CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

205437Orig1s000

STATISTICAL REVIEW(S)

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

Statistical Review

CLINICAL STUDY

NDA / Sequence Number:	NDA 205437 / Seq 0000
Drug Name:	Apremilast
Proposed Indication:	Psoriatic arthritis
Applicant:	Celgene
Date(s):	Received: 03-21-2013 PDUFA Due Date: 03-21-2014
Review Priority:	Standard
Biometrics Division:	Division of Biometrics II
Statistical Reviewer:	Robert Abugov, Ph.D.
Concurring Reviewer:	Joan Buenconsejo, Ph.D.
Statistics Supervisor:	Thomas Permutt, Ph.D. (Division Director)
Medical Division:	Division of Pulmonary, Allergy, and Rheumatology Products
Clinical Team:	Keith Hull, M.D. (Medical Reviewer) Janet Maynard, M.D. (Acting Team Leader) Sarah Yim, M.D. (Supervisory Associate Director) Badrul A Chowdhury, M.D., Ph.D. (Medical Division Director)
Project Manager:	Michelle Jordan Garner

Keywords: NDA review, Clinical Studies

Contents

Figures

Tables

1 EXECUTIVE SUMMARY

Celgene has proposed apremilast 30 mg tablets two times daily (bid) for the treatment of adult patients with active psoriatic arthritis (PsA). Efficacy and safety of this phosphodiesterase-4 (PDE4) inhibitor were examined in three Phase 3 clinical trials. The primary efficacy endpoint was modified ACR 20[1] response.

This submission demonstrates benefits of apremilast 30 mg tablets compared to placebo for the treatment of adult patients with active PsA. Three randomized double blinded placebo controlled parallel arm Phase 3 trials show that apremilast 30 mg provides statistically significant benefits compared to placebo for the primary endpoint ACR 20 at week 16 (average 18%) as well as for the key secondary endpoint ΔHAQ-DI at week 16 (average -0.14).

Evidence for additional efficacy benefits of apremilast 30 mg over apremilast 20 mg are suggestive but not conclusive or even consistent, with effects of apremilast 20 mg compared to placebo statistically significant for the primary endpoint ACR 20 at week 16 in all three phase 3 studies, and statistically significant for key secondary endpoint HAQ-DI in two of the three phase 3 studies. Approval of apremilast 20 mg rather than apremilast 30 mg may therefore be justifiable if apremilast 30 mg poses large additional risks to safety compared to apremilast 20 mg.

Claims of effectiveness for endpoints at week 24 are considered in this review as claims for sustained effect beyond week 16. Such claims were confirmed for ACR 20 but undermined for other endpoints by the loss of adequate control; approximately 70% of placebo patients discontinued initial randomized treatment prior to week 24.

(b) (4)

A large number of proposed claims were based on endpoints in the analysis hierarchy below failed significance tests on enthesitis at week 16. Because analyses of these endpoints were only exploratory, tests of their statistical significance were considered only nominal, with p-values underestimating the probability of Type I error.

[1] A patient who is an ACR 'N' responder has a reduction of at least N% in the number of swollen joints, a reduction of at least N% in the number of tender joints, and a reduction of N% in three of the following five parameters: physician global assessment of disease, patient global assessment of disease, patient assessment of pain, C-reactive protein or erythrocyte sedimentation rate, or Health Assessment Questionnaire Disability Score (HAQ-DI).

2 INTRODUCTION

2.1 Overview

2.1.1 Drug Class and Indication

Apremilast is a PDE4 inhibitor proposed for the treatment of adult patients with active psoriatic arthritis.

2.1.2 History of Drug Development

The apremilast clinical development program for PSA was introduced to the Agency under IND 101761.

Design and analysis of the three Phase 3 studies (Table 1) was discussed at the End-of-Phase 2 teleconference held on March 25, 2010. The Agency agreed that nonclinical studies completed at that time were sufficient to support initiation of Phase 3 studies in PsA. The sponsor proposed three multicenter, international, randomized, double-blind, parallel group Phase 3 studies, CC-10004-PSA -002, -003, and -004 (studies 2, 3, and 4), to compare to placebo (P), after 24 weeks of treatment, two doses of apremilast, 20 mg bid (A20) and 30 mg bid (A30), with primary endpoint ACR20 analyzed using the Cochran-Mantel-Haenszel test and key secondary endpoint change from baseline HAQ-DI analyzed using analysis of covariance (ANCOVA). Provisions for escape therapy after week 16, prior to the primary endpoint at week 24, were made for patients in the placebo arm. To control Type I error for testing multiple doses, the sponsor proposed using the Hochberg procedure, with an analysis hierarchy to control Type 1 error when testing statistical significance of multiple endpoints.

The Agency responded that the proposed Phase 3 studies exceeded FDA requirements for two Phase 3 studies with a controlled duration of 12 weeks, and that the statistical analyses proposed were reasonable. The Agency also agreed with the proposed test doses A20 and A30, noting that a previously conducted Phase 2 trial CC-10004 PSA-001 had shown statistically similar ACR20 responses for 20 mg bid and 40 mg bid doses.

In written response communicated to the sponsor on June 29, 2012 regarding a meeting request sent April 12, 2012, the Agency agreed with the a revised plan to test the primary and secondary endpoints at week 16 rather than week 24. The Agency also noted that, because other effective therapies were available, the benefit-risk profile of apremilast would be a review issue. The Agency agreed with the sponsor's proposal to impute ACR20 non-response for patients who discontinued the study prior to week 16, (b) (4) and

that the statistical analysis plan should justify an imputation method based on careful examination of potential mechanisms by which data could be missing.

In a pre-NDA meeting held December 19, 2012, the Agency agreed that, although there was adequate efficacy data to support the filing of an NDA for apremilast as a treatment for PsA, preliminary analyses submitted by the sponsor indicated observed treatment benefits of questionable clinical relevance, with minimal differences between A20 and A30. The sponsor replied that the enrolled patient population had already shown an inadequate response to previously approved DMARDs, and that many of the enrolled patients were administered apremilast or placebo as an add-on to DMARD therapy, reducing the expected difference between treatment and placebo. The sponsor also noted that, although there were no statistically significant differences between A20 and A30 for ACR20 response rates, in all three Phase 3 studies, response rates were numerically higher for A30 compared to A20, and that, across all Phase 3 studies, more secondary endpoints achieved statistically significant difference from placebo in the A30 arms than in the A20 arms.

The sponsor also detailed safety tables to be submitted, with percent and exposure adjusted incidence rates for each arm from 0-4, 0-6 and 0-12 months of treatment, regardless of when the patients began treatment.

2.1.3 Current Submission

The applicant's proposed indication for the treatment of active PsA is based on three similar parallel arm placebo-controlled studies, CC-10004-PSA-002, CC-10004-PSA-003 and CC-10004-PSA-004, hereafter referred to as studies 2, 3, and 4 (Table 1). Each study enrolled approximately 495 patients and randomized equal numbers of patients to P, A20, or A30.

Study 2 began enrolling patients on 02 June 2010, and the last patient completed the week 24 visit on 26 March 2012; study 3 began enrolling patients on 27 September 2010, and the last patient completed the week 24 visit on 04 July 2012; study 4 began enrolling patients on 11 October 2011, and the last patient completed the week 24 visit on 9 July 2012.

Patients in study 2 were enrolled from Australia, Austria, Canada, France, Germany, Hungary, New Zealand, Poland, the Russian Federation, South Africa, Spain, the United Kingdom, and the United States; patients in study 3 were enrolled from Belgium, Bulgaria, Canada, Czech Republic, Estonia, France, Germany, Hungary, Italy, Poland, Russian Federation, South Africa, Spain, Taiwan, Province of China, the United Kingdom, and the United States; patients in study 4 were enrolled from Australia, Austria, Canada, France, Germany, Hungary, New Zealand, Poland, the Russian Federation, South Africa, Spain, the United Kingdom, and the United States.

Final database lock dates for studies 2, 3, and 4 were 21 June 2012, 26 July 2012, and 21 August 2012. The final statistical analysis plans were updated on 03 July 2012.

Table 1. Phase 3 Studies in Current Submission

Study[1]	Design	Population	Endpoints
PSA002 (Palace 1) *Study 2* PSA003 (Palace 2) *Study 3* PSA004 (Palace 3) *Study 4*	A30 A20 P Parallel arm DB EE W16 P to W24	Adults Active PsA Inadequately controlled by previous DMARDs May have current stable DMARD therapy Study PSA004: Qualifying Psoriatic Skin Lesion[2] N=495 1:1:1	Primary: Modified[3] ACR20 W16 Key Secondary: Δ HAQ-DI W16

1. Study names in parentheses cross reference to label.
2. Lesion ≥ 2 cm
3. Modified ACR20 includes distal interphalangeal joints, for a total of 78 joints examined for tenderness and 76 joint examined for swelling.

2.2 Data Sources

Data for all three studies was provided by the sponsor and is currently located at:

\\Cdsesub1\evsprod\NDA205437\0000\m5\datasets .

3 STATISTICAL EVALUATION

3.1 Data and Analysis Quality

The original submission omitted programs for the statistical tests. An information request to the sponsor satisfactorily resolved this issue.

3.2 Evaluation of Efficacy

3.2.1 Study Design and Endpoints

The three parallel arm, double blind, double dummy, multinational Phase III studies (Table 1) randomized adult patients who had active PsA despite prior treatment with DMARDs to A20, A30, or P, 1:1:1, for 24 weeks. Treatment was by oral tablet two times daily, titrated[2] during the first seven days of treatment with blinding maintained by providing visually identical blister cards to all subjects. During the study, patients were allowed to continue prior DMARD therapy on methotrexate, sulfasalazine, or leflunomide, with one increase permitted after week 24. The proportion of the study population that had experienced therapeutic failure with TNF blockers was limited to no more than 10% of the patients enrolled. Treatment assignments were stratified according to whether patients were using the aforementioned DMARDs, with at least 25 subjects in each study taking either leflunomide or sulfasalazine.

The primary (ACR20) and key secondary (change from baseline HAQ-DI) endpoints were measured at week 16. To further characterize treatment effect on HAQ-DI, the proportion of subjects achieving a decrease from baseline HAQ-DI ≥ 0.3 was summarized. Additional secondary endpoints not included in Table 1 are provided in the Appendix.

At week 16, early escape was provided for all patients with < 20% improvement from baseline in either tender or swollen joint count, with early escape patients initially randomized to P rerandomized 1:1 in blinded fashion[3] to A20 or A30. At week 24 the placebo controlled phase of each trial was terminated, with an open label phase in which all patients on placebo rerandomized 1:1 to A20 or A30.

Visit windows for week 4 and 28 visits were ±4 days, and visit windows for other days up to week 52 were ±7 days.

[2] During the first week, apremilast dosage was ramped from 10 mg to 20 mg for patients randomized to A20, and was ramped from 10 mg to 20 mg to 30 mg for patients randomized to A30. Blinding was maintained by providing doses in a blister card containg tablets identical in appearance. Use of the blister card was continued throughout the study.

[3] Patients whose swollen and painful joint scores had not improved by ≥ 20% at 16 weeks were told that they were going into early escape. Placebo patients were rerandomized in a blinded fashion by an Interactive Voice Response System (IVRS) 1:1 to A20 or A30. Apremilast patients were blindly "rerandomized" by IVRS to the same dose group to which they were originally assigned. All early escape patients received identically appearing blister cards.

3.2.2 Statistical Methodologies

Statistical analyses were conducted on all randomized subjects at the two sided 0.05 level of significance, using Cochran-Mantel-Haenszel (CMH) tests for discrete endpoints and analysis of covariance (ANCOVA) for continuous endpoints. The ANCOVA included baseline reading as a covariate, and both the CMH and ANCOVA tests controlled for baseline DMARD usage (Yes/No). For study 4, the statistical analyses additionally controlled for $\geq$ 3% body surface area with psoriasis at baseline.

Control of Type I error within each endpoint was maintained using the Hochberg procedure. In particular, pairwise comparisons were made between A30 and P and between A20 and P, with differences considered statistically significant if both comparisons were significant at the 0.05 level or if one comparison was significant at the 0.025 level. Endpoints were tested in a hierarchy, with the primary endpoint tested first, then the key secondary endpoint, followed by other secondary endpoints in the order listed in the Appendix, Section 6.

For ACR response endpoints, the primary analysis used non-responder imputation for patients missing data at week 16. At week 24, ACR non-responder imputation was applied not only to patients missing data, but also to patients who escaped early Analyses of other binary endpoints not involving joint counts were based on last observation carried forward (LOCF). For patients discontinuing initially assigned treatment, missing data for continuous endpoints at weeks 16 and 24 was imputed using LOCF, with sensitivity analyses at week 16 based on baseline observation carried forward (BOCF). Efficacy analyses were also performed at week 52, according to original randomized treatment.

Unassessed joints classified permanently unassessable at baseline were excluded from the analyses, while those which were not assessed for other reasons were classified using BOCF.

3.2.3 Patient Disposition, Demographic and Baseline Characteristics

There were no obvious differences between treatments for baseline characteristics in the three submitted Phase III studies (Table 2, Table 3, and Table 4).

Table 2. Baseline Demographics, n (%), Study 2

Variable	Class	P	A20	A30
Full Analysis Set		168	168	168
Age	< 40	30 (18)	34 (20)	30 (18)
	40 - < 65	119 (71)	123 (73)	116 (69)
	65 - < 75	14 (8)	11 (7)	20 (12)
	75 - < 85	5 (3)	0 (0)	2 (1)
Sex	F	80 (48)	83 (49)	92 (55)
	M	88 (52)	85 (51)	76 (45)
Country	USA	48 (29)	44 (26)	43 (26)
	NOT USA	120 (71)	124 (74)	125 (74)
DMARD	Yes	108 (64)	107 (64)	108 (64)
	No	60 (36)	61 (36)	60 (36)
Race	WHITE	153 (91)	150 (89)	152 (90)
	AMERICAN INDIAN OR ALASKA NATIVE	1 (1)	2 (1)	0 (0)
	ASIAN	8 (5)	8 (5)	8 (5)
	BLACK OR AFRICAN AMERICAN	0 (0)	2 (1)	0 (0)
	NATIVE HAWAIIAN OR OTHER PACIFIC ISLANDER	1 (1)	1 (1)	1 (1)
	OTHER	5 (3)	5 (3)	7 (4)
Ethnicity	HISPANIC OR LATINO	4 (2)	2 (1)	3 (2)
	NOT HISPANIC OR LATINO	164 (98)	166 (99)	165 (98)

source: Demographics 2013 06 12.sas

Table 3. Baseline Demographics, n (%), Study 3

Variable	Class	P	A20	A30
Full Analysis Set		159	163	162
Age	< 40	22 (14)	30 (18)	30 (19)
	40 - < 65	121 (76)	119 (73)	114 (70)
	65 - < 75	13 (8)	12 (7)	17 (10)
	75 - < 85	3 (2)	2 (1)	1 (1)
Sex	F	85 (53)	95 (58)	95 (59)
	M	74 (47)	68 (42)	67 (41)
Country	USA	18 (11)	27 (17)	30 (19)
	NOT USA	141 (89)	136 (83)	132 (81)
DMARD	Yes	110 (69)	113 (69)	112 (69)
	No	49 (31)	50 (31)	50 (31)
Race	WHITE	152 (96)	151 (93)	157 (97)
		1 (1)	0 (0)	0 (0)
	ASIAN	3 (2)	9 (6)	1 (1)
	BLACK OR AFRICAN AMERICAN	2 (1)	1 (1)	1 (1)
	OTHER	1 (1)	2 (1)	3 (2)
Ethnicity		1 (1)	0 (0)	0 (0)
	HISPANIC OR LATINO	1 (1)	1 (1)	2 (1)
	NOT HISPANIC OR LATINO	157 (99)	162 (99)	160 (99)

source: Demographics 2013 06 12.sas

Table 4. Baseline Demographics, n (%), Study 4

Variable	Class	P	A20	A30
Full Analysis Set		169	169	167
Age	< 40	36 (21)	35 (21)	30 (18)
	40 - < 65	119 (70)	117 (69)	122 (73)
	65 - < 75	11 (7)	15 (9)	14 (8)
	75 - < 85	3 (2)	2 (1)	1 (1)
Sex	F	91 (54)	90 (53)	88 (53)
	M	78 (46)	79 (47)	79 (47)
Country	USA	40 (24)	48 (28)	42 (25)
	NOT USA	129 (76)	121 (72)	125 (75)
DMARD	Yes	101 (60)	102 (60)	100 (60)
	No	68 (40)	67 (40)	67 (40)
Race	WHITE	158 (93)	161 (95)	163 (98)
	ASIAN	7 (4)	6 (4)	2 (1)
	BLACK OR AFRICAN AMERICAN	2 (1)	0 (0)	0 (0)
	NATIVE HAWAIIAN OR OTHER PACIFIC ISLANDER	1 (1)	0 (0)	0 (0)
	OTHER	1 (1)	2 (1)	2 (1)
Ethnicity	HISPANIC OR LATINO	9 (5)	9 (5)	9 (5)
	NOT HISPANIC OR LATINO	160 (95)	160 (95)	158 (95)

source: Demographics 2013 06 12.sas

Patterns of patient disposition at week 16 did not contradict efficacy of apremilast (Table 5). With roughly equal numbers of patients in each treatment arm, the percent of patients who entered early escape at week 16 was numerically higher among patients randomized to placebo than among those randomized to A20 or A30. In studies 2 and 3, and the number of patients withdrawing due to adverse events (AE) was slightly higher at week 16 among patients assigned A30 than among patients randomized to placebo or A20.

By week 24, the number of adverse events was numerically higher among patients assigned A30 than among patients randomized to placebo or A20 only in study 3 (Table 6). Withdrawal due to lack of efficacy did not appear to vary by treatment arm (Table 5 and Table 6).

Table 5. Patient Disposition, n (%), Studies 2, 3, and 4, at Week 16.

Study	Disposition Status	Pbo	A20	A30
2	Full Analysis Set	168	168	168
	Discontinue Treatment Not Early Escape	10 (6)	10 (6)	14 (8)
	ADVERSE EVENT	5 (3)	5 (3)	9 (5)
	DEATH	0 (0)	1 (1)	0 (0)
	LACK OF EFFICACY	3 (2)	2 (1)	2 (1)
	NON-COMPLIANCE WITH STUDY DRUG	0 (0)	0 (0)	1 (1)
	OTHER	0 (0)	1 (1)	0 (0)
	PROTOCOL VIOLATION	1 (1)	0 (0)	0 (0)
	WITHDRAWAL BY SUBJECT	1 (1)	1 (1)	2 (1)
	Early Escape	107 (64)	78 (46)	58 (35)
3	Full Analysis Set	159	163	162
	Discontinue Treatment Not Early Escape	11 (7)	12 (7)	13 (8)
	ADVERSE EVENT	3 (2)	4 (2)	11 (7)
	LACK OF EFFICACY	2 (1)	2 (1)	0 (0)
	LOST TO FOLLOW-UP	1 (1)	0 (0)	0 (0)
	OTHER	0 (0)	1 (1)	0 (0)
	PROTOCOL VIOLATION	0 (0)	0 (0)	1 (1)
	WITHDRAWAL BY SUBJECT	5 (3)	5 (3)	1 (1)
	Early Escape	88 (55)	59 (36)	64 (40)
4	Full Analysis Set	169	169	167
	Discontinue Treatment Not Early Escape	13 (8)	12 (7)	11 (7)
	ADVERSE EVENT	6 (4)	6 (4)	5 (3)
	LACK OF EFFICACY	2 (1)	3 (2)	2 (1)
	LOST TO FOLLOW-UP	0 (0)	0 (0)	1 (1)
	OTHER	3 (2)	0 (0)	1 (1)
	PROTOCOL VIOLATION	0 (0)	0 (0)	1 (1)
	WITHDRAWAL BY SUBJECT	2 (1)	3 (2)	1 (1)
	Early Escape	97 (57)	76 (45)	53 (32)

Source: Disposition 2013 07 09.sas

* Note – in study 3, four patients were randomized in error and did not receive study drug

Table 6. Patient Disposition, Studies 2, 3, and 4, at Week 24.

Study	Disposition Status	Pbo	A20	A30
2	Full Analysis Set	168	168	168
	Discontinue Treatment Not Early Escape	14 (8)	19 (11)	21 (13)
	ADVERSE EVENT	8 (5)	8 (5)	11 (7)
	DEATH	0 (0)	1 (1)	0 (0)
	LACK OF EFFICACY	3 (2)	2 (1)	4 (2)
	LOST TO FOLLOW-UP	0 (0)	1 (1)	0 (0)
	NON-COMPLIANCE WITH STUDY DRUG	0 (0)	1 (1)	2 (1)
	OTHER	0 (0)	1 (1)	1 (1)
	PROTOCOL VIOLATION	1 (1)	0 (0)	0 (0)
	WITHDRAWAL BY SUBJECT	2 (1)	5 (3)	3 (2)
	Early Escape	107 (64)	78 (46)	58 (35)
3	Full Analysis Set	159	163	162
	Discontinue Treatment Not Early Escape	16 (10)	19 (12)	21 (13)
	ADVERSE EVENT	3 (2)	5 (3)	12 (7)
	LACK OF EFFICACY	2 (1)	3 (2)	2 (1)
	LOST TO FOLLOW-UP	1 (1)	1 (1)	2 (1)
	NON-COMPLIANCE WITH STUDY DRUG	1 (1)	0 (0)	0 (0)
	OTHER	1 (1)	2 (1)	0 (0)
	PROTOCOL VIOLATION	0 (0)	0 (0)	1 (1)
	WITHDRAWAL BY SUBJECT	8 (5)	8 (5)	4 (2)
	Early Escape	88 (55)	59 (36)	64 (40)
4	Full Analysis Set	169	169	167
	Discontinue Treatment Not Early Escape	13 (8)	12 (7)	11 (7)
	ADVERSE EVENT	10 (6)	12 (7)	11 (7)
	LACK OF EFFICACY	5 (3)	5 (3)	6 (4)
	LOST TO FOLLOW-UP	1 (1)	0 (0)	3 (2)
	OTHER	3 (2)	1 (1)	2 (1)
	PROTOCOL VIOLATION	0 (0)	0 (0)	1 (1)
	WITHDRAWAL BY SUBJECT	3 (2)	4 (2)	1 (1)
	Early Escape	97 (57)	76 (45)	53 (32)

Source: Disposition 2013 07 09.sas

* Note – in study 3, four patients were randomized in error and did not receive study drug

3.2.4 Results and Conclusions

3.2.4.1 Primary Endpoint

3.2.4.1.1 ACR 20 at Week 16

For all three studies, the primary analysis showed differences between placebo and A30 which were significant at the 0.05 level (Table 7). Although not shown, differences between placebo and A20 and between placebo and the apremilast arms combined (A20 plus A30) were also statistically significant. While the average difference from placebo in percent response was 13% for A20 and 18% for A30, the proportion of patients who met the ACR20 response criteria in the active treatment groups was less than half of the patients randomized to the groups (i.e., 28 to 37% in those patients taking A20, and 32 to 41% in patients taking A30).

Table 7. ACR20 at Week 16. Percent Responders, A30 versus Placebo, Primary Analysis

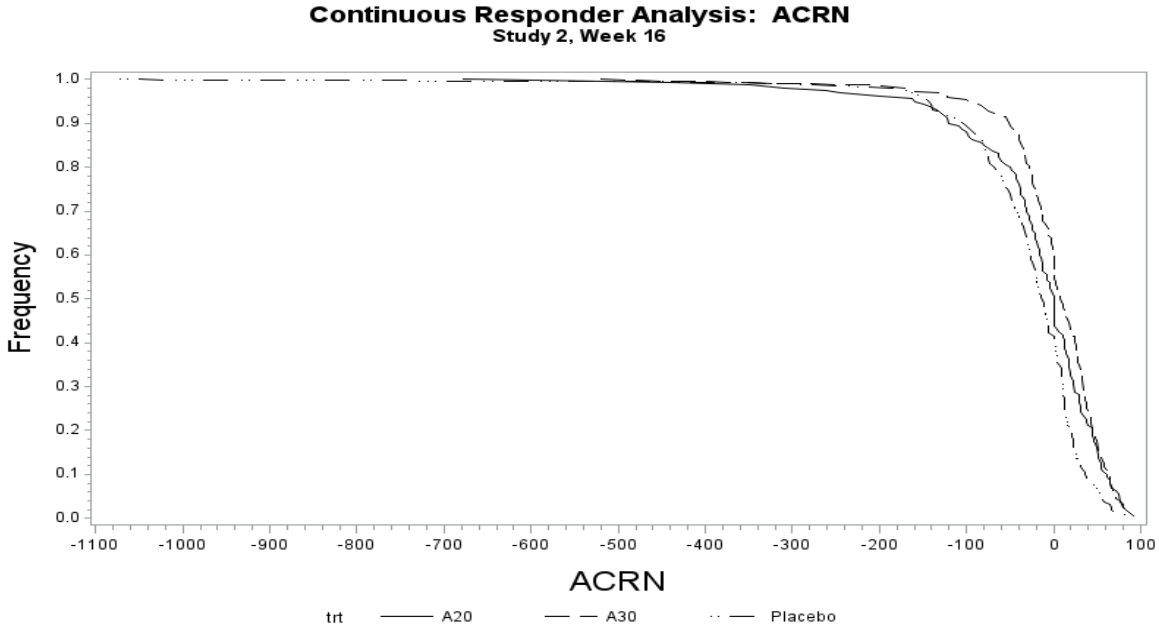
Study	Treatment			Treatment Difference %			P-Value	
	P	A20	A30	A20-P	A30-P	A30-A20	A30-P	A30-A20
2	19 (32/168)	30 (51/168)	38 (64/168)	11	19	8	0.0001	0.1456
3	19 (30/159)	37 (61/163)	32 (52/162)	19	13	-5	0.006	0.3132
4	18 (31/169)	28 (48/169)	41 (68/167)	10	22	12	<.0001	0.0172

Source: mainline.sas
Patients who discontinued treatment were considered non-responders

Kolmogorov-Smirnov tests on showed nominally significant differences between A30 and P for all three studies (Figure 1, Figure 2, and Figure 3), with a nominally significant difference between A30 and A20 only in study 4. The cumulative responder functions in Figure 1, Figure 2, and Figure 3, which are 1 – the cumulative distribution functions, suggest that use of apremilast improves response in a majority of patients, but does not consistently prevent extreme deteriorations or provide extreme benefits. Further exploratory analyses using t-tests showed significant differences in mean ACRn between A30 and placebo for all studies and between A20 and placebo in study 3 but not in studies 2 and 4.

