CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER: 21-302

STATISTICAL REVIEW(S)

STATISTICAL REVIEW AND EVALUATION

Date Review Completed: 10/05/01

NDA #: 21-302

Date of Received by the Center: 12/15/00

Name of Drug: Elidel (1% SDZ ASM 981) Cream 1%

Indication: Treatment of atopic dermatitis

Applicant: Novartis

Documents Reviewed: Volumes 1.1, 1.50; 1.156-1-235, 1.291-1.323, dated 12.15.00,
Volumes 5.1-5.74 dated 5.25.01, and
Volumes 9.1-9.12 dated 9.12.01.

Project Manager: Mildred Wright (HFD-540)

Medical reviewer: Denise Cook, M.D. (HFD-540)

Statistical reviewer: Valeria Freidlin, Ph.D. (HFD-725)

Statistical Team Leader: Mohamed Alosh, Ph.D. (HFD-725)

Table of Contents:

INTRODUCTION AND BACKGROUND

The sponsor submitted reports of three pivotal trials (Studies B305, B307, and 316) to demonstrate that Elidel (1% SDZ ASM 981 cream) is safe and effective in the treatment of atopic dermatitis (AD) [redacted] The US Studies B305 and B307 enrolled pediatric subjects age 2 to 17 years old and had identical designs. A foreign Study 0316 enrolled subjects 3 to 23 months old and had a similar design. In addition, the sponsor submitted report of a Phase 2 dose-ranging

study B202. As Studies B305, B307, and 0316 had similar designs, only design of Study B305 is presented. Throughout this review, the term "ASM cream" is used instead of Elidel (1% SDZ ASM 981 cream).

2. DESIGN OF STUDY B305

The primary objective of Study B305 was to demonstrate efficacy and safety of ASM cream compared to vehicle after 6 weeks double-blind treatment in pediatric subjects 2-17 years old with mild to moderate AD.

Overall study design

This multicenter study in pediatric subjects from 2 to 17 years of age with mild to moderate AD had two phases: a 6-week double-blind (DB) phase and a 20-week open label (OL) phase.

In the 6-week, double-blind, vehicle-controlled, parallel-group phase, pediatric subjects were randomized in a 2:1 ratio to receive twice daily either ASM cream or corresponding vehicle. Study drug treatment continued until complete clearance of AD, or completion of the 6-week, double-blind phase. Subjects subsequently entered the 20-week, open-label phase to treat ongoing dermatitis and/or disease recurrence(s) with ASM cream as needed. If all lesions were cleared prior to 20 weeks, the subjects continued to be evaluated but applied study drug only as needed. Further follow-up continued for approximately 4 weeks after completing the open-label phase to assess safety.

Treatment assignment and blinding

Subjects who met the eligibility criteria, were randomized to one of two treatment groups at Day 1 and were given a unique 7-digit subject number that remained with the subject throughout the study. The first 3 digits represented the center code assigned by the sponsor, and the next 4 digits represented a unique subject identification number assigned sequentially by the investigator. Study drug was dispensed by the study center sequentially by the randomization numbers as subjects came in for their Baseline study visit. A 2:1 randomization ratio of active to vehicle treatment was used, and treatment assignments were balanced within and between centers. The study medications for the 6-week, double-blind phase were identical in appearance, smell and texture.

Sponsor's **primary efficacy variable** was Investigator's Global Assessment (IGA). The IGA score provided a static overall evaluation of AD. Subjects were to have Baseline scores of 2 (mild) or 3 (moderate) for study inclusion. In agreement with the Division's recommendation, treatment success was defined as a score of 0 or 1 at the end of the double-blind phase. Table 305-f presents the IGA scores and descriptions.

Table 305-1. IGA scores and descriptions

Score	Description	Binary category
0 = Clear	No inflammatory signs of atopic dermatitis	Treatment Success
1 = Almost clear	Just perceptible erythema, and just perceptible papulation/infiltration	
2 = Mild disease	Mild erythema, and mild papulation/infiltration	Treatment Failure
3 = Moderate disease	Moderate erythema, and moderate papulation/infiltration	
4 = Severe disease	Severe erythema, and severe papulation/infiltration	
5 = Very severe disease	Severe erythema, and severe papulation/infiltration with oozing/crusting	

Sponsor's secondary efficacy variables were:

- All-category analysis of IGA at intermediate time points.
- EASI defined as a composite score that evaluated the severity of the four key signs of AD (ie, erythema, infiltration/papulation, excoriation and lichenification), and the extent of disease in each of four body regions (ie, head, trunk, upper limbs, and lower limbs). The formula, which weights body region proportions, was adjusted for the subject's age.
- Investigator's assessment at each visit for the absence or presence of oozing/crusting, hyperpigmentation, hypopigmentation, dry skin/xerosis and other signs.
- Severity score of pruritus

Key inclusion criteria

- Any gender or race, from 2 to 17 years of age,
- A clear diagnosis of AD that affected ≥5% of TBSA (total body surface area) and with a Baseline IGA of AD scored as 2 (mild) or 3 (moderate).

Sponsor's Statistical Methods

Populations

The following subject populations were defined:

- **Randomized** - all subjects who were randomized.
- **Safety** - all subjects randomized, who were dispensed study medication.
- **ITT** - all subjects randomized, who were dispensed study medication.

- PP - all subjects in the ITT population, who adhered to the protocol without any major deviations, adhered to the inclusion/exclusion criteria and did not violate the protocol in any way that effected efficacy evaluation. Primary efficacy analysis was based on the ITT population. A PP analysis was performed to support the primary efficacy analysis.

Background and demographic characteristics

The age of the subject was defined as the age (in years) at the pre-treatment visit. The age categories were defined as follows: <2 years, 2 - <12 years, 12 - <18 years.

For qualitative variables (e. g., gender, race) Fisher's exact test was employed, while a Wilcoxon test was used for quantitative data.

Efficacy evaluation

Primary efficacy analysis

The primary efficacy variable was the IGA, the primary efficacy population was ITT.

A dichotomized IGA (0, 1 = success; 2–5 inclusive = failure) at Endpoint was defined as the primary efficacy variable. The Endpoint was defined as the end of double-blind, randomized treatment, or the last post-Baseline efficacy measure in case of early discontinuation, or Baseline measure if no post-Baseline available.

The primary statistical null hypothesis, of no difference in proportions between the ASM and vehicle treatment groups, was to be analyzed using a CMH test stratified for center.

Sponsor' s secondary efficacy variables

Secondary analysis on IGA

The primary analysis (CMH) for the primary efficacy variable (dichotomized IGA) was repeated at the intermediate time points. An analysis of the all-category frequency distribution of IGA scores (0-5) at each visit was also performed, using the CMH row mean score test for treatment difference

Pruritus

The subject's assessment of pruritus on an ordinal 0-3 scale was analyzed descriptively over time and at endpoint (double-blind phase), and presented as frequency tables by treatment group over time. Differences between treatments were assessed by the CMH row mean score test stratifying by center. Dichotomization of the scores into absence (score = 0) and presence of pruritus (score of 1, 2 or 3) were presented and descriptively analyzed using the CMH general association test stratifying by center. This procedure was repeated using a dichotomization of absence/mild presence (score of 0 or 1) and presence (2 or 3), an addition to the protocol.

Safety

When calculating the number of AEs within a group, the number of subjects with the event, rather than the total number of events, was always calculated. For multiple occurrences of the same AE, summaries were made according to the 'worst severity'.

Interim analyses

No formal interim analysis was planned or performed for Studies 305 and 307. However, the main statistical analysis was performed once all data from the double-blind phase of the study were locked. This incorporated all the efficacy analyses and all available safety analyses.

There were three locks: one after all double-blind data is in; an administrative lock on open-label data for the submission; and a final database lock once the study completes.

For infant Study 0316, according to Amendment 1, dated June 23, 2000, an interim analysis was to be performed to evaluate efficacy data of the subjects who were randomized by May 31, 2000 and completed by July 12, 2000. In the Amendment, the significance level for the interim analysis was set at 0.005 (two-sided) with the nominal significance level for the final analyses at 0.047 (two-sided).

Reviewer's Comments:

1. In agreement with the medical division, the reviewer's primary efficacy variable was the same as the sponsor's, the proportion of subjects with the score 0 or 1 in IGA at Week 6. It was analyzed using a CMH test stratified for center. Homogeneity of the odds was assessed using the Breslow-Day test. The reviewer's primary efficacy results were the same as the sponsor's. As many subjects responded to the ASM cream treatment by Week 3, this review also shows the results at Week 3

2. In agreement with the Medical Division, this reviewer did not use the sponsor's secondary efficacy variable EASI. This reviewer used the following secondary efficacy variables:

- **Four individual signs of AD: erythema, induration/papulation, excoriations and lichenification. Each sign assessment was dichotomized into "less or equal to 1" versus "greater than 1".**
- **The subject's assessment of pruritus. The subject's assessment of pruritus was based on an ordinal 0-3 scale. This assessment was dichotomized into "absence" (score = 0) and "presence of pruritus" (score of 1, 2 or 3). Another dichotomization of pruritus assessment was also examined based on "absence/mild presence" (score of 0 or 1) and "presence" (2 or 3).**
- **In addition, all-category analysis of the IGA at Week 6 was used to support the analysis of the dichotomized IGA at Week 6. All-category analysis of the IGA was performed using the CMH row mean score test.**

3. **In the safety analysis, for the difference between the treatment groups relative to the adverse events, this reviewer used the Chi-square test if all cells had expected count of 5 or greater. Otherwise, the Fisher's Exact test was used.**
4. **In Study 0316, Amendment 1 specified significance level of 0.005 for the interim analysis and 0.047 for the final analysis. As the level of 0.047 for the final analysis is more conservative than the level of 0.048 in the O'Brian/ Fleming method, this reviewer accepts the Amendment's specifications.**
5. **For the integrated efficacy subgroup analysis, the reviewer's analysis was different from the sponsor's analysis. This reviewer used the Chi-square test if all cells had expected count of 5 or greater. Otherwise, the Fisher's exact test was used.**
6. **The sponsor did not provide p-values for the key individual signs of AD (erythema, induration/papulation, excoriations and lichenification), because they were components of the EASI score. This review shows the reviewer's analysis for the key individual signs of AD.**
7. **The sponsor did not provide p-values for safety analysis. This review shows the reviewer's safety results.**

3. RESULTS OF STUDY B305

Patient disposition

Table 305-2. Subjects discontinued from study double-blind phase (ITT population)

	ASM 1% (N=130) n (%)	Vehicle (N=68) n (%)	Total (N=198) n (%)	P-Value
Completed	114 (87.7)	48 (70.6)	162 (81.8)	0.004
All discontinuations	16 (12.3)	20 (29.4)	36 (18.2)	
Primary reasons for discontinuation				
Lack of efficacy	6 (4.6)	16 (23.5)	22 (11.1)	0.001
Lost to follow-up	6 (4.6)	0	6 (3.0)	0.1

A summary of subject disposition is provided in Table 305-2. Statistically significantly more patients discontinued in the vehicle group than in the ASM group (p=0.004). The predominant reasons for discontinuation from the double-blind phase in the ASM group was lack of efficacy and lost to follow-up. The rate of discontinuation for lack of efficacy was statistically significantly

(p=0.001) higher in subjects treated with vehicle than in those treated with ASM 1% and accounts for the higher rate of discontinuation in this group. No subjects on vehicle were lost to follow-up. The other reasons for discontinuation, protocol violation and AEs were comparable between the treatment groups.

Patient populations

Three populations were defined for analysis and these are summarized in Table 305-3. The PP population was defined for use in the double-blind phase of the study only.

Table 305-3. Number of subjects by population type in Study B305

Analysis population	ASM 1% n (%)	Vehicle n (%)	Total n (%)	P-value
Randomized	130	68	198	
ITT	130 (100.0)	68 (100.0)	198 (100.0)	-
PP	79 (60.8)	36 (52.9)	115 (58.1)	0.3

Denominator for percentages is randomized subjects in each treatment group.
ITT and Safety populations are identical.
ITT: All subjects randomized, who were dispensed study medication.
PP: All subjects in the ITT population, who adhered to the protocol without any major deviations, adhered to the inclusion/exclusion criteria and did not violate the protocol in any way that would affect efficacy evaluation. This population was defined only for the double-blind phase.

The ITT population comprised all subjects who were randomized. The numbers of subjects excluded from the PP population was numerically higher in the vehicle-treated group (p=0.3).

Baseline demographic and background characteristics

A summary of the Baseline demographics is provided in Table 305-4. There were no statistically significant differences between the treatment groups in age, gender, or race ($p \geq 0.32$).

Table 305-4. Baseline demographics (ITT population of Study B305)

Parameter	Parameter	ASM 1% (N=130)	Vehicle (N=68)	p-value
Age (years)	Mean ± SD range	6.9 ± 4.2 1 –17	6.4 ± 4.3 1 –16	0.320[†]
	Median	6.0	5.0	
Age group (years) N (%)				
	<2	3 (2.3)	3 (4.4)	
	2-<12	107 (82.3)	54 (79.4)	
	12-<18	20 (15.4)	11 (16.2)	
Gender (n, %)	Male Female	63 (48.5) 67 (51.5)	35 (51.5) 33 (48.5)	0.765[‡]
Race (n, %)	Caucasian Black Oriental Other	76 (58.5) 19 (14.6) 13 (10.0) 22 (16.9)	34 (50.0) 12 (17.6) 8 (11.8) 14 (20.6)	0.690[‡]

[†]Wilcoxon rank sum test
[‡]Fishers exact test

Disease characteristics at Baseline are summarized in Table 305-5. There were no statistically significant differences between the treatment groups in the severity of AD (in IGA) prior to treatment or in the %TBSA ($p \geq 0.6$). The majority of subjects (84.3%) had mild or moderate AD at Baseline, with the majority having moderate AD at Baseline (IGA score of 3). According to the protocol, subjects were to have an IGA score of 2 or 3 at Baseline, but some investigators enrolled subjects with severe (12.1%) or very severe (3.5%) disease at Baseline (IGA scores of 4 and 5). At baseline, there were no statistically significant differences between the two treatment groups relative to number of patients with IGA score ≥ 4 ($p=0.7$) and $\geq 30\%$ TBSA involved ($p=0.8$).

Table 305-5. Disease characteristics at Baseline (ITT population of Study 305)

Parameter	Parameter	ASM 1% (N=130) n (%)	Vehicle (N=68) n (%)	p-value
IGA score	2 (mild)	28 (21.5)	18 (26.5)	0.874†
	3 (moderate)	83 (63.8)	38 (55.9)	
	4 (severe)	16 (12.3)	8 (11.8)	
	5 (very severe)	3 (2.3)	4 (5.9)	
% TBSA involved	<= 5%	13 (10.0)	6 (8.8)	0.598†
	> 5% - <= 15%	33 (25.4)	20 (29.4)	
	> 15% - <= 30%	34 (26.2)	18 (26.5)	
	> 30% - <= 60%	30 (23.1)	17 (25.0)	
	> 60%	20 (15.4)	7 (10.3)	

†Mantel-Haenszel Chi-square test

Efficacy results in Study B305

Primary efficacy results

Treatment success defined as an IGA score of 0 or 1 (clear or almost clear of disease signs) is summarized in Table 305-6 for the ITT LOCF population.

