CENTER FOR DRUG EVALUATION AND RESEARCH AND CENTER FOR BIOLOGICS EVALUATION AND RESEARCH

APPLICATION NUMBER:

125031/0

STATISTICAL REVIEW(S)

COPY

Biostatistical Review

BLA:	**STN# 125031/0** *Filgrastim-SD/01 (Pegfilgrastim) for decreasing the duration of severe neutropenia in patients with non-myeloid malignancies receiving myelosuppressive anti-cancer drugs associated with a clinically significant incidence of febrile neutropenia (FN)* Submission received 2nd Quarter of 2001 Amgen, Inc. Thousand Oaks, CA

Date:	January 10, 2002
Reviewer:	***Clare Gnecco, Ph.D.***
Through:	***Ghanshyam Gupta, Ph.D.*** ***Branch Chief, Therapeutics Group***

1-15-02

cc:

HFM-99/DCC: BLA #125031/0
HFM-573/ Dr. Gootenberg
HFM-570 / Dr. Keegan
HFM-588/Ms. Giuliani / Mr. Crim
HFM-215/Chron – File: BLAPEGF.DOC

STATISTICAL REVIEW ISSUES / SUMMARY: The sponsor's major efficacy analyses were investigated and statistical findings confirmed for the two pivotal studies, #980226 and #990749, which are the subject of this review. However, a statistically significant claim for superiority in terms of FN incidence for study #980226 is not

warranted since an endpoint multiplicity adjustment was not applied to the sponsor's reported confidence interval.

BACKGROUND:

Pegfilgrastim (Filgrastim-SD/01) was developed as a sustained duration form of Filgrastim that would allow a single injection within a chemotherapy cycle in contrast to once daily injections through ANC (absolute neutrophil count) recovery. The clinical development of Pegfilgrastim focused on demonstrating non-inferiority as compared to Filgrastim and the development program was initiated in 1997. The Integrated Summary of Efficacy included data from 599 subjects, 289 who received Filgrastim and 310 who received Pegfilgrastim. There were 390 subjects who participated in the two pivotal trials and 209 in the four supportive trials. The trials were conducted in the US, Canada, Europe, and Australia. Both pivotal studies were conducted in subjects with breast cancer (high-risk stage II and stages III/IV). Both trials used doxorubicin and docetaxel chemotherapy. Dosages used were 60 mg/m^2 and 75 mg/m^2, respectively.

SUMMARY OF STUDY #980226: This study was entitled ***"A Blinded, Randomized, Multicenter Study to Evaluate Single Administration Filgrastim-SD/01 per Cycle versus Daily Filgrastim as an Adjunct to Chemotherapy in Subjects with High-risk Stage II or Stage III/IV Breast Cancer."*** It was conducted at 62 study sites in the US. Subjects were randomized to receive, in conjunction with up to 4 cycles of doxorubicin/docetaxel chemotherapy, either Filgrastim 5 μg/kg once daily through ANC recovery or a single dose of Filgrastim-SD/01 100 μg/kg followed by daily placebo injections through ANC recovery. This was a weight-dose study. Study enrollment was 310 subjects, 156 randomized to Filgrastim and 154 to Filgrastim-SD/01. Of the 310 randomized subjects, 150/156 (96%) Filgrastim subjects received at least one dose of study drug while 151/154 (98%) Filgrastim-SD/01 subjects received at least one dose of study drug. Of the 310 subjects randomized, 284 (92%) subjects completed the active treatment period.

Study Endpoints:

The **primary efficacy endpoint** was the duration of severe neutropenia (DSN) in cycle 1. Severe neutopenia is defined as ANC < 0.5 x 10^9/L. There were seven **protocol specified secondary efficacy endpoints**. These are: (i) the DSN in each of cycles 2 through 4 and (2) the depth of ANC nadir in each of cycles 1 through 4. Additional efficacy endpoints analyzed were rates of FN by cycle and across all cycles, time to ANC recovery by cycle and across all cycles, and ANC-time profiles by cycle.

Reviewer's Comment: It is noted that the rates of FN by cycle and across all cycles and the times to ANC recovery in cycles 2 through 4 were not mentioned in the Statistical Analysis Plan (SAP).

Randomization: Randomization to study treatment was carried out centrally via a centralized telephone system. Subjects were randomized (1:1) to one of the two blinded

treatment groups. The randomization was stratified by study center and prior chemotherapy utilizing a stratified permuted-block randomization schedule.
Sample Size: Sample size estimates were calculated using a non-inferiority design and the normal approximation for two independent groups. The planned enrollment for this study was 300 subjects, 150 per treatment group. The average DSN in the absence of cytokine support was based on literature reports on doxorubicin and docetaxel chemotherapy and was found to be similar to that observed in Amgen study #8801, which was pivotal for the approval of NEUPOGEN® (Filgrastim) in the setting of chemotherapy-induced neutropenia. At the time of study design, available data were limited on breast cancer patients treated with this chemotherapy regimen in the presence of cytokine support and mainly consisted of data from patients who had received no previous chemotherapy. Thus, the pooled variance was estimated from Amgen study #8801. Using these estimates and a conservative estimate of 2.17 days for the standard deviation of DSN, a planned sample size of 150 subjects/treatment group yields 95% power to support a conclusion of non-inferiority based on the 97.5% UCL of the difference being < 1 day. A 15% dropout rate during the first cycle was assumed. In addition, the protocol states that this sample size will also give an $\alpha = 0.025$ test of the null hypothesis that the true incidence of febrile neutropenia with Filgrastim-SD/01 is more than 14 percentage points higher than the estimated 12 % with Filgrastim, and an $\alpha = 0.025$ test of the null hypothesis that the true incidence of bone pain with Filgrastim-SD/01 is more than 18.5 percentage points higher than the estimated 25% with Filgrastim.