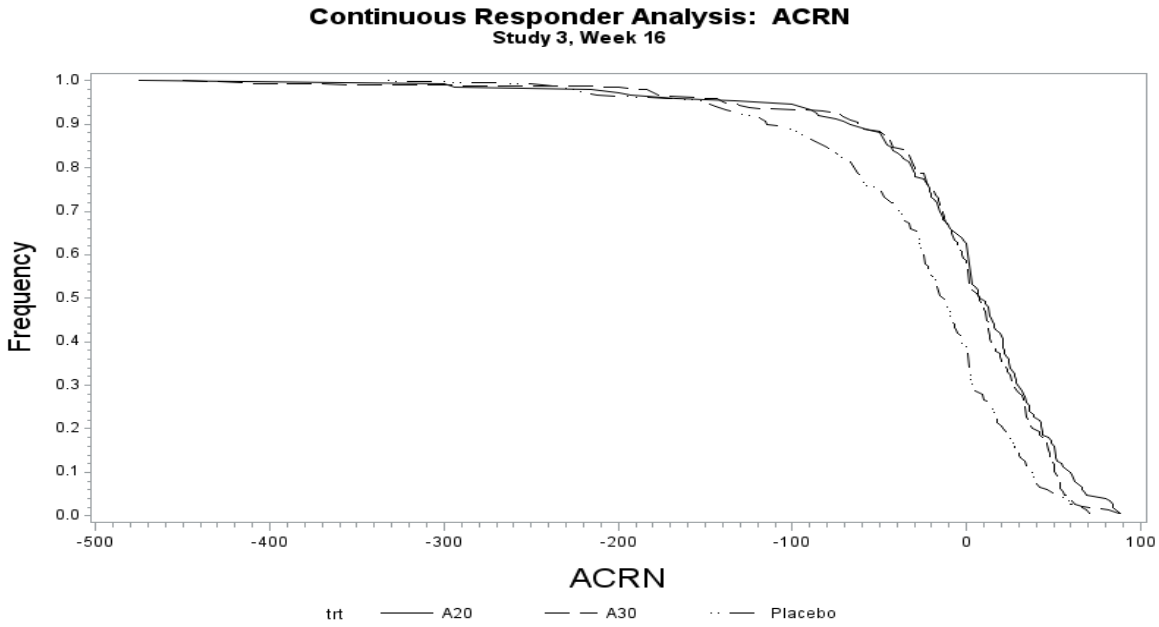
A spot quality check on some of the lower values suggests that the ACRn were calculated correctly. For example, the lowest value in Figure 1 was -1075 for a placebo patient who had 4 swollen joints at baseline and 47 swollen joints at week 16. Similarly, an A30 patient with an ACRn value of -520, had 5 tender joints at baseline and 31 at week 16.

Figure 1. ACRn. Continuous Responder Analysis, Study 2, Week 16.



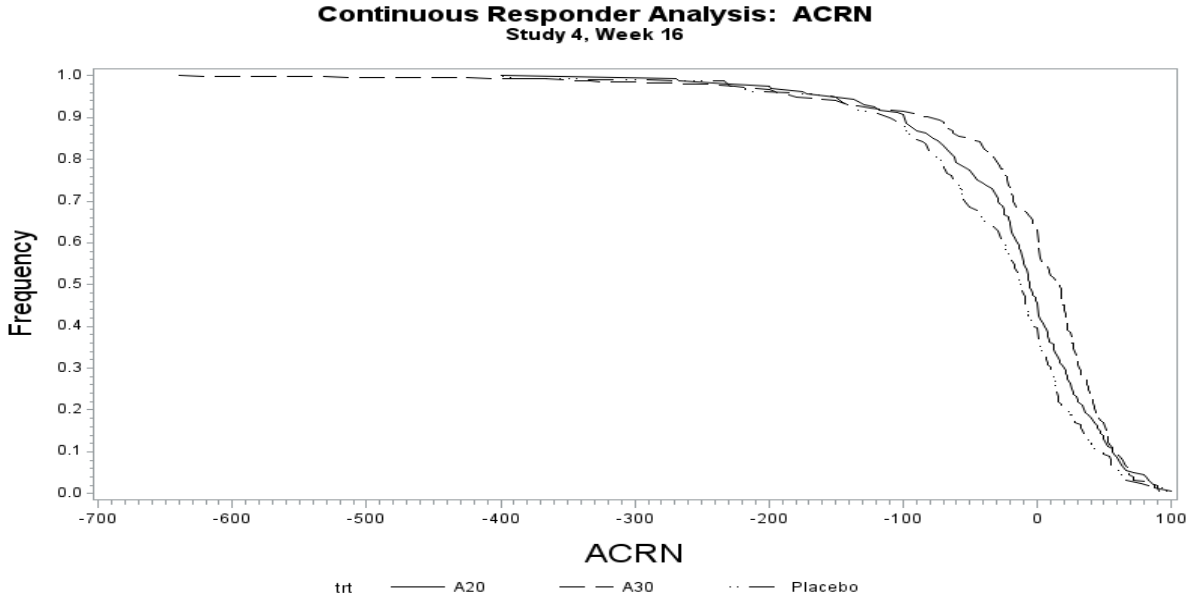
Kolmogorov-Smirnov Test Apremilast 30 mg vs Placebo: P-Value = <.0001
Kolmogorov-Smirnov Test Apremilast 30 mg vs 20 mg: P-Value = 0.1619

Figure 2. ACRn. Continuous Responder Analysis, Study 3, Week 16.



Kolmogorov-Smirnov Test Apremilast 30 mg vs Placebo: P-Value = 0.0008
Kolmogorov-Smirnov Test Apremilast 30 mg vs 20 mg: P-Value = 0.9118
Source: mainline.sas

Figure 3. ACRn. Continuous Responder Analysis, Study 4, Week 16.



Kolmogorov-Smirnov Test Apremilast 30 mg vs Placebo: P-Value = <.0001
Kolmogorov-Smirnov Test Apremilast 30 mg vs 20 mg: P-Value = 0.0097
Source: mainline.sas

3.2.4.1.2 Difference between A20 and A30

The sponsor proposed approval for A30 rather than A20. However, the lower dose, A20, also provided statistically significant improvements over placebo (Table 8), and whether there was even a numerical advantage of A30 over A20 for ACR response rate at week 16 is debatable. In particular, the difference between A20 and A30 was statistically significant in only one of three studies and the response to A30 was numerically smaller than A20 in study 3 (Table 8).

Table 8. ACR20 at Week 16. Percent Responders, A20 versus Placebo, Primary Analysis

Study	**Treatment**			**Treatment Difference %**			**P-Value**	
	P	**A20**	**A30**	**A20-P**	**A30-P**	**A30-A20**	**A20-P**	**A30-A20**
2	19 (32/168)	30 (51/168)	38 (64/168)	11	19	8	0.0166	0.1456
3	19 (30/159)	37 (61/163)	32 (52/162)	19	13	-5	0.0002	0.3132
4	18 (31/169)	28 (48/169)	41 (68/167)	10	22	12	0.0295	0.0172

Source: mainline.sas
Patients who discontinued treatment were considered non-responders

In summary, the Phase 3 trials demonstrated statistically significant differences between A30 and P for primary response variable ACR20. Further analyses suggested that use of apremilast improves ACRn in a majority of patients, but does not consistently prevent extreme deteriorations or provide extreme benefits. Evidence for additional efficacy benefit of A30 over A20 is inconsistent.

3.2.4.1.3 ACR Components

Compared to placebo, percent improvement is observed in patients treated with A30 and A20 for all ACR components at week 16 which supports the primary endpoint of ACR20 (Table 9).

The treatment effect of apremilast was not consistently greater for the higher dose than for the the lower apremilast dose:

1. In studies 2 and 4, except for CRP in study 4, improvements compared to placebo were numerically larger for the higher apremilast dose than for the lower apremilast dose.
2. In study 3, improvements compared to placebo of the higher apremilast dose were numerically smaller than those of the lower dose for CRP, patient global assessment, physician global assessment, tender joint count.

As discussed in section 3.2.2, missing data due to treatment discontinuation were handled using last observation carried forward. Since only 6 – 8% of patients discontinued treatment prior to week 16 and had missing data, the results should not be affected by the imputation strategy used. This was confirmed when baseline observation carried forward was used to impute missing data, assuming that patients' score at week 16 reverted back to its baseline measure (or bad score) when they discontinued treatment; the estimated effects were similar. All patients who escaped at week 16 have observed data for these components and should not be affected by missing data.

Table 9. ACR Components. Percent Change, Week 16. Negative Values Imply Improvement

Study	Variable	Percent Change (N)			Percent Difference		
		P	A20	A30	A30-P	A20-P	A30-20
2	CRP	-1 (166)	-8 (167)	-12 (167)	-11	-6	-5
	HAQ-DI	-8 (165)	-14 (163)	-20 (159)	-12	-7	-6
	Pain	-10 (165)	-17 (163)	-26 (159)	-16	-7	-9
	Patient Global	-9 (165)	-18 (163)	-20 (159)	-11	-9	-2
	Physician Global	-13 (158)	-34 (160)	-42 (159)	-29	-21	-8
	Swollen Joint Count	-17 (166)	-39 (164)	-50 (164)	-33	-22	-11
	Tender Joint Count	-9 (166)	-24 (164)	-43 (164)	-34	-15	-19
3	CRP	5 (157)	-14 (162)	-8 (161)	-13	-19	6
	HAQ-DI	-7 (153)	-13 (159)	-20 (154)	-13	-5	-8
	Pain	-5 (151)	-22 (157)	-24 (152)	-19	-17	-2
	Patient Global	-6 (151)	-17 (157)	-16 (152)	-10	-12	1
	Physician Global	-15 (150)	-42 (156)	-36 (146)	-21	-27	6
	Swollen Joint Count	-33 (154)	-50 (158)	-54 (155)	-21	-17	-4
	Tender Joint Count	-9 (154)	-36 (158)	-33 (155)	-25	-27	3

Table 9 (continued)

Study	Variable	Percent Change (N)			Percent Difference		
		P	A20	A30	A30-P	A20-P	A30-20
4	CRP	-6 (168)	-17 (168)	-3 (165)	3	-11	14
	HAQ-DI	-7 (163)	-11 (163)	-20 (160)	-13	-4	-9
	Pain	-3 (164)	-10 (163)	-24 (161)	-21	-7	-14
	Patient Global	-3 (164)	-10 (163)	-17 (161)	-14	-7	-7
	Physician Global	-13 (159)	-23 (156)	-40 (156)	-27	-9	-18
	Swollen Joint Count	-20 (165)	-35 (164)	-50 (161)	-30	-15	-15
	Tender Joint Count	-8 (165)	-29 (164)	-43 (161)	-35	-22	-14

source: mainline.sas
Missing data due to treatment discontinuation were imputed using last observation carried forward. Only 6 – 8% patients have missing data. The results from baseline observation carried forward were consistent.

Describing benefit in terms of percent improvement from baseline, as provided in Table 9, is problematic for two reasons. First, absolute change is more appropriate than percent change when evaluating benefit against risk. For example, benefit is greater in a patient whose number of tender joints decreases from 12 to 0 than in a patient whose number of tender joints decreases from 1 to 0. However, in both patients, the percent change is the same value, 100%, obscuring the higher benefit in the patient with greater reduction in tender joint count. Second, percent improvement itself is undefined when baseline is zero. For example, 16 patients from study 2 recorded zero HAQ-DI at baseline; the sponsor's analysis simply excluded their data, ignoring any HAQ-DI deteriorations such patients may have experienced.

Measuring absolute rather than percent change in ACR components provides quantitative estimates of improvement (Table 10). For example, compared to placebo, in study 2, A30 reduced the mean number of tender joints in each patient by 3.5, a straightforward metric for improvement which can be compared to risks. Also note that, in contrast to the percentage changes of Table 9the absolute measures in Table 10 show, for study 3, that the higher apremilast dose is associated with numerically larger improvements in CRP than the lower dose, but is also associated with numerically smaller improvements in swollen joint count and pain.

Compared to placebo, A30 improved nearly all ACR components. Significant differences between A30 and placebo were seen for all components (Table 10) except for CRP in studies 3 and 4 and pain in study 3, for which improvements were numerical but were not statistically significant.

Table 10. ACR Components. Absolute Mean Change, Week 16. Negative Values Imply Improvement

Study	Endpoint	Treatment (N)			Treatment Difference			P-Value	
		P	A20	A30	A20-P	A30-P	A30-20	A30-P	A30-20
2	CRP	0.1 (166)	-0.1 (167)	-0.1 (167)	-0.2	-0.3	0	0.0231	0.8238
	HAQ-DI	-0.1 (165)	-0.2 (163)	-0.2 (159)	-0.1	-0.2	0	0.0017	0.3634
	Pain	-5.7 (165)	-11.5 (163)	-13.5 (159)	-5.8	-7.9	-2.1	0.0023	0.4205
	Patient Global	-3 (165)	-9.1 (163)	-10.1 (159)	-6.1	-7.1	-1	0.0092	0.713
	Physician Global	-7.6 (158)	-16.3 (160)	-18.1 (159)	-8.7	-10.5	-1.8	<.0001	0.4882
	Swollen Joint Count	-1.7 (166)	-4.2 (164)	-5.2 (164)	-2.6	-3.5	-0.9	<.0001	0.2869
	Tender Joint Count	-1.8 (166)	-5.4 (164)	-7.2 (164)	-3.6	-5.4	-1.8	<.0001	0.1961

Table 10 (continued)

Study	Endpoint	Treatment (N)			Treatment Difference			P-Value	
		P	A20	A30	A20-P	A30-P	A30-20	A30-P	A30-20
3	CRP	-0.1 (157)	-0.1 (162)	-0.2 (161)	-0.1	-0.1	-0.1	0.3938	0.6341
	HAQ-DI	-0.1 (153)	-0.2 (159)	-0.2 (154)	-0.1	-0.1	0	0.0042	0.4507
	Pain	-7 (151)	-12.5 (157)	-11.9 (152)	-5.5	-4.9	0.7	0.0648	0.8014
	Patient Global	-4.6 (151)	-8.9 (157)	-8.8 (152)	-4.3	-4.2	0.1	0.1065	0.9802
	Physician Global	-8.8 (150)	-18 (156)	-16.8 (146)	-9.2	-8	1.2	0.0014	0.6201
	Swollen Joint Count	-2.4 (154)	-4.3 (158)	-3.9 (155)	-1.9	-1.5	0.4	0.0154	0.5732
	Tender Joint Count	-1.1 (154)	-5.6 (158)	-4.1 (155)	-4.4	-3	1.5	0.0122	0.2178
4	CRP	-0.1 (168)	-0.3 (168)	-0.1 (165)	-0.2	0	0.2	0.7996	0.1726
	HAQ-DI	-0.1 (163)	-0.1 (163)	-0.2 (160)	-0.1	-0.1	-0.1	0.0073	0.1952
	Pain	-4.9 (164)	-8.6 (163)	-12.7 (161)	-3.6	-7.8	-4.2	0.0021	0.0996
	Patient Global	-3.3 (164)	-5.7 (163)	-8.8 (161)	-2.4	-5.5	-3.1	0.0382	0.2429
	Physician Global	-7.8 (159)	-13.6 (156)	-19.3 (156)	-5.7	-11.5	-5.8	<.0001	0.0165
	Swollen Joint Count	-1.3 (165)	-2.3 (164)	-3.5 (161)	-1	-2.2	-1.2	0.01	0.1568
	Tender Joint Count	-0.8 (165)	-3.7 (164)	-6.1 (161)	-2.9	-5.3	-2.4	<.0001	0.029

source: mainline.sas

Missing data due to treatment discontinuation were imputed using last observation carried forward. Only 6 – 8% patients have missing data. The results from baseline observation carried forward were consistent.

3.2.4.2 Key Secondary Endpoint: ΔHAQ-DI at Week 16

Statistically significant differences between A30 and placebo for change from baseline HAQ-DI at week 16 were seen in all studies with an average reduction by A30 of 0.14 (Table 11) and an average reduction by A20 of 0.09. Numerical differences between A30 and A20 favored A30 in all three studies but the differences were not statistically significant.

Differences between A20 and placebo were significant in studies 2 (p=0.025) and 3 (p=0.032) but not in study 4 (p=0.162).

Like the component scores, missing data due to treatment discontinuation were handled using last observation carried forward. Since only 6 – 8% of patients discontinued treatment prior to week 16 and had missing data, the results should not be affected by the imputation strategy used. This was confirmed when baseline observation carried forward was used to impute missing data assuming that patients' score at week 16 reverted back to its baseline measure (or bad score) when they discontinued treatment; the estimated effects were similar All patients who escaped at week 16 have observed data for this endpoint and should not be affected by missing data.

Table 11. HAQ-DI. Mean Change from Baseline, Week 16

Study	**ΔHAQ-DI**			**Treatment Difference**			**P-Value**	
	P	**A20**	**A30**	**A20-P**	**A30-P**	**A30-A20**	**A30-P**	**A30-A20**
2	-0.09 (165)	-0.2 (163)	-0.24 (159)	-0.11	-0.16	-0.05	0.0017	0.3634
3	-0.05 (153)	-0.16 (159)	-0.19 (154)	-0.10	-0.14	-0.04	0.0042	0.4507
4	-0.07 (163)	-0.13 (163)	-0.19 (160)	-0.07	-0.13	-0.06	0.0073	0.1952

source: mainline.sas

Missing data due to treatment discontinuation were imputed using last observation carried forward. Only 6 – 8% patients have missing data. The results from baseline observation carried forward were consistent.

Differences between A30 and placebo for HAQ-DI response (ΔHAQ-DI ≤ -0.3) were statistically significant in only two of the Phase 3 studies (Table 12). The average difference in percent HAQ-DI response between A30 and placebo was 11%. Numerical differences favored A30 over A20 in all three studies, but differences between the two doses were not statistically significant (Table 12).

Differences between A20 and placebo were not statistically significant in any of the three studies.

Table 12. HAQ-DI Improvement ≥ 0.3. Percent Response, Week 16

Study	HAQ-DI Percent Response			Treatment Difference			P-Value	
	P	A20	A30	A20-P	A30-P	A30-A20	A30-P	A30-A20
2	27 (45/168)	33 (55/168)	38 (64/168)	6	12	5	0.0249	0.2982
3	25 (39/159)	32 (52/163)	38 (61/162)	7	13	6	0.011	0.2766
4	28 (45/163)	33 (54/163)	35 (56/160)	5	7	1	0.1577	0.7821

source: mainline.sas
Patients who discontinued treatment were considered non-responders

3.2.4.3 Other Secondary Endpoints (based on pre-specified hierarchy)

To further support the efficacy of apremilast, the applicant also examined week 24 ACR20 response and HAQ-DI, as well as other endpoints including SF36 component and domain scores, Psoriasis Area and Severity Index (PASI) 75 scores, Maastricht Ankylosing Spondylitis Entheses Score (MASES), ACR50, ACR70, and DAS28-CRP.

The week 24 data is to a certain extent problematic because a large proportion of the patients in each study had by week 24 discontinued initially assigned treatment, either by meeting the escape criteria at week 16, or because of adverse events, lack of efficacy or patient withdrawal (Table 13). For ACR responses at week 24, we will present a summary of responses by patient status, according to whether patients are responders, non-responders due to escape or dropout, or non-responders according to data recorded at week 24. For continuous outcomes like HAQ-DI, because of the challenge of patients escaping or discontinuing treatment, we will not attempt to analyze week 24 data. Instead, we will provide summaries using continuous responder plots to describe the improvement in HAQ-DI as well as responder analyses using a cut-off of 0.3.

Table 13. Patient Disposition, Studies 2, 3, and 4, at Week 24. Number (percent)

Study	Disposition Status	Pbo	A20	A30
2	Full Analysis Set	168	168	168
	Discontinue Treatment Not Early Escape	14 (8)	19 (11)	21 (13)
	Early Escape	107 (64)	78 (46)	58 (35)
3	Full Analysis Set	159	163	162
	Discontinue Treatment Not Early Escape	16 (10)	19 (12)	21 (13)
	Early Escape	88 (55)	59 (36)	64 (40)
4	Full Analysis Set	169	169	167
	Discontinue Treatment Not Early Escape	13 (8)	12 (7)	11 (7)
	Early Escape	97 (57)	76 (45)	53 (32)

Source: Disposition 2013 07 09.sas
* Note – in study 3, four patients were randomized in error and did not receive study drug

3.2.4.3.1 Percent ACR 20 Response, Week 24

As noted earlier, approximately 55% to 66% of placebo patients entered escape and about 32% to 46% in the active group entered escape at week 16. These individuals were considered ACR 20 non-responders in the week 16 and week 24 analyses. Patients who discontinued treatment were also considered non-responders. Compared to placebo, patients treated with A30 achieved a higher ACR 20 response at week 24 in all three studies (Table 14).

Table 14. ACR 20 Percent Response by Escape and Treatment Discontinuation Status, Week 24

Study	Week		P	A20	A30	Trt Diff (p-value)	
2		N	168	168	168	A20 - P	A30 – P
	16	Responder	32 (19%)	51 (30%)	64 (38%)	11% (0.0166)	19% (0.0001)
	24	Responder	22 (13%)	43 (26%)	59 (35%)	13% (0.0038)	22% <.0001
		Non-responder	25 (15%)	28 (17%)	29 (17%)		
		NR – Escape/drop	121 (72%)	97 (58%)	80 (48%)		
3		N	159	163	162		
	16	Responder	30 (19%)	61 (37%)	52 (32%)	19% (0.0002)	13% (0.006)
	24	Responder	25 (16%)	51 (31%)	40 (25%)	16% (0.0009)	9% (0.0394)
		Non-responder	30 (19%)	34 (21%)	37 (23%)		
		NR – Escape/drop	104 (65%)	78 (48%)	85 (52%)		
4		N	169	169	167		
	16	Responder	31 (18%)	48 (28%)	68 (41%)	10% 0.0295	22% 0.0173
	24	Responder	26 (15%)	46 (27%)	52 (31%)	12 (0.0110)	16 (0.0007)
		Non-responder	24 (14%)	26 (15%)	39 (23%)		
		NR – Escape/drop	119 (70%)	97 (57%)	76 (46%)		

Responder = number (percent) of patients who achieved ACR20 at Week 16 (Week 24)
Non-responder = number (percent) of patients who are still taking their assigned treatment at week 24 and who did not achieve ACR20 response status.
NR-escape/drop = number (percent) of patients who have non-responder status because of escape or treatment discontinuation
For inference, patients who entered escape or discontinued treatment prior are considered non-responders.

The analysis provided in Table 14 shows the proportion of responders at week 24 given that they did not escape or withdraw from treatment. A metric which may be more relevant to sustained response, and which may be more readily interpretable, is the proportion of individuals who had an ACR20 response both at weeks 16 and 24. Compared to placebo, percent ACR 20 responders was higher for A30 and for A20 in all three studies (Table 15).

Table 15. Percent ACR 20 Responders at both Week 16 and Week 24

Study	**Percent Response**			**Trt Diff (p-value)**	
	P	**A20**	**A30**	**A20-P**	**A30-P**
2	11 (18/168)	20 (34/168)	29 (49/168)	9 (0.0162)	19 (<.0001)
3	10 (16/159)	26 (42/163)	20 (33/162)	16 (0.0002)	10 (0.0087)
4	12 (21/169)	22 (37/169)	26 (43/167)	9 (0.0213)	13 (0.0020)

source: mainline.sas

3.2.4.3.2 HAQ-DI Change from Baseline, Week 24

Because many placebo patients escaped from their initially treatment assignment prior to week 24, it is impossible to predict what their HAQ-DI score would have been at week 24 had they remained on placebo. Nevertheless, to help visualize the data for each study, we provide continuous responder profiles (Figure 4, Figure 5, and Figure 6). On each each graph, the x-axis provides a potential cutoff for a responder analysis, the y-axis represents proportion of responders at that cutoff, with each curve representing a particular dose. Visually, Figure 4, Figure 5, and Figure 6 show that, regardless of the cutoff used, response rate is numerically higher in A30 and A20 than in placebo.

For traditional HAQ-DI response (ΔHAQ-DI ≤ 0.3) at week 24, we tabulate the data according to response and withdrawal status (Table 16).

Figure 4. HAQ-DI, Continuous Responder Profile, Study 2, Week 24

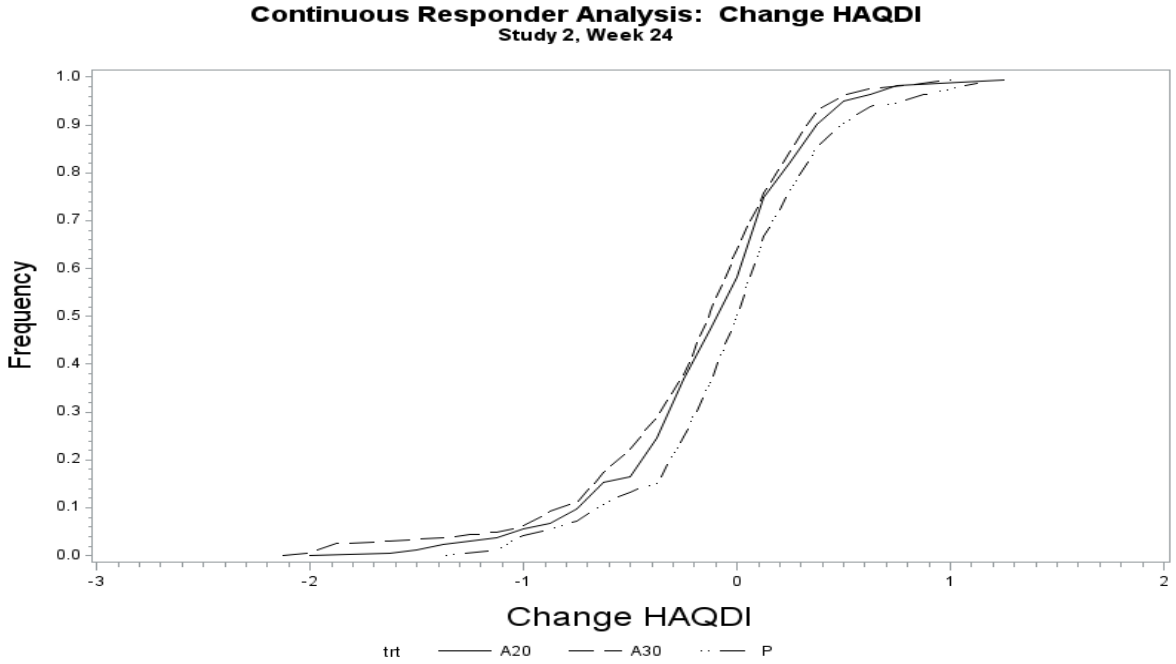


Figure 5. HAQ-DI, Continuous Responder Profile, Study 3, Week 24

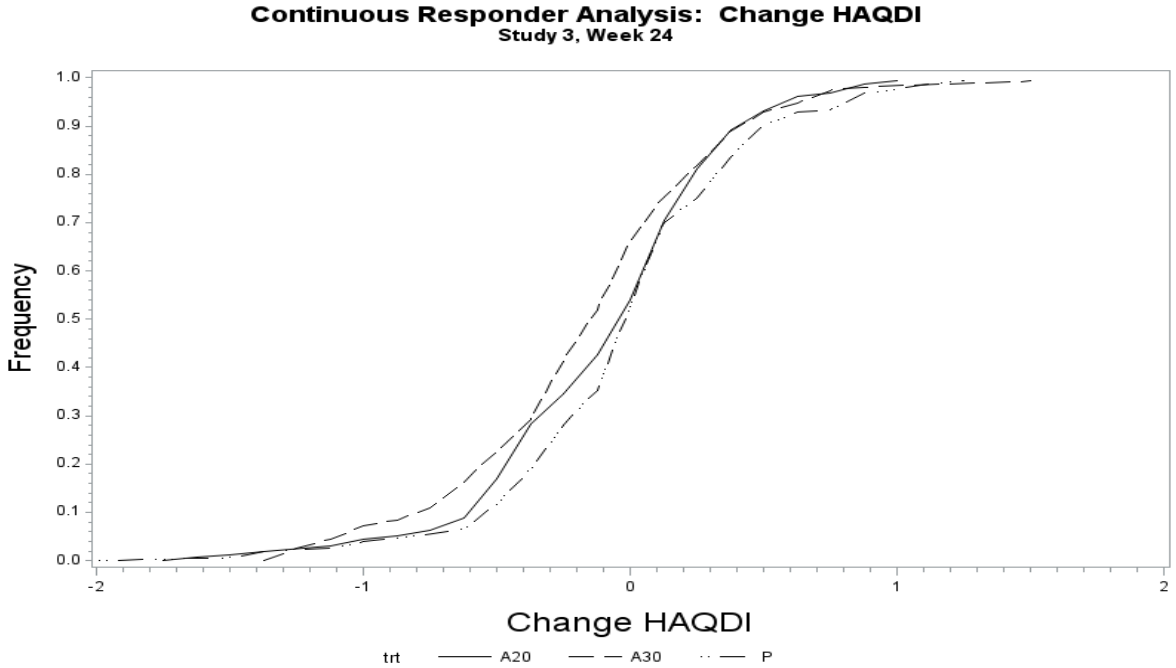
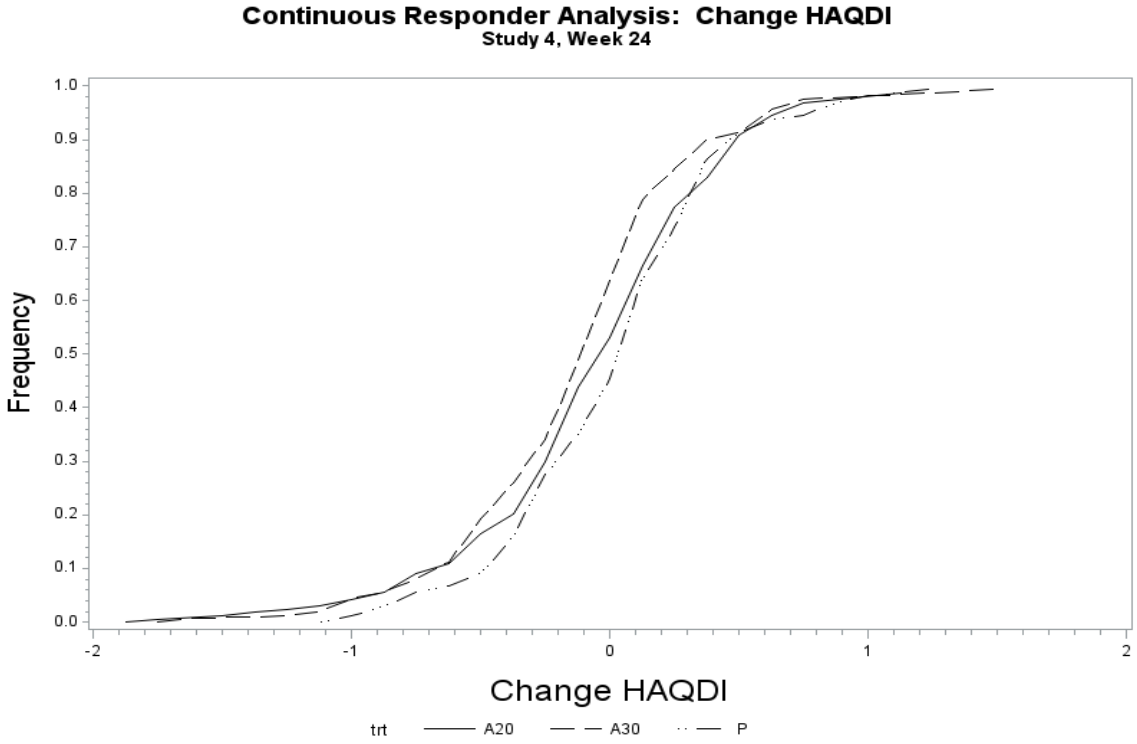


