be robust with sensitivity analyses. The conclusion drawn from these results is that a regimen consisting of mycophenolate mofetil, daclizumab, corticosteroids and low-dose tacrolimus as planned and used in this study is efficacious in preventing biopsy confirmed acute rejection and leads to adequate kidney function in the first year.

Given this is an investigator driven study, the quality of the data does not meet the usual standards of a trial in support of a product. In this reviewer's opinion, the flaws in the data submitted do not have the potential to affect the conclusion on efficacy. However, lack of reproducible safety data on adverse events such as PTDM (post transplant diabetes mellitus) and infection may be a concern on the validity of the safety conclusion of the study. In addition, as a result that the ELiTE-Symphony study was conducted outside of US, the study population is not representative of the transplant patients in the US. Therefore, generalization of the results to the US transplant population should be made with caution.

1.2 Brief Overview of Clinical Studies

The ELiTE- Symphony study report and data sets were submitted to the Agency by Hoffmann-La Roche in response to a FDA request, which enlisted the help of sponsors of transplant drugs in obtaining information and data regarding concomitant use of tacrolimus and mycophenolate-containing products. In the March 31, 2008 submission, Astellas cited a letter of authorization dated December 19, 2007 from Roche to the FDA. In the letter, Roche authorized FDA to cross reference to NDA 50722, NDA 50723, NDA 50758 and NDA 50759 specifically for and limited to the clinical study reports and case report tabulations/datasets from studies ML16979/Symphony and ML17386/FDCC in support of Astellas's supplemental New Drug Application for the use of the immunosuppressant Prograf® (tacrolimus) plus mycophenolate mofetil as an adjunct therapy for the prophylaxis of organ rejection in kidney transplant patients.

The ELiTE-Symphony study was a prospective, randomized, open-label, multi-center, four parallel arm study. The primary objective of the study was to compare the renal function, as expressed by the glomerular filtration rate (GFR), 12 months after primary renal transplantation in patients receiving four different immunosuppressive treatments. Current standard immunosuppression (group A) was compared to regimens with mycophenolate mofetil, daclizumab and corticosteroids as mainstay immunosuppression in combination with low-dose cyclosporine (group B), low-dose tacrolimus (group C) or low-dose sirolimus (group D). The primary efficacy parameter specified in the protocol was the glomerular filtration rate (GFR), 12 months after transplantation, determined from serum creatinine using the Cockcroft-Gault formula to estimate the creatinine clearance. In our analysis, efficacy failure defined by rate of BPAR with graft loss, death or loss to follow-up was considered co-primary.

1.3 Statistical Issues and Findings

The statistical issues of the study include the method used to impute missing values, adjustment for multiple pairwise comparisons, sample size increase without pre-specification, generalization of the study result to the US population and quality of safety data.

- The primary efficacy parameter specified in the protocol was the glomerular filtration rate (GFR), 12 months after transplantation, determined from serum creatinine using the Cockcroft-Gault formula to estimate the creatinine clearance. The formula also uses variables weight, age and sex to estimate GFR. For the sponsor's primary efficacy analysis in the ITT population, missing values were handled according to the following procedure: the last observation carry forward (LOCF) was applied from the month 3 visit on, including some cases where the subject had died or lost their graft. When the weight was missing, the weight was imputed by LOCF to calculate Cockcroft-Gault GFR. In all other cases, a value of 10 ml/min was imputed. The sponsor's method of handling missing month 12 creatinine values was a concern.

  The sponsor's efficacy analysis of estimated GFR at month 12 after transplantation in the ITT population (with imputation of missing values at month 12 visit) showed that group C had significantly higher GFR values than the other three groups with and without adjustment for multiple comparisons. The analyses using two alternative imputation methods for the ITT population give the same conclusion as the sponsor's ITT analyses. Weaker evidence is seen in the two analyses which remove patients instead of imputing values, in per protocol population (without imputation of missing values) and in the analysis of measured GFR. This raises questions on whether imputing the missing Month 12 creatinine values had an effect on the stronger evidence of superiority of group C in ITT population. Note that one reason for missing data is due to a subject's death or loss of their graft. Since group C had the smallest number of subjects with death or graft loss, it may be a large reason for the lack of significance in the analyses which exclude these patients. It would not seem appropriate to exclude these patients in an analysis; however, even with exclusion of these patients, there is still the trend with group C doing better than the other groups. The superiority of group C

compared with groups A, B and D seems fairly robust in terms of primary endpoint of GFR at month 12 after renal transplantation.

- There were 4 study arms in the Symphony study. This constitutes 6 pairwise comparisons to investigate the difference of outcome between the groups. The sponsor proposed in the protocol to use a Bonferroni-adjusted alpha-level of 0.05/6. However, in the study report, the sponsor reported the results without any adjustment for multiple comparisons. The primary efficacy analyses in this review are presented with 95% CI as well as the 99.2% for the difference of outcome measures between the groups. The 99.2% confidence interval is presented for adjusting multiple comparisons using the Bonferroni method (1-0.05/6=99.2%). The superiority of group C on estimated GFR and efficacy failure compared with other groups was maintained with the adjustment of multiple comparisons.

- The sponsor in protocol amendment 2 dated September 22, 2003 proposed to increase the sample size to a total of 1760 patients from the originally proposed 1300 patients. The reasons for the sample size increase stated in amendment 2 were to improve the precision of the estimates, to increase the power of the planned statistical tests and to increase the probability of reaching the goals of the study, i.e., of identifying statistically significant differences between the four treatments. Since the sponsor did not intend to use the study in support of a product, the study protocol and its amendments were not submitted to the Agency prior to submission of the study report under the Agency's Aug 18, 2007 request. In addition, sample size re-estimation was not specified in the protocol of this open label study. There is generally a concern about increasing sample size during an ongoing study where blindness is not preserved and steps to control type I error are not explained. In order to study the possible impact on the study results due to the increased sample size, efficacy analysis for the first 1300 randomized patients in the ITT population were performed. With the exception of 99.2% CI between group B and group C, the 95% and 99.2% confidence intervals for difference of overall failures between group C and groups A, B and D do not contain 0, Group C is superior to group A and D with and without Bonferroni correction for multiplicity. The 99.2% confidence interval for the difference of efficacy failure between group B and group C was (-0.5%,18.1%), which did not show a statistically significant difference between group B and C after the conservative Bonferroni correction for multiple comparisons, but still showed the trend with group C doing better than group B with respect to the efficacy failure. For the estimated GFR, the conclusion of this group of patients did not change from the ITT population as a whole. Group C is shown to be superior to other treatment groups with respect to the estimated GFR.

- The study population in Elite-Symphony study is not representative of US transplant patients. As shown in Table 3.1.3A, the population studied was primarily white. The numbers of subjects for black, Asian and other races are too small to perform any meaningful subgroup analyses by race. The lack of non-white patients is a concern as to generalization of the results to the US transplant population.

- The ELiTE-Symphony study was not designed as a trial that would be used to support a product. As a result, the quality of the datasets does not meet the usual standards of a

registration trial. Numerous errors were found by this reviewer while examining the data and trying to reproduce the sponsor's results. The flaws in the dataset would not likely affect the conclusions on efficacy because the results are fairly robust. However, the data on some important adverse events are not readily available. For instance, the results reported in the study report on PTDM and infections were not reproducible using the data provided by the sponsor.

2. INTRODUCTION

2.1 Overview

The Symphony study report was submitted to the Agency by Hoffmann-La Roche in response to the DSPTP's letter dated August 18, 2007, which enlisted the help of sponsors of transplant drugs in obtaining information and data regarding concomitant use of tacrolimus and mycophenolate-containing products. In the March 31, 2008 submission, Astellas cited a letter of authorization dated December 19, 2007 from Roche to the FDA. In the letter, Roche authorized FDA to cross reference to NDA 50722, NDA 50723, NDA 50758 and NDA 50759 specifically for and limited to the clinical study reports and case report tabulations/datasets from studies ML16979/Symphony and ML17386/FDCC in support of Astellas's supplemental New Drug Application for the use of the Prograf® (tacrolimus) plus MMF as an adjunct therapy for the prophylaxis of organ rejection in kidney transplant patients.

Prior to August 2007, the division had received a number of requests to study new treatments for the prophylaxis of organ rejection in patients receiving allogeneic kidney or liver transplants using a combination of tacrolimus and mycophenolate as a comparator regimen, which is not a FDA-approved regimen. The labeling for CellCept® and Myfortic®, the two currently marketed mycophenolate preparations, both specifically state in the Indications and Usage section that mycophenolate should be used concomitantly with cyclosporine and corticosteroids, because these were the combinations studied that served as the basis for the approval of these regimens. The labeling does not mention tacrolimus because a regimen containing the combination of tacrolimus and mycophenolate for treating patients with allogeneic kidney or liver transplants has not been approved by FDA.

It has been FDA's interpretation that the safety and efficacy of tacrolimus/mycophenolate combinations in the setting of allogeneic kidney and allogeneic liver transplantation remain uncertain, that a safe and effective regimen in either or both indications has not been identified and approved. However, the division is aware that tacrolimus and mycophenolate combinations are currently the most commonly used initial regimens for the prophylaxis of organ rejection in patients receiving an allogeneic kidney or liver transplant in the United States, based on published literature, registry data and submissions. In addition, there are ongoing or completed clinical studies that have used or are using tacrolimus/mycophenolate combinations as one or more of the study arms. The division believes that these studies may provide important

information for assessing tacrolimus and mycophenolate doses and exposures that are potentially safe and effective in patients with kidney or liver transplants.

The ELiTE-Symphony study has been identified by the division as one of the clinical trials that would be informative and valuable to this effort of investigating concomitant use of tacrolimus and mycophenolate-containing products. The study was sponsored by Prof. H. Ekberg, University Hospital of Lund University in Sweden and Prof. P. Halloran, University of Alberta in Canada, and supported by Hoffmann-La Roche. The principal investigator was Prof. Herink Ekberg. Hoffmann-La Roche assisted the sponsor in local country support, monitoring and serious adverse event reporting, but did not intend to submit the study in support of any application.

2.2 Data Sources

Data sets for the Symphony study were submitted electronically. The full electronic path according to the CDER EDR naming convention is as follows:

\\FDSWA150\NONECTD\N50759\N_000\2007-12-12\crt\datasets\ml16979

The ELite-Symphony study was not designed as a registration trial. Roche noted at the time of its submission that it regarded the Elite-Symphony as an investigator-driven trial. While Roche provided significant financial backing to conduct the study, it did not ensure that the data were compiled with the rigor and precision Roche would demand of a trial intended for submission to the FDA. Although the datasets were adequately documented, the quality of the datasets does not meet the usual standards of a trial in support of a product. Numerous errors were found by this reviewer while examining the data and trying to reproduce the sponsor's results. It is this reviewer's opinion that the flaws in the dataset do not have the potential to affect the conclusions on efficacy because the results are fairly robust. However, lack of data on some important adverse events might be a concern for safety outcomes. For instance, the results reported in the study report on PTDM and infections were not reproducible with the data provided by the sponsor.

3. STATISTICAL EVALUATION

3.1 Evaluation of Efficacy

3.1.1 Introduction

This section presents and discusses the details of sponsor's ELiTE-Symphony study, entitled

Efficacy Limiting Toxicity Elimination (ELiTE) – Symphony: Evaluating the safety and efficacy of mycophenolate mofetil, daclizumab and corticosteroids as mainstay immunosuppression in combination with low-dose cyclosporine, tacrolimus or sirolimus in comparison to current standard immunosuppression (mycophenolate mofetil, cyclosporine and corticosteroids) in renal transplantation

3.1.2 Objectives and Study Design

The Symphony study was a prospective, randomized, open-label, multi-center, four parallel arm study. The patients were randomized to receive one of the following four treatments prior to transplantation:

A: MMF, normal dose cyclosporine, corticosteroids
- Cyclosporine in a standard dosage (3-5mg/kg twice daily PO, IV permitted, adjusted to achieve target trough level of 150–300 ng/ml for the first 3 months and 100–200 ng/ml thereafter)

B: daclizumab, MMF, low-dose cyclosporine, corticosteroids
- Cyclosporine in low dosage (1-2 mg/kg twice daily PO, IV permitted, adjusted to achieve target trough level of 50-100 ng/ml)

C: daclizumab, MMF, low-dose tacrolimus, corticosteroids
- Tacrolimus in low dosage (0.1 mg.kg (in 2 daily doses PO, IV permitted) adjusted to achieve a target trough level of 3–7 ng/ml)

D: daclizumab, MMF, low-dose sirolimus, corticosteroid
- Sirolimus in a low dosage (9 mg/day for 3 days and 3 mg/kg thereafter, PO, adjusted to achieve trough levels of 4-8 ng/ml)

All groups received MMF 1g twice daily PO. First dose of 2 mg/kg Daclizumab was administered IV within 24 hours before the transplantation, followed by four doses of 1 mg/kg every 2 weeks for a total of 5 doses. Corticosteroids were administered according to center practice with a minimum dosage defined in a tapering schedule.

The **primary objective** of the study was to compare renal function, as expressed by the glomerular filtration rate (GFR), 12 months after primary renal transplantation in patients receiving four different immunosuppressive treatments. Current standard immunosuppression (group A) was compared to regimens with mycophenolate mofetil, daclizumab and corticosteroids as mainstay immunosuppression in combination with low-dose cyclosporine (group B), low-dose tacrolimus (group C) or low-dose sirolimus (group D).

The primary efficacy parameter specified in the protocol was the glomerular filtration rate (GFR), 12 months after transplantation, determined from serum creatinine using the Cockcroft-Gault formula to estimate the creatinine clearance. According to the protocol, the effect of the treatment type on the primary efficacy endpoint was to be tested by means of an ANOVA model. The ANOVA model would include the factors center and treatment-by-center interaction. In addition, donor age will be included in the model, as well as baseline covariates assumed to influence the outcomes substantially, and which turn out to be unbalanced between treatment

groups. Nonparametric methods will be employed for data not fulfilling the normality assumptions. If the global hypothesis can be rejected, the differences between the treatment arms will be compared pairwise, with a Bonferroni-adjusted alpha-level of 0.05/6. (The Statistical Analysis Plan changed this to the test of Scheffe).

Note: In the study report, the sponsor presented the results from the nonparametric Kruskal-Wallis test for overall global comparison and Wilcoxon rank-sum test for pairwise comparisons. The sponsor reported the results without adjustment for multiple comparisons. See details in Section 3.1.4.

Secondary objectives were to determine the acute rejection rate at 6 and 12 months, patient and graft survival rates at 12 months and the proportion of patients who experienced treatment failure at 12 months, to evaluate the renal function as expressed by the serum creatinine and the calculated GFR over the course of the study, and to characterize the safety (including hypertension, hyperlipidemia, etc.) of these combinations at 6 and 12 months post-transplant. Biopsy proven acute rejections (BPAR) were determined using the modified BANFF criteria.

Reviewer's comment: In our analysis, BPAR with graft loss, death or loss to follow-up will be considered co-primary.

The analysis populations for the study used in this review were defined as follows:

- **Safety population**: the main population for safety analyses, in which patients were analyzed according to the treatment actually received. It included all patients who received at least one dose of study specific medication from day 1 before until 3 days after transplantation or as starting dose after transplantation, regardless of the treatment group they were randomized to. If treatment received in this time period did not correspond to any of the study specific treatment groups, patients were handled separately and analyzed as group O.
- **Intent-to-Treat population**: it included all patients of the Safety population who underwent transplantation. The ITT population is the primary analysis population of the study. Patients were analyzed in the groups they were randomized to.
- **Per-protocol (PP) population:** a subset of the ITT population of patients without violations of inclusion and exclusion criteria who were not prematurely withdrawn from the study due to: use of additional maintenance immunosuppressive medication not specified in the assigned treatment group (such as switching from cyclosporine to tacrolimus), discontinuation of any of the assigned immunosuppressants for more than 14 consecutive days or 30 cumulative days, necessity for treatment with other investigational drug or other medications prohibited by protocol.

Reviewer's comment: The analyses in this review will be performed mainly in the ITT population.

The parameter considered for the sample size calculation was the acute rejection rate at month 6 (i.e. the proportion of patients who experienced at least one acute rejection in the first 6 months after transplantation). The sample size was originally designed to be able to detect a 10%

difference in acute rejection rate between group D and group A at a 5% significance level and 80% power. Assuming the rejection rates of 30% and 20% in groups D and A respectively, the original sample size proposed by the sponsor was 325 patients per group, i.e., a total of 1300 patients.

The sponsor in protocol amendment 2 dated September 22, 2003 proposed to increase the sample size to 1760 total number of patients, which would provide a statistical power slightly exceeding 90% to detect a 10% difference in acute rejection rate between group D and group A. The reasons for the sample size increase stated in amendment 2 were to improve the precision of the estimates, to increase the power of the planned statistical tests and to increase the probability of reaching the goals of the study, i.e., of identifying statistically significant differences between the four treatments.