Table 305-6. Treatment success† (IGA), ITT population of Study 305

	ASM 1% (N=130)	Vehicle (N=68)	P-value‡
Baseline	0	0	-
Week 3	35 (26.9)	2 (2.9)	<0.001
Week 6	**49 (37.7)**	**11 (16.2)**	**0.002**

†Defined as a score of 0 or 1 (clear or almost clear)
‡CMH test stratified by center

The primary efficacy analysis of Study B305 showed that ASM 1% treatment group was statistically significantly better (p=0.002) than the vehicle group relative to the primary efficacy variable, proportion of subjects clear or almost clear of disease at Week 6. Supportive analysis produced similar result: p=0.047 in the PP population.

Secondary efficacy results

All-category analysis (Distribution of IGA scores)

The distribution of IGA scores is summarized in Table 305-7. There was a statistically significant ($p<0.001$) difference in favor of ASM cream between the treatment groups in the all category analysis of IGA scores at Week 6.

Table 305-7. Frequency distribution of IGA at Day 43 (ITT population of Study 305)

Time point	Group N	IGA score						p-value‡
		0	1	2	3	4	5	
Week 3	ASM N=130	8 (6.2%)	27 (20.8%)	44 (33.8%)	42 (32.3%)	8 (6.2%)	1 (0.8%)	<0.001
	Vehicle N=68	0	2 (2.9%)	17 (25.0%)	32 (47.1%)	13 (19.1%)	4 (5.9%)	
Week 6	ASM N=130	13 (10.0%)	36 (27.7%)	32 (24.6%)	38 (29.2%)	9 (6.9%)	2 (1.5%)	**<0.001**
	Vehicle N=68	0	11 (16.2%)	11 (16.2%)	29 (42.6%)	15 (22.1%)	2 (2.9%)	

IGA categories: 0=Clear, 1=Almost Clear, 2=Mild, 3=Moderate, 4=Severe, 5=Very Severe
‡CMH row mean score test, stratified by center

Individual key signs of AD

The reviewer's analysis of the proportion of subjects with sign scores of 1 or less (mild or absent symptoms) for each of the key signs of AD, erythema, induration/papulation, excoriation and lichenification is shown in Table 305-8. At Baseline, there were no statistically significant differences between the treatment groups in the proportion of subjects with mild or absent signs of erythema, induration/papulation, and lichenification ($p\geq0.329$). At Baseline, statistically significantly ($p=0.017$) more subjects treated with ASM 1% had mild or absent symptoms of excoriation compared with vehicle.

Table 305-8. Number (%) of subjects with mild or absent key signs of AD in the double-blind phase, (ITT population of Study 305, LOCF)

Visit	Treatment group	N	Erythema n (%)	Induration/ papulation n (%)	Excoriation n (%)	Lichenification n (%)
Baseline	ASM 1%	130	24 (18.5)	28 (21.5)	53 (40.8)	43 (33.1)
	Vehicle	68	14 (20.6)	14 (20.6)	17 (25.0)	18 (26.5)
	P-value*		0.639	0.767	0.017	0.329
Week 3	ASM 1%	130	74(56.9)	75 (57.7)	76 (58.5)	65 (50.0)
	Vehicle	68	15 (22.1)	19 (27.9)	23 (33.8)	26 (38.2)
	P-value*		<0.001	<0.001	<0.001	0.110
Week 6	ASM 1%	130	74 (56.9)	70 (53.8)	70 (53.8)	68 (52.3)
	Vehicle	68	16 (23.5)	23 (33.8)	29 (42.6)	28 (41.2)
	P-value*		**0.001**	**0.005**	0.131	0.137

* - CMH test (general association).

At Week 3, ASM cream was statistically significantly better than vehicle relative to proportion of patients with mild or absent signs of erythema, induration/papilation, and excoriation ($p<0.001$). At the primary time point, Week 6, ASM cream was statistically significantly better than vehicle relative to proportion of patients with mild or absent signs of erythema and induration/papilation ($p\leq0.005$).

Pruritus

A summary of pruritus assessment is provided in Table 305-9. At Baseline, there was no difference between the treatment groups in severity of pruritus ($p=0.644$), but statistically significantly ($p=0.016$) more subjects on vehicle had no pruritus (score of 0) at Baseline compared with subjects treated with ASM 1%. The majority of subjects in each treatment group had moderate to severe pruritus at Baseline.

At Week 3, ASM cream was statistically significantly better than vehicle relative to the proportion of patients with no pruritus ($p=0.006$) and proportion of patients with absent or mild pruritus ($p=0.002$). At the primary time point, Week 6, ASM was statistically significantly better than vehicle relatiye to the proportion of patients with no pruritus ($p=0.001$) and proportion of patients with absent or mild pruritus ($p=0.019$).

Table 305-9. Frequency table of pruritus assessment (ITT population, LOCF)

	Treatment group	N	0 Absent n (%)	1 Mild n (%)	2 Moderate n (%)	3 Severe n (%)	p-value†	
							0	0 or 1
Baseline	ASM 1%	130	0	22 (16.9)	55 (42.3)	53 (40.8)	0.016	0.674
	Vehicle	68	3 (4.4)	7 (10.3)	25 (36.8)	33 (48.5)		
Week 3	ASM 1%	130	16 (12.3)	55 (42.3)	39 (30.0)	20 (15.4)	0.006	0.002
	Vehicle	68	1 (1.5)	21 (30.9)	18 (26.5)	28 (41.2)		
Week 6	ASM 1%	130	18 (13.8)	47 (36.2)	44 (33.8)	21 (16.2)	**0.001**	**0.019**
	Vehicle	68	0	22 (32.4)	18 (26.5)	28 (41.2)		

†p-value for pruritus score of 0 (absent) or 0, 1 (absent to mild) based on CMH general association test adjusted for center

Safety results

Table 305-10. Overall summary of treatment emergent AEs (Safety population of Study 305)

Parameter	ASM /ASM n (%)	Vehicle/ASM n (%)	P-value†
Double-blind phase	(N=130)	(N=68)	
At least 1 AE	91 (70.0)	47 (69.1)	1.000
At least 1 local AE	38 (29.2)	22 (32.4)	0.745
Any drug-related AE	23 (17.7)	17 (25.0)	0.264
Open-label phase	(N=112)	(N=48)	
At least 1 AE	83 (74.1)	30 (62.5)	0.184
At least 1 local AE	31 (27.7)	9 (18.8)	0.319
Any drug-related AE	6 (5.4)	4 (8.3)	0.489
Both phases	(N=130)	(N=68)	
At least 1 AE	109 (83.8)	53 (77.9)	0.335
At least 1 local AE	56 (43.1)	28 (41.2)	0.880
Any drug-related AE	28 (21.5)	19 (27.9)	0.379

† Chi-square test if all cells have expected count of 5 or greater. Otherwise, the Fisher's Exact test was used.

A summary of adverse events (AEs) is provided in Table 305-10. Denominator for open-label phase is the number of subjects who entered this phase. There was no statistically significant difference between the two treatment groups relative the proportion of patients with any AEs, local AEs, or drug related AEs in the DB phase, in the OL phase, or in both phases combined ($p \geq 0.184$).

A summary of common AEs in the DB phase is provided by body system in Table 305-11. There were no statistically significant differences between the treatment groups relative to common AEs in the DB phase ($p \geq 0.147$).

Table 305-11. Incidence rates of common (≥2% in any treatment group) treatment emergent AEs double-blind phase (Safety population of Study 305)

Organ class	ASM 1% (N=130) n (%)	Vehicle (N=68) n (%)	P-value†
At least one AE	91 (70.0)	47 (69.1)	1.000
Infections and infestations	56 43.1%)	27 (39.7%)	0.762
General disorders and administration site conditions	24 (18.5%)	19 (27.9%)	0.147
Respiratory, thoracic and mediastinal disorders	24 (18.5%)	8 (11.8%)	0.309
Nervous system disorders	18 (13.8%)	6 (8.8%)	0.365
Gastrointestinal disorders	12 (9.2%)	9 (13.2%)	0.467
Immune system disorders	7 (5.4%)	4 (5.9%)	1.000

† Chi-square test if all cells have expected count of 5 or greater. Otherwise, the Fisher's Exact test was used.

A summary of common local treatment emergent AEs in the DB phase is provided in Table 305-12. There were no statistically significant differences between the treatment groups relative to common local treatment emergent AEs in the DB phase ($p \geq 0.136$).

Table 305-12. Incidence rates of common (≥2% in any treatment group) local treatment emergent AEs double-blind phase (Safety population)

AE Class	ASM 1% (N=130)	Vehicle (N=68)	P-value†
At least one local AE	38 (29.2)	22 (32.4)	0.745
At least 1 common local AE	22 (16.9)	18 (26.5)	0.136
General disorders and administration site conditions	18 (13.8)	11 (16.2)	0.676
Infections and infestations	18 (13.8)	12 (17.6)	0.533

† Chi-square test if all cells have expected count of 5 or greater. Otherwise, the Fisher's Exact test was used.

Conclusions on Study 305

The primary efficacy variable is the proportion of subjects with the score 0 or 1 (clear or almost clear) in the Investigator's Global Assessment (IGA) at Week 6. The secondary efficacy variables in this review are the subject's assessment of pruritus and four individual signs of AD (erythema, induration/papulation, excoriations and lichenification) at Week 6. The all-category analysis of the IGA at Week 6 is also used to support the dichotomized analysis of the IGA.

The ITT population of Study 305 included 130 subjects in the ASM cream group and 68 subjects in the vehicle group. In the primary efficacy analysis, ASM cream was statistically significantly better than vehicle relative to the proportion of subjects with the score 0 or 1 in the IGA at Week 6 ($p=0.002$). All-category analysis of the IGA supported the primary efficacy results ($p<0.001$).

The secondary efficacy analysis showed that at Week 6, ASM cream was statistically significantly better than vehicle relative to proportion of patients with mild or absent signs of erythema and induration/papilation ($p\leq0.005$). At Week 6, ASM cream was statistically significantly better than vehicle relative to the proportion of patients with no pruritus ($p=0.001$) and proportion of patients with absent or mild pruritus ($p=0.019$). There was no statistically significant difference between ASM cream and vehicle relative to the proportion of patients with adverse events ($p>0.05$).

4. STUDY 307

Patient disposition

A summary of subject disposition is provided in Table 307-2. Statistically significantly more subjects in the vehicle group did not complete the study compared with the ASM group ($p=0.047$). The predominant reason for discontinuation in the vehicle group was unsatisfactory therapeutic effect ($p=0.01$). The numbers of subjects lost to follow up was comparable between the treatment groups (0.608).

Table 307-2. Subjects discontinued from study double-blind phase
(ITT population of Study 307)

	ASM 1% (N=137) n (%)	Vehicle (N=68) n (%)	Total (N=205) n (%)	P-Value
Completed	123 (89.8)	54 (79.4)	177 (86.3)	0.047
All discontinuations	14 (10.2)	14 (20.6)	28 (13.7)	0.047
Primary reasons for discontinuation				
Adverse event (s)	3 (2.2)	2 (2.9)	5 (2.4)	0.746
Lost to follow up	6 (4.4)	2 (2.9)	8 (3.9)	0.608
Unsatisfactory therapeutic effect	1 (0.7)	5 (7.4)	6 (2.9)	0.01

Patient populations

The efficacy populations are summarized in Table 307-3. There was no statistically significant difference between the treatment groups relative to the number of patients in the Per Protocol population (p=0.121).

Table 307-3. Number of subjects by population type

Analysis population	ASM 1% n (%)	Vehicle n (%)	Total n (%)	P-value
Randomized	137	68	205	
ITT	137 (100.0)	68 (100.0)	205 (100.0)	-
PP	103 (75.2)	44 (64.7)	147 (71.7)	0.121

Baseline demographic and background characteristics

A summary of demographic characteristics is provided in Table 307-4. There were significantly more males in the ASM group compared to the vehicle group (p=0.037). The majority of subjects (82.4%) were between 2 to 12 years of age.

Table 307-4. Baseline demographics (ITT population of Study 307)

Parameter	Units	ASM 1% (N=137)	Vehicle (N=68)	p-value
Age (years)	Mean ± SD	6.7 ± 4.05	6.9 ± 4.29	0.844†
	Range	1 -17	1 –17	
	Median	6.0	7.0	
Age group (years) n (%)				
	<2	1 (0.7)	2 (2.9)	
	2-<12	113 (82.5)	56 (82.4)	
	12-<18	16 (16.8)	10 (14.7)	
Gender (n, %)	Male	77 (56.2)	27 (39.7)	0.037‡
	Female	60 (43.8)	41 (60.3)	
Race (n, %)	Caucasian	70 (51.1)	32 (47.1)	0.737‡
	Black	38 (27.7)	23 (33.8)	
	Oriental	5 (3.6)	1 (1.5)	
	Other	24 (17.5)	12 (17.6)	

†Wilcoxon rank sum test
‡Fishers exact test

Disease characteristics at baseline are summarized in Table 307-5. There were no statistically significant differences between the treatment groups in the severity of AD (IGA) prior to treatment or in the %TBSA ($p \geq 0.389$). The majority of subjects had mild or moderate AD at Baseline, with the majority having moderate AD at Baseline (IGA score of 3). Subjects were to have an IGA score of 2 or 3 at Baseline, but some investigators enrolled subjects with severe disease at Baseline (IGA scores of 4).

Table 307-5. Disease characteristics at Baseline (ITT population of Study 307)

Parameter	Units	ASM 1% (N=137) n (%)	Vehicle (N=68) n (%)	p-value
IGA score, n (%)	2 (mild)	52 (38.0)	25 (36.8)	0.953†
	3 (moderate)	78 (56.9)	40 (58.8)	
	4 (severe)	7 (5.1)	3 (4.4)	
% TBSA involved	<= 5%	11 (8.0)	7 (10.3)	0.389‡
	> 5% - <= 15%	55 (40.1)	20 (29.4)	
	> 15% - <= 30%	40 (29.2)	16 (23.5)	
	> 30% - <= 60%	19 (13.9)	23 (33.8)	
	> 60%	12 (8.8)	2 (2.9)	

†Mantel-Haenszel Chi-square test

Efficacy results

Primary efficacy results

Investigators global assessment

Table 307-6. Treatment success† (IGA) double-blind phase (ITT population of Study 307)

Time point	ASM 1% (N=137)	Vehicle (N=68)	P-value‡
Baseline	0	0	
Week3	37 (27.0%)	8 (11.8%)	0.009
Week 6	44 (32.1%)	14 (20.6%)	0.076

†Defined as a score of 0 or 1 (clear or almost clear)
‡CMH test stratified by center

Treatment success, defined as an IGA score of 0 or 1 (clear or almost clear of disease), is summarized in Table 307-6. At Week 3, the ASM group was statistically significantly better than the vehicle group relative to the proportion of patients clear or almost clear of disease signs (p=0.009). At Week 6, the ASM group was only numerically better than the vehicle group relative to the proportion of patients clear or almost clear of disease signs (32% vs. 21%, p=0.076). Analysis of the Per Protocol population showed similar results: p= 0.019 at Week 3 and p=0.281 at Week 6.

Secondary efficacy results

The distribution of IGA scores is summarized in Table 307-7. There was a statistically significant difference between the treatment groups in the all-category analysis of IGA scores both at Week 3 and Week 6 ($p \leq 0.002$). Similar results were observed for the Per Protocol population ($p \leq 0.016$).