Figure 6. HAQ-DI, Continuous Responder Profile, Study 4, Week 24



source: mainline.sas

Table 16. HAQ-DI Response, by Treatment Discontinuation Status, Week 24

Study	Week		P	A20	A30	Trt Diff	
2		N	168	168	168	A20 - P	A30 – P
	16	Responder	45 (27%)	55 (33%)	64 (38%)	6% (0.2360)	12% (0.0249)
	24[1]	Responder	17 (10%)	39 (23%)	48 (29%)	13%	18%
		R – Escape	24 (14%)	21 (13%)	13 (8%)	-2%	-7%
		Non-responder	29 (17%)	31 (18%)	40 (24%)	1%	7%
		NR – Escape/drop	98 (58%)	77 (46%)	67 (40%)	-13%	-18%
3		N	159	163	162		
	16	Responder	39 (25%)	52 (32%)	61 (38%)	7% (0.1400)	13% (0.0110)
	24	Responder	24 (15%)	40 (25%)	39 (24%)	9%	9%
		R – Escape	19 (12%)	15 (9%)	25 (15%)	-3%	3%
		Non-responder	31 (19%)	45 (28%)	38 (23%)	8%	4%
		NR – Escape/drop	85 (53%)	63 (39%)	60 (37%)	-15%	-16%
4		N	169	169	167		
	16	Responder	45 (28%)	54 (33%)	56 (35%)	5% (0.2904)	7% (0.7821)
	24	Responder	19 (11%)	29 (17%)	38 (23%)	6%	12%
		R – Escape	26 (15%)	20 (12%)	17 (10%)	-4%	-5%
		Non-responder	31 (18%)	42 (25%)	52 (31%)	7%	13%
		NR – Escape/drop	93 (55%)	78 (46%)	60 (36%)	-9%	-19%

source: mainline.sas

1. Week 24 Responder or Non-responder: HAQ responder or non-responder with no early escape or treatment discontinuation, R- or NR – Escape/drop: HAQ responder or non-responder with early escape or treatment discontinuation

3.2.4.3.3 SF-36 Change from Baseline, Week 16

The sponsor examined SF-36 domain and component scores at week 16. Mean change from baseline among patients randomized to A30 differed significantly from placebo in all studies for physical function domain and physical component score (Table 17). The average difference between A30 and placebo was 2.3.

The claim for improved physical function is reinforced by nominally significant improvements associated with A30 compared to placebo of physical component score, physical function, role physical, and bodily pain component score, physical component score in all three studies and for general health in one of three studies (Table 17).

Statistically significant differences between A30 and placebo for the mental component score were seen only in study 3 (Table 18). For study 3 domains of the mental component score, statistically significant differences were seen for mental health and vitality, but not for social functioning or role emotional.

Table 17. SF-36 Physical. Mean Change from Baseline, Week 16

Metric	Study	Change from Baseline			Treatment Difference			P-Value
		P	A20	A30	A20-P	A30-P	A30-20	A30 – P
PCS	2	2.4 (168)	3.5 (168)	4.6 (168)	1.1	2.2	1.1	0.0097
	3	2 (159)	3.2 (163)	3.7 (162)	1.3	1.7	0.5	0.0335
	4	1.3 (169)	3.2 (169)	3.4 (167)	2	2.1	0.1	0.006
PF	2	1.8 (168)	3.5 (168)	4.2 (168)	1.7	2.4	0.7	0.0056
	3	0.8 (159)	2.2 (163)	2.9 (162)	1.4	2.1	0.7	0.0237
	4	1.1 (169)	2.3 (169)	3.5 (167)	1.1	2.3	1.2	0.0053
RP	2	2 (168)	2.8 (168)	4.2 (168)	0.9	2.2	1.3	0.0218
	3	1.4 (159)	1.3 (163)	3.5 (162)	-0.1	2.1	2.2	0.0247
	4	0.8 (169)	2.9 (169)	3.2 (167)	2.2	2.4	0.3	0.0049
BP	2	2 (168)	3.9 (168)	4.3 (168)	1.8	2.2	0.4	0.0083
	3	2 (159)	3.5 (163)	3.8 (162)	1.5	1.8	0.3	0.0337
	4	1.1 (169)	2.7 (169)	3.5 (167)	1.6	2.4	0.8	0.0027
GH	2	1.5 (168)	1.7 (168)	2.5 (168)	0.2	1	0.8	0.2435
	3	1.1 (159)	2.5 (163)	2.7 (162)	1.4	1.6	0.2	0.0473
	4	1 (169)	1.9 (169)	1.6 (167)	0.9	0.6	-0.3	0.4024

source: mainline.sas

PCS Physical Component Score, PF Physical Function, RP Role Physical, BP Bodily Pain, GH General Health

Table 18. SF-36 Mental. Mean Change from Baseline, Week 16

Metric	Study	Change from Baseline			Treatment Difference			P-Value
		P	A20	A30	A20-P	A30-P	A30-20	A30 - P
MCS	2	0.1 (168)	0.4 (168)	0.7 (168)	0.3	0.6	0.3	0.4932
	3	-1 (159)	-1 (163)	0.9 (162)	0	1.9	1.9	0.0326
	4	0.2 (169)	-0.3 (169)	1.2 (167)	-0.5	1	1.5	0.1989
MH	2	1.1 (168)	1.5 (168)	1.9 (168)	0.4	0.8	0.4	0.436
	3	-0.8 (159)	-0.3 (163)	1.5 (162)	0.5	2.3	1.8	0.0369
	4	0.4 (169)	0.4 (169)	2.3 (167)	0	1.9	1.8	0.0396
VT	2	2 (168)	1.7 (168)	3.5 (168)	-0.3	1.5	1.8	0.1451
	3	0.7 (159)	1.2 (163)	3 (162)	0.5	2.3	1.8	0.0136
	4	0.8 (169)	1.9 (169)	2.6 (167)	1.1	1.8	0.7	0.0378
SF	2	0.7 (168)	0.5 (168)	0.3 (168)	-0.2	-0.4	-0.2	0.4227
	3	-0.2 (159)	-0.2 (163)	-0.4 (162)	0.1	-0.1	-0.2	0.7999
	4	0.5 (169)	0.9 (169)	1.2 (167)	0.3	0.7	0.4	0.2281
RE	2	-0.4 (168)	1.6 (168)	1.6 (168)	2	2	0.0	0.0647
	3	-0.3 (159)	-0.2 (163)	2.7 (162)	0	2.9	2.9	0.0071
	4	0.2 (169)	0 (169)	1.7 (167)	-0.2	1.5	1.7	0.1277

source: mainline.sas

MCS Mental Component Score, MH Mental Health , VT Vitality, SF Social Functioning, RE Role Emotional

3.2.4.3.4 PASI 75 Response, Week 16

Analysis of PASI 75 among patients with psoriasis involved body surface area (BSA) ≥ 3% was preplanned only in study 4; therefore indications of statistical significance in studies 2 and 3 are only nominal, with true statistical significance only in study 4 (Table 19). The average difference between A30 and placebo for PASI 75 response at week 16 was 17%.

Table 19. PASI75. Percent Response Among Patients with Psoriasis BSA ≥ 3%, Week 16

Study	PASI 75 Response (%)			Treatment Difference		P-Value A30-P
	P	A20	A30	A20-P	A30-P	
2	4 (3/68)	21 (16/77)	22 (18/82)	16	18	0.0022
3	2.7 (2/74)	19 (15/80)	22 (17/77)	16	20	0.0002
4	7.9 (7/89)	21 (19/91)	22 (20/90)	13	15	0.0062

source: mainline.sas

3.2.4.3.5 MASES Reduction from Baseline, Week 16

For reduction from baseline of Maastricht Ankylosing Spondylitis Entheses Score (MASES) at week 16 among patients with pre-existing enthesitis, differences between A30 and placebo were not statistically significant in any of the three Phase 3 studies (Table 20). The average difference between A30 and placebo for reduction from baseline MASES was 0.32.

Table 20. MASES. Mean Change from Baseline Among Patients with Pre-Existing Enthesitis, Week 16

Study	ΔMASES			Treatment Difference			P-Value A30-P
	P	A20	A30	A20-P	A30-P	A30-A20	
2	-0.9 (95)	-1.5 (100)	-1.3 (108)	-0.5	-0.4	0.2	0.36
3	-1.0 (100)	-0.9 (105)	-1.4 (97)	0.1	-0.4	-0.4	0.35
4	-0.7 (106)	-0.7 (93)	-1.0 (107)	0.1	-0.2	-0.3	0.53

source: mainline.sas

Because the reductions from baseline MASES were not statistically significant in any of the three studies, any p-values from analyses of secondary variables lower in the analysis hierarchy (i.e. below item 7, Section 6) underestimate true Type I error. Analyses of such variables are only exploratory, and any indications of statistical significance are only nominal.

3.2.4.4 Exploratory Analyses: Other Claims

Endpoints analyzed in this section are included on the proposed product label. However, none of the differences between apremilast and placebo are statistically significant in any formal sense because, in the analysis hierarchy, they are below the MASES endpoint which failed for statistical significance. Calculated p-values provided therefore underestimate true Type I error, and any indications of statistical significance are only nominal.

3.2.4.4.1 ACR 50 Response

For ACR 50 response, differences between A30 and placebo were nominally significant at week 16 in study 2 but not in studies 3 and 4 (Table 21). The average difference between A30 and placebo was 8% at week 16. Because ACR50 failed for significance at week 16, results for this endpoint were not reviewed at week 24.

Table 21. ACR 50 Percent Response

Week	Study	ACR50 Response (%)			Treatment Difference		P-Value A30-P
		P	A20	A30	A20-P	A30-P	
16	2	6 (10/168)	16 (26/168)	16 (27/168)	10	10	0.0027
	3	5 (8/159)	15 (24/163)	11 (17/162)	10	6	0.0589
	4	8 (14/169)	12 (21/169)	15 (25/167)	4	7	0.0520

source: mainline.sas

3.2.4.4.2 ACR 70 Response

For ACR70 response, differences between A30 and placebo were not nominally significant at week 16 in any of the three studies, and were nominally significant at week 24 only in study 2 (Table 22). The average difference between A30 and placebo was 2% at week Because ACR70 failed for significance at week 16, results for this endpoint were not reviewed at week 24.

Table 22. ACR70 Percent Response

Week	Study	ACR70 Response (%)			Treatment Difference		P-Value A30-P
		P	A20	A30	A20-P	A30-P	
16	2	1 (2/168)	6 (10/168)	4 (7/168)	5	3	0.0792
	3	1 (1/159)	4 (6/163)	1 (2/162)	3	1	0.5620
	4	2.4 (4/169)	5 (8/169)	4 (6/167)	2	1	0.5154

source: mainline.sas

(b) (4)

(b) (4)

1 Page has been Withheld in Full as b4 (CCI/TS) immediately following this page

3.2.4.4.4 Dactylitis

None of the three studies showed nominally significant differences between placebo and A30 with regard to dactylitis (Table 24).

Table 24. Dactylitis. Mean Change from Baseline, Week16

Week	Study	Δ Dactylitis			Treatment Difference			P-Value
		P	A20	A30	A20-P	A30-P	A30-A20	A30-P
16	2	-1.4	-1.9	-1.7	-0.5	-0.3	0.2	0.3978
		(63)	(56)	(66)				
	3	-1.12	-0.79	-1.3	0.3	-0.2	-0.6	0.5438
		(63)	(75)	(70)				
	4	-1.3	-1.7	-2.1	-0.4	-0.8	-0.4	0.0720
		(67)	(70)	(76)				

source: mainline.sas

3.3 Evaluation of Safety

Safety evaluations for this submission were conducted by the Medical Reviewer, Keith Hull, M.D. and are provided in his review.

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

To examine the impact of subgroups on treatment efficacy, logistic regression was conducted on the primary response variable, week 16 ACR20 response. Independent factor included treatment (P, A30), the stratification variables DMARD use and percent body surface area psoriasis involvement (study 4 only), the subgroup under examination, stratification variable by treatment interactions, and the subgroup by treatment interaction. Subgroup induced changes in treatment efficacy were examined by testing treatment by subgroup interactions at a nominal 0.05 level of significance, without correction for analysis of multiple endpoints.

4.1 Gender, Race, Age, and Geographic Region

No significant subgroup effects on efficacy of were seen for race (White, non-White), age class (<65, ≥65), or geographic region (USA, not USA).

A nominally significant effect of sex on treatment effect was seen in study 2 (p=0.009) but not in study 3 (p=0.70) or study 4 (p=0.32). In all three studies, the difference between A30 and P was numerically larger among male patients than among female patients (Table 25).

Table 25. ACR20 Percent Response by Sex. A30 versus Placebo, Week 16

Study	F		M		A30 - P		A30-P	P-Value*
	P	A30	P	A30	F	M	M - F	F - M
2	21	33	13	53	12	40	27	0.009
3	19	29	18	33	11	15	4	0.702
4	19	41	16	50	22	33	11	0.320

source: mainline.sas
P-Value for sex*treatment interaction

In a model with treatments A20 and P, the treatment by sex interaction was nominally significant in study 2 (p=0.01) but not in study 3 (p=0.26) or study 4 (p=0.63). In all three studies, the difference between A20 and P was numerically larger among male patients than among female patients; The difference between males and females for A20 - P equal in studies 2, 3, and 4 to 25%, 13%, and 4% respectively. Numerical values of A30 - P and A20 - P were positive among males and among females in all three studies.

4.2 Other Special/Subgroup Populations

4.2.1.1.1 Baseline DMARD Usage

A nominally significant effect of baseline DMARD usage on treatment effect was seen in study 2 (p=0.008) but not in study 3 (p=0.60) or study 4 (p=0.18). Back-transformed means from the model are provided in Table 26. Patients not taking DMARDS at baseline exhibited a numerically higher response to A30 compared to placebo in studies 2 and 4, but not in study 3.

Table 26. ACR20 Percent Response by Baseline DMARD Usage. A30 versus Placebo, Week 16

Study	**Y**		**N**		**A30 - P**		**A30-P**	**P-Value***
	P	**A30**	**P**	**A30**	**Y**	**N**	**N - Y**	**N - Y**
2	24	33	10	47	9	36	27	0.008
3	20	36	15	22	16	7	-9	0.600
4	22	43	14	48	21	33	13	0.182

source: mainline.sas
P-Value for DMARD*treatment interaction

In a model with treatments A20 and P, the treatment by baseline DMARD usage interaction was not nominally significant in any of the three studies.

4.2.1.1.2 Baseline DMARD Usage and Prior Biologic Use

The sponsor included a table on the proposed label which details from study 4 the joint effects on apremilast efficacy of DMARD usage and prior treatment with biologics. The statistical test for the interaction of treatment, prior biologic usage, and DMARD usage on ACR20 response was not significant, with a p-value of 0.088.

5 SUMMARY AND CONCLUSIONS

5.1 Statistical issues

In respective studies 2, 3, and 4, 72%, 65%, and 65% of placebo patients discontinued randomized treatment before week 24, obscuring maintenance of effects to week 24 for endpoints other than ACR 20.

5.2 Collective evidence

This submission demonstrates statistically significant benefits of apremilast 30 mg tablets two times daily (bid) for the treatment of adult patients with active psoriatic arthritis (PsA). Compared to placebo, patients in the three Phase 3 trials randomized to apremilast 30 mg experienced statistically significant improvements for the primary response variable, ACR20 at week 16, and the key secondary response variable, HAQ-DI score at week 16.

Statistical significance of differences between A30 and placebo for endpoints associated with all proposed label claims is summarized in Table 27 and Table 28. (b) (4)

Table 27. Statistical Significance of Proposed Label Claims. Preplanned Endpoints, Week 16

Claim	Significant Studies (3 total)	Average Effect (A30 - P)*
ACR20 W16	3	18.2%
Δ HAQ-DI W16	3	-0.14
		(b) (4)

Table 28. Nominal Significance of Proposed Label Claims. Exploratory Endpoints, Week 16

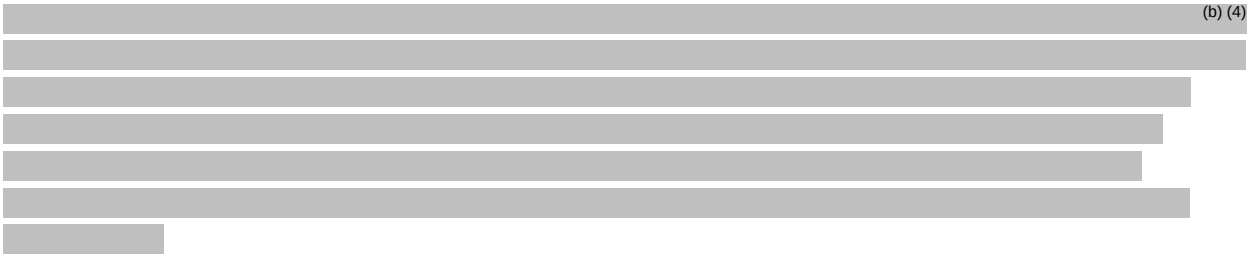
Claim	Significant studies (3 total)	Average Effect (A30 - P)
HAQ-DI response W16	2	10.6%
ACR 50 W16	1	7.6%
ACR 70 W16	0	1.6%
		(b) (4)

5.3 Conclusions and Recommendations

This submission demonstrates benefits of apremilast 30 mg tablets compared to placebo for the treatment of adult patients with active PsA. Three randomized double blinded placebo controlled parallel arm Phase 3 trials show that apremilast 30 mg provides statistically significant benefits compared to placebo for the primary endpoint ACR20 at week 16 (average 18%) as well as for the key secondary endpoint ΔHAQ-DI at week 16 (average -0.14).

Evidence for additional efficacy benefits of apremilast 30 mg over apremilast 20 mg are suggestive but not conclusive or even consistent, with effects of apremilast 20 mg compared to placebo statistically significant for the primary endpoint ACR 20 at week 16 in all three phase 3 studies, and statistically significant for key secondary endpoint HAQ-DI in two of the three phase 3 studies. Approval of apremilast 20 mg rather than apremilast 30 mg may therefore be justifiable if apremilast 30 mg poses large additional risks to safety compared to apremilast 20 mg.

Claims of effectiveness for endpoints at week 24 are considered in this review as claims for sustained effect beyond week 16. Such claims were confirmed for ACR 20 but undermined for other endpoints by the loss of adequate control; approximately 70% of placebo patients discontinued initial randomized treatment prior to week 24.



5.4 Labeling Recommendations

6 APPENDIX

6.1 Secondary Endpoints Weeks 16 and 24

1. Change from baseline in physical function HAQ-DI after 16 weeks of treatment
2. Proportion of subjects who achieve ACR 20 after 24 weeks of treatment
3. Change from baseline in physical function HAQ-DI after 24 weeks of treatment
4. Change from baseline in the physical function domain score of the Medical Outcome Study Short Form 36-Item Health Survey, Version 2 (SF-36) after 16 weeks of treatment
5. Proportion of subjects who achieve the modified Psoriatic Arthritis Response Criteria (PsARC) after 16 weeks of treatment

5a. (Study 4 only) Proportion of subjects in each treatment group, whose psoriasis body surface area (BSA) at baseline was ≥ 3%, that achieves PASI-75 after 16 weeks of treatment

6. Change from baseline in subject's assessment of pain (VAS) after 16 weeks of treatment
7. Change from baseline in the Maastricht Ankylosing Spondylitis Entheses Score (MASES) in subjects with pre-existing enthesopathy after 16 weeks of treatment
8. Change from baseline in the dactylitis severity score in subjects with pre-existing dactylitis after 16 weeks of treatment
9. Change from baseline in the Clinical Disease Activity Index (CDAI) after 16 weeks of treatment
10. Change from baseline in the Disease Activity Score (DAS28) after 16 weeks of treatment
11. Change from baseline in the Functional Assessment of Chronic Illness Therapy – Fatigue (FACIT-Fatigue) score after 16 weeks of treatment
12. Change from baseline in the physical function domain score of the SF-36 after 24 weeks of treatment
13. Proportion of subjects who achieve the modified PsARC after 24 weeks of treatment

13a. (Study 4 only) Proportion of subjects in each treatment group, whose psoriasis body surface area (BSA) at baseline was ≥ 3%, that achieves PASI-75 after 24 weeks of treatment

14. Change from baseline in subject's assessment of pain (VAS) after 24 weeks
15. Change from baseline in the MASES in subjects with pre-existing enthesopathy after 24 weeks of treatment
16. Change from baseline in the dactylitis severity score in subjects with pre-existing dactylitis after 24 weeks of treatment
17. Change from baseline in the CDAI after 24 weeks of treatment
18. Change from baseline in the DAS28 after 24 weeks of treatment
19. Change from baseline in the FACIT-Fatigue score after 24 weeks of treatment
20. Proportion of subjects with pre-existing enthesopathy whose MASES improves by ≥ 20% after 16 weeks of treatment
21. Proportion of subjects with pre-existing dactylitis whose dactylitis severity score

improves by ≥ 1 after 16 weeks of treatment
22. Proportion of subjects with a good or moderate European League Against Rheumatism (EULAR) response after 16 weeks of treatment
23. Proportion of subjects with pre-existing enthesopathy whose MASES improves by ≥ 20% after 24 weeks of treatment
24. Proportion of subjects with pre-existing dactylitis whose dactylitis severity score improves by ≥ 1 after 24 weeks of treatment

6.2 Secondary Endpoints: Week 52, Studies 2 and 3

1. Proportion of subjects with a good or moderate EULAR response after 24 weeks of treatment
2. Proportion of subjects who achieve an ACR 50, compared with baseline, after
3. 16 weeks of treatment
4. Proportion of subjects who achieve an ACR 70, compared with baseline, after 16 weeks of treatment
5. Proportion of subjects who achieve an ACR 50, compared with baseline, after 24 weeks of treatment
6. Proportion of subjects who achieve an ACR 70, compared with baseline, after 24 weeks of treatment
7. Proportion of subjects with pre-existing enthesopathy whose MASES improves to 0 after 16 weeks of treatment
8. Proportion of subjects who achieve the ACR 20, compared with baseline, after 52 weeks of treatment
9. Change from baseline in physical function (HAQ-DI) after 52 weeks of treatment
10. Change from baseline in the physical function domain score of the SF-36 after 52 weeks of treatment
11. Proportion of subjects who achieve the modified PsARC after 52 weeks of treatment

11a. (Study 4 only) Proportion of subjects in each treatment group, whose psoriasis body surface area (BSA) at baseline was ≥ 3%, that achieves PASI-75 after 52 weeks of treatment

12. Change from baseline in subject's assessment of pain (VAS) after 52 weeks of treatment
13. Change from baseline in the MASES in subjects with pre-existing enthesopathy after 52 weeks of treatment
14. Change from baseline in the dactylitis severity score subjects with pre-existing dactylitis after 52 weeks of treatment
15. Change from baseline in the CDAI after 52 weeks of treatment
16. Change from baseline in the DAS28 after 52 weeks of treatment
17. Change from baseline in the FACIT-Fatigue score after 52 weeks of treatment
18. Proportion of subjects with pre-existing enthesopathy whose MASES improves by ≥ 20% after 52 weeks of treatment

19. Proportion of subjects with pre-existing dactylitis whose dactylitis severity score improves by ≥ 1 after 52 weeks of treatment
20. Proportion of subjects with a good or moderate EULAR response after 52 weeks of treatment
21. Proportion of subjects who achieve an ACR 50, compared with baseline, after 52 weeks of treatment
22. Proportion of subjects who achieve an ACR 70, compared with baseline, after 52 weeks of treatment
23. Proportion of subjects with pre-existing enthesopathy whose MASES improves to 0 after 52 weeks of treatment
24. Proportion of subjects with pre-existing dactylitis whose dactylitis severity score improves to 0 after 52 weeks of treatment

---

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

/s/

ROBERT ABUGOV
11/20/2013

JOAN K BUENCONSEJO
11/20/2013
I concur

THOMAS J PERMUTT
11/25/2013
I concur.