SUMMARY OF STUDY #990749: This study was entitled ***"A Blinded, Randomized, Multicenter Study to Evaluate Fixed Dose Single Administration Filgrastim-SD/01 per Cycle versus Daily Filgrastim as an adjunct to Chemotherapy in Subjects with High-risk Stage II or Stage III/IV Breast Cancer."*** It was conducted in the US, Europe, and Australia at 37 study sites. This study is identical to #980226 with the exception that a fixed 6 mg dose of Filgrastim-SD/01 was employed. In fact, 34 centers randomized a total of 157 subjects; 80 were randomized to Filgrastim-SD/01 and 77 to Filgrastim. Of the 157 randomized subjects, 155 (99%) received at least one dose of study drug. Ninety two percent (145 subjects) completed the treatment phase of this study (cycles 1 through 4). Completing 3 months of follow-up were 146 subjects.
Study Endpoints:
The **primary efficacy endpoint** was the duration of severe neutropenia (DSN) in cycle 1.
The **protocol specified secondary efficacy endpoints** included: (i) duration of severe neutropenia in each of cycles 2 through 4 and (ii) depth of ANC nadir in each of cycles 1 through 4. In addition, cycle 1 incidence of febrile neutropenia, and cumulative incidence of severe neutropenia were also analyzed.
Randomization: Randomization to study treatment was carried out centrally. Subjects were randomized (1:1) to one of the two blinded treatment groups. The randomization was stratified by body weight (< 50 kg, ≥ 50 kg and < 80 kg, or ≥ 80kg), prior chemotherapy exposure, and geographic location (US vs. non-US).
Sample Size: Sample size estimates were calculated using a non-inferiority design and the normal approximation for two independent groups. The average duration and standard deviation for severe neutropenia (SN) was based on that observed in Amgen

study #980147. Using these estimates and a conservative estimate for the standard deviation of duration of 1.5 days, a sample size of 75 subjects/treatment arm was planned. This yields a 95% probability that a one-sided upper 97.5% CL would show non-inferiority of Filgrastim-SD/01 with < 1 day difference in mean DSN to Filgrastim under the assumption of no difference between the treatments. This sample size allowed for a 20% dropout rate during the first cycle, which was higher than that observed in #980147.

Interim Analysis: No comparative interim analyses for efficacy were conducted for study #990749. An external SDMC periodically reviewed blinded aggregate safety data. For study #980226, the SDMC examined antibody assay data, hematopoietic recovery data, reported adverse events, and demographics in a blinded fashion. If concerns about trends in the data (e.g., unexpected differences in ANC profile were encountered), the SDMC could request unblinding of the data and were required to make recommendations regarding study continuation. To protect against the SDMC prematurely recommending study discontinuation, a p-value of 0.0001 was specified in the protocol for study #980226.. Therefore, adjustment of the final significance or confidence level was not required at the end of the study. For both studies, after accrual of half of the subjects, the variance of the DSN in cycle 1 was calculated on the blinded data. If the resulting estimated variance was ≥ 33% of the estimate on which the original sample size was based, the sample size was to be recalculated using the updated estimate. If the variance was < 3.00, the sample size was not to be recalculated and recruitment continued until the original planned number of subjects had been enrolled. In both cases, it was found that sample size augmentation was not needed. No adjustment was made in the confidence level for the CI's for the final analysis, since this type of sample size re-estimation has a negligible effect on the α-level of the final analysis (Wittes and Brittain, 1990).

Analysis Populations: Since the primary goal was to demonstrate non-inferiority, the designated primary analysis group for both studies was the per-protocol (pP) population. The one exception was in the analysis of febrile neutropenia, which used the modified intent-to-treat (m-ITT) group. Due to cycle exclusions, different subjects are represented in the various cycles and the m-ITT analysis was a more comprehensive representation of cumulative incidence. The pP subset consists of those subjects who were randomized and who had not deviated from key eligibility criteria and cycle-specific protocol requirements that could affect the efficacy endpoints. The m-ITT subset consists of those subjects who were randomized and were exposed to assigned study-drug treatment and who received study drug within a cycle and who had at least one post-cycle chemotherapy ANC.

Missing Data: Except for two special cases, missing ANC values between the first and last observed values in each cycle were estimated using linear interpolation. The two exceptions to this rule were the following: (i) Any missing ANC values within the expected period of severe neutropenia that were part of a consecutive series of missing values were replaced by values indicating severe neutropenia and (ii) Two or more consecutive missing ANC values bounded by an ANC < 0.5×10^9/L and another ANC < 0.5×10^9/L were replaced by values indicating severe neutropenia.

