CENTER FOR DRUG EVALUATION AND RESEARCH

APPROVAL PACKAGE FOR:

APPLICATION NUMBER

BLA 125103/0

Statistical Review(s)

12-2-04

STATISTICAL REVIEW AND EVALUATION

DEPARTMENT OF HEALTH AND HUMAN SERVICES
PUBLIC HEALTH SERVICE
FOOD AND DRUG ADMINISTRATION
CENTER FOR DRUG EVALUATION AND RESEARCH

Medical Division: Biologic Oncology Drug Products (HFD-107)
Biometrics Division: Biologic Therapeutics Statistical Staff (HFD-711)

STATISTICAL KEY WORDS:	**Generalized Cochran-Mantel-Haenszel statistic (CMH)**
NDA NUMBER:	**BLA125103**
DRUG NAME:	**Kepivance™ (palifermin)**
INDICATION:	**Oral mucositis**
SPONSOR:	**Amgen**
STATISTICAL REVIEWER:	**Yuan-Li Shen, Dr. P.H. (HFD-711)**
STATISTICAL TEAM LEADER:	**Mark Rothmann, Ph.D. (HFD-711)**
BTSS DIRECTOR:	**Aloka Chakravarty, Ph.D. (HFD-711)**
CLINICAL REVIEWERS:	**Patricia Dinndorf, M.D. (HFD-107)**
CLINICAL TEAM LEADER:	**Joseph Gootenberg, M.D. (HFD-107),**
PROJECT MANAGER:	**Susan Giuliani (HFD-109)**

Distribution: BLA 125123
HFD-107/Dinndorf
HFD-107/Gootenberg
HFD-109/Giuliani
HFD-711/Shen
HFD-711/Rothmann
HFD-711/Chakravarty
HFD-700/Anello

File Directory: C:/BLA_2004/KGF/Palifermin_statreview_2004.doc

Table of Contents

Appears This Way
On Original

Table of Contents

STATISTICAL REVIEW AND EVALUATION

1 Executive Summary of Statistical Findings

The sponsor is seeking labeling claims based on clinical outcomes and patient-reported outcomes. This review provides a summary of the clinical efficacy results, statistical issues and an overview of the studies submitted in this application for palifermin. Dr. Lisa Kammerman will evaluate the patient-reported outcomes.

1.1 Recommendations and Conclusions

Based on review of the Phase III (study 2000162) and Phase II ((study 980231) randomized, placebo-controlled study results, both results support the efficacy of palifermin by demonstrating a shorter duration of WHO grade 3 or 4 oral mucositis in palifermin group as compared with the placebo group (p-value <0.001 for the Phase III study and p-value =0.004 for the Phase II study).

In study 20000162, the sponsor also demonstrated a statistical significance in two key efficacy endpoints: use of parenteral or transdermal opioid analgesics and duration of WHO grade 2, 3, and 4 oral mucositis. However, the results from Study 980231 do not provide confirmatory evidence due to exploratory nature of the analyses.

1.2 Brief Overview of Clinical Studies

The sponsor submitted two study results (Phase III study : 20000162 and Phase II study 980231) for patients with hematologic malignancies undergoing Total Body Irradiation (TBI) and high-dose chemotherapy followed by autologous peripheral blood progenitor cell (PBPC) transplantation to provide evidence of efficacy and safety for palifermin.

Both studies were randomized, double-blind, placebo-controlled trials. Both studies were based on stratified randomization: Study 2000162 was randomized (stratified by center and hematologic malignancy) in a 1:1 ratio to receive placebo or palifermin 60 μg/kg/day IV for 3 consecutive days before the conditioning regimen began and for 3 consecutive days after PBPC transplant ('pre-post'). Study 980231 was randomized (stratified by center) to 3 dose regimens:

- 'Pre-post': palifermin 60 μg/kg/day IV for 3 consecutive days before the conditioning regimen began and for 3 consecutive days after PBPC transplant
- 'Pre-': palifermin 60 μg/kg/day IV for 3 consecutive days before the conditioning regimen began and placebo 60 μg/kg/day for 3 consecutive days after PBPC transplant
- placebo: placebo 60/kg/day IV for 3 consecutive days before conditioning regimen and for 3 consecutive days after PBPC transplant

Originally, the dosing schedule included one additional dose at day -5 after TBI. After the July 9, 1999 safety monitoring committee meeting, the committee recommended the

termination of the day -5 dose for possible increasing risk of oral mucositis. Thirty four subjects in study 980231 received this additional dose at day -5 after TBI.

For consistency of the review, this review will focus on review of the 6-dose schedule. The results from the 7-dose schedule will be discussed if deemed appropriate.

The number of patients randomized, treated and early discontinued of study medication is summarized in the following table:

Table 1 Summary of the Number of Patients

Study	Duration	Population	Placebo	Palifermin (Pre)	Palifermin (Pre-post)
20000162	3/23/01-10/23/02	Randomized	107	NA	107
		Treated	106		106
		Treatment discontinuation	2		3
		Study discontinuation	3		2
980231	2/23/99 – 7/024/00	**6-dose schedule :**			
		Randomized	40	46	48
		Treated	40	43	46
		Treatment discontinuation	0	1	3
		Study discontinuation	0	1	2
		7-dose schedule :			
		Randomized	11	11	12
		Treated	11	11	12
		Treatment discontinuation	0	2	0
		Study discontinuation	1	0	0

The primary efficacy endpoint from both studies is the duration of WHO grade 3 or 4 oral mucositis. It was analyzed by the generalized Cochran-Mantel Haenszel (CMH) method stratified by pooled center and incorporated the standardized mid-ranks (also known as the modified Ridit scores) in the test statistic.

1.3 Statistical Issues and Findings

In general, the primary efficacy results from two studies are robust after independent evaluation based on different statistical methods and with or without any imputation schemes. There are only a few statistical issues related to the analysis:

- The highly skewed data caused by the larger zero duration in the palifermin group may create analytical challenge. However, the generalized Cochran-Mantel Haenszel method appears to be adequate.
- Using mean duration to impute those missing data due to early discontinuation may not be the most optimum method. However, due to the small percentages of patients who had early withdrawal or die in these studies, the impact of the imputation is not felt to be important.
- Some of the pre-irradiation treatment related AE (e.g., skin, facial, or oral-related AE) may potentially unblind the results. However, there is not enough data in the current trials to demonstrate the effect.
- The lack of pre-specify rules of dealing with comparison of multiple arms and multiple endpoints in study 980231 resulted in difficulties of interpreting the results of the secondary efficacy endpoints.

2 Introduction

This section provides an overview of the submitted trials.

2.1 Overview

This subsection provides a background of the design of the submitted trials, the data analyzed and the source, and any major statistical issues.

2.1.1 Background

The sponsor submitted the results from one phase III and one phase II study to compare the palifermin and placebo in patients who receive the conditioning regimen (consisting of fractionated total-body irradiation [TBI], total-dose etoposide, and high-dose cyclophosphamide) followed by peripheral blood progenitor cell (PBPC) support for the treatment of hematological malignancies (non-Hodgkin's hymphoma [NHL], Hodgkin's disease, acute myeloid leukemia [AML], acute lymphoblastic leukemia [ALL], chronic myeloid leukemia [CML], chronic lymphocytic leukemia [CLL], or multiple myeloma).

This submission provides safety and efficacy evidence of palifermin in reducing the incidence, duration, and severity of oral mucositis and related sequelae (mouth and throat soreness, use of opioid analgesics, and to improve patient functioning (ability to swallow, eat, drink, talk, and sleep) in patients with hematologic malignancies that requires PBPC.

The primary efficacy endpoint of both studies is the duration of WHO grade 3 or 4 mucositis. The primary comparison of interest compares placebo group and the palifermin pre-post group.

2.1.2 Major Statistical Issues

Several statistical issues with respect to the analysis are summarized below:

- Since many patients in the palifermin group did not develop WHO grade 3 or 4 mucositis, the duration of WHO grade 3 or 4 contains many zero duration which may cause skewness of the data, more in the palifermin group than in the placebo group (2% and 37% for placebo group and palifermin group, respectively, for study 20000162; and 20%, 28% and 33% for placebo and palifermin pre- and pre-post groups, respectively, for study 980231). However, the generalized Cochran-Mantel Haenszel method appears to be adequate in handling such situation.

- Using mean duration to impute the data for the patients who had unresolved WHO grade 3 or 4 by the time of early withdrawal or death may not be the optimum method and may cause reduction of the variability. However, given the small percentages of patients with early discontinuation, the impact of the imputation scheme may not be important.

- Some of the pre-irradiation treatment related AE (e.g., skin, facial, or oral-related AE) may potentially unblind the results. However, there is not enough data in the current trials to demonstrate the effect.

- The lack of prespecified rules of dealing with comparison of multiple arms and multiple endpoints in study 980231 resulted in difficulties of interpreting the results of the secondary efficacy endpoints.

2.2 Data Sources

Data used for review is from the electronic submission received on 6/15/04. The network path is \\cdsesub1\evsprod\bla125103".

The efficacy analysis results will be presented in this section for protocols 20000162 and 980231, respectively.

3 Statistical Evaluation

The efficacy analysis results will be presented in this section for protocols 20000162 and 980231, respectively.

3.1 Evaluation of Efficacy

3.1.1 Study 20000162

This subsection will present the efficacy evaluation for study 20000162. This

includes the background information, efficacy endpoints, sample size determination, the efficacy methods used, and the statistical findings.

3.1.1.1 Introduction

Study 20000162 is a phase 3, randomized, double-blind, placebo-controlled trial of recombinant human keratinocyte growth factor (rHuKGF) for reduction of mucositis in patients with hematologic malignancies undergoing total body irradiation (TBI) and high-dose chemotherapy with autologous peripheral blood progenitor cell (PBPC) transplantation.

Patients received 3 consecutive daily doses of palifermin or placebo 60 μg/kg/day IV before the myeloablative conditioning regimen (TBI plus high-dose chemotherapy) and 3 consecutive doses after PBPC transplant ('pre-post'). Filgrastim (60 μg/kg/day) was administered from day 0 after transplant until neutrophil recovery (absolute neutrophil count [ANC] > 1.0×10^9 for 3 consecutive days or > 1.0×10^9 for 1 day, or day 21, whichever occurred first).

The study day was numbered with respect to the starting day of the stem cell transplantation, which is defined as day 0. This study includes a pre-study and a treatment phase including the post-transplantation observation up to the end of the study (day 28).

Starting study day -8, oral cavity assessment was performed daily during hospitalization by one of the trained assessors at the site. Oral mucositis assessments were performed daily during the study using the WHO, Radiation Therapy Oncology Group (RTOG), and Western Consortium of cancer Nursing Research (WCCNR) rating scales. Daily oral mucositis assessment was not discontinued before the oral mucositis grade returns to a WHO grade ≤ 2.

Patient randomization was stratified by center and type of hematologic malignancy (non-Hodgkin's lymphoma, Hodgkin's disease, leukemia [acute or chronic], or multiple myeloma). The randomization was performed centrally by an Interactive Voice Response System (IVRS) vendor. All patients were to be on study from the first day of the study medication to the end of study visit.

The primary objective of this study is to evaluate the efficacy of palifermin in reducing the duration of severe oral mucositis (WHO grade 3 or 4) induced by total body irradiation (TBI) and high-dose chemotherapy in subjects with hematologic malignancies. The secondary objectives of this study are to evaluate:

- The efficacy of palifermin in

    1. Reducing oro-pharyngeal mucositis

2. Reducing sequelae of mucositis

- The safety of palifermin.

Other than a regular safety monitoring, there is no formal interim analysis planned for this study.

3.1.1.2 Efficacy Endpoints

The primary efficacy endpoint for this study was the duration of severe oral mucositis (WHO grades 3 or 4). This duration was calculated based on the number of days a subject experienced WHO grade 3 to 4 mucusitis.

If a subject's mucositis is WHO grade 3 or 4 on the day of discharge (last day of daily assessment), the subject will be followed up until his/her oral mucositis resolved to WHO grade 2 or less. For subjects who dies or withdraws from study, resolution of severe mucositis is defined as observing two consecutive assessments of WHO grade 2 or less after the last reading of a WHO grade 3 or 4. The date of resolution will be the first of the two consecutive days with WHO mucositis grade 2 or less.

The grading of the WHO toxicity criteria for oral mucositis is described in the following table.

Table 2 WHO Toxicity Criteria for Oral Mucositis

WHO Grade				
0	1	2	3	4
None	Soreness and erythema	Erythema, ulcers, ability to eat solids	Ulcers, requires liquid diet	Alimentation not possible

For this primary efficacy endpoint, the imputation scheme was indicated in the analysis plan. If a subject did not experience any WHO grade 3 or 4 oral mucositis, the duration of severe oral mucositis was assigned as 0 days.

If a subject's mucositis is not resolved upon early withdrawal or death, he/she was given the mean duration of severe mucositis (WHO grade 3 or 4) among subjects who experienced at least the same duration of severe mucositis as this subject.

When there was a gap between 2 observed grades, the missing grades were assigned the larger adjacent grade. The same rule was applied to gaps after a patient is discharged with unresolved WHO grade 3 or 4 mucositis (but the follow-up assessments were not done daily).

In addition, the sponsor also conducted "worst case" analysis to demonstrate the robustness of the results. In the "worst case" analysis, the missing grades were replaced with the higher adjacent value in the palifermin group, and the missing grades in the placebo group were replaced with the lower adjacent values.

The secondary efficacy endpoints include the following assessment

- Patient Report Outcome (PRO) assessment of mouth and throat soreness
- Use of parenteral or transdermal opioid analgesics (in mg morphine equivalents)
- Incidence of WHO grade 4 oral mucositis
- Duration (days) of WHO grade 2, 3, and 4 oral mucositis
- Duration (days) of oral mucositis based on WCCNR lesion domain, grades 2 and 3, and RTOG rating, grades 3 and 4

The total dose of analgesic was computed based on daily doses (reported doses converted to morphine equivalent doses [see Analysis Plan section 9.5.2.3 Endpoint Calculations and Derivations for equivalent dose conversion] and adjusted based on the dosing frequency) multiplied by the number of days on the analgesic. For any missing analgesic specific total dose due to missing dosing record, the midpoint imputation (average of the 2 adjacent morphine equivalent doses) was used to impute the missing data. These derived doses were then added up over the duration of the study for each patient. For subjects whose records of analgesics use are incomplete due to death or withdrawal, they were given the worst amount of analgesic used observed among subjects with the same type of hematologic disease regardless of the treatment assignments.

The incidence of WHO grade 4 oral mucositis is calculated based on the number of patients who had at least one of the WHO grade 4 oral mucositis over the course of daily assessments divided by the number patients in each treatment group.

The calculation of the duration of WHO grade 2, 3, and 4 oral mucositis is similar to the calculation of the duration of WHO grade 3 or 4. The imputation scheme for these two endpoints are also similar except that for calculation of the duration of WHO grade 2, 3, and 4 mucositis, if a patient's mucositis is not resolved (to less than grade 2) on the day of discharge, linear interpolation will be done for the days between the discharge day assessment and the end of study day assessment.