Reviewer's comment: Note that this is not a trial intended for support of a product and there were no regulatory procedures to follow in conducting this study. The study protocol and its amendments were not submitted to Agency prior to submission of the study report. In addition, sample size re-estimation was not specified in the protocol. The first enrollment of the study was November 28, 2002, which was 10 months prior to amendment 2. In order to study the possible impact on the study results due to the increased sample size, efficacy results for the first 1300 randomized patients in the ITT population will be presented in section 3.1.4.

3.1.3 Patient Disposition, Demographic and Baseline Characteristics at Entry

A total of 1645 patients were randomized into the four treatment groups of this study, which lasted from November 28, 2002 to December 5, 2005. Patients were recruited in 85 centers from 15 countries: Australia, Austria, Belgium, Brazil (11%), Canada, Czech Republic, Germany (19%), Greece, Israel, Mexico, Poland, Spain (17%), Sweden, Turkey (16%), and United Kingdom. European countries contributed to 76.5% (1285) of the randomized subjects. Germany was the country with the most recruited patients (19.2%) followed by Spain (16.7%), Turkey (15.7%) and Brazil (10.9%). Of the 1645 randomized patients, 1589 subjects received a transplant and the study treatments. Forty patients who did not receive any study specific medication, 13 patients who did not undergo transplantation and 3 subjects whose records were lost were excluded from the ITT population. Among the 1589 subjects in the ITT population, 390 were randomized to Group A, 399 were in group B, 401 were in group C and 399 subjects were randomized to receive treatment in group D. Table 3.1.3A shows the reasons of patients excluded from the per protocol population. There were relatively more patients in group C than in other groups who did not fulfill all inclusion/exclusion criteria. There were fewer patients in group C than in the other 3 groups who were not compliant or withdrew from study medication or were treated with other medication. Since patients who withdrew or treated with other medication were more likely to have encountered problems, this indicates that the per protocol population is more likely biased against group C than other 3 groups.

Table 3.1.3A: Number of Patients (%) excluded from the Per Protocol analysis Population

	A	B	C	D
Randomized patients	410	413	411	411
Did not receive any dose of study medication	16 (4)	8 (2)	7 (2)	9 (2)
Patients records lost by hospital	1 (0)	2 (0)	0	0
Patients not transplanted	3 (1)	4 (1)	3 (1)	3 (1)
Patient did not fulfill all inclusion and exclusion criteria	15 (4)	15 (4)	22 (5)	15 (4)
Patient not compliant or withdrew temporarily (>14 days) or permanently from study medication or treated with other medication	70 (17)	66 (16)	39 (9)	126 (31)
PP analysis population	305 (74)	318 (77)	340 (83)	258 (63)

General demographic information for patients in the ITT population is listed in Table 3.1.3B below. More male patients were included in the study with male to female ratio close to 2:1. The racial distributions are similar across the treatment groups with the vast majority of the patients being Caucasian. The distribution of the ethnic groups in the study is not representative of the US renal transplant patient population since all subjects were recruited from centers outside of US. The mean, median and standard deviation of patient's age and weight were comparable across the treatment groups.

Table 3.1.3B Demographic Information at Entry, ITT population

	Group A N=390	Group B N=399	Group C N=401	Group D N=399
Gender				
Male	243 (62.3%)	265 (66.4%)	264 (65.8%)	266 (66.7%)
Female	147 (37.7%)	134 (33.6%)	137 (34.2%)	133 (33.3%)
Age				
mean±SD	45.9±13.8	47.2±13.5	45.4±14.7	44.9±14.5
median	47.1	47.7	46.2	45.8
Min – Max	18.2 - 72.5	18.4 - 75.8	18.1 - 75.1	18.1 - 74.6
Race				
Caucasian	359(92.1%)	368(92.2%)	377(94.0%)	376(94.2%)
Black	8(2.1%)	9(2.3%)	4(1.0%)	5(1.3%)
Asian	5(1.3%)	3(0.8%)	3(0.8%)	2(0.5%)
other	18(4.6%)	19(4.8%)	17(4.2%)	16(4.0%)
Weight (kg)				
mean±SD	69.7±13.3	70.1±15.0	70.0±15.1	70.8±15.6
median	70	70	68.5	69
Min – Max	36 - 105	38 - 126	37 - 120	40 - 137

3.1.4 Efficacy Results

Estimated GFR at Month 12

The primary efficacy parameter specified in the protocol was the glomerular filtration rate (GFR), 12 months after transplantation, determined from serum creatinine using the Cockcroft-Gault formula to estimate the creatinine clearance. The following is the Cockcroft-Gault formula to calculate the creatinine clearance:

- In men:

$$\text{Creatinine Clearance } [ml/min] = \frac{(140 - \text{Age } [years]) \times \text{Body Weight } [kg]}{\text{Serum Creatinine } [mg/dl] \times 72}$$

- In women:

$$\text{Creatinine Clearance } [ml/min] = \frac{(140 - \text{Age } [years]) \times \text{Body Weight } [kg]}{\text{Serum Creatinine } [mg/dl] \times 72} \times 0.85$$

For the sponsor's primary efficacy analysis in the ITT population, missing values were handled according to the following procedure: the last observation carry forward (LOCF) was applied from the month 3 visit on (including some cases of graft loss and death). When the weight was missing, the weight was imputed by LOCF to calculate Cockcroft-Gault GFR. In all other cases, a value of 10 ml/min was imputed. The sponsor reported that the numbers of values replaced by LOCF method in the treatment groups A, B, C and D were 23, 22, 19 and 27 and the number of imputations by 10 ml/min was 37, 25, 24 and 38 respectively. A detailed inspection of the data found that the numbers reported for LOCF were for missing creatinine values only. The numbers replaced by LOCF for weight at month 12 were 43, 40, 34 and 45 for groups A, B, C and D. The smallest numbers of imputations were made to group C compared to other groups in all three categories.

A note about the dataset: There were three variables in the dataset D.CREA.xpt that seem to flag for missing creatinine LOCF, missing weight LOCF and imputation of 10 respectively. But the variables were not well defined in the definition file.

Table 3.1.4A summarizes the sponsor's primary efficacy results with imputation procedure described above. The 99.2% (=1-0.05/6) confidence interval is listed for adjusting for multiple comparisons using the Bonferroni method. The 95% and 99.2% confidence intervals for the difference of mean GFR between Group C and other groups all excluded 0. This indicates that group C had significantly higher GFR values than the other three groups both with and without adjustment for multiple comparisons. The p values for the pairwise comparisons are p=0.0001 for A vs C, p=0.0011 for B vs C and p=0.0001 for C vs D. Note that these p values are <0.05/6 (= 0.0083), which also indicate statistical significance with Bonferroni correction.

Table 3.1.4A Summary Statistics for Estimated GFR at Month 12 with Sponsor's Imputation, ITT

Group	GFR [ml/min] at Month 12							
	N	MEAN	SD	MIN	Q1	MEDIAN	Q3	MAX
A	390	57.07	25.10	8.60	43.49	57.04	73.67	126.62
B	399	59.39	25.05	4.89	44.89	60.94	75.36	143.25
C	401	65.40	27.03	8.40	49.33	66.18	83.50	160.52
D	399	56.68	26.88	8.40	39.11	57.45	73.63	143.56
Total	1589	59.66	26.25	4.89	44.06	60.59	77.05	160.52

Pairwise Comparisons between the Groups						
	A vs B	A vs C	A vs D	B vs C	B vs D	C vs D
95% CI	(-5.83,1.18)	(-11.98,-4.69)	(-3.25,4.02)	(-9.63,-2.40)	(-0.90,6.32)	(4.98,12.47)
99.2% CI	(-7.05,2.40)	(-13.25,-3.43)	(-4.52,5.29)	(-10.89,-1.14)	(-2.16,7.58)	(3.68,13.77)
P value*	0.1658	0.0001	0.8087	0.0011	0.1171	0.0001

*Wilcoxon rank-sum test

Reviewer's comment: The study report noted that the planned ANOVA models in the protocol were not performed because the normality assumption of the data was rejected even after standard transformations and therefore the validity of the model was highly questionable. This reviewer verified the above claim and considers the use of nonparametric tests in this case appropriate. In addition, the conclusion from the ANOVA model is the same as from the nonparametric tests.

The p value is <0.0001 based on nonparametric Kruskal-Wallis test for overall comparison. The p values for pairwise comparisons are based on Wilcoxon rank-sum test. The results listed in Table 3.1.4A were consistent with the results presented in the study report.

In verifying sponsor's results, this reviewer found that the sponsor also imputed age of the recipient in addition to weight and creatinine values when calculating the estimated GFR using Crockcroft-Gault formula shown above. For instance, if the creatinine at month 12 was LOCF from the month 3 visit, the subject's age at month 3 was used in the formula instead of the age at month 12. This resulted a slightly higher estimated GFR for all subjects whose missing values were LOCF, since the age of a patient at month 3 is a smaller number than the patient's real age at month 12. The effect of age LOCF seems balanced across the groups and is too small to change the conclusion on estimated GFR. In order to verify the sponsor's results, this reviewer used the estimated GFR provided by the sponsor, but noted that age was also imputed.

With the sponsor's imputation method, there were 38 patients (10 in group A, 10 in group B, 5 in group C and 13 in group D) who died and/or had graft loss and who were imputed with a >10 GFR by the LOCF method. We consider that all deaths/graft losses (41, 27, 23 and 42 in groups A, B, C and D) should be imputed with a small number for GFR. An alternative imputation

method was performed as follows: 1) all death/graft loss were imputed with GFR of 10 ml/min; 2) LOCF from month 3 on for missing creatinine, weight will also be imputed if missing; 3) patients whose last recorded creatinine values were prior to month 3 visit were also imputed by 10 for estimated GFR at month 12. With this imputation method the numbers of creatinine values at month 12 replaced by LOCF method were for group A 11 (1 from month 3 visit, 4 from month 6 visit and 6 from month 9 visit), for group B 12 (6 from month 3 visit, 3 from month 6 visit and 3 from month 9 visit), for group C 15 (3 from month 3 visit, 7 from month 6 visit and 5 from month 9 visit) and for group D 19 (1 from month 3 visit, 12 from month 6 visit and 6 from month 9 visit). 35 patients (10 in group A, 9 in group B, 7 in group C and 9 in group D) were excluded from this analysis.

The results with this imputation method are listed in Table 3.1.4B. The 99.2% confidence interval is listed for adjusting multiple comparisons using the Bonferroni method. The 95% and 99.2% confidence intervals for the difference of mean GFR between Group C and other groups all excluded 0. This indicates that group C had significantly higher GFR values than the other three groups with and without adjustment for multiple comparisons. The p values for the pairwise comparisons are $p<0.0001$ for A vs C, $p=0.0010$ for B vs C and $p<0.0001$ for C vs D. Note that these p values are $<0.05/6$ (0.0083), which also indicate statistical significance with Bonferroni correction for multiple pairwise comparisons. The conclusion for superiority of group C compared with other three groups regarding GFR did not change with this alternative imputation method.

Table 3.1.4B Summary Statistics for GFR at Month 12 with Alternative Imputation

Group	GFR [ml/min] at Month 12							
	N	MEAN	SD	MIN	Q1	MEDIAN	Q3	MAX
A	390	56.45	25.79	8.60	43.39	56.90	73.60	126.62
B	399	58.94	25.65	4.89	43.75	60.94	75.36	143.25
C	401	65.14	27.42	8.40	49.33	66.18	83.50	160.52
D	399	56.17	27.37	10.00	38.66	57.28	73.63	143.56
Total	1589	59.20	26.80	4.89	43.56	60.48	77.04	160.52
Pairwise Comparisons between the Groups								
	A vs B	A vs C	A vs D	B vs C	B vs D	C vs D		
95% CI	(-6.08,1.11)	(-12.41,-4.97)	(-3.43,4.01)	(-9.89,-2.52)	(-0.92,6.46)	(5.17,12.78)		
99.2% CI[#]	(-7.33,2.36)	(-13.7,-3.68)	(-4.73,5.30)	(-11.18,-1.24)	(-2.2,7.74)	(3.85,14.10)		
P value[*]	0.1499	<0.0001	0.8115	0.0010	0.1088	<0.0001		

[*] Wilcoxon rank-sum test

[#] Adjusted by multiple comparisons, 1-0.05/6=99.2%

Table 3.1.4C lists the results for estimated GFR at Month 12 for the per protocol population where all subjects without a 12 month estimated GFR were excluded from the calculation including most subjects who died or had a graft loss. Note that the 95% CI for the difference of mean GFR between group C and other three groups still excluded 0. However, all three corresponding 99.2% CI contain 0 ((-9.55, 0.29) for A – C, (-9.32, 0.28) for B – C and (-0.80, 9.99) for C – D). The 99.2% confidence intervals are listed for adjusting for multiple

comparisons using the Bonferroni method. This indicates that the differences of GFR between group C and other groups are not statistically significant after adjusting for multiple pairwise comparisons using Bonferroni method. The p values for the pairwise comparisons using Wilcoxon rank-sum test are 0.0173 for A vs C, 0.0175 for B vs C and 0.0199 for C vs D, which are all >0.0083=0.05/6. Since no imputations were performed for the per protocol population, the results for per protocol population raise questions on whether imputing the missing Month 12 creatinine values had an effect on the stronger evidence of superiority of group C in ITT population.

Table 3.1.4C Summary Statistics for GFR at Month 12 without Imputation, Per Protocol Population

Group	GFR [ml/min] at Month 12								
	N	NMiss	MEAN	SD	MIN	Q1	MEDIAN	Q3	MAX
A	252	53	66.16	20.04	24.94	51.01	65.49	80.05	126.62
B	282	36	66.27	20.47	8.74	52.95	65.67	77.88	143.25
C	296	44	70.79	22.94	9.21	54.55	70.18	85.69	160.52
D	214	44	66.20	22.34	19.62	51.21	65.05	79.64	143.56
Total	1044	177	66.30	21.55	8.74	52.41	66.30	81.28	160.52

Pairwise Comparisons between the Groups						
	A vs B	A vs C	A vs D	B vs C	B vs D	C vs D
95% CI	(-3.56,3.34)	(-8.28,-0.99)	(-3.90,3.82)	(-8.08,-0.97)	(-3.72,3.86)	(0.59,8.59)
99.2% CI[#]	(-4.77,4.55)	(-9.55,0.29)	(-5.25,5.17)	(-9.32,0.28)	(-5.05,5.19)	(-0.80,9.99)
P value[*]	0.7676	0.0173	0.9137	0.0175	0.7552	0.0199

[*] Wilcoxon rank-sum test

[#] Adjusted by multiple comparisons, 1-0.05/6=99.2%

To further investigate the effect of imputation methods on the outcome of estimated GFR, a method of imputing all missing estimated GFR with a value of 40 was suggested at the mid-cycle review meeting. This idea is based on the information from prior analyses that most patients who failed treatment had missing outcome values and that there were more failed patients in the other 3 groups than group C. Therefore an imputation of 40 for the missing estimated GFR would favor other groups more than group C. Table 3.1.4D summarizes the results for this imputation method. Note that the mean estimated GFR values for groups A, C and D were higher than the corresponding values in Table 3.1.4A with the sponsor's imputation, and the mean estimated GFR for group C with this imputation is lower than the corresponding value with the sponsor's imputation shown in Table 3.1.4A. The 95% and 99.2% CI for difference of mean GFR between group C and groups A, B and D all excluded 0. With this imputation method, group C is still shown to be superior to groups A, B and D in terms of estimated GFR at month 12 post-transplantation.