Statistical Review and Evaluation
CARCINOGENICITY STUDIES

IND/NDA Number:	NDA 205437
Drug name:	Apremilast
Indication(s):	Active psoriatic arthritis
Applicant:	Celgene Corporation
Documents Reviewed:	Electronic submission Electronically submitted dataset Dated: 2013-03-21
Review Priority:	Normal
Biometrics Division:	Division of Biometrics 6
Statistical Reviewer:	Matthew Jackson, PhD
Concurring Reviewer:	Karl Lin, PhD
Medical Division:	Division of Pulmonary, Allergy, and Rheumatology Products
Reviewing Pharmacologist:	Steve Leshin, Ph.D.
Project Manager:	Michelle Jordan Garner
Keywords:	Animal Mouse Rat Carcinogenicity

Contents

List of Tables

List of Figures

Background

In this submission the sponsor included reports of two animal carcinogenicity studies, in mice and rats, to assess the carcinogenic potential of Apremilast when administered by gavage, once daily at appropriate drug levels for about 104 weeks. Results of this review have been discussed with the reviewing pharmacologist, Steve Leshin, Ph.D..

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 or decreasing mortality or tumor incidence rate as dose increases.

Chapter 1

Summary of findings

1.1 Mouse study

Whether or not the female mouse experiment should be considered positive depends on the significance thresholds used. By the standards of the new critical values recommended by the Pharmacology and Toxicology statistics team, the experiment is positive for the combination of osteomas and osteosarcomas. However, under the more conservative old standards, this is a negative finding.

The male mouse experiment is a negative experiment.

In both experiments, high levels of mortality in the treated groups lead to extensive readjustments of dosing levels. Ultimately, while a significant dose-related effect on survival was noted among the male mice, no such effect was found for the female mice. Nonetheless, it seems reasonable to assume that the dose levels were indeed high, and the absence of a mortality effect is simply a consequence of the dose readjustment process. Despite the high mortality levels, at least 20 animals in each group survived to 90 weeks, so it is reasonable to conclude that the dose levels were not quite so high as to interfere with our analyses, even for late developing tumors.

Autolysis levels were generally acceptable, except for the gallbladder. Likewise, large numbers of animals were reported as having the parathyroids and Peyer's patches unexamined. In the case of tumors of any of these three organs, both experiments should be considered inconclusive rather than negative.

1.2 Rat study

Both the female and male rat experiments are negative. The dose levels were clearly adequate, but the early mortality rates, especially for the male animals, mean that the appropriate survival adjusted populations for such tumors might well be too small to draw definite conclusions. In addition, the rate at which the Zymbal's gland in male rats was left unanalyzed, and the autolysis rates for the jejunum for both female and male rats mean that the studies should be considered inconclusive rather than negative for tumors associated with these endpoints.

Also of concern is the possibility that the early cessation of dosing will have masked genuine tumorigenic effects, especially among the high dose animals, and especially with late onset tumors.

Chapter 2

Mouse Study

2.1 Experimental design

The mouse study comprised two separate experiments; one in female mice and one in male mice. In each experiment, the animals used were CD1 mice. Two hundred and eighty animals of each sex were used, divided into four groups of seventy; a control group, who received the vehicle, and three treated groups; the low, mid, and high dose groups.

The dose levels were originally planned as 100, 300, and 1000 mg/kg for the three treated groups, in both female and male mice. However, after significant dose related toxicity was observed, all of the treated groups except the low dose male animals had their dose levels reduced, and in some cases eliminated. Table 2.1 describes the dose levels received by the animals in the different groups at different times. The *average* dose level in this table is the mean daily dose received by an animal who survived until termination. Note that the true average dose level received by a particular animal will depend on when it died, and will be higher for animals who died earlier than for those who lived longer.

The vehicle for the test was 1.0% sodium carboxymethylcellulose prepared with deionized water, in which the test article, Apremilast, was suspended. The dose volume was consistently 10mL/kg of bodyweight, although animals in the treated groups whose dose levels were changed to 0 appear not to have continued to receive this vehicle.

From the sponsor's report:

> All animals were observed twice daily, once in the morning and once in the afternoon, for mortality and moribundity. Detailed physical examinations were conducted on all . . . animals approximately weekly, beginning during acclimation upon individual housing. The absence or presence of findings was recorded for individual animals at the scheduled intervals. . . A separate computer protocol was used to record any observations noted outside of the above-specified intervals for the toxicology group animals.
>
> . . .
>
> All animals were examined weekly for the presence of palpable masses. The time of onset, location, size, appearance, and progression of each mass were recorded throughout the study period.
>
> Body weights were recorded weekly, beginning during acclimation, through study week 14 and biweekly thereafter.
>
> . . .
>
> Individual food consumption was recorded weekly, beginning during acclimation, through study week 14 and biweekly thereafter. Food intake was calculated as g/animal/day for the corresponding body weight intervals.

After death, whether premature or after sacrifice, all animals underwent a full necroscopy.

Table 2.1: Dose levels administered during mouse study

Sex	Group	Dose (mg/kg)	Period (weeks)	Termination	Average dose
Female	Control	0	1 — 102	102	0
	Low	100	1 — 99	102	73.7
		0	100 — 102		
	Mid	300	1 — 73	102	259.8
		200	74 — 96		
		0	97 — 192		
	High	1000	1 — 73	102	857.8
		0	74 — 102		
Male	Control	0	1 — 103	103	0
	Low	100	1 — 103	103	100
	Mid	300	1 — 73	103	261.2
		200	74 — 98		
		0	99 — 103		
	High	1000	1 — 73	103	708.7
		0	74 — 103		

2.2 Sponsor's analysis

2.2.1 Survival analysis

From the sponsor's report:

> A log-rank dose response trend test of survival rates was performed utilizing ordinal coefficients. In addition, a log-rank test for survival was used to make pairwise comparisons of each treated group with the control group. All tests were conducted at the 0.05 significance level.
>
> Survival times in which the status of the animals death was classified as an accidental death, planned interim sacrifice or terminal sacrifice, were considered censored values for the purpose of the Kaplan-Meier estimates and survival rate analyses.

The sponsor found a significant dose related reduction in survival among the male animals ($p = 0.0279$), although none of the individual dose groups experienced a significant reduction in survival compared with the vehicle control group. Among female mice, the only significant finding was the pairwise comparison between the mid dose and control animals ($p = 0.0277$). The sponsor does not discuss the significance of these findings.

2.2.2 Tumor analysis

Each distinct organ–tumor pair reported in at least two treated animals was analyzed using Peto's method [6]. Both dose response and pairwise tests were conducted. In addition, a number of combination endpoints were assessed. The combination endpoints considered are listed in table 2.2.

The sponsor reports no statistically significant findings.

Table 2.2: Combination endpoints considered in sponsor's analysis (mouse study)

Combination number	Sex considered	Organ–tumor pairs included	
		Organ	Tumor
1	Both	Liver	Hepatocellular carcinoma
			Hepatocellular adenoma
2	Male	Lung	Bronchiolo-alveolar carcinoma
			Bronchiolo-alveolar adenoma
3	Female	Adrenal cortex	Carcinoma
			Adenoma
4	Female	Bone	Osteosarcoma
			Osteoma
5	Female	Ovary	All granulosa cell tumors
6	Both	All	All

2.3 CDER reviewer's analysis

2.3.1 Survival analysis

The Kaplan-Meier survival plots are shown as figures 2.1 and 2.2. The numbers and proportions of animals surviving to various times are presented in table A.1. The results of log-rank tests of heterogeneity of survival and of dose response across the groups are presented in table A.2, and the results of log-rank survival tests comparing the treated groups with the control group are presented in table A.3.

Commentary When the survival data are viewed in isolation, there is no compelling evidence of a dose related mortality effect among female mice (the test of trend does not show a statistically significant dose related increase in mortality). Although the effect is clearer among male mice ($p = 0.0136$), no one group has experienced a mortality rate significantly higher than the control group.

However, these negative and weak findings have to be balanced against the fact of the changes in the dose levels. Note, for instance, that among female mice, the mortality rate among the mid dose group exceeds that of the high dose group just after the high dose group stopped receiving Apremilast, even as the mid dose group merely received a reduction in their dose levels. It thus seems reasonably safe to conclude that had the doses not been adjusted, we would have observed evidence of dose-related mortality (the goal of the dose reductions was, after all, to reduce the mortality rate in the higher dose groups).

2.3.2 Tumor analysis

Endpoints

Analyses have been conducted using the sponsor's submitted dataset, and the sponsor's chosen nomenclature. In this dataset, organs or tissue types are described as being either tumorous, examined but found unusable due to autolysis, or unexamined. An organ that has been examined but was not found to be tumorous is not mentioned in the dataset.

From these data, we can infer the numbers of animals for which each organ or tissue type was examined, but only in those cases where at least one anomalous finding (i.e., a tumor was found, or a sample that was planned to be analyzed could not be, either because no sample was taken or because the sample was unusable due to autolysis) was reported. Organs which can thus be deduced to have been successfully analyzed in the majority of animals are, for the purposes of this review, considered *primary*. The lists of primary organs in the experiments on female and male mice respectively are presented in tables A.4 and A.5.

Organ or tissue types which were examined in only a few animals are considered *secondary*.

Figure 2.1: Survival curves for female mice

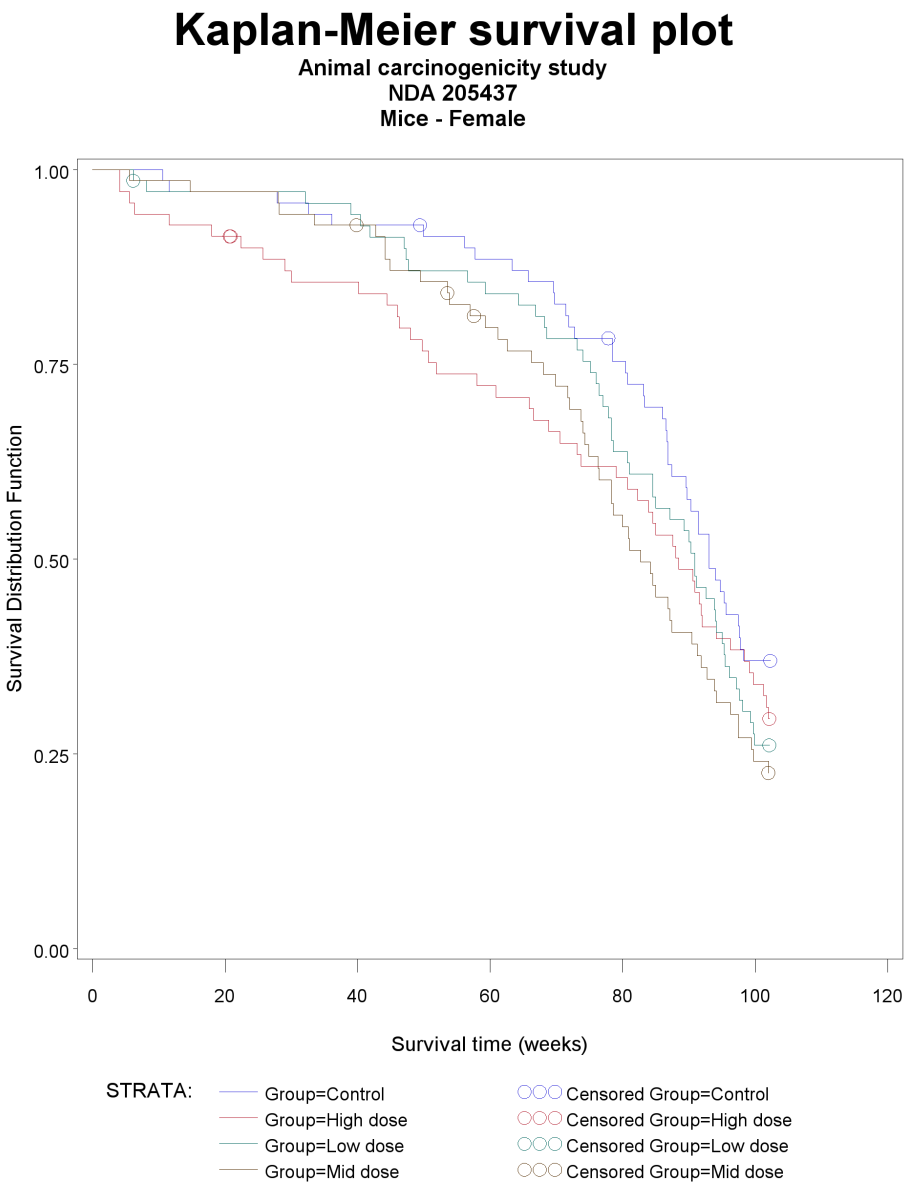
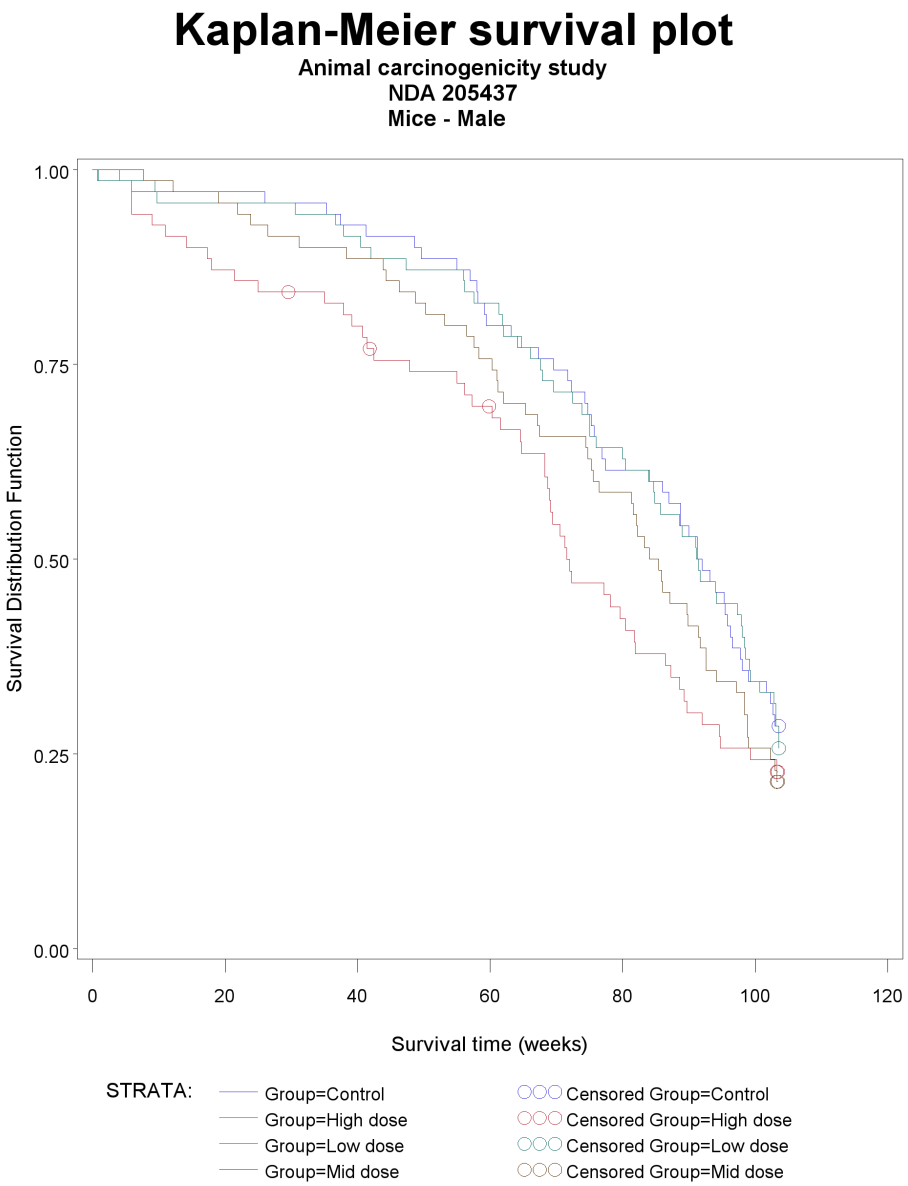


Figure 2.2: Survival curves for male mice



In the mouse study, there are no secondary organs.

Each tumor type found in a primary organ of at least one animal is considered a primary endpoint. In addition, in consultation with Steve Leshin, Ph.D., a list of combination endpoints has been drawn up. This list is presented in table A.6.

Statistical procedure

The tumor data were analyzed for dose response relationships and pairwise comparisons of tumor incidence in each of the treated groups versus the control group. Both the dose response relationship tests and pairwise comparisons were performed using the poly-k method described in the paper of Bailer and Portier[1] and developed in the paper of Bieler and Williams[2]. In this method, given a tumor type T, an animal h that lives the full study period (w_m) or dies before the terminal sacrifice with at least one tumor of type T gets a score of $s_h = 1$. An animal that dies at week w_h before the end of the study without such a tumor gets a score of

$$s_h = \left(\frac{w_h}{w_m}\right)^k < 1.$$

The adjusted group size is defined as $\sum_h s_h$. As an interpretation, an animal with score $s_h = 1$ can be considered as a whole animal while an animal with score $s_h < 1$ can be considered as a partial animal. The adjusted group size $\sum s_h$ is equal to N (the original group size) if all animals live up to the end of the study or if each animal develops at least one tumor of type T, otherwise the adjusted group size is less than N. These adjusted group sizes are then used for the dose response relationship (or the pairwise) tests using the Cochran-Armitage test. The test is repeated for each tumor type T.

One critical point to consider in the application of the poly-k test is the choice of the appropriate value of k, which depends on the relationship between tumor onset time and increased dose. For long term 104 week standard rat and mouse studies, a value of $k = 3$ is suggested in the literature, and so has been used in this review. For the calculation of p-values, the exact permutation method was used.

When testing so many endpoints, there is a danger of inflation of type I error. To control against this, the current draft guidance recommends making adjustments in the significance thresholds. In order to best manage the trade-off between control of type I and type II error, and to allow for the relative rarity of some tumors, it is recommended that a distinction be drawn between rare tumors (with a background incidence rate below 1%) and common tumors. For a two year study of two species, the currently proposed significance thresholds are given in table 2.3. It is expected that these adjustments will suffice to keep the submission-wide false positive rate at a nominal level of approximately 10%.

However, it is also understood that the ECAC is currently exercising its prerogative to adhere to the old thresholds presented in table 2.4 rather than the new thresholds recommended by the Pharmacology and Toxicology statistical review team.

It should be noted that the FDA guidance for multiple testing for dose response relationship is based on a publication by Lin and Rahman [5]. In this work the authors investigated the use of this rule for Peto analysis. However, in a later work Rahman and Lin [7] showed that this rule for multiple testing for dose response relationship is also suitable for poly-k tests.

Table 2.3: Critical p-values used to determine statistical significance

Type of test	Rare tumor	Common tumor
Trend	0.025	0.005
Pairwise test between placebo and high dose	0.10	0.05

In this particular study, there is an additional problem; what values to use for the dose levels in the trend tests. We have chosen to use the average daily dose level, as shown in table 2.1. However,

Table 2.4: Old critical p-values used to determine statistical significance

Type of test	Rare tumor	Common tumor
Trend	0.025	0.005
Pairwise test between placebo and high dose	0.05	0.01

the poly-k method is not very sensitive to these choices, so it is unlikely that alternative analyses would have been very different.

The results of the statistical analyses of tumor incidence in primary endpoints are presented in tables A.7 (female mice) and A.8 (male mice). The results of analyses of customized endpoints (see table A.6) are presented in tables A.9 and A.10.

Noteworthy results

No tests of individual tumor types in female mice were conducted which yielded p-values below 0.05. Combination tumor types for which tests yielding p-values below 0.05 were conducted are presented in table A.11, which is excerpted from table A.9. No statistical tests were conducted in the male mouse experiment which resulted in p-values below 0.05.

Osteomas and osteosarcoms in female mice Three animals developed osteomas or osteosarcomas; all were high dose female animals. The statistical tests yield p-values of 0.0128 (trend test) and 0.0918 (pairwise test). These are sufficient to justify a positive finding for a rare tumor type according to the currently recommended significance thresholds (table 2.3), but not according to the old standards (table 2.4). This is is therefore a positive finding only if one considers these to be rare tumors (which seems reasonable) and if one accept the appropriateness of the currently recommended standards. Otherwise, this should be considered a narrowly negative finding.

2.3.3 Analysis of unexamined and autolytic organs

Unexamined animals

No animals have been reported as completely unexamined.

Organs reported autolytic

The numbers of organs found in female mice to be autolytic to the extent that analysis of collected tissue was not possible are presented in table A.12. The numbers of such organs found in male mice are presented in table A.13.

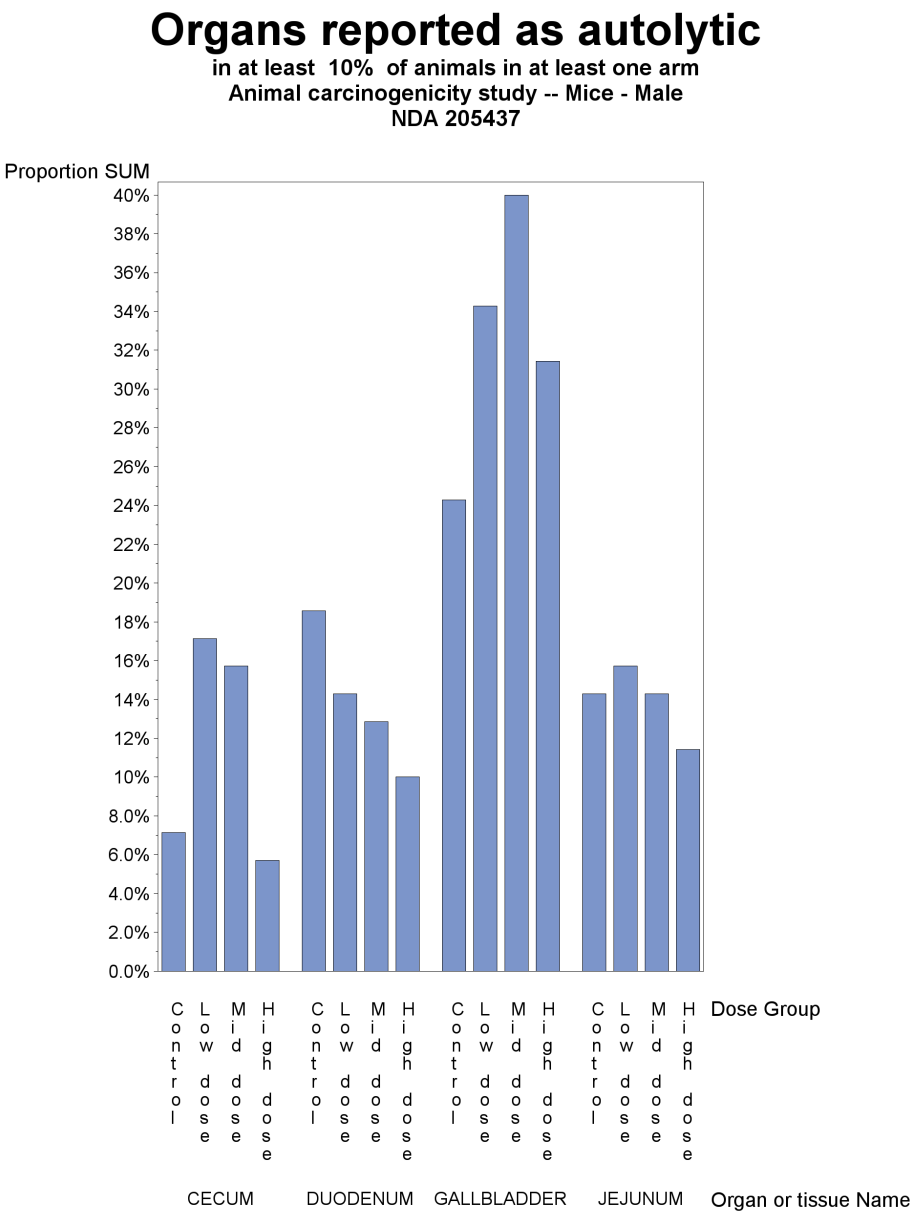
Autolysis rates were generally acceptable, except for the gallbladder; 28% of female animals and 33% of male animals had this organ autolyzed to the extent that a usable sample could not be obtained. The study should therefore be viewed as inconclusive for tumores of the gallbladder, rather than negative.

Aside from the gallbladder, the only frequently autolyzed organs were the jejunum, the duodunum, and the cecum in male mice, which were reported as autolyzed in between 11% and 14% of animals. However, the fact that the autolysis was concentrated in the low and mid dose groups (see figure 2.3 means that the impact on our analyses was slight.

Organs reported as unexamined

The numbers of animals with organs reported as being unexamined are presented in tables A.14 and A.15. The only organs for which large numbers of animals have been reported as unexamined are the parathyroids and Peyer's patches in both the female and male mice; in each case more than 30% of animals have been recorded as having these organs unexamined. While such rates for these organs are no cause for concern about the conduct of the study, it is still the case that the study should be regarded as inconclusive for tumors of these organs, rather than negative.

Figure 2.3: Autolysis rates for male mice



Chapter 3

Rat Study

3.1 Experimental design

The rat study comprised two separate experiments; one in female rats and one in male rats. In each experiment, the animals used were Crl:CD(SD) rats. Two hundred and eighty animals of each sex were used, divided into four groups of seventy; a control group, who received the vehicle, and three treated groups; the low, mid, and high dose groups.

The dose levels for the female experiment were originally planned as 0.3, 1, and 3 mg/kg for the three dose groups. The corresponding levels for the male rat experiment were were 3, 10, and 20 mg/kg. However, after significant dose related toxicity was observed, all six treated groups had their dose levels reduced, and in some cases eliminated. Table 3.1 describes the dose levels received by the animals in the different groups at different times. The *average* dose level in this table is the mean daily dose received by an animal who survived until termination. Note that the true average dose level received by a particular animal will depend on when it died, and will be higher for animals who died earlier than for those who lived longer.

The vehicle for the test was 1.0% sodium carboxymethylcellulose prepared with deionized water, in which the test article, Apremilast, was suspended. The dose volume was consistently 10mL/kg of bodyweight, although animals in the treated groups whose dose levels were changed to 0 appear not to have continued to receive this vehicle.

From the sponsor's report:

> All animals were observed twice daily, once in the morning and once in the afternoon, for mortality and moribundity.
>
> Detailed physical examinations were conducted on all ... animals approximately weekly, beginning during acclimation upon individual housing. A separate computer protocol was used to record any observations noted outside of the above-specified intervals.
>
> ...
>
> All animals were examined weekly for the presence of palpable masses. The time of onset, location, size, appearance, and progression of each mass were recorded throughout the study period.
>
> Body weights were recorded weekly, beginning during acclimation, through study week 14 and biweekly thereafter.
>
> ...
>
> Individual food consumption was recorded weekly, beginning during acclimation, through study week 14 and biweekly thereafter. Food intake was calculated as g/animal/day for the corresponding body weight intervals.

After death, whether premature or after sacrifice, all animals underwent a full necroscopy.

Table 3.1: Dose levels administered during rat study

Sex	Group	Dose (mg/kg)	Period (weeks)	Termination	Average dose
Female	Control	0	1 — 104	104	0
	Low	0.3	1 — 103	104	0.30
	Mid	1	1 — 101	104	0.97
		0	102 — 104		
	High	3	1 — 94	104	2.71
		0	95 — 104		
Male	Control	0	1 — 102	102	0
	Low	3	1 — 91	100	2.73
		0	92 — 100		
	Mid	10	1 — 66	98	8.14
		6	67 — 89		
		0	90 — 98		
	High	20	1 — 66	95	13.89
		0	67 — 95		

3.2 Sponsor's analysis

3.2.1 Survival analysis

A log-rank dose response trend test of survival rates was performed utilizing ordinal coefficients. In addition, a log-rank test for survival was used to make pairwise comparisons of each treated group with the control group. All tests were conducted at the 0.05 significance level.