The rules described for the duration of WHO grade 3, 4 mucositis were applied to the calculation of duration of grade 3 or 4 RTOG and duration of grade 2 or 3 WCCNR lesion domain.

The patient report outcome assessment will be evaluated by Dr. Lisa Kammerman.

3.1.1.3 Sample Size Consideration

Approximately 210 subjects were planned for this study. The null hypothesis for this study is that the palifermin and placebo groups had the same duration (in days) of severe oral mucositis. The alternative hypothesis is that the duration of severe oral mucositis in the two treatment groups will be different. With 105 subjects per group, this sample size was calculated based on 2 sample t-test to detect a 3-day difference on WHO grade 3 or 4 oral mucositis between palifermin and placebo groups with 90 % power, assuming the standard deviation of the duration is 6.6 days and 5% significance level (2-sided).

The sponsor indicates that the primary analysis method, though different from the t-test approach, is not expected to have lower power since the stratified analysis is expected to be more efficient.

In the sponsor's analysis plan, it stated that during the study, the standard deviation for the duration of severe mucositis (WHO grade 3 or 4) will be calculated periodically (approximately every 20 subjects) to examine the assumption of 6.6 days used in the sample size calculation. If the calculated standard deviation is at least 1 day larger than that used in the assumption, the sample size will be re-estimated. However, based on this rule, the sample size has not been re-estimated during the course of the study.

3.1.1.4 Efficacy Analysis Method

The generalized Cochran-Mantel Haenszel (CMH) method stratified by pooled center was used for the primary efficacy analyses. Centers with less than six subjects were pooled into one 'super' center. The standardized mid-rank (also known as the modified Ridit scores) was used for the test statistic. This generalized CMH test follows a Chi-square distribution with 1 degree of freedom.

The primary efficacy analysis dataset was based on the modified intend-to-treat (mITT) population defined as all randomized subjects who receive at least one dose of study drug and who are analyzed according to their randomized treatment assignment.

Three sensitivity analyses were proposed for the imputation of the grade 3 or 4 mucositis in patients who have died or withdrawn from the study in the analysis plan :

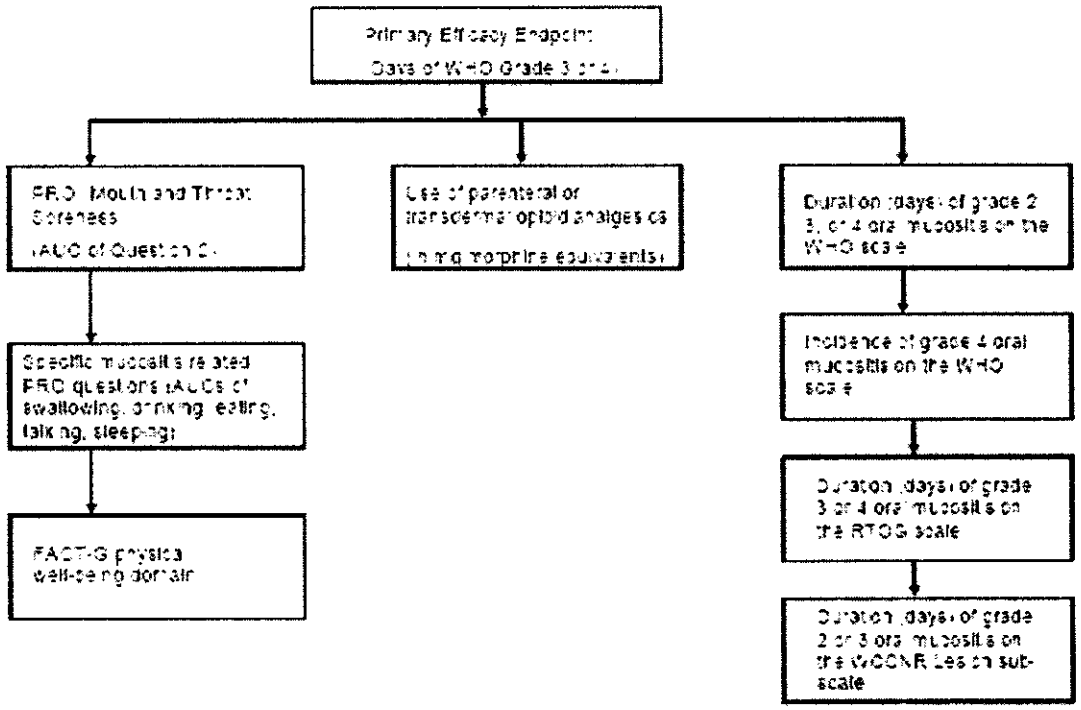
- Patients with missing values should be censored.
- The duration of grade 3 or 4 mucositis in patients with missing values should extend to day 28.
- The duration of grade 3 or 4 mucositis in patients with missing values should be assigned with the longest duration of grade 3 or 4 mucositis observed on study.

However, since only handful of patients had early discontinuation from the study (3 in placebo and 2 in palifermin group), the original proposal was not adopted. Instead, a

sensitivity analysis was performed (worst-case) by assigning worse adjacent value for subjects who received palifermin and better adjacent value for the subjects who received placebo.

To control the type 1 error rate, the sponsor's analysis plan indicated that the primary treatment comparison will be performed between the palifermin group and the placebo group at the 2-sided 0.05 significance level. After the primary efficacy endpoint was found to be statistical significant, 3 secondary efficacy endpoints were to be tested simultaneously (each at 0.05 significance level) – mouth and throat soreness (Patient Reported Outcome [PRO]); use of parenteral or transdermal opioid analgesics; and duration of WHO grade 2, 3, and 4 oral mucositis. Once a significant result was obtained for the duration of WHO grade 2, 3, and 4, a sequence of secondary efficacy endpoints were performed (see the following figure). Similarly, once a significant result in mouth and throat soreness was found, a test of subsequent secondary endpoints was performed (see the following figure).

Figure 1 Sponsor's Summary of Testing Relationship of all Endpoints



3.1.1.5 Sponsor's Results and Statistical Reviewer's Findings/ Comments

A summary of subject disposition is presented in the following table. Among all randomized patients, only 1 patient in each treatment group did not take study medication. Over 97% of the patients completed the study. About 3% and 2% of the patients for placebo and palifermin groups withdrew from the study prematurely.

Table 3 Sponsor's Summary of Primary Analysis Subset (Study 20000162)

	Placebo n	Palifermin n
Subjects Randomized	107	107
mITT Population	106	106
Per-Protocol Efficacy Analysis Population	77	88
Safety Analysis Population	106	106

Table 4 Sponsor's Summary of Subject Disposition (Study 20000162)

	Placebo n (%)	Palifermin n (%)	Total n (%)
Subjects randomized	106	106	212
Test Article Accounting			
Subjects who never received test article	0 (0)	0 (0)	0 (0)
Subjects who received test article	106 (100)	106 (100)	212 (100)
Subjects who completed test article	104 (98)	103 (97)	207 (98)
Subjects who discontinued test article	2 (2)	3 (3)	5 (2)
Study Completion Accounting			
Subjects who completed study	103 (97)	104 (98)	207 (98)
Subjects who discontinued study	3 (3)	2 (2)	5 (2)

3.1.1.5.1 Baseline Characteristics

Summaries of patients' demographic information and baseline characteristics are presented in the following 3 tables. The distribution of all demographic information (gender, race, age, height, weight, Karnofsky Performance Status) appears to be compatible. The study population includes more men than women in either treatment group. Over 80% of the patients had a score above 90 of the Karnosky performance status (i.e. able to carry on normal activity: minor symptoms of disease).

Table 5 Sponsor's Summary of Demographic Information (Study 20000162)

	Placebo (N = 106)	Palifermin (N = 106)
Sex – n (%)		
Men	72 (68)	59 (56)
Women	34 (32)	47 (44)
Race/Ethnicity – n (%)		
White	89 (84)	78 (74)
Black	7 (7)	11 (10)
Hispanic	7 (7)	11 (10)
Asian	1 (1)	4 (4)
Japanese	0 (0)	0 (0)
Native American	0 (0)	1 (1)
Native Hawaiian	0 (0)	1 (1)
Other	2 (2)	0 (0)
Age (years)		
N	106	106
Mean (SD)	46 (12)	46 (12)
Median (Min, Max)	49 (19, 68)	48 (18, 69)
Q1, Q3	38, 54	37, 57
Height (cm)		
n	106	106
Mean (SD)	174 (10)	171 (9.0)
Median (Min, Max)	175 (147, 195)	170 (150, 188)
Q1, Q3	166, 181	165, 178
Weight (kg)		
n	105	106
Mean (SD)	87 (18)	84 (20)
Median (Min, Max)	83 (46, 164)	82 (48, 149)
Q1, Q3	73, 97	69, 96
Karnofsky Performance Status (%) - n(%)		
70	1 (1)	3 (3)
80	19 (18)	15 (14)
90	58 (55)	59 (56)
100	28 (26)	29 (27)

Most of the baseline characteristics also seem to be compatible. Majority of the patients are Non-Hodgkin's Lymphoma patients (65% and 68% for the placebo and palifermin group, respectively). CD34+ cells were mobilized with cytokines and chemotherapy for over 70% of the patients. The reason of PBPC transplantation for most patients is for chemotherapy sensitive relapse (31% and 37% for the placebo and palifermin group, respectively). All patients had prior chemotherapy. Also, over 85% of the patients did not receive prior radiotherapy.

Table 6 Sponsor's Summary of Baseline Disease Characteristics (Study 20000162)

	Placebo (N=106)	**Palifermin** (N=106)
Type of Diagnosis -n(%)		
Hodgkin's Disease	23 (22)	21 (20)
Non-Hodgkin's Lymphoma	69 (65)	72 (68)
Multiple Myeloma	9 (8)	11 (10)
Leukemia	5 (5)	2 (2)
Mobilization -n(%)		
Cytokines only	30 (28)	26 (25)
Chemotherapy only	0 (0)	1 (1)
Cytokines and Chemotherapy	76 (72)	79 (75)
Total Number of CD34+ Cells		
n	106	106
Mean (SD)	7.0 (7.8)	8.6 (12)
Median (Min, Max)	5.0 (1.5, 41)	5.2 (1.8, 87)
Q1, Q3	3.1, 7.3	3.4, 7.4
Reason for PBPC Transplantation -n(%)		
Consolidation	9 (8)	4 (4)
First complete remission	22 (21)	15 (14)
Second complete remission	2 (2)	3 (3)
First partial remission	12 (11)	19 (18)
Chemotherapy sensitive relapse	33 (31)	39 (37)
Induction failure	0 (0)	4 (4)
Primary refractory disease	24 (23)	20 (19)
Other	4 (4)	2 (2)
The Current Transplantation is the Second in a Tandem Transplantation Regimen - n(%)		
Yes	1 (1)	0 (0)

No	105 (99)	106 (100)
Subjects with Prior Chemotherapy -n(%)		
Yes	106 (100)	106 (100)
Subjects with Prior Radiotherapy -n(%)		
Yes	9 (8)	13 (12)
No	97 (92)	93 (88)

3.1.1.5.2 Primary Efficacy Endpoint Analyses

The sponsor's primary efficacy endpoint analysis results are summarized in the following table. The median duration of WHO Grade 3 or 4 mucositis was 9 and 3 for placebo and palifermin treated groups, respectively (p-value <0.001). The palifermin treated group had a significant lower duration of the WHO Grade 3 or 4 mucositis. Such trend was also demonstrated in the group of patients who developed the WHO Grade 3 or 4 mucositis.

Table 7 Sponsor's Summary of the Primary Efficacy Endpoint – WHO Grade 3 or 4 Mucositis (Study 20000162)

	Placebo (N=106)	**Palifermin** (N=106)	**P value** (CMH test [a])
Duration of Severe Mucositis days (WHO grade 3,4)			<0.001
n	106	106	
Mean (SD)	10.4 (6.2)	3.7 (4.1)	
Median (Min, Max)	9.0 (0.0, 27.0)	3.0 (0.0, 21.7)	
Q1, Q3	6.0, 13.0	0.0, 6.0	

[a] Generalized Cochran-Mantel-Haenszel test based on the standardized mid-ranks (modified Ridit scores) within each stratum.

Table 8 Duration (days) of WHO Grade 3 or 4 Mucositis for Subjects Who Developed Such Toxicity During Study (Study 20000162)

	Placebo (N=106)	**Palifermin** (N=106)	**P value** (CMH test [a])
Duration of Severe Mucositis (WHO Grade 3,4) (days)			<0.001
n	104	67	
Mean	10.6 (6.1)	5.9 (3.6)	
Median (Min, Max)	9.0 (1.0, 27.0)	6.0 (1.0, 21.7)	
Q1, Q3	6.0, 13.0	3.0, 8.0	

[a] Generalized Cochran-Mantel-Haenszel test based on the standardized mid-ranks (modified Ridit scores) within each stratum.

The treatment difference in the WHO Grade 3 or 4 mucositis was also significant based on a sensitivity analysis results using the "worst-case" scenario (see the following table).

Table 9 Sponsor's Summary of Primary Efficacy Endpoint – WHO Grade 3 or 4 Mucositis (Worst case) (Study 20000162)

	Placebo (N=106)	Palifermin (N=106)	P value (CMH test [a])
Duration of Severe Mucositis (WHO Grade 3,4) (days)			<0.001
n	106	106	
Mean (SD)	9.4 (5.1)	3.7 (4.0)	
Median (Min, Max)	8.0 (0.0, 27.0)	3.0 (0.0, 20.0)	
Q1, Q3	6.0, 12.0	0.0, 6.0	

[a] Generalized Cochran-Mantel-Haenszel test based on the standardized mid-ranks (modified Ridit scores) within each statum.

3.1.1.5.3 Secondary Efficacy Endpoint Analyses

The sponsor's summary of cumulative dose of opioid analgesic use for mucositis indicates that there was a significant difference in opioid analgesic doses. The palifermin treated group had smaller total dose of opioid analgesic uses (median = 211.6 mg) as compared with the placebo treated group (median= 534.9mg). The palifermin treated group also had shorter duration of analgesic use for mucositis (median days on analgesic for mucositis is 8 days for the palifermin group and 14.3 days for placebo group).

Table 10 Sponsor's Summary of Cumulative Dose of Opioid Analgesic Use for Mucositis (Subset of Subjects with Opioid Analgesic Use for Mucositis) (Study 20000162)

	Placebo (N=106)	Palifermin (N=106)	P value (CMH test [a])
Dose of Opioid Analgesics (mg)			<0.001
n	106	106	
Mean (SD)	1146.5 (1702.1)	699.5 (1749.3)	
Median (Min, Max)	534.9 (0.0, 9418.4)	211.6 (0.0, 9418.4)	
Q1, Q3	268.7, 1429.0	3.0, 558.4	

[a] Generalized Cochran-Mantel-Haenszel test based on the standardized mid-ranks (modified Ridit scores) within each stratum.