Table 3.1.4D Summary Statistics for GFR at Month 12 with Another Imputation Method

Group	GFR [ml/min] at Month 12							
	N	MEAN	SD	MIN	Q1	MEDIAN	Q3	MAX
A	390	58.74	20.55	8.60	40	54.40	72.26	126.62
B	399	60.28	21.81	8.74	40	59.71	74.25	143.25
C	401	64.80	23.78	9.21	42.82	62.57	82.13	160.52
D	399	58.41	22.14	11.50	40	54.59	72.93	143.56
Total	1589	60.57	22.24	8.60	40	57.60	75.61	160.52

Pairwise Comparisons between the Groups						
	A vs B	A vs C	A vs D	B vs C	B vs D	C vs D
95% CI	(-4.51,1.42)	(-9.16,-2.95)	(-2.65,3.32)	(-7.68,-1.35)	(-1.18,4.93)	(3.20, 9.58)
99.2% CI[#]	(-5.54,2.45)	(-10.25,-1.87)	(-3.69,4.36)	(-8.79,-0.25)	(-2.24,5.99)	(2.09,10.69)
P value[*]	0.2746	0.0004	0.5309	0.0092	0.1181	<0.0001

[*] Wilcoxon rank-sum test

[#] Adjusted by multiple comparisons, 1-0.05/6=99.2%

The primary efficacy endpoint of the study was GFR estimated by Cockcroft-Gault formula. Ideally, GFR measured at a lab is a more accurate measure of kidney function. Estimated GFR was used mostly for cost consideration. Out of 1589 ITT subjects, 508 (32.0%) did not have measured GFR at Month 12 in the dataset. There were also no records of previous measurements of GFR for any of the subjects in the study. Table 3.1.4E summarizes the results for measured GFR for subjects in the ITT population who had measured GFR at Month 12. The second column in the table is the number of subjects whose GFR were measured at Month 12, and the third column is the number of subjects whose GFR were missing (not measured) in the dataset LAB.xpt. Note that the majority of subjects who died or had a graft loss had missing GFR at month 12. The Kruskal-Wallis test for overall comparison has a p-value of 0.0580, which indicated there was no statistically significant difference among the treatment groups in the subgroup of patients whose GFRs were measured at Month 12. The pairwise comparisons were listed for comparison purposes. The differences between groups were smaller with measured GFR than the estimated GFR by Cockcroft-Gault method and the superiority of group C is not evident in the subgroup of the ITT population who had measured GFR at Month 12.

Table 3.1.4E Summary Statistics for Measured GFR, missing values excluded

Group	Measured GFR at Month 12								
	N	NMiss	MEAN	SD	MIN	Q1	MEDIAN	Q3	MAX
A	263	127	61.24	25.31	0.92	43	59	77.67	163
B	288	111	63.03	26.71	0.31	46	61	77.60	180
C	262	139	66.57	27.89	0.52	50	64.9	81.54	184.38
D	268	131	61.47	27.46	0.60	43	60	76.96	174
Total	1081	508	63.07	26.90	0.31	45.01	61	78	184.38

Pairwise Comparisons between the Groups						
	A vs B	A vs C	A vs D	B vs C	B vs D	C vs D
95% CI	(-6.16,2.57)	(-9.89,-0.76)	(-4.73,4.27)	(-8.11,1.04)	(-2.95,6.08)	(0.38,9.82)
99.2% CI[#]	(-7.68,4.10)	(-11.49,0.83)	(-6.31,5.85)	(-9.71,2.63)	(-4.53,7.65)	(-1.27,11.47)
P value[*]	0.3341	0.0185	0.9619	0.1410	0.3368	0.0191

[*] Wilcoxon rank-sum test

[#] Adjusted by multiple comparisons, 1-0.05/6=99.2%

Reviewer's comment: The sponsor's efficacy analysis of estimated GFR at month 12 after transplantation in the ITT population (with imputation of missing values at month 12 visit) showed that group C had significantly higher GFR values than the other three groups with and without adjustment for multiple comparisons. The analyses using two alternative imputation methods for the ITT population give the same conclusion as the sponsor's ITT analyses. Weaker evidence is seen in the two analyses which remove patients instead of imputing values, in per protocol population (without imputation of missing values) and in the analysis of measured GFR. This raises questions on whether imputing the missing Month 12 creatinine values had an effect on the stronger evidence of superiority of group C in ITT population. Note that one reason for missing data is due to a subject's death or loss of their graft. Since group C had the smallest number of subject with death or graft loss (see below), it may be a large reason for the lack of significance in the analyses which exclude these patients. Missing data due to a subject having a graft loss or death is very different than due to a subject's GFR being merely unknown. Some might argue that GFR is not missing for those who died or lost their graft since their outcome is known. Therefore, it would not seem appropriate to exclude these patients in an analysis; however, even with exclusion of these patients, there is still the trend with group C doing better than the other groups. The superiority of group C compared with groups A, B and D seems fairly robust in terms of primary endpoint of GFR at month 12 after renal transplantation.

BPAR/graft loss/death/loss to follow-up

A commonly used primary endpoint in transplant clinical trials is efficacy failure defined by biopsy proven acute rejection (BPAR), graft loss, death or loss to follow-up. The sponsor of the Symphony study did not present the result of this endpoint in the study report. Table 3.1.4F lists the overall failure rates for this endpoint as well as failure rate for each component of the endpoint by treatment groups in the ITT population. Group C had the lowest overall failure rate

compared with groups A, B and D. The 95% and 99.2% confidence intervals for difference of overall failures between the treatment groups are listed in Table 3.1.4G. Group C had significantly lower rate of overall failure and lower rate of BPAR than groups A, B and D. There was no statistically significant difference of mortality rates between the treatment groups. For graft loss excluding death, the 95% CIs for difference of rate between group A and C, and between group C and D do not contain 0, but both the corresponding 99.2% CI contain 0. The rates for graft loss excluding death, death and LTFU combined were 11.8% (46/390), 8.5% (34/399), 7.0% (28/401) and 12.0% (48/399) for Groups A, B, C, and D. The rates for graft loss/death for each group are presented in Table 3.1.4H.

Table 3.1.4F Rate of BPAR, Graft Loss, Death or Loss to follow-up

	A N=390	B N=399	C N=401	D N=399
Overall Failure	141 (36.2)	126 (31.6)	82 (20.4)	185 (46.4)
BPAR	113 (29.0)	106 (26.6)	60 (15.0)	152 (38.1)
Graft loss excluding death	28 (7.2)	20 (5.0)	12 (3.0)	30 (7.5)
Mortality	13 (3.3)	7 (1.8)	11 (2.7)	12 (3.0)
Loss to follow-up	5 (1.3)	7 (1.8)	5 (1.3)	6 (1.5)

Table 3.1.4G Pairwise Comparisons of Overall Failure, BPAR, Graft Loss and Mortality

	A - B	A - C	A – D	B – C	B – D	C - D
Overall Failure						
95% CI	(-2.3,11.4)	(9.3,22.1)	(-17.3,-3.1)	(4.8,17.4)	(-21.7,-7.8)	(-32.5,-19.4)
99.2% CI	(-4.6,13.8)	(7.1,24.3)	(-19.7,-0.7)	(2.7,19.5)	(-24.1.,-5.5)	(-34.7,-17.2)
P value	0.1996	<0.0001	0.0045	0.0005	<0.0001	<0.0001
BPAR						
95% CI	(-4.1, 8.9)	(8.1, 20.0)	(-15.9,-2.3)	(5.8, 17.4)	(-18.2,-4.8)	(-29.3,-17.0)
99.2% CI	(-6.3, 11.1)	(6.0, 22.0)	(-18.2, 0.0)	(3.8, 19.4)	(-20.5,-2.6)	(-31.4,-14.9)
P value	0.4993	<0.0001	0.0084	<0.0001	0.0007	<0.0001
Graft loss excluding death						
95% CI	(-1.4, 5.8)	(0.9, 7.5)	(-4.2, 3.6)	(-0.9, 5.0)	(-6.1, 1.1)	(-7.9, -1.2)
99.2% CI	(-2.6, 6.9)	(-0.2, 8.6)	(-5.5, 4.8)	(-1.9, 5.9)	(-7.3, 2.3)	(-8.9, 0.1)
P value	0.2609	0.0116	0.9632	0.2015	0.1886	0.0067
Mortality						
95% CI	(-0.9, 4.0)	(-2.1, 3.2)	(-2.4, 3.0)	(-3.3, 1.3)	(-3.6, 1.1)	(-2.8, 2.3)
99.2% CI	(-1.6, 4.8)	(-2.9, 4.1)	(-3.2, 3.9)	(-4.0, 2.0)	(-4.4, 1.9)	(-3.6, 3.1)
P value	0.2363	0.7822	0.9538	0.4811	0.3530	0.9903

The calculations here were based on the datasets provided by the sponsor. The study report states that in the ITT population, 133 patients lost their graft or died during the first post-

transplantation year, 41, 27, 23 and 42 in groups A, B, C and D, respectively. There were 15 reported GL/death after month 12, which were included in the last row of Table 3.1.4H below.

Table 3.1.4H Number of Graft Loss/Death (and rates by Kaplan-Meier Estimates)

	Group A N=390	Group B N=399	Group C N=401	Group D N=399
Week 4	16 (4.10%)	11 (2.77%)	11 (2.75%)	19 (4.77%)
Week 8	23 (5.92%)	13 (3.28%)	14 (3.51%)	22 (5.53%)
Month 3	30 (7.74%)	15 (3.78%)	17 (4.27%)	28 (7.06%)
Month 6	37 (9.59%)	20 (5.07%)	20 (5.03%)	32 (8.08%)
Month 9	37 (9.59%)	23 (5.84%)	21 (5.29%)	37 (9.39%)
Month 12	41 (10.66%)	27 (6.89%)	23 (5.83%)	42 (10.73%)
Total # of GL/Death	45	31	27	45

Table 3.1.4H lists the number of graft loss/death and rates at different time points estimated by Kaplan-Meier survival curve based on time to GL/death. Group B and group C had similar rates of graft loss/death over the time. Group A and group D were similar and were higher than Groups B and C with respect to rate of graft loss/death. The p values for pairwise comparisons by Log-rank test were 0.0195 for A vs C, 0.7217 for B vs C and 0.0253 for C vs D. None of the pairwise comparisons would be considered statistically significant with the adjustment for multiple comparisons (p value>0.0083=0.05/6).

Table 3.1.4I Number of BPAR by Groups (Rates by Kaplan-Meier Estimates)

Time	Group A N=390	Group B N=399	Group C N=401	Group D N=399
Week 4	75 (19.54%)	57 (14.44%)	35 (8.82%)	54 (13.80%)
Week 8	89 (23.28%)	65 (16.50%)	41 (10.37%)	79 (20.37%)
Month 3	96 (25.19%)	77 (19.59%)	44 (11.16%)	107 (27.87%)
Month 6	107 (28.25%)	97 (24.79%)	55 (14.07%)	145 (38.17%)
Month 9	110 (29.10%)	104 (26.63%)	58 (14.87%)	150 (39.56%)
Month 12	113 (29.98%)	106 (27.18%)	60 (15.42%)	152 (40.12%)

Table 3.1.4I lists the number of BPAR and rates at different time points estimated by Kaplan-Meier survival curve based on time to first biopsy proven acute rejection (BPAR). Group C had the lowest rate of BPAR at all time points. The p values for pairwise comparisons by Log-rank test were <0.0001 for A vs C, B vs C and C vs D. Group C had statistically significantly lower rate of BPAR than groups A, B and D even with the adjustment for multiplicity.

Table 3.1.4J Distribution of Number of BPAR Per Subjects

	A	B	C	D
Number of subjects with rejection	113	106	60	152
1 rejection	89	80	50	93
2 rejections	17	20	8	42
3 rejections	5	4	2	11
4 rejections	2	2	0	3
5 rejections	0	0	0	2
7 rejections	0	0	0	1

Table 3.1.4K Types of BPAR: Total Number (multiple per subject)

	A	B	C	D
Total number of rejection (inc. borderline)	136	142	70	226
Borderline	25	23	15	26
Grade I (mild acute)	62	57	26	131
Grade I (mild chronic transplant nephropathy)	3	5	3	2
Grade IIA (moderate acute)	31	28	13	43
Grade II B (moderate acute)	7	15	5	16
Grade II (moderate chronic transplant nephropathy)	0	5	4	1
Grade III (severe acute)	8	8	3	6
Grade III (severe chronic transplant nephropathy)	0	1	1	1

Table 3.1.4J lists the distribution of number of subjects who had BPAR by the number of rejections they had. The numbers were obtained from the dataset D_ACREJ.xpt. The proportions of the number of rejections seem to be similar across the 4 treatment groups. Table 3.1.4K shows the number of rejections by BANFF grade I-III including borderlines (based on dataset ACREJ.xpt). There are no concerning trends shown among the groups.

Reviewer's Comment: The data from D_ACREJ.xpt and ACREJ.xpt do not totally match. It is not clear which dataset provided more accurate information. However, the trends are not concerning based on the numbers obtained from the two datasets.

Efficacy outcome in the first 1300 randomized subjects in ITT population

The original sample size proposed by the sponsor was 325 patients per group, i.e., a total of 1300 patients. The sponsor in protocol amendment 2 dated September 22, 2003 proposed to increase the sample size to a total of 1760 patients, which would provide a statistical power slightly exceeding 90% to detect a 10% difference in acute rejection rates between group D and group A. The reasons for the sample size increase stated in amendment 2 were to improve the precision of the estimates, to increase the power of the planned statistical tests and to increase the probability of reaching the goals of the study, i.e., of identifying statistically significant differences between the four treatments. Given that this is an investigator-driven study, the division did not have knowledge of the protocol and its amendments prior to the submission of the study report. In

addition, sample size re-estimation was not pre-specified in the protocol. The first enrollment of the study was November 28, 2002, which was 10 months prior to amendment 2. In order to study the possible impact on the study results due to the increased sample size, efficacy analyses for the first 1300 randomized patients in the ITT population were performed.

Table 3.1.4L Summary Statistics for Estimated GFR at Month 12 with Imputation, first 1300 subjects in ITT population

Group	GFR [ml/min] at Month 12							
	N	MEAN	SD	MIN	Q1	MEDIAN	Q3	MAX
A	319	57.37	25.10	8.60	43.39	56.32	73.78	126.62
B	326	59.22	24.56	4.89	45.25	61.32	75.07	134.75
C	329	64.92	26.86	8.40	49.12	65.14	83.89	160.52
D	326	56.41	26.65	10.00	39.77	57.32	73.40	133.93
Total	1300	59.50	26.00	4.89	44.08	60.48	77.13	160.52

Pairwise comparisons between the groups						
	A vs B	A vs C	A vs D	B vs C	B vs D	C vs D
95% CI	(-5.68,2.00)	(-11.55,-3.53)	(-3.04,4.97)	(-9.65,-1.75)	(-1.13,6.75)	(4.40,12.61)
99.2% CI#	(-7.02,3.34)	(-12.95,-2.13)	(-4.43,6.37)	(-11.03,-0.37)	(-2.51,8.12)	(2.97,14.04)
P value*	0.2393	0.0003	0.6867	0.0067	0.1212	<0.0001

Table 3.1.4L presents summary statistics for estimated GFR at month 12 by Cockcroft-Gault in the first 1300 patients in the ITT population. The conclusion for this group of patients did not change from the ITT population as a whole. The Kruskal-Wallis test for overall comparison resulted a p=0.0002. The pairwise comparisons by Wilcoxon rank-sum test resulted p values of 0.0003 for groups A vs C, 0.0067 for groups B vs C and <0.0001 for groups C vs D. The 95% and 99.2% CI excluded 0 for comparisons between groups A vs C, B vs C and C vs D.

Table 3.1.4M Rate of BPAR, Graft Loss, Death or Loss to Follow-up, first 1300 Subjects in ITT Population

Group A N=319	Group B N=326	Group C N=329	Group D N=326
113 (35.4%)	98 (30.1%)	70 (21.3%)	153 (46.9%)

Pairwise comparisons between groups						
	A-B	A-C	A-D	B-C	B-D	C-D
95% CI	(-2.2, 12.9)	(7.0, 21.3)	(-19.4,-3.7)	(1.8,15.7)	(-24.5,-9.2)	(-33.0,-18.4)
99.2% CI	(-4.7, 15.5)	(4.6, 23.7)	(-22.0,-1.0)	(-0.5,18.1)	(-27.1,-6.6)	(-35.4,-15.9)
P value	0.1716	<0.0001	0.0039	0.0130	<0.0001	<0.0001

Table 3.1.4M shows the outcome for the rate of BPAR, graft loss, death or loss to follow-up for the subgroup of first 1300 subjects in ITT population. With the exception of 99.2% CI between group B and group C, the 95% and 99.2% confidence intervals for difference of overall failures between group C and groups A, B and D do not contain 0, Group C is superior to group A and D

with and without Bonferroni correction for multiplicity. The 99.2% confidence interval for the difference of efficacy failure between group B and group C was (-0.5%,18.1%), which did not show a statistically significant difference between group B and C after the conservative Bonferroni correction for multiple comparisons, but still showed the trend with group C doing better than group B with respect to efficacy failure.