Survival times in which the status of the animals death was classified as an accidental death, planned interim sacrifice or terminal sacrifice, were considered censored values for the purpose of the Kaplan-Meier estimates and survival rate analyses.

The sponsor found a significant dose related reduction in survival among the male animals ($p = 0.0023$), and a significant decrease in survival in the high dose group compared with the vehicle control group ($p = 0.0020$). Among female rats, no statistically significant findings were reported. The sponsor does not discuss the significance of these findings.

3.2.2 Tumor analysis

Each distinct organ–tumor pair reported in at least two treated animals was analyzed using Peto's method [6]. Both dose response and pairwise tests were conducted. In addition, a number of combination endpoints were assessed. The combination endpoints considered are listed in table 3.2.

The sponsor reports no statistically significant findings.

Table 3.2: Combination endpoints considered in sponsor's analysis (rat study)

Combination number	Sex considered	Organ–tumor pairs included	
		Organ	Tumor
1	Both	Liver	Hepatocellular carcinoma
			Hepatocellular adenoma
2	Both	All	All

3.3 CDER reviewer's analysis

3.3.1 Survival analysis

The Kaplan-Meier survival plots are shown as figures 3.1 and 3.2. The numbers and proportions of animals surviving to various times are presented in table B.1. The results of log-rank tests of heterogeneity of survival and of dose response across the groups are presented in table B.2, and the results of log-rank survival tests comparing the treated groups with the control group are presented in table B.3.

Commentary Among both the female rats ($p = 0.0305$) and the male rats (0.0011), there is statistically significant evidence of a dose-related reduction in survival, although only the only group with significantly reduced survival relative to the control group in a pairwise test is the high dose male group ($p = 0.0035$). The mortality rates are especially high for the mid and high dose male animals, for whom just 33 and 25 animals (respectively) survived to the 78th week, although the mortality rates in these groups seem to have eased after the approximate time that the doses were reduced (week 66).

3.3.2 Tumor analysis

Endpoints

As in the mouse study, organs have been classed as either primary or secondary (see Section 2.3.2). The lists of organs adduced to be primary are presented in tables B.4 and B.5. In the rat study, there are no secondary organs.

The same customized endpoints have been analyzed as were considered in the mouse study (see table A.6).

Statistical procedure

The same statistical procedures were used to assess tumor incidence in rats as were used in mice (see Section 2.3.2). Note that the critical p-values used to determine significance are presented in table 2.3 (if the new critical values are preferred) and table 2.4 (otherwise).

The results of the statistical analyses of tumor incidence in primary endpoints are presented in tables B.6 (female rats) and B.7 (male rats). The results of analyses of customized endpoints (see table A.6) are presented in tables B.8 and B.9.

As with the mouse experiment, in this particular study, we face the additional problem of what values to use for the dose levels in the trend tests. as before, we have chosen to use the average daily dose level, as shown in table 3.1. However, the poly-k method is not very sensitive to these choices, so it is unlikely that alternative analyses would have been very different.

Noteworthy results

No tests of individual tumor types in female rats were conducted which yielded p-values below 0.05. Combination tumor types for which tests yielding p-values below 0.05 were conducted are presented in table B.10, which is excerpted from table B.8. No statistical tests were conducted in the male rat experiment which resulted in p-values below 0.05.

Figure 3.1

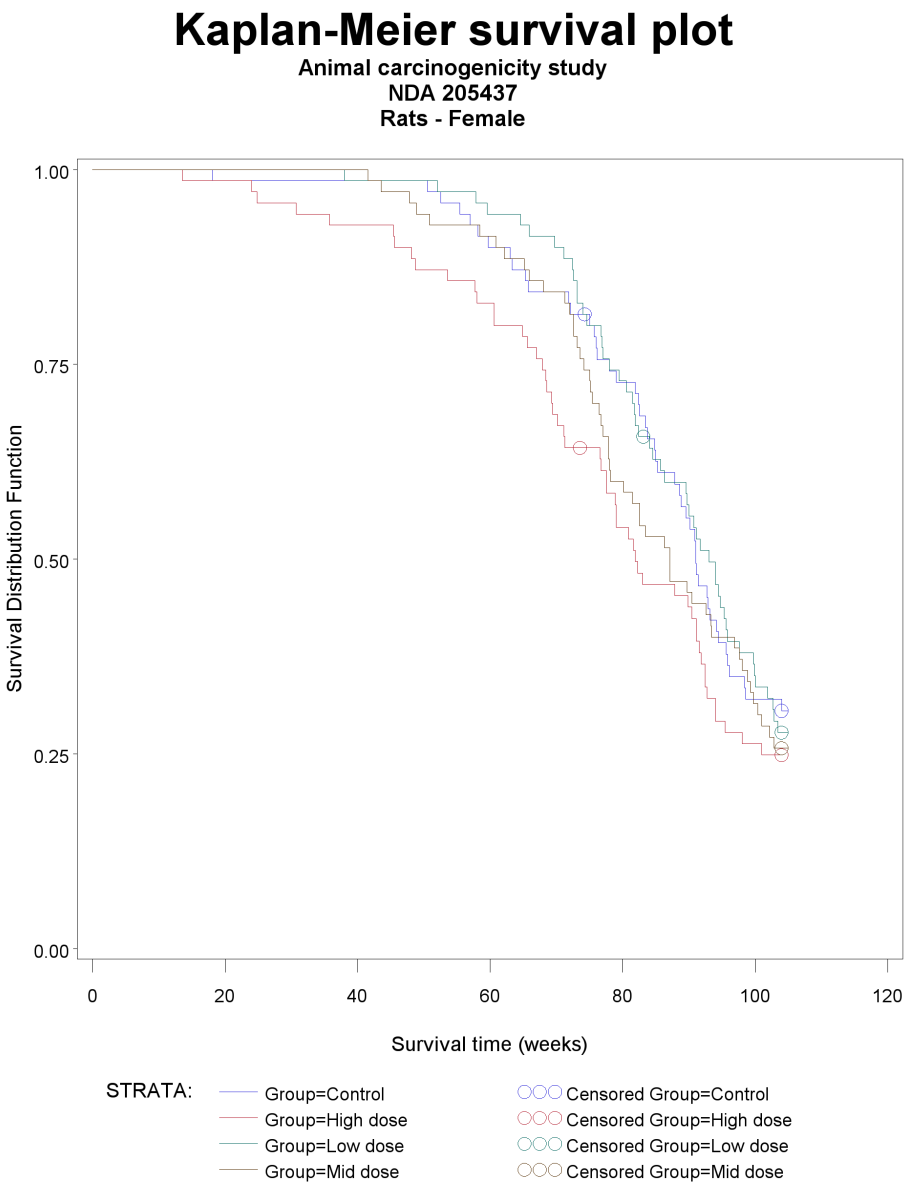
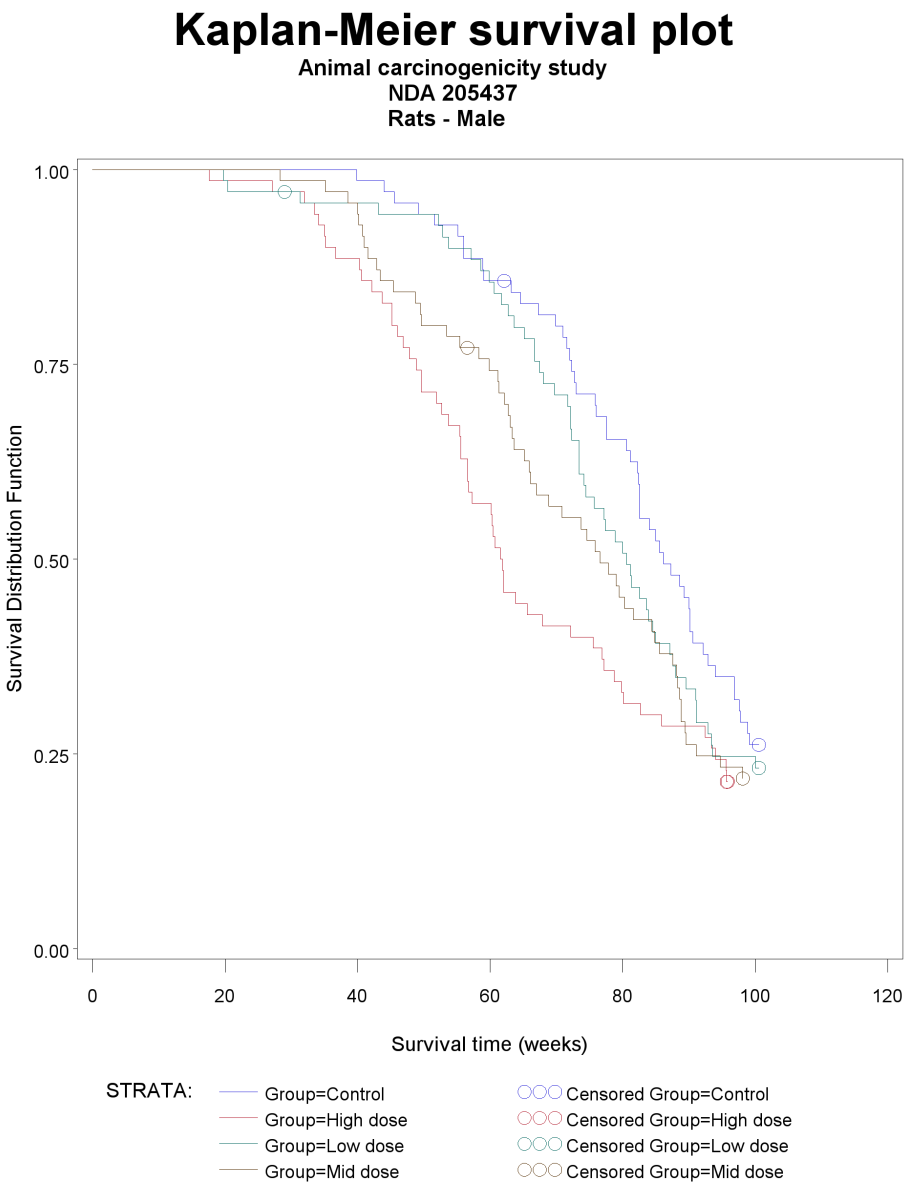

Kaplan-Meier survival plot
Animal carcinogenicity study
NDA 205437
Rats - Female
Survival Distribution Function
1.00
0.75
0.50
0.25
0.00
0
20
40
60
80
100
120
Survival time (weeks)
STRATA:
Group=Control
Group=High dose
Group=Low dose
Group=Mid dose
Censored Group=Control
Censored Group=High dose
Censored Group=Low dose
Censored Group=Mid dose

Figure 3.2


Kaplan-Meier survival plot
Animal carcinogenicity study
NDA 205437
Rats - Male
Survival Distribution Function
1.00
0.75
0.50
0.25
0.00
0
20
40
60
80
100
120
Survival time (weeks)
STRATA:
Group=Control
Group=High dose
Group=Low dose
Group=Mid dose
Censored Group=Control
Censored Group=High dose
Censored Group=Low dose
Censored Group=Mid dose

Sertoli cell tumors in female rats The only endpoint for which any consideration is required is Sertoli cell tumors in female rats. In this case, the p-value for the test of trend is 0.0456. However, this is not sufficient for a positive finding, even for a rare tumor. Additionally, the pairwise test if far from significant ($p = 0.2066$). Thus we must consider this a negative finding.

3.3.3 Analysis of unexamined and autolytic organs

Unexamined animals

No animals have been reported as completely unexamined.

Organs reported autolytic

The numbers of organs found in female rats to be autolytic to the extent that analysis of collected tissue was not possible are presented in table B.11. The numbers of such organs found in male rats are presented in table B.12.

The only organ for which the autolysis levels are high enough to be problematic is the jejunum; among both female and male rats, the autolysis rate for this organ is 18% across all groups. Given our concern about the small number of animals at risk, due to the high levels of toxicity, these levels are sufficient to prevent us from drawing any conclusions regarding the tumorigenic effect of Apremilast on tumors of the jejunum.

Organs reported as unexamined

The numbers of animals with organs reported as being unexamined are presented in tables B.13 and B.14. As with the jejunum, a substantial number of animals (15% of male animals) are reported to have had their zymbal's glands unexamined. As with the jejunum above, thi is sufficient to deny us any conclusion for tumors of this organ.

Chapter 4

Assessment of the validity of a negative study

4.1 Issues of concern when selecting the dose levels

The selection of an appropriate dose level for the high dose group is made difficult by the need to satisfy two competing imperatives: on the one hand, if the dose level is insufficiently high, then genuine carcinogenicity effects may not be apparent, but on the other hand, if the dose level is too high, then there is a risk of non-carcinogenic toxic effects killing the animals before they have a chance to demonstrate a carcinogenicity effect.

Haseman [4] suggested that a satisfactory balance between these two imperatives has been found when the following two conditions are both satisfied:

1. Were enough animals exposed, for a sustained amount of time, to the risk of late developing tumors?

2. Were dose levels high enough to pose a reasonable tumor challenge to the animals?

There is no consensus among experts regarding the number of animals and length of time at risk, although most carcinogenicity studies are designed to run for two years with fifty animals per treatment group. The following are some rules of thumb regarding these two issues as suggested by experts in this field:

Haseman [4] has done an investigation on the first issue. He gathered data from 21 studies using Fischer 344 rats and B6C3Fl mice conducted at the National Toxicology Program (NTP). It was found that, on the average, approximately 50% of the animals in the high dose group survived the two year study period. Also, in a personal communication with Dr. Karl Lin of Division of Biometrics-6, Haseman suggested that, as a rule of thumb, a 50% survival of 50 initial animals or 20 to 30 animals still alive in the high dose group, between weeks 80—90, would be considered as a sufficient number and adequate exposure. In addition Chu, Cueto and Ward [3], suggested that "to be considered adequate, an experiment that has not shown a chemical to be carcinogenic should have groups of animals with greater than 50% survival at one year."

It appears, from these three sources that the proportions of survival at 52 weeks, 80–90 weeks, and two years are of interest in determining the adequacy of exposure and number of animals at risk.

Regarding the question of adequate dose levels, it is generally accepted that the high dose should be close to the maximum tolerated dose (MTD). In the paper of Chu, Cueto and Ward [3], the following criteria are mentioned for dose adequacy. A high dose is considered as close to MTD if any of the criteria is met:

1. A dose is considered adequate if there is a detectable loss in weight gain of up to 10% in a dosed group relative to the controls.

2. The administered dose is also considered an MTD if dosed animals exhibit clinical signs or severe histopathologic toxic effects attributed to the chemical.

3. In addition, doses are considered adequate if the dosed animals show a slight increased mortality compared to the controls.

4.2 Assessment of the validity of the mouse study

The question of whether or not the female mouse experiment is a negative study depends on the choice of threshold used to determine statistical significance. The male mouse experiment is unambiguously negative. It is therefore appropriate in each case to consider whether the experiments posed an adequate challenge to a sufficient number of animals.

In each case, the need to reduce dose levels to reduce mortality rates is strong evidence that the dose levels were indeed at or above the MTD. In each group, at least twenty animals survived to week 90, so according to Haseman's guidelines, the number of surviving animals is adequate, albeit only barely. The dose level can be considered to be just short of excessive.

4.3 Assessment of the validity of the rat study

Both experiments are negative, and so it is appropriate to ask whether the a sufficient number of animals faced a sufficient dose for a sufficiently long time to allow us to draw genuine negative conclusions.

In both sexes, there is ample evidence that the dose levels were adequate, and indeed, most likely above the MTD.

The question of whether sufficient animals survived long enough is harder to address, especially for the male rat experiment. While twenty animals of each group (except the mid dose male group) did survive until week 90, the pattern of survival was atypical, with a higher rate of early death than the 90-week survival rate would normally suggest. Accordingly, the poly-3 adjusted population is quite small. Ultimately, it is reasonable to consider the sample sizes to be just adequate for most tumors, but to be very careful about drawing negative conclusions about those tumors which are typically found in a predominantly geriatric population.

In addition, the fact that large numbers of animals stopped being dosed after just 66 weeks raises concerns about whether tumorigenic effects might be halted with the premature cessation of dosing.

Appendix A

Tables from mouse study

A.1 Survival analysis

Table A.1

Survival rates at key times
NDA 205437
Animal carcinogenicity study
Mice

Species and Sex	*Dose Group*	*Dose (mg per kg)*	*Number at start*	*Number alive after 52 weeks*	*Percentage alive after 52 weeks*	*Number alive after 78 weeks*	*Percentage alive after 78 weeks*	*Number alive after 90 weeks*	*Percentage alive after 90 weeks*	*Number sacrificed*	*Percentage sacrificed*	*Maximum surivial (weeks)*
Mice - Female	Control	0	70	63	90%	53	76%	39	56%	25	36%	102
	Low dose	97.1	70	60	86%	47	67%	37	53%	18	26%	102
	Mid dose	259.8	70	59	84%	40	57%	27	39%	15	21%	102
	High dose	857.8	70	50	71%	42	60%	33	47%	20	29%	102
Mice - Male	Control	0	70	62	89%	43	61%	38	54%	20	29%	104
	Low dose	100	70	61	87%	45	64%	37	53%	18	26%	104
	Mid dose	261.2	70	57	81%	41	59%	29	41%	15	21%	103
	High dose	708.7	70	50	71%	30	43%	20	29%	15	21%	103

Table A.2

Log-rank tests of survival
NDA 205437
Animal carcinogenicity study
Mice

Sex	*Test of homogeneity: chi squared statistic*	*Test of homogeneity: degrees of freedom*	*Number of groups*	*Test of homogeneity: p-value*	*Test of trend (two tailed): p-value*	*Test of trend (one tailed): p-value*
Female	4.6168	3	4	0.2021	0.3945	0.1973
Male	5.2055	3	4	0.1574	0.0272	0.0136

Table A.3

Pairwise comparisons (log-rank) of survival between treated groups and controls
NDA 205437
Animal carcinogenicity study
Mice

Species and Sex	*Quantity*	*Low dose*	*Mid dose*	*High dose*
Mice - Female	Chi squared test statistic	1.6081	4.5658	1.8391
	p-value of comparison with control	0.2048	0.0326	0.1751
Mice - Male	Chi squared test statistic	0.0388	1.5367	3.8341
	p-value of comparison with control	0.8438	0.2151	0.0502

A.2 Tumor analysis

Table A.4

Primary organs in study of female mice
NDA 205437
Animal carcinogenicity study

Organ or tissue name
ADRENAL CORTEX
ADRENAL MEDULLA
AORTA
BONE
BRAIN
CECUM
CERVIX
CLITORAL GLANDS
COLON
DUODENUM
GALLBLADDER
HARDERIAN GLANDS
ILEUM
JEJUNUM
LAC. GLAND EXOR
LARYNX
LIVER
LUNGS
LYMPH NODE, MAND
LYMPH NODE, MED
LYMPH NODE, MES
LYMPH NODE, TR/B
MAMMARY GLAND
NASAL LEVEL IV
NERVE, SCIATIC
OVARIES
OVIDUCTS
PANCREAS
PARATHYRO DS
PEYER'S PATCHES
PITUITARY
SAL. GLAND MAND
SKIN
SOFT TISSUE, ABD
SPINAL CORD
SPLEEN
STOMACH, GLAN
STOMACH, NON

Primary organs in study of female mice
NDA 205437
Animal carcinogenicity study

Organ or tissue name
SYSTEMIC TUMORS
TEETH
THORACIC CAVITY
THYMUS
THYROID GLANDS
TRACHEA
URINARY BLADDER
UTERUS
VAGINA
ZYMBAL'S GLANDS

Table A.5

Primary organs in study of male mice
NDA 205437
Animal carcinogenicity study

Organ or tissue name
ADRENAL CORTEX
ADRENAL MEDULLA
AORTA
BILE DUCT
BONE
BRAIN
CECUM
COLON
DUODENUM
EPIDIDYMIDES
GALLBLADDER
HARDERIAN GLANDS
ILEUM
JEJUNUM
KIDNEYS
LAC. GLAND EXOR
LARYNX
LIVER
LUNGS
LYMPH NODE, MAND
LYMPH NODE, MES
LYMPH NODE, POP
LYMPH NODE, REN
MAMMARY GLAND
MARROW, STERN
NASAL LEVEL I
NASAL LEVEL II
NASAL LEVEL III
NASAL LEVEL IV
PANCREAS
PARATHYRO DS
PEYER'S PATCHES
PHARYNX
PITUITARY
PREPUTIAL GLANDS
PROSTATE
RECTUM
SAL. GLAND MAND

Primary organs in study of male mice
NDA 205437
Animal carcinogenicity study

Organ or tissue name
SEMINAL VESICLES
SKIN
SOFT TISSUE, THO
SPINAL CORD
SPLEEN
STOMACH, GLAN
STOMACH, NON
SYSTEMIC TUMORS
TESTES
THYMUS
THYROID GLANDS
TONGUE
TRACHEA
URETERS
URINARY BLADDER
ZYMBAL'S GLANDS

Table A.6

Customized and combination endpoints analyzed
NDA 205437
Animal carcinogenicity study

Composite endpoint
A cell tumors of the adrenal cortex
Acinar cell tumors
Adenocarcinomas and carcinomas of the mandibular lymph node
Adenomas and carcinomas of the adrenal cortex (excluding A cell tumors)
All cervical and uterine polyps
All hibernomas
All leiomyosarcomas
All lipomas
All schwannomas
Bronchiolo-alveolar tumors
C-cell tumors
Cervical and uterine endometrial stromal sarcomas
Cervical and vaginal carcinomas
Cervical, uterine, and vaginal fibromas
Cervical, uterine, and vaginal fibromas and fibrosarcomas
Endometrial stromal sarcomas and polyps of the cervix and uterus
Fibromas and fibrosarcomas of the skin (and tail)
Follicular cell tumors
Gastrointestinal adenocarcinomas
Gastrointestinal adenomas
Gastrointestinal adenomas and adenocarcinomas
Glial cell tumors
Granular cell tumors
Harderian gland adenomas and adenocarcinomas
Hemangiomas and hemangiosarcomas
Hepatocellular tumors
Histiocytoma and reticulosis
Internal squamous cell papillomas and carcinomas
Islet cell tumors
Leiomyomas and leiomyosarcomas of the uterus
Leiomyosarcomas and leiomyomas
Mammary adenoma, adenocarcinomas, and adenocanthomas
Meningiomas and meningeal sarcomas
Osteomas and osteosarcomas
Ovarian Sertoli cell tumors and tubulostromal adenomas
Ovarian luteomas and thecomas

Table A.6

Customized and combination endpoints analyzed
NDA 205437
Animal carcinogenicity study

Composite endpoint
Sarcoma, undifferentiated in reproductive tissues
Sarcoma, undifferentiated of the uterus and vagina
Sarcoma, undifferentiated, of the skin and paws
Sarcomas of the cervix, uterus, and vagina (excluding fibrosarcomas)
Sarcomas of the cervix, uterus, and vagina (including fibrosarcomas)
Sebaceous cell tumors
Sertoli cell tumors
Squamous cell papillomas of the skin and tail
Tumors of the pars distalis and pars intermedia
Uterine fibromas and fibrosarcomas
Zymbal glands tumors

Table A.7

Table of reported tumors in Mouse Study
NDA 205437
Animal carcinogenicity study
Female mice

Organ or tissue name	*Tumor name*	*Quantity*	*Control Size = 70*	*Low dose Size = 70*	*Mid dose Size = 70*	*High dose Size = 70*
ADRENAL CORTEX	ADENOMA	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
	ADENOMA, A CELL	P-value of test of trend or comparison	.1709		.1944	.4512
		Number of animals reported with tumor	0	0	2	1
	CARCINOMA, A CELL	P-value of test of trend or comparison	.2342			.4512
		Number of animals reported with tumor	0	0	0	1
ADRENAL MEDULLA	PHEOCHROMOCYTOMA	P-value of test of trend or comparison	.5827	.4824	.4430	
		Number of animals reported with tumor	0	1	1	0
BONE	OSTEOMA	P-value of test of trend or comparison	.2342			.4512
		Number of animals reported with tumor	0	0	0	1
	OSTEOSARCOMA	P-value of test of trend or comparison	.0560			.2066
		Number of animals reported with tumor	0	0	0	2
BRAIN	OLIGODENDROGLIOMA	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
CERVIX	CARCINOMA	P-value of test of trend or comparison	.7197	.4824		
		Number of animals reported with tumor	0	1	0	0
	FIBROMA	P-value of test of trend or comparison	.2357			.4568
		Number of animals reported with tumor	0	0	0	1
	GRANULAR CELL TUMOR	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
	LEIOMYOSARCOMA	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
DUODENUM	ADENOCARCINOMA	P-value of test of trend or comparison	.0565			.2108
		Number of animals reported with tumor	0	0	0	2
HARDERIAN GLANDS	ADENOCARCINOMA	P-value of test of trend or comparison	.8533	.8613	.6034	1
		Number of animals reported with tumor	2	1	2	0
	ADENOMA	P-value of test of trend or comparison	.3166	.4474	.2522	.3902

Table A.7

Table of reported tumors in Mouse Study
NDA 205437
Animal carcinogenicity study
Female mice

Organ or tissue name	*Tumor name*	*Quantity*	*Control Size = 70*	*Low dose Size = 70*	*Mid dose Size = 70*	*High dose Size = 70*
		Number of animals reported with tumor	3	4	5	4
JEJUNUM	ADENOCARCINOMA	P-value of test of trend or comparison	.7347	.5000		
		Number of animals reported with tumor	0	1	0	0
LIVER	ADENOMA, HEPATOCELLULAR	P-value of test of trend or comparison	.3856	.2244	.1944	.4512
		Number of animals reported with tumor	0	2	2	1
	CARCINOMA, HEPATOCELLULAR	P-value of test of trend or comparison	.5994	.2084	.5501	.5779
		Number of animals reported with tumor	3	6	3	3
LUNGS	ADENOMA, BRONCHIOLO-ALVEOLAR	P-value of test of trend or comparison	.9530	.5600	.9642	.9673
		Number of animals reported with tumor	7	7	2	2
	CARCINOMA, BRONCHIOLO-ALVEOLAR	P-value of test of trend or comparison	.6937	.7007	.9986	.8087
		Number of animals reported with tumor	10	8	1	6
LYMPH NODE, MAND	ADENOCARCINOMA; UNKNOWN	P-value of test of trend or comparison	.2357			.4512
		Number of animals reported with tumor	0	0	0	1
	CARCINOMA; UNKNOWN	P-value of test of trend or comparison	.4586		.4375	
		Number of animals reported with tumor	0	0	1	0
MAMMARY GLAND	ADENOACANTHOMA	P-value of test of trend or comparison	.4734	.7286	.2452	.7080
		Number of animals reported with tumor	1	1	3	1
	ADENOCARCINOMA	P-value of test of trend or comparison	.4594	.1883	.1602	.4173
		Number of animals reported with tumor	2	5	5	3
NASAL LEVEL IV	NEUROBLASTOMA, OLFACTORY	P-value of test of trend or comparison	.7179	.4762		
		Number of animals reported with tumor	0	1	0	0
OVARIES	ADENOMA, TUBULOSTROMAL	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
	CHORIOCARCINOMA	P-value of test of trend or comparison	.7152	.4767		
		Number of animals reported with tumor	0	1	0	0
	CYSTADENOMA	P-value of test of trend or comparison	.4147	1	1	.7019
		Number of animals reported with tumor	1	0	0	1

Table A.7

Table of reported tumors in Mouse Study
NDA 205437
Animal carcinogenicity study
Female mice