Statistical Methodology: The statistical analysis plans for the pivotal studies called for the calculation of a one-sided confidence limit (CL) for the difference in the means of each endpoint between the Filgrastim-SD/01 and Filgrastim treatment groups. The 95% CI for the duration of severe neutropenia in the pivotal studies was calculated as the 2.5 and 97.5 percentiles of the empirical sampling distribution of the difference in treatment group means resulting from 10,000 independent samples, drawn with replacement, from the observed data (bootstrapping). The software used for this analysis was the validated and commercially available Resampling Stats (version 5.0). Stratification by body weight and prior chemotherapy was employed for study #990749 and by prior chemotherapy for study #980226. For the resampling analysis for study #990749, stratification by body weight, prior chemotherapy, and geographic location would not have provided sufficient numbers of subjects within each stratum. Therefore, the location strata were not used because they were considered to be the least clinically relevant factor. The non-inferiority margin for the pivotal studies was defined as the upper bound of the 95% CI of the mean duration of severe neutropenia in the Filgrastim-SD/01 group being less than 1 day longer than that of the Filgrastim group. Analyses of the primary endpoint, DSN, were also conducted by age group (≤ 50 vs. > 50), sex, and race. Rates of febrile neutropenia (FN) were also analyzed by cycle and all cycles combined. CI's for other than DSN endpoints were calculated using the normal theory approximation and no stratification was employed.

SPONSOR's EFFICACY RESULTS FOR PIVOTAL STUDIES:

In study #980226, 146/154 (95%) of Filgrastim-SD/01 subjects and 141/156 (90%) of Filgrastim subjects were evaluable for the pP analysis subset in at least one cycle. In study #990749, 71/80 (89%) of Filgrastim-SD/01 subjects and 64/77 (83%) of Filgrastim subjects were evaluable for the pP analysis in at least one cycle.

Patients were well balanced on baseline demographic and medical characteristics at study entry.

Primary Efficacy Endpoint: Results for the primary efficacy endpoint of DSN in cycle 1 indicate that Filgrastim-SD/01 is statistically comparable to Filgrastim in both pivotal studies as the following table with the sponsor's findings indicates:

TABLE 1. Cycle 1 Duration of Severe Neutropenia (DSN) in Pivotal Trials

Study #980226			Study #990749	
N = 301			N = 155	
	Peg-F	F	Peg-F	F
Mean days of SN	1.7	1.6	1.8	1.6
Δ in Means	0.09		0.18	
95% CI for Δ	(-0.23, 0.40)		(-0.23, 0.61)	

Reviewer's Comment: The confidence intervals are based on bootstrap re-sampling as previously described. This reviewer verified the sponsor's findings that the one-day non-inferiority margin was not exceeded for either study. Thus, the statistical non-inferiority claim is justified.

Three study sites in study #980226 were found to have sub-investigators with substantial financial interests. Summary statistics for the primary endpoint of DSN in cycle one are presented for these sites individually and for the data set excluding these sites.

TABLE 2. Cycle 1 DSN Summary for High Financial Disclosure Study Sites for Study #980226

Site#	Treatment	n	Median	Range	Mean	Std. Dev
55	F	4	1.5		1.5	1.29
	Peg-F	5	2.0		2.2	0.84
47	F	8	0.5		0.8	0.89
	Peg-F	8	0.5		1.0	1.19
43	F	6	2.0		1.8	1.33
	Peg-F	6	0.0		0.8	1.60
All Others	F	110	2.0		1.6	1.19
	Peg-F	112	2.0		1.7	1.42

***Reviewer's Comment*:** Given the small sample sizes, if one were to exclude any of these from the bootstrap re-sampling CI determination, it would make little difference. Note that excluding the three centers yields mean estimates identical to what the sponsor found for the full per protocol data set analysis.

TABLE 3. Incidence of Febrile Neutropenia (FN) Including All Cycles

Study #980226			Study #990749	
N = 301			N = 155	
	Peg-F	F	Peg-F	F
Incidence	9%	18%	13%	20%
Δ	-9%		-7%	
95% CI for Δ	(-17%, -1%)		(-19%, 5%)	

***Reviewer's Comment*:** Neither protocol for the pivotal studies pre-specifies FN as a major secondary endpoint. For study #980226, a statistically significant claim for lower FN incidence favoring Filgrastim-SD/01 (Peg-F) is not warranted even though the CI for

the difference in proportions excludes zero. FN was one of a number of efficacy endpoints analyzed and no multiplicity adjustment was applied to the reported confidence interval. If one were applied, the CI would be wider and include zero.

OVERALL SUMMARY AND CONCLUSIONS:

This reviewer's analyses of the major efficacy endpoints for the two pivotal studies, based on the electronic database provided, confirm the sponsor's reported statistical findings. However, a statistically significant claim for superiority in terms of FN incidence for study #980226 is not warranted since an endpoint multiplicity adjustment was not applied to the reported confidence interval.

APPEARS THIS WAY
ON ORIGINAL