Table 11 Sponsor's Summary of Duration of Opioid Analgesic Use for Mucositis (Study 20000162)

	Placebo (N=106)	Palifermin (N=106)	P value (CMH test [a])
Days of Opioid Analgesic Use			
n	106	106	<0.001
Mean (SD)	11.8 (5.6)	6.7 (5.5)	
Median (Min, Max)	11.0 (0.0, 32.0)	7.0 (0.0, 23.0)	
Q1, Q3	8.0, 14.0	1.0, 10.0	

[a] Generalized Cochran-Mantel-Haenszel test based on the standardized mid-ranks (modified Ridit scores) within each stratum.

The sponsor's summary of duration of WHO Grade 2, 3 or 4 mucositis (see the following table) also shows that the palifermin group had shorter duration of milder mucositis.

Table 12 Sponsor's Summary of Duration of Mucositis: WHO Grade 2, 3 or 4 (Study 20000162)

	Placebo (N=106)	Palifermin (N=106)	P value (CMH test [a])
Duration of WHO Grade 2, 3, 4 mucositis (days)			<0.001
n	106	106	
Mean (SD)	15.7 (7.8)	8.4 (5.8)	
Median (Min, Max)	14.3 (0.0, 37.0)	8.0 (0.0, 28.0)	
Q1, Q3	11.0, 19.0	4.0, 12.0	

[a] Generalized Cochran-Mantel-Haenszel test based on the standardized mid-ranks (modified Ridit scores) within each stratum.

In addition, the sponsor shows significant treatment differences in incidence and duration of WHO Grade 4 mucositis.

Table 13 Sponsor's Summary of Incidence and Duration of WHO Grade 4 Mucositis (Study 20000162)

	Placebo (N=106)	Palifermin (N=106)	P value (CMH test [a])
Duration of Mucositis (WHO Grade 4) (days)			
n	106	106	<0.001
Mean (SD)	3.9 (5.1)	0.7 (1.7)	
Median (Min, Max)	2.0 (0.0, 37.0)	0.0 (0.0, 9.0)	
Q1, Q3	0.0, 6.0	0.0, 0.0	

Incidence of WHO Grade 4 Mucositis – n (%)	66 (62)	21 (20)	<0.001

[a] Generalized Cochran-Mantel-Haenszel test based on the standardized mid-ranks (modified Ridit scores) within each stratum.

The significant treatment difference in duration based on WHO grade was also confirmed in favor of the palifermin treated group based on RTOG grades and WCCNR grades (see the next 2 tables).

Table 14 Sponsor's Summary of Duration of Severe Mucositis (RTOG Grade 3 or 4) (Study 20000162)

	Placebo (N=106)	**Palifermin** (N=106)	**P value** (CMH test [a])
Duration of Severe Mucositis (RTOG Grade 3,4) (days)			<0.001
n	106	106	
Mean (SD)	8.1 (8.5)	2.1 (3.5)	
Median (Min, Max)	6.0 (0.0, 54.0)	0.0 (0.0, 24.0)	
Q1, Q3	3.0, 11.0	0.0, 4.0	

[a] Generalized Cochran-Mantel-Haenszel test based on the standardized mid-ranks (modified Ridit scores) within each stratum.

Table 15 Sponsor's Summary of Duration of Severe Mucositis (WCCNR Grade 2 or 3) (Study 20000162)

	Placebo (N=106)	**Palifermin** (N=106)	**P value** (CMH test [a])
Duration of WCCNR Grade 2 or 3 Lesion (days)			
n	106	106	<0.001
Mean (SD)	9.9 (9.3)	3.2 (5.2)	
Median (Min, Max)	7.0 (0.0, 56.0)	1.0 (0.0, 36.0)	
Q1, Q3	4.0, 13.2	0.0, 5.0	

[a] Generalized Cochran-Mantel-Haenszel test based on the standardized mid-ranks (modified Ridit scores) within each stratum.

3.1.1.6 Sponsor's Conclusions and Reviewer's Conclusions/Comments

Based on the sponsor's analysis, the median (25^{th} percentile, 75^{th} percentile) duration of WHO Grade 3 or 4 mucositis was 3.0 (0, 6) days for the palifermin group which is significantly shorter than 9 (6, 13) days for the placebo group (p-value <0.001). Similar results (median [25^{th} percentile, 75^{th} percentile] duration was 8.0 (6, 12) and 3.0 (0, 6) for the placebo and palifermin, respectively) were observed in sensitivity analyses using worst-case imputation rules. These significant shorter duration of WHO grade 3 or 4 mucositis was supported by other criteria, including the duration of WHO Grade 2, 3, 4 mucositis (for less severe mucositis), WCCNR grade 2 or 3 mucositis and RTOG Grade 3 or 4 mucositis.

This reviewer performed two sensitivity analyses to independently validate the sponsor's results. Since there are only handful of patients with early withdrawal (3% and 2% for the placebo and palifermin group respectively) and a few missing visits that require imputation (4% and 0.5% among all potential numbers of days of evaluation for placebo and palifermin group, respectively) for grade 3 or 4, the impact of the imputation may not be important. In the sensitivity analyses, this reviewer did not implement any imputation scheme. In addition to counting days for WHO grade 3 or 4 mucositis, this reviewer computed the duration of mucositis based on the time span between the first date when the mucositis became grade 3 or 4 to the last date when the mucositis was still of grade 3 or 4.

Since there are about 2% and 37% subjects having zero duration (i.e. not developing grade 3 or 4 mucositis) for the placebo and palifermin group, respectively, distribution-free statistical method such as the CMH method that the sponsor proposed and Wilcoxon rank sum test that does not require distribution assumption beyond randomization requirement would be more appropriate. The results are shown in the following table. It appears to be close to the results that the sponsor provided.

Table 16 Reviewer's Summary of Duration of Severe Mucositis (WHO Grade 3 or 4) (Study 20000162)

Duration of WHO Grade 3 or 4	Placebo (N=106)	Palifermin (N=106)	P value [a]
1. Counting days of Grade 3 or 4			
n	106	106	<0.001
Mean (SD)	9.3 (5.0)	3.5 (3.7)	
Median (Min, Max)	8.0 (0.0, 27.0)	3.0 (0.0, 13.0)	
Q1, Q3	(6, 12)	(0, 6)	
2. Counting duration of Grade 3 or 4 [b]			
n	106	106	<0.001
Mean (SD)	10.3 (6.1)	3.8 (4.1)	
Median (Min, Max)	9.0 (0.0, 29.0)	3.0 (0.0, 19.0)	
Q1, Q3	(6, 13)	(0, 7)	

[a] Both Generalized Cochran-Mantel-Haenszel test based on the standardized mid-ranks (modified Ridit scores) within each stratum and Wilcoxon rank sum test were performed. Both p-values were less than 0.001.

[b] Duration of mucositis was computed based on the time span between the first date when the mucositis became grade 3 or 4 to the last date when the mucositis grade was still 3 or 4

Based on the same strategy, similar results were obtained for analyses of duration of grade 2, 3, or 4 on the WHO scale, duration of grade 3 or 4 on the RTOG scale and duration of grade 2 or 3 on WCCNR subscale Grade.

To evaluate the distribution of the WHO Grade, the sponsor summarized the worst WHO grade over time for each patient. The results indicate that despite the palifermin

group having lower incidence in WHO Grade 4 or Grade 3 or 4 mucositis (i.e. more severe mucositis), the palifermin group seems to pick up more in Grade 2 mucositis. It shows that over 90% of the patient population, either the placebo or palifermin treated, developed at least WHO Grade 2 mucositis. However, the palifermin treated group seems to have milder mucositis (Grade 2 or less).

Table 17 Sponsor's Summary of the Worst WHO Grades for Oral Mucositis (Study 20000162)

	Placebo (N=106)	**Palifermin** (N=106)	**P value** (CMH test [a])
Worst Score for WHO Grade Oral Mucositis			
Grade 4	66(62)	21(20)	
Grade 3	38(36)	46(43)	
Grade 2	1(1)	30(28)	
Grade 1	0(0)	8(8)	
Grade 0	1(1)	1(1)	
Worst Score for WHO Grade Oral Mucositis			
n	106	106	<0.001
Mean (SD)	3.6 (0.6)	2.7 (0.9)	
Median (Min, Max)	4.0 (0.0, 4.0)	3.0 (0.0, 4.0)	
Q1,Q3	3.0, 4.0	2.0, 3.0	

[a] Generalized Cochran-Mantel-Haenszel test based on the standardized mid-ranks (modified Ridit scores) within each stratum.

The sponsor's calculation of morphine equivalent doses are based on the following subsets among all analgesic taken:

- Fentanyl, morphine, hydromorphone or meperidine IV doses or Fentanyl patches
- Include non-prophylaxis use only
- Include oral mucositis or dysphagia only.

To confirm the sponsor's finding, similar analyses were performed. This analysis did not incorporate any imputation scheme for missing dosing information. The results are shown in the following table. The median, first and third quartile of the morphine use for mucositis appear to be similar to the sponsor's results. In addition, an analysis including all analgesic doses, regardless of indications, was performed. Again, a significant treatment difference in the morphine dose use was confirmed (p-value <0.001).

The medical officer addressed the concern to the sponsor about how the duration was counted for patients who wore 2 transdermal patches on the same day in a teleconference call on October 25, 2004. The sponsor agrees to perform a sensitivity analysis by assigning ½ day to each of the 2 patches. The sponsor sent in the results in

November 4, 2004. The sponsor found a total of 12 patients falling in this category (10 pacebo, 2 palifermin). This sensitivity analysis shows very similar results to those from the previous analysis.

Comments: there were 3% and 22 % zero morphine equivalent doses for the placebo and palifermin treated group, respectively. Due to the skewness of the data, a distribution method such as generalized CMH or Wilcoxon rank sum test is more appropriate.

Table 18 Reviewer's Summary of Cumulative Dose of Opioid Analgesic Use for Mucositis (Subset of Subjects with Opioid Analgesic Use for Mucositis)

	Placebo (N=106)	Palifermin (N=106)	P value [a]
Morphine equivalent dose			
n	106	106	<0.001
Mean (SD)	1049.3 (1487.3)	520.4 (1251.9)	
Median (Min, Max)	527.2 (0.0 , 9357.6)	211.6 (0.0 , 9418.4)	
Q1,Q3	(268.7, 1362.5)	(3, 516.0)	

[a] Both Generalized Cochran-Mantel-Haenszel test based on the standardized mid-ranks (modified Ridit scores) within each stratum and Wilcoxon rank sum test were performed. Both p-values were less than 0.001.

According to the sponsor's pre-specified decision rule, given a significant result on the duration of WHO grade 3 or 4 mucositis, 3 secondary efficacy endpoints were to be tested simultaneously (each at 0.05 significance level) – mouth and throat soreness (Patient Reported Outcome [PRO]); use of parenteral or transdermal opioid analgesics; and duration of WHO grade 2, 3, and 4 oral mucositis. Although the simultaneous testing at 0.05 significance level may not be satisfactory for controlling the type I error, if a more conservative testing scheme was applied (eg, Bonferroni), the significant findings in duration of WHO grade 2, 3, and 4 oral mucositis and the total doses of opioid analgesic use (note: PRO data will be reviewed by other reviewers) may still hold. These significant findings should be confirmed with an independent well controlled study with pre-specified rules.

According to sponsor's plan based on study 980231, the placebo group is expected to have approximately 70% to 80% WHO grades 3 or 4 oral mucositis. In this study, the WHO grades 3 or 4 mucositis (98%) in placebo group is much higher than expected, but the grades 3 or 4 mucositis rate in palifermin group is similar to the previous study (63% in study 20000162 and 67% in study 980231). This finding somewhat explains the finding in the duration of WHO grade 3 or 4.

3.1.2 Study 980231

This sub-subsection will present the efficacy evaluation for study 980231. This

includes the background, efficacy endpoints, sample size determination, the efficacy methods used, and the statistical findings.

3.1.2.1 Introduction

Study 980231 is a phase 2, randomized, double-blind, placebo-controlled trial of recombinant human keratinocyte growth factor (rHuKGF) for reduction of mucositis in patients with hematologic malignancies undergoing total body irradiation (TBI) and high-dose chemotherapy with autologous peripheral blood progenitor cell (PBPC) transplantation.

The original protocol was dated November 17, 1998. There were 2 amendments for this protocol : Amendment 1, dated May 27, 1999, allowed for the inclusion of patients 12 years and older, patients with multiple myeloma, and patients with a tandem-transplant regimen. Amendment 2, dated August 18, 1999, adopted a 6-dose treatment schedule by dropping the day -5 dose from the original 7-dose schedule. For the purpose of this statistical evaluation, the analysis will be focused on the 6-dose schedule. The analysis including the 7-dose schedule will be presented if deemed appropriate.

The randomization was stratified by center and was performed centrally by an Interactive Voice Response System (IVRS) vendor. Eligible subjects were randomly assigned at a 1:1:1 ratio into one of the three study arms:

- Six doses of placebo,
- Three doses rHuKGF 60 μg/kg/day followed by three doses of placebo (Pre-), or
- Six doses of rHuKGF 60 μg/kg/day (pre-post).

For the 6-dose schedule, patients received 3 consecutive daily doses of the palifermin or placebo 60 μg/kg/day IV before the myeloablative conditioning regimen (TBI plus high-dose chemotherapy) and 3 consecutive doses after PBPC transplant ('pre-post'). Filgrastim (60 μg/kg/day) was administered from day 0 after transplant until neutrophil recovery (absolute neutrophil count [ANC] > 1.0×10^9 for 3 consecutive days or > 1.0×10^9 for 1 day, or day 21, whichever occurred first). For the 7-dose schedule, patients received additional palifermin (pre- for pre-post arm) or placebo dose on day -5.

The study day was numbered with respect to the starting day of the stem cell transplantation, which is defined as day 0. This study includes a pre-study and a treatment phase including the post-transplantation observation up to the end of the study (day 28).

Starting study day -8, oral cavity assessment was performed daily during hospitalization by one of the trained assessors at the site. Oral mucositis assessments were performed daily during the study using the WHO, Radiation Therapy Oncology Group (RTOG), and Western Consortium of cancer Nursing Research (WCCNR) rating scales. Daily

oral mucositis assessment was not discontinued before the oral mucositis grade returns to a WHO grade ≤ 2.

The primary objective of this study is to evaluate the efficacy of palifermin in reducing the duration of severe oral mucositis (WHO grade 3 or 4) induced by total body irradiation (TBI) and high-dose chemotherapy in subjects with hematologic malignancies. The secondary objectives of this study are to evaluate :

- The efficacy of palifermin in

    1. Improving PRO as measured by mouth and throat soreness and its sequelae;
    2. Reducing clinical sequelae related to oral mucositis;
    3. Reducing oral mucositis as measured by different severity grades of interest and by other visual assessment scales
    4. Assess the feasibility, reliability and validity including sensitivity of the Oral Mucositis Daily Questionnaire (OMDQ)

- The safety of palifermin

One interim analysis was performed for 55 patients randomized to the amended, 6-dose schedule and conducted at an alpha level of 0.001. The protocol also specified that a second interim analysis could be performed. However, the second interim analysis was never performed.