Organ or tissue name	*Tumor name*	*Quantity*	*Control Size = 70*	*Low dose Size = 70*	*Mid dose Size = 70*	*High dose Size = 70*
	GRANULOSA CELL TUMOR	P-value of test of trend or comparison	.4812	.4450	1	.6056
		Number of animals reported with tumor	2	3	0	2
	LUTEOMA	P-value of test of trend or comparison	.4571	.1042	.4375	.4512
		Number of animals reported with tumor	0	3	1	1
	SARCOMA, UNDIFFERENTIATED	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
	SERTOLI'S CELL TUMOR	P-value of test of trend or comparison	.9202	.7291	1	1
		Number of animals reported with tumor	1	1	0	0
PANCREAS	ADENOMA, ISLET CELL	P-value of test of trend or comparison	.7714	.2244		
		Number of animals reported with tumor	0	2	0	0
PITUITARY	ADENOMA, PARS DISTALIS	P-value of test of trend or comparison	.8305	.1799	.8273	.8398
		Number of animals reported with tumor	2	5	1	1
SKIN	ADENOCARCINOMA	P-value of test of trend or comparison	.7152	.4767		
		Number of animals reported with tumor	0	1	0	0
	FIBROSARCOMA	P-value of test of trend or comparison	.2342			.4512
		Number of animals reported with tumor	0	0	0	1
	SARCOMA, UNDIFFERENTIATED	P-value of test of trend or comparison	.7891	.9629	.6701	.9528
		Number of animals reported with tumor	4	1	3	1
SOFT TISSUE, ABD	NEUROENDOCRINE CELL TUMOR	P-value of test of trend or comparison	.7134	.4706		
		Number of animals reported with tumor	0	1	0	0
SPINAL CORD	MENINGIOMA	P-value of test of trend or comparison	.7152	.4767		
		Number of animals reported with tumor	0	1	0	0
	SARCOMA, UNDIFFERENTIATED	P-value of test of trend or comparison	.2342			.4512
		Number of animals reported with tumor	0	0	0	1
STOMACH, GLAN	ADENOMA	P-value of test of trend or comparison	.2342			.4512
		Number of animals reported with tumor	0	0	0	1
	NEUROENDOCRINE CELL TUMOR	P-value of test of trend or comparison	.7053	1	.6867	1

Table A.7

Table of reported tumors in Mouse Study
NDA 205437
Animal carcinogenicity study
Female mice

Organ or tissue name	*Tumor name*	*Quantity*	*Control Size = 70*	*Low dose Size = 70*	*Mid dose Size = 70*	*High dose Size = 70*
		Number of animals reported with tumor	1	0	1	0
STOMACH, NON	PAPILLOMA, SQUAMOUS CELL	P-value of test of trend or comparison	.4625		.1944	
		Number of animals reported with tumor	0	0	2	0
SYSTEMIC TUMORS	FIBROUS HISTIOCYTOMA	P-value of test of trend or comparison	.4591		.4444	
		Number of animals reported with tumor	0	0	1	0
	HEMANGIOMA	P-value of test of trend or comparison	.8041	.1042		
		Number of animals reported with tumor	0	3	0	0
	HEMANGIOSARCOMA	P-value of test of trend or comparison	.8823	.7537	.8394	.9552
		Number of animals reported with tumor	4	3	2	1
	LEUKEMIA, GRANULOCYTIC	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	2	0	0	0
	LYMPHOMA	P-value of test of trend or comparison	.6847	.1735	.1908	.5807
		Number of animals reported with tumor	13	18	16	11
	SARCOMA, HISTIOCYTIC	P-value of test of trend or comparison	.5875	.9629	.6838	.8488
		Number of animals reported with tumor	4	1	3	2
TEETH	ODONTOMA, AMELOBLASTIC	P-value of test of trend or comparison	.4591		.4444	
		Number of animals reported with tumor	0	0	1	0
THYROID GLANDS	CARCINOMA, FOLLICULAR CELL	P-value of test of trend or comparison	.4591		.4444	
		Number of animals reported with tumor	0	0	1	0
UTERUS	FIBROMA	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
	FIBROSARCOMA	P-value of test of trend or comparison	.4591		.4444	
		Number of animals reported with tumor	0	0	1	0
	GRANULAR CELL TUMOR	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
	LEIOMYOMA	P-value of test of trend or comparison	.3451	1	.6944	.7019
		Number of animals reported with tumor	1	0	1	1

Table A.7

Table of reported tumors in Mouse Study
NDA 205437
Animal carcinogenicity study
Female mice

Organ or tissue name	*Tumor name*	*Quantity*	*Control Size = 70*	*Low dose Size = 70*	*Mid dose Size = 70*	*High dose Size = 70*
	LEIOMYOSARCOMA	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
	POLYP, ENDOMETRIAL STROMAL	P-value of test of trend or comparison	.8761	.9976	.9777	.9799
		Number of animals reported with tumor	8	1	2	2
	SARCOMA, ENDOMETRIAL STROMAL	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
	SARCOMA, UNDIFFERENTIATED	P-value of test of trend or comparison	.7152	.4767		
		Number of animals reported with tumor	0	1	0	0
ZYMBAL'S GLANDS	ADENOMA	P-value of test of trend or comparison	.4610		.4474	
		Number of animals reported with tumor	0	0	1	0
	CARCINOMA	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0

Table A.8

Table of reported tumors in Mouse Study
NDA 205437
Animal carcinogenicity study
Male mice

Organ or tissue name	*Tumor name*	*Quantity*	*Control Size = 70*	*Low dose Size = 70*	*Mid dose Size = 70*	*High dose Size = 70*
ADRENAL CORTEX	ADENOMA	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	2	0	0	0
	ADENOMA, A CELL	P-value of test of trend or comparison	.9490	.9449	.9252	1
		Number of animals reported with tumor	3	1	1	0
	CARCINOMA, A CELL	P-value of test of trend or comparison	.5453	.5060	.4675	
		Number of animals reported with tumor	0	1	1	0
BONE	SCHWANNOMA	P-value of test of trend or comparison	.4474		.4750	
		Number of animals reported with tumor	0	0	1	0
COLON	ADENOCARCINOMA	P-value of test of trend or comparison	.4400		.4684	
		Number of animals reported with tumor	0	0	1	0
EPIDIDYMIDES	ADENOMA, INTERSTITIAL CELL	P-value of test of trend or comparison	.7237	.5059		
		Number of animals reported with tumor	0	1	0	0
GALLBLADDER	PAPILLOMA	P-value of test of trend or comparison	.6934	.7385	.3697	1
		Number of animals reported with tumor	1	1	2	0
HARDERIAN GLANDS	ADENOCARCINOMA	P-value of test of trend or comparison	.7219	.5000		
		Number of animals reported with tumor	0	1	0	0
	ADENOMA	P-value of test of trend or comparison	.8935	.6088	.9608	.9133
		Number of animals reported with tumor	8	8	3	3
ILEUM	ADENOCARCINOMA	P-value of test of trend or comparison	.7260	.5062		
		Number of animals reported with tumor	0	1	0	0
	ADENOMA	P-value of test of trend or comparison	.1986			.4203
		Number of animals reported with tumor	0	0	0	1
JEJUNUM	ADENOCARCINOMA	P-value of test of trend or comparison	.2061			.4286
		Number of animals reported with tumor	0	0	0	1
	ADENOMA	P-value of test of trend or comparison	.7252	.5000		
		Number of animals reported with tumor	0	1	0	0
KIDNEYS	ADENOMA, RENAL TUBULE	P-value of test of trend or comparison	.6889	1	.7142	1

Table A.8

Table of reported tumors in Mouse Study
NDA 205437
Animal carcinogenicity study
Male mice

Organ or tissue name	*Tumor name*	*Quantity*	*Control Size = 70*	*Low dose Size = 70*	*Mid dose Size = 70*	*High dose Size = 70*
		Number of animals reported with tumor	1	0	1	0
	CARCINOMA, RENAL TUBULE	P-value of test of trend or comparison	.1987			.4167
		Number of animals reported with tumor	0	0	0	1
LIVER	ADENOMA, HEPATOCELLULAR	P-value of test of trend or comparison	.6850	.7408	.9753	.7659
		Number of animals reported with tumor	11	9	4	6
	CARCINOMA, HEPATOCELLULAR	P-value of test of trend or comparison	.2804	.0906	.4229	.1978
		Number of animals reported with tumor	6	13	7	8
LUNGS	ADENOMA, BRONCHIOLO-ALVEOLAR	P-value of test of trend or comparison	.8754	.7967	.7310	.9444
		Number of animals reported with tumor	9	7	7	3
	CARCINOMA, BRONCHIOLO-ALVEOLAR	P-value of test of trend or comparison	.1574	.9377	.6795	.4165
		Number of animals reported with tumor	5	2	4	5
SEMINAL VESICLES	ADENOMA	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
SKIN	PAPILLOMA, SQUAMOUS CELL	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
	SARCOMA, UNDIFFERENTIATED	P-value of test of trend or comparison	.0677	.5059		.1702
		Number of animals reported with tumor	0	1	0	2
SYSTEMIC TUMORS	FIBROUS HISTIOCYTOMA	P-value of test of trend or comparison	.4437		.4684	
		Number of animals reported with tumor	0	0	1	0
	HEMANGIOMA	P-value of test of trend or comparison	.3755	.7471	.4532	.6564
		Number of animals reported with tumor	1	1	2	1
	HEMANGIOSARCOMA	P-value of test of trend or comparison	.7172	.8200	.7700	.8866
		Number of animals reported with tumor	3	2	2	1
	LYMPHOMA	P-value of test of trend or comparison	.9765	.9758	.9917	.9917
		Number of animals reported with tumor	14	6	4	3
	SARCOMA, HISTIOCYTIC	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	2	0	0	0

Table A.8

Table of reported tumors in Mouse Study
NDA 205437
Animal carcinogenicity study
Male mice

Organ or tissue name	*Tumor name*	*Quantity*	*Control Size = 70*	*Low dose Size = 70*	*Mid dose Size = 70*	*High dose Size = 70*
TESTES	ADENOMA, INTERSTITIAL CELL	P-value of test of trend or comparison	.2192	.5000	.4684	.4167
		Number of animals reported with tumor	0	1	1	1
THYROID GLANDS	ADENOMA, FOLLICULAR CELL	P-value of test of trend or comparison	.0762	.4940	.4684	.1769
		Number of animals reported with tumor	0	1	1	2
URINARY BLADDER	SUBMUCOSAL MESENCHYMAL TUMOR	P-value of test of trend or comparison	.5037	.5000	.7269	.6700
		Number of animals reported with tumor	1	2	1	1

Table A.9

Table of reported tumors in Mouse Study
NDA 205437
Animal carcinogenicity study
Female mice
Composite endpoints

Composite endpoint	*Quantity*	*Control Size = 70*	*Low dose Size = 70*	*Mid dose Size = 70*	*High dose Size = 70*
A cell tumors of the adrenal cortex	P-value of test of trend or comparison	.0559		.1944	.2005
	Number of animals reported with tumor	0	0	2	2
Adenocarcinomas and carcinomas of the mandibular lymph node	P-value of test of trend or comparison	.1601		.4375	.4512
	Number of animals reported with tumor	0	0	1	1
Adenomas and carcinomas of the adrenal cortex (excluding A cell tumors)	P-value of test of trend or comparison	1	1	1	1
	Number of animals reported with tumor	1	0	0	0
All cervical and uterine polyps	P-value of test of trend or comparison	.8761	.9976	.9777	.9799
	Number of animals reported with tumor	8	1	2	2
All leiomyosarcomas	P-value of test of trend or comparison	1	1	1	1
	Number of animals reported with tumor	2	0	0	0
Bronchiolo-alveolar tumors	P-value of test of trend or comparison	.9577	.6976	.9993	.9706
	Number of animals reported with tumor	17	14	3	7
Cervical and uterine endometrial stromal sarcomas	P-value of test of trend or comparison	1	1	1	1
	Number of animals reported with tumor	1	0	0	0
Cervical and vaginal carcinomas	P-value of test of trend or comparison	.7152	.4767		
	Number of animals reported with tumor	0	1	0	0
Cervical, uterine, and vaginal fibromas	P-value of test of trend or comparison	.4147	1	1	.7019
	Number of animals reported with tumor	1	0	0	1
Cervical, uterine, and vaginal fibromas and fibrosarcomas	P-value of test of trend or comparison	.3451	1	.6944	.7019
	Number of animals reported with tumor	1	0	1	1
Endometrial stromal sarcomas and polyps of the cervix and uterus	P-value of test of trend or comparison	.9151	.9989	.9874	.9887
	Number of animals reported with tumor	9	1	2	2
Fibromas and fibrosarcomas of the skin (and tail)	P-value of test of trend or comparison	.2342			.4512
	Number of animals reported with tumor	0	0	0	1
Follicular cell tumors	P-value of test of trend or comparison	.4591		.4444	
	Number of animals reported with tumor	0	0	1	0

Table A.9

Table of reported tumors in Mouse Study
NDA 205437
Animal carcinogenicity study
Female mice
Composite endpoints

Composite endpoint	*Quantity*	*Control Size = 70*	*Low dose Size = 70*	*Mid dose Size = 70*	*High dose Size = 70*
Gastrointestinal adenocarcinomas	P-value of test of trend or comparison	.7170	.4828		
	Number of animals reported with tumor	0	1	0	0
Gastrointestinal adenomas and adenocarcinomas	P-value of test of trend or comparison	.7170	.4828		
	Number of animals reported with tumor	0	1	0	0
Glial cell tumors	P-value of test of trend or comparison	1	1	1	1
	Number of animals reported with tumor	1	0	0	0
Granular cell tumors	P-value of test of trend or comparison	1	1	1	1
	Number of animals reported with tumor	2	0	0	0
Harderian gland adenomas and adenocarcinomas	P-value of test of trend or comparison	.5520	.5694	.2637	.6507
	Number of animals reported with tumor	5	5	7	4
Hemangiomas and hemangiosarcomas	P-value of test of trend or comparison	.9395	.3404	.8394	.9552
	Number of animals reported with tumor	4	6	2	1
Hepatocellular tumors	P-value of test of trend or comparison	.5474	.0781	.2522	.4049
	Number of animals reported with tumor	3	8	5	4
Histiocytoma and reticulosis	P-value of test of trend or comparison	.5908	.9629	.5150	.8488
	Number of animals reported with tumor	4	1	4	2
Internal squamous cell papillomas and carcinomas	P-value of test of trend or comparison	.4625		.1944	
	Number of animals reported with tumor	0	0	2	0
Islet cell tumors	P-value of test of trend or comparison	.7714	.2244		
	Number of animals reported with tumor	0	2	0	0
Leiomyomas and leiomyosarcomas of the uterus	P-value of test of trend or comparison	.5251	1	.8337	.8398
	Number of animals reported with tumor	2	0	1	1
Leiomyosarcomas and leiomyomas	P-value of test of trend or comparison	.6709	1	.9104	.9148
	Number of animals reported with tumor	3	0	1	1
Mammary adenoma, adenocarcinomas, and adenocanthomas	P-value of test of trend or comparison	.5566	.3052	.0657	.5643
	Number of animals reported with tumor	3	5	8	3

Table A.9

Table of reported tumors in Mouse Study
NDA 205437
Animal carcinogenicity study
Female mice
Composite endpoints

Composite endpoint	*Quantity*	*Control Size = 70*	*Low dose Size = 70*	*Mid dose Size = 70*	*High dose Size = 70*
Meningiomas and meningeal sarcomas	P-value of test of trend or comparison	.7152	.4767		
	Number of animals reported with tumor	0	1	0	0
Osteomas and osteosarcomas	P-value of test of trend or comparison	.0128			.0918
	Number of animals reported with tumor	0	0	0	3
Ovarian Sertoli cell tumors and tubulostromal adenomas	P-value of test of trend or comparison	.9780	.8613	1	1
	Number of animals reported with tumor	2	1	0	0
Ovarian luteomas and thecomas	P-value of test of trend or comparison	.4571	.1042	.4375	.4512
	Number of animals reported with tumor	0	3	1	1
Pheochromocytomas or neuroendocrine cell tumors	P-value of test of trend or comparison	.7876	.7291	.6867	1
	Number of animals reported with tumor	1	1	1	0
Sarcoma, undifferentiated in reproductive tissues	P-value of test of trend or comparison	.9202	.7291	1	1
	Number of animals reported with tumor	1	1	0	0
Sarcoma, undifferentiated of the uterus and vagina	P-value of test of trend or comparison	.7152	.4767		
	Number of animals reported with tumor	0	1	0	0
Sarcoma, undifferentiated, of the skin and paws	P-value of test of trend or comparison	.7891	.9629	.6701	.9528
	Number of animals reported with tumor	4	1	3	1
Sarcomas of the cervix, uterus, and vagina (excluding fibrosarcomas)	P-value of test of trend or comparison	.9940	.9298	1	1
	Number of animals reported with tumor	3	1	0	0
Sarcomas of the cervix, uterus, and vagina (including fibrosarcomas)	P-value of test of trend or comparison	.9561	.9298	.9104	1
	Number of animals reported with tumor	3	1	1	0
Sertoli cell tumors	P-value of test of trend or comparison	.9202	.7291	1	1
	Number of animals reported with tumor	1	1	0	0
Tumors of the pars distalis and pars intermedia	P-value of test of trend or comparison	.8305	.1799	.8273	.8398
	Number of animals reported with tumor	2	5	1	1
Uterine fibromas and fibrosarcomas	P-value of test of trend or comparison	.7090	1	.6944	1
	Number of animals reported with tumor	1	0	1	0

Table A.9

Table of reported tumors in Mouse Study
NDA 205437
Animal carcinogenicity study
Female mice
Composite endpoints

Composite endpoint	*Quantity*	*Control Size = 70*	*Low dose Size = 70*	*Mid dose Size = 70*	*High dose Size = 70*
Zymbal glands tumors	P-value of test of trend or comparison	.7111	1	.6979	1
	Number of animals reported with tumor	1	0	1	0

Table A.10

Table of reported tumors in Mouse Study
NDA 205437
Animal carcinogenicity study
Male mice
Composite endpoints

Composite endpoint	*Quantity*	*Control Size = 70*	*Low dose Size = 70*	*Mid dose Size = 70*	*High dose Size = 70*
A cell tumors of the adrenal cortex	P-value of test of trend or comparison	.9591	.8277	.9252	1
	Number of animals reported with tumor	3	2	1	0
Adenomas and carcinomas of the adrenal cortex (excluding A cell tumors)	P-value of test of trend or comparison	1	1	1	1
	Number of animals reported with tumor	2	0	0	0
All schwannomas	P-value of test of trend or comparison	.4474		.4750	
	Number of animals reported with tumor	0	0	1	0
Bronchiolo-alveolar tumors	P-value of test of trend or comparison	.6049	.9460	.8065	.8210
	Number of animals reported with tumor	14	8	10	8
Follicular cell tumors	P-value of test of trend or comparison	.0762	.4940	.4684	.1769
	Number of animals reported with tumor	0	1	1	2
Gastrointestinal adenocarcinomas	P-value of test of trend or comparison	.2192	.5000	.4684	.4167
	Number of animals reported with tumor	0	1	1	1
Gastrointestinal adenomas	P-value of test of trend or comparison	.2477	.5000		.4167
	Number of animals reported with tumor	0	1	0	1
Gastrointestinal adenomas and adenocarcinomas	P-value of test of trend or comparison	.1218	.2470	.4684	.1702
	Number of animals reported with tumor	0	2	1	2
Harderian gland adenomas and adenocarcinomas	P-value of test of trend or comparison	.9107	.5000	.9608	.9133
	Number of animals reported with tumor	8	9	3	3
Hemangiomas and hemangiosarcomas	P-value of test of trend or comparison	.5977	.7834	.5717	.7945
	Number of animals reported with tumor	4	3	4	2
Hepatocellular tumors	P-value of test of trend or comparison	.6165	.3180	.9228	.5737
	Number of animals reported with tumor	16	20	9	12
Histiocytoma and reticulosis	P-value of test of trend or comparison	.8479	1	.8498	1
	Number of animals reported with tumor	2	0	1	0
Renal tubule tumors	P-value of test of trend or comparison	.2988	1	.7142	.6564
	Number of animals reported with tumor	1	0	1	1

Table A.10

Table of reported tumors in Mouse Study
NDA 205437
Animal carcinogenicity study
Male mice
Composite endpoints

Composite endpoint	*Quantity*	*Control Size = 70*	*Low dose Size = 70*	*Mid dose Size = 70*	*High dose Size = 70*
Sarcoma, undifferentiated, of the skin and paws	P-value of test of trend or comparison	.0677	.5059		.1702
	Number of animals reported with tumor	0	1	0	2
Squamous cell papillomas of the skin and tail	P-value of test of trend or comparison	1	1	1	1
	Number of animals reported with tumor	1	0	0	0

Table A.11

Table of tumors reported significant (alpha < 0.05) in at least one arm - Mouse Study
NDA 205437
Animal carcinogenicity study
Female mice
Composite endpoints

Composite endpoint	*Quantity*	*Control*	*Low dose*	*Mid dose*	*High dose*
Osteomas and osteosarcomas	P-value of test of trend or comparison	.0128			.0918
	Number of animals reported with tumor	0	0	0	3
	Poly-3 adjusted incidence rate	0.0%	0.0%	0.0%	7.9%
	95% CI for poly-3 adjusted incidence rate (%)	(0,7.9)	(0,8.6)	(0,10.0)	(1.62,21.4)
	Poly-3 adjusted number of animals at risk	45.0	41.2	35.7	38.1

A.3 Unexamined and autolytic organs

Table A.12

Organs reported as autolytic
NDA 205437
Animal carcinogenicity study
Female Mice

Organ or tissue name	*Control(count)*	*Control(%)*	*Low dose(count)*	*Low dose(%)*	*Mid dose(count)*	*Mid dose(%)*	*High dose(count)*	*High dose(%)*	*Total(count)*	*Total(%)*
CECUM	3	4.3%	3	4.3%	.	.	3	4.3%	9	3.2%
COLON	1	1.4%	.	.	.	.	1	1.4%	2	0.7%
DUODENUM	3	4.3%	3	4.3%	3	4.3%	2	2.9%	11	3.9%
GALLBLADDER	18	26%	18	26%	20	29%	21	30%	77	28%
ILEUM	3	4.3%	1	1.4%	1	1.4%	2	2.9%	7	2.5%
JEJUNUM	11	16%	4	5.7%	5	7.1%	3	4.3%	23	8.2%
MAMMARY GLAND	.	.	1	1.4%	.	.	.	.	1	0.4%
PEYER'S PATCHES	.	.	.	.	1	1.4%	2	2.9%	3	1.1%
SPINAL CORD	.	.	.	.	1	1.4%	.	.	1	0.4%
URINARY BLADDER	1	1.4%	1	1.4%	.	.	.	.	2	0.7%

Table A.13

Organs reported as autolytic
NDA 205437
Animal carcinogenicity study
Male Mice

Organ or tissue name	*Control(count)*	*Control(%)*	*Low dose(count)*	*Low dose(%)*	*Mid dose(count)*	*Mid dose(%)*	*High dose(count)*	*High dose(%)*	*Total(count)*	*Total(%)*
ADRENAL CORTEX	.	.	1	1.4%	.	.	.	.	1	0.4%
ADRENAL MEDULLA	.	.	1	1.4%	.	.	.	.	1	0.4%
BRAIN	.	.	1	1.4%	.	.	.	.	1	0.4%
CECUM	5	7.1%	12	17%	11	16%	4	5.7%	32	11%
COLON	1	1.4%	.	.	.	.	1	1.4%	2	0.7%
DUODENUM	13	19%	10	14%	9	13%	7	10%	39	14%
GALLBLADDER	17	24%	24	34%	28	40%	22	31%	91	33%
ILEUM	4	5.7%	2	2.9%	2	2.9%	3	4.3%	11	3.9%
JEJUNUM	10	14%	11	16%	10	14%	8	11%	39	14%
LAC. GLAND EXOR	.	.	1	1.4%	.	.	.	.	1	0.4%
LARYNX	.	.	1	1.4%	.	.	.	.	1	0.4%
LYMPH NODE, MAND	.	.	1	1.4%	.	.	.	.	1	0.4%
LYMPH NODE, MES	1	1.4%	1	1.4%	.	.	1	1.4%	3	1.1%
MARROW, STERN	.	.	1	1.4%	.	.	.	.	1	0.4%
NASAL LEVEL I	.	.	1	1.4%	.	.	.	.	1	0.4%
NASAL LEVEL II	.	.	1	1.4%	.	.	.	.	1	0.4%
NASAL LEVEL III	.	.	1	1.4%	.	.	.	.	1	0.4%
NASAL LEVEL IV	.	.	1	1.4%	.	.	.	.	1	0.4%
PANCREAS	.	.	1	1.4%	1	1.4%	.	.	2	0.7%
PARATHYROIDS	.	.	1	1.4%	.	.	.	.	1	0.4%
PEYER'S PATCHES	1	1.4%	6	8.6%	4	5.7%	2	2.9%	13	4.6%
PITUITARY	.	.	1	1.4%	.	.	.	.	1	0.4%
PREPUTIAL GLANDS	.	.	1	1.4%	.	.	.	.	1	0.4%
PROSTATE	.	.	1	1.4%	.	.	.	.	1	0.4%
SAL. GLAND MAND	.	.	1	1.4%	.	.	.	.	1	0.4%
SEMINAL VESICLES	1	1.4%	.	.	.	.	.	.	1	0.4%
SPINAL CORD	.	.	1	1.4%	.	.	.	.	1	0.4%

Table A.13

Organs reported as autolytic
NDA 205437
Animal carcinogenicity study
Male Mice

Organ or tissue name	*Control(count)*	*Control(%)*	*Low dose(count)*	*Low dose(%)*	*Mid dose(count)*	*Mid dose(%)*	*High dose(count)*	*High dose(%)*	*Total(count)*	*Total(%)*
STOMACH, GLAN	1	1.4%	1	1.4%	1	1.4%	1	1.4%	4	1.4%
STOMACH, NON	1	1.4%	.	.	.	.	.	.	1	0.4%
THYROID GLANDS	.	.	2	2.9%	.	.	.	.	2	0.7%
TONGUE	.	.	1	1.4%	.	.	.	.	1	0.4%
TRACHEA	1	1.4%	1	1.4%	.	.	.	.	2	0.7%
URETERS	.	.	1	1.4%	.	.	.	.	1	0.4%
URINARY BLADDER	2	2.9%	2	2.9%	.	.	1	1.4%	5	1.8%
ZYMBAL'S GLANDS	.	.	1	1.4%	.	.	.	.	1	0.4%

Table A.14

Organs reported as unexamined
NDA 205437
Animal carcinogenicity study
Female Mice