Table 307-7. Frequency distribution of IGA by visit, double-blind phase (ITT population, LOCF, Study 307)

Time point	N	IGA score						p-value‡
Group		0	1	2	3	4	5	
Week 3								<0.001
ASM 1%	137	6 (4.4)	31 (22.6)	53 (38.7)	44 (32.1)	3 (2.2)	0	
Vehicle	68	2 (2.9)	6 (8.8)	21 (30.9)	33 (48.5)	6 (8.8)	0	
Week 6								**0.002**
ASM 1%	137	15 (10.9)	29 (21.2)	55 (40.1)	35 (25.5)	3 (2.2)	0	
Vehicle	68	5 (7.4)	9 (13.2)	19 (27.9)	28 (41.2)	7 (10.3)	0	

IGA categories: 0=Clear, 1=Almost Clear, 2=Mild, 3=Moderate, 4=Severe, 5=Very Severe
‡CMH row mean score test, stratified by center

Key signs of AD

Table 307-8. Number (%) of subjects with mild or absent key signs of AD (double-blind phase, ITT population, LOCF, Study 307)

Visit	Treatment group	N	Erythema n (%)	Induration/ papulation n (%)	Excoriation n (%)	Lichenification n (%)
Baseline	ASM 1%	137	47 (34.3)	50 (36.5)	71 (51.8)	68 (49.6)
	Vehicle	68	28 (41.2)	33 (48.5)	43 (63.2)	35 (51.5)
	P-value		0.244	0.062	0.099	0.795
Week 3	ASM 1%	137	81 (59.1)	93 (67.9)	95 (69.3)	93 (67.9)
	Vehicle	68	29 (42.6)	31 (45.6)	44 (64.7)	36 (52.9)
	P-value		0.022	0.002	0.439	0.031
Week 6	ASM 1%	137	89 (65.0)	88 (64.2)	109 (79.6)	99 (72.3)
	Vehicle	68	34 (50.0)	36 (52.9)	44 (64.7)	40 (58.8)
	P-value		**0.037**	0.123	**0.013**	**0.030**

Reviewer's results on the proportion of subjects with sign scores of 1 or less (mild or absent symptoms) for each of the key signs of AD, erythema, induration/papulation, excoriation and lichenification are shown in Table 307-8.

At Baseline, there were no statistically significant differences between the treatment groups in the proportion of subjects with mild or absent signs of four key signs of AD ($p \geq 0.062$). At Week 3, there was a statistically significantly greater proportion of subjects with mild or absent symptoms of erythema, induration/papilation, and lichenification ($p \leq 0.031$) after treatment with ASM cream, compared to vehicle. At Week 6, there was a statistically significantly greater proportion of subjects with mild or absent symptoms of erythema, excoriation, and lichenification ($p \leq 0.037$) after treatment with ASM cream, compared to vehicle.

Pruritus

The frequency distribution of the pruritus assessment is provided in Table 307-9. At Baseline, there were no differences between the treatment groups in the numbers of subjects with no more than mild pruritus ($p=0.198$). At Week 6, statistically significantly more ASM subjects had no pruritus ($p=0.009$) or no more than mild pruritus ($p<0.001$).

Table 307-9. Frequency table of pruritus assessment (ITT population, LOCF, Study 307)

	Treatment group	N	0 Absent n (%)	1 Mild n (%)	2 Moderate n (%)	3 Severe n (%)	p-value†	
							0	0, 1
Baseline	ASM 1%	137	0	18 (13.1)	62 (45.3)	57 (41.6)	N/A	0.198
	Vehicle	68	0	5 (7.4)	33 (48.5)	30 (44.1)		
Week 3	ASM 1%	137	15 (10.9)	67 (48.9)	36 (26.3)	19 (13.9)	0.124	<0.001
	Vehicle	68	3 (4.4)	15 (22.1)	32 (47.1)	18 (26.5)		
Week 6	ASM 1%	137	24 (17.5)	62 (45.3)	32 (23.4)	19 (13.9)	**0.009**	**<0.001**
	Vehicle	68	3 (4.4)	21 (30.9)	27 (39.7)	17 (25.0)		

†p-value for pruritus score of 0 (absent) or 0, 1 (absent to mild) based on CMH general association test adjusted for center.

Safety results

Overall experience of AEs

An overall summary of AEs is summarized in Table 307-10. In the double-blind phase, there were no significant differences between the treatment groups relative to the proportion of patients with at least one AE or at least one local AE ($p \geq 0.086$). The ASM cream group had significantly smaller proportion of patients with any drug-related AEs ($p=0.033$).

Table 307-10. Overall summary of treatment-emergent AEs (Safety population of Study 307)

Parameter	ASM 1%/ASM n (%)	Vehicle/ASM n (%)	P-value
Double-blind phase	**(N=137)**	**(N=68)**	
At least 1 AE	91 (66.4)	50 (73.5)	0.297
At least 1 local AE	38 (27.7)	27 (39.7)	0.086
Any drug-related AE	24 (17.5)	21 (30.9)	0.033
Open-label phase	**(N=121)**	**(N=54)**	
At least 1 AE	86 (71.1)	41 (75.9)	0.0503
At least 1 local AE	36 (29.8)	15 (27.8)	0.790
Any drug-related AE	9 (7.4)	7 (13.0)	0.254
Both phases	**(N=137)**	**(N=68)**	
At least 1 AE	114 (83.2)	54 (79.4)	0.509
At least 1 local AE	58 (42.3)	31 (45.6)	0.658
Any drug-related AE	26 (19.0)	23 (33.8)	0.021

Denominator for open-label phase is the number of subjects who entered this phase
Drug-related events are those considered as "suspected' by the investigator to be related to study medication.

Double-blind phase

Table 307-11. Incidence rates of common (≥2% in any treatment group) local treatment-emergent AEs (double-blind phase, Safety population of Study 307)

Organ Class Preferred term	ASM 1% (N=137) n (%)	Vehicle (N=68) n (%)	P-value[†]
At least 1 local AE	38 (27.7)	27 (39.7)	0.086
At least 1 common local AE	29 (21.2%)	22 (32.4%)	0.085
General disorders and administration site conditions	24 (17.5%)	18 (26.5%)	0.141
Infections & infestations	11 (8.0%)	10 (14.7%)	0.147
Skin & subcutaneous tissue disorders	7 (5.1%)	4 (5.9%)	0.818

[†] Chi-square test if all cells have expected count of 5 or greater. Otherwise, the Fisher's Exact test was used.

The incidence rates of common (≥2%) local AEs is summarized in Table 307-11. There were no significant differences between the treatment groups relative to local treatment-emergent AEs (p≥ 0.085).

A summary of common AEs is provided by body system in Table 307-12. There were no statistically significant differences between the treatment groups (p≥ 0.063) relative to common treatment-emergent AEs.

Table 307-12. Incidence rates of common (≥2% in any treatment group) treatment-emergent AEs double-blind phase (Safety population of Study 307)

Organ Class	ASM 1% (N=137) n (%)	Vehicle (N=68) n (%)	P-value†
At least 1 AE	91 (66.4)	50 (73.5)	0.297
At least 1 common AE total	80 (58.4)	44 (64.7)	0.384
Infections and infestations	46 (33.6)	32 (47.1)	0.063
General disorders and administration site conditions	33 (24.1)	22 (32.4)	0.213
Gastrointestinal disorders	23 (16.8)	13 (19.1)	0.681
Respiratory and thoracic disorders	22 (16.1)	12 (17.6)	0.774
Nervous system disorders	21 (15.3)	6 (8.8)	0.182
Immune system disorders	10 (7.3)	3 (4.4)	0.411
Skin & subcutaneous tissue disorders	8 (5.8)	6 (8.8)	0.557
Injury & poisoning	7 (5.1)	3 (4.4)	1.000

† Chi-square test if all cells have expected count of 5 or greater. Otherwise, the Fisher's Exact test was used.

Conclusions on Study 307

The ITT population of Study 307 included 137 subjects in the ASM cream group and 68 subjects in the vehicle group. Primary efficacy analysis showed that ASM cream was only numerically better than vehicle at Week 6 (32% vs. 21%, p=0.076). Analysis of the Per Protocol population at Week 6 showed similar result (p=0.281). All-category analysis of the IGA at Week 6 showed that ASM cream was statistically significantly better than vehicle (p=0.002). Similar results were observed for the Per Protocol population (p=0.016).

Secondary efficacy analysis showed that at Week 6, ASM cream was statistically significantly better than vehicle relative to proportion of patients with mild or absent signs of erythema, excoriation, and lichenification (p≤0.037). At Week 6, statistically significantly more ASM subjects had no pruritus (p=0.009) or no more than mild pruritus (p<0.001).

There was no statistically significant difference between ASM cream and vehicle relative to the proportion of patients with adverse events ($p>0.05$).

5. STUDY 316

The design of Study 316 was similar to that of Study B305 with the two major differences:

- An infant Study 316 was a foreign study with 20 centers enrolling patients 3-23 months old. The age categories were defined as follows: 3 months - <1 year, 1 - <2 years.
- According to Amendment 1, dated June 23, 2000, an interim analysis was to be performed to evaluate efficacy data of the subjects who were randomized by May 31, 2000 and completed by July 12, 2000. In the Amendment, the significance level for the interim analysis was set at 0.005 (two-sided) with the nominal significance level for the final analyses at 0.047 (two-sided).

Reviewer's Comment:

A planned interim analysis was performed for subjects who were randomized by May 31, 2000, and completed the double-blind phase by July 12, 2000. The nominal significance level was set for the interim analysis at 0.005 and for the final analysis at 0.047 (two-sided). As the level of 0.047 for the final analysis is more conservative than the level of 0.048 in the O'Brian/ Fleming method, this reviewer accepts the Amendment's specifications.

Patient disposition in Study 316

Table 316-2 presents the reasons for discontinuation in the intent-to-treat population. The rate of discontinuation due to unsatisfactory therapeutic effect was statistically significantly higher in subjects treated with vehicle than in those treated with ASM cream ($p=0.001$).

Table 316-2. Subjects discontinued from study DB phase (ITT population of Study 316)

	ASM 1% (N=123)		Vehicle (N=63)		P-value
	n	(%)	n	(%)	
Completed	109	(88.6)	33	(52.4)	0.001
All discontinuations	14	(11.4)	30	(47.6)	
Primary Reasons for Discontinuation					
Unsatisfactory Therapeutic Effect	8	(6.5)	26	(41.3)	0.001
Protocol violation	2	(1.6)	1	(1.6)	0.984
Lost to follow up	2	(1.6)	1	(1.6)	0.984

Analysis Populations

Two populations were defined for analysis and are summarized in Table 316-3. The ITT population was comprised all subjects who were randomized. The PP population was defined for use in the double-blind phase only. The percentage of subjects excluded from the PP population was numerically higher in the vehicle group (p=0.080).

Table 316-3. Number of subjects by analysis population and treatment group

Analysis population	ASM 1%	Vehicle	P-value
Randomized – n	123	63	
Intent-to-treat – n (%)	123 (100.0)	63 (100.0)	-
Per protocol – n (%)	95 (77.2)	41 (65.1)	0.080

Baseline demographic and background characteristics

Baseline demographics and background characteristics are provided in Table 316-4. There were no statistically significant differences between the treatment groups for age, gender, or race ($p \geq 0.153$).

Table 316-4. Baseline demographics (ITT population of Study 316)

		ASM 1% (N=59)	Vehicle (N=24)	P-value
Age (months)	mean ± SD	12.6 ± 6.25	12.7 ± 6.29	0.891
	range	3 - 24	3 - 23	
Gender (n, %)	male	68 (55.3)	34 (54.0)	0.878
	female	55 (44.7)	29 (46.0)	
Race (n, %)	Caucasian	65 (52.8)	44 (69.8)	0.153
	Black	264 (13.0)	4 (6.3)	
	Oriental	3 (2.4)	1 (1.6)	
	Other	39 (31.7)	14 (22.2)	

Baseline disease characteristics by treatment group are shown in Table 316.5. There were no statistically significant differences between treatment groups ($p \geq 0.135$) relative to the % TBSA or IGA at baseline.

Table 316-5. Baseline disease characteristics (ITT population of Study 316)

		ASM 1% (N=123)	Vehicle (N=63)	P-value
TBSA involved (%)	mean ± SD	27.4 ± 20.79	23.0 ± 18.63	0.135
	range			
IGA score (n, %)	2 (mild)	40 (32.5)	21 (33.3)	0.911
	3 (moderate)	83 (67.5)	42 (66.7)	

Efficacy results in Study 316

Primary efficacy results

Treatment success, defined as an IGA score of 0 (clear) or 1 (almost clear), is summarized in Table 316-6 for the ITT population of Study 316. Primary efficacy analysis showed that Week 6, there were statistically significantly more subjects clear or almost clear as assessed by IGA in the ASM cream group as compared with subjects treated with vehicle ($p<0.001$). Analysis in the Per Protocol population supported the results in the ITT population ($p=0.002$). All-category analysis of the IGA scores supported the analysis of the dichotomized IGA ($p<0.001$).

Table 316-6. Treatment success: Investigator's Global Assessment
(ITT population of Study 316)

Time point	ASM 1% (N=123)		Vehicle (N=63)		p-value[†]
	n	(%)	N	(%)	
Week 3	54	(43.9)	11	(17.5)	**<0.001**
Week 6	67	(54.5)	15	(23.8)	**<0.001**

[†] CMH test, stratified by center;

Secondary efficacy analysis in Study 316

Proportion of subjects with scores of 1 or less (mild or absent signs) for each of the key signs of AD, erythema, infiltration/papulation, excoriation and lichenification is provided in Table 316-8.

At Baseline, the proportion of subjects with mild or absent signs of atopic dermatitis in the two treatment groups was balanced ($p \geq 0.87$). At Week 6, there was a statistically significantly greater proportion of subjects with mild or absent signs for all 4 key signs of atopic dermatitis after treatment with ASM compared with vehicle ($p \leq 0.008$).

Table 316-8. Subjects with mild or absent key signs of atopic dermatitis
(ITT population of Study 316)

Visit	Treatment group		Erythema		Infiltration/ papulation		Excoriation		Lichenification	
		N	n	(%)	n	(%)	N	(%)	n	(%)
Week 3	ASM	123	82	(66.7)	86	(69.9)	102	(82.9)	97	(78.9)
	Vehicle	63	26	(41.3)	26	(41.3)	40	(63.5)	41	(65.1)
	P-value		<0.001		<0.001		<0.001		0.009	
Week 6	ASM	123	88	(71.5)	86	(69.9)	101	(82.1)	99	(80.5)
	Vehicle	63	19	(30.2)	24	(38.1)	37	(58.7)	41	(65.1)
	P-value		**<0.001**		**<0.001**		**0.001**		**0.008**	

Pruritus

The frequency distribution of the pruritus assessment is provided in Table 316-9. At Baseline, there was no differences between the treatment groups in the proportion of subjects with absent pruritus ($p=0.596$). At Week 6, ASM cream was statistically significantly better than vehicle relative to the proportion of subjects with absent pruritus or no more than mild pruritus ($p<0.001$).