3.1.2.2 Efficacy Endpoints

The primary efficacy endpoint for this study was the duration of severe oral mucositis (WHO grades 3 or 4). This duration was calculated as the number of days a subject experienced WHO grade 3 to 4 mucusitis during the efficacy evaluation period (day -8 to day 28 [end of study]).

For the primary efficacy endpoint, the imputation scheme was indicated in the analysis plan. If a subject did not experience any WHO grade 3 or 4 oral mucositis, the duration of severe oral mucositis was assigned as 0 days.

If a subject's mucositis is not resolved upon early withdrawal or death, he/she was given the mean duration of severe mucositis (WHO grade 3 or 4) among subjects who experience at least the same duration of severe mucositis as this subject.

When there was a gap between 2 observed grades, the missing grades were assigned the higher adjacent grade. The same rule was applied to gaps after a patient is discharged with WHO grade 3 or 4 mucositis (but the follow-up assessments were not done daily).

In addition to the imputation scheme, the sponsor also conducted "worst case" analysis to demonstrate the robustness of the results. In the "worst case" analysis, the missing grades were replaced with the higher adjacent value in the palifermin group, and the missing grades in the placebo group were replaced with the lower adjacent values.

The secondary efficacy endpoints related to patient reported outcome (PRO) and clinical sequelae related to oral mucositis include

- Subject's self assessment of mouth and throat soreness as summarized by AUC of the daily VDS score (question 2 in Oral Mucositis Daily Questionnaire [OMDQ]) over the study period;
- Incidence, duration and cumulative dose of parenteral or transdermal opioid analgesics over the study period;
- Incidence and duration of grade 3 or 4 diarrhea as determined using the revised NCI CTC scale;
- Incidence and duration of febrile neutropenia defined as ANC $<0.5 \times 10^9$/L, fever $\geq$ 38.5^0C;
- Incidence of severe neutropenia defined as ANC $< 0.5 \times 10^9$/L;
- Incidence and duration of treatment with IV antifungals or IV antibiotics for documented infections caused by fungi or bacteria.

The secondary efficacy endpoint related to oral mucositis as measured by different severity grades of interest and by other visual assessment scale:

- Incidence of WHO grade 3 or 4 oral mucositis;
- Incidence and duration of WHO grade 4 oral mucositis;
- Incidence and duration of WHO grade 2, 3, or 4 oral mucositis;
- Incidence and duration of WCCNR grade 2 or 3 oral mucositis on Lesion sub-scale.

A subject is considered to have the WHO grade 4 oral mucositis if he/she had at least one WHO grade 4 mucositis at any time. The duration of WHO grade 4 oral mucositis is computed based on the days of grade 4 mucositis. If a patient's grade 4 mucositis is not resolved (resolution is defined as a reading of $\leq$grade 3), he or she will be given mean of grade 4 oral mucositis among subjects who experience at least the same duration of grade 4 oral mucositis as this patient regardless of treatment groups.

For computation of the incidence or duration of WHO grade 2, 3, or 4 oral mucositis, the same rules used for the WHO grade 3 or 4 were applied.

The incidence and duration of grade 2 or 3 WCCNR legion domain were computed based on similar definitions as those for incidence and duration of WHO grades.

The cumulative dose of parenteral or transdermal opioid analgesics over the study period were computed based on similar algorithm described in section 3.1.1.2. For patients who

had missing records due to incomplete records, death or withdrawal, the mean analgesic doses among patients who had taken at least the same amount of analgesic use irrespective of treatment groups.

The incidence of parenteral or transdermal opioid analgesics use was computed based on the number of patients who took at least one non-prophylactic parenteral or transdermal opioid analgesic at anytime during the study divided by the number of patients in each treatment group. The duration of parenteral or trandermal opioid analgesics use was calculated based on the number of days on which a patient had non-prophylactic records of parenteral or transdermal opioid analgesics use.

The patient report outcome assessment will be evaluated by Dr. Lisa Kammerman.

3.1.2.3 Sample Size Consideration

It was planned to have approximately 37 subjects per treatment arm which leads to a total sample size of at least 111 subjects. The null hypothesis for this study is that palifermin and placebo groups had the same duration (in days) of severe oral mucositis. With 37 subjects per group, this sample size was calculated based on 2 sample t-test to detect a 50 % reduction from an average of 10 days of severe mucositis with 80 % power, assuming the standard deviation of the duration is 7.5 days and 5% significance level (2-sided).

3.1.2.4 Efficacy Analysis Method

The generalized Cochran-Mantel-Haenszel (CMH) method stratified by pooled center was used for the primary efficacy analyses. Centers with less than six subjects were pooled into one 'super' center. The standardized mid-rank (also known as the modified Ridit scores) was used for the test statistic. The generalized CMH test follows a Chi-square distribution with 1 degree of freedom.

The primary efficacy analysis dataset was based on modified intend-to-treat (mITT) population defined as all randomized patients who receive at least one dose of study drug and who are analyzed according to their randomized treatment assignment. However protocol amendment 2 eliminated the day -5 dose from the 7-dose schedule, the sponsor performed the analysis based on the following 2 sets :

- Subject with the 7-dose schedule
- Subject with the 6-dose schedule

In the April 2004 analysis plan, controlling of type 1 error was discussed. It indicates that the treatment comparison will be performed first between the palifermin pre-post group and the placebo at 0.05 significance level. A p-value <0.05 result will lead to the

further comparison between the palifermin pre- group and the placebo at 0.05 significance level.

3.1.2.5 Sponsor's Results and Statistical Reviewer's Findings/ Comments

To be consistent with the presentation for Study 20000162, this review will primarily focus on the 6-dose schedule unless specified otherwise

The main comparison under consideration in this review will be the comparison between the placebo group versus the palifermin pre-post group. The patient distribution in the 6-dose sub-population seems to be balanced across treatment groups, however as described in the baseline characteristics section, the distribution of several baseline characteristics appear to be somehow imbalanced across the treatment groups. The study results may be confounded due to the imbalanced baseline characteristics. Also, it is noted that this subgroup analysis may lose power due to the reduction of sample size (i.e. analysis only include 79% [129 out of the 163] of the randomized and treated patients).

The counts of all analysis subsets are presented in the next two tables for 6-dose and 7-dose subsets, respectively.

Table 19 Sponsor's Summary of Primary Analysis Subset – 7-dose Schedule (Study 980231)

		Palifermin	
	Placebo	**Pre**	**Pre-Post**
Subjects Randomized	12	12	11
mITT Population	11	12	11
Safety Analysis Population	11	12	11

Table 20 Sponsor's Summary of Primary Analysis Subset – 6-dose Schedule (Study 980231)

		Palifermin	
	Placebo	**Pre**	**Pre-Post**
Subjects Randomized	40	46	48
mITT Population	40	43	46
Safety Analysis Population	40	43	46

A summary of the subject disposition for 6-dose schedule is provided in the following table. There are only a few patients who had early discontinuation of the study agent or had early withdrawal from the study. Among 48 pre-post patients randomized, 2 of them never received test medication.

Table 21 Sponsor's Summary of Subject Disposition in the Primary Analysis Subset - 6-dose Schedule (Study 980231)

	Placebo	Palifermin Pre	Palifermin Pre-Post	Total
Subjects randomized	40	43	46	89
Test Article Accounting				
Subjects who never received test article	0 (0)	0 (0)	0 (0)	0 (0)
Subjects who received test article	40 (100)	43 (100)	46 (100)	89 (100)
Subjects who completed test article	40 (100)	42 (98)	43 (93)	85 (96)
Subjects who discontinued test article	0 (0)	1 (2)	3 (7)	4 (4)
Study Completion Accounting				
Subjects who completed study	40 (100)	42 (98)	44 (96)	86 (97)
Subjects who discontinued study	0 (0)	1 (2)	2 (4)	3 (3)

3.1.2.5.1 Baseline Characteristics

Sponsor's summaries of the demographic information and baseline disease characteristics are presented in the following two tables. Most of the demographic data appears to be balanced between treatment groups, except that placebo group seems to have better ECOG status at baseline (70% placebo treated patients had ECOG status 0, while 51% and 59% had status 0 for the pre- and pre-post palifermin treated patients, respectively).

Table 22 Sponsor's Summary of Demographic Information (Study 980231)

	Placebo (N=40)	Palifermin Pre (N=43)	Palifermin Pre-Post (N=46)
Sex - n(%)			
Male	20 (50)	30 (70)	26 (57)
Female	20 (50)	13 (30)	20 (43)
Race - n(%)			
Caucasian	30 (75)	30 (70)	33 (72)
Black	5 (13)	7 (16)	4 (9)
Hispanic	4 (10)	3 (7)	7 (15)
Asian	1 (3)	2 (5)	1 (2)
Other	0 (0)	1 (2)	1 (2)

Unknown	0 (0)	0 (0)	0 (0)
Age (years)			
n	40	43	46
Mean (SD)	42 (13)	46 (12)	44 (12)
SD	13	12	12
Median (Min, Max)	44 (18, 63)	50 (18, 65)	44 (18, 64)
Q1, Q3	30, 55	34, 55	35, 55
Height (cm)			
n	40	43	46
Mean (SD)	170	172	170
SD	9.8	12	11
Median (Min, Max)	170 (155, 200)	173 (145, 191)	170 (152, 193)
Q1, Q3	163, 175	163, 182	163, 178
Weight (kg)			
n	40	43	46
Mean (SD)	80 (17)	85 (24)	87 (21)
Median (Min, Max)	79 (53, 128)	85 (50, 158)	85 (48, 136)
Q1, Q3	68, 88	67, 95	68, 105
ECOG Performance Status - n(%)			
0	28 (70)	22 (51)	27 (59)
1	12 (30)	20 (47)	18 (39)
2	0 (0)	1 (2)	0 (0)
Unknown	0 (0)	0 (0)	1 (2)

The baseline characteristics do not seem to be balanced between treatment groups. The pre- and pre-post palifermin group had more non-Hodgkin's Lymnphoma (70% and 61% for the pre- and pre-post palifermin group, respectively) as compared with the placebo group (45%); while placebo group (35%) had more Hodgkin's disease as compared to the 2 palifermin groups (12% and 17% for the pre- and pre-post palifermin group, respectively). CD34+ cells were mobilized with cytokines and chemotherapy for majority of patients, more in placebo group (98%) than in the palifermin groups (84% palifermin pre, 78% palifermin pre-post). Majority of the patients did not have prior mediastinal radiotherapy or prior radiotherapy.

Table 23 Sponsor's Summary of Baseline Characteristics (Study 980231)

		Palifermin	
	Placebo (N=40)	**Pre** (N=43)	**Pre-Post** (N=46)
Type of Diagnosis -n(%)			
Acute Lymphoblastic Leukemia	1 (3)	0 (0)	2 (4)
Acute Myelogenous Leukemia	2 (5)	0 (0)	7 (15)
Hodgkin's Disease	14 (35)	5 (12)	8 (17)
Non-Hodgkin's Lymphoma	18 (45)	30 (70)	28 (61)
Multiple Myeloma	5 (13)	8 (19)	1 (2)
Mobilization -n(%)			
Cytokines only	2 (5)	7 (16)	10 (22)
Cytokines and Chemotherapy	38 (95)	36 (84)	36 (78)
Reason for PBPC Transplantation -n(%)			
Consolidation	18 (45)	22 (51)	19 (41)
Chemotherapy sensitive relapse	18 (45)	15 (35)	21 (46)
Induction failure	4 (10)	6 (14)	6 (13)
Days From Harvest to PBPC Infusion			
n	40	43	46
Mean (SD)	47 (42)	53 (117)	50 (67)
Median (Min, Max)	35 (14, 265)	31 (15, 794)	29 (13, 318)
Q1, Q3	22, 59	21, 43	22, 44
Subjects with Prior Chemotherapy -n(%)			
Yes	40 (100)	43 (100)	46 (100)
Subjects with Prior Mediastinal Radiotherapy -n(%)			
Yes	0 (0)	0 (0)	1 (2)
No	40 (100)	43 (100)	45 (98)
Subjects with Prior Radiotherapy -n(%)			
Yes	4 (10)	6 (14)	6 (13)
No	35 (88)	37 (86)	40 (87)
Unknown	1 (3)	0 (0)	0 (0)

3.1.2.5.2 Primary Efficacy Analyses

The sponsor's primary efficacy analysis results for the 6-dose schedule indicated a significant shorter duration of WHO Grade 3 or 4 mucositis for the palifermin group (medians= 4 days for pre- or pre-post palifermin group) than the duration for the placebo group (median= 6.0 days). Such trend (shorter duration of mucositis in palifermin group) was consistent in the subgroup of patients who developed the WHO Grade 3 or 4 mucositis. The primary analysis results are also confirmed based on the sponsor's worst case analysis.

Table 24 Sponsor's Summary of Duration of WHO Grade 3 or 4 (days) – 6-dose Schedule (Study 980231)

		Palifermin	
	Placebo (N=40)	**Pre** (N=43)	**Pre-Post** (N=46)
Duration - WHO Grade 3 or 4 (days)			
N	40	43	45
Mean (SD)	8.6 (8.2)	5.2 (6.1)	4.7 (5.7)
Median (Min,Max)	6.0 (0, 31)	4.0 (0, 27)	4.0 (0, 32)
Q1, Q3	3.5 , 11.0	0.0 , 7.0	0.0 , 7.0
P-Value [a]		0.003	0.004

[a] Generalized Cochran-Mantel-Haenszel test based on the standardized mid-ranks (modified Ridit scores) within each stratum.

Reviewer's comment : Subject 153 (palifermin pre-post) experienced a respiratory infection that delayed the beginning of her TBI. This delay prevented her from being on the same schedule as all other subjects. The sponsor made administrative decision to remove her from study participation. There was no mucositis grade measure after the first dose for this patient, therefore, the analysis size for the pre-post palifermin group is 45 throughout the sponsor's summary of incidence or duration of mucositis grades.

In the analysis of the duration of WHO grade 3 or 4 for the 7-dose schedule subgroup, the duration is slightly shorter in the palifermin group (medians= 11 and 9 days for pre- and pre-post palifermin groups, respectively) as compared with that for the placebo group (median=11.6 days). However, the difference in duration between the placebo and either one of palifermin group is not statistically significant.

Table 25 Sponsor's Summary of Duration of WHO Grade 3 or 4 (days) – 7-dose Schedule (Study 980231)

		Palifermin	
	Placebo (N=11)	**Pre** (N=12)	**Pre-Post** (N=11)
Duration of Severe Mucositis (WHO Grade 3,4) (days)			
n	11	12	11
Mean	12.1 (7.9)	13.1 (9.9)	13.0(9.8)
Median (Min, Max)	11.6 (1.0 , 30.0)	11.0 (0.0 , 30.0)	9.0 (0.0 , 26.0)
Q1,Q3	7.0 , 13.0	7.5 , 21.0	8.0 , 25.0
P-Value [a] (Compared to Placebo)		0.974	0.574

[a] Generalized Cochran-Mantel-Haenszel test based on the standardized mid-ranks (modified Ridit scores) within each stratum.