Organ or tissue name	*Control(count)*	*Control(%)*	*Low dose(count)*	*Low dose(%)*	*Mid dose(count)*	*Mid dose(%)*	*High dose(count)*	*High dose(%)*	*Total(count)*	*Total(%)*
ADRENAL MEDULLA	1	1.4%	.	.	1	1.4%	.	.	2	0.7%
AORTA	2	2.9%	.	.	1	1.4%	.	.	3	1.1%
CERVIX	2	2.9%	.	.	.	.	.	.	2	0.7%
CLITORAL GLANDS	1	1.4%	2	2.9%	.	.	1	1.4%	4	1.4%
LAC. GLAND EXOR	1	1.4%	1	1.4%	.	.	1	1.4%	3	1.1%
LARYNX	1	1.4%	2	2.9%	1	1.4%	1	1.4%	5	1.8%
LYMPH NODE, MAND	.	.	1	1.4%	.	.	.	.	1	0.4%
LYMPH NODE, MED	.	.	.	.	.	.	3	4.3%	3	1.1%
LYMPH NODE, MES	.	.	1	1.4%	1	1.4%	2	2.9%	4	1.4%
LYMPH NODE, TR/B	.	.	1	1.4%	.	.	.	.	1	0.4%
MAMMARY GLAND	1	1.4%	.	.	.	.	1	1.4%	2	0.7%
NASAL LEVEL IV	1	1.4%	1	1.4%	1	1.4%	.	.	3	1.1%
NERVE, SCIATIC	.	.	.	.	.	.	1	1.4%	1	0.4%
OVIDUCTS	4	5.7%	.	.	1	1.4%	2	2.9%	7	2.5%
PARATHYROIDS	16	23%	20	29%	16	23%	20	29%	72	26%
PEYER'S PATCHES	26	37%	27	39%	22	31%	29	41%	104	37%
PITUITARY	1	1.4%	.	.	.	.	.	.	1	0.4%
SAL. GLAND MAND	.	.	1	1.4%	.	.	.	.	1	0.4%
SOFT TISSUE, ABD	.	.	1	1.4%	1	1.4%	.	.	2	0.7%
SPLEEN	.	.	.	.	.	.	1	1.4%	1	0.4%
THORACIC CAVITY	.	.	.	.	.	.	1	1.4%	1	0.4%
THYMUS	3	4.3%	4	5.7%	4	5.7%	2	2.9%	13	4.6%
TRACHEA	.	.	1	1.4%	.	.	.	.	1	0.4%
URINARY BLADDER	.	.	.	.	1	1.4%	.	.	1	0.4%
VAGINA	1	1.4%	1	1.4%	1	1.4%	.	.	3	1.1%
ZYMBAL'S GLANDS	4	5.7%	2	2.9%	4	5.7%	2	2.9%	12	4.3%

Table A.15

Organs reported as unexamined
NDA 205437
Animal carcinogenicity study
Male Mice

Organ or tissue name	*Control(count)*	*Control(%)*	*Low dose(count)*	*Low dose(%)*	*Mid dose(count)*	*Mid dose(%)*	*High dose(count)*	*High dose(%)*	*Total(count)*	*Total(%)*
ADRENAL CORTEX	1	1.4%	.	.	1	1.4%	2	2.9%	4	1.4%
ADRENAL MEDULLA	1	1.4%	.	.	2	2.9%	2	2.9%	5	1.8%
AORTA	.	.	2	2.9%	1	1.4%	1	1.4%	4	1.4%
BILE DUCT	.	.	.	.	1	1.4%	.	.	1	0.4%
GALLBLADDER	1	1.4%	2	2.9%	1	1.4%	.	.	4	1.4%
LAC. GLAND EXOR	2	2.9%	1	1.4%	1	1.4%	.	.	4	1.4%
LARYNX	.	.	1	1.4%	1	1.4%	.	.	2	0.7%
LYMPH NODE, MAND	1	1.4%	1	1.4%	2	2.9%	1	1.4%	5	1.8%
LYMPH NODE, MES	1	1.4%	2	2.9%	5	7.1%	.	.	8	2.9%
LYMPH NODE, POP	1	1.4%	.	.	.	.	1	1.4%	2	0.7%
LYMPH NODE, REN	.	.	1	1.4%	.	.	.	.	1	0.4%
MAMMARY GLAND	1	1.4%	1	1.4%	4	5.7%	8	11%	14	5.0%
NASAL LEVEL III	.	.	.	.	1	1.4%	.	.	1	0.4%
NASAL LEVEL IV	3	4.3%	.	.	2	2.9%	4	5.7%	9	3.2%
PARATHYROIDS	16	23%	22	31%	27	39%	23	33%	88	31%
PEYER'S PATCHES	26	37%	27	39%	23	33%	19	27%	95	34%
PHARYNX	1	1.4%	.	.	.	.	.	.	1	0.4%
PITUITARY	1	1.4%	.	.	.	.	.	.	1	0.4%
PREPUTIAL GLANDS	2	2.9%	2	2.9%	2	2.9%	2	2.9%	8	2.9%
PROSTATE	.	.	.	.	1	1.4%	.	.	1	0.4%
RECTUM	.	.	.	.	1	1.4%	2	2.9%	3	1.1%
SOFT TISSUE, THO	1	1.4%	.	.	.	.	.	.	1	0.4%
SPLEEN	.	.	1	1.4%	1	1.4%	.	.	2	0.7%
STOMACH, NON	1	1.4%	.	.	.	.	.	.	1	0.4%
THYMUS	6	8.6%	8	11%	3	4.3%	5	7.1%	22	7.9%
TRACHEA	.	.	.	.	1	1.4%	.	.	1	0.4%

Table A.15

Organs reported as unexamined
NDA 205437
Animal carcinogenicity study
Male Mice

Organ or tissue name	*Control(count)*	*Control(%)*	*Low dose(count)*	*Low dose(%)*	*Mid dose(count)*	*Mid dose(%)*	*High dose(count)*	*High dose(%)*	*Total(count)*	*Total(%)*
URINARY BLADDER	.	.	1	1.4%	.	.	.	.	1	0.4%
ZYMBAL'S GLANDS	9	13%	6	8.6%	5	7.1%	5	7.1%	25	8.9%

Appendix B

Tables from rat study

B.1 Survival analysis

Table B.1

Survival rates at key times
NDA 205437
Animal carcinogenicity study
Rats

Species and Sex	*Dose Group*	*Dose (mg per kg)*	*Number at start*	*Number alive after 52 weeks*	*Percentage alive after 52 weeks*	*Number alive after 78 weeks*	*Percentage alive after 78 weeks*	*Number alive after 90 weeks*	*Percentage alive after 90 weeks*	*Number sacrificed*	*Percentage sacrificed*	*Maximum surivial (weeks)*
Rats - Female	Control	0	70	68	97%	52	74%	38	54%	21	30%	105
	Low dose	0.3	70	69	99%	53	76%	39	56%	19	27%	105
	Mid dose	0.97	70	65	93%	44	63%	32	46%	18	26%	105
	High dose	2.71	70	61	87%	40	57%	30	43%	16	23%	105
Rats - Male	Control	0	70	65	93%	45	64%	31	44%	18	26%	101
	Low dose	2.73	70	65	93%	37	53%	23	33%	16	23%	101
	Mid dose	8.14	70	56	80%	33	47%	18	26%	15	21%	98
	High dose	13.89	70	49	70%	25	36%	20	29%	15	21%	96

Table B.2

Log-rank tests of survival
NDA 205437
Animal carcinogenicity study
Rats

Sex	*Test of homogeneity: chi squared statistic*	*Test of homogeneity: degrees of freedom*	*Number of groups*	*Test of homogeneity: p-value*	*Test of trend (two tailed): p-value*	*Test of trend (one tailed): p-value*
Female	3.5967	3	4	0.3084	0.0611	0.0305
Male	9.5108	3	4	0.0232	0.0022	0.0011

Table B.3

Pairwise comparisons (log-rank) of survival between treated groups and controls
NDA 205437
Animal carcinogenicity study
Rats

Species and Sex	*Quantity*	*Low dose*	*Mid dose*	*High dose*
Rats - Female	Chi squared test statistic	0.0007	0.5388	2.5403
	p-value of comparison with control	0.9791	0.4629	0.1110
Rats - Male	Chi squared test statistic	0.8788	3.1688	8.5113
	p-value of comparison with control	0.3485	0.0751	0.0035

B.2 Tumor analysis

Table B.4

Primary organs in study of female rats
NDA 205437
Animal carcinogenicity study

Organ or tissue name
ADRENAL CORTEX
ADRENAL MEDULLA
BILE DUCT
BONE
BRAIN
CECUM
CERVIX
CLITORAL GLANDS
DUODENUM
EYES/OPTIC N.
GINGIVA
ILEUM
JEJUNUM
KIDNEYS
LARYNX
LIVER
LUNGS
LYMPH NODE, MAND
LYMPH NODE, MES
MAMMARY GLAND
MARROW, FEMUR
MARROW, STERN
MESENTERY
NASAL LEVEL II
NERVE, SCIATIC
OVARIES
OVIDUCTS
PANCREAS
PARATHYRO DS
PAWS
PEYER'S PATCHES
PHARYNX
PITUITARY
RECTUM
SKELETAL MUSCLE
SKIN
SOFT TISSUE- ABD
SOFT TISSUE- OC

Primary organs in study of female rats
NDA 205437
Animal carcinogenicity study

Organ or tissue name
SOFT TISSUE- THO
SPINAL CORD
SPLEEN
STERNUM
STOMACH, NON
SYSTEMIC TUMORS
TAIL
THYMUS
THYROID GLANDS
TRACHEA
UTERUS
VAGINA
ZYMBAL'S GLANDS

Table B.5

Primary organs in study of male rats
NDA 205437
Animal carcinogenicity study

Organ or tissue name
ADIPOSE TISSUE
ADRENAL CORTEX
ADRENAL MEDULLA
BILE DUCT
BONE
BRAIN
CECUM
COLON
DUODENUM
EARS
ESOPHAGUS
EYES/OPTIC N.
HEART
ILEUM
JEJUNUM
KIDNEYS
LIVER
LUNGS
LYMPH NODE, MAND
LYMPH NODE, MES
LYMPH NODE, REN
MAMMARY GLAND
MARROW, STERN
NASAL LEVEL III
NASAL LEVEL IV
PANCREAS
PARATHYRO DS
PAWS
PENIS
PEYER'S PATCHES
PITUITARY
PREPUTIAL GLANDS
PROSTATE
RECTUM
SAL. GLAND MAND
SEMINAL VESICLES
SKELETAL MUSCLE
SKIN

Primary organs in study of male rats
NDA 205437
Animal carcinogenicity study

Organ or tissue name
SOFT TISSUE- ABD
SOFT TISSUE- THO
SPLEEN
STERNUM
STOMACH, GLAN
STOMACH, NON
SYSTEMIC TUMORS
TAIL
TESTES
THYMUS
THYROID GLANDS
TRACHEA
URINARY BLADDER
ZYMBAL'S GLANDS

Table B.6

Table of reported tumors in Rat Study
NDA 205437
Animal carcinogenicity study
Female rats

Organ or tissue name	*Tumor name*	*Quantity*	*Control Size = 70*	*Low dose Size = 70*	*Mid dose Size = 70*	*High dose Size = 70*
ADRENAL CORTEX	ADENOMA	P-value of test of trend or comparison	.1428	.8389	1	.4049
		Number of animals reported with tumor	3	2	0	4
	CARCINOMA	P-value of test of trend or comparison	.4651		.4886	
		Number of animals reported with tumor	0	0	1	0
ADRENAL MEDULLA	PHEOCHROMOCYTOMA	P-value of test of trend or comparison	.9901	.9877	.9827	1
		Number of animals reported with tumor	5	1	1	0
BONE	OSTEOSARCOMA	P-value of test of trend or comparison	.7384	.5109		
		Number of animals reported with tumor	0	1	0	0
BRAIN	ASTROCYTOMA	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	2	0	0	0
	GRANULAR CELL TUMOR	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	3	0	0	0
CERVIX	FIBROMA	P-value of test of trend or comparison	.4651		.4886	
		Number of animals reported with tumor	0	0	1	0
	GRANULAR CELL TUMOR	P-value of test of trend or comparison	.8367	.6816	.9332	.9154
		Number of animals reported with tumor	3	3	1	1
	POLYP	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
	POLYP, ENDOMETRIAL STROMAL	P-value of test of trend or comparison	.7384	.5109		
		Number of animals reported with tumor	0	1	0	0
	SARCOMA, ENDOMETRIAL STROMAL	P-value of test of trend or comparison	.9304	.7581	1	1
		Number of animals reported with tumor	1	1	0	0
	SCHWANNOMA	P-value of test of trend or comparison	.4792	.7581	1	.7031
		Number of animals reported with tumor	1	1	0	1
EYES/OPTIC N.	MELANOMA, AMELANOTIC	P-value of test of trend or comparison	.2717	.5109		.4512
		Number of animals reported with tumor	0	1	0	1
GINGIVA	CARCINOMA, SQUAMOUS CELL	P-value of test of trend or comparison	.5838	.5109	.4886	

Table B.6

Table of reported tumors in Rat Study
NDA 205437
Animal carcinogenicity study
Female rats

Organ or tissue name	*Tumor name*	*Quantity*	*Control Size = 70*	*Low dose Size = 70*	*Mid dose Size = 70*	*High dose Size = 70*
		Number of animals reported with tumor	0	1	1	0
JEJUNUM	ADENOCARCINOMA	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
KIDNEYS	CARCINOMA, RENAL TUBULE	P-value of test of trend or comparison	.4651		.4886	
		Number of animals reported with tumor	0	0	1	0
	LIPOMA	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
LIVER	CARCINOMA, HEPATOCELLULAR	P-value of test of trend or comparison	.7384	.5109		
		Number of animals reported with tumor	0	1	0	0
LUNGS	ADENOMA, BRONCHIOLO-ALVEOLAR	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
MAMMARY GLAND	ADENOCARCINOMA	P-value of test of trend or comparison	.9925	.6381	.7823	.9952
		Number of animals reported with tumor	15	15	12	4
	ADENOMA	P-value of test of trend or comparison	.3030	1	.4828	.7019
		Number of animals reported with tumor	1	0	2	1
	FIBROADENOMA	P-value of test of trend or comparison	.9247	.9759	.9885	.9828
		Number of animals reported with tumor	27	19	16	14
MESENTERY	LIPOMA	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
NASAL LEVEL II	CARCINOMA, SQUAMOUS CELL	P-value of test of trend or comparison	.7384	.5109		
		Number of animals reported with tumor	0	1	0	0
OVARIES	GRANULAR CELL TUMOR	P-value of test of trend or comparison	.7384	.5109		
		Number of animals reported with tumor	0	1	0	0
	GRANULOSA CELL TUMOR	P-value of test of trend or comparison	.4651		.4886	
		Number of animals reported with tumor	0	0	1	0
	SERTOLI'S CELL TUMOR	P-value of test of trend or comparison	.1535		.4886	.4512
		Number of animals reported with tumor	0	0	1	1

Table B.6

Table of reported tumors in Rat Study
NDA 205437
Animal carcinogenicity study
Female rats

Organ or tissue name	*Tumor name*	*Quantity*	*Control Size = 70*	*Low dose Size = 70*	*Mid dose Size = 70*	*High dose Size = 70*
	THECOMA	P-value of test of trend or comparison	.7384	.5109		
		Number of animals reported with tumor	0	1	0	0
PANCREAS	ADENOMA, ISLET CELL	P-value of test of trend or comparison	.8910	.5122	1	.9154
		Number of animals reported with tumor	3	4	0	1
	CARCINOMA, ISLET CELL	P-value of test of trend or comparison	.5702	.5245	1	.7091
		Number of animals reported with tumor	1	2	0	1
PARATHYROIDS	ADENOMA	P-value of test of trend or comparison	.3865	1	1	.7057
		Number of animals reported with tumor	1	0	0	1
PITUITARY	ADENOMA, PARS DISTALIS	P-value of test of trend or comparison	.7430	.1777	.6972	.6495
		Number of animals reported with tumor	56	66	59	54
	ADENOMA, PARS INTERMEDIA	P-value of test of trend or comparison	.2197			.4578
		Number of animals reported with tumor	0	0	0	1
	CARCINOMA, PARS DISTALIS	P-value of test of trend or comparison	.2771	.5109		.4578
		Number of animals reported with tumor	0	1	0	1
RECTUM	ADENOMA	P-value of test of trend or comparison	.7384	.5109		
		Number of animals reported with tumor	0	1	0	0
SKIN	BASAL CELL TUMOR	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	2	0	0	0
	FIBROMA	P-value of test of trend or comparison	.7883	.2637		
		Number of animals reported with tumor	0	2	0	0
	FIBROSARCOMA	P-value of test of trend or comparison	.3849	1	1	.7019
		Number of animals reported with tumor	1	0	0	1
	KERATOACANTHOMA	P-value of test of trend or comparison	.5838	.5109	.4886	
		Number of animals reported with tumor	0	1	1	0
	LIPOMA	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
	MAST CELL TUMOR	P-value of test of trend or comparison	.4651		.4886	

Table B.6

Table of reported tumors in Rat Study
NDA 205437
Animal carcinogenicity study
Female rats

Organ or tissue name	*Tumor name*	*Quantity*	*Control Size = 70*	*Low dose Size = 70*	*Mid dose Size = 70*	*High dose Size = 70*
		Number of animals reported with tumor	0	0	1	0
	PAPILLOMA, SQUAMOUS CELL	P-value of test of trend or comparison	.7384	.5109		
		Number of animals reported with tumor	0	1	0	0
	SARCOMA, UNDIFFERENTIATED	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
SOFT TISSUE- ABD	SCHWANNOMA	P-value of test of trend or comparison	.4678		.4886	
		Number of animals reported with tumor	0	0	1	0
SOFT TISSUE- THO	HIBERNOMA	P-value of test of trend or comparison	.2398	.9415	.9302	.5688
		Number of animals reported with tumor	3	1	1	3
SPINAL CORD	GLIOMA, ANAPLASTIC	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
SPLEEN	SERTOLI'S CELL TUMOR	P-value of test of trend or comparison	.2151			.4512
		Number of animals reported with tumor	0	0	0	1
STOMACH, NON	CARCINOMA, SQUAMOUS CELL	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
SYSTEMIC TUMORS	HEMANGIOSARCOMA	P-value of test of trend or comparison	.7384	.5109		
		Number of animals reported with tumor	0	1	0	0
	LYMPHOMA	P-value of test of trend or comparison	.8984	.8790	.8620	1
		Number of animals reported with tumor	2	1	1	0
	SARCOMA, HISTIOCYTIC	P-value of test of trend or comparison	.7889	.2582		
		Number of animals reported with tumor	0	2	0	0
THYROID GLANDS	ADENOMA, C-CELL	P-value of test of trend or comparison	.9982	.7514	.9435	1
		Number of animals reported with tumor	7	6	3	0
	ADENOMA, FOLLICULAR CELL	P-value of test of trend or comparison	.4828	.7635	1	.7091
		Number of animals reported with tumor	1	1	0	1
	CARCINOMA, C-CELL	P-value of test of trend or comparison	.2151			.4512
		Number of animals reported with tumor	0	0	0	1

Table B.6

Table of reported tumors in Rat Study
NDA 205437
Animal carcinogenicity study
Female rats

Organ or tissue name	*Tumor name*	*Quantity*	*Control Size = 70*	*Low dose Size = 70*	*Mid dose Size = 70*	*High dose Size = 70*
	CARCINOMA, FOLLICULAR CELL	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
UTERUS	FIBROMA	P-value of test of trend or comparison	.4651		.4886	
		Number of animals reported with tumor	0	0	1	0
	GRANULAR CELL TUMOR	P-value of test of trend or comparison	.4651		.4886	
		Number of animals reported with tumor	0	0	1	0
	POLYP, ENDOMETRIAL STROMAL	P-value of test of trend or comparison	.4278	.6298	.3554	.5716
		Number of animals reported with tumor	7	7	9	6
	SARCOMA, ENDOMETRIAL STROMAL	P-value of test of trend or comparison	.1571		.4886	.4578
		Number of animals reported with tumor	0	0	1	1
VAGINA	CARCINOMA, SQUAMOUS CELL	P-value of test of trend or comparison	.4651		.4886	
		Number of animals reported with tumor	0	0	1	0
	FIBROMA	P-value of test of trend or comparison	.7864	.2582		
		Number of animals reported with tumor	0	2	0	0
	GRANULAR CELL TUMOR	P-value of test of trend or comparison	.7153	.1944	.7414	.7091
		Number of animals reported with tumor	1	4	1	1
	SARCOMA, UNDIFFERENTIATED	P-value of test of trend or comparison	.7384	.5161		
		Number of animals reported with tumor	0	1	0	0
ZYMBAL'S GLANDS	CARCINOMA	P-value of test of trend or comparison	.2189			.4568
		Number of animals reported with tumor	0	0	0	1

Table B.7

Table of reported tumors in Rat Study
NDA 205437
Animal carcinogenicity study
Male rats

Organ or tissue name	*Tumor name*	*Quantity*	*Control Size = 70*	*Low dose Size = 70*	*Mid dose Size = 70*	*High dose Size = 70*
ADIPOSE TISSUE	LIPOMA	P-value of test of trend or comparison	.1890			.3810
		Number of animals reported with tumor	0	0	0	1
ADRENAL CORTEX	CARCINOMA	P-value of test of trend or comparison	.6953	.4730		
		Number of animals reported with tumor	0	1	0	0
ADRENAL MEDULLA	PHEOCHROMOCYTOMA	P-value of test of trend or comparison	.7719	.3979	.2904	.9140
		Number of animals reported with tumor	9	10	10	3
BONE	CHONDROSARCOMA	P-value of test of trend or comparison	.6953	.4730		
		Number of animals reported with tumor	0	1	0	0
	OSTEOMA	P-value of test of trend or comparison	.6953	.4730		
		Number of animals reported with tumor	0	1	0	0
BRAIN	ASTROCYTOMA	P-value of test of trend or comparison	.2038	.4658	.4348	.3906
		Number of animals reported with tumor	0	1	1	1
	GRANULAR CELL TUMOR	P-value of test of trend or comparison	.6775	.1010	.4348	
		Number of animals reported with tumor	0	3	1	0
	MENINGIOMA	P-value of test of trend or comparison	.2831		.1925	
		Number of animals reported with tumor	0	0	2	0
	OLIGODENDROGLIOMA	P-value of test of trend or comparison	.4252		.4348	
		Number of animals reported with tumor	0	0	1	0
	RETICULOSIS	P-value of test of trend or comparison	.1953			.3906
		Number of animals reported with tumor	0	0	0	1
	SARCOMA, MENINGEAL	P-value of test of trend or comparison	.1953			.3906
		Number of animals reported with tumor	0	0	0	1
EARS	PAPILLOMA, SQUAMOUS CELL	P-value of test of trend or comparison	.3946	1	1	.6131
		Number of animals reported with tumor	1	0	0	1
ESOPHAGUS	PAPILLOMA, SQUAMOUS CELL	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
EYES/OPTIC N.	MELANOMA, AMELANOTIC	P-value of test of trend or comparison	1	1	1	1

Table B.7

Table of reported tumors in Rat Study
NDA 205437
Animal carcinogenicity study
Male rats

Organ or tissue name	*Tumor name*	*Quantity*	*Control Size = 70*	*Low dose Size = 70*	*Mid dose Size = 70*	*High dose Size = 70*
		Number of animals reported with tumor	1	0	0	0
HEART	PARAGANGLIOMA	P-value of test of trend or comparison	.4252		.4348	
		Number of animals reported with tumor	0	0	1	0
	SCHWANNOMA	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
JEJUNUM	ADENOMA	P-value of test of trend or comparison	.7103	.4918		
		Number of animals reported with tumor	0	1	0	0
	LEIOMYOSARCOMA	P-value of test of trend or comparison	.4299		.4655	
		Number of animals reported with tumor	0	0	1	0
KIDNEYS	ADENOMA, RENAL TUBULE	P-value of test of trend or comparison	.6715	1	.6841	1
		Number of animals reported with tumor	1	0	1	0
	CARCINOMA, RENAL TUBULE	P-value of test of trend or comparison	.6953	.4730		
		Number of animals reported with tumor	0	1	0	0
	LIPOMA	P-value of test of trend or comparison	.6250	.2203	.4348	
		Number of animals reported with tumor	0	2	1	0
	RENAL MESENCHYMAL TUMOR	P-value of test of trend or comparison	.8276	1	.8195	1
		Number of animals reported with tumor	2	0	1	0
LIVER	ADENOMA, HEPATOCELLULAR	P-value of test of trend or comparison	.3304	.2203		.3810
		Number of animals reported with tumor	0	2	0	1
	CARCINOMA, HEPATOCELLULAR	P-value of test of trend or comparison	.4565	.8113	.2325	.8268
		Number of animals reported with tumor	5	3	7	2
LUNGS	ADENOMA, BRONCHIOLO-ALVEOLAR	P-value of test of trend or comparison	.9040	.7112	1	1
		Number of animals reported with tumor	1	1	0	0
	HIBERNOMA ; UNDETERMINED	P-value of test of trend or comparison	.4252		.4348	
		Number of animals reported with tumor	0	0	1	0
MAMMARY GLAND	ADENOCARCINOMA	P-value of test of trend or comparison	.9092	.7309	1	1
		Number of animals reported with tumor	1	1	0	0

Table B.7

Table of reported tumors in Rat Study
NDA 205437
Animal carcinogenicity study
Male rats

Organ or tissue name	*Tumor name*	*Quantity*	*Control Size = 70*	*Low dose Size = 70*	*Mid dose Size = 70*	*High dose Size = 70*
	ADENOMA	P-value of test of trend or comparison	.6957	.4776		
		Number of animals reported with tumor	0	1	0	0
	FIBROADENOMA	P-value of test of trend or comparison	.3969	1	1	.6272
		Number of animals reported with tumor	1	0	0	1
PANCREAS	ADENOMA, ACINAR CELL	P-value of test of trend or comparison	.2925	1	.6841	.6206
		Number of animals reported with tumor	1	0	1	1
	ADENOMA, ISLET CELL	P-value of test of trend or comparison	.6665	.7765	.2224	1
		Number of animals reported with tumor	3	2	5	0
	CARCINOMA, ACINAR CELL	P-value of test of trend or comparison	.6953	.4730		
		Number of animals reported with tumor	0	1	0	0
	CARCINOMA, ISLET CELL	P-value of test of trend or comparison	.7556	.7765	.9003	.8562
		Number of animals reported with tumor	3	2	1	1
PARATHYROIDS	ADENOMA	P-value of test of trend or comparison	.1136	1	1	.3015
		Number of animals reported with tumor	1	0	0	2
PAWS	SARCOMA, UNDIFFERENTIATED	P-value of test of trend or comparison	.4252		.4348	
		Number of animals reported with tumor	0	0	1	0
PITUITARY	ADENOMA, PARS DISTALIS	P-value of test of trend or comparison	.9383	.4248	.9193	.9024
		Number of animals reported with tumor	43	44	27	26
	ADENOMA, PARS INTERMEDIA	P-value of test of trend or comparison	.4252		.4348	
		Number of animals reported with tumor	0	0	1	0
SAL. GLAND MAND	SCHWANNOMA	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
SEMINAL VESICLES	CARCINOMA, SQUAMOUS CELL	P-value of test of trend or comparison	.6953	.4730		
		Number of animals reported with tumor	0	1	0	0
	SCHWANNOMA	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
SKELETAL MUSCLE	SCHWANNOMA	P-value of test of trend or comparison	.7402	.2270		

Table B.7

Table of reported tumors in Rat Study
NDA 205437
Animal carcinogenicity study
Male rats

