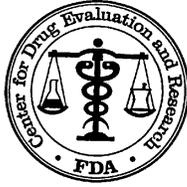


**CENTER FOR DRUG EVALUATION AND
RESEARCH**

APPLICATION NUMBER:

202192Orig1s000

STATISTICAL REVIEW(S)



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

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/BLA Serial Number: 212192/001

Drug Name: Ruxolitinib

Indication(s): Myelofibrosis

Applicant: Incyte Corp.

Date(s): Submitted on 6/3/2011, PDUFA date 12/3/2011

Review Priority: Priority

Biometrics Division: Division of Biometrics V (HFD-711)

Statistical Reviewer: Dr. Laura Lu, Ph.D

Concurring Reviewers: Dr. Mark Rothmann, Ph.D, Team Leader
Dr. Rajeshwari Sridhara, Ph.D, Division Director

Medical Division: Division of Hematology Drug Products (HFD-107)

Clinical Team: Dr. Al Deisseroth, MD, Medical Reviewer
Dr. Ann Farrell, MD, Division Director

Project Manager: Ms. Amy Baird

Keywords:

Response rate, missing data

Table of Contents

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES	1
FOOD AND DRUG ADMINISTRATION	1
STATISTICAL REVIEW AND EVALUATION	1
LIST OF TABLES.....	3
1. EXECUTIVE SUMMARY	4
2. INTRODUCTION	4
2.1 OVERVIEW.....	4
2.2 DATA SOURCES	5
3. STATISTICAL EVALUATION	5
3.1 DATA AND ANALYSIS QUALITY	5
3.2 EVALUATION OF EFFICACY	5
3.2.1 Study 351	5
3.2.2 Study 352	10
3.3 EVALUATION OF SAFETY	13
4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS	13
4.1 GENDER, RACE, AGE, AND GEOGRAPHIC REGION	13
4.2 OTHER SPECIAL/SUBGROUP POPULATIONS	14
5. SUMMARY AND CONCLUSIONS	16
5.1 STATISTICAL ISSUES AND COLLECTIVE EVIDENCE	16
5.2 CONCLUSIONS AND RECOMMENDATIONS	17
CHECK LIST	18

LIST OF TABLES

Table 1. Patient Disposition (Study 351).....	7
Table 2. Patient Demographics (Study 351).....	7
Table 3. Baseline Characteristics (Study 351).....	8
Table 4. Proportion of subjects achieved a \geq 35% reduction from baseline at Week 24.....	9
Table 5. Results of secondary endpoints (Study 351).....	9
Table 6. Patient disposition (Study 352).....	11
Table 7. Patient Demographics (Study 352).....	12
Table 8. Baseline characteristics (Study 352).....	12
Table 9. Results of primary and key secondary endpoint (Study 352).....	13
Table 10. Subgroup analyses by gender, age and race (Study 351).....	14
Table 11. Subgroup analyses by gender, age and race (Study 352).....	14
Table 12. Subgroup analyses by baseline characteristics (Study 351).....	15
Table 13. Subgroup analyses by baseline characteristics (Study 352).....	15
Table 14. Results of primary endpoints in Study 351 and Study 352.....	16

1. EXECUTIVE SUMMARY

NDA202192 was submitted to support the approval of ruxolitinib on the treatment of myelofibrosis. Two randomized, controlled studies (Study 351 and Study 352) were conducted. The control arms were placebo and best available therapy (BAT) in Study 351 and Study 352, respectively. The results of these two studies demonstrated that, compared with those in the control arm, patients in the ruxolitinib arm were more likely to achieved at least a 35% reduction in spleen volume from baseline at Week 24 in Study 351 (42% vs. 1%) and Week 48 in Study 352 (28% vs. 0%).

Across subgroups, response rate of the primary endpoint is consistently higher in the ruxolitinib group than those in the placebo group. However, the response rate in female group is much higher than that in the male group within the ruxolitinib arm in both studies (25% vs. 59% respectively in Study 351; 25% vs. 33% respectively in Study 351). And the response rate in V617F positive group is much higher than that in the V617F negative group within the ruxolitinib arm in both studies (48% vs. 28% respectively in Study 351; 33% vs. 14% respectively in Study 351).

2. INTRODUCTION

2.1 Overview

Two randomized, controlled studies (Study 351 and Study 352) have been conducted for ruxolitinib on the treatment of patients with high or intermediate-2 risk myelofibrosis (including primary myelofibrosis, post-polycythemia vera myelofibrosis, and post-essential thrombocythemia myelofibrosis). The majority of the patients in Study 351 were from US (76%) and all patients in Study 352 were from Europe.

Study 351 is a randomized, double-blind, placebo-controlled study comparing the efficacy and safety of ruxolitinib to placebo in subjects with Primary Myelofibrosis (PMF), Post Polycythemia-myelofibrosis (PPV-MF) or Post Essential Thrombocythemia myelofibrosis (PET-MF). Subjects were randomized with a 1:1 ratio to receive ruxolitinib or matching placebo tablets. The primary endpoint of Study 351 was the proportion of subjects achieving >35% reduction in spleen volume from baseline to Week 24 as measured by MRI (or by CT for applicable). Secondary endpoints include the proportion of subjects who have a 50% reduction from baseline to Week 24 in the total symptom score, change from baseline to Week 24 in the total symptom score and overall survival. The trial was originally planned with a sample size of 240 subjects, which provided a 97% power to detect a treatment difference in the primary endpoint at two-sided alpha level of 0.05 using the chi-square test.