Reviewer's comments : this is the subgroup that was later dropped after July 9, 1999 safety committee review. The sponsor said that review of pooled and individual subject data suggested that palifermin administration on day -5 (approximately 14 hours before initiation of etoposide on day -4) had no beneficial effects on oral mucositis. Considering the pharmacodynamic information, the safety committee proposed that there were (theoretical) risks of eliminating benefit or worsening oral mucositis when the time between palifermin administration and chemotherapy was ≤ 24 hours. However, such concern may not be supported by this clinical study result, either duration of WHO grade 3, 4 or incidence of WHO grade 3, 4 (these are 100%, 83% and 82% for the placebo group, pre- and pre-post palifermin group respectively). Both are trending in favor of palifermin groups, although in a smaller magnitude as compared with those comparisons from the 6-dose schedule.

Table 26 Duration (days) of WHO Grade 3 or 4 Mucositis for Subjects Who Developed Such Toxicity During Study (Study 980231)

		Palifermin	
	Placebo (N=40)	**Pre (N=43)**	**Pre-Post (N=46)**
Duration of Mucositis (WHO Grade 3,4) (days)			
n	32	31	31
Mean (SD)	10.8 (7.7)	7.2 (6.0)	6.9 (5.7)
Median (Min, Max)	8.5 (1.0, 31.0)	6.0 (1.0, 27.4)	6.0 (2.0, 32.0
Q1, Q3	5.5, 13.0	4.0, 9.0	4.0, 8.0

Table 27 Sponsor's Summary of Duration of WHO Grade 3 or 4 (days) – 6-dose Schedule (Worst Case) (Study 980231)

		Palifermin	
	Placebo (N=40)	Pre (N=43)	Pre-Post (N=46)
Duration - WHO Grade 3 or 4 (days)			
N	40	43	45
Mean (SD)	7.3 (6.3)	5.2 (6.0)	4.7 (5.7)
Median (Min,Max)	6.0 (0, 24.5)	4.0 (0, 27)	4.0 (0, 32)
Q1, Q3	0, 7.0	0, 7.0	0, 7.0
P-Value [a]		0.0194	0.0102

[a] Generalized Cochran-Mantel-Haenszel test based on the standardized mid-ranks (modified Ridit scores) within each stratum.

3.1.2.5.3 Secondary Efficacy Analyses

Since the sponsor did not pre-specify the decision rule for the secondary endpoints, the following analyses are more exploratory in nature. The only nominally significant results (nominal p-value < 0.05) that might provide supporting information for the Phase III pivotal result are the duration of WHO grade 2, 3 or 4 mucositis, incidence and duration of WHO grade 4 mucositis. However, these 2 endpoints are not prespecified in order of importance.

Table 28 Sponsor's Summary of Incidence and Duration (days) of WHO Grades – 6-dose Schedule (Study 980231)

		Palifermin	
	Placebo (N=40)	Pre (N=43)	Pre-Post (N=46)
Incidence - WHO Grade 3 or 4 - n (%)	32 (80)	31 (72)	31 (67)
P-Value [a]		0.184	0.159
Incidence – WHO Grade 2, 3, or 4 - n (%)	39 (98)	41 (95)	39 (85)
P-Value [a]		0.634	0.099
Duration - WHO Grade 2, 3 or 4 (days)			
N	40	43	45
Mean (SD)	17.7 (11.3)	10.8 (8.4)	11.1 (9.7)
Median (Min,Max)	12.5 (0, 37)	9.0 (0, 34)	9.0 (0, 34)
Q1, Q3	10.0, 27.5	5.0, 14.0	4.0, 15.0
P-Value [a]		0.002	0.001
Incidence - WHO Grade 4 - n (%)	20 (50)	14 (33)	12 (26)

P-Value [a]		0.029	0.025
Duration - WHO Grade 4 (days)			
N	40	43	45
Median (Min,Max)	0.5 (0, 13)	0.0 (0, 13)	0.0 (0, 21)
Q1, Q3	0.0, 4.5	0.0, 3.0	0.0, 2.0
P-Value [a]		0.026	0.022

[a] Generalized Cochran-Mantel-Haenszel test based on the standardized mid-ranks (modified Ridit scores) within each stratum.

Reviewer's comment : The sponsor analyzed the incidence data based on the generalized CMH using modified ridit score. This reviewer used the CMH method for binary data instead. When the CMH method for binary data was used, the p-values are 0.6020, 0.4028 and 0.1085 for incidences of WHO Grade 2, 3, 4, Grade 3, 4 and Grade 4, respectively, for the comparison between palifermin pre- group and the placebo group. The p-values are 0.1056, 0.1658 and 0.0255 for incidences of WHO Grade 2, 3, 4, Grade 3, 4 and Grade 4, respectively, for the comparison between palifermin pre-post group and the placebo group. The only noticeable change is in the comparison of the palifermin pre- group versus placebo. The p-value was changed from the sponsor's result: 0.029 to 0.1085.

The duration of WHO grade 3 or 4 results were further supported by the duration of WCCNR grade 2 or 3 results (see the following table).

Table 29 Sponsor's Summary of Duration (days) of WCCNR Grade 2 or 3 Mucositis (Study 980231)

		Palifermin	
	Placebo (N=40)	**Pre (N=43)**	**Pre-Post (N=46)**
Duration (days) of WCCNR Grade 2 or 3 Mucositis			
n	40	43	45
Mean (SD)	8.8 (9.1)	3.4 (5.3)	2.9 (3.5)
Median (Min, Max)	6.0 (0.0, 32.0)	0.0 (0.0, 26.8)	0.0 (0.0, 11.0)
Q1, Q3	3.0, 10.0	0.0, 6.0	0.0, 6.0
P-Value [a] (Compared to Placebo)		<0.001	<0.001

[a] Generalized Cochran-Mantel-Haenszel test based on the standardized mid-ranks (modified Ridit scores) within each stratum.

The results from the sponsor's analysis of the cumulative dose of opioid analgesic use for mucositis showed a nominally significant lower total dose of analgesic use for the palifermin pre- and pre-post groups. Similarly, the palifermin pre- and pre-post groups seem to have shorter duration of analgesic use for mucositis as compared with placebo.

Table 30 Cumulative Dose of Opioid Analgesic Use for Mucositis – 6-dose Schedule (Subset of Subjects with Opioid Analgesic Use for Mucositis) (Study 980231)

		Palifermin	
	Placebo (N=40)	Pre (N=43)	Pre-Post (N=46)
Cumulative Dose of Opioid Analgesic Use for Mucositis			
n	40	43	46
Mean (SD)	1163 (1776)	505.3 (859.1)	514.4 (922.2)
Median (Min, Max)	523.9 (0.0, 8472)	216.7 (0.0, 4574)	204.9 (0.0, 5213)
Q1, Q3	236.6, 1380	6.0,562.1	16.0,574.5
P-Value [a] (Compared to Placebo)		0.002	0.004

[a] Generalized Cochran-Mantel-Haenszel test based on the standardized mid-ranks (modified Ridit scores) within each stratum.

Table 31 Duration of Opioid Analgesic Use for Mucositis – 6-dose Schedule (Study 980231)

		Palifermin	
	Placebo (N=40)	Pre (N=43)	Pre-Post (N=46)
Duration (days) of Opioid Analgesic Use			
n	40	43	46
Mean	10.6 (5.5)	7.1 (5.1)	7.0 (5.4)
Median	10.0 (0.0, 25.0)	8.0 (0.0, 19.0)	7.0 (0.0, 28.0)
Q1, Q3	8.0, 12.0	3.0, 11.0	3.0, 9.0
P-Value [a] (Compared to Placebo)		0.007	<0.001

[a] Generalized Cochran-Mantel-Haenszel test based on the standardized mid-ranks (modified Ridit scores) within each stratum.

3.1.2.6 Sponsor's Conclusions and Reviewer's Conclusion/Comments

Based on the sponsor's analysis, the median (25th percentile, 75th percentile) duration of WHO Grade 3 or 4 mucositis was 6.0 (3.5, 11.0) days for the placebo group, and 4.0 (0, 7.0) days and 4.0 (0, 7.0) for the palifermin pre- and pre-post groups, respectively. Both comparisons (the placebo versus palifermin pre- group and the placebo versus palifermin pre-post group) in the duration of WHO Grade 3 or 4 mucositis group are significant (p-value<0.01) in favor of the palifermin group. These results were

supported by the "worst case" analysis and WCCNR grade 2 or 3 mucositis. The duration of WHO grade 2, 3, 4 mucositis also had the similar trend.

Since there were only a few missing visits that requires imputation of WHO grades 3 or 4 (6%, 2% and 4% patients among all potential number of days of evaluation for placebo, palifermin pre and pre-post groups, respectively), the impact of imputation may not be important. Without implementing any imputation scheme, this reviewer performed sensitivity analyses using the similar strategy as shown in section 3.1.1.6 and the results were shown in the following table. The first analysis is using the number of days for WHO grade 3 or 4 mucositis and the second analysis is using the duration of mucositis computed based on the time span between the first date when the mucositis became grade 3 or 4 to the last date when the mucositis was still of grade 3 or 4. P-values from 2 statistics were presented: the generalized CMH and the Wilcoxon rank sum test.

Table 32 Reviewer's Summary of Duration of Severe Mucositis (WHO Grade 3 or 4) (Study 980231)

	Placebo	Palifermin	
		Pre	Pre-post
	(N=40)	(N=43)	(N=46)
Duration of WHO Grade 3 or 4			
1. Counting days of Grade 3 or 4			
n	40	43	46
Mean (SD)	7.1 (5.9)	4.6 (4.9)	3.6 (3.8)
Median (Min, Max)	6.0 (0.0, 21.0)	4.0 (0.0, 22.0)	3.5 (0.0, 16.0)
Q1, Q3	(3, 10.5)	(0.0, 7.0)	(0.0, 6.0)
P-value [a]		1) 0.0029	1) 0.0012
(compared with placebo group)		2) 0.0417	2) 0.0040
2. Counting duration [b] of Grade 3 or 4			
n	40	43	46
Mean (SD)	7.8 (7.3)	4.9 (5.7)	4.1 (4.1)
Median (Min, Max)	6.0 (0.0, 31.0)	4.0 (0.0, 27.0)	4.0 (0.0, 16.0)
Q1, Q3	(3.5, 11)	(0.0, 7.0)	(0.0, 6.0)
P value [a]		1) 0.0069	1) 0.0041
(compared with placebo group)		2) 0.0453	2) 0.0089

[a] Duration of mucositis was computed based on the time span between the first date when the mucositis became grade 3 or 4 to the last date when the mucositis grade was still 3 or 4.

[b] The first p-value was based on Generalized Cochran-Mantel-Haenszel test with the standardized mid-ranks (modified Ridit scores) within each stratum while the second p-value was based on Wilcoxon rank sum test.

Both analysis results for comparison of the duration of grade 3 or 4 WHO grades between the placebo group and either one of palifermin group appear to be similar to the

sponsor's results. The Wilcoxon rank sum test seems to provide more conservative results.

The sponsor's analysis results are based on subgroup analyses (i.e. based on 6-dose and 7-dose schedule, separately). If the 2 schedule data were combined, the result for the primary efficacy analysis was not changed (p-value=0.005 for the palifermin pre-post group versus placebo and p-value=0.009 for the palifermin pre- group versus placebo).

The study was finished in October, 2002, however the analysis plan was finalized in April 23, 2004. Some modification/changes of the analysis plan from the original plan include

- In the August, 1999 version of protocol, the primary efficacy analysis dataset was not determined. It stated that the decision on whether a mITT analysis set or per protocol analysis subset to use for the efficacy analysis will be based on the examination of whether there is a difference between subjects who received the originally schedules day -5 dose from those who do not. However, in the final analysis, mITT was used and the analysis was based on 6-dose and 7-dose schedule, separately.

- The sponsor indicated in the August 1999 protocol that no multiplicity adjustment for the two comparisons (pre-post versus placebo; pre- versus placebo) of the duration of severe mucositis between treatments will be made. However, a decision rule with regard to the two comparisons was indicated in the April, 2004 analysis plan to control the type 1 error (see section 3.1.2.4 Efficacy Analysis Method of this review for details).

- The imputation method for the primary efficacy endpoint was changed. The original plan indicated in the August, 1999 protocol is to assign the average background duration of ten days or the actual observed value, which ever is longer for those patients whose primary endpoint is not completely observed due to subject withdrawal or death. However, in the sponsor's report, for a subject who had early discontinuation or died, his grade 3 or 4 mucosa duration was assigned with the mean duration of severe mucositis (WHO grade 3 or 4) among subjects who experience at least the same duration of severe mucositis as this subject.

- The sponsor did not perform analysis based on AUC endpoint for mucositis (specified in August, 1999 protocol). The incidence and duration of WHO Grade 2, 3, and 4 oral mucositis and of Grade 4 oral mucositis were analyzed instead. In addition, the incidence and duration, as well as cumulative dose of opioid analgesic were analyzed, rather than the AUC of the opioid analgesic use specified in the August, 1999 protocol.

- The incidence and duration of NCI CTC grade 3 or 4 diarrhea were analyzed instead of grade 2, 3, or 4 (specified in the August, 2004 protocol).

In addition, there was no mention of the decision rule with regard to the secondary endpoints. The analysis of the secondary endpoints was more exploratory in nature. Therefore, the evidence of the supporting information is not clear.

3.2 Evaluation of Safety

Safety efficacy endpoints include incidence and severity of adverse events, change of hematology and clinical chemistries (especially serum amylase and lipase) and incidence of serum anti-palifermin antibody formation were evaluated. The details of the safety data results were provided in the clinical review.

4 Findings in Special/Subgroup Populations

This section only includes descriptive statistics (mean, standard deviation, median, maximum, minimum, 25th percentile [Q1] and 75th percentile [Q3]) for all the subgroup analysis. Only the results from study 20000162 will be provided since the sample size from study 980231 is too small to have meaningful presentation.

4.1 Gender

Summaries of the duration based on WHO, WCCNR or RTOG criteria and WHO grade 4 incidence, as well as the morphine equivalent doses are presented in the following table. The palifermin group appears to have shorter duration of mucositis based on different criteria either in male or female subgroup. The palifermin group also had lower grade 4 WHO grade in each gender group. In addition, the palifermin group had consistently less cumulative opioid analgesic use across gender groups.