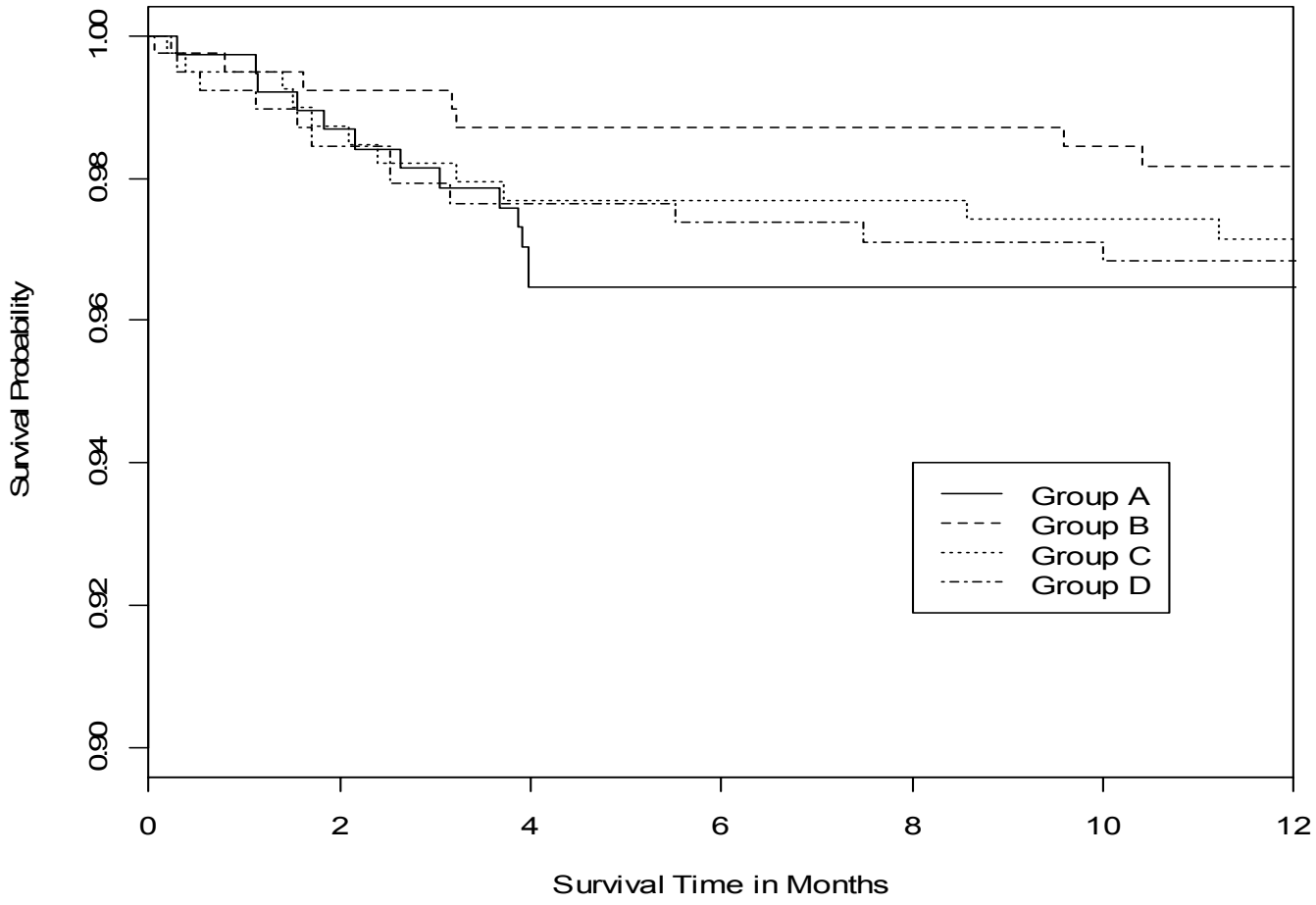
3.2 Evaluation of Safety

A total of 49 deaths (14, 8, 12 and 15 for groups A, B, C and D) were reported for the patients who participated in the Symphony study. 43 deaths occurred by month 12 after renal transplantation, and 6 occurred after 12 months. Table 3.2A lists the number of deaths and mortality rates at different time points estimated by Kaplan-Meier survival curve (shown in Figure 1) based on time to death from transplantation. There is no statistically significant difference in mortality rates among the treatment groups (p=0.3448, log-rank test). Note that the y-axis in Figure 1 is truncated to 0.90 to 1.

Table 3.2A Number of Deaths by Groups (Mortality Rates by Kaplan-Meier Estimates)

	Group A	Group B	Group C	Group D
Week 4	1 (0.26%)	2 (0.51%)	2 (0.50%)	3 (0.77%)
Week 8	5 (1.32%)	3 (0.77%)	5 (1.27%)	6 (1.56%)
Month 3	7 (1.87%)	3 (0.77%)	7 (1.79%)	8 (2.08%)
Month 6	13 (3.53%)	5 (1.29%)	9 (2.31%)	10 (2.62%)
Month 9	13 (3.53%)	5 (1.29%)	10 (2.58%)	11 (2.89%)
Month 12	13 (3.53%)	7 (1.83%)	11 (2.85%)	12 (3.17%)
Total # of Death	14	8	12	15

Figure 1: Survival Curve of the Study by Groups



One of the adverse events of special interest is PTDM (post transplant diabetes mellitus). However, there were no variables in any of the datasets identifying this specific adverse event. There was a field in D_ADEV describing adverse events and there were a number of such descriptions containing the word Diabetes. It is not clear which can be defined as PTDM. The incidence of PTDM was an issue raised by some of the readers of the sponsor's published article in NEJM (Ekberg H, Tedesco-Silva H, Demirbas A, et al. Reduced exposure to calcineurin inhibitors in renal transplantation. N Engl J Med 2007; 357:2562-75). The comments noted that new-onset diabetes after transplantation is considered an important adverse effect after renal transplantation. Therefore, readers need to know how new-onset diabetes after transplantation was defined in the present trial. The question was asked whether the definition was based on the American Diabetes Association–World Health Organization guidelines as suggested in 2003. The authors responded that the Symphony study was designed in 2002, when there was no established consensus definition for this condition, and before publication of the American Diabetes Association–World Health Organization guidelines. The authors stated that new-onset

diabetes after transplantation in the ELITE–Symphony study was defined as adverse-event reports that included the term "diabetes" or "hyperglycemia", and that only a small fraction of patients became insulin-dependent.

The sponsor reported in the study report that after 12 months the rates of PTDM (post transplant diabetes mellitus) based on Kaplan Meier estimates were 6.4% in group A, 4.7% in group B, 10.6% in group C and 7.8% in group D, with statistical significance among the four groups in an overall log-rank test (p=0.0182). It was noted in the study report, for PTDM, the preferred terms included Diabetes Mellitus, Diabetes mellitus (non-)insulin-dependent and Diabetes mellitus inadequate control. Based on dataset D_ADEV.xpt, 6.8% (26/384) subjects in group A, 5.6% (23/408) in group B, 11.2% (45/403) in group C, 8.4% (32/380) in group D and 3.7% (1/27) in group O (p=0.0370) had the adverse events with the preferred terms noted above. Additionally, if the preferred terms Glucose tolerance impaired and Hyperglycemia were also included (as mentioned in the response of the authors to the readers of NEJM), the rates of PTDM for groups A, B, C, D and O were 11.5% (44/384), 8.3% (34/408), 16.1% (65/403), 13.4% (51/380) and 3.7% (1/27) (p=0.0071).

The sponsor stated in the study report that 2132 infections in 931 patients (58.1%) were reported. The highest rate was observed in group A followed by group D with groups B and C exhibiting the lowest rates; 241 patients (62.8%) in group A, 230 patients (56.4%) in group B, 229 patients (56.8%) in group C, 226 patients (59.5%) in group D and 5 patients (18.5%) in group O. Group O was defined as the group of subjects in the safety population whose treatment did not correspond to any of the study specific treatment groups. The numbers quoted above for infections do not match the information obtained from dataset D_ADEV.xpt submitted by the sponsor. However, the overall conclusions are consistent. According to the dataset, there were 2062 infections in 897 patients in the safety population. The rates of infection for groups A, B, C, D and O were 58.1% (223/384), 55.4% (226/408), 55.8% (225/403), 57.1% (217/380) and 22.2% (6/27).

Comment: Please see section 4.1 for adverse events infection and PTDM by gender. The reported numbers for both PTDM and infection in the study report were not reproducible with the dataset provided by the sponsor.

For overview of the safety results, please see details in the clinical review by the medical officer.

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Race and Age

As noted in Dr. Patrick Archdeacon's clinical review, subpopulations of general interest in renal transplant clinical studies include those related to gender, age, race, diabetes, and hepatitis C status. However, due to the demographics and the conduct of the Symphony-ELiTE study, meaningful subgroup analyses could not be conducted for all the usual subpopulations:

Caucasians comprised the only substantial group according to race, instances of post-transplant diabetes were both rare and also poorly documented, and infection with hepatitis C constituted grounds for exclusion from the study. Subgroup analyses were conducted according to gender and age.

Table 4.1A Efficacy Outcome by Gender

	Group A N=390	Group B N=399	Group C N=401	Group D N=399	P value
Estimated GFR at Month 12					
Male mean ± SD median	N=243 60.15±24.38 59.26	N=265 61.35±24.96 63.32	N=264 68.46±27.24 69.50	N=266 58.09±27.46 59.61	<0.0001
Female mean ± SD median	N=147 51.97±25.53 52.54	N=134 55.50±24.86 55.72	N=137 59.51±25.72 59.97	N=133 53.86±25.54 53.46	0.0892
All patients mean ± SD median	N=390 57.07±25.10 57.04	N=399 59.39±25.05 60.94	N=401 65.40±27.03 66.18	N=399 56.68±26.88 57.45	<0.0001
Rate of BPAR/GL/Death/LTFU					
Male N=1038	86/243 (35.4%)	89/265 (33.6%)	52/264 (19.7%)	136/266 (51.1%)	<0.0001
Female N=551	55/147 (37.4%)	37/134 (27.6%)	30/137 (21.9%)	49/133 (36.8%)	0.0119
All patients N=1589	141/390 (36.2%)	126/399 (31.6%)	82/401 (20.5%)	185/399 (46.4%)	<0.0001
Rate of BPAR					
Male N=1038	74/243 (30.5%)	77/265 (29.1%)	39/264 (14.8%)	111/266 (41.7%)	<0.0001
Female N=551	39/147 (26.5%)	29/134 (21.6%)	21/137 (15.3%)	41/133 (30.8%)	0.0182
Rate of Graft Loss excluding Death					
Male N=1038	16/243 (6.6%)	13/265 (4.9%)	6/264 (2.3%)	23/266 (8.6%)	0.0122
Female N=551	12/147 (8.2%)	7/134 (5.2%)	6/137 (4.4%)	7/133 (5.3%)	0.5404
Mortality Rate					
Male N=1038	4/243 (1.7%)	6/265 (2.3%)	5/264 (1.9%)	10/266 (3.8%)	0.3936
Female N=551	9/147 (6.1%)	1/134 (0.8%)	6/137 (4.4%)	2/133 (1.5%)	0.0396

The number of males in the study was almost twice as high as the number of females. Males had higher mean and median values of GFR than females in the same treatment groups. There were statistically significant differences of estimated GFR among males across the treatment groups

(P<0.0001, Kruskal-Wallis test). Group C had the highest average and median value of estimated GFR compared with other treatment groups for both males and females, but the p value for the overall comparison in females across the treatment groups is 0.0892. This of course could be due to the smaller number of patients than required to reach the power designed to detect the statistically significant difference. However, the difference of estimated GFR in males across the treatment groups is more pronounced than in females. This is also true for endpoint of efficacy failure defined as BPAR, graft loss, death or loss to follow up. Group C had the lowest rate of BPAR/GL/death/LTFU than groups A, B and D for both males and females. The difference among the treatment groups is more pronounced in males than in females. The same conclusion can be made for rates of BPAR and graft loss excluding death. There were higher mortality rates in females in groups A and C compared to groups B and D, though the numbers are too small to make meaningful conclusions. The rates of mortality across males were more consistent, with the highest rate occurring in group D and the lowest in group A.

Reviewer's Comment: The higher mortality for females in group C compared to group B is concerning given the finding of higher mortality rate in Prograf/MMF group compared to the Neoral/MMF group in study 158 (see statistical review of Study 158 under Prograf NDAs 50708/S027 and 50709/S021 (b) (4) *The medical reviewer for study 158 found that there were more deaths associated with serious infections in the Prograf group compared to the Neoral group. This reviewer looked at serious infections by gender using the data from D_ADEV.xpt. The rate of serious infection for females in group C (22/138 or 15.9%) was actually lower than the corresponding rate for females in group B (31/139 or 22.3%). In addition, incidence of PTDM by gender does not seem to explain the higher mortality for females in group C either (see Table 4.1D).*

The analysis of mortality by gender from study 158 did not show a similar pattern in female subjects. Note that the study regimens were slightly different in this study, including different dosing of Prograf and Neoral. Mortality in males was 4.4% (6/136) for Prograf/MMF and 1.5% (2/130) for Neoral/MMF and in females it was 3.9% (3/76) for Prograf/MMF and 3.7% (3/82) for Neoral/MMF.

In Table 4.1B, Age is categorized by quartiles. The first quartile for age of patients at the time of transplantation was 34.5 years, median was 47.0 and the third quartile is 56.8 years. The youngest patient was 18.1 years old and the oldest patient was 75.8 years old at the time of transplantation. The patients in younger age groups generally had better GFR values and lower rates of BPAR/GL/death/LTFU. The advantage of group C compared to other treatment groups seemed to carry through all age groups with respect to both estimated GFR and rate of BPAR/GL/death/LTFU. However, the difference seems more pronounced in younger age groups than older age groups. In particular, the overall comparisons for the estimated GFR across the treatment groups resulted p values of 0.0007, 0.0406, 0.0826 and 0.0749 (Kruskal-Wallis test) from the first quartile to the last quartile of the age groups.

Table 4.1B Efficacy Outcome by Age Groups

	Group A N=390	Group B N=399	Group C N=401	Group D N=399	P value
Estimated GFR at month 12					
min - Q1	N=93	N=78	N=110	N=117	
mean ± SD	62.98±26.22	64.78±23.82	75.20±26.86	64.80±25.21	0.0007
median	67.22	66.42	77.63	67.98	
Q1 - median	N=101	N=111	N=95	N=90	
mean ± SD	65.33±23.47	66.83±26.24	70.30±24.68	59.61±27.08	0.0406
median	65.57	67.69	71.78	62.41	
Median – Q3	N=105	N=102	N=96	N=94	
mean ± SD	53.75±23.73	59.08±22.56	62.58±25.74	57.03±29.20	0.0826
median	54.44	61.88	65.35	57.75	
Q3 - Max	N=91	N=108	N=100	N=98	
mean ± SD	45.68±22.39	48.14±23.05	52.70±25.38	43.94±21.50	0.0749
median	47.71	48.47	51.00	43.02	
All patients	N=390	N=399	N=401	N=399	
mean ± SD	57.07±25.10	59.39±25.05	65.40±27.03	56.68±26.88	<0.0001
median	57.04	60.94	66.18	57.45	
Rate of BPAR/GL/Death/LTFU					
Min – Q1	27/93 (29.0%)	20/78 (25.6%)	20/110 (18.2%)	54/117 (46.2%)	<0.0001
Q1 - median	36/101 (35.6%)	41/111 (36.9%)	19/95 (20.0%)	48/90 (53.3%)	<0.0001
Median – Q3	43/105 (41.0%)	34/102 (33.3%)	19/96 (19.8%)	39/94 (41.5%)	0.0041
Q3 - max	35/91 (38.5%)	31/108 (28.7%)	24/100 (24%)	44/98 (44.9%)	0.0080
All Patients N=1589	141/390 (36.2%)	126/399 (31.6%)	82/401 (20.5%)	185/399 (46.4%)	<0.0001

In addition to the efficacy outcome, some safety outcomes are of interest, such as the incidence of infection and PTDM. The infection rates by gender are listed in Table 4.1C below. The infection rates appeared to be higher in female patients than in male patients in all treatment groups. The rates were similar for the same gender across the treatment groups.

Table 4.1C Rate of Infection by Gender

	Group A N=384	Group B N=408	Group C N=403	Group D N=380	P value
Male	125/239 (52.3%)	137/269 (50.9%)	140/265 (52.8%)	139/254 (54.7%)	0.8560
Female	98/145 (67.6%)	89/139 (64.0%)	85/138 (61.6%)	78/126 (61.9%)	0.7079
All patients	223/384 (58.1%)	226/408 (55.4%)	225/403 (55.8%)	217/380 (57.1%)	0.8691

The incidences of PTDM by gender and treatment groups are listed in Table 4.1 D below. The patients with the following preferred terms of adverse events in the dataset D_ADEV.xpt were included for the analysis: Diabetes Mellitus, Diabetes mellitus (non-)insulin-dependent and Diabetes mellitus inadequate control, Glucose tolerance impaired and Hyperglycemia. The rates of PTDM were higher for males than females in groups B and C. But the trend reversed in groups A and D. The difference of PTDM incidence among the groups appears to be driven more by the male patients.