Study 352 is an open-label, randomized study comparing the efficacy and safety of ruxolitinib tablets versus investigator-selected BAT in subjects with MF with splenomegaly of at least 5 cm below the costal margin by manual palpation, and either 2 (Intermediate-2 risk category) or 3 or

more (High risk category) prognostic factors according to the International Working Group for Myelofibrosis Research and Treatment (IWG-MRT). A total of 219 patients were randomized with a ratio of 2:1 to ruxolitinib and BAT. The primary endpoint of Study 352 was the proportion of subjects achieving >35% reduction in spleen volume from baseline to Week 48 as measured by MRI (or by CT for applicable). The key secondary endpoint for this study was the proportion of subjects achieving a $\geq 35\%$ reduction of spleen volume as measured by MRI or CT from baseline at Week 24.

2.2 Data Sources

The study reports and data for this NDA are located at <\\Cdsub1\evsprod\NDA202192\0000\m5\datasets>. Efficacy evaluation in this NDA was mainly based on the following electronic datasets:

1. Aseff3.xpt at <\\Cdsub1\evsprod\NDA202192\0000\m5\datasets\incb-18424-351\listings>; and
2. Aseff3.xpt at <\\Cdsub1\evsprod\NDA202192\0000\m5\datasets\incb-18424-351\listings>.

Note that these two analysis datasets were submitted to the wrong folders. They should be included in the 'analysis' data folder instead of the 'listings' folder.

3. STATISTICAL EVALUATION

3.1 Data and Analysis Quality

The raw data and derived data were submitted in the folder <\\Cdsub1\evsprod\NDA202192\0000\m5\datasets>. The analysis datasets were documented well. The raw data can be verified through Case Report Forms. The value of the derived variables can be verified by the raw data.

3.2 Evaluation of Efficacy

3.2.1 Study 351

3.2.1.1 Study Design and Endpoints

Study 351 is a randomized, double-blind, placebo-controlled study comparing the efficacy and safety of ruxolitinib to placebo in subjects with Primary Myelofibrosis (PMF), Post Polycythemia-myelofibrosis (PPV-MF) or Post Essential Thrombocythemia myelofibrosis (PET-MF). Subjects were randomized with a 1:1 ratio to receive ruxolitinib or matching placebo tablets. The starting dose was determined based on baseline platelet count. Subjects with baseline platelet count > 200,000/ μL began a dose regimen of 20 mg twice daily. Subjects with baseline platelet count of 100,000/ μL to 200,000/ μL (inclusive) began a dose regimen of 15 mg twice daily. The dose could be adjusted by the investigator based on a standardized dosing paradigm, which was used to determine dose adjustments for safety and efficacy so that each subject was titrated to their most appropriate dose. Doses were not to exceed 25 mg twice daily. Subjects

randomized to placebo made the same adjustments in the number of matching placebo tablets to maintain the blind.

When half of the subjects remaining in the study (including subjects continuing on randomized treatment and those who had crossed over to ruxolitinib from placebo) completed the Week 36 visit and all subjects enrolled completed Week 24 or discontinued, the database was frozen and the primary analysis was conducted. Once this was complete, all subjects were unblinded. After the study was unblinded, subjects who had been randomized to placebo were given the opportunity to crossover to ruxolitinib treatment, provided hematology laboratory parameters were adequate; and subjects who had been randomized to ruxolitinib remained in the study if they were obtaining benefit from treatment.

The primary efficacy endpoint is the proportion of subjects achieving >35% reduction in spleen volume from baseline to Week 24 as measured by MRI (or by CT for applicable). The treatment difference in the proportion was tested by Fisher's exact test. A subject must have a baseline spleen volume in order to be included in the primary efficacy analysis. A subject with a missing Week 24 spleen volume was considered as having not achieved the >35% reduction. Subjects who dropped out of the study due to lack of efficacy or treatment-related adverse events, or made an early crossover to active treatment (placebo subjects only) prior to Week 24 visit were considered as having not achieved the >35% reduction.

Secondary endpoints include the proportion of subjects who have a 50% reduction from baseline to Week 24 in the total symptom score, change from baseline to Week 24 in the total symptom score and overall survival. The secondary efficacy endpoints was analyzed only when the study had reached the efficacy objective in the primary endpoint and was tested in a fixed sequence-testing procedure with each at the 0.05 alpha level in the order below:

1. The proportion of subjects who have a 50% reduction from baseline to Week 24 in the total symptom score
2. Change from baseline to Week 24 in the total symptom score; and
3. Overall Survival

Proportion of subjects achieving $\geq 50\%$ reduction from baseline in the Week 24 total symptom score were then calculated by treatment group. The 2 proportions were compared using the Chi-squared test. Only subjects who had a baseline total symptom score were included in the sponsor's analysis. A subject with a missing Week 24 total symptom score was considered as having not achieved the $> 50\%$ reduction. Subjects who dropped out of the study due to lack of efficacy or treatment related adverse events, or were unblinded prior to Week 24 for the early crossover to active treatment were considered as having not achieved the $> 50\%$ reduction.

The change from baseline to Week 24 in the total symptom score was analyzed using both parametric and non-parametric methods. The Wilcoxon Rank-Sum Test was used for a between-group comparison in the median change from baseline. Patients with missing baseline or Week 24 total symptom scores were not included in the analysis. The overall survival variable was analyzed using the Kaplan-Meier method.

Assuming that at least 30% of the active subjects would achieve a >35% reduction from baseline to Week 24, and that rate for the placebo subjects would be no more than 10%, a sample size of

240 subjects (120 per group) would provide a 97% power to detect a treatment difference in the primary endpoint at two-sided alpha level of 0.05 using the chi-square test.

3.2.1.2 Patient disposition

A total of 309 patients were randomized into the study (155 in the roxolitinib arm and 154 in the placebo arm). The number of patient withdrew from study early and the reasons of withdrawal are presented in Table 1 below.