Table 33 Summary of Durations (days) and Incidences of WHO, WCCNR and RTOG Grades, Total Doses of Opioid Analgesic Use by Gender (Study 20000162)

	Gender	Statistics	Placebo	Palifermin
Duration of WHO Grade 3,4	Female	N	34	47
		Mean (SD)	10.2 (6.7)	3.7 (4.4)
		Median (Min,Max)	8.0 (0.0,27.0)	3.0 (0.0,21.7)
		Q1, Q3	(5.0,14.0)	(0.0,6.0)
	Male	N	72	59
		Mean (SD)	10.5 (6.0)	3.7 (3.8)
		Median (Min,Max)	10.0 (0.0,27.0)	3.0 (0.0,13.0)
		Q1, Q3	(6.5,13.0)	(0.0,7.0)

Duration of WHO Grade 2, 3, 4	Female	N	34	47
		Mean (SD)	15.9 (8.7)	8.5 (5.5)
		Median (Min,Max)	14.0 (0.0,37.0)	8.0 (0.0,20.0)
		Q1, Q3	(9.0,22.0)	(5.0,12.0)
	Male	N	72	59
		Mean (SD)	15.6 (7.4)	8.3 (6.1)
		Median (Min,Max)	14.8 (1.0,37.0)	7.0 (0.0,28.0)
		Q1, Q3	(11.0,19.0)	(4.0,12.0)
Incidence of WHO Grade 4	Female	N (%)	19 (55.9)	6 (12.8)
	Male	N (%)	47 (65.3)	15 (25.4)
Duration of WCCNR Grade 2, 3	Female	N	34	47
		Mean (SD)	8.4 (8.0)	3.1 (5.9)
		Median (Min,Max)	7.0 (0.0,30.0)	0.0 (0.0,36.0)
		Q1, Q3	(2.0,12.0)	(0.0,5.0)
	Male	N	72	59
		Mean (SD)	10.6 (9.8)	3.3 (4.5)
		Median (Min,Max)	7.5 (0.0,56.0)	2.0 (0.0,25.0)
		Q1, Q3	(4.0,15.0)	(0.0,5.0)
Duration of RTOG Grade 3, 4	Female	N	34	47
		Mean (SD)	6.9 (6.5)	1.6 (2.4)
		Median (Min,Max)	5.0 (0.0,22.0)	0.0 (0.0,8.0)
		Q1, Q3	(1.0,11.0)	(0.0,3.0)
	Male	N	72	59
		Mean (SD)	8.7 (9.4)	2.5 (4.2)
		Median (Min,Max)	6.0 (0.0,54.0)	0.0 (0.0,24.0)
		Q1, Q3	(3.0,11.0)	(0.0,4.0)
Total doses of Opioid Analgesic use (mg)	Female	N	34	47
		Mean (SD)	811.5 (796.3)	681.0 (1905.8)
		Median (Min,Max)	451.6 (0.0,2901.8)	168.6 (0.0,9418.4)
		Q1, Q3	(192.1,1300.5)	(2.0,476.0)
	Male	N	72	59
		Mean (SD)	1304.7 (1977.6)	714.3 (1630.5)
		Median (Min,Max)	580.9 (0.0,9418.4)	270.0 (0.0,9418.4)
		Q1, Q3	(290.0,1435.2)	(4.0,662.0)

4.2 Race

Summaries of WHO grade 4 incidence and durations of mucositis based on WHO, WCCNR and RTOG grades as well as total doses of analgesics, by race (white and other) are presented in the following table. Again, the palifermin group had consistently a shorter grade 3 or 4 mucositis duration, a lower grade 4 incidence and less analgesic use across white and non-white subpopulations.

Table 34 Summary of Incidence and Duration (days) of WHO, WCCNR and RTOG Grades, Total Doses of Opioid Analgesic Use by Race (Study 20000162)

	Race	Statistics	Placebo	Palifermin
Duration of WHO Grade 3,4	White	N	89	78
		Mean (SD)	10.3 (6.4)	3.5 (4.2)
		Median (Min,Max)	9.0 (0.0,27.0)	3.0 (0.0,21.7)
		Q1, Q3	(6.0,13.0)	(0.0,6.0)
	Other	N	17	28
		Mean (SD)	11.0 (5.2)	4.2 (3.8)
		Median (Min,Max)	10.0 (4.0,22.0)	5.0 (0.0,10.0)
		Q1, Q3	(7.0,15.6)	(0.0,7.0)
Duration of WHO Grade 2, 3, 4	White	N	89	78
		Mean (SD)	15.3 (7.5)	8.6 (5.9)
		Median (Min,Max)	14.5 (0.0,37.0)	8.0 (0.0,28.0)
		Q1, Q3	(11.0,18.0)	(4.0,12.0)
	Other	N	17	28
		Mean (SD)	17.5 (9.4)	7.8 (5.5)
		Median (Min,Max)	12.0 (6.0,37.0)	7.0 (0.0,24.0)
		Q1, Q3	(11.0,25.8)	(5.0,9.0)
Incidence of WHO Grade 4	White	N (%)	53 (59.6)	17 (21.8)
	Other	N (%)	13 (76.5)	4 (14.3)
Duration of WCCNR Grade 2, 3	White	N	89	78
		Mean (SD)	10.0 (9.7)	3.3 (5.6)
		Median (Min,Max)	7.0 (0.0,56.0)	1.0 (0.0,36.0)
		Q1, Q3	(4.0,15.0)	(0.0,5.0)
	Other	N	17	28
		Mean (SD)	9.2 (6.6)	2.8 (3.6)
		Median (Min,Max)	9.0 (0.0,25.0)	1.0 (0.0,12.0)

		Q1, Q3	(5.0,12.0)	(0.0,5.0)
Duration of RTOG Grade 3, 4	White	N	89	78
		Mean (SD)	8.4 (9.1)	2.1 (3.6)
		Median (Min,Max)	6.0 (0.0,54.0)	0.0 (0.0,24.0)
		Q1, Q3	(3.0,11.0)	(0.0,4.0)
	Other	N	17	28
		Mean (SD)	6.7 (4.6)	2.0 (3.2)
		Median (Min,Max)	7.0 (0.0,16.0)	0.0 (0.0,12.0)
		Q1, Q3	(4.0,10.0)	(0.0,3.5)
Total doses of Opioid Analgesic use (mg)	White	N	89	78
		Mean (SD)	1105.4 (1666.4)	820.9 (2008.3)
		Median (Min,Max)	507.3 (0.0,9418.4)	204.3 (0.0,9418.4)
		Q1, Q3	(256.7,1362.5)	(3.0,558.4)
	Other	N	17	28
		Mean (SD)	1361.9 (1918.7)	361.3 (486.3)
		Median (Min,Max)	672.0 (30.0,8162.7)	262.5 (0.0,2313.7)
		Q1, Q3	(409.5,1639.2)	(3.0,559.0)

4.3 Age

The duration of oral mucositis based on WHO, WCCNR and RTOG grades are consistently shorter for the palifermin group in patients older than 65 or younger than 65. Similarly, there was lower WHO grade 4 mucositis in the palifermin group. Due to the small size in the patients of ≤ 65 years old, some variations of treatment difference are observed, e.g., larger difference of duration of WHO Grade 3 or 4 in the > 65 years old patient group than that in the $\leq$ 65 years old patient group. This differential result can not be further confirmed in the current study.

Table 35 Summary of Incidence and Duration (days) of WHO, WCCNR and RTOG Grades, Total Doses of Opioid Analgesic Use by Age (Study 20000162)

	Age	Statistics	Placebo	Palifermin
Duration of WHO Grade 3,4	< 65	N	101	103
		Mean (SD)	10.2 (6.2)	3.7 (4.0)
		Median (Min,Max)	9.0 (0.0,27.0)	3.0 (0.0,21.7)
		Q1, Q3	(6.0,13.0)	(0.0,6.0)
	≥65	N	5	3
		Mean (SD)	14.6 (6.4)	3.3 (5.8)
		Median (Min,Max)	18.0 (5.0,20.0)	0.0 (0.0,10.0)

		Q1, Q3	(11.0,19.0)	(0.0,10.0)
Duration of WHO Grade 2, 3, 4	< 65	N	101	103
		Mean (SD)	15.5 (7.9)	8.3 (5.8)
		Median (Min,Max)	14.0 (0.0,37.0)	8.0 (0.0,28.0)
		Q1, Q3	(10.0,18.0)	(4.0,12.0)
	≥ 65	N	5	3
		Mean (SD)	17.8 (5.8)	11.0 (5.6)
		Median (Min,Max)	21.0 (11.0,23.1)	12.0 (5.0,16.0)
		Q1, Q3	(12.0,22.0)	(5.0,16.0)
Incidence of WHO Grade 4	< 65	N (%)	64 (63.3)	20 (19.4)
	≥ 65	N (%)	2 (40.0)	1 (33.3)
Duration of WCCNR Grade 2, 3	< 65	N	101	103
		Mean (SD)	9.5 (9.3)	3.2 (5.2)
		Median (Min,Max)	7.0 (0.0,56.0)	1.0 (0.0,36.0)
		Q1, Q3	(4.0,12.0)	(0.0,5.0)
	≥ 65	N	5	3
		Mean (SD)	17.4 (5.2)	3.3 (3.5)
		Median (Min,Max)	18.0 (9.0,23.0)	3.0 (0.0,7.0)
		Q1, Q3	(17.0,20.0)	(0.0,7.0)
Duration of RTOG Grade 3, 4	< 65	N	101	103
		Mean (SD)	7.7 (8.5)	2.1 (3.5)
		Median (Min,Max)	6.0 (0.0,54.0)	0.0 (0.0,24.0)
		Q1, Q3	(2.0,10.0)	(0.0,4.0)
	≥ 65	N	5	3
		Mean (SD)	16.4 (6.2)	2.0 (3.5)
		Median (Min,Max)	17.0 (6.0,22.0)	0.0 (0.0,6.0)
		Q1, Q3	(17.0,20.0)	(0.0,6.0)
Total doses of Opioid Analgesic use (mg)	< 65	N	101	103
		Mean (SD)	1150.8 (1727.5)	714.9 (1771.9)
		Median (Min,Max)	532.0 (0.0,9418.4)	213.2 (0.0,9418.4)
		Q1, Q3	(284.5,1429.0)	(4.0,602.0)
	≥ 65	N	5	3
		Mean (SD)	1059.3 (1197.4)	172.0 (297.9)
		Median (Min,Max)	684.5 (166.0,3085.1)	0.0 (0.0,516.0)
		Q1, Q3	(230.0,1131.0)	(0.0,516.0)

4.4 Other Special/Subgroup Populations

Sponsor provided subgroup analysis for duration of WHO grade 3 or 4 mucositis by type of hematologic malignancy. In general, the palifermin group had consistently shorter duration of WHO grade 3 or 4 mucositis across different subgroup.

Table 36 Sponsor's Summary of Duration (days) of WHO Grade 3 or 4 Mucositis by Type of Hematologic Malignancy (Study 2000162)

	Placebo (N=106)	Palifermin (N=106)
Non-Hodgkins Lymphoma		
n	69	72
Mean (SD)	10.8 (6.6)	3.9 (3.9)
Median (Min,Max)	10.0 (0.0 , 27.0)	3.0 (0.0 , 13.0)
Q1,Q3	6.0 , 14.0	0.0 , 7.0
Hodgkins Disease		
n	23	21
Mean (SD)	9.7 (5.0)	3.2 (5.0)
Median (Min,Max)	8.0 (4.0 , 25.0)	2.0 (0.0 , 21.7)
Q1,Q3	6.0 , 13.0	0.0 , 5.0
Multiple Myeloma		
n	9	11
Mean (SD)	9.8 (4.8)	3.1 (3.4)
Median (Min,Max)	8.0 (5.0 , 18.0)	3.0 (0.0 , 9.0)
Q1,Q3	6.0 , 13.0	0.0 , 7.0
Leukemia		
n	5	2
Mean (SD)	10.3 (9.1)	6.0 (5.7)
Median (Min,Max)	6.0 (5.0 , 26.3)	6.0 (2.0 , 10.0)
Q1,Q3	5.0 , 9.0	2.0 , 10.0

The sponsor provided evaluation of whether the development of pre-irradiation treatment, skin, facial, or oral-related AE may un-blind the results. However, only 30 (1.4%, 5 in placebo group and 25 in the palifermin group) of the patients in study 20000162 had such condition. The result from this evaluation is inconclusive.

5 Summary and Conclusions

The sponsor submitted two randomized, placebo-controlled clinical studies to evaluate effect of palifermin for reduction of oral mucositis in patients with hematologic malignancies undergoing total body irradiation (TBI) and high-dose chemogherapy with autologous peripheral blood progenitor cell transplantation (PBPC).

In study 20000162, 212 patients were randomized and received either palifermin or placebo. The study medication was administered as a daily IV injection of 60 μg/kg for 3 consecutive days prior to TBI and for 3 consecutive days after PBPC (pre-post group). Randomization for this Phase III study was stratified by center and type of hematologic malignancy (non-Hodgkin's lymphoma, Hodgkin's disease, leukemia or multiple myeloma).

In study 980231, 163 subjects were randomized and received placebo, pre- and pre-post palifermin. In the palifermin pre- group, patients only received palifermin prior to initiation of TBI. Originally, in addition to the 3 doses prior to TBI and 3 doses after PBPC, the Phase II study also includes a day -5 dose immediately after the last dose of TBI. After July 9, 1999 internal Safety Committee meeting, the safety committee decided to terminate the Day -5 dose since there were possible risk of worsening oral mucositis if the time between palifermin admistration and chemotherapy was less than 24 hours. This Phase II study was stratified by study center. Analysis for this Phase II study was based on 6-dose (n=129) and 7-dose schedule (n=34).

5.1 Statistical Issues and Collective Evidence

The primary efficacy endpoint for both studies is the duration of WHO grade 3 or 4 oral mucositis. The primary efficacy analysis was based on the generalized Cochran-Mantel-Haenszel test using the standardized mid-ranks (modified Ridit scores) within each stratum.

In study 20000162, the sponsor prospectively specified the key secondary endpoints : mouth and throat soreness (Patient Reported Outcome [PRO]); use of parenteral or transdermal opioid analgesics; and duration of WHO grade 2, 3, and 4 oral mucositis and stated that if the duration of WHO grade 3 or 4 is significant, each of these endpoints will be tested at a 0.05 significance level. However, there is no pre-specified plan implemented in study 980231. The lack of prespecified rules of dealing with comparison of multiple arms and multiple endpoints in study 980231 resulted in difficulties of interpreting the results of the secondary efficacy endpoints.

A summary of the primary efficacy endpoint of these 2 studies is presented in the following table.

Table 37 Summary of the Primary Efficacy Endpoint – WHO Grade 3 or 4 Mucositis

Duration of Severe Mucositis days (WHO grade 3,4)	Placebo	Palifermin	
Study 20000162		**Pre-**	**Pre-post**
	(N=106)		**(N=106)**
N	106	NA	106
Mean (SD)	10.4 (6.2)		3.7 (4.1)
Median (Min,Max)	9.0 (0.0, 27.0)		3.0 (0.0, 21.7)
Q1, Q3	6.0, 13.0		0.0, 6.0
P-Value [a] (compare with placebo)			< 0.001
Study 980231 – 6-dose schedule	**(N=40)**	**(N=43)**	**(N=46)**
N	40	43	45
Mean (SD)	8.6 (8.2)	5.2 (6.1)	4.7 (5.7)
Median (Min,Max)	6.0 (0, 31)	4.0 (0, 27)	4.0 (0, 32)
Q1, Q3	3.5, 11.0	0.0, 7.0	0.0, 7.0
P-Value [a] (compare with placebo)		0.003	0.004
Study 980231 – 7-dose schedule	**(N=11)**	**(N=12)**	**(N=11)**
N	11	12	11
Mean (SD)	12.1 (7.9)	13.1 (9.9)	13.0(9.8)
Median (Min,Max)	11.6 (1.0 , 30.0)	11.0 (0.0 , 30.0)	9.0 (0.0 , 26.0)
Q1, Q3	7.0, 13.0	7.5, 21.0	8.0, 25.0
P-Value [a] (compare with placebo)		0.974	0.574

[a] Generalized Cochran-Mantel-Haenszel test based on the standardized mid-ranks (modified Ridit scores) within each stratum.