Table 4.1D Incidence of PTDM by Gender

	Group A N=384	Group B N=408	Group C N=403	Group D N=380	P value
Male	25/239 (10.5%)	26/269 (9.7%)	47/265 (17.7%)	33/254 (13.0%)	0.0249
Female	19/145 (13.1%)	8/139 (5.8%)	18/138 (13.0%)	18/126 (14.3%)	0.1041
All patients	44/384 (11.5%)	34/408 (8.3%)	65/403 (16.1%)	51/380 (13.4%)	0.0070

4.2 Other Special/Subgroup Populations

As noted in Dr. Patrick Archdeacon's clinical review, subpopulations of general interest in renal transplant clinical studies include those related to gender, age, race, diabetes, and hepatitis C status. Because instances of post-transplant diabetes were both rare and also poorly documented, and infection with hepatitis C constituted grounds for exclusion from the study, subgroup analyses on PTDM and hepatitis C status were not feasible.

5. SUMMARY AND CONCLUSIONS

5.1 Statistical Issues and Collective Evidence

The statistical issues of the study include the method used to impute missing values, adjustment for multiple pairwise comparisons, sample size increase without pre-specification, generalization of the study result to the US population and lack of safety data.

- The primary efficacy parameter specified in the protocol was the glomerular filtration rate (GFR), 12 months after transplantation, determined from serum creatinine using the Cockcroft-Gault formula to estimate the creatinine clearance. The formula also uses variables weight, age and sex to estimate GFR. For the sponsor's primary efficacy analysis in the ITT population, missing values were handled according to the following procedure: the last observation carry forward (LOCF) was applied from the month 3 visit on, including some cases where the subject had died or lost their graft. When the weight was missing, the weight was imputed by LOCF to calculate Cockcroft-Gault GFR. In all other cases, a value of 10 ml/min was imputed. The sponsor's method of handling missing month 12 creatinine values was a concern.

  The sponsor's efficacy analysis of estimated GFR at month 12 after transplantation in the ITT population (with imputation of missing values at month 12 visit) showed that group C had significantly higher GFR values than the other three groups with and without adjustment for multiple comparisons. The analyses using two alternative imputation methods for the ITT population give the same conclusion as the sponsor's ITT analyses. Weaker evidence is seen in the two analyses which remove patients instead of imputing values, in per protocol population (without imputation of missing values) and in the analysis of measured GFR. This raises questions on whether imputing the missing Month 12 creatinine values had an effect on the stronger evidence of superiority of group C in ITT population. Note that one reason for missing data is due to a subject's death or loss of their graft. Since group C had the smallest number of subjects with death or graft loss, it may be a large reason for the lack of significance in the analyses which exclude these patients. It would not seem appropriate to exclude these patients in an analysis; however, even with exclusion of these patients, there is still the trend with group C doing better than the other groups. The superiority of group C compared with groups A, B and D seems fairly robust in terms of primary endpoint of GFR at month 12 after renal transplantation.

- There were 4 study arms in the Symphony study. This constitutes 6 pairwise comparisons to investigate the difference of outcome between the groups. The sponsor proposed in the protocol to use a Bonferroni-adjusted alpha-level of 0.05/6. However, in the study report, the sponsor reported the results without any adjustment for multiple comparisons. The primary efficacy analyses in this review are presented with 95% CI as well as the 99.2% for the difference of outcome measures between the groups. The 99.2% confidence interval is presented for adjusting multiple comparisons using the Bonferroni method (1-0.05/6=99.2%).

The superiority of group C on estimated GFR and efficacy failure compared with other groups was maintained with the adjustment of multiple comparisons.

- The sponsor in protocol amendment 2 dated September 22, 2003 proposed to increase the sample size to a total of 1760 patients from the originally proposed 1300 patients. The reasons for the sample size increase stated in amendment 2 were to improve the precision of the estimates, to increase the power of the planned statistical tests and to increase the probability of reaching the goals of the study, i.e., of identifying statistically significant differences between the four treatments. Since the sponsor did not intend to use the study in support of a product, the study protocol and its amendments were not submitted to the Agency prior to submission of the study report under the Agency's Aug 18, 2007 request. In addition, sample size re-estimation was not specified in the protocol of this open label study. There is generally a concern about increasing sample size during an ongoing study where blindness is not preserved and steps to control type I error are not explained. In order to study the possible impact on the study results due to the increased sample size, efficacy analysis for the first 1300 randomized patients in the ITT population were performed. With the exception of 99.2% CI between group B and group C, the 95% and 99.2% confidence intervals for difference of overall failures between group C and groups A, B and D do not contain 0, Group C is superior to group A and D with and without Bonferroni correction for multiplicity. The 99.2% confidence interval for the difference of efficacy failure between group B and group C was (-0.5%,18.1%), which did not show a statistically significant difference between group B and C after the conservative Bonferroni correction for multiple comparisons, but still showed the trend with group C doing better than group B with respect to the efficacy failure. For the estimated GFR, the conclusion of this group of patients did not change from the ITT population as a whole. Group C is shown to be superior to other treatment groups with respect to the estimated GFR.

- The study population in ELiTE-Symphony study is not representative of US transplant patients. As shown in Table 3.1.3A, the population studied was primarily white. The numbers of subjects for black, Asian and other races are too small to perform any meaningful subgroup analyses by race. The lack of non-white patients is a concern as to generalization of the results to the US transplant population.

- The ELiTE-Symphony study was not designed as a trial that would be used to support a product. As a result, the quality of the datasets does not meet the usual standards of a registration trial. Numerous errors were found by this reviewer while examining the data and trying to reproduce the sponsor's results. The flaws in the dataset would not likely affect the conclusion on efficacy outcome because the results are fairly robust. However, the data on some important adverse events are not readily available. For instance, the results reported in the study report on PTDM and infections were not reproducible using the data provided by the sponsor.

5.2 Conclusions and Recommendations

In the ELiTE-Symphony study, four immunosuppression regimens with mycophenolate mofetil, and corticosteroids were compared for efficacy and safety. A standard regimen of mycophenolate mofetil, cyclosporine and corticosteroids (group A) was compared to regimens with mycophenolate mofetil, daclizumab and corticosteroids as mainstay immunosuppression in combination with low-dose cyclosporine (group B), low-dose tacrolimus (group C) or low-dose sirolimus (group D). In the assessment of this reviewer, group C was shown to be superior to groups A, B and D on the primary efficacy endpoint defined by the sponsor, the glomerular filtration rate (GFR) 12 months after transplantation which was determined from serum creatinine using the Cockcroft-Gault formula. Group C was also shown to be superior compared with other three groups on the endpoint of efficacy failure defined by rate of BPAR (biopsy proven acute rejection), graft loss, death or loss to follow-up. The results seem to be robust with sensitivity analyses. The conclusion drawn from these results is that a regimen consisting of mycophenolate mofetil, daclizumab, corticosteroids and low-dose tacrolimus as planned and used in this study is efficacious in preventing biopsy confirmed acute rejection and leads to adequate kidney function in the first year.

Given this is an investigator driven study, the quality of the data does not meet the usual standards of a trial in support of a product. In this reviewer's opinion, the flaws in the data submitted do not have the potential to affect the conclusion on efficacy. However, lack of reproducible safety data on adverse events such as PTDM (post transplant diabetes mellitus) and infection may be a concern on the validity of the safety conclusion of the study. In addition, as a result that the ELiTE-Symphony study was conducted outside of US, the study population is not a representative of the transplant patients in the US. Therefore, generalization of the results to the US transplant population should be made with caution.

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

/s/

Hong-Ling Zhou
5/15/2009 05:32:27 PM
BIOMETRICS

Karen Higgins
5/15/2009 06:28:42 PM
BIOMETRICS

STATISTICAL NDA REVIEW AND EVALUATION

Medical Division: **Division of Special Pathogen and Transplant Products**

Biometrics Division: **Division of Biometrics IV**

NDA:	50-708/S027, 50-709/S021
NAME OF DRUG:	Prograf® (tacrolimus) capsules 0.5 mg, 1 mg, and 5 mg and Prograf® (tacrolimus) injection, 5 ng/ml
INDICATION(S):	Prophylaxis of organ rejection in kidney transplantation
APPLICANT:	Astellas Pharma US, Inc. (APUS)
DOCUMENTS REVIEWED:	Briefing Package for Type A Meeting Request
STATISTICAL REVIEWER:	LaRee Tracy, MA
STATISTICAL TEAM LEADER:	Karen Higgins, ScD
BIOMETRICS DIVISION DIRECTOR:	Mohammad Huque, PhD
CLINICAL REVIEWER:	Hui-Hsing Wong, MD, JD
PROJECT MANAGER:	Hyun Son, PharmD

Summary

This submission contains a briefing package for a Type A meeting in reference to the Approvable action taken by the Division on March 14, 2007 regarding the supplemental application to expand Prograf labeling to provide for the adjunct use of MMF in kidney transplant recipients. This is a brief summary of the Division's preliminary responses to the questions in the briefing document (sent on 5/29/07) and relevant discussion during the May 30, 2007 face-to-face (FTF) meeting.

The basis of the approvable action taken was that a safe and effective dosage regimen of MMF as an adjunct therapy with Prograf has not been established in study 02-0-158. Although this study demonstrated non-inferiority (based on a 10% margin) of Prograf/MMF to CsA/MMF with respect to incidence of graft loss, death, BCAR and lost to follow-up at one-year post-transplantation, the risk/benefit analysis did not support inclusion of this study in the CLINICAL STUDIES section of the label because of more AEs known to be associated with the use of MMF as well and an increase in deaths considered to be related to over immunosuppression and infections in the Prograf/MMF arm.

Correspondence from APUS on March 22, 2007 attempted to address the Division's concerns by (b) (4) Specifically, APUS proposed the following:

- (b) (4)

In this submission briefing package, APUS is seeking the Division's input on the proposed changes. Additionally, APUS is proposing to provide the following information as part of an anticipated NDA amendment:

- (b) (4)

(b)
(4)

- All available data from Roche studies of Prograf in combination with MMF (including the Symphony and OptiCept studies)
    - Symphony Trial: compared standard CsA/MMF (1g bid) without antibody induction (n=385), CsA (low)/MMF (1 g bid)/daclizumab (n=399), Prograf/MMF (1 g bid)/daclizumab (n=402), and Sirolimus/MMF (1 g bid)/daclizumab (n=399) in de novo kidney transplant recipients over 1 yr
      Reviewer's Comment: The reference (Ekberg, 2006) cited was not provided nor can it be located in PubMed suggesting that this summary is based on information presented at a scientific meeting only.
    - OptiCept Trial: An open-label, randomized, multicenter study comparing concentration-controlled (CC) MMF/standard level calcineurin inhibitor (CNISL) MMF (CC)/CNI reduced level and fixed-dose MMF/CNISL in renal transplant recipients
      Reviewer's Comments: The study summary notes that this study was conducted in 718 renal transplant recipients; however, interim results from the American Transplant Congress are stated in only 297 patients. Also, a reference (Bloom et al, 2006) as cited; however, the actual publication was not provided nor can one be located in PubMed suggesting that this summary is based only from information presented at a scientific meeting.

- Analysis of the United Network of Organ Sharing (UNOS) databases regarding causes of death, as well as current results of the Collaborative Transplant Study (CTS) regarding Prograf/MMF, Prograf/AZA and CsA/AZA
  Reviewer's Comment: Data from UNOS are limited in that doses/duration of treatment are not collected. These data will be of limited use in the assessment of a safe and effective MMF dose to be given with Prograf.

- Additional PK data from studies and information from the literature to better characterize the MMF dose in combination with Prograf that would provide better similar MPA exposure as that given by 1 g bid MMF in combination with Neoral

- Other safety information to support Prograf in combination with MMF

The following is a summary of the Division's written responses to questions submitted in this meeting package (the complete responses are available in DFS):

Question 1: Our correspondence of March 22, 2007 reiterated differences in our interpretation of Study 02-0-158 from that of the Division. APUS offered a proposed package insert based upon consideration of the Division's comments. Does the Division agree with, or have any comments on, the proposed labeling that was appended to the March 22, 2007 letter?

Response: The Division responded that they did not agree with the proposed labeling and referred APUS to the following FDA Guidance documents: Adverse Reactions Section of Labeling for Human Prescription Drug (final 1/18/06) and Biological Products and Clinical Studies of Labeling for Human Prescription Drug and Biological Products (final 1/18/06). These guidance documents describe the type of information appropriate to include in the package insert. The review of the Prograf/MMF regimen as studied in 02-0-158 (study 158) concluded that this regimen resulted in an unacceptable safety profile. The Division then concluded that it is inappropriate to use study 158 as the basis for the indication of Prograf with MMF and to include information about this study in any section other the WARNINGS section as proposed initially by the Division.

Question 2: APUS believes that a safe and effective dosage regimen has been established for MMF as an adjunct therapy with Prograf in the prophylaxis of organ rejection in renal transplant patients based on the results of Study 02-0-158 and the consistency of these results with other studies reported in the peer reviewed literature and the accumulated data reported by UNOS, as well as the approved labeling for CellCept. APUS believes that the (b) (4)

Does the Division concur?

Response: The Division noted that upon preliminary review of the information in the current submission package, (b) (4)

Question 3: Does the Division believe that the proposed amended dossier would allow the differences between APUS and the Division regarding scientific and regulatory matters with

respect to the use of MMF in combination with Prograf to be resolved at the divisional level, or does the Division consider that formal Dispute Resolution must be pursued at a higher level?

Response: This will be discussed during the FTF meeting.

The following is a brief summary of key points discussed during the FTF meeting (the complete minutes are available in DFS):

- The Division noted that the level of data needed for the re-submission needs to be detailed and not just in summary form. Analysis of the UNOS registry data on outcome does not provide detailed data on dosing regimens and therefore such data would be considered supportive only. APUS responded that they do not intend to use the registry data to determine a dose.
- The Symphony study included additional PK information and the starting dose of MMF was 2 g/day. The sponsor also noted that this study had a better safety profile than that observed in study 158.
- The Division recommended that the sponsor systematically compile data to assess the safety and risk benefit of using 1 g/bid of MMF with Prograf.
- APUS noted that they are negotiation with Roche to obtain all patient-level data for the Symphony study, which will be submitted in a re-submission.

Reviewer's Comment: A safe and effective dose of MMF to be used in combination with Prograf could not be determined from the review of study 158 (see original reviews). It is unclear at present if the information that the sponsor intends to provide in a re-submission will be sufficient to determine a safe and effective regimen. Although the sponsor plans to submit patient level data on the Symphony study, the Division has not received the protocol or analysis plan so the quality of the study can not be determine until a full review is performed.

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

/s/

LaRee Tracy
8/20/2007 03:14:23 PM
BIOMETRICS

Karen Higgins
8/20/2007 04:08:05 PM
BIOMETRICS

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/Serial Number:	NDA 50-708/S-027 NDA 50-709/S-021
Drug Name:	Prograf® (tacrolimus) Capsules, 0.5 mg, 1 mg, and 5 mg, and Prograf® (tacrolimus) Injection, 5 mg/ml
Indication(s):	Prophylaxis of organ rejection in kidney transplantation
Applicant:	Astellas Pharma US, Inc
Date(s):	Application: 02/13/06 PDUFA Due Date: 03/14/07
Review Status:	Standard 10-Month 3-month extension due to major amendment
Biometrics Division:	Division of Biometrics IV
Statistical Reviewer:	LaRee Tracy, M.A.
Concurring Reviewers:	Karen Higgins, Sc.D.
Medical Division:	Division of Special Pathogen and Transplant Products
Clinical Team:	Team Leader: Marc Cavaille-Coll, MD, PhD Medical Reviewer: Hui Hsing Wong, MD, JD
Project Manager:	Hyun Son, Pharm D

Keywords: class immunosuppressant, indication kidney, liver and heart transplantation

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

1. EXECUTIVE SUMMARY

1.1 Conclusions and Recommendations

Study 02-0-158 demonstrated non-inferiority (within in a 10% margin) of Prograf/MMF to Neoral/MMF with respect to incidence of efficacy failure (graft loss, death, BCAR or lost to follow-up) at one year post transplantation in *de novo* kidney transplantation. Specifically, the incidence of efficacy failure in the Prograf/MMF group was 15.1% (32/212) and 17.0% (36/212) in the Neoral/MMF resulting in a difference of -1.9% (Prograf-Neoral), 95.2% CI of -9.0 to 5.2, p-value=0.6 (Table 3.3). Twice as many patients in the Neoral/MMF group (16/212, 13.7%) experienced an episode of BCAR compared to the Prograf/MMF group (16/212, 7.5%), difference (Prograf-Neoral) -6.1, 95 % CI [-12.2, -0.2], p-value=0.04. Despite these differences, the 1-yr incidence of graft loss, death or loss to follow-up was numerically greater in the Prograf/MMF group, 8.5% (18/212) compared to the incidence in the Neoral/MMF group (4.7% , 10/212) resulting in a difference (Prograf-Neoral) of 3.8%, 95.2% CI of -1.05 to 8.92, p-value=0.12 (Table 3.4). This difference was driven, in large part, to a numerically higher number of deaths in the Prograf group (n=9) compared to the Neoral group (n=5).