Table 1. Patient Disposition (Study 351)

Treatment Arm	Ruxolitinib (N=155) n (%)	Placebo (N=154) n (%)
Number subjects withdrawn from study	21 (13.5)	37 (24.5)
Reasons for withdrawal from study		
Death	9 (5.8)	9 (6.0)
Adverse event	8 (5.2)	8 (5.3)
Disease progression	3 (1.9)	12 (7.9)
Consent withdrawn	1 (0.6)	5 (2.0)
Other	0 (0)	3 (2.0)

3.2.1.3 Demographics and baseline characteristic

Patient demographics and baseline characteristics are presented in Table 2 and Table 3 below.

Table 2. Patient Demographics (Study 351)

Variable	Ruxolitinib (N = 155)	Placebo (N = 154)
Age (yrs)		
<=65 years, n (%)	70 (45.2)	52 (33.8)
> 65 years, n (%)	85 (54.8)	102 (66.2)
Sex, n (%)		
Male	79 (51.0)	88 (57.1)
Female	76 (49.0)	65 (42.2)
Unknown	0 (0)	1 (0.6)
Race, n (%)		
Black or African American	6 (3.9)	7 (4.5)
White	138 (89.0)	139 (90.3)
Asian	5 (3.2)	4 (2.6)
Native Hawaiian or Other Pacific Islander	1 (0.6)	0 (0)
Other	5 (3.2)	4 (2.6)

Table 3. Baseline Characteristics (Study 351)

Variable	Ruxolitinib (N = 155)	Placebo (N = 154)
Disease subtype, n (%)		
PMF	70 (45.2)	84 (54.5)
PPV-MF	50 (32.3)	47 (30.5)
PET-MF	35 (22.6)	22 (14.3)
Unknown	0 (0)	1 (0.6)
Years since initial diagnosis		
Mean (STD)	4.9 (6.1)	4.6 (6.2)
Fibrosis grade at baseline, n (%)		
0	2 (1.3)	1 (0.6)
1	14 (9.0)	18 (11.7)
2	63 (40.6)	51 (33.1)
3	65 (41.9)	71 (46.1)
Unknown	11 (7.2)	13 (8.4)
Spleen volume, cm ³		
Mean (STD)	2745.7 (1247.0)	2797.6 (1388.5)
Palpable spleen length below the left costal margin, (cm)		
N	155	153
Mean (STD)	16.1 (5.7)	16.4 (6.3)
ECOG Performance Status		
0	47 (31.1)	38 (25.5)
1	87 (57.6)	82 (55.0)
2	14 (9.3)	25 (16.8)
3	3 (2.0)	4 (2.7)
Unknown	4 (2.6)	5 (3.3)
IWG risk category		
High	90 (58.1)	99 (64.3)
Intermediate 2	64 (41.3)	54 (35.1)
Unknown	1 (0.6)	1 (0.6)
Percent V617F at baseline		
Yes	113 (72.9)	123 (79.9)
No	40 (25.8)	27 (17.7)
Unknown	1 (0.6)	4 (0.3)

3.2.1.2 Efficacy Results

Primary endpoint

A significantly larger proportion of subjects in the ruxolitinib group achieved a $\geq 35\%$ reduction from baseline at Week 24 compared with the placebo group (41.9% vs. 0.7%, $p <$

0.0001 by Fisher's Exact test). Table 4 below presents a summary of the results of the primary endpoint.

Table 4. Proportion of subjects achieved a $\geq 35\%$ reduction from baseline at Week 24 (Study 351)

	Ruxolitinib (N = 155)	Placebo (N = 154)
Subjects achieving a $\geq 35\%$ reduction from baseline in spleen volume, n (%)	65 (41.9)	1 (0.7)
% Difference between treatments (95% CI ^a)	41.2 (32.8, 48.7)	
p-value (by Fisher's Exact test)	<0.0001	

a: By Agresti-Caffo method (Agresti and Caffo; The American Statistician, Vol. 54, No. 4, (Nov., 2000), pp.280-288)

Secondary endpoints

Results for the secondary endpoints are presented in Table 5 below. Survival data were not mature. Ten (6.5%) subjects in the ruxolitinib group and 14 (9.1%) subjects in the placebo group died. This includes 20 subjects who died during randomized treatment, after crossover, and up to 28 days after study withdrawal, and 4 subjects who died more than 28 days after withdrawal from the study and represents a hazard ratio of 0.668 ($p = 0.3268$).

Table 5. Results of secondary endpoints (Study 351)

	Ruxolitinib (N = 155)	Placebo (N = 154)
Number of evaluable subjects	148 (95.5)	152 (98.7)
Subjects achieving a $\geq 50\%$ improvement from baseline in total symptom score, n (%)		
Yes	68 (45.9)	8 (5.3)
No	80 (54.1)	144 (94.7)
p-value (by Fisher's Exact test)	< 0.0001	
Change from baseline to Week 24 in total symptom score		
N	131	105
Mean (SD)	-8.6 (10.0)	3.2 (9.4)
p-value (by Wilcoxon rank-sum test)	<0.0001	
Overall Survival		
Number of events	10 (6.5%)	14 (9.1%)
HR	0.668	
P-value (by log-rank test)	0.33	

Reviewers comment: Note that 24% (73/309) of the patients have missing value in change from baseline to Week 24 in total symptom score. To assess the impact of the missing data, this reviewer did a worst-case carried forward by imputing the missing scores in the following way:

1. *imputing the worst score in Week 4, Week 8, Week 12, Week 16 and Week 20 to the missing score on Week 24 in the ruxolitinib arm; and*
2. *imputing the best score in Week 4, Week 8, Week 12, Week 16 and Week 20 to the missing score on Week 24 in the placebo arm.*

After missing value imputation, there are still 6 patients missing in the ruxolitinib arm and 4 missing in the placebo arm. The means become -7.9 in ruxolitinib arm and 2.0 in the BAT arm, which is still statistically significantly different ($p < 0.0001$). Therefore, the result of this sensitivity analysis supports that of the primary analysis.