Due to larger zero duration of grade 3 or 4 mucositis, particularly in the palifermin group, the distribution is quite skewed. The distribution free method, such as generalized CMH or Wilcoxon rank sum test, is a more appropriate method.

The sponsor's imputation scheme, eg, using mean duration to impute the data for the patients who had unresolved WHO grade 3 or 4 by the time of early withdrawal or death, may not be the optimum method. However, given the small percentages of patients with early discontinuation, the impact of the imputation scheme may not be important.

There is a concern about the possible unblinding effect of some pre-irradiation treatment related AE (e.g., skin, facial, or oral-related AE). However, due to the small incidences of these AE, the potential impact of these AE can not be confirmed.

5.2 Conclusions and Recommendations

Based on study 20000162, the median duration of WHO grade 3 or 4 mucositis were 9 days and 3 days for placebo and palifermin group, respectively. These results were confirmed by the 6-dose subgroup from study 980231, either based on the comparison between palifermin pre-post group and placebo or between palifermin pre- group and placebo. The result based on the 7-dose schedule did not reach statistical significance.

Based on all data provided, the results supported the efficacy claim of palifermin in reduction of duration of WHO grade 3 or 4 mucositis based on 6-dose dosing schedule.

The sponsor also showed statistical significance in the use of parenteral or transdermal opioid analgesics; and duration of WHO grade 2, 3, and 4 oral mucositis in study 2000162. However, the results from Study 980231 do not provide confirmatory evidence due to exploratory nature of the analyses.

6 SIGNATURES/DISTRIBUTION LIST PAGE

Primary Statistical Reviewer: **Date:** December 2, 2004

Yuan-Li Shen 12/2/04

Dr. Yuan-Li Shen
Mathematical Statistician

Concurrence:

Mark Rothmann 12/2/04

Dr. Mark Rothmann
Statistical Team Leader

Alaka Chakravarty 12/2/04

Dr. Aloka Chakravarty
Staff Director

CC:
HFD-109/ Ms. Giuliani
HFD-107/ Dr. Dinndorf
HFD-107/ Dr. Gootenberg
HFD-711/ Dr. Rothmann
HFD-711/ Dr. Shen
HFD-711/ Dr. Chakravarty
HFD-700/ Dr. Anello
HFD-700/ Dr. Dubey

This review consists of 48 pages (43 pages of text)
C:\BLA_2004\KGF\Palifermin_statreview 2004.doc

12/1/04

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Pharmacoepidemiology and Statistical Science
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

BLA/Serial Number: 125103 / 0

Drug Name: Kepivance (palifermin, recombinant human Keratinocyte Growth Factor (rKGF))

Indication(s): [redacted]

Applicant: Amgen

Date(s): PDUFA: 12/16/2004

Review Priority: Priority

Biometrics Division: DBII

Statistical Reviewer: Lisa A. Kammerman, Ph.D.

Concurring Reviewers: Not applicable

Medical Division: Office of Therapeutics Research and Review (HFM-99)

Clinical Team: Pat Dinndorf, MD

Project Manager: Susan Giuliani, RN, MS

Keywords: clinical studies, patient-reported outcomes

Table of Contents

1 EXECUTIVE SUMMARY

The applicant is proposing the following indication:

[ ]

Patient-reported outcomes of the oral mucositis daily diary are the subject of this review. Dr. Yuan-Li Shen is reviewing the efficacy and safety of palifermin.

The development and validation of the oral mucositis daily diary are the major review issues.

FDA and Amgen have discussed these review issues over several years prior to the submission of this BLA. As agreed at the pre-BLA meeting, Amgen has documented the development and validation history of the daily diary questionnaire. The information is provided within appendices to each of the three study reports, starting with the Dose Escalation Trial (Study 960189), then the Phase 2 Trial (Study 980231) and ending with the Phase 3 Trial (Study 20000162).

The daily diary was not properly developed for use in reporting the impact of mucositis on a patient's quality of life. Appropriate steps were implemented in developing the pilot diary and its revision, which was used in the Phase 1 study.

The results from Phase 1 were not used in a scientific approach to creating a shorter form or for constructing new composite scales. Steps specified in the Phase 1 protocol for further development of the diary were not implemented.

An *ad hoc* analysis of the Phase 1 data led to the deletion of questions, most of which were diarrhea-related. Poor compliance was the primary reason for deleting the questions. Subsequent analyses suggest some of these questions were poorly constructed. Steps to investigate reasons for the poor compliance and, potentially, rewording the questions were not undertaken. The revised diary was used in Phase 2 and Phase 3.

The validation analysis cited "redundancy" for the deletion of three items. Two of these, however, were not redundant to other questions within the diary. The questions, which asked what a subject did to treat MTS and diarrhea were "eliminated due to redundancy with … an item in the clinical case report form." Eliminating questions because the information is captured elsewhere is not appropriate.

Because the *ad hoc* approach taken to revising the diary used in Phase 1 simply deleted questions, the diary may not adequately represent the implications of mouth and throat soreness

[1] Cover letter dated May 14, 2004

and of diarrhea on daily activities. An iterative process, as stated in the Phase 1 protocol, for revising the diary and creating new scales was not implemented.

For the above reasons, the daily diary is not validated for assessing the impace of oral mucositis on a patient's quality of life.

2 INTRODUCTION

2.1 Overview

The submission contains three controlled studies:

1. **Phase 1, Study 960189:** A randomized, double blind, placebo controlled, dose escalation trial of the safety of recombinant human Keratinocyte growth factor (rHuKGF) in Hodgkin's disease and non-Hodgkin's lymphoma patients undergoing high dose chemotherapy with autologous peripheral blood progenitor cell transplantation.

2. **Phase 2, Study 980231**: Recombinant Human Keratinocyte Growth Factor (rHuKGF) for Reduction of Oral Mucositis in Patients with Hematologic Malignancies Undergoing Total Body Irradiation (TBI) and High- Dose Chemotherapy with Autologous Peripheral Blood Progenitor Cell (PBPC) Transplantation.

3. **Phase 3, Study 20000162**: Trial of Recombinant Human Keratinocyte Growth Factor (rHuKGF) for Reduction of Mucositis in Patients with Hematologic Malignancies Undergoing Total Body Irradiation (TBI) and High- dose Chemotherapy with Autologous Peripheral Blood Progenitor Cell (PBPC) Transplantation.

This review focuses on the development and validation of the daily diary questionnaire, which the applicant calls "Oral Mucositis Daily Questionnaire (OMDQ)". Interestingly, the study protocols for all three studies refer to the questionnaire as a "health related quality-of-life assessment"; the instrument itself is labeled "Daily Diary". The name "Oral Mucositis Daily Questionnaire (OMDQ)" appears for the first time in this submission.

The Phase 2 and Phase 3 studies used the FACT-G (Functional Assessment of Cancer Therapy-General) as an additional QoL instrument.

Dr. Yuan-Li Shen is reviewing the efficacy results for the primary endpoint.

2.2 Development of daily diary

The daily diary used in the Phase 2 and Phase 3 studies is the result of several revisions to a pilot, daily-diary, heath related quality of life (HRQOL) questionnaire.

Daily Diary used in a pilot study
Documentation for the development and revision of the pilot instrument was not provided in the submission because, apparently, the documentation is incomplete or may be missing. Instead, appendices to the clinical study reports attempt to provide an overview.

According to the study report[2] for the Phase 1 study (KGF 960189) and the report on the post-hoc validation analyses of the questionnaire used in the Phase 1 study, the pilot instrument was developed through a serious of focus groups and one-on-one cancer patient interviews to be a mucositis-specific patient reported questionnaire. The pilot was developed in consultation with Drs. Ron Hays and Karen Syrjala.

The pilot, a copy of which was not submitted, was administered to a group of patients with hematological and solid tumor cancers in different settings and modified for use in the Phase 1, dose-escalation study. The modified instrument contained 10 questions; APPENDIX 1: Daily Diary used in Phase 1 contains a copy of the instrument. An ad hoc analysis was performed following the completion of the Phase 1 study, which led to the revised version of the questionnaire which was used in the Phase 2 and Phase 3 studies.

Daily Diary used in Phase 1
The Phase 1 study protocol states the intent of the daily diary HRQOL questionnaire was to assess the patient's perception of functional status, general well-being and pain. The investigators theorized the symptoms of moderate to severe mucositis could affect the patient's ability to function at work and home, the patient's mood and level of stress. They wanted to identify the activities of daily life that are impacted by mucositis, the assessment of the duration of the impact, and the description of the relationship between pain and functional status.

The Phase 1 study hypothesis for the HRQOL aspect of the study was "rHuKGF can improve the HRQOL of patients receiving chemotherapy and compare the validity and sensitivity of specific questions and schedule of administration to measure HRQOL"[3]. Exploratory analyses would examine the hypothesis.

The study protocol and the Statistical Analysis Plan[4] specified three objectives of the exploratory HRQOL assessments:

- Test the feasibility of conducting a HRQOL assessment by examining patient compliance and determining the optimal assessment schedule,
- Explore relationships between the items in order to create a shorter form or to construct new composite scales, and
- Assess the external and internal validity for the scales or questions that represent functional status, well-being and pain.

[2] Study report for KGF 960189: Appendix 14 (page 1860) and Appendix 15 (page 1907).
[3] Study report for KGF 960189: Study protocol, (page 1336).
[4] Study report for KGF 960189: Statistical Analysis Plan (page 1428).

The study report's description of the results of the questionnaire analyses[5], on the other hand, indicates the objectives of the questionnaire used in the dose escalation study were to explore

- the impact of palifermin on the subject's overall health, mouth and throat soreness (MTS), and diarrhea;
- the impact of MTS and diarrhea on the subject's ability to perform daily activities (eg, sleeping, eating, drinking, and getting together with friends) and
- activities undertaken by subjects to deal with diarrhea and MTS.

Based on the results of the Phase 1 study, four questions were eliminated. The resulting six question questionnaire was used in Phase 2 and Phase 3 studies; APPENDIX 2: Daily Diary used in Phase 2 and APPENDIX 3: Daily Diary used in Phase 3 contain copies of the diaries.

How this item reduction occurred is not known. The clinical study report states "the original validation analysis, conducted immediately after the study completed in 1999 was not based on a formal prespecified analysis plan and the analysis was not well-documented[6]."

Daily diary used in Phase 2 and Phase 3

Instruments used

The Phase 2 and Phase 3 studies included both the daily diary and the FACT-G instruments.

Both protocols state the daily diary questionnaire

> "assesses the subject's perception of the effects of chemo-radiotherapy and its clinical manifestation of mucositis. The questions included in this questionnaire have been used in other Amgen rHuKGF studies and are designed to address specific QoL issues related to mucositis such as mouth and throat soreness, eating, swallowing, drinking, talking, sleeping, diarrhea, and overall health. This questionnaire should require approximately 3 minutes to complete"[7,8].

Study objectives and endpoints

The Phase 2 protocol lists "health related quality-of-life assessment for rHuKGF, specifically the domain of mouth and throat soreness" as a secondary endpoint[9]. A secondary objective of the study was to investigate this endpoint.

The Phase 3 protocol, however, does not include HRQoL assessments as either a primary or secondary study objective although "HRQoL assessment of mouth and throat soreness" is included as a secondary endpoint.

Table 3.1 summarizes these points.

---

[5] Study report for KGF 960189: Appendix #14, Patient Reported Outcomes (page 1859).
[6] Study report for KGF 960189: Appendix #14, Patient Reported Outcomes (page 1903).
[7] Study report for KGF 980231: Study protocol (page 1164)
[8] Study report for KGF 20000162: Study protocol (page 884)
[9] Study report for KGF 980231: Study protocol (page 1128)

Proposed analyses
The Phase 2 protocol indicates formal hypothesis testing would be done for the primary efficacy endpoint only. There would be no adjustments for multiplicity among the secondary endpoints; they would be analyzed "one-by-one".

The Phase 3 protocol states "descriptive statistics will be calculated for all secondary endpoints and compared between the treatment groups".[10] The Statistical Analysis Plan would discuss the testing sequence of the secondary endpoints and 95% confidence intervals would be provided for the treatment difference for each secondary endpoint.

2.3 Validation of daily diary

To establish the validity of the daily diary, the applicant did a retrospective, validation analysis of the questionnaire used in Phase 1. The statistical analysis plan (SAP) specified the validation analyses[11].

Consistent with the protocol for the Phase 1 study, the SAP called for the evaluation of compliance, test-retest reliability, internal validity and external validity. The report states further the validation analysis plan for the study was finalized after the study was unblinded and the clinical results analyzed. The validation analyses and the analyses of comparisons between treatment groups were exploratory.

The SAP specified three criteria for eliminating questions from the daily diary. The general categories were

- questions with limited variability in responses over the study period,
- redundant questions, defined by a correlation of >0.75 with the same or similar constructs or questions measured on the same rating scale, and
- poorly constructed questions meaning, for example, the responses were contradictory to other questions.

The analyses were conducted as if the results from the studies were unknown. The results of these retrospective analyses were compared with the modifications made to the daily diaries at the time the studies were completed.

2.4 Results of validation analyses for daily diary used in the Phase 1 study

Based on the ad hoc analysis of the Phase 1 study results, fourteen questions and sub-questions were deleted; see Table 2.1. The retrospective analysis identified all but three (8a, 8c, 9) for elimination from the questionnaire. Had the validation rules been followed, the three questions would have been retained.

---

[10] Study report for KGF 20000162: Study protocol (page 814)
[11] Study report for KGF 960189: Statistical Analysis Plan (page 1438).

The results of the validation analyses led the applicant to conclude[12]:

(1) "the compliance for the MTS- related questions was generally higher than that for diarrhea- related questions;
(2) most MTS- related questions displayed good test-retest reliability, criterion validity (both internal and external), discriminative validity, and evaluative validity (sensitivity) with respect to the WHO mucositis scale;
(3) most diarrhea-related questions also displayed a reasonable level of test- retest reliability and internal criterion validity; and
(4) the external criterion validity, discriminative validity, and evaluative validity of diarrhea-related questions with respect to the WHO mucositis scale was either poor or inconclusive."

The conclusions note the majority of the deleted questions were related to diarrhea severity and limitations. The report claims "because these questions were poorly correlated with both the WHO mucositis scale and the MTS-related questions, and diarrhea was not a common or severe symptom in this treatment setting, deleting them is unlikely" to affect the ability of the OMDQ to capture the impact of palifermin on MTS severity and limitations.