Table 316-9. Frequency table of pruritus assessment (ITT population of Study 316)

		N	0 Absent	1 Mild	2 Moderate	3 Severe	p-value† Score =	
			n (%)	n (%)	n (%)	n (%)	0	0, 1
Week 3	ASM	123	44 (35.8)	51 (41.5)	17 (13.8)	11 (8 .9)	<0.001	<0.001
	Vehicle	63	6 (9 .5)	17 (27.0)	18 (26.6)	22 (34.9)		
Week 6	ASM	123	55 (44.7)	34 (27.6)	20 (16.3)	14 (11.4)	**<0.001**	**<0.001**
	Vehicle	63	6 (9.5)	15 (23.8)	16 (25.4)	26 (41.3)		

†p-value for pruritus score of 0 (absent) or 0, 1 (absent to mild).

Safety results in Study 316

Overall experience of adverse events (AEs)

An overall summary of adverse events is provided in Table 316-10. The ASM group had a marginally statistically significantly greater proportion of subjects with at least one adverse event in the double-blind phase (79% vs. 65%, p=0.052).

Table 316-10. Overall summary of treatment-emergent adverse events in the double-blind phase (Safety population of Study 316)

	ASM 1% (N=123)		Vehicle (N=63)		P-value
	n	(%)	n	(%)	
At least 1 AE	97	(78.9)	41	(65.1)	0.052
At least 1 local AE	27	(22.0)	18	(26.6)	0.367
Any drug related AE†	7	(5.7)	8	(12.7)	0.152

A summary of common AEs is provided by body system in Table 316-11. Compared to vehicle, statistically significantly more ASM subjects had pyrexia (p=0.003) and diarrhea (p=0.017).

Table 316-11. Incidence rates of common (≥2% in any treatment group) treatment-emergent AEs in the double-blind phase (Safety population of Study 316)

Organ Class Preferred term	ASM 1% (N=123) n (%)	Vehicle (N=63) n (%)	P-value[†]
At least 1 AE	97 (78.9)	41 (65.1)	**0.052**
At least 1 common AE	84 (68.3)	39 (61.9)	0.386
Infections and infestations			
Upper respiratory tract infections (NOS)	29 (23.6)	9 (14.3)	0.128
Nasopharyngitis	18 (14.6)	5 (7.9)	0.175
Gastroenteritis	9 (7.3)	2 (3.2)	0.338
Otitis media NOS	5 (4.1)	0	0.169
Influenza	7 (5.7)	2 (3.2)	0.720
General disorders and administration site conditions			
Pyrexia	39 (31.7)	8 (12.7)	**0.003**
Gastrointestinal disorders			
Teething	10 (8.1)	3 (4.8)	0.548
Diarrhea NOS	10 (8.1)	0	**0.017**
Respiratory and thoracic disorders			
Asthma	7 (5.7)	2 (3.2)	0.720
Psychiatric disorders			
Restlessness	10 (8.1)	3 (4.8)	0.548
Skin & subcutaneous tissue disorders			
Dermatitis contact	4 (3.3)	1 (1.6)	0.664
Injury & poisoning			
Abrasion NOS	4 (3.3)	0	0.302

[†] Chi-square test if all cells have expected count of 5 or greater. Otherwise, the Fisher's Exact test was used.

Conclusions on Study 316

A planned interim analysis was performed for subjects who were randomized by May 31, 2000 and completed the double-blind phase by July 12, 2000. In the protocol, the nominal significance level for the final analyses was set at 0.047 (two-sided). The ITT population of Study 316 included 123 subjects in the ASM cream group and 63 subjects in the vehicle group. In the primary efficacy analysis, ASM cream was statistically significantly better than vehicle relative to the proportion of subjects with the score 0 or 1 in the IGA at Week 6 ($p<0.001$).

The secondary efficacy analysis of Study 316 showed that at Week 6, ASM cream was statistically significantly better than vehicle relative to the proportion of subjects with absent or mild key signs

of AD ($p \leq 0.008$), no pruritus or absent or mild pruritus ($p<0.001$). Compared to vehicle, the ASM group had a marginally statistically significantly greater proportion of subjects with at least one adverse event ($p=0.052$). Statistically significantly more ASM subjects had pyrexia ($p=0.003$) and diarrhea ($p=0.017$).

6. STUDY B202

Study B202 was a foreign, 3-week, Phase 2, dose ranging study. The primary objective of the study was to determine the safety and efficacy of four concentrations (0.05%, 0.2%, 0.6% and 1.0%) of ASM cream in moderate atopic dermatitis in comparison to vehicle and 0.1% betamethasone-17-valerate cream (BMV). This reviewer's objective was to examine the dose selection development and the evidence of the dose response.

Study design

This was a randomized, multicenter, double-blind, parallel-group, vehicle- and active-controlled trial of ASM cream in the treatment of moderate atopic dermatitis for 3 weeks. A total of 260 subjects were randomly allocated to receive either one of the four concentrations (0.05%, 0.2%, 0.6%, 1.0%) of ASM cream, corresponding vehicle, or BMV. The study medication was applied twice daily to the areas treated by the subjects for up to 3 weeks. New lesions occurring during treatment were to be treated along with the lesions already receiving medication. If the atopic dermatitis cleared (Hanifin = 0 and overall evaluation 100% clear) before the full 21 days of treatment, the patient was instructed to stop treatment.

Subject population

The target population for this study was adult subjects, of either gender, with moderate atopic dermatitis effecting 5-30% of the total body surface area and a target area score of at least 8 in the Hanifin scale.

Efficacy assessments

Sponsor's primary efficacy variable

The primary efficacy variable in the protocol was the percentage change from baseline to Week 3 in the Hanifin target score of the most severely affected lesion at baseline. The Hanifin score was defined as the sum of the six ordinal scales of erythema, oozing/crust, papulation, lichenification, excoriations and pruritus. Each was on a 4-point scale, from 0 (absence) to 3 (severe stage). Scoring of half steps was allowed. The Hanifin Score was determined by the investigator immediately before the first application of the trial medication (Day 1) and at each weekly visit. The visit schedule is shown in Table 202-1.

Table 202-1. Visit Schedule

	Pre-treatment	Week 1	Week 2	Week 3	Post-treatment
Hanifin score: Target Site	X	X	X	X	X
EASI	X	X	X	X	X
Overall evaluation score by Investigator					X

Primary analysis was based on the ITT population. In cases of early discontinuation the last recorded Hanifin score (endpoint) was to be carried forward. Intent to Treat (ITT) Population was defined as all subjects randomized, who received at least one dose of the trial medication and provided baseline and at least one documented assessment of the primary efficacy variable (Hanifin score) post-baseline. In this trial, all subjects who were randomized and received at least one dose of study medication, subsequently provided at least one post-baseline safety and primary efficacy evaluations. Therefore, the safety and ITT populations were equivalent and consistent with the Division's definition of ITT as 'all subjects randomized who were dispensed study medication'.

Sponsor's secondary efficacy variables

Pruritus

Along with the evaluation of pruritus as a component of the Hanifin score at a target site, an overall pruritus evaluation was performed. Categories of 'absent' and 'mild' were collapsed in the analysis. Summary statistics, in terms of a frequency distribution of the pruritus assessment were provided, including the proportion of subjects with a score of 0 or 1 (mild pruritus or better) at each visit.

Overall Evaluation

An overall improvement score was recorded at Week 3 (or the day of complete clearance as observed by the investigator). This was a 7-point scale varying from 0=complete clearing of the areas (i.e. normal skin), to a worsening of the disease. When this evaluation was missing, subjects were assumed as worst cases - a score of 6.

Reviewer's Comments:

- In Study 202, the sponsor's primary efficacy variable, Hanifin score, is different from the primary efficacy variable in the Phase 3 studies, success rate in the investigator's global evaluation (IGA). In order to compare results of Study 202 with the results of the Phase 3 Studies 305, 307, and 316, this reviewer used the investigator's overall improvement score at Week 3 as a primary efficacy variable.
- Investigator's overall evaluation of improvement in Study 202 was evaluated only at Week 3 and was not evaluated at baseline. The scale of the overall evaluation is as follows:

0 – 100% (completely clear of signs and symptoms of dermatitis);
1 - 90%-99% (almost clear with few signs and symptoms of dermatitis remaining at treated areas;
2 - 75%-89% (markedly clear with some signs and symptoms of dermatitis remaining at treated areas;
3 - 50%-74% (moderately clear with notable signs and symptoms of dermatitis remaining at treated areas;
4 - 25%-49% (minimally clear with signs and symptoms of dermatitis remaining at treated areas;
5 - 0%-24% (unclear with very marked signs and symptoms of dermatitis remaining at treated areas;
6 – Extreme involvement with gross signs and symptoms of dermatitis.

- The protocol called the investigator's overall evaluation "evaluation of improvement over baseline", but the scale was actually a static evaluation of dermatitis status at the endpoint. For this reason, in agreement with the medical reviewer, the primary efficacy variable in the statistical review of Study 202 is success rate in investigator' overall evaluation at Week 3. The success rate is defined as the percent of patients with the score of 0 or 1 (clear or almost clear) in the investigator's overall evaluation. This primary efficacy variable is consistent with the primary efficacy variable in the Phase 3 Studies B305, B307, and 316.
- As a secondary efficacy variable in Study 202, this reviewer used percent of patients with pruritus score of 0 or 1(absent or mild) at Week 3.
- In the dose–ranging Study 202, to avoid a p-value adjustment for multiple comparisons, the sponsor used a closed test procedure. This procedure is based on the assumption that there is a monotonic non-decreasing dose response to the study drug. As the sponsor did not provide evidence to support this assumption, this reviewer instead used the Bonferroni adjustment for multiple comparisons. Using Bonferroni adjustment for multiple comparisons, the significance level for the 5 multiple comparisons versus vehicle should be 0.05/5 = 0.01.

Results in the reviewer's efficacy analysis

Reviewer's primary efficacy analysis

The primary efficacy variable in the reviewer' analysis was success rate in the investigator's overall evaluation. Table 202-2 shows the percentage of subjects in each treatment group classified as clear or almost clear in the overall evaluation by investigator at Week 3 (grades 0 or 1). There was a dose response trend indicating increasing proportion of patients clear or almost clear as dosage increased, but ASM 1% was only numerically better than vehicle (11% vs. 0%, p=0.056).

Table 202-2. Percentage of Subjects with Normal or Almost Clear Overall Evaluation Score by Investigator at Week 3 (ITT population of Study 202)

Treatment Group	Number (%) of patients	P-value versus vehicle *
Vehicle (n=43)	**0 (0%)**	-
ASM 0.05% (n=42)	0 (0%)	-
ASM 0.2 % (n=46)	1 (2%)	1.00
ASM 0.6 % (n=42)	2 (5%)	0.241
ASM 1.0 % (n=45)	**5 (11%)**	**0.056**
BMV (n=42)	21 (50%)	<0.001

* The significance level for 5 multiple comparisons versus vehicle is 0.05/5 = 0.01.

Reviewer's secondary efficacy analysis

Pruritus

Table 202-3 shows the number and percentage of patients with absent and mild (0 and 1) pruritus score at baseline and at Week 3. ASM 1% cream was statistically significantly better than vehicle (p=0.007) at endpoint.

Table 202-3. Subjects with Absent and Mild (0 and 1) Pruritus Score at Baseline and Week 3 (ITT population of Study 202).

Treatment group	Baseline	Week 3	P-value versus vehicle* At Week 3
Vehicle (n = 43)	**2 (4.7%)**	**8 (18.6 %)**	-
ASM 0.05% (n = 42)	2 (4.8%)	10 (23.8%)	0.604
ASM 0.2 % (n = 46)	4 (8.7%)	17 (37.0%)	0.063
ASM 0.6 % (n = 42)	5 (11.9%)	22 (52.4%)	0.002
ASM 1.0 % (n = 45)	**3 (6.7%)**	**21 (46.7%)**	**0.007**
BMV (n = 42)	5 (11.9%)	34 (81.0%)	<0.001

* The significance level for 5 multiple comparisons versus vehicle is 0.05/5 = 0.01

Sponsor's Primary Efficacy Results

The sponsor' primary s efficacy analysis relative to the mean reduction in the Hanifin score at endpoint is shown in Table 202-4. ASM 1% cream was statistically significantly better than

vehicle ($p<0.001$) at endpoint.

Table 202-4. Hanifin Scores at Endpoint (ITT population of Study 202)

Treatment comparison	Estimate of mean treatment difference	P-Value
ASM 1.0% against vehicle	-2.60	<0.001
ASM 0.6% against vehicle	-2.96	<0.001
ASM 0.2% against vehicle	-2.04	0.0029
ASM 0.05% against vehicle	-0.41	not significant

Reviewer's Conclusions on Study B202

Study B202 was a Phase 2, six-arm, 3-week, dose ranging study in 260 adult patients with moderate AD. The vehicle and ASM 1% arms had 43 and 45 patients, respectively. The sponsor's primary efficacy variable was the mean Hanifin score. The reviewer's primary efficacy variable was the proportion of patients with grade 0 or 1 (clear or almost clear) in the investigator's overall evaluation, which was consistent with the primary efficacy variable in the Phase 3 Studies B305, B307, and 316. In this review, to adjust for multiple comparisons of five active arms against vehicle, the significance level in this study is 0.01. The primary efficacy analysis showed a dose response trend. ASM 1% was numerically better than vehicle relative to the proportion of patients clear or almost clear (11% vs. 0%, p=0.056). Secondary efficacy analysis showed that at ASM 1% cream was statistically significantly better than vehicle relative to the proportion of patients with absent or mild pruritus (p=0.007).

7. INTEGRATED EFFICACY SUBGROUP ANALYSIS

Integrated subgroup analysis in this review is based on the pooled data from the three pivotal short-term pediatric studies, B305, B307, and 0316. A total of 589 subjects were enrolled in these three studies with 390 subjects treated with ASM cream and 199 subjects treated with vehicle.

Age subgroup analysis

Table ISS-1. IGA treatment success at endpoint by subject age in pediatric subjects of the pooled ITT populations of Studies B305, B307, and 0316.

Age Category	n/N (%) Treatment Success*		P-value†
	ASM Cream	Vehicle	
< 2 years	69/126 (54.8%)	17/68 (25.0%)	**<0.001**
2 - <12 years	74/221 (33.5%)	20/110 (18.2%)	**<0.003**
12 - <18 years	17/43 (39.5%)	3/21 (14.3%)	**0.033**

* Treatment success defined as Investigator Global of 0=Clear or 1 = Almost Clear.

† Chi-square test.

Table ISS-1 presents the IGA treatment success by subject age in pediatric subjects of the pooled ITT populations of Studies B305, B307, and 316. ASM cream was statistically significantly better than vehicle in any of the three age subgroups ($p \leq 0.033$).

Gender subgroup analysis

Table ISS-2 presents the IGA treatment success by subject gender in pediatric subjects of pooled ITT populations of Studies B305, B307, and 316. ASM cream was statistically significantly better than vehicle in both males and females ($p \leq 0.002$).

Table ISS-2. IGA treatment success by subject gender in pediatric subjects of the pooled ITT populations of Studies B305, B307, and 316.

Gender	n/N (%) Treatment Success*		P-value†
	ASM Cream	Vehicle	
Male	81/209 (38.8%)	20/95 (21.1%)	**0.002**
Female	79/181 (43.6%)	20/104 (19.2%)	**<0.001**

* Treatment success defined as Investigator Global of 0=Clear or 1 = Almost Clear.

† Chi-square test.

Race subgroup analysis

Table ISS-3 presents the IGA treatment success by subject race in pediatric subjects of pooled ITT populations of Studies B305, B307, and 316. ASM cream was statistically significantly better than vehicle in both Caucasian and non-Caucasian subgroups ($p<0.001$).