3.2.2 Study 352

3.2.2.1 Study Design and Endpoints

Study 352 is an open-label, 2:1 randomized study comparing the efficacy and safety of ruxolitinib tablets versus investigator-selected BAT in subjects with MF with splenomegaly of at least 5 cm below the costal margin by manual palpation, and either 2 (Intermediate-2 risk category) or 3 or more (High risk category) prognostic factors according to the International Working Group for Myelofibrosis Research and Treatment (IWG-MRT). The starting dose of ruxolitinib was determined based on baseline platelet count; the maximum dose on study did not exceed 25 mg twice daily.

Patients were randomized 2:1 to the ruxolitinib and BAT arms. Randomization was stratified by baseline prognostic risk level in the following manner:

- Stratum 1: Intermediate risk level 2 (2 risk factors)
- Stratum 2: High risk (3 or more risk factors)

The duration of this study consisted of an enrollment period of 27 weeks, a period of 48 weeks before the primary analysis, and a period of 96 weeks after the last visit of the last ongoing subject for the primary endpoint. Patients in the BAT arm could cross over to ruxolitinib after progression.

The primary efficacy endpoint of this study was the proportion of subjects achieving $\geq 35\%$ reduction in spleen volume from baseline at Week 48 as measured by MRI or CT (for subjects unable to undergo MRI). The key secondary endpoint for this study was the proportion of subjects achieving a $\geq 35\%$ reduction of spleen volume as measured by MRI or CT from baseline at Week 24. Other secondary endpoints included the duration of maintenance of spleen volume reduction (DoMSR) $\geq 35\%$ reduction from baseline, the time to achieve a first $\geq 35\%$ reduction in spleen volume from baseline, progression-free survival (PFS), leukemia-free survival (LFS), overall survival (OS), transfusion dependency/independency, and a change in bone marrow histomorphology.

The percent change from baseline at Week 48 in spleen volume was calculated only for subjects who had an evaluable spleen volume at baseline. The proportion of subjects achieving a $\geq 35\%$ reduction in spleen volume from baseline at Week 48 was then calculated by treatment group. A subject was required to have a baseline spleen volume measurement to be included in the primary efficacy analysis. A subject with a missing Week 48 spleen volume measurement was

considered as not having achieved the $\geq 35\%$ reduction. Subjects who dropped out of the study due to any reason or who had a protocol-defined qualifying event of disease progression prior to Week 48 visit were considered as not having achieved the $\geq 35\%$ reduction. The two proportions were compared using the Cochran-Mantel-Haenszel (CMH) test stratified by prognostic category (Intermediate-2 or High risk).

The definition and analysis of the key secondary endpoint was identical to the definition and main analysis of the primary endpoint, the only difference being the timing of evaluation, at 24 weeks. This key secondary endpoint will be tested at 2-sided alpha level of 0.05 if the primary endpoint is statistically significant at two-sided alpha level of 0.05. The DoMSR was evaluated using Kaplan-Meier estimates for each treatment arm. The time to achieve $\geq 35\%$ reduction in spleen volume from baseline was performed on subjects who achieved a 35% reduction in spleen volume. A separate analysis of DoMSR was performed employing a different definition for duration of response, where the start date was defined as the first spleen volume measurement with $\geq 35\%$ reduction from baseline, and the end date was defined as the first scan that was no longer equal to a 35% reduction and that was a $>25\%$ increase over nadir. PFS, LFS, and OS were summarized using Kaplan-Meier estimates for each treatment arm. Bone marrow histomorphology was noted as fibrosis density and was tabulated by fibrosis grade at baseline and post-baseline. Descriptive statistics (number of subjects and subject percentages) were used.

The sample size of this study was originally calculated based on the primary efficacy variable, the proportion of subjects achieving 35% reduction in spleen volume from baseline at Week 48 as measured by MRI. This endpoint was analyzed using the Chi-square test for a treatment comparison. Assuming at least 35% of the active subjects would achieve a 35% reduction from baseline to Week 48, and that rate for the control subjects would be no more than 10%, a sample size of 150 subjects (100 in active and 50 in control) would provide at least 90% power to detect a treatment difference in the primary endpoint at a 2-sided alpha level of 0.05 using the Chi-square test.

3.2.1.2 Patient disposition

A total of 219 patients were randomized into the study (146 in the roxolitinib arm and 73 in the placebo arm). The number of patients who withdrew from the study early and the reasons of withdrawal are presented in Table 6 below.

Table 6. Patient disposition (Study 352)

	Ruxolitinib N=146	BAT N=73
Discontinued Randomized Treatment Phase	55 (37.7)	42 (57.5)
Reasons for discontinuation from Randomized Treatment Phase		
Entered extension phase ruxolitinib	29 (19.9)	18 (24.7)
Adverse event(s)	12 (8.2)	4 (5.5)
Consent withdrawn	2 (1.4)	9 (12.3)
Disease Progression	1 (0.7)	3 (4.1)

Protocol deviation	2 (1.4)	0
Non-compliance with study medication	2 (1.4)	0
Non-compliance with study procedures	0	1 (1.4)
Other	7 (4.8)	7 (9.6)

3.2.1.2 Demographics and baseline characteristics

Patient demographics and baseline characteristics are presented in Table 7 and Table 8 below.