Death or lost to follow-up at the 2-year follow-up was 17/212 (8.02%) in the Prograf/MMF group and 8/212 (3.77%) in the Neoral/MMF group (lost to follow-up accounted for 5 and 1 patients in the Prograf and Neoral groups respectively) (2-sided Fisher's Exact test p-value=0.097). The medical reviewer (see (b) (4) medical review) concluded that 7 deaths in the Prograf/MMF group were due to over-immunosuppression compared to two deaths in the Neoral/MMF group. These results are further explained by findings from the clinical pharmacology review (see reviews by Dr. Seong Jang, PhD) suggesting that the mean daily MMF dose received in the Prograf group was consistently less than the target 1 gram/twice daily approved dose. Although MMF doses were lower than 2 g/day in the Prograf group, the measured MPA (active moiety of MMF) trough concentrations for Prograf were consistently higher than Neoral concentrations at 30, 192 and 365 days (Table 3.6 & Table 3.7). Specifically, MPA trough concentrations in Prograf group were approximately 76% higher at one month and 29% higher at 12 months than in the Neoral arm. A finding of higher incidence of serious infection in the Prograf/MMF, with some infections associated with mortality, suggests a relationship between higher MPA exposure and over-immunosuppression (refer to medical review for detailed discussion).

Study 02-0-158 demonstrated that the incidence of BCAR was significantly lower in the Prograf/MMF group compared to the Neoral/MMF group contributing to an overall finding of non-inferiority with respect to efficacy failure in the Prograf/MMF group. The numerically higher number of deaths observed at 1 and 2-years follow-up in the Prograf/MMF group compared to the Neoral/MMF control group raises concern regarding the safety of the Prograf/MMF regimen. Appropriate doses of MMF to be given with Prograf in *de novo* kidney transplantation can not be accurately assessed from these data since a safe and effective MMF dose was not assessed.

1.2 Brief Overview of Clinical Studies

Study 02-0-158 was an open-label, randomized, non-inferiority study evaluating the safety and efficacy of Prograf (b) (4) (b) (4) Prograf or Neoral all given in combination with CellCept® (MMF) in *de novo* kidney transplant recipients. The primary endpoint, measured at 1-year post-transplant, was efficacy failure defined as the first occurrence of graft loss, death, biopsy-confirmed acute rejection (BCAR) or lost to follow-up. The study was designed to demonstrate non-inferiority of both Prograf MR/MMF and Prograf/MMF to Neoral/MMF within a 10% margin.

1.3 Statistical Issues and Findings

- Patient survival (defined as no death or lost to follow-up) in study 02-0-158 at **1-year** was **93.9%** (*199/212*) and **97.2%** (*206/212*), difference **-3.3**, 95% CI [**-7.8, 0.7**], p-value=0.1 and at **2-years** was **92.0%** (*195/212*) and **96.2%** (*204/212*), difference **-4.2,** 95% CI [**-9.2, 0.3**], p-value=0.06 in the Prograf and Neoral groups respectively. Graft survival (where graft loss is defined as death, graft failure or lost to follow-up) at 1-year post-transplant was 91.5% (*194/212*) and 95.3% (*202/212*), difference **-3.8**, 95% CI [**-8.9, 0.9**], p-value=0.1 and at 2-years was 89.6% (*190/212*) and 92.4% (*196/212*), difference **-2.8,** 95% CI [**-8.5, 2.7**], p-value=0.3. These findings indicate a consistently higher mortality rate in the Prograf group compared to the Neoral group resulting in lower 95% CI bounds approaching -10%. The Division considers a clinically acceptable margin for non-inferiority (test-control) with respect to patient or graft survival between -5 and -10% for which these results are approaching the -10% limit. Additionally, the medical reviewer found that there were more deaths associated with serious infections in the Prograf group compared to the Neoral group. Lastly, these findings are possibly related to higher MMF exposures observed in the Prograf group compared to the Neoral group (see last bullet below). The totality of these findings warrants concern regarding the safety of the studied Prograf/MMF regimen.

- Study 02-0-158 was open-label and allowed for treatment crossover to an alternative primary calcineurin inhibitor regimen (either to the Prograf/MMF or Neoral/MMF treatment arms). Cross-over was allowed due to an adverse event which led to randomized study drug discontinuation or in the case of severe or refractory rejection. Crossover to the Prograf MR/MMF treatment arm was not permitted and crossover was not considered as the basis for efficacy failure alone. At one-year post-transplant, the crossover rates were 2.8% (6/212), 4.7% (10/214), and 18.4% (39/212) in the Prograf/MMF, Prograf MR/MMF and Neoral/MMF groups respectively. Among patients who crossed over to another treatment and were considered efficacy failures, there were 5 (2.4%) in the Neoral/MMF arm and 2 (0.9%) in the Prograf/MMF who crossed over before reaching the primary endpoint. Of these patients, 3 (1.4%) and 2 (0.9%) respectively crossed over at least 49 days before the reaching the primary efficacy endpoint. The ability to crossover before reaching the primary event, especially in an un-blinded study, introduces potential bias.

- Kidney biopsies were performed based on suspicion of acute rejection and not per protocol. Although this is the standard approach in kidney transplantation, there is concern of bias in assessment of acute rejection rates given the open-label nature of this study.

- Thirty patients were randomized but never received study treatment. The distribution of patients was approximately equal across groups; however, reasons for not receiving treatment were potentially clinically biased against the Neoral treatment group. Refer to the clinical review for details. Additionally, graft and patient outcomes of these 30 patients were not provided in the submission.

- Among the randomized patients who received treatment (ITT, n=638), there were more cadaveric donor types (n=111, 51.9%) compared to living donor types (n=103, 48.1%) in the Prograf MR group. The opposite was true for the Neoral group (i.e. n=101, 47.6% cadaveric; n=111, 52.4% living) and equally distributed in the Prograf group (n=106, 50% each). Among the 30 randomized patients who never received treatment (non-ITT), the donor type was unknown for 3, 6 and 1 patients, living for 3, 1, and 1 patients, and was cadaveric in 6, 4 and 5 patients in the MR, Neoral and Prograf groups respectively. The distribution of donor types among groups and between ITT and non-ITT groups is suspect given the open-label nature of this study. ***Bias in the decisions to not administer treatment in the 30 randomized patients based on donor characteristics (among those that actually received the transplant) can not be ruled out.***

- A DSMB convened with the primary responsibility of monitoring the study for safety; however, the board was also charged with reviewing the primary efficacy endpoint (efficacy failure at 1-year post-transplant). The statistical analysis plan specified two interim analyses after approximately 10% and 50% of the patients had completed the one-year efficacy follow-up. In actuality, the DSMB convened on three separate occasions; however, the only formal interim analysis on efficacy was performed for the third and final DSMB meeting. This interim analysis was based on 45% of patients completing 1-year follow-up. This is reported in the clinical study report and, although it differs from the analysis plan, the amount of alpha spent for the interim analysis (0.002) is appropriate (final analysis type I error=0.048). Estimation of the alpha spent on this interim analysis was based on the Lan and DeMets method (Lan and DeMets, 1983) using an approximate O'Brien-Fleming boundary.

- Ten patients did not receive the full course of protocol-specified induction therapy (n=5) or received no therapy at all (n=5).

- Two grams/day MMF was not consistently achieved in the Prograf group due to elevated MPA exposure requiring reduction of MMF during the study. Despite reductions in doses of MMF received in the Prograf group, increased exposures were observed due to differing interactions between Prograf and MMF compared to Neoral and MMF. These increased exposures most likely contributed to observed toxicities in the Prograf group. This study fails to provide sufficient data on MMF dosing and exposure to allow for accurate determination of a safe and effective dose to be given with Prograf.

2. INTRODUCTION

2.1 Overview

Prograf® (tacrolimus capsules and tacrolimus injection) is approved in the United States for prophylaxis against organ rejection in heart, liver and kidney transplantation. This supplement to NDA 50-708 Prograf® (tacrolimus) capsules (0.5, 1 and 5 mg) and NDA 50-709 Prograf® (tacrolimus) injection 5 mg/ml is to add the adjunct therapy mycophenolate mofetil (MMF, CellCept®) to the Prograf package insert for the indication of prophylaxis of organ rejection in kidney transplant patients (proposed labeling changes below). Currently, MMF is approved at a fixed dose of 1 g/b.i.d. in combination with cyclosporine in kidney transplantation.

The primary confirmatory study supporting this application is study 02-0-158. This study was a prospective, randomized, confirmatory phase III open-label, non-inferiority study comparing Prograf/MMF, Prograf MR/MMF and Neoral/MMF immunosuppressive regimens in *de novo* kidney transplantation. This study was previously reviewed in detail in the statistical review of NDA (b) (4)

Select Labeling Changes Proposed by Applicant:
As noted in bold, underline below, Astellas Pharma US, Inc (Applicant) is proposing the inclusion of MMF to the therapeutic regimen in *de novo* renal transplantation.

INDICATIONS AND USAGE

Prograf is indicated for the prophylaxis of organ rejection in patients receiving allogeneic liver or kidney transplants. (b) (4)

It is recommended that Prograf be used concomitantly with adrenal corticosteroids. Because of the risk of anaphylaxis, Prograf injection should be reserved for patients unable to take Prograf capsules orally

DOSAGE AND ADMINISTRATION

Kidney Transplantation

The recommended starting oral dose of Prograf is 0.2 mg/kg/day (b) administered every 12 hours in two divided doses.

(b) (4)

The Applicant's proposed label changes do not include a recommended MMF dose for use in combination with Prograf; however, the approved MMF regimen in de novo kidney transplantation is 1 g twice a day[1].

2.2 Data Sources

Original submission: \\Cdsesub1\n50708\S_027\2006-02-13
Reviewer requested information:
\\Cdsesub1\n50708\S_027\2006-08-25\060825
\\Cdsesub1\n50708\S_027\2006-09-07
\\Cdsesub1\n50708\S_027\2006-09-07A
\\Cdsesub1\n50708\S_027\2006-09-15
\\Cdsesub1\n50708\S_027\2006-11-15

(b) (4)

3. STATISTICAL EVALUATION

3.1 Evaluation of Efficacy

3.1.1 Study Design and Endpoints

The following is a brief summary of the study design and endpoints for study 02-0-158. Please refer to

1 CellCept® label: http://www.fda.gov/cder/foi/label/2005/050722s013,050723s010,050758s012,050759s015lbl.pdf

the statistical review for ND (b) (4)

Study 02-0-158 was entitled: "*A phase III randomized, open-label, comparative, multi-center study to assess the safety and efficacy of Prograf® (tacrolimus)/MMF, modified release (MR4) tacrolimus/MMF, and Neoral® (cyclosporine)/MMF in de novo kidney transplant recipients*". This study enrolled male and female patients who are 12 years of age or older and undergoing primary or re-transplanted non-HLA-identical living or cadaveric kidney transplantation. Previous and current transplants of an organ other than a kidney were basis for exclusion. The study was conducted between July 2003 and April 2005 in 60 clinical sites across the United States, Canada and Brazil. Study duration was one-year followed by a clinical continuation phase for a minimum of 2 years.

The primary efficacy endpoint of ***efficacy failure*** was defined as death, graft failure [permanent return to dialysis (> 30 days) or re-transplant], biopsy-confirmed (Banff Grade ≥ I) acute rejection (BCAR) based on local assessment, or lost to follow-up at 1-year post-transplant. All suspected rejection must be biopsy confirmed before treatment for rejection is initiated or within 48 hours of initiation of treatment for acute rejection. Routine or surveillance kidney biopsies were not allowed given that this is not considered standard of practice.

Reviewer's Comment: The incidence of graft loss or death equally as important as the primary efficacy endpoint of efficacy failure. A non-inferiority margin for this endpoint within 5-10% is generally considered acceptable by the Division; however, reasons for death or graft loss are considered as well. Statistical adjustments for analysis of this key secondary endpoint are not required.

Treatment failure was defined as *discontinuation of randomized study drug for any reason*. Patients who met the definition of treatment failure were to be followed throughout the 12-month treatment period.

Patients were randomized to Prograf/MMF, MR/MMF or Neoral/MMF in a 1:1:1 fashion stratified by donor type (living or cadaveric) and transplant history (primary or re-transplant). All MMF doses were to be 1.0 g BID (currently labeled dose); however, African American patients were allowed to receive 1.5 g MMF BID if necessary. MMF dose changes for adverse events were permitted at the investigator's discretion if clinically indicated. Target trough MMF levels (MPA) were not identified; therefore, dose adjustments were not to be made based on MPA trough concentration.

Patients were allowed to crossover to an alternate primary calcineurin inhibitor regimen in the case of an adverse event or severe refractory rejection leading to study drug discontinuation. Crossover was to be discussed with the Applicant's medical monitor for approval. Crossover to the MR/MMF group was not allowed. All patients were followed regardless of treatment status. Premature treatment discontinuation of MMF was allowed at investigator's discretion. Crossover in the absence of one of the primary endpoints was not considered efficacy failure.

Reviewer's Comment: There is concern of bias in the incidence of crossover due to the open-label nature of this study design. Conclusions regarding regimen tolerance as a function of crossover rate should be interpreted with caution. (b) (4) (b) (4)

(b) (4)

The primary endpoint was measured at one year; however, patients were followed for efficacy and safety every six months up to a minimum of two years. The 120-day safety update contains efficacy and safety data up to cut-off day 780 (2-years follow-up).

3.1.2 Statistical Methods

A brief outline of the statistical methods used in study 02-0-158 follows. (b) (4)

For the 1-year efficacy failure rate, the two primary treatment group differences were calculated as the experimental regimen minus the comparator (Prograf/MMF minus Neoral/MMF; MR/MMF minus Neoral/MMF). Only data collected up to and including day 365 post-randomization were used in the primary analysis. For the primary analysis, the treatment differences (Prograf/MMF – Neoral/MMF and MR/MMF – Neoral/MMF) for the efficacy failure rate were calculated and confidence intervals of the treatment differences were constructed using a normal approximation. Interim efficacy analyses performed by the DSMB necessitated adjustment of the type I error rate. The applicant performed this using the method of Lan and DeMets [1983] with an O'Brien-Fleming boundary. The applicant performed one formal interim analysis when approximately 45% of patients completed the one-year follow-up visit for the primary endpoint. The amount of alpha spent during this analysis was calculated to be 0.002 yielding an overall type I error of 0.048 (0.05-0.002). Statistical adjustments required (to preserve the overall type I error of 5%) for the two primary efficacy comparisons (i.e. MR/MMF vs. Neoral/MMF and Prograf/MMF vs. Neoral/MMF) were performed using the Hochberg method [1988]. This method essentially calculates the two-sided 95.2% confidence intervals (i.e. (1- 0.048)*100) for each comparison. Both upper bounds of the 95.2% confidence intervals were required to be less than 10% to achieve non-inferiority of both regimens to control. If the least significant comparison failed to demonstrate non-inferiority, the other comparison was to be considered using a 97.6% confidence interval (i.e. (1-0.048/2)*100).

Non-Inferiority Margin Justification

The inclusion of a calcineurin inhibitor (CNI), specifically cyclosporine or tacrolimus, as part of an immunosuppressive regimen has led to marked decreases in acute rejection rates and significant improvements in 1-year graft survival in kidney transplantation. Previous clinical trials comparing regimens of cyclosporine (Sandimmune®, FDA approval 1983) plus steroids to non-CNI regimens (predominantly azathioprine plus steroids) demonstrated statistically significant decreases in the number of acute rejection episodes in the CNI treated patients. Acute rejection rates were 4 to 10 times higher in the azathioprine/steroid groups compared to the CNI/steroid regimen. Graft survival was also higher in the CNI treated patients [Canadian Multicenter Transplant Study Group, 1983]. Neoral®, a microemulsion formulation of cyclosporine, has been shown to be equivalent to cyclosporine and is also FDA approved for use in kidney transplantation (NDAs 50-715 and 50-716).

CNI-based immunosuppressive therapy in kidney transplantation is well-established with over 20 years of clinical use. Recent studies suggest increased acute rejection with CNI withdrawal and decreased graft loss when including a CNI in the therapeutic regimen. A 10% non-inferiority margin for the composite endpoint of acute rejection, patient and graft loss and lost to follow-up is viewed as considerably lower

than the benefit of immunosuppressive therapy over placebo alone. A more restrictive margin, between 5 and 10%, is typically used to assess the outcome of graft loss and death.