Organ or tissue name	*Tumor name*	*Quantity*	*Control Size = 70*	*Low dose Size = 70*	*Mid dose Size = 70*	*High dose Size = 70*
		Number of animals reported with tumor	0	2	0	0
SKIN	ADENOMA, SEBACEOUS CELL	P-value of test of trend or comparison	.6929	.4658		
		Number of animals reported with tumor	0	1	0	0
	BASAL CELL TUMOR	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	2	0	0	0
	CARCINOMA, SEBACEOUS CELL	P-value of test of trend or comparison	.9055	.7189	1	1
		Number of animals reported with tumor	1	1	0	0
	FIBROMA	P-value of test of trend or comparison	.3350	.9206	.5340	.6437
		Number of animals reported with tumor	3	1	3	2
	FIBROSARCOMA	P-value of test of trend or comparison	.7468	.7180	.6841	1
		Number of animals reported with tumor	1	1	1	0
	KERATOACANTHOMA	P-value of test of trend or comparison	.3062	.5661	.8202	.3644
		Number of animals reported with tumor	4	4	2	4
	LIPOMA	P-value of test of trend or comparison	.9394	.9244	.9049	1
		Number of animals reported with tumor	3	1	1	0
	PAPILLOMA, SQUAMOUS CELL	P-value of test of trend or comparison	.6953	.4730		
		Number of animals reported with tumor	0	1	0	0
	PILOMATRICOMA	P-value of test of trend or comparison	.2259	.4730		.3810
		Number of animals reported with tumor	0	1	0	1
	SARCOMA, UNDIFFERENTIATED	P-value of test of trend or comparison	.4653	.7189	1	.6131
		Number of animals reported with tumor	1	1	0	1
	SCHWANNOMA	P-value of test of trend or comparison	.8674	.8537	.8195	1
		Number of animals reported with tumor	2	1	1	0
SOFT TISSUE- ABD	HIBERNOMA	P-value of test of trend or comparison	.6953	.4730		
		Number of animals reported with tumor	0	1	0	0
SOFT TISSUE- THO	HIBERNOMA	P-value of test of trend or comparison	.8685	.7754	.9821	.8714
		Number of animals reported with tumor	6	4	1	2

Table B.7

Table of reported tumors in Rat Study
NDA 205437
Animal carcinogenicity study
Male rats

Organ or tissue name	*Tumor name*	*Quantity*	*Control Size = 70*	*Low dose Size = 70*	*Mid dose Size = 70*	*High dose Size = 70*
	SCHWANNOMA	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
SPLEEN	SARCOMA, UNDIFFERENTIATED	P-value of test of trend or comparison	.6929	.4658		
		Number of animals reported with tumor	0	1	0	0
STOMACH, NON	LEIOMYOSARCOMA	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
SYSTEMIC TUMORS	FIBROUS HISTIOCYTOMA	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
	HEMANGIOMA	P-value of test of trend or comparison	.1292		.4348	.3906
		Number of animals reported with tumor	0	0	1	1
	HEMANGIOSARCOMA	P-value of test of trend or comparison	.7468	.7180	.6841	1
		Number of animals reported with tumor	1	1	1	0
	LYMPHOMA	P-value of test of trend or comparison	.5827	1	1	.7738
		Number of animals reported with tumor	2	0	0	1
	MESOTHELIOMA	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0
	SARCOMA, HISTIOCYTIC	P-value of test of trend or comparison	.1292		.4348	.3906
		Number of animals reported with tumor	0	0	1	1
TAIL	FIBROMA	P-value of test of trend or comparison	.6929	.4658		
		Number of animals reported with tumor	0	1	0	0
	PAPILLOMA, SQUAMOUS CELL	P-value of test of trend or comparison	.6929	.4658		
		Number of animals reported with tumor	0	1	0	0
TESTES	ADENOMA, INTERSTITIAL CELL	P-value of test of trend or comparison	.5858	.9171	.6162	.8354
		Number of animals reported with tumor	5	2	4	2
THYMUS	THYMOMA	P-value of test of trend or comparison	.6917	.4714		
		Number of animals reported with tumor	0	1	0	0
THYROID GLANDS	ADENOMA, C-CELL	P-value of test of trend or comparison	.8447	.5829	.8202	.9137

Table B.7

Table of reported tumors in Rat Study
NDA 205437
Animal carcinogenicity study
Male rats

Organ or tissue name	*Tumor name*	*Quantity*	*Control Size = 70*	*Low dose Size = 70*	*Mid dose Size = 70*	*High dose Size = 70*
		Number of animals reported with tumor	4	4	2	1
	ADENOMA, FOLLICULAR CELL	P-value of test of trend or comparison	.7372	.9124	.9664	.8182
		Number of animals reported with tumor	5	2	1	2
	CARCINOMA, C-CELL	P-value of test of trend or comparison	.6929	.4658		
		Number of animals reported with tumor	0	1	0	0
	CARCINOMA, FOLLICULAR CELL	P-value of test of trend or comparison	.5237	.4730	.4348	
		Number of animals reported with tumor	0	1	1	0
ZYMBAL'S GLANDS	CARCINOMA	P-value of test of trend or comparison	1	1	1	1
		Number of animals reported with tumor	1	0	0	0

Table B.8

Table of reported tumors in Rat Study
NDA 205437
Animal carcinogenicity study
Female rats
Composite endpoints

Composite endpoint	*Quantity*	*Control Size = 70*	*Low dose Size = 70*	*Mid dose Size = 70*	*High dose Size = 70*
Adenomas and carcinomas of the adrenal cortex (excluding A cell tumors)	P-value of test of trend or comparison	.1548	.8389	.9361	.4049
	Number of animals reported with tumor	3	2	1	4
All cervical and uterine polyps	P-value of test of trend or comparison	.5341	.6075	.4410	.6536
	Number of animals reported with tumor	8	8	9	6
All hibernomas	P-value of test of trend or comparison	.2398	.9415	.9302	.5688
	Number of animals reported with tumor	3	1	1	3
All lipomas	P-value of test of trend or comparison	1	1	1	1
	Number of animals reported with tumor	3	0	0	0
All schwannomas	P-value of test of trend or comparison	.4516	.7581	.7416	.7031
	Number of animals reported with tumor	1	1	1	1
Bronchiolo-alveolar tumors	P-value of test of trend or comparison	1	1	1	1
	Number of animals reported with tumor	1	0	0	0
C-cell tumors	P-value of test of trend or comparison	.9843	.7514	.9435	.9934
	Number of animals reported with tumor	7	6	3	1
Cervical and uterine endometrial stromal sarcomas	P-value of test of trend or comparison	.4509	.7581	.7357	.7031
	Number of animals reported with tumor	1	1	1	1
Cervical and vaginal carcinomas	P-value of test of trend or comparison	.4651		.4886	
	Number of animals reported with tumor	0	0	1	0
Cervical, uterine, and vaginal fibromas	P-value of test of trend or comparison	.6876	.2582	.2359	
	Number of animals reported with tumor	0	2	2	0
Cervical, uterine, and vaginal fibromas and fibrosarcomas	P-value of test of trend or comparison	.6876	.2582	.2359	
	Number of animals reported with tumor	0	2	2	0
Endometrial stromal sarcomas and polyps of the cervix and uterus	P-value of test of trend or comparison	.4631	.5000	.3389	.5481
	Number of animals reported with tumor	8	9	10	7
Fibromas and fibrosarcomas of the skin (and tail)	P-value of test of trend or comparison	.5618	.5245	1	.7019
	Number of animals reported with tumor	1	2	0	1

Table B.8

Table of reported tumors in Rat Study
NDA 205437
Animal carcinogenicity study
Female rats
Composite endpoints

Composite endpoint	*Quantity*	*Control Size = 70*	*Low dose Size = 70*	*Mid dose Size = 70*	*High dose Size = 70*
Follicular cell tumors	P-value of test of trend or comparison	.6339	.8830	1	.8407
	Number of animals reported with tumor	2	1	0	1
Gastrointestinal adenocarcinomas	P-value of test of trend or comparison	1	1	1	1
	Number of animals reported with tumor	1	0	0	0
Gastrointestinal adenomas	P-value of test of trend or comparison	.7384	.5109		
	Number of animals reported with tumor	0	1	0	0
Gastrointestinal adenomas and adenocarcinomas	P-value of test of trend or comparison	.9327	.7635	1	1
	Number of animals reported with tumor	1	1	0	0
Glial cell tumors	P-value of test of trend or comparison	1	1	1	1
	Number of animals reported with tumor	3	0	0	0
Granular cell tumors	P-value of test of trend or comparison	.9313	.5165	.9047	.9438
	Number of animals reported with tumor	6	7	3	2
Hemangiomas and hemangiosarcomas	P-value of test of trend or comparison	.7384	.5109		
	Number of animals reported with tumor	0	1	0	0
Hepatocellular tumors	P-value of test of trend or comparison	.7384	.5109		
	Number of animals reported with tumor	0	1	0	0
Histiocytoma and reticulosis	P-value of test of trend or comparison	.7889	.2582		
	Number of animals reported with tumor	0	2	0	0
Internal squamous cell papillomas and carcinomas	P-value of test of trend or comparison	.8105	.5165	.4828	1
	Number of animals reported with tumor	1	2	2	0
Islet cell tumors	P-value of test of trend or comparison	.8574	.3976	1	.8488
	Number of animals reported with tumor	4	6	0	2
Mammary adenoma, adenocarcinomas, and adenocanthomas	P-value of test of trend or comparison	.9951	.7163	.8397	.9973
	Number of animals reported with tumor	16	15	12	4
Osteomas and osteosarcomas	P-value of test of trend or comparison	.7384	.5109		
	Number of animals reported with tumor	0	1	0	0

Table B.8

Table of reported tumors in Rat Study
NDA 205437
Animal carcinogenicity study
Female rats
Composite endpoints

Composite endpoint	*Quantity*	*Control Size = 70*	*Low dose Size = 70*	*Mid dose Size = 70*	*High dose Size = 70*
Ovarian Sertoli cell tumors and tubulostromal adenomas	P-value of test of trend or comparison	.1535		.4886	.4512
	Number of animals reported with tumor	0	0	1	1
Ovarian luteomas and thecomas	P-value of test of trend or comparison	.7384	.5109		
	Number of animals reported with tumor	0	1	0	0
Renal tubule tumors	P-value of test of trend or comparison	.4651		.4886	
	Number of animals reported with tumor	0	0	1	0
Sarcoma, undifferentiated in reproductive tissues	P-value of test of trend or comparison	.7399	.5161		
	Number of animals reported with tumor	0	1	0	0
Sarcoma, undifferentiated of the uterus and vagina	P-value of test of trend or comparison	.7399	.5161		
	Number of animals reported with tumor	0	1	0	0
Sarcoma, undifferentiated, of the skin and paws	P-value of test of trend or comparison	1	1	1	1
	Number of animals reported with tumor	1	0	0	0
Sarcomas of the cervix, uterus, and vagina (excluding fibrosarcomas)	P-value of test of trend or comparison	.5396	.5161	.7357	.7031
	Number of animals reported with tumor	1	2	1	1
Sarcomas of the cervix, uterus, and vagina (including fibrosarcomas)	P-value of test of trend or comparison	.5396	.5161	.7357	.7031
	Number of animals reported with tumor	1	2	1	1
Sertoli cell tumors	P-value of test of trend or comparison	.0456		.4886	.2066
	Number of animals reported with tumor	0	0	1	2
Squamous cell papillomas of the skin and tail	P-value of test of trend or comparison	.7384	.5109		
	Number of animals reported with tumor	0	1	0	0
Tumors of the pars distalis and pars intermedia	P-value of test of trend or comparison	.7764	.0809	.6972	.6284
	Number of animals reported with tumor	56	67	59	56
Uterine fibromas and fibrosarcomas	P-value of test of trend or comparison	.4651		.4886	
	Number of animals reported with tumor	0	0	1	0
Zymbal glands tumors	P-value of test of trend or comparison	.2189			.4568
	Number of animals reported with tumor	0	0	0	1

Table B.9

Table of reported tumors in Rat Study
NDA 205437
Animal carcinogenicity study
Male rats
Composite endpoints

Composite endpoint	*Quantity*	*Control Size = 70*	*Low dose Size = 70*	*Mid dose Size = 70*	*High dose Size = 70*
Acinar cell tumors	P-value of test of trend or comparison	.3890	.7257	.6841	.6206
	Number of animals reported with tumor	1	1	1	1
Adenomas and carcinomas of the adrenal cortex (excluding A cell tumors)	P-value of test of trend or comparison	.6953	.4730		
	Number of animals reported with tumor	0	1	0	0
All hibernomas	P-value of test of trend or comparison	.8685	.7754	.9821	.8714
	Number of animals reported with tumor	6	4	1	2
All leiomyosarcomas	P-value of test of trend or comparison	.6715	1	.6841	1
	Number of animals reported with tumor	1	0	1	0
All lipomas	P-value of test of trend or comparison	.7346	.6094	.7292	.8619
	Number of animals reported with tumor	3	3	2	1
All schwannomas	P-value of test of trend or comparison	.9948	.8811	.9817	1
	Number of animals reported with tumor	6	3	1	0
Bronchiolo-alveolar tumors	P-value of test of trend or comparison	.9040	.7112	1	1
	Number of animals reported with tumor	1	1	0	0
C-cell tumors	P-value of test of trend or comparison	.8667	.4318	.8202	.9137
	Number of animals reported with tumor	4	5	2	1
Fibromas and fibrosarcomas of the skin (and tail)	P-value of test of trend or comparison	.5820	.7257	.6664	.7544
	Number of animals reported with tumor	4	3	3	2
Follicular cell tumors	P-value of test of trend or comparison	.7174	.8126	.8749	.8182
	Number of animals reported with tumor	5	3	2	2
Gastrointestinal adenomas	P-value of test of trend or comparison	.6953	.4730		
	Number of animals reported with tumor	0	1	0	0
Gastrointestinal adenomas and adenocarcinomas	P-value of test of trend or comparison	.6953	.4730		
	Number of animals reported with tumor	0	1	0	0
Glial cell tumors	P-value of test of trend or comparison	.1261	.4658	.0775	.3906
	Number of animals reported with tumor	0	1	3	1

Table B.9

Table of reported tumors in Rat Study
NDA 205437
Animal carcinogenicity study
Male rats
Composite endpoints

Composite endpoint	*Quantity*	*Control Size = 70*	*Low dose Size = 70*	*Mid dose Size = 70*	*High dose Size = 70*
Granular cell tumors	P-value of test of trend or comparison	.6775	.1010	.4348	
	Number of animals reported with tumor	0	3	1	0
Hemangiomas and hemangiosarcomas	P-value of test of trend or comparison	.3277	.7180	.4134	.6324
	Number of animals reported with tumor	1	1	2	1
Hepatocellular tumors	P-value of test of trend or comparison	.3890	.5448	.2325	.6507
	Number of animals reported with tumor	5	5	7	3
Histiocytoma and reticulosis	P-value of test of trend or comparison	.1044	1	.6770	.3274
	Number of animals reported with tumor	1	0	1	2
Internal squamous cell papillomas and carcinomas	P-value of test of trend or comparison	.6532	.8537	1	.7629
	Number of animals reported with tumor	2	1	0	1
Islet cell tumors	P-value of test of trend or comparison	.7851	.7863	.4362	.9677
	Number of animals reported with tumor	6	4	6	1
Leiomyosarcomas and leiomyomas	P-value of test of trend or comparison	.6715	1	.6841	1
	Number of animals reported with tumor	1	0	1	0
Mammary adenoma, adenocarcinomas, and adenocanthomas	P-value of test of trend or comparison	.8964	.4659	1	1
	Number of animals reported with tumor	1	2	0	0
Meningiomas and meningeal sarcomas	P-value of test of trend or comparison	.0956		.1925	.3906
	Number of animals reported with tumor	0	0	2	1
Osteomas and osteosarcomas	P-value of test of trend or comparison	.6953	.4730		
	Number of animals reported with tumor	0	1	0	0
Renal tubule tumors	P-value of test of trend or comparison	.7450	.7257	.6841	1
	Number of animals reported with tumor	1	1	1	0
Sarcoma, undifferentiated, of the skin and paws	P-value of test of trend or comparison	.3830	.7189	.6770	.6131
	Number of animals reported with tumor	1	1	1	1
Sebaceous cell tumors	P-value of test of trend or comparison	.8936	.4494	1	1
	Number of animals reported with tumor	1	2	0	0

Table B.9

Table of reported tumors in Rat Study
NDA 205437
Animal carcinogenicity study
Male rats
Composite endpoints

Composite endpoint	*Quantity*	*Control Size = 70*	*Low dose Size = 70*	*Mid dose Size = 70*	*High dose Size = 70*
Squamous cell papillomas of the skin and tail	P-value of test of trend or comparison	.7409	.2203		
	Number of animals reported with tumor	0	2	0	0
Tumors of the pars distalis and pars intermedia	P-value of test of trend or comparison	.9383	.4248	.9193	.9024
	Number of animals reported with tumor	43	44	27	26
Zymbal glands tumors	P-value of test of trend or comparison	1	1	1	1
	Number of animals reported with tumor	1	0	0	0

Table B.10

Table of tumors reported significant (alpha < 0.05) in at least one arm - Rat Study
NDA 205437
Animal carcinogenicity study
Female rats
Composite endpoints

Composite endpoint	*Quantity*	*Control*	*Low dose*	*Mid dose*	*High dose*
Sertoli cell tumors	P-value of test of trend or comparison	.0456		.4886	.2066
	Number of animals reported with tumor	0	0	1	2
	Poly-3 adjusted incidence rate	0.0%	0.0%	2.3%	5.2%
	95% CI for poly-3 adjusted incidence rate (%)	(0,7.9)	(0,7.5)	(0.06,12.3)	(0.63,17.7)
	Poly-3 adjusted number of animals at risk	45.6	47.3	43.6	38.3

B.3 Unexamined and autolytic organs

Table B.11

Organs reported as autolytic
NDA 205437
Animal carcinogenicity study
Female Rats

Organ or tissue name	*Control(count)*	*Control(%)*	*Low dose(count)*	*Low dose(%)*	*Mid dose(count)*	*Mid dose(%)*	*High dose(count)*	*High dose(%)*	*Total(count)*	*Total(%)*
CECUM	1	1.4%	.	.	1	1.4%	1	1.4%	3	1.1%
DUODENUM	2	2.9%	1	1.4%	1	1.4%	1	1.4%	5	1.8%
ILEUM	6	8.6%	5	7.1%	5	7.1%	4	5.7%	20	7.1%
JEJUNUM	11	16%	12	17%	15	21%	12	17%	50	18%
LARYNX	.	.	.	.	.	.	1	1.4%	1	0.4%
PEYER'S PATCHES	.	.	2	2.9%	4	5.7%	3	4.3%	9	3.2%
RECTUM	.	.	1	1.4%	.	.	.	.	1	0.4%
THYMUS	.	.	.	.	.	.	1	1.4%	1	0.4%
TRACHEA	.	.	.	.	.	.	1	1.4%	1	0.4%

Table B.12

Organs reported as autolytic
NDA 205437
Animal carcinogenicity study
Male Rats

Organ or tissue name	*Control(count)*	*Control(%)*	*Low dose(count)*	*Low dose(%)*	*Mid dose(count)*	*Mid dose(%)*	*High dose(count)*	*High dose(%)*	*Total(count)*	*Total(%)*
CECUM	1	1.4%	3	4.3%	3	4.3%	2	2.9%	9	3.2%
COLON	.	.	2	2.9%	.	.	1	1.4%	3	1.1%
DUODENUM	3	4.3%	.	.	1	1.4%	4	5.7%	8	2.9%
ILEUM	4	5.7%	3	4.3%	2	2.9%	5	7.1%	14	5.0%
JEJUNUM	19	27%	10	14%	8	11%	12	17%	49	18%
PEYER'S PATCHES	3	4.3%	1	1.4%	.	.	3	4.3%	7	2.5%
RECTUM	1	1.4%	.	.	.	.	.	.	1	0.4%
URINARY BLADDER	.	.	.	.	.	.	1	1.4%	1	0.4%

Table B.13

Organs reported as unexamined
NDA 205437
Animal carcinogenicity study
Female Rats

Organ or tissue name	*Control(count)*	*Control(%)*	*Low dose(count)*	*Low dose(%)*	*Mid dose(count)*	*Mid dose(%)*	*High dose(count)*	*High dose(%)*	*Total(count)*	*Total(%)*
ADRENAL MEDULLA	1	1.4%	1	1.4%	2	2.9%	.	.	4	1.4%
BILE DUCT	1	1.4%	.	.	.	.	.	.	1	0.4%
CLITORAL GLANDS	.	.	3	4.3%	3	4.3%	.	.	6	2.1%
LYMPH NODE, MAND	.	.	1	1.4%	.	.	.	.	1	0.4%
LYMPH NODE, MES	1	1.4%	.	.	.	.	.	.	1	0.4%
MARROW, FEMUR	.	.	1	1.4%	.	.	.	.	1	0.4%
MARROW, STERN	.	.	.	.	2	2.9%	.	.	2	0.7%
NERVE, SCIATIC	.	.	.	.	.	.	2	2.9%	2	0.7%
OVIDUCTS	1	1.4%	.	.	1	1.4%	.	.	2	0.7%
PARATHYROIDS	7	10%	6	8.6%	2	2.9%	4	5.7%	19	6.8%
PAWS	.	.	.	.	.	.	1	1.4%	1	0.4%
PEYER'S PATCHES	2	2.9%	3	4.3%	3	4.3%	7	10%	15	5.4%
PHARYNX	1	1.4%	.	.	.	.	.	.	1	0.4%
SKELETAL MUSCLE	1	1.4%	.	.	.	.	.	.	1	0.4%
SOFT TISSUE- ABD	.	.	1	1.4%	1	1.4%	.	.	2	0.7%
SOFT TISSUE- OC	.	.	.	.	1	1.4%	.	.	1	0.4%
SPLEEN	.	.	.	.	.	.	2	2.9%	2	0.7%
STERNUM	.	.	.	.	2	2.9%	.	.	2	0.7%
TAIL	.	.	.	.	.	.	1	1.4%	1	0.4%
THYMUS	3	4.3%	.	.	3	4.3%	3	4.3%	9	3.2%
VAGINA	.	.	.	.	1	1.4%	.	.	1	0.4%
ZYMBAL'S GLANDS	2	2.9%	2	2.9%	3	4.3%	2	2.9%	9	3.2%

Table B.14

Organs reported as unexamined
NDA 205437
Animal carcinogenicity study
Male Rats

Organ or tissue name	*Control(count)*	*Control(%)*	*Low dose(count)*	*Low dose(%)*	*Mid dose(count)*	*Mid dose(%)*	*High dose(count)*	*High dose(%)*	*Total(count)*	*Total(%)*
BILE DUCT	1	1.4%	4	5.7%	.	.	1	1.4%	6	2.1%
LYMPH NODE, MAND	1	1.4%	.	.	1	1.4%	.	.	2	0.7%
LYMPH NODE, MES	.	.	.	.	1	1.4%	1	1.4%	2	0.7%
LYMPH NODE, REN	.	.	1	1.4%	.	.	.	.	1	0.4%
MAMMARY GLAND	5	7.1%	4	5.7%	5	7.1%	3	4.3%	17	6.1%
MARROW, STERN	1	1.4%	.	.	.	.	.	.	1	0.4%
NASAL LEVEL III	.	.	.	.	.	.	1	1.4%	1	0.4%
NASAL LEVEL IV	.	.	.	.	1	1.4%	.	.	1	0.4%
PANCREAS	1	1.4%	.	.	.	.	.	.	1	0.4%
PARATHYROIDS	5	7.1%	4	5.7%	7	10%	11	16%	27	9.6%
PENIS	.	.	.	.	1	1.4%	.	.	1	0.4%
PEYER'S PATCHES	2	2.9%	2	2.9%	4	5.7%	3	4.3%	11	3.9%
PREPUTIAL GLANDS	3	4.3%	1	1.4%	3	4.3%	4	5.7%	11	3.9%
PROSTATE	1	1.4%	.	.	.	.	.	.	1	0.4%
SOFT TISSUE- ABD	1	1.4%	.	.	.	.	.	.	1	0.4%
STERNUM	1	1.4%	.	.	.	.	.	.	1	0.4%
STOMACH, GLAN	.	.	1	1.4%	.	.	.	.	1	0.4%
STOMACH, NON	.	.	1	1.4%	.	.	.	.	1	0.4%
TAIL	.	.	2	2.9%	.	.	.	.	2	0.7%
THYMUS	5	7.1%	5	7.1%	5	7.1%	2	2.9%	17	6.1%
TRACHEA	.	.	1	1.4%	.	.	.	.	1	0.4%
ZYMBAL'S GLANDS	10	14%	13	19%	11	16%	9	13%	43	15%

Bibliography

[1] A J Bailer and C J Portier. Effects of treatment-induced mortality and tumor-induced mortality on tests for carcinogenicity in small samples. *Biometrics*, 44(2):417–31, 1988.

[2] G S Bieler and R L Williams. Ratio estimates, the delta method, and quantal response tests for increased carcinogenicity. *Biometrics*, 49(3):793–801, 1993.

[3] KC Chu, C Cueto, and JM Ward. Factors in the evaluation of 200 National Cancer Institute carcinogen bioassays. *Journal of toxicology and environmental health*, 8(1-2):251–280, 1981.

[4] J K Haseman. A reexamination of false-positive rates carcinogenesis studies. *Fundamental and applied toxicology*, 3(4):334–339, 1983.

[5] K K Lin and M A Rahman. Overall false positive rates in tests for linear trend in tumor incidence in animal carcinogenicity studies of new drugs. *J Biopharm Stat*, 8(1):1–15; discussion 17–22, 1998.

[6] R Peto, M C Pike, N E Day, R G Gray, P N Lee, S Parish, J Peto, S Richards, and J Wahrendorf. Guidelines for simple, sensitive significance tests for carcinogenic effects in long-term animal experiments. *IARC Monogr Eval Carcinog Risk Chem Hum Suppl*, NIL(2 Suppl):311–426, 1980.

[7] M A Rahman and K K Lin. A comparison of false positive rates of peto and poly-3 methods for long-term carcinogenicity data analysis using multiple comparison adjustment method suggested by Lin and Rahman. *J Biopharm Stat*, 18(5):949–58, 2008.

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

/s/

MATTHEW T JACKSON
11/15/2013

KARL K LIN
11/15/2013
Concur with review

STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

NDA Number: 205437 **Applicant: Celgene** **Stamp Date: 3/21/2013**

Drug Name: Apremilast **NDA/BLA Type: Standard**

On **initial** overview of the NDA/BLA application for RTF:

	Content Parameter	Yes	No	NA	Comments
1	Index is sufficient to locate necessary reports, tables, data, etc.	x			
2	ISS, ISE, and complete study reports are available (including original protocols, subsequent amendments, etc.)	x			
3	Safety and efficacy were investigated for gender, racial, and geriatric subgroups investigated (if applicable).	x			
4	Data sets in EDR are accessible and do they conform to applicable guidances (e.g., existence of define.pdf file for data sets).	x			

IS THE STATISTICAL SECTION OF THE APPLICATION FILEABLE? Yes

If the NDA/BLA is not fileable from the statistical perspective, state the reasons and provide comments to be sent to the Applicant.

Please identify and list any potential review issues to be forwarded to the Applicant for the 74-day letter.

Content Parameter (possible review concerns for 74-day letter)	Yes	No	NA	Comment
Designs utilized are appropriate for the indications requested.	x			
Endpoints and methods of analysis are specified in the protocols/statistical analysis plans.	x			
Interim analyses (if present) were pre-specified in the protocol and appropriate adjustments in significance level made. DSMB meeting minutes and data are available.	x			
Appropriate references for novel statistical methodology (if present) are included.	x			
Safety data organized to permit analyses across clinical trials in the NDA/BLA.	x			
Investigation of effect of dropouts on statistical analyses as described by applicant appears adequate.	x			

STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

Robert Abugov, Ph.D.	5/6/2013
Reviewing Statistician	Date

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

/s/

ROBERT ABUGOV
05/06/2013

JOAN K BUENCONSEJO
05/15/2013
I concur