Table 7. Patient Demographics (Study 352)

Variable	Ruxolitinib (N = 146)	BAT (N = 73)
Age (yrs)		
<=65 years, n (%)	69 (47.2)	36 (49.3)
> 65 years, n (%)	77 (52.8)	37 (50.7)
Sex, n (%)		
Male	83 (56.8)	42 (57.5)
Female	63 (43.2)	31 (42.5)
Race, n (%)		
White	118 (80.8)	67 (91.8)
Other	0	1 (1.4)
	28 (19.2)	33 (15)
Unknown	28 (19.2)	5 (6.8)

Table 8. Baseline characteristics (Study 352)

	Ruxolitinib (N = 146)	BAT (N = 73)
Type of MF – n (%)		
PMF	77 (52.7)	39 (53.4)
PPV-MF	48 (32.9)	20 (27.4)
PET-MF	21 (14.4)	14 (19.2)
Palpable spleen size (cm) below costal margin		
Mean (SD)	14.9 (6.45)	15.8 (6.71)
Spleen volume (cm ³)		
Mean (SD)	2662.1 (1351.26)	2631.1 (1405.27)
Prior hydroxyurea use – n (%)		
Yes	110 (75.3)	50 (68.5)
No	36 (24.7)	23 (31.5)
Prior splenic radiotherapy – n (%)		

Yes	0	4 (5.5)
No	146 (100)	69 (94.5)
V617F at baseline		
Yes	108 (74.0)	48 (65.8)
No	36 (24.7)	24 (32.9)
Unknown	2 (1.4)	1 (1.4)

3.2.1.2 Efficacy Results

A significantly larger proportion of subjects in the ruxolitinib group achieved a $\geq 35\%$ reduction from baseline at Week 48 and Week 24 compared with the placebo group ($p < 0.0001$ by Fisher's Exact test). Table 9 below presents a summary of results of the primary and key secondary endpoint.

Table 9. Results of primary and key secondary endpoint (Study 352)

	Ruxolitinib N=146	BAT N=73	% difference (95% CI ^a)	P value ^b
Primary Endpoint: % SVR $\geq 35\%$ at 48 weeks, n (%)	41 (28.1%)	0 (0)	28.1 (19.3, 34.8)	<0.0001
Key Secondary Endpoint: % SVR $\geq 35\%$ at 24 weeks, n (%)	46 (31.9%)	0 (0)	31.9 (22.5, 38.4)	<0.0001

a: By Agresti-Caffo method (Agresti and Caffo; The American Statistician, Vol. 54, No. 4, (Nov., 2000), pp.280-288)

b: Fisher's Exact test

Reviewers comment:

As specified in the protocol, patients with missing value for Week 48 and Week 24 spleen volumes were treated as non-responders for the primary and key secondary endpoints.

This study was originally planned with a sample size of 150. The actual number of patients enrolled is 219, which is 43% more than planned. However, even if the trial only accrued the first 150 patients, the result for the primary endpoint is still highly statistically significant (response rates are 26% and 0% in the ruxolitinib arm and BAT arm, respectively with p -value <0.0001).

Survival data are not mature. A total of 6 (4.1%) patients in the ruxolitinib arm and 4 (5.5%) patients in the BAT arm died.

3.3 Evaluation of Safety

Please refer Dr. Deisseroth's clinical review for efficacy evaluation.

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Race, Age, and Geographic Region

Results of subgroup analysis in age, gender and race for Study 351 and Study 352 are presented in Table 10 and 11 below. The proportions of subjects with $\geq 35\%$ reduction from baseline at Week 24 (Study 351) and Week 48 (Study 352) in the ruxolitinib group are consistently higher than those in the placebo group in all subgroups. However, with the response rate equal or close to 0 arm in both male and female groups within the placebo arm, the response rate in female group is much higher than that in the male group within the ruxolitinib arm (25% vs. 59% respectively in Study 351, $p < 0.0001$; 25% vs. 33% respectively in Study 351, $p = 0.26$).

Table 10. Subgroup analyses by gender, age and race (Study 351)

Subgroups	Ruxolitinib (n/N (%))	Placebo (n/N (%))	% difference	95% CI ^a for % difference
Gender				
Male	20/79 (25)	1/87 (1)	24	(14, 34)
Female	45/76 (59)	0/65 (0)	59	(46, 69)
Age				
≤ 65	32/70 (46)	0/52 (0)	46	(23, 56)
> 65	33/85 (39)	0/101 (0)	39	(28, 49)
Race				
White	56/138 (41)	1/138 (1)	40	(31, 48)
Non-White	9/17 (53)	0/14 (0)	53	(21, 72)

a: By Agresti-Caffo method (Agresti and Caffo; The American Statistician, Vol. 54, No. 4, (Nov., 2000), pp.280-288)

Table 11. Subgroup analyses by gender, age and race (Study 352)

Subgroups	Ruxolitinib (n/N (%))	BAT (n/N (%))	% difference	95% CI ^a for % difference
Gender				
Male	20/81 (25)	0/42 (0)	25	(13, 33)
Female	21/63 (33)	0/30 (0)	33	(21, 45)
Age				
≤ 65	20/68 (29)	0/36 (0)	29	(15, 39)
> 65	21/76 (28)	0/36 (0)	28	(14, 37)
Race				
White	33/116 (28)	0/67 (0)	28	(19, 36)
Non-White	0/0	0/1 (0)	NA	NA

a: By Agresti-Caffo method (Agresti and Caffo; The American Statistician, Vol. 54, No. 4, (Nov., 2000), pp.280-288)

4.2 Other Special/Subgroup Populations

Results of subgroup analysis based on important baseline characteristics are presented in Table 12 and 13 below. The proportions of subjects with $\geq 35\%$ reduction from baseline at Week 24

(Study 351) and Week 48 (Study 352) in the ruxolitinib group are consistently higher than those in the placebo group in all subgroups. However, with the response rate equal or close to 0 arm in both V617F positive and negative groups within the placebo arm, the response rate in V617F positive group is much higher than that in the V617F negative group within the ruxolitinib arm (48% vs. 28% respectively in Study 351, $p=0.03$; 33% vs. 14% respectively in Study 352, $p=0.03$).