Table ISS-3. IGA treatment success by subject race in pediatric subjects of the pooled ITT populations of Studies B305, B307, and 316.

Race	n/N (%) Treatment Success*		P-value†
	ASM Cream	Vehicle	
Caucasian	95/211 (45.0%)	26/110 (23.6%)	**<0.001**
Non-Caucasian	65/179 (36.3%)	14/89 (15.7%)	**<0.001**

* Treatment success defined as Investigator Global of 0=Clear or 1 = Almost Clear.

† Chi-square test.

Influence of area affected

Table ISS-4 presents the IGA treatment success by the percent of total body surface area affected (TBSA) in pediatric subjects of pooled ITT populations of Studies B305, B307, and 316. ASM cream was statistically significantly better than vehicle in all subgroups ($p \leq 0.032$) except for the subgroup of patients with >60% TBSA involved.

Table ISS-4. IGA treatment success by percent of area affected in pediatric subjects of the pooled ITT populations of Studies B305, B307, and 316.

Percent of area affected (TBSA)	n/N (%) Treatment Success*		P-value†
	ASM Cream	Vehicle	
≤5%	19/39 (48.7%)	2/17 (11.8%)	**0.005**
>5% - ≤15%	58/113 (51.3%)	23/66 (34.8%)	**0.032**
>15% - ≤30%	50/112 (44.6%)	8/50 (16.0%)	**0.001**
>30% - ≤60%	29/83 (34.9%)	6/52 (11.5%)	**0.002**
>60%	4/43 (9.3%)	1/14 (7.1%)	1.000

* Treatment success defined as Investigator Global of 0=Clear or 1 = Almost Clear.

† Chi-square test if all cells have expected count 5 or greater. Otherwise, the Fisher's Exact test was used.

Influence of baseline disease severity

Table ISS-5 presents the IGA treatment success by baseline IGA in pediatric subjects of pooled ITT populations of Studies B305, B307, and 316. ASM cream was statistically significantly better than vehicle in the subgroups with the baseline IGA equal 2 and 3 ($p \leq 0.012$) but not in the subgroup of patients with IGA equal 4 or 5 ($p=0.287$).

Table ISS-5. IGA treatment success by baseline IGA in pediatric subjects of the pooled ITT populations of Studies B305, B307, and 316.

Baseline IGA	n/N (%) Treatment Success*		P-value†
	ASM Cream	Vehicle	
2	70/120 (58.3%)	25/64 (39.1%)	**0.012**
3	87/244 (35.7%)	15/120 (12.5%)	**0.001**
4 and 5	3 /26 (11.5%)	0/15 (0%)	0.287

* Treatment success defined as Investigator Global of 0=Clear or 1 = Almost Clear.

† Chi-square test if all cells have expected count 5 or greater. Otherwise, the Fisher's Exact test was used.

8. REVIEWER'S CONCLUSIONS

The sponsor submitted reports of three pivotal trials (Studies B305, B307, and 316) to demonstrate that Elidel (1% SDZ ASM 981 cream) is safe and effective in the treatment of atopic dermatitis — The US Studies B305 and B307 enrolled pediatric subjects age 2 to 17 years old and had identical designs. A foreign Study 0316 enrolled subjects 3 to 23 months old and had a similar design. In addition, the sponsor submitted a report of a Phase 2 dose-ranging study B202. The primary efficacy variable in the three pivotal studies, used both by the sponsor and by the reviewer, was the proportion of subjects with the score 0 or 1 (clear or almost clear) in the Investigator's Global Assessment (IGA) at Week 6. The secondary efficacy variables in this review are four individual signs of AD (erythema, induration/papulation, excoriations and lichenification)

and the subject's assessment of pruritus at Week 6. The all-category analysis of the IGA at Week 6 is also used to support the dichotomized analysis of the IGA.

Study B305

The ITT population of Study B305 included 130 subjects in the ASM cream group and 68 subjects in the vehicle group. In the primary efficacy analysis, ASM cream was statistically significantly better than vehicle at Week 6 relative to the proportion of subjects with the score 0 or 1 in the IGA ($p=0.002$). All-category analysis of the IGA supported the primary efficacy results ($p<0.001$).

The secondary efficacy analysis showed that at Week 6, ASM cream was statistically significantly better than vehicle relative to proportion of patients with mild or absent signs of erythema and induration/papilation ($p\leq0.005$). At Week 6, ASM cream was statistically significantly better than vehicle relative to the proportion of patients with no pruritus ($p=0.001$) and proportion of patients with absent or mild pruritus ($p=0.019$). There was no statistically significant difference between ASM cream and vehicle relative to the proportion of patients with adverse events ($p>0.05$).

Study B307

The ITT population of Study B307 included 137 subjects in the ASM cream group and 68 subjects in the vehicle group. Primary efficacy analysis showed that ASM cream was only numerically better than vehicle at Week 6 (32% vs. 21%, $p=0.076$). Analysis of the Per Protocol population at Week 6 showed similar results ($p=0.281$). All-category analysis of the IGA at Week 6 showed that ASM cream was statistically significantly better than vehicle ($p=0.002$). Similar results were observed for the all-category analysis in the Per Protocol population ($p=0.016$).

Secondary efficacy analysis showed that at Week 6, ASM cream was statistically significantly better than vehicle relative to the proportion of patients with mild or absent signs of erythema, excoriation, and lichenification ($p\leq0.037$). At Week 6, statistically significantly more ASM subjects had no pruritus ($p=0.009$) or no more than mild pruritus ($p<0.001$). There was no statistically significant difference between ASM cream and vehicle relative to the proportion of patients with adverse events ($p>0.05$).

Study 0316

The ITT population of Study 0316 included 123 subjects in the ASM cream group and 63 subjects in the vehicle group. As a planned interim analysis was performed in this study, for the final analysis, the nominal significance level was set at 0.047 (two-sided). In the primary efficacy analysis, ASM cream was statistically significantly better than vehicle relative to the proportion of subjects with the score 0 or 1 in the IGA at Week 6 ($p<0.001$).

The secondary efficacy analysis of Study 0316 showed that at Week 6, ASM cream was statistically significantly better than vehicle relative to the proportion of subjects with absent or mild key signs of AD ($p \leq 0.008$), no pruritus or absent or mild pruritus ($p<0.001$). Compared to vehicle, the ASM group had a marginally statistically significantly greater proportion of subjects with at least one adverse event (79% vs. 65%, p=0.052). Statistically significantly more ASM subjects had pyrexia (32% vs. 13%, p=0.003) and diarrhea (8% vs. 0%, p=0.017).

Study B202

Study B202 was a Phase 2, six-arm, 3-week, dose ranging study in 260 adult patients with moderate atopic dermatitis. The vehicle and ASM 1% arms had 43 and 45 patients, respectively. The reviewer's primary efficacy variable was the proportion of patients with grade 0 or 1 (clear or almost clear) in the investigator's overall evaluation at Week 3. In this review, to adjust for multiple comparisons of five active arms against vehicle, the significance level in this study is 0.01. The primary efficacy analysis showed a dose response trend. ASM 1% was numerically better than vehicle relative to the proportion of patients clear or almost clear (11% vs. 0%, p=0.056). Secondary efficacy analysis showed that ASM 1% cream was statistically significantly better than vehicle relative to the proportion of patients with absent or mild pruritus (p=0.007).

Integrated efficacy subgroup analysis

Integrated efficacy subgroup analysis in this review is based on the pooled data from the three pivotal pediatric studies, B305, B307, and 0316. A total of 589 subjects were enrolled in these three studies with 390 subjects treated with ASM cream and 199 subjects treated with vehicle.

ASM cream was statistically significantly better than vehicle ($p \leq 0.033$) in any of the three age subgroups (<2, 2-11, and 12-17 years), both in males and females ($p \leq 0.002$), and both in Caucasians and non-Caucasians ($p<0.001$). ASM cream was statistically significantly better than vehicle both in the subgroup with the baseline IGA score of mild and moderate ($p \leq 0.012$), but not in the subgroup with the baseline IGA score of severe or very severe (p=0.287). ASM cream was statistically significantly better than vehicle in each subgroup of total body surface area involved ($p \leq 0.032$) except for the subgroup of patients who had involvement of more than 60% (p=1.0).

Overall Conclusions

Overall, primary efficacy analyses of Studies B305 and 0316 showed that ASM cream was statistically significantly better than vehicle relative to the proportion of subjects with the score 0 or 1 in Investigator's Global Assessment at Week 6 ($p<0.002$). Primary efficacy analysis of Study B307 showed that ASM cream was numerically better than vehicle at Week 6 (32% vs. 21%, p=0.076).

Safety analyses of Studies B305 and B307 showed that there was no statistically significant difference between ASM cream and vehicle relative to the proportion of patients with adverse events (p>0.05). In the infant Study 0316, the ASM group was marginally statistically significantly worse than vehicle relative to the proportion of subjects with at least one adverse event (79% vs. 65%, p=0.052). In this study, the ASM group was statistically significantly worse than vehicle relative to the proportion of subjects with pyrexia (32% vs. 13%, p=0.003) and diarrhea (8% vs. 0%, p=0.017).

[ ]

Valeria Freidlin, Ph.D.
Mathematical Statistician, Biometrics III

Concur: Mohamed Alosh, Ph.D.
Team Leader, Biometrics III

cc: NDA 21-302
HFD-540
HFD-540/Mrs. Wright
HFD-540/Dr. Cook
HFD-540/Dr. Luke
HFD-540/Dr. Wilkin
HFD-725/Dr. Alosh
HFD-725/Dr. Freidlin
HFD-725/Dr. Huque
HFD-700/Dr. Anello
This review contains 37 pages.

APPEARS THIS WAY
ON ORIGINAL

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

/s/

Valeria Freidlin
10/19/01 09:55:51 AM
BIOMETRICS

Mohamed Alosh
10/22/01 06:16:59 PM
BIOMETRICS
Concur with review

APPEARS THIS WAY
ON ORIGINAL

Statistical Review and Evaluation

NDA #:	21-302
Applicant:	Novartis
Name of Drug:	Elidel (Pimecrolimus) Cream 1%
Documents Reviewed:	Submission volumes 82 to 86 and 98 to 106 of December 15, 2000 for rat and Mouse dermal carcinogenicity study. Data on CD supplied by the sponsor
Reviewing Pharmacologist:	Barbara Hill, Ph.D.
Reviewing Biostatistician:	M. Atiar Rahman, Ph.D.

1. Background

In this submission the sponsor included reports of two animal carcinogenicity studies, one in rats and one in mice. These studies were intended to assess the carcinogenic potential of Elidel (Pimecrolimus) Cream 1% on rats and mice. The route of administration was dermal (back shaved once a week), applied daily with appropriate dose levels. The lengths of both studies were designed to be 104 weeks. The results of this review have been discussed with Dr. Hill.

2. The Rat study

Design: Two separate experiments, one in males and one in females were conducted. In each of these two experiments there were three treated and two control groups. Three hundred —— :WIST (SPF) Wistar rats of each sex were equally (50 sex/group) randomly allocated to controls and treated groups. The dose levels for the treated group was 0.2% (Low), 0.6% (Medium), and 1.0% (High) body weight/day for three treated groups. Animals in one control group were treated with saline (0.9% NaCl) (Control-1), and those in the other control group were administered a placebo without the test article (Control- 2).

Mortality, clinical signs, nodules and masses, food consumption and body weight were recorded periodically during acclimatization and treatment. All protocol designated organs and tissues obtained from the animals of the two control groups and the high dose group, died during the study, killed moribund, or sacrificed at the end of the study, were histopathologically examined.

2.1 Sponsor's analyses

Survival analysis: Number of animals died during the study in different dose groups was presented in tabular form. Kaplan-Miere survival curves for death rates were presented. The sponsor concluded that there was no effect on survival of animals, which were treated with the test article when compared with animals treated with saline or placebo.

Tumor data analysis: Tumor data were analyzed using the method described in the paper of Peto et al. (Guidelines for sample sensitive significance test for carcinogenic effects in long-term animal experiments, *Long term and short term screening assays for carcinogens: A critical appraisal*, International agency for research against cancer monographs, Annex to supplement, World Health

Organization, Geneva, 311-426, 1980). The sponsor did not provide details of any mortality adjustment during this test. Adjustment for multiple tests were done using the method suggested by Haseman (Haseman, 1983. A re-examination of false-positive rates for carcinogenesis studies, *Fundamental and Applied Toxicology*, 3: 334-339), which recommends to use p=0.01 for common[1] tumor type, and p=0.05 for rare[2] tumor type, in order to keep the false-positive rate at the nominal level of approximately five percent.

Statistical analysis did not show statistically significant increased incidence of any tumor type in the high dose group when compared with either the saline or the placebo control.

2.2 Reviewer's analyses

This reviewer independently performed survival and tumor data analyses. For survival analysis the methods described in the papers of Cox (Regression models and life tables, *Journal of the Royal Statistical Society*, B, 34, 187-220, 1972), and of Gehan (A generalized Wilcoxon test for comparing arbitrarily singly censored samples, *Biometrika*, 52, 203-223,1965) were used. The tumor data analyses were performed using the methods described in the paper of Peto et al. (1980), and the age-adjusted Fisher exact test. Data used in this reviewer's analyses were provided by the sponsor on a CD.

Survival analysis: The intercurrent mortality data are given in Tables 1A and 1B for males and females, respectively. The Kaplan-Meyer curves for death rate are given in Figures 1A and 1B for males and females, respectively. The homogeneity of survival distributions of animals in Control-1, Control-2, Low, Medium, and High dose groups was tested separately for males and females using the Cox test (Cox, 1972) and the Generalized Wilcoxon test (Gehan, 1965). Results of the tests are given in Tables 2A and 2B for males and females, respectively. The statistical analysis showed no statistically significant (at .05 level) dose-response relationship[3] in survival across treatment groups in either sex.

Tumor data analysis: On the request of Dr. Hill, this reviewer performed only the pairwise comparisons of the high dose group with Control-1 and Control-2. This was done using the age adjusted Fisher exact test. Since the sponsor classified the tumor types as 'cause of death' and 'not a cause of death', following Peto et al. (1980) this reviewer applied the 'death rate method' and the 'prevalence method' for these two categories of tumors respectively. For tumor types occurring in both categories, a combined test of 'death rate method' and the 'prevalence method' was performed. The time intervals used were 0 - 52, 53 - 78, 79 - 91, 92 - 104 weeks, and terminal sacrifice for males and 0 - 52, 53 - 78, 79 - 91, 92 - 105 weeks, and terminal sacrifice for females. The tumor rates and the P-values for pairwise comparisons are given in Tables 3A and 3B for males and

[1] A tumor type is known as a rare tumor if it has a background rate of less than or equal to one percent.

[2] A tumor type is known as a common tumor if it has a background rate of more than one percent.

[3] In this review, the phrase "dose-response relationship" refers to the linear component of the effect of treatment, and not necessarily to a strictly increasing mortality or tumor rate as dose increases.

females, respectively.

Multiple testing adjustment: Adjustment for multiple comparisons was done using the results of Haseman (1983), which recommends to use p=0.01 for common tumor type, and p=0.05 for rare tumor type, in order to keep the false-positive rate at the nominal level of approximately five percent.

Based on the results of Haseman, Thyroid gland/Follicular cell adenoma (considering it as a rare tumor) in males showed statistically significant increased incidence in the high dose group when compared with the placebo control.