Three questions (2g, 3, 7) were deemed "redundant" and, therefore, could be deleted. However, only one (2g, brushing teeth) was correlated with other questions. The other two (treatment for MTS, treatment of diarrhea) were "eliminated due to redundancy with ... an item in the clinical case report form"[13].

[12] Study report for KGF 960189: Appendix 15 (page 1959)
[13] Study report for KGF 960189: Appendix 15 (page 1951)

Table 2.1. Questions deleted from the Phase 1 daily diary version based on an ad hoc analysis. The criterion/rationale column lists reasons identified by the retrospective validation analysis. "None" denotes the items that would not have been deleted if the criterion/rationale had been followed.

Table 5-2. OMDQ Questions Deleted by Criterion/Rationale

Question Deleted	Criterion/Rationale
2f—MTS limitation on entertainment	Lack of variability
2g—MTS limitation on brushing teeth	Redundancy
2h—MTS limitation on kissing	Lack of variability
2i—MTS limitation on leaving home	Lack of variability
2j—MTS limitation on getting together	Lack of variability
3—Treatment for MTS	Redundancy
5—Number of BM	Poorly constructed
6—Description of BM	Poorly constructed
7—Treatment of diarrhea	Redundancy
8a—Diarrhea limitation on sleeping	None
8b—Diarrhea limitation on drinking	Lack of variability
8c—Diarrhea limitation on eating	None
8d—Diarrhea limitation on entertainment	Lack of variability
8e—Diarrhea limitation on taking care of oneself	Lack of variability
8f—Diarrhea limitation on getting together	Lack of variability
8g—Diarrhea limitation on leaving home	Lack of variability
9—Overall discomfort from diarrhea	None

Source: Study report for KGF 960189: Validation analysis of patient-reported outcomes (Table 5-2, page 1952)

3 EVALUATION OF DEVELOPMENT PLAN FOR DAILY DIARY

The development of the pilot daily diary and its revision, which was used in the Phase 1 study, appeared to follow sound principles. The plan included focus groups, one-on-one interviews with cancer patients, and the evaluation of a pilot daily diary administered to a group of patients with hematological and solid tumor cancers in different settings.

Following the conclusion of the Phase 1 study, analyses and revisions to the diary appear *ad hoc*. The applicant is seeking claims based on the results from the revised diary, which was used in Phase 2 and Phase 3.

The concept to be measured by the diary is never clearly specified. The study objectives and hypotheses appear to emphasize health-related quality of life issues (Table 3.1) and are consistently deemed as one of several secondary study objectives to be assessed by exploratory analyses. The Phase 3 protocol doesn't include the diary as a secondary objective; it's included as a secondary endpoint only[14]. The Phase 3 Statistical Analysis Plan, however, specifies Question 2 (During the past 24 hours, how much mouth and throat soreness did you have?) as a secondary endpoint.

Table 3.1. Hypothesis and study objectives for each study as specified in the study protocols.

Study		Hypothesis or Study Objectives for Daily Diary Usage	Secondary Endpoint	Analyses
Pilot				
Phase 1		"rHuKGF can improve HRQoL of patients receiving chemotherapy and [to] compare the validity and sensitivity of specific questions and schedule of administration to measure HRQoL"	Yes	Exploratory
Phase 2		Investigate "health related quality-of-life assessment for rHuKGF, specifically the domain of mouth and throat soreness" as a secondary endpoint.	Yes	Exploratory
Phase 3		None.	HRQoL assessment of mouth and throat soreness	Exploratory

In today's terminology, which has changed since these protocols were written, these diaries perhaps could be called disease-specific health-related quality of life instruments. Even within the context of being a mucositis health-related quality of life instrument, however, the development of the daily diary is problematic.

The study protocol for Phase 1 stated three objective of the exploratory assessments:

- Test the feasibility of conducting a HRQOL assessment by examining patient compliance and determining the optimal assessment schedule,
- Explore relationships between the items in order to create a shorter form or to construct new composite scales, and
- Assess the external and internal validity for the scales or questions that represent functional status, well-being and pain.

[14] KGF 20000162: Protocol (page 845)

There is no evidence indicating these assessments were done. A pre-specified plan for deleting the four questions from this version of the diary is absent. Instead, the method for eliminating questions appears *ad hoc*.

Most of the deleted questions were diarrhea-related. The retrospective validation analysis for the Phase 1 study indicates compliance for the diarrhea-related questions was lower than compliance for the other questions. The report of the validation analyses states "questions 5 and 6 were intended to assess the severity of diarrhea, but these questions were poorly worded[15]". The report argues because diarrheas was not common, elimination of the diarrhea questions likely did not affect the usefulness of the daily diary.

Two questions were deleted due to "redundancy". These asked subjects if they took any medications for their mucositis or diarrhea. The reports suggests these deletions were appropriate because the information was captured on the case report forms by the study coordinator. This is not a sufficient reason for deleting the questions from the diary, because an intent of the diary is to capture the impact of mucositis. Only those items unrelated to intervention should be deleted.

The absence of the assessments stated in the Phase 1 protocol is a crucial piece missing from the development process. The initial development identified issues important to patients with oral mucositis. Among them were diarrhea-related aspects of oral-mucositis. Simply deleting them because of poor compliance is not an adequate reason, especially in light of the conclusions that some of the diarrhea-related questions were poorly worded. Apparently, investigators did not explore reasons for poor compliance. This could have been accomplished by speaking with the subjects. This information could have been used to revise the questions.

There is no evidence supporting a rigorous approach to creating a shorter form or developing new composite scales. In addition to eliminating questions, a shorter form can be obtained by combining questions or developing new scales. This was not done.

Because of the manner in which questions were deleted, the diary may not adequately represent the implications of mouth and throat soreness and of diarrhea on daily activities.

[15] Study report for KGF 960189: Appendix 15 (page 1953)

4 SIGNATURES/DISTRIBUTION LIST PAGE

Primary Statistical Reviewer: **Date:** December 1, 2004

Lisa A. Kammerman 12/1/04

Lisa A. Kammerman, Ph.D.
Mathematical Statistician

Concurrence: not needed

cc:
HFD-109 / Susan Giuliani, R.N., M.S.
HFM-99 / Pat Dinndorf, M.D.
HFD-711 / Mark Rothmann, Ph.D.
HFD-711 / Aloka Chakravarty, M.D.
HFD-715 / Lisa Kammerman, Ph.D.
HFD-700 / Charles Anello, Sc.D.

5 APPENDIX 1: Daily Diary used in Phase 1

10. ORAL MUCOSITIS DAILY QUESTIONNAIRE (OMDQ)

KGF 960189

D D MON Y Y	Patient initials	Patient No.	Day -9	Serial No. 0247

1. During the PAST 24 HOURS, how much **MOUTH AND THROAT SORENESS** did you have? *(Mark an "X" in only one box)*

No soreness	❑ [0]	*If you marked "No soreness", please skip to Question 5*
A little soreness	❑ [1]	
Moderate soreness	❑ [2]	
Quite a lot of soreness	❑ [3]	
Extreme soreness	❑ [4]	

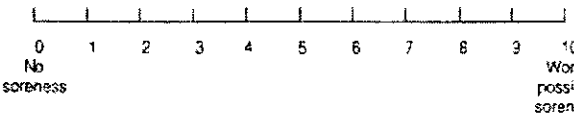
2. During the PAST 24 HOURS, how much did **MOUTH OR THROAT SORENESS** limit you in each of the following activities? *(Mark an "X" in one box on each line)*

	Not at all limited	A little limited	Limited	Very limited	Unable to do	Not applicable
a. Sleeping	❑	❑	❑	❑	❑	❑
b. Swallowing	❑	❑	❑	❑	❑	❑
c. Drinking	❑	❑	❑	❑	❑	❑
d. Eating	❑	❑	❑	❑	❑	❑
e. Talking	❑	❑	❑	❑	❑	❑
f. Things you do for entertainment	❑	❑	❑	❑	❑	❑
g. Brushing your teeth	❑	❑	❑	❑	❑	❑
h. Kissing	❑	❑	❑	❑	❑	❑
i. Leaving your home	❑	❑	❑	❑	❑	❑
j. Getting together with friends or relatives	❑	❑	❑	❑	❑	❑

3. During the PAST 24 HOURS what did you do for **MOUTH AND THROAT SORENESS**? *(Mark an "X" in one box on each line)*

	Yes	No	Not applicable
a. I used mouth rinses for my mouth and throat soreness	❑	❑	❑
b. I used pain medications that my doctor prescribed or nurse gave to me	❑	❑	❑
c. I used pain medications that were not prescribed by my doctor	❑	❑	❑
d. I did nothing at all	❑	❑	❑

4. On a scale from 0 to 10, how would you rate your **OVERALL MOUTH AND THROAT SORENESS** during the PAST 24 HOURS? *(Circle one number)*



KGF 960189

D D MMC Y Y | Patient Initials | Patient No. | Day -9 | Serial No. 0267

5. During the PAST 24 HOURS, **HOW MANY BOWEL MOVEMENTS** did you have? *(Mark an "X" in only one box)*

0	☐
1 - 3	☐
4 - 6	☐
7 - 9	☐
More than 9	☐

If you marked 0, please skip to question 10

6. Which of the following statements best describes most of your **BOWEL MOVEMENTS** during the PAST 24 HOURS? *(Mark an "X" in one box on each line)*

	Yes	No
a. My bowel movements were liquid (diarrhea)	☐	☐
b. My bowel movements were very soft, but not liquid	☐	☐
c. My bowel movements were well formed or hard	☐	☐

If you did not have diarrhea, please skip to question 10

7. During the PAST 24 HOURS, what did you do for **DIARRHEA**? *(Mark an "X" in one box on each line)*

	Yes	No	Not applicable
a. I took non prescription medication for my diarrhea	☐	☐	☐
b. I took medication my doctor prescribed or my nurse gave me	☐	☐	☐
c. I ate certain foods	☐	☐	☐
d. I had diarrhea, but did nothing for it	☐	☐	☐

8. During the PAST 24 HOURS, how much did **DIARRHEA** limit you in each of the following activities? *(Mark an "X" in one box on each line)*

	Not at all limited	A little limited	Limited	Very limited	Unable to do	Not applicable
a. Sleeping	☐	☐	☐	☐	☐	☐
b. Drinking	☐	☐	☐	☐	☐	☐
c. Eating	☐	☐	☐	☐	☐	☐
d. Things you do for entertainment	☐	☐	☐	☐	☐	☐
e. Taking care of yourself	☐	☐	☐	☐	☐	☐
f. Getting together with friends or relatives	☐	☐	☐	☐	☐	☐
g. Leaving your home	☐	☐	☐	☐	☐	☐

9. On a scale from 0 to 10, how would you rate your **OVERALL DISCOMFORT FROM DIARRHEA** during the PAST 24 HOURS? *(Circle one number)*

0 1 2 3 4 5 6 7 8 9 10
No discomfort — Worst possible discomfort

10. How would you rate your **OVERALL HEALTH** during the PAST 24 HOURS? *(Circle one number)*

0 1 2 3 4 5 6 7 8 9 10
Worst possible (as bad as being dead) — Half-way between worst possible and perfect health — Perfect health

6 APPENDIX 2: Daily Diary used in Phase 2

AMGEN
KGF 980231

DAILY DIARY

1. How would you rate your **OVERALL HEALTH** during the PAST 24 HOURS? *(Circle one number)*

0 1 2 3 4 5 6 7 8 9 10
Worst possible — Half-way between worst possible and perfect health — Perfect health

2. During the PAST 24 HOURS, how much **MOUTH AND THROAT SORENESS** did you have? *(Circle one number)*

No soreness	0 ➡
A little soreness	1
Moderate soreness	2
Quite a lot of soreness	3
Extreme soreness	4

If you circle 0, please skip to question 5.

3. During the PAST 24 HOURS, how much did **MOUTH OR THROAT SORENESS** limit you in each of the following activities? *(Circle one number on each line)*

	Not Limited	Limited A Little	Limited Some	Limited A Lot	Unable To Do
a. Swallowing	0	1	2	3	4
b. Drinking	0	1	2	3	4
c. Eating	0	1	2	3	4
d. Talking	0	1	2	3	4
e. Sleeping	0	1	2	3	4

4. On a scale from 0 to 10, how would you rate your **OVERALL MOUTH AND THROAT SORENESS** during the PAST 24 HOURS? *(Circle one number)*

0 1 2 3 4 5 6 7 8 9 10
No soreness — Worst possible soreness

5. During the PAST 24 HOURS, how much **DIARRHEA** did you have? *(Circle one number)*

No diarrhea	0 ➡
A little diarrhea	1
Moderate diarrhea	2
Quite a lot of diarrhea	3
Severe diarrhea	4

If you circle 0, please STOP HERE

6. On a scale from 0 to 10, how would you rate your **OVERALL DIARRHEA** during the PAST 24 HOURS? *(Circle one number)*

0 1 2 3 4 5 6 7 8 9 10
No diarrhea — Worst possible diarrhea

Distribution: White - Amgen; Pink - Site

7 APPENDIX 3: Daily Diary used in Phase 3

AMGEN KGF 20000162	Date: Day / Month / Year	Site No.	Subject No.	Subject Initials

1. How would you rate your **OVERALL HEALTH** during the PAST 24 HOURS?
(Circle one number)

0 1 2 3 4 5 6 7 8 9 10
Worst possible — Half-way between worst possible and perfect health — Perfect health

2. During the PAST 24 HOURS, how much **MOUTH AND THROAT SORENESS** did you have?
(Circle one number)

No soreness	0	*If you circled 0, please skip to question 5.*
A little soreness	1	
Moderate soreness	2	
Quite a lot of soreness	3	
Extreme soreness	4	

3. During the PAST 24 HOURS, how much did **MOUTH AND THROAT SORENESS** limit you in each of the following activities?
(Circle one number on each line)

	Not Limited	Limited A Little	Limited Some	Limited A Lot	Unable To Do
a. Swallowing	0	1	2	3	4
b. Drinking	0	1	2	3	4
c. Eating	0	1	2	3	4
d. Talking	0	1	2	3	4
e. Sleeping	0	1	2	3	4

4. On a scale from 0 to 10, how would you rate your **OVERALL MOUTH AND THROAT SORENESS** during the PAST 24 HOURS? *(Circle one number)*

0 1 2 3 4 5 6 7 8 9 10
No soreness — Worst possible

5. During the PAST 24 HOURS, how much **DIARRHEA** did you have?
(Circle one number)

No diarrhea	0	*If you circled 0, STOP here.*
A little diarrhea	1	
Moderate diarrhea	2	
Quite a lot of diarrhea	3	
Severe diarrhea	4	

6. On a scale from 0 to 10, how would you rate your **OVERALL DIARRHEA** during the PAST 24 HOURS? *(Circle one number)*

0 1 2 3 4 5 6 7 8 9 10
No diarrhea — Worst possible diarrhea