Table 12. Subgroup analyses by baseline characteristics (Study 351)

Subgroups	Ruxolitinib (n/N (%))	Placebo (n/N (%))	% difference	95% CI ^a for % difference
HU-use				
User	29/52 (56)	1/51 (2)	54	(38, 66)
Non-user	36/103 (35)	0/102 (0)	35	(25, 44)
Tumor type				
PMF	27/70 (39)	1/82 (1)	37	(25, 48)
Post-PV	25/50 (50)	0/47 (0)	50	(34, 62)
Country				
USA	49/115 (43)	1/121 (1)	42	(38, 66)
Non-USA	16/40 (40)	0/32 (0)	40	(25, 44)
Baseline risk group				
High	32/90 (36)	1/98 (1)	35	(32, 50)
Intermediate	33/64 (52)	0/54 (0)	52	(22, 53)
V617F at baseline				
Positive	54/113 (48)	1/122 (1)	47	(37, 56)
Negative	11/40 (28)	0/27 (0)	28	(10, 40)

a: By Agresti-Caffo method (Agresti and Caffo; The American Statistician, Vol. 54, No. 4, (Nov., 2000), pp.280-288)

Table 13. Subgroup analyses by baseline characteristics (Study 352)

Subgroups	Ruxolitinib (n/N (%))	BAT (n/N (%))	% difference	95% CI ^a for % difference
HU-use				
User	28/108 (26)	0/49 (0)	26	(15, 33)
Non-user	13/36 (36)	0/23 (0)	36	(16, 50)
Tumor type				
PMF	14/76 (18)	0/38 (0)	18	(7, 27)
Post-PV	20/48 (42)	0/20 (0)	42	(21, 54)
PETM	7/20 (35)	0/14 (0)	35	(7, 53)
Baseline risk group				
High	17/70 (24)	0/35 (0)	24	(11, 34)
Intermediate	24/74 (32)	0/37 (0)	32	(19, 42)
V617F at baseline				
Positive	36/108 (33)	0/48 (0)	33	(27, 41)

Negative	5/35 (14)	0/20 (0)	14	(-3, 26)
----------	-----------	----------	----	----------

a: By Agresti-Caffo method (Agresti and Caffo; The American Statistician, Vol. 54, No. 4, (Nov., 2000), pp.280-288)

5. SUMMARY AND CONCLUSIONS

5.1 Statistical Issues and Collective Evidence

Ruxolitinib arm had higher proportions of subjects with $\geq 35\%$ reduction from baseline at Week 24 (Study 351) and Week 48 (Study 352) than the control arm. The results for the primary endpoint of Study 351 and 352 are presented in Table 14 below. The results of secondary endpoints, except for overall survival, in both studies are supportive for that of the primary endpoint. Survival results are not mature in both studies.

Table 14. Results of primary endpoints in Study 351 and Study 352

	Ruxolitinib	Placebo
Study 351		
Subjects achieving a $\geq 35\%$ reduction from baseline in spleen volume at Week 24, n (%)	65 (41.9)	1 (0.7)
% Difference between treatments (95% CI ^a)	41.2 (32.8, 48.7)	
p-value (by Fisher's Exact test)	<0.0001	
Study 352		
Subjects achieving a $\geq 35\%$ reduction from baseline in spleen volume at Week 48, n (%)	65 (41.9)	1 (0.7)
% Difference between treatments (95% CI ^a)	41.2 (32.8, 48.7)	
p-value (by Fisher's Exact test)	<0.0001	

a: By Agresti-Caffo method (Agresti and Caffo; The American Statistician, Vol. 54, No. 4, (Nov., 2000), pp.280-288)

The results of secondary endpoints, except for overall survival, in both studies are supportive for that of the primary endpoint (See Table 5 and Table 9). Survival results are not mature in both studies. Results of subgroup analyses are also consistent with that of the primary endpoint (see Tables 10-13).

Across subgroups, response rate of the primary endpoint is consistently higher in the ruxolitinib group than those in the placebo group. However, with the response rate equal or close to 0 arm in both male and female groups within the placebo arm, the response rate in female group is much higher than that in the male group within the ruxolitinib arm in Study 351 and Study 352 (25% vs. 59% respectively in Study 351, $p < 0.0001$; 25% vs. 33% respectively in Study 351, $p = 0.26$). Similarly, the response rate in V617F positive group is much higher than that in the V617F negative group within the ruxolitinib arm in both Study 351 and Study 352 (48% vs. 28% respectively in Study 351, $p = 0.03$; 33% vs. 14% respectively in Study 351, $p = 0.03$).

5.2 Conclusions and Recommendations

Study 351 and Study 352 demonstrated that, compared with those in the control arm, patients in the ruxolitinib arm were more likely to achieved at least a 35% reduction in spleen volume from baseline at Week 24 in Study 351 (42% vs. 1%) and Week 48 in Study 352 (28% vs. 0%).

Across subgroups, response rate of the primary endpoint is consistently higher in the ruxolitinib group than those in the placebo group. However, the response rate in female group is much higher than that in the male group within the ruxolitinib arm in both studies (25% vs. 59% respectively in Study 351; 25% vs. 33% respectively in Study 351). And the response rate in V617F positive group is much higher than that in the V617F negative group within the ruxolitinib arm in both studies (48% vs. 28% respectively in Study 351; 33% vs. 14% respectively in Study 351).