3. The Mouse study

Design: Two separate experiments, one in male and one in female were conducted. In each of these two experiments there were three treated and two control groups. Three hundred CD-1®(ICR) BR mice of each sex were equally (50 sex/group) randomly allocated to controls and treated groups. The dose levels for the treated group was 0.04% (Low), 0.4% (Medium), and 4.0% (High) mg/kg/day for three treated groups. Animals in one control group (Control-1) were treated with vehicle, and those in the other control group (Control-2) remained untreated. All surviving female animals were sacrificed after 99 weeks, while the males received treatment up to week 104.

Mortality and clinical signs were recorded periodically during the treatment period. Palpation of possible masses was carried out every two weeks, from the sixth month of the study. All protocol designated organs and tissues obtained from all animals died during the study, killed moribund or sacrificed at the end of the study were histopathologically examined.

3.1 Sponsor's analyses

Survival analysis: Number of animals died during the study in different dose groups was presented in tabular form. Kaplan-Miere survival curves for death rates were presented. The sponsor concluded that there was no effect on survival of animals, which were treated with the test article when compared with animals from either control.

Tumor data analysis: Tumor data were analyzed using the method described in the paper of Peto et al (1980). Score used were 1, 2, 3, 4, and 5 for Control-1, Control-2, Low, Medium, and High, respectively. Adjustment for multiple tests were done using the method suggested by Haseman (1983), which recommends to use p=0.01 for common tumor type, and p=0.05 for rare tumor type, in order to keep the false-positive rate at the nominal level of approximately five percent.

Statistical analysis did not show statistically significant treatment-related neoplastic or non-neoplastic findings.

3.2 Reviewer's analyses

This reviewer independently performed survival and tumor data analyses. For survival analysis the methods described in the papers of Cox (1972), and of Gehan (1965) were used. The tumor data analyses were performed using the methods described in the paper of Peto et al. (1980), the method of Exact Permutation trend test, and the age-adjusted Fisher exact test. Data used in this reviewer's analyses were provided by the sponsor on a CD.

Survival analysis: The intercurrent mortality data are given in Tables 4A and 4B for males and females, respectively. The Kaplan-Meyer curves for death rate are given in Figures 2A and 2B for males and females, respectively. The homogeneity of survival distributions of animals in Control-1, Control-2, Low, Medium, and High dose groups was tested separately for males and females using the Cox test (Cox, 1972) and the Generalized Wilcoxon test (Gehan, 1965). Results of the tests are given in Tables 5A and 5B for males and females, respectively. The statistical analysis showed no statistically significant (at .05 level) dose-response relationship in survival across the treatment groups in either sex.

Tumor data analysis: This reviewer performed the dose-response tests and pairwise comparisons of the high dose group with the controls. Since the sponsor classified the tumor types as 'cause of death' and 'not a cause of death', following Peto et al. (1980), this reviewer applied the 'death rate method' and the 'prevalence method' for these two categories of tumors respectively, to test the dose-response relationship. For tumor types occurring in both categories a combined test of 'death rate method' and the 'prevalence method' was performed. For the calculation of p-values, the Exact Permutation method was used. Scores used were 0, 0, 0.04, 0.4, and 4.0 for Control-1, Control-2 Low, Medium, and High dose groups, respectively. The time intervals used were 0 - 52, 53 - 78, 79 - 91, 92 - 104 weeks, and terminal sacrifice for males and 0 - 52, 53 - 78, 79 - 91, 92 - 99 weeks, and terminal sacrifice for females. The tumor rates and the p-values of the tumor types tested for the dose-response relationship are given in Table 6A and 6B for males and females, respectively. Pairwise comparisons of tumor rates between the high dose group with Control-1 and Control-2 were also performed using the age adjusted Fisher exact test. The tumor rates and the p-values for pairwise tests are given in Tables 7A and 7B for males and females, respectively.

Multiple testing adjustment: Adjustment for multiple trend tests was done using the results of Lin and Rahman (Lin K.K. and Rahman M.A., 1998. Overall false positive rates in tests for linear trend in tumor incidence in animal carcinogenicity studies of new drugs, *Journal of Biopharmaceutical Statistics,* 8(1), 1-15), which recommends to use p=0.025 for rare tumor type, and p=0.005 for common tumor type, in order to keep the false-positive rate at the nominal level of approximately five percent. Adjustment for multiple pairwise comparisons was done using the results of Haseman (1983), which recommends to use p=0.01 for common tumor type, and p=0.05 for rare tumor type, in order to keep the false-positive rate at the nominal level of approximately five percent.

Based on the results of Lin and Rahman for dose-response, and that of Haseman for pairwise comparisons no statistically significant dose-response or increased incidence of any tumor type in

the high dose group was found when compared with either control.

5. Summary

In this submission the sponsor included reports of two animal carcinogenicity studies. These studies were intended to assess the carcinogenic potential of Elidel (Pimecrolimus) Cream 1% in rats and mice with dermal administration. The length of the two studies was designed to be 104 weeks.

In this review, the phrase "Dose-response" refers to the linear component of the effect of treatment, and not necessarily to a strictly increasing or decreasing mortality or tumor rate as dose increases.

Rat Study: This study had 5 treatment groups namely, saline control, placebo control, 0.2%, 0.6%, and 1.0% body weight/day. The statistical analysis showed no statistically significant (at .05 level) dose-response in survival across treatment groups in either sex. Thyroid gland/Follicular cell adenoma in males showed statistically significant increased incidence in the high dose group when compared with the placebo control.

Mouse Study: This study had 5 treatment groups namely, vehicle control, untreated control, 0.04, 0.4, and 4.0 mg/kg/day. The statistical analysis showed no statistically significant (at .05 level) dose-response relationship in survival across treatment groups in either sex. No statistically significant dose-response or increased incidence of any tumor type was found in the high dose group when compared with either control.

M. Atiar Rahman, Ph.D.
Mathematical Statistician, Biometrics III

Concur: Mohamed Alosh, Ph.D.
Team Leader, Biometrics III

cc:
Archival NDA 21-302
HFD-540/Division File
HFD-540/Dr. Wilkin
HFD-540/Dr. Cook
HFD-540/Dr. Hill
HFD-540/Dr. Walker
HFD-540/Dr. Jacobs
HFD-540/Ms. Wright
HFD-725/ Chron
HFD-725/ Dr. Huque
HFD-725/ Dr. Alosh
HFD-725/ Dr. Rahman
HFD-700/ Dr. Anello

Table 1A: Intercurrent Mortality Rate
Male Wistar Rat

Week	Control-1*		Control-2*		Low (2 mg)		Medium (5 mg)		High	
	No. of Death	Cum. %	No. of Death	Cum. %	No. of Death	Cum. %	No. of Death	Cum. %	No. of Death	Cum. %
0 - 52	2	4.0	2	4.0	0	0.0	2	4.0	3	6.0
53 - 78	8	20.0	6	16.0	6	12.0	8	20.0	2	10.0
79 - 91	6	32.0	4	24.0	6	24.0	3	26.0	10	30.0
92 - 104	10	52.0	7	38.7	3	30.0	6	38.0	9	48.0
Term. Sac.	24	48.0	31	62.3	35	70.0	31	62.0	26	52.0

*Control-1: Saline control, Control-2: Placebo control

Table 1B: Intercurrent Mortality Rate
Female Wistar Rat

Week	Control-1*		Control-2*		Low (2 mg)		Medium (5 mg)		High	
	No. of Death	Cum. %	No. of Death	Cum. %	No. of Death	Cum. %	No. of Death	Cum. %	No. of Death	Cum. %
0 - 52	2	4.0	5	10.0	2	4.0	1	4.0	2	6.0
53 - 78	9	22.0	4	18.0	5	14.0	4	20.0	6	16.0
79 - 91	2	26.0	3	24.0	4	22.0	5	26.0	6	28.0
92 - 105	10	46.0	6	36.7	6	34.0	8	38.0	7	42.0
Term. Sac.	27	54.0	32	64.3	33	66.0	32	62.0	29	58.0

*Control-1: Saline control, Control-2: Placebo control

Table 2A: Intercurrent Mortality Comparison
Male Wistar Rat

Method	Test	Statistic	P-value
Cox	Homogeneity	5.52	0.2382
Kruskal-Wallis	Homogeneity	4.56	0.3351

Table 2B: Intercurrent Mortality Comparison
Female Wistar Rat

Method	Test	Statistic	P-value
Cox	Homogeneity	1.84	0.7655
Kruskal-Wallis	Homogeneity	1.63	0.8029

Table 3A

Pairwise comparisons p-values of Tested Tumors in Rat Study
Male Wistar Rats, Fed Over 104 Weeks

Organ	Tumor	Cont. 1	Cont. 2	1.0%	P-value using Cont. 1	P-value using Cont. 2
ADRENAL MEDU	Benign pheochromocyt	2	0	1	0.8868	0.4561
	Malignant pheochromo	1	1	1	0.7530	0.8250
BODY CAVITIES	Malignant Schwannoma	1	0	0	N/A	
	Malignant mesothelio	2	2	0	N/A	N/A
CEREBELLUM	Meningeal sarcoma	0	0	1	0.5316	0.5122
CEREBRUM	Granular cell tumor	2	1	0	1.0000	1.0000
	Malignant reticulosi	1	0	0	1.0000	
	Meningeal sarcoma	1	0	1	0.7000	0.7143
	Oligodendroglioma	0	0	1	0.5000	0.4531
COLON	Adenocarcinoma, muci	0	1	0		1.0000
DUODENUM	Leiomyoma	0	1	0		1.0000
EYES	Amelanotic melanoma	1	0	0	1.0000	
HINDFOOT	Hemangiosarcoma	0	1	0		1.0000
	Squamous cell papill	0	1	0		1.0000
JEJUNUM	Leiomyosarcoma	0	0	1	0.4923	0.4776
KIDNEYS	Lipomatous tumor	1	0	0	1.0000	
	Tubular cell adenoma	1	0	0	1.0000	
	Tubular cell carcino	1	1	1	0.7600	0.7361
LIVER	Hepatocellular adeno	0	1	1	0.5200	0.7086
	Hepatocellular carci	0	1	0		1.0000
LUNG	Alveolar/brohchiolar	0	0	1	0.6250	0.7143
	Squamous cell carcin	0	0	1	0.4923	0.4706
LYMPH NODE	Hemangiosarcoma	1	0	0	1.0000	
LYMPHORETIC SY	Histiocytic sarcoma	0	0	1	0.5366	0.5176
	Malignant fibrous hi	0	1	0		1.0000
	Malignant lymphoma	0	2	1	0.5072	0.8524
MESENT. LN	Hemangioma	3	8	8	0.1717	0.6910
PANCREAS	Islet cell adenoma	2	3	0	1.0000	1.0000
	Islet cell carcinoma	1	0	0	1.0000	
	Mixed acinar-islet c	1	0	0	1.0000	
PARATHYROID GL	Adenoma	2	1	3	0.5957	0.4214
PITUITARY GL	Adenoma/pars distali	29	17	11	0.9999	0.8858
	Adenoma/pars interme	2	0	0	1.0000	
SKIN (UNTRT)	Fibrosarcoma	0	0	2	0.2626	0.2351
	Keratoacanthoma	0	1	1	0.5200	0.7825
	Liposarcoma	0	0	1	0.4925	0.4783
	Malignant Schwannoma	0	1	0		1.0000
	Sarcoma (not otherwi	0	1	2	0.2600	0.4369
	Squamous cell papill	1	0	0	1.0000	
SPLEEN	Hemangiosarcoma	1	0	0	1.0000	
	Sarcoma (not otherwi	0	0	1	0.5200	0.4561

Cont. 1: Saline control, Cont. 2: Placebo control

Table 3A (Continued)

Pairwise comparisons p-values of Tested Tumors in Rat Study
Male Wistar Rats, Fed Over 104 Weeks

Organ	Tumor	Cont. 1	Cont. 2	1.0%	P-value using Cont. 1	P-value using Cont. 2
STOMACH	Leiomyosarcoma	0	0	1	0.5200	0.4561
TESTES	Benign Leydig cell t	1	1	4	0.2004	0.1264
	Hemangioma	1	0	0	1.0000	
THYMUS	Benign thymoma	1	0	2	0.5156	0.2096
THYROID GL	C-cell adenoma	3	5	0	1.0000	1.0000
	C-cell carcinoma	1	0	0	1.0000	
	Follicular cell aden	1	0	5	0.1323	0.0357

Cont. 1: Saline control, Cont. 2: Placebo control

Table 3B

Pairwise comparisons p-values of Tested Tumors in Rat Study
Female Wistar Rats, Fed Over 104 Weeks

Organ	Tumor	Cont. 1	Cont. 2	1.0%	P-value using Cont. 1	P-value using Cont. 2
ADRENAL MEDU	Benign pheochromocyt	3	0	0	1.0000	
BODY CAVITIES	Hemangioma	0	1	0		N/A
	Malignant Schwannoma	0	0	1	N/A	0.6667
CEREBELLUM	Granular cell tumor	3	0	1	0.9307	0.5385
CEREBRUM	Astrocytoma	0	0	2	0.2589	0.2480
	Granular cell tumor	1	1	2	0.5273	0.4625
DUODENUM	Leiomyoma	0	1	0		1.0000
HEART	Malignant endocardia	0	1	0		1.0000
HINDFOOT	Hemangioma	1	0	0	1.0000	
KIDNEYS	Tubular cell adenoma	1	0	0	1.0000	
	Tubular cell carcino	1	2	0	1.0000	1.0000
LIVER	Hepatocellular adeno	0	1	1	0.5179	0.7290
LUNG	Alveolar/bronchiolar	1	0	0	1.0000	
LYMPHORETIC SY	Malignant lymphoma	1	2	3	0.3237	0.4824
MAMMARY AREA	Adenocarcinoma	1	6	0	1.0000	1.0000
	Adenoma	0	1	0		1.0000
	Fibroadenoma	7	6	3	0.9708	0.9248
MESENT. LN	Hemangioma	2	3	5	0.1953	0.3362
OVARIES	Benign granulosa cel	2	2	2	0.6390	0.6777
	Benign granulosa-the	0	0	1	0.5179	0.4754
	Malignant granulosa	2	0	0	1.0000	
	Yolk sac carcinoma	0	0	1	0.5179	0.4754
OVIDUCTS	Cystadenocarcinoma	0	0	1	N/A	N/A
PANCREAS	Islet cell adenoma	0	2	0		1.0000
	Islet cell carcinoma	0	1	0		1.0000
PITUITARY GL	Adenoma/pars distali	42	41	34	0.9654	0.9270
	Adenoma/pars interme	0	2	0		1.0000
	Pituicytoma	0	0	1	0.5179	0.4754
SKIN (UNTRT)	Benign fibrous histi	1	0	0	1.0000	
	Fibrosarcoma	1	0	0	1.0000	
	Keratoacanthoma	0	0	1	0.5179	0.4754
	Lipoma	0	0	1	0.3750	0.5000
	Malignant Schwannoma	1	0	0	1.0000	
	Squamous cell papill	0	0	1	0.3750	0.5000
SPLEEN	Hemangiosarcoma	1	0	0	1.0000	
	Sarcoma (not otherwi	0	0	1	0.5179	0.4754
STOMACH	Adenoma, polypoid	1	0	0	1.0000	
THYMUS	Benign thymoma	4	2	2	0.9133	0.6555
THYROID GL	C-cell adenoma	2	2	4	0.3043	0.3304
	C-cell carcinoma	0	1	0		1.0000