Reviewer's Comment: An upper CI around the difference in patient and graft loss nearing 10% is of concern. If such an outcome occurs, causes of patient and graft loss will be taken into consideration along with efficacy and other safety findings in a thorough risk benefit analysis.

Reviewer's Comment: A randomized controlled study by Ashberg et al. [2006] used MMF/steroids and antibody induction with and without CNI is the most relevant published study for this current study. A CI around the observed difference was calculated to be [-65.1, -16.4]. This CI would exclude a margin of 10% considered appropriate for BCAR, graft and patient survival in kidney transplantation. Note however that the width of this CI is a reflection of the small sample size of the Ashberg et al. study, which occurred due to premature halting of patient enrollment from safety concerns.

3.1.3 Demographics and Patient Disposition

Study 02-0-158

Six hundred sixty-eight (668) patients were randomized among 60 clinical sites in three countries (US, n=50; Canada, n=5; and Brazil, n=5). Of these 95%, 97% and 95% in the MR/MMF, Prograf/MMF and Neoral/MMF groups respectively received study drug. The largest number of patients enrolled at one site was 42 (Brazil) and one site enrolled only one patient. The mean enrollment was 11 patients per site. The majority (548/668, 82%) of study patients were enrolled in sites within the United States. Ninety-two (13.8%) patients were enrolled in clinical sites in Brazil and the remaining 28 patients were enrolled in Canadian sites.

At 1-yr post-transplantation, 14.5%, 15.6% and 28.8% of patients prematurely discontinued treatment in the MR, Prograf and Neoral groups (Table 3.1). These differences were driven largely by the significantly different rates of treatment cross-over particularly from the Neoral group to the Prograf group. As discussed in other sections of this review, these findings are suspect given the open-label study design and broad protocol criteria allowing for treatment changes. The primary reason for premature treatment discontinuation was adverse event.

Table 3.1: Patient Disposition (Study 02-0-158)

	MR/MMF	Prograf/MMF	Neoral/MMF
Randomized	226	219	223
Full Analysis Set*	214 (94.7)	212 (96.8)	212 (95.1)
Crossover**	10 (4.4)	6 (2.7)	39 (27.5)
Completed 1-yr of therapy	183 (85.5)	179 (84.4)	151 (71.2)
Discontinued Randomized Therapy	31 (14.5)	33 (15.6)	61 (28.8)
AE	19 (8.9)	23 (10.8)	37 (17.5)
Rejection	1 (0.5)	0	16 (7.5)
Non-Compliance	2 (0.9)	4 (1.9)	5 (2.4)
Graft Failure	2 (0.9)	3 (1.4)	1 (0.5)
Withdrawal of Consent	4 (1.9)	0	1 (0.5)
LTF	0	1 (0.5)	0
Other#	3 (1.4)	2 (0.9)	1 (0.5)

** Randomized and received at least one dose of study treatment*
*** Protocol allowed crossover to another study treatment except to MR/MMF regimen*
 # Prograf/MMF: Converted to rapamycin, acute tubular necrosis, MR/MMF: incorrect study drug dispensed, investigator discretion/possible toxicity, improper drug absorption, Neoral/MMF: subsequent pancreas transplant after kidney transplant

There were no statistically significant differences among treatment groups in any baseline characteristics; however, there were more cadaveric donors compared to living donors in the MR group and the opposite was true in the Neoral group (Table 3.2). This discordance adds to the concern of bias associated with the open-label nature of this study along with the fact that 30 randomized patients did not receive study drug. The donor type breakdown for these 30 patients was *unknown donor type*: 3, 6, and 1; *cadaveric donor type*: 6, 4, and 5; and *living donor type*: 3, 1, and 1 in the MR/MMF, Neoral/MMF and Prograf/MMF groups respectively. The majority of randomized patients (614/638; 96.2%) received a primary kidney transplant upon entry into the study, with approximately half receiving a kidney from a deceased donor and half receiving a kidney from a living donor. The majority of patients (491/638; 77.0%) had ≥ 3 HLA mismatches. There were no significant differences across treatment groups in terms of baseline status of hypertension, diabetes type I or II, or hyperlipidemia. There were no clinically significant differences in the history/type or duration of pre-study dialysis among the three treatment groups. The majority of patients (539/638; 84.5%) underwent kidney dialysis (hemodialysis, peritoneal dialysis, or both) prior to study entry, with the median duration of pre-study dialysis being 29 months. There were significantly more male donors in the MR/MMF group (113/214, 52.8%) compared to the Prograf/MMF group (86/212, 40.6%) (2-sided chi-square p-value=0.011).

There were no significant differences across treatment groups for any of the other donor demographic parameters (race, age, and age group). It was noted that numerically more males in the Prograf/MMF treatment group (85/136; 62.5%) received a kidney from a female donor than males in the MR/MMF treatment group (68/138; 49.3%) or Neoral/MMF treatment group (73/130; 56.2%).

Table 3.2: Patient/Donor Demographics (Study 02-0-158)

Variable	*MR/MMF (n=214)*	*Prograf/MMF (n=212)*	*Neoral/MMF (n=212)*
Country			
Brazil	31 (14.4)	30 (14.1)	30 (14.1)
Canada	10 (4.7)	7 (3.3)	10 (4.7)
United States	173 (80.8)	175 (82.5)	172 (81.1)
Males	138 (34.1)	136 (33.7)	130 (32.2)
Donor Males	113 (52.8)	86 (40.6)	96 (45.3)
Race			
Black	41 (19.1)	51 (24.0)	36 (17.0)
Caucasian	160 (74.8)	152 (7.2)	163 (7.7)
Asian	5 (2.3)	5 (2.3)	8 (3.8)
Other	8 (3.7)	4 (1.9)	5 (2.3)
Mean Age (range)	47.8 (17-77)	48.6 (19-74)	47.6 (17-77)
Mean Donor age (range)	38.2 (2-72)	39.0 (0-68)	39.9 (17-63)
Donor Type*			
Cadaver	111 (51.9)	106 (50.0)	101 (47.6)
Living	103 (48.1)	106 (50.0)	111 (52.4)
Living non-related	*41 (19.2)*	*41 (19.3)*	*37 (17.4)*
Living related	*62 (29.0)*	*65 (30.7)*	*74 (34.9)*

Values in parenthesis are percentages unless otherwise noted
** There were numerically more cadaveric donors in the MR group compared to living donors. The opposite was observed in the Neoral group.*

Please refer to the statistical review for (b) (4) *located in DFS (submitted on* (b) (4) *for a more comprehensive review.*

3.1.4 Efficacy Analyses

Study 02-0-158
One year efficacy failure (first occurrence of BCAR, graft loss, death or loss to follow-up) was **14%** (30/214), **15.1%** (32/212) and **17.0%** (36/212) in **MR/MMF**, **Prograf/MMF** and **Neoral/MMF** groups respectively yielding differences of -3.0, 95.2% CI [-10.0, 4.0], p-value=0.4 (MR-Neoral) and -1.9, 95.2% CI [-9.0, 5.2], p-value=0.6 (Prograf-Neoral). The upper 95.2% confidence bound for both comparisons fell below the clinically relevant 10% margin demonstrating non-inferiority of these regimens to Neoral/MMF. See Table 3.3 below.

When adjusting for donor type (living or cadaveric) the 95.2% confidence interval around MR-Neoral difference is (-10.2, 3.6) and around the Prograf-Neoral difference is (-9.1, 4.8).

Incidence of Patient and Graft Survival
The lower bounds around the 95% confidence interval closely approach -10% for the difference in patient and graft survival, which includes lost to follow-up as failure, between Prograf/MMF and Neoral/MMF (Table 3.4). Additionally, at day 365 three times more deaths, 9/212 (4.2%) vs. 3/214 (1.4%) occurred in the Prograf/MMF group compared to the MR/MMF groups, which is clinically and near statistically relevant. This difference is notable given the open-label nature of this study where patients were given the option to switch to an alternative therapy. In the Neoral/MMF group, five patients died at one year of follow-up. Sepsis was more commonly the cause of death in the Prograf/MMF group compared to the other two groups. There were no notable differences in death rates across countries or clinical sites. Additional analyses of patient survival are presented in section 3.2.

Table 3.3: Efficacy Failure at day 365 (Study 02-0-158)

Day 365	MR/MMF (n=214)	Prog/MMF (n=212)	Neoral/MMF (n=212)	Difference, 95.2% CI, p-value		
				MR-Neoral	Prograf-Neoral	MR-Prograf
Efficacy failure[1]	**30 (14.0)**	**32 (15.1)**	**36 (17.0)**	**-3.0, [-10.0, 4.0], 0.4**	**-1.9, [-9.0, 5.2], 0.6**	**-1.1, [-7.9, 5.8], 0.8**
Reason for failure:						
BCAR*	22 (10.3)	16 (7.5)	29 (13.7)	-3.4, [-9.7, 2.8], 0.3	-6.1, [-12.2, -0.2], 0.04	2.7, [-2.8, 8.4], 0.4
Graft Failure[2]	4 (1.9)	6 (2.8)	4 (1.9)			
Death[3]*	3 (1.4)	9 (4.2)	5 (2.4)	-0.9, [-4.2, 2.0], 0.5	1.9, [-1.7, 5.8], 0.3	-2.8, [-6.6, 0.3], 0.08
LTF[4]	3 (1.4)	4 (1.9)	1 (0.5)			

[1]Incidence of the first occurrence of graft failure, death, local confirmed acute rejection or loss to follow-up
[2] Permanent (>30 days) return to dialysis or re-transplant <u>not</u> resulting in death. Note that 1 subject in each arm had a BCAR prior to experiencing graft failure.
[3] Note that 2 Neoral/MMF subjects experienced BCAR prior to death, 2 patients in Neoral group died after crossing over to Prograf and one patient in the Prograf/MMF group died after crossing over to Neoral/MMF
[4] One subject in the MR/MMF and 1 subject in the Prograf/MMF were LTF after experiencing an acute rejection. An additional patient on the Prograf/MMF arm was LTF after experiencing graft failure.
****95% CI and p-value calculated using Fisher's Exact Test***

Table 3.4: Patient and Graft Survival (Study 02-0-158)

				Difference, 95.2% CI, p-value		
	MR/MMF (n=214)	**Prog/MMF (n=212)**	**Neoral/MMF (n=212)**	**MR-Neoral**	**Prograf-Neoral**	**MR-Prograf**
Day 365 (1 yr follow-up)						
Patient Survival	**208 (97.2)**	**199 (93.9)**	**206 (97.2)**	0, [-3.6, 3.6], 0.99	-3.3, [-7.8, 0.7], 0.1	3.3, [-6.9, 7.8], 0.1
Reason:						
Death*	3 (1.4)	9 (4.2)	5 (2.3)	-0.9, [-4.2, 2.0], 0.5	1.9, [-1.7, 5.8], 0.3	-2.8, [-6.6, 0.4], 0.08
LTF	3 (1.4)	4 (1.9)	1 (0.5)			
Graft Survival	**204 (95.3)**	**194 (91.5)**	**202 (95.3)**	0, [-4.3, 4.4], 0.99	-3.8, [-8.9, 0.9], 0.1	3.8, [-0.9, 8.9], 0.1
Reason						
Death	3 (1.4)	9 (4.2)	5 (2.3)			
Graft Failure[1]	5 (2.3)	9 (4.2)	4 (1.9)			
LTF[2]	3 (1.4)	4 (1.9)	1 (0.5)			
Day 547 (1.5 yr follow-up)						
Patient Survival	**207 (96.7)**	**195 (92.0)**	**204 (96.2)**	0.5, [-3.3, 4.4], 0.8	-4.2, [-9.2, 0.3], 0.06	4.7, [0.3, 9.6], 0.03
Reason:						
Death*	4 (1.9)	12 (5.7)	7 (3.3)	-1.4, [-5.0, 1.8], 0.5	2.4, [-1.7, 6.7], 0.2	-3.8, [-7.9, -0.2], 0.04
LTF	3 (1.4)	5 (2.3)	1 (0.5)			
Graft Survival	**202 (94.4)**	**190 (89.6)**	**198 (93.3)**	1.0, [-3.8, 5.9], 0.7	-3.8, [-9.4, 1.6], 0.2	4.8, [-0.4, 10.3], 0.07
Reason						

Death	4 (1.9)	12 (5.7)	7 (3.3)			
Graft Failure[1]	6 (2.8)	9 (4.2)	6 (2.8)			
LTF[2]	3 (1.4)	5 (2.3)	1 (0.5)			
Day 780 (2 yr follow-up)						
Patient Survival	**206 (96.3)**	**195 (92.0)**	**204 (96.2)**	0, [-3.9, 4.0], 0.99	-4.2, [-9.2, 0.3], 0.06	4.3, [-0.2, 9.2], 0.06
Reason:						
Death	5 (2.3)	12 (5.7)	7 (3.3)	-1.0, [-4.6, 2.5], 0.6	2.4, [-1.7, 6.7], 0.2	-3.8, [-7.9, -0.2], 0.04
LTF	3 (1.4)	5 (2.3)	1 (0.5)			
Graft Survival	**198 (92.5)**	**190 (89.6)**	**196 (92.4)**	0, [-5.2, 5.3], 0.99	-2.8, [-8.5, 2.7], 0.3	2.9, [-2.6, 8.6], 0.3
Reason						
Death	5 (2.3)	12 (5.7)	7 (3.3)			
Graft Failure[1]	9 (4.2)	9 (4.2)	8 (3.8)			
LTF[2]	3 (1.4)	5 (2.3)	1 (0.5)			

[1]Permanent (>30 days) return to dialysis or re-transplant <u>not</u> resulting in death.
Note: 1 patient in MR/MMF and 3 patients in Prograf/MMF died after experiencing graft failure
[2] Note that 1 subject in the Prograf/MMF was LTF after experiencing graft failure.
****95% CI and p-value calculated using exact methods***

Incidence of Treatment Crossover:

There were 10 (4.7%), 6 (2.8%) and 39 (18.4%) patients in the MR/MMF, Prograf/MMF and Neoral/MMF groups respectively who crossed over to alternative therapy by study day 365. Of these 0, 2, and 3 patients in the MR, Prograf and Neoral groups respectively reached the primary endpoint AFTER crossing over to an alternative therapy. Two, zero and twenty patients in the MR, Prograf and Neoral reached the primary endpoint PRIOR to crossover. The remaining 8, 4, and 16 patients in the MR, Prograf and Neoral groups respectively crossed over but did not reach the primary endpoint up to day 365 (Table 3.5). Although crossover rates were disproportionate among treatment groups, the number of primary events occurring after time of crossover was negligible. Sensitivity analysis treating any cross over as a failure would bias the results in favor of the Prograf treatment groups.

Reviewer's Comment: The applicant reports a statistically significantly lower incidence of treatment discontinuation and crossover in both Prograf groups compared to the Neoral group. These rates should be interpreted with caution given the open-label nature of this study and that the study protocol explicitly allowed for treatment crossover. Treatment crossover did not necessarily imply efficacy failure, i.e. a patient could be switched to an alternate therapy without having reached the primary efficacy endpoint. Given this it can be argued that the decision to switch treatment occurred for less severe reasons since crossover did not result in a study outcome. Additionally, data suggest that tacrolimus is more commonly used in kidney transplantation, about 72% of transplant recipients at discharge, compared to Neoral [Meier-Krieschea et al., 2006]. Given this, there is potential for investigator bias when deciding to switch patients from Neoral to Prograf during the study. Lastly, patients who switched to Prograf or to Neoral based regimen were also provided drug free of charge for duration of the study therefore removing the incentive to stay on randomized therapy for cost reasons.

The Division recommended a blinded study design during the protocol development phase of this study (see communication on 7/28/03 and 9/22/03). The applicant responded that such a design was not feasible given logistical (over-encapsulating capsules) and clinical (therapeutic drug monitoring and differing adverse event profiles between calcineurin inhibitor regimens) hurdles.

Table 3.5: Incidence of Treatment Cross Over by Day 365 (Study 02-0-158)

	MR/MMF (n=214)	**Prograf/MMF (n=212)**	**Neoral/MMF (n=212)**
Crossed Over By Day 365			
Yes	10 (4.7)	6 (2.8)	39 (18.4)
Reached Event Before Crossing Over	*2 (0.9)*	*0 (0)*	*20 (9.0)*
*Reached Event after Crossing Over**	*0 (0)*	*2 (0.9)*	*3 (1.9)*
Never Reached Event	*8 (3.7)*	*4 (1.9)*	*16 (7.5)*
No	204 (95.3)	206 (97.2)	173 (81.6)
Reached Event	*28 (13.1)*	*30 (14.1)*	*13 (6.1)*

Time (days) to primary event after crossover: Prograf: 48, 86; Neoral: 5, 152, 244

MMF Exposure

The study protocol specified 2 g/day (1 g/BID) dose was not consistently achieved in either of the Prograf treatment groups. As discussed in greater detail in the clinical pharmacology review, calculated mean daily doses of MMF in both the Prograf and MR groups were consistently less than 2 g/day (Table 3.6); however, MPA trough concentrations for Prograf and MR were consistently higher than Neoral concentrations at 30, 192 and 365 days (Table 3.7). Specifically, MPA trough concentrations in both tacrolimus groups were approximately 76% higher at one month and 29% higher at 12 months than in the Neoral arm. Appropriate doses of MMF to be given with Prograf can not be fully assessed from these data.