CHECK LIST

Number of Pivotal Studies: 2

Study 351

Trial Specification

Specify for each trial:

Protocol Number (s): 351

Phase: 3

Control: Placebo Control

Blinding: double blind

Number of Centers: 89

Region(s) (Country): US, Canada, Australia

Duration: Five cycles. Five weeks in each cycle.

Treatment Arms: ruxolitinib

Treatment Schedule: Twice daily with varying dose

Randomization: Yes

Ratio: 1:1

Method of Randomization: central randomization without stratification

Primary Endpoint: Proportion of subjects achieved a >35% reduction from baseline at Week 24.

Primary Analysis Population: ITT

Statistical Design: Superiority

Primary Statistical Methodology: Fisher's exact test

Interim Analysis: No

Sample Size: 309 (planned at 240 subjects)

Sample Size Determination: Was it calculated based on the primary endpoint variable and the analysis being used for the primary variable? Yes

Statistic = Fisher's exact

Power= 0.97

Response rate: 30% vs. 10%

α = 0.05

- Were there any major changes, such as changing the statistical analysis methodology or changing the primary endpoint variable? No.
- Did the Applicant perform **Sensitivity Analyses**? No.

How were the **Missing Data** handled? A subject with a missing Week 24 spleen volume will be considered as having not achieved the >35% reduction. Subjects who drop out of the study due to lack of efficacy or treatment-related adverse events, or make an early crossover to active treatment (placebo subjects only), prior to Week 24 visit will all be considered as having not

achieved the >35% reduction. Missing value for total symptoms scores are excluded from the analysis.

- Was there a **Multiplicity** involved? Yes.
- **Multiple Secondary Endpoints**: Are they being included in the label? Yes. Gate keeper method is used to adjust for multiplicity.

Were Subgroup Analyses Performed? Yes

- Were there any **Discrepancies** between the protocol/statistical analysis plan vs. the study report? No.
- Overall, was the study positive? Yes.

Study 352

Trial Specification

Specify for each trial:

Protocol Number (s): 352

Phase: 3

Control: Placebo Control

Blinding: Open-label

Number of Centers: 57

Region(s) (Country): Europe

Duration: Five cycles. Five weeks in each cycle.

Treatment Arms: ruxolitinib

Treatment Schedule: Twice daily with varying dose

Randomization: Yes

Ratio: 2:1

Method of Randomization: central randomization stratified by

- Stratum 1: Intermediate risk level 2 (2 risk factors)
- Stratum 2: High risk (3 or more risk factors)

Primary Endpoint: Proportion of subjects achieved a >35% reduction from baseline at Week 48.

Primary Analysis Population: ITT

Statistical Design: Superiority

Primary Statistical Methodology: Fisher's exact test

Interim Analysis: No

Sample Size: 219 (planned at 150 subjects)

Sample Size Determination: Was it calculated based on the primary endpoint variable and the analysis being used for the primary variable? Yes

Statistic = Fisher's exact

Power= 0.9

Response rate: 35% vs. 10%

$\alpha = 0.05$

- Were there any major changes, such as changing the statistical analysis methodology or changing the primary endpoint variable? No.

- Did the Applicant perform **Sensitivity Analyses**? No.

How were the **Missing Data** handled? A subject with a missing Week 48 spleen volume will be considered as having not achieved the >35% reduction. Subjects who drop out of the study due to lack of efficacy or treatment-related adverse events, or make an early crossover to active treatment (placebo subjects only), prior to Week 24 visit will all be considered as having not achieved the >35% reduction.

- Was there a **Multiplicity** involved? No

- **Multiple Secondary Endpoints**: Are they being included in the label? No.

- **Were Subgroup Analyses Performed**? Yes

- Were there any **Discrepancies** between the protocol/statistical analysis plan vs. the study report? No.

- Overall, was the study positive? Yes.

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

/s/

HONG LU
09/19/2011

MARK D ROTHMANN
09/19/2011

RAJESHWARI SRIDHARA
09/19/2011

Statistical Review and Evaluation CARCINOGENICITY STUDIES



IND/NDA Number: NDA 202-192
Drug name: INCB018424 (Phosphate salt)
Indication(s): Myelofibrosis
Applicant: Incyte
Documents Reviewed: Electronic submission
Electronically submitted dataset
Dated: 2011-06-03
Review Priority: Normal
Biometrics Division: Division of Biometrics 6
Statistical Reviewer: Matthew Jackson, PhD
Concurring Reviewer: Karl Lin, PhD
Medical Division: DHP
Reviewing Pharmacologists: Miyun M. Tsai-Turton, PhD
Wei Chen, PhD
Project Manager:
Keywords: Animal Mouse Carcinogenicity

Contents

0.1	Background	4
0.2	Overview	4
1	Carcinogenicity study	5
1.1	Sponsor's analyses	5
1.1.1	Survival analysis	5
1.1.2	Tumor data analysis	5
1.2	Reviewer's analysis	5
1.2.1	Survival analysis	5
1.2.2	Tumor analysis	9
1.3	Unexamined and Autolytic organs	16
1.3.1	Missing animals	16
1.3.2	Unexamined organs	16
1.3.3	Autolytic organs	16
2	Urethane study	17
2.1	Sponsor's analyses	17
2.1.1	Survival analysis	17
2.1.2	Tumor data analysis	17
2.2	Reviewer's analysis	17
2.2.1	Survival analysis	17
2.2.2	Tumor analysis	19
3	Evaluation of the validity of negative studies	23
3.1	Issues of concern when selecting the dose levels	23
3.2	Assessment of the validity of the carcinogenicity study	23
3.3	Assessment of the validity of the urethane study	24
4	Conclusion	25
4.1	Carcinogenicity study	25
4.1.1	Tumor findings	25
4.2	Urethane study	25
A	Tumor incidence calculations with all treated groups combined	26