Cont. 1: Saline control, Cont. 2: Placebo control

Table 3B (Continued)

Pairwise comparisons p-values of Tested Tumors in Rat Study
Female Wistar Rats, Fed Over 104 Weeks

Organ	Tumor	Cont. 1	Cont. 2	1.0%	P-value using Cont. 1	P-value using Cont. 2
THYROID GL	Follicular cell carc	1	0	1	0.7721	0.4754
UTERUS	Adenocarcinoma	3	3	2	0.8456	0.8136
	Granular cell tumor	1	0	0	1.0000	
	Hemangioma	0	1	0		1.0000
	Hemangiosarcoma	0	1	1	0.4835	0.7377
	Malignant Schwannoma	1	0	2	0.5005	0.2429
	Polyp/endometrial-st	4	4	5	0.5827	0.4918
VAGINA	Granular cell tumor	1	1	0	1.0000	1.0000
	Hemangioma	1	0	0	1.0000	

Cont. 1: Saline control, Cont. 2: Placebo control

Table 4A: Intercurrent Mortality Rate
Male CD-1 Mice

Week	Control-1		Control-2		Low (2 mg)		Medium (5 mg)		High	
	No. of Death	Cum. %	No. of Death	Cum. %	No. of Death	Cum. %	No. of Death	Cum. %	No. of Death	Cum. %
0 - 52	3	6.0	0	0.0	1	2.0	2	4.0	1	2.0
53 - 78	5	16.0	3	6.0	2	6.0	6	16.0	2	6.0
79 - 91	8	32.0	6	18.0	8	22.0	6	28.0	10	26.0
92 - 104	7	46.0	11	40.0	9	40.0	10	48.0	5	36.0
Term. Sac.	27	54.0	30	60.0	30	60.0	26	52.0	32	64.0

Conttrol-1: Untreated control, Control-2: Vehicle control

Table 4B: Intercurrent Mortality Rate
Female CD-1 Mice

Week	Control-1		Control-2		Low (2 mg)		Medium (5 mg)		High	
	No. of Death	Cum. %	No. of Death	Cum. %	No. of Death	Cum. %	No. of Death	Cum. %	No. of Death	Cum. %
0 - 52	4	8.0	0	0.0	4	84.0	2	4.0	2	4.0
53 - 78	8	24.0	4	8.0	8	24.0	2	8.0	5	14.0
79 - 91	5	34.0	6	20.0	8	40.0	9	26.0	6	26.0
92 - 99	5	44.0	6	32.7	5	50.0	9	44.0	4	34.0
Term. Sac.	28	56.0	34	68.3	25	50.0	28	56.0	33	66.0

Conttrol-1: Untreated control, Control-2: Vehicle control

Table 5A: Intercurrent Mortality Comparison
Male CD1 Mice

Method	Test	Statistic	P-value
Cox	Homogeneity	2.27	0.6861
Kruskal-Wallis	Homogeneity	2.70	0.6088

Table 5B: Intercurrent Mortality Comparison
Female CD1 Mice

Method	Test	Statistic	P-value
Cox	Homogeneity	5.43	0.2455
Kruskal-Wallis	Homogeneity	6.20	0.1849

Table 6A

Tumor Rates and Dose Response p-values of Tested Tumors in Mouse Study
Male CD-1 Mice, Fed Over 104 Weeks

Organ	Tumor	Cont. 1	Cont. 2	0.04 mg	0.4 mg	4 mg	P-value using Cont. 1	P-value using Cont. 2
ADRENAL GLANDS	ADENOMA, CORTICAL	0	1	1	0	1	0.3490	0.5704
	TUMOR, MEDULLARY, BE	0	1	0	0	0		1.0000
BRAIN	MENINGIOMA	0	0	0	0	1	0.2683	0.2543
EPIDIDYMIDES	GRANULAR CELL TUMOR	0	1	0	0	0		1.0000
	TUBULAR ADENOMA	0	0	1	0	0	0.7652	0.7458
EYES	MELANOMA, MALIGNANT	0	0	0	0	1	0.3000	0.3103
HARDERIAN GLAND	ADENOMA	3	3	3	4	4	0.3781	0.2857
HEMOLYMPHORET.	HISTIOCYTIC SARCOMA	0	1	0	0	1	0.2593	0.4301
	MALIGNANT LYMPHOMA	6	2	3	2	5	0.3922	0.1212
KIDNEYS	ADENOMA, RENAL TUBUL	0	0	0	1	0	0.5043	0.4915
LIVER	ADENOMA, HEPATOCELLU	4	8	5	1	3	0.7473	0.8279
	CARCINOMA, HEPATOCEL	2	3	0	1	0	0.8773	0.9296
	HEMANGIOMA	1	0	0	0	0	1.0000	
	HEMANGIOSARCOMA	1	1	0	0	0	1.0000	1.0000
	HEPATOCELLULAR ADENO	1	0	0	0	0	1.0000	
	HEPATOCELLULAR ADENO	1	0	2	0	1	0.6719	0.4565
LUNGS WITH BRON	ADENOMA, BRONCHIO-AL	7	12	9	8	8	0.4912	0.7082
	BRONCHOALVEOLAR ADEN	1	1	1	0	2	0.2837	0.2684
	CARCINOMA, BRONCHIO-	7	7	8	8	9	0.3796	0.3408
MESENT. LYMPH N	HEMANGIOMA	0	1	0	0	0		1.0000
PITUITARY GLAND	ADENOMA OF PARS DIST	1	1	0	0	0	1.0000	1.0000
	ADENOMA OF PARS INTE	0	0	1	0	0	0.7652	0.7458
SKIN	HISTIOCYTOMA, FIBROU	0	0	0	1	0	0.5043	0.4915
SPLEEN	HEMANGIOMA	0	0	1	0	0	0.7652	0.7458
	HEMANGIOSARCOMA	0	0	0	1	0	0.5043	0.4915
TESTES	ADENOMA, INTERSTITIA	3	3	3	1	3	0.5222	0.4582
	ADENOMA, RETE TESTIS	0	0	0	0	1	0.2783	0.2712
THORACIC CAVITY	MESOTHELIOMA	0	0	1	0	0	0.3333	N/A
THYROID GLANDS	ADENOMA, FOLLICULAR	0	0	1	0	0	0.7586	0.7857

Cont. 1: Untreated control, Cont. 2: Vehicle control

Table 6B

Tumor Rates and Dose Response p-values of Tested Tumors in Mouse Study
Female CD-1 Mice, Fed Over 104 Weeks

Organ	Tumor	Cont. 1	Cont. 2	0.04 mg	0.4 mg	4 mg	P-value using Cont. 1	P-value using Cont. 2
ADRENAL GLANDS	ADENOMA, CORTICAL	0	0	0	1	0	0.5351	0.5083
	TUMOR, MEDULLARY, BE	0	0	0	1	0	0.5351	0.5083
BONE	OSTEOSARCOMA	0	0	0	0	1	N/A	N/A
FEMORAL BONE(+A	LIPOMA	0	1	0	0	0		1.0000
HARDERIAN GLAND	ADENOMA	0	3	0	4	1	0.3933	0.7031
HEMOLYMPHORET.	HISTIOCYTIC SARCOMA	1	2	2	0	0	0.9481	0.9757
	MALIGNANT LYMPHOMA	9	10	6	6	6	0.7819	0.7697
	MYELOID LEUKEMIA	0	0	0	1	0	0.5394	0.5145
LIVER	ADENOMA, HEPATOCELLU	1	0	0	1	1	0.4561	0.2034
	CARCINOMA, HEPATOCEL	0	0	0	1	1	0.2251	0.2022
	HEMANGIOMA	1	1	0	0	0	1.0000	1.0000
	HEMANGIOSARCOMA	0	0	0	0	1	0.2727	0.2625
LUNGS WITH BRON	ADENOMA, BRONCHIO-AL	5	5	5	4	7	0.2927	0.2615
	BRONCHOALVEOLAR ADEN	0	0	1	1	2	0.1577	0.1321
	CARCINOMA, BRONCHIO-	2	1	4	5	2	0.7799	0.6952
MAMMARY AREA	ADENOCARCINOMA	2	1	0	2	1	0.6019	0.3963
MESENT. LYMPH N	HEMANGIOMA	2	0	0	0	0	1.0000	
OVARIES	ADENOMA, TUBULOSTROM	0	0	1	0	0	0.7544	0.7167
	CARCINOMA	0	1	0	0	0		1.0000
	LUTEOMA, BENIGN	1	0	2	2	0	0.8918	0.7689
	THECOMA, MALIGNANT	1	0	0	0	0	1.0000	--
SKIN	CARCINOMA, SQUAMOUS	0	0	2	0	0	0.8472	0.8228
	HISTIOCYTOMA, FIBROU	1	2	1	0	3	0.1036	0.1621
	KERATOACANTHOMA	0	0	0	1	0	0.5318	0.5111
SPLEEN	HEMANGIOMA	0	0	0	1	0	0.5351	0.5083
	HEMANGIOSARCOMA	0	0	0	0	1	0.2714	0.2585
STOMACH	LEIOMYOSARCOMA	1	0	0	0	0	1.0000	
SUBCUTANEOUS TI	OSTEOSARCOMA	0	0	1	0	1	0.3303	0.3548
THYROID GLANDS	ADENOCARCINOMA, FOLL	0	0	0	1	0	0.5351	0.5083
	ADENOMA, FOLLICULAR	1	0	0	0	0	1.0000	
UTERUS	CARCINOMA, ADENOSQUA	0	3	0	0	0		1.0000
	FIBROMA	0	0	0	0	1	0.2174	0.2632
	FIBROSARCOMA	6	1	1	1	1	0.9338	0.5267
	HEMANGIOMA	0	0	1	0	0	0.7607	0.7236
	LEIOMYOMA	1	2	0	2	0	0.7218	0.8336
	METASTATIC OSTEOSARC	0	0	0	1	0	0.5652	0.5417

Cont. 1: Untreated control, Cont. 2: Vehicle control

Table 7A

Pairwise comparisons p-values of Tested Tumors in Mouse Study
Male CD-1 Mice, Fed Over 104 Weeks

Organ	Tumor	Cont. 1	Cont. 2	4 mg	P-value using Cont. 1	P-value using Cont. 2
ADRENAL GLANDS	ADENOMA, CORTICAL	0	1	1	0.5424	0.7781
	TUMOR, MEDULLARY, BE	0	1	0		1.0000
BRAIN	MENINGIOMA	0	0	1	0.5432	0.4889
EPIDIDYMIDES	GRANULAR CELL TUMOR	0	1	0		1.0000
EYES	MELANOMA, MALIGNANT	0	0	1	0.5625	0.6000
HARDERIAN GLAND	ADENOMA	3	3	4	0.5971	0.4044
HEMOLYMPHORET.	HISTIOCYTIC SARCOMA	0	1	1	0.5250	0.7279
	MALIGNANT LYMPHOMA	6	2	5	0.7763	0.2273
LIVER	ADENOMA, HEPATOCELLU	4	8	3	0.8523	0.9474
	CARCINOMA, HEPATOCEL	2	3	0	1.0000	1.0000
	HEMANGIOMA	1	0	0	1.0000	
	HEMANGIOSARCOMA	1	1	0	1.0000	1.0000
	HEPATOCELLULAR ADENO	1	0	0	1.0000	
	HEPATOCELLULAR ADENO	1	0	1	0.7949	0.5161
LUNGS WITH BRON	ADENOMA, BRONCHIO-AL	7	12	8	0.5403	0.8625
	BRONCHOALVEOLAR ADEN	1	1	2	0.5645	0.5246
	CARCINOMA, BRONCHIO-	7	7	9	0.4471	0.3847
MESENT. LYMPH N	HEMANGIOMA	0	1	0		1.0000
PITUITARY GLAND	ADENOMA OF PARS DIST	1	1	0	1.0000	1.0000
TESTES	ADENOMA, INTERSTITIA	3	3	3	0.7429	0.6285
	ADENOMA, RETE TESTIS	0	0	1	0.5424	0.5161

Cont. 1: Untreated control, Cont. 2: Vehicle control

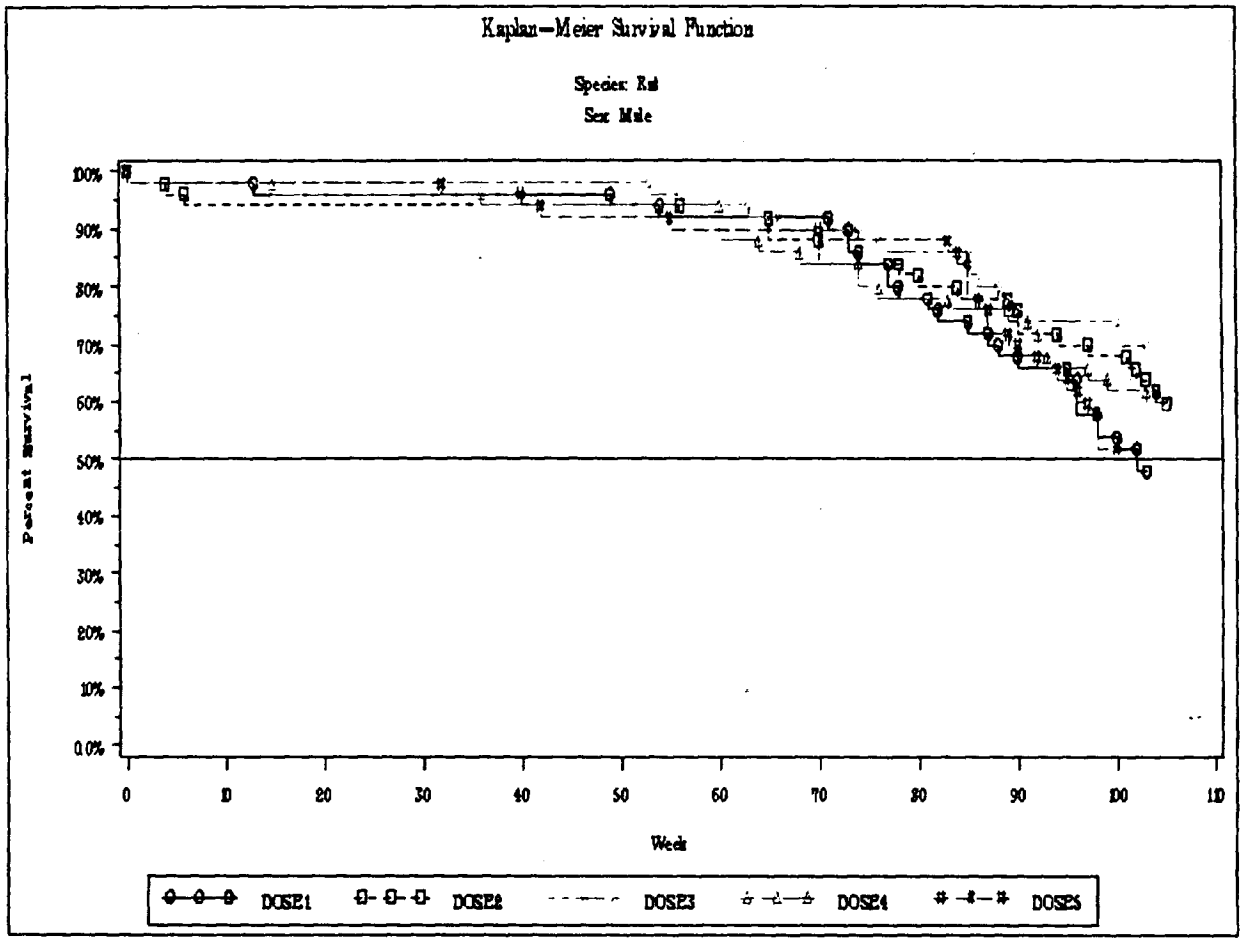
Table 7B

Pairwise comparisons p-values of Tested Tumors in Mouse Study
Female CD-1 Mice, Fed Over 104 Weeks