Table 3.6: Mean daily dose (mg/day) of MMF (Study 02-0-158)

Time period (Days)		Treatment Group		
		MR/MMF	**Prograf/MMF**	**Neoral/MMF**
1-30	Mean±SD	1871±344	1851±396	1962±410
	N	214	212	210
	Median [Range]	1967 [633-3000]	1967 [250-2950]	1967 [166.7-5500]
31-90	Mean±SD	1748±477	1737±538	1881±439
	N	208	206	200
	Median [Range]	2000 [0-3000]	2000 [0-3000]	2000 [0-3000]
91-183	Mean±SD	1635±552	1571±639	1765±526
	N	203	201	185
	Median [Range]	1973 [0-3000]	1978 [0-3000]	2000 [0-3000]
183-365	Mean±SD	1489±646	1405±669	1708±560
	N	186	187	170
	Median [Range]	1671 [0-3000]	1500 [0-3000]	2000 [0-3000]

Source: Table adapted from clinical pharmacology review

Table 3.7: MPA trough concentrations (μg/mL) (Study 02-0-158)

Time (Days)	Prograf/MMF	MR4/MMF	Neoral/MMF
30	3.7±3.1 (N=152)	3.3±2.2 (N=145)	2.1±2.1 (N=151)
	2.75 [0-17.5]	2.8 [0-13.5]	1.5 [0-12.7]
182	3.4±3.0 (N=147)	3.3±2.5 (N=131)	2.7±3.0 (N=153)
	2.7 [0-22.8]	2.9 [0-15.3]	1.9 [0-18]
365	3.1±2.6 (N=134)	3.0±2.6 (N=126)	2.4±2.3 (N=153)
	2.5 [0-15]	2.5 [0-18.4]	1.7 [0-13.8]

Mean±SD (N), Median [range]
Source: Table adapted from clinical pharmacology review

3.2 Evaluation of Safety

Study 02-0-158

Safety analysis was performed in all randomized patients who received at least one dose of study drug and were on therapy or had discontinued (or crossed over to an alternate therapy) within 10 days. Data beyond the cutoffs established for the efficacy and safety analyses were included in the data listings if available. The following events occurred more frequently than what is typically observed in this patient population but this is not an exhaustive list. Refer to the clinical review for a more detailed discussion and analysis of safety findings.

Patient Survival

In study 02-0-158, the 1-year patient survival in patients receiving kidneys from living donors was 100%, 97.2% and 98.2% in MR, Prograf and Neoral respectively. Survival in recipients of kidneys from deceased donors was 97.3%, 94.3% and 97.0% in MR, Prograf and Neoral groups respectively. *These rates do not account for loss to follow-up or graft failure not resulting in death for comparison against registry data.* In comparison, the Organ and Procurement and Transplantation Network (OTPN), estimates 1-year unadjusted patient survival rates of 94.6% (n=14,476) and 97.9% (n=11,339) in kidney transplant recipients from deceased and living donors respectively (transplant 2002-2003)[1]**.** These data are shown below in Table 3.8. Although the patient survival rates observed in study 02-0-158 were similar to registry data from 2002-2003, these rates are somewhat less than expected in the context of an open-label randomized clinical trial. Generally, patients enrolled in such clinical trials are healthier and are managed more closely than the overall population. It is therefore expected that patient survival should be considerably greater than that observed in the overall population. This was not observed in the Prograf/MMF group. Additionally, this sensitivity analysis does not treat lost to follow-up as a failure (to allow for comparison against registry data). When including lost to follow-up, patient survival in study 02-0-158 is lower than reported in this sensitivity analysis (refer to Table 3.4).

Table 3.8: Patient Survival* in Study 02-0-158 and UNOS Data at 1-Year

	Living Donor	**Deceased Donor**
MR	100% (n=103)	97.3% (n=111)
Prograf/MMF	97.2% (n=106)	94.3% (n=106)
Neoral/MMF	98.2% (n=111)	97.0% (n=101)
OPTN/UNOS Data (unadjusted)	97.9% (n=11,339)	94.6% (n=14,476)

**** Ignores lost to follow-up***

1 http://www.ustransplant.org/annual_reports/current/113_surv-new_dh.htm

Analysis of 2-yr mortality in study 02-0-158:

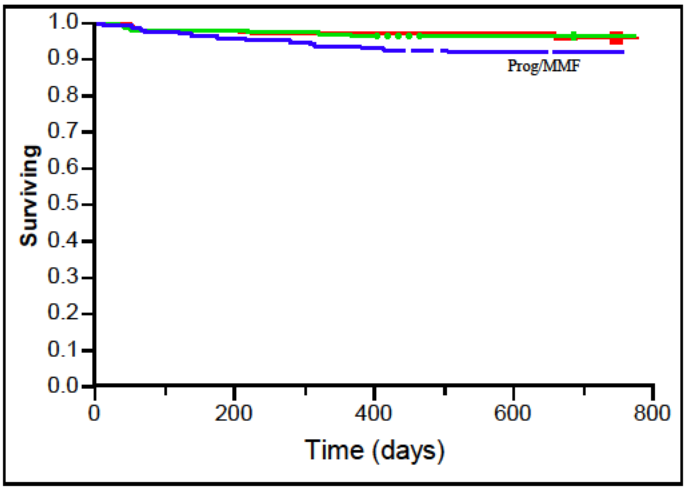
Analysis at day 780 treating death and lost to follow-up in the intent-to-treat population as failure:

Table 3.9: 2-year Patient Survival (Study 02-0-158)

Day 780	MR/MMF (n=214)	Prog/MMF (n=212)	Neoral/MMF (n=212)
Death or LTF	**8 (3.74)**	**17 (8.02)**	**8 (3.77)**
Death	5 (2.3)	12 (5.7)	7 (3.3)
Cadaveric Donor Type	4	9	5
Living Donor Type	1	3	2
LTF	3 (1.4)	5 (2.3)	1 (0.5)
Mean Days to Event (+/- SD)	646.1 (+/-6.9)	485.9 (+/- 6.5)	461.5 (+/- 4.5)
Logrank Test (vs. Neoral/MMF)	P=0.96	P=0.06	NA

Time to death or lost to follow-up was not statistically significant among treatment groups (log rank test) (Figure 1). This is an intent-to-treat analysis and therefore does not account for the high rate of treatment cross-over observed in the Neoral group.

Figure 1: Two-year patient survival (Study 02-0-158)



The following are primary reasons for death in study 02-0-158 (adapted from the medical officer's review of NDA (b) (4), Dr. Hui-Hsing, Wong, MD, JD) (see review in DFS) at the 2-year follow-up. Reasons in bold were suspected (as per the medical review) deaths associated with serious infection.

Table 3.10 Reported Deaths (Study 02-0-158)

2-year follow-up	Death (days)	Primary Cause of Death
Prograf/MMF		
Patient ID		
01811002	1	Cardiac Arrest
10202007	57	**Tissue invasive strongyloidsis**
01652002	69	**Nephrectomy/sepsis**
00352003	123	**Sepsis related to vancomycin-resistant enterococci infection**
00442004	141	Possible pulmonary embolism
10211002	142	**Recurrent disease/miliar[y] tuberculosis *(crossed-over to Neoral group on day 42)***
00512003	178	Subdural bleed after fall
10181003	218	**Chronic allograft nephropathy/sepsis**
02082013	311	Homicide
07502001	374	**Metastatic renal carcinoma with hemothorax**
00432002	383	Cardiac Arrest
04062008	415	**Septic shock**
MR/MMF		
00512005	53	**Lymphocytic Choriomeningitis**
00292003	57	**Cardiac and respiratory arrest** (*review of records suggests death from infection*)
01092001	227	Chronic allograft nephropathy/stroke
00711010	429	Probable myocardial infarction
00712002	663	**Sepsis**
Neoral/MMF		
10931013	35	Myocardial infarction
10222001	45	Pulmonary edema
00712001	55	Probably pulmonary embolus
10212009	222	Diverticulitis ***(crossed-over to Prograf group on day 129)***
00321006	324	Myocardial infarction
01812009	371	**Encephalitis**
00252003	472	Suicide

Incidence of Diarrhea

The incidence (by patient) of diarrhea was statistically significantly higher in both the MR/MMF group (100/214, **47%**, p-value <0.0001, 95% CI 7.3, 25.5) and Prograf/MMF group (94/212, **44%**, p-value 0.002, 95% CI 5.0, 23.1) compared to Neoral/MMF (64/212, **30%**) (Table 3.11).

Table 3.11: Incidence of Diarrhea up to 365 days (Study 02-01-58)

	MR4 N=214	**Prograf N=212**	**Neoral N=212**
Incidence of diarrhea			
Patients	100 (46.7%)	94 (44.3%)	64 (30.2%)
P-value, 95% CI (vs. Neoral)	*<0.001, [7.3, 25.5]*	*0.002, [5.0, 23.1]*	
Events*	141 (6)	135 (2)	89 (19)
Duration of diarrhea (by event)			
Mean no. of days	27.3	22.1	18.0
Median no. of days	8	9	5.5
Severity of Diarrhea (by event)			
Mild	96	90	55
Moderate	42	40	31
Severe	3	5	3

** Events occurring after treatment cross-over in parenthesis*
Analysis using ADV.xpt file, diarrhea term based on MedRA preferred term, event day <=365 days

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Race and Age

Please refer to the statistical review for (b) (4) *located in DFS (submitted on* (b) (4) *a more comprehensive review.*

4.2 Other Special/Subgroup Populations

Please refer to the statistical review for (b) (4) *located in DFS (submitted on* (b) (4) *a more comprehensive review.*

5. SUMMARY AND CONCLUSIONS

5.1 Statistical Issues and Collective Evidence

Patient and graft survival was consistently numerically lower in the Prograf group compared to the Neoral group with the lower bound of the 95% confidence interval approaching -10%. There were more deaths associated with serious infections noted in the Prograf group compared to the Neoral group. Additionally, higher MMF exposures were observed in the Prograf group compared to the Neoral group. The higher exposures are potentially associated with the higher incidence of serious infections observed in the Prograf group.

The open-label nature of study 02-0-158 introduces potential bias in study conduct and assessment of patients outcomes. Specifically, there is concern regarding ascertainment and assessment bias of non-protocol specified kidney biopsies therefore raising concerns regarding differences observed in rates of BCAR. Additionally, the open-label nature of this study introduces bias with respect to treatment options and decisions to receive, administer, modify or stop study treatment. This bias challenges firm conclusions regarding any differences in premature treatment discontinuation and crossover.

As discussed in greater detail in the clinical pharmacology and clinical reviews, the mean daily doses of MMF in the Prograf group were consistently lower than the Neoral group and below the protocol specified dose of 2 g/day in study 02-0-158. These inconsistencies in MMF dosing present challenges in assessing comparability of MMF dose between the Neoral comparator group and the Prograf group. Given that MMF is a secondary immunosuppressive often given in combination with tacrolimus these data do not provide sufficient information to conclude an appropriate safe and effective dose of MMF.

In conclusion, this single phase III study demonstrated that a Prograf/MMF regimen was non-inferior (within a 10% margin), with respect to the primary efficacy endpoint (incidence of graft loss, acute rejection, death, and lost to follow-up) to a Neoral/MMF regimen. One and two-year patient survival rates were surprisingly lower in the Prograf group compared to the Neoral group. The efficacy of the Prograf/MMF regimen was demonstrated; however, there are several inconsistencies regarding MMF dosing and safety findings that require further study.

5.2 Conclusions and Recommendations

Study 02-0-158 demonstrated non-inferiority (within in a 10% margin) of Prograf/MMF to Neoral/MMF with respect to incidence of efficacy failure (graft loss, death, BCAR or lost to follow-up) at one year post transplantation in *de novo* kidney transplantation. Specifically, the incidence of efficacy failure in the Prograf/MMF group was 15.1% (32/212) and 17.0% (36/212) in the Neoral/MMF resulting in a difference of -1.9% (Prograf-Neoral), 95.2% CI of -9.0 to 5.2, p-value=0.6 (Table 3.3). Twice as many patients in the Neoral/MMF group (16/212, 13.7%) experienced an episode of BCAR compared to the Prograf/MMF group (16/212, 7.5%), difference (Prograf-Neoral) -6.1, 95 % CI [-12.2, -0.2], p-value=0.04. Despite these differences, the 1-yr incidence of graft loss, death or loss to follow-up was numerically greater in the Prograf/MMF group, 8.5% (18/212) compared to the incidence in the Neoral/MMF group (4.7% , 10/212) resulting in a difference (Prograf-Neoral) of 3.8%, 95.2% CI of -1.05 to 8.92, p-value=0.12 (Table 3.4). This difference was driven, in large part, to a numerically higher number of deaths in the Prograf group (n=9) compared to the Neoral group (n=5).

Death or lost to follow-up at the 2-year follow-up was 17/212 (8.02%) in the Prograf/MMF group and 8/212 (3.77%) in the Neoral/MMF group (lost to follow-up accounted for 5 and 1 patients in the Prograf and Neoral groups respectively) (2-sided Fisher's Exact test p-value=0.097). The medical reviewer (see (b) (4) medical review) concluded that 7 deaths in the Prograf/MMF group were due to over-immunosuppression compared to two deaths in the Neoral/MMF group. These results are further explained by findings from the clinical pharmacology review (see reviews by Dr. Seong Jang, PhD) suggesting that the mean daily MMF dose received in the Prograf group was consistently less than the target 1 gram/twice daily approved dose. Although MMF doses were lower than 2 g/day in the Prograf

group, the measured MPA (active moiety of MMF) trough concentrations for Prograf were consistently higher than Neoral concentrations at 30, 192 and 365 days (Table 3.6 & Table 3.7). Specifically, MPA trough concentrations in Prograf group were approximately 76% higher at one month and 29% higher at 12 months than in the Neoral arm. A finding of higher incidence of serious infection in the Prograf/MMF, with some infections associated with mortality, suggests a relationship between higher MPA exposure and over-immunosuppression (refer to medical review for detailed discussion).

Study 02-0-158 demonstrated that the incidence of BCAR was significantly lower in the Prograf/MMF group compared to the Neoral/MMF group contributing to an overall finding of non-inferiority with respect to efficacy failure in the Prograf/MMF group. The numerically higher number of deaths observed at 1 and 2-years follow-up in the Prograf/MMF group compared to the Neoral/MMF control group raises concern regarding the safety of the Prograf/MMF regimen. Appropriate doses of MMF to be given with Prograf in *de novo* kidney transplantation can not be accurately assessed from these data since a safe and effective MMF dose was not assessed.

6. REFERENCES

The Canadian Multicenter Transplant Study Group. A randomized clinical trial of cyclosporine in cadaveric renal transplantation. *N Engl J Med* 1983;309(14):809-15.

Asberg A, Midtvedt K, Line PD, Narverud J, Holdaas H, Jenssen T, et al. Calcineurin inhibitor avoidance with daclizumab, mycophenolate mofetil, and prednisolone in DR-matched de novo kidney transplant recipients. *Transplantation* 2006;82(1):62-8.

Meier-Krieschea HU, Li S, Gruessnerc RWG, Fung JJ, Bustamib RT, Barre ML and Leichtman AB. Immunosuppression: Evolution in Practice and Trends, 1994–2004. *Am J Trans*. 2006; 6 (Part 2): 1111–1131

Vincenti F, Ramos E, Brattstrom C, Cho S, Ekberg H, Grinyo J, et al. Multicenter trial exploring calcineurin inhibitors avoidance in renal transplantation. *Transplantation* 2001;71(9):1282-7.

Tran HT, Acharya MK, McKay DB, Sayegh MH, Carpenter CB, Auchincloss H, et al. Avoidance of cyclosporine in renal transplantation: Effects of daclizumab, mycophenolate mofetil, and steroids. *J Am Soc Nephrol* 2000;11:1903-9

SIGNATURES/DISTRIBUTION LIST

Primary Statistical Reviewer: LaRee Tracy, M.A.
Date: March 14, 2007

Concurring Reviewers:
Statistical Team Leader: Karen Higgins, Sc.D.

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

/s/

LaRee Tracy
3/14/2007 11:16:22 AM
BIOMETRICS

Karen Higgins
3/14/2007 11:26:12 AM
BIOMETRICS