Organ	Tumor	Cont. 1	Cont. 2	4 mg	P-value using Cont. 1	P-value using Cont. 2
BONE	OSTEOSARCOMA	0	0	1	N/A	N/A
FEMORAL BONE(+A	LIPOMA	0	1	0		1.0000
HARDERIAN GLAND	ADENOMA	0	3	1	0.5410	0.9395
HEMOLYMPHORET.	HISTIOCYTIC SARCOMA	1	2	0	1.0000	1.0000
	MALIGNANT LYMPHOMA	9	10	6	0.9138	0.8925
LIVER	ADENOMA, HEPATOCELLU	1	0	1	0.7934	0.4925
	CARCINOMA, HEPATOCEL	0	0	1	0.5410	0.4925
	HEMANGIOMA	1	1	0	1.0000	1.0000
	HEMANGIOSARCOMA	0	0	1	0.5385	0.5000
LUNGS WITH BRON	ADENOMA, BRONCHIO-AL	5	5	7	0.4234	0.3582
	BRONCHOALVEOLAR ADEN	0	0	2	0.2885	0.2388
	CARCINOMA, BRONCHIO-	2	1	2	0.7160	0.4886
MAMMARY AREA	ADENOCARCINOMA	2	1	1	0.9041	0.7368
MESENT. LYMPH N	HEMANGIOMA	2	0	0	1.0000	
OVARIES	CARCINOMA	0	1	0		1.0000
	LUTEOMA, BENIGN	1	0	0	1.0000	
	THECOMA, MALIGNANT	1	0	0	1.0000	
SKIN	HISTIOCYTOMA, FIBROU	1	2	3	0.3347	0.4627
SPLEEN	HEMANGIOSARCOMA	0	0	1	0.5352	0.4872
STOMACH	LEIOMYOSARCOMA	1	0	0	1.0000	
SUBCUTANEOUS TI	OSTEOSARCOMA	0	0	1	0.5333	0.6154
THYROID GLANDS	ADENOMA, FOLLICULAR	1	0	0	1.0000	
UTERUS	CARCINOMA, ADENOSQUA	0	3	0		1.0000
	FIBROMA	0	0	1	0.3846	0.5556
	FIBROSARCOMA	6	1	1	0.9952	0.7463
	LEIOMYOMA	1	2	0	1.0000	1.0000

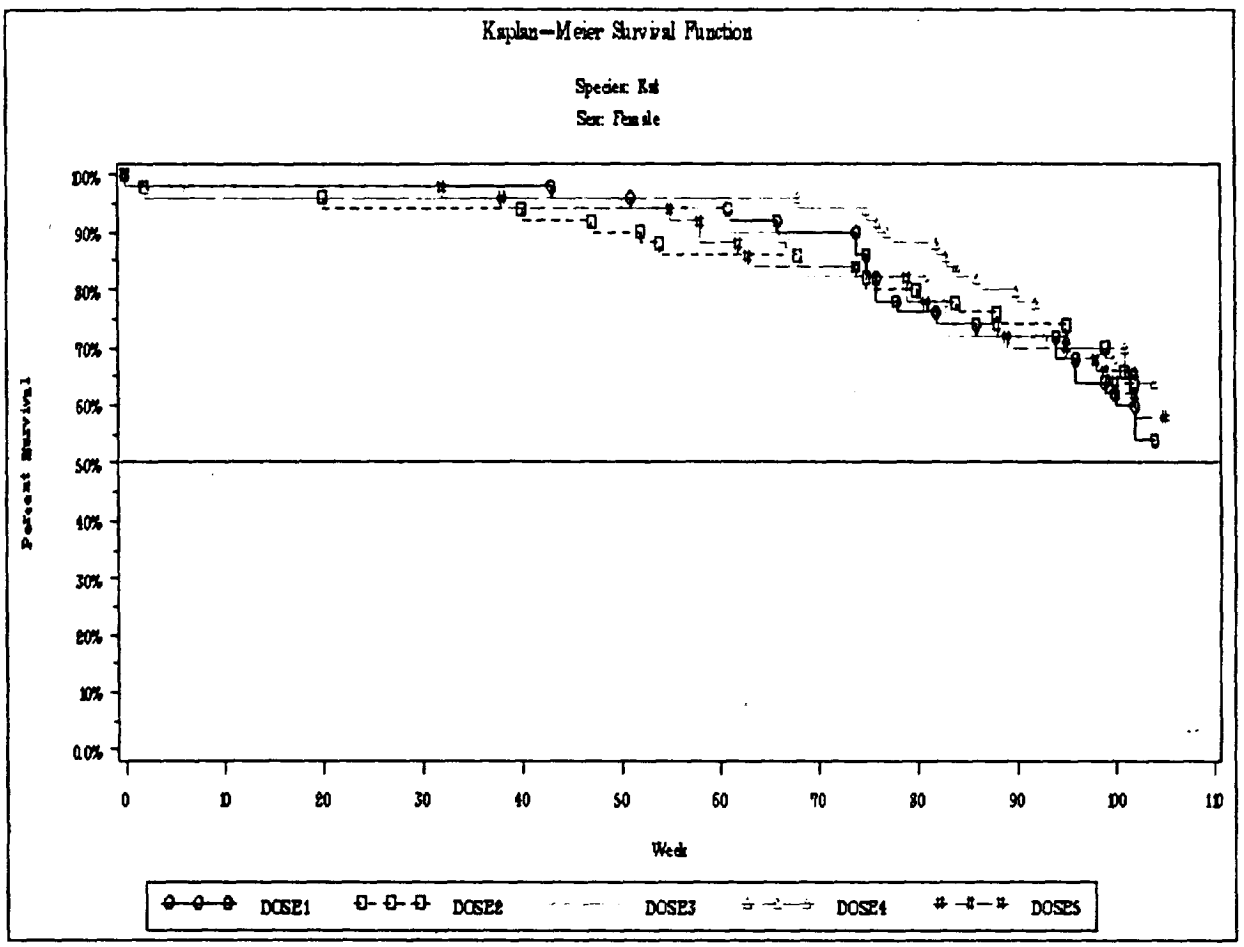
Cont. 1: Untreated control, Cont. 2: Vehicle control

Figure 1A
Kaplan-Meier Survival Curve
Male Rat



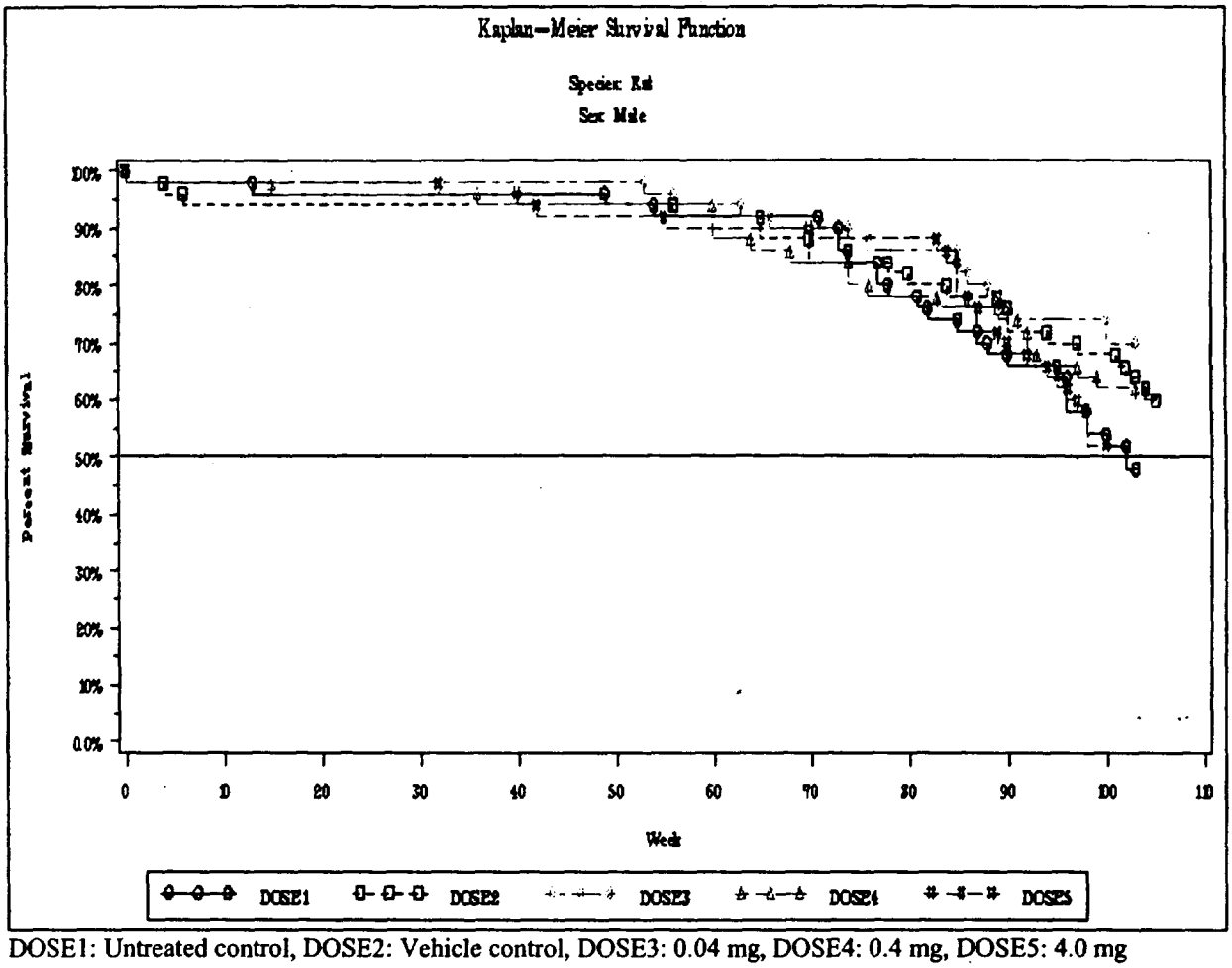
DOSE1: Saline control, DOSE2: Placebo control, DOSE3: 0.2%, DOSE4: 0.6%, DOSE5: 1.0%

Figure 1B
Kaplan-Meier Survival Curve
Female Rat



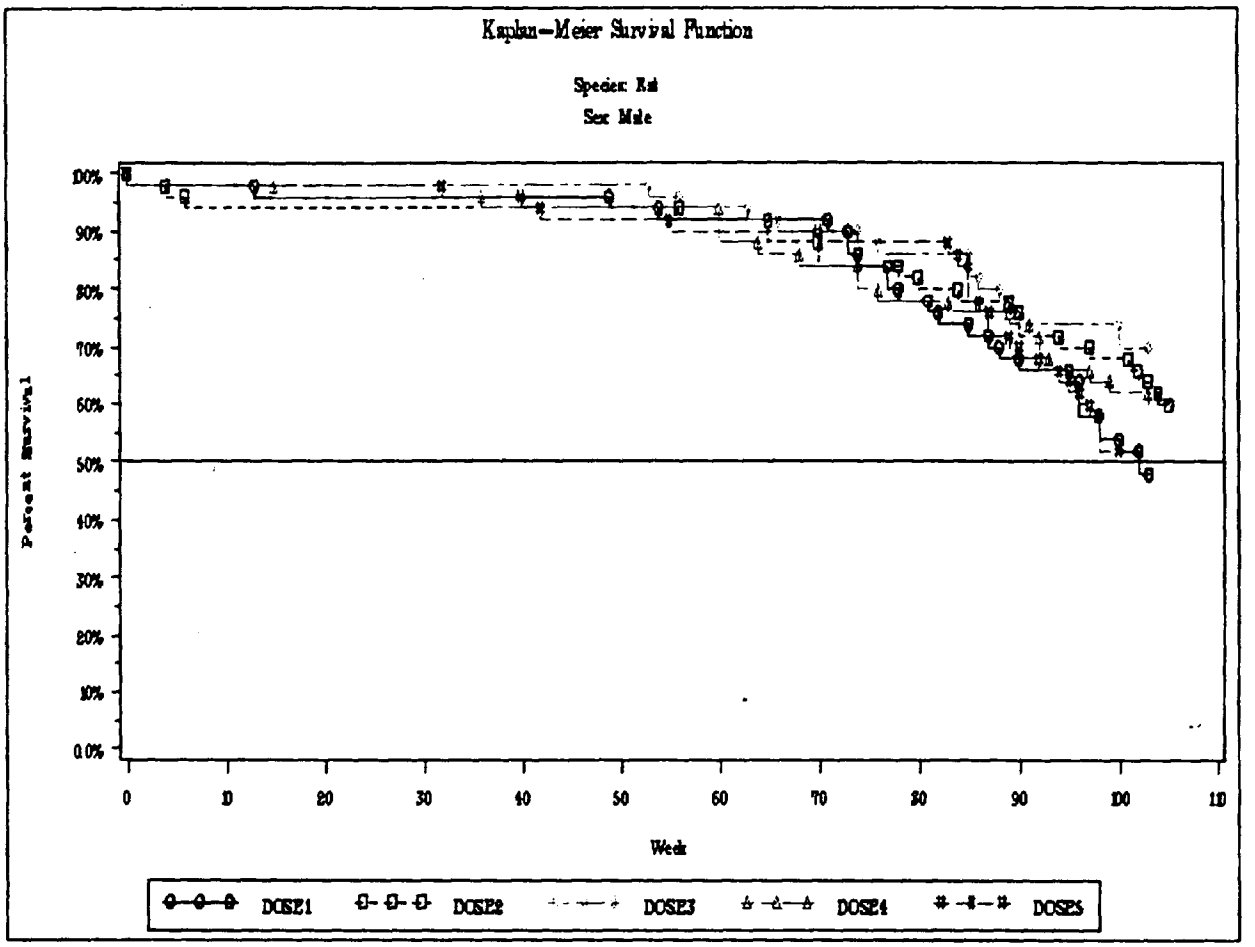
DOSE1: Saline control, DOSE2: Placebo control, DOSE3: 0.2%, DOSE4: 0.6%, DOSE5: 1.0%

Figure 2A
Kaplan-Meier Survival Curve
Male Mouse



DOSE1: Untreated control, DOSE2: Vehicle control, DOSE3: 0.04 mg, DOSE4: 0.4 mg, DOSE5: 4.0 mg

Figure 2B
Kaplan-Meier Survival Curve
Female Mouse



DOSE1: Untreated control, DOSE2: Vehicle control, DOSE3: 0.04 mg, DOSE4: 0.4 mg, DOSE5: 4.0 mg

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

/s/

Atiar Rahman
6/8/01 12:09:12 PM
BIOMETRICS

Mohamed Alosh
6/8/01 05:13:04 PM
BIOMETRICS
Concur with review

APPEARS THIS WAY
ON ORIGINAL