List of Tables

1.1	Survival rates at key times	7
1.2	Results of log-rank tests of survival	7
1.3	Log-rank tests of survival between treated groups and control	8
1.4	Organs reported as being analyzed in most animals	11
1.5	Combination endpoints tested	11
1.6	Reported neoplastic tumors in study of female mice	12
1.7	Reported neoplastic tumors in study of male mice	13
1.8	Reported combinations in study of female mice	15
1.9	Reported combinations in study of male mice	15
2.1	Survival rates at key times (urethane study)	18
2.2	Log-rank tests of survival between urethane group and control (urethane study)	18
2.3	Reported neoplastic tumors in study of female mice (urethane study)	21
2.4	Reported neoplastic tumors in study of male mice (urethane study)	21
3.1	Weight changes by group	24
A.1	Reported neoplastic tumors in study of female mice (treated groups combined)	27
A.2	Reported neoplastic tumors in study of male mice (treated groups combined)	28
A.3	Reported combinations in study of female mice (treated groups combined)	30
A.4	Reported combinations in study of male mice (treated groups combined)	30

List of Figures

1.1	Survival curves for female mice	9
1.2	Survival curves for male mice	10
2.1	Survival curves for female mice (urethane study)	19
2.2	Survival curves for male mice (urethane study)	20

0.1 Background

In this submission the sponsor included reports of one animal carcinogenicity study, in Tg.rasH2 Mice. This study were intended to assess the carcinogenic potential of INCB018424 when administered by dermally, once daily at appropriate drug levels for about 26 weeks. Results of this review have been discussed with the reviewing pharmacologist, Dr. Tsai-Turton.

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.

0.2 Overview

Two separate studies were conducted; one each in male and female mice. Each study involved five groups of twenty five animals. One group was the control group, receiving by gavage a daily dose (10mL per kilogram of bodyweight) of the the vehicle, 0.5% methylcellulose. Three groups received doses of the test article, INCB018424, at daily doses of 15, 45, and 125 mg per kilogram of body weight, in the same vehicle (10 mL/kg). The fifth group was a positive control group, which recieved intraperitoneal injections of urethane (1g/kg) on days 1, 3, and 5. Animals in this group were sacrificed after 119 days (females) or 121 days (males), rather than at the end of the 26 week study.

Chapter 1

Carcinogenicity study

This chapter concerns the comparison of outcomes between the groups of animals treated with INCB018424 and the vehicle control.

1.1 Sponsor's analyses

1.1.1 Survival analysis

Kaplan-Meier survival estimates were constructed for each group within each sex, and the generalized Wilcoxon test was used to test for heterogeneity of survival between the INCB018424 treated groups and the control. Among males, no statistically significant findings were reported. Among females, the mid dose group were found to have significantly reduce survival compared with the control group, although no p -values are presented.

1.1.2 Tumor data analysis

The sponsor tested each reported tumor type, in each sex, separately, using Peto's [5] mortality prevalence method. For each tumor type, each treated group was compared with control, and a 1-tailed trend test was conducted across all four groups. Exact versions of the tests were used for tumor types with lower prevalence (how low is not specified), and tests were conducted at both the 1% and 5% significance levels.

No significant results were reported.

1.2 Reviewer's analysis

To verify the sponsor's analyses and to perform additional analysis suggested by the reviewing pharmacologist, independent survival and tumor data analyses were performed. Data used in these analyses were provided by the sponsor electronically.

1.2.1 Survival analysis

Intercurrent mortality data are presented in table 1.1. The results of the log-rank tests of survival (both tests of trend and of heterogeneity) are presented in table 1.2, and the results of log-rank tests of survival between individual treated groups and the control group are presented in table 1.3. Kaplan-Meier survival plots are displayed as figures 1.1 and 1.2. In both plots, the curve representing the control group cannot be seen because it coincides exactly with the curve for the high dose group; in both the studies of male and female mice, no early deaths were reported in either the control or high dose group.

Among females, there is evidence ($p = 0.0253$) of heterogeneity of survival across the dose groups, and for both sexes there is a statistically significant indication of increased mortality ($p = 0.0090$

for females and $p = 0.0311$ for males) in the mid dose group compared with the control group. However, there is no evidence of a dose related increase in mortality in males or females in either study, and it should be noted that the mortality rates were extremely low in both sexes — only nine mice in total (five female and four male) died before the end of the 26 week study. The significant results noted above are therefore based on the analyses of very small numbers of deaths.

It should also be noted that because deaths are so rare in this study, the various statistical tests used have very low power to detect even substantial increases in relative risk.

Table 1.1: Survival rates at key times

Species and Sex	Dose Group	Dose (mg per kg)	Number and percentage alive								
			Start	12 weeks (%)	19 weeks (%)	22 weeks (%)	Termination (%)				
Mice — Female	Control	0	25	25	25	25	25	100%	100%	25	100%
	Low dose	15	25	25	25	25	25	100%	100%	24	96%
	Mid dose	45	25	25	24	24	24	96%	96%	21	84%
	High dose	125	25	25	25	25	25	100%	100%	25	100%
Mice — Male	Control	0	25	25	25	25	25	100%	100%	25	100%
	Low dose	15	25	25	25	25	24	100%	96%	24	96%
	Mid dose	45	25	25	24	24	24	96%	96%	22	88%
	High dose	125	25	25	25	25	25	100%	100%	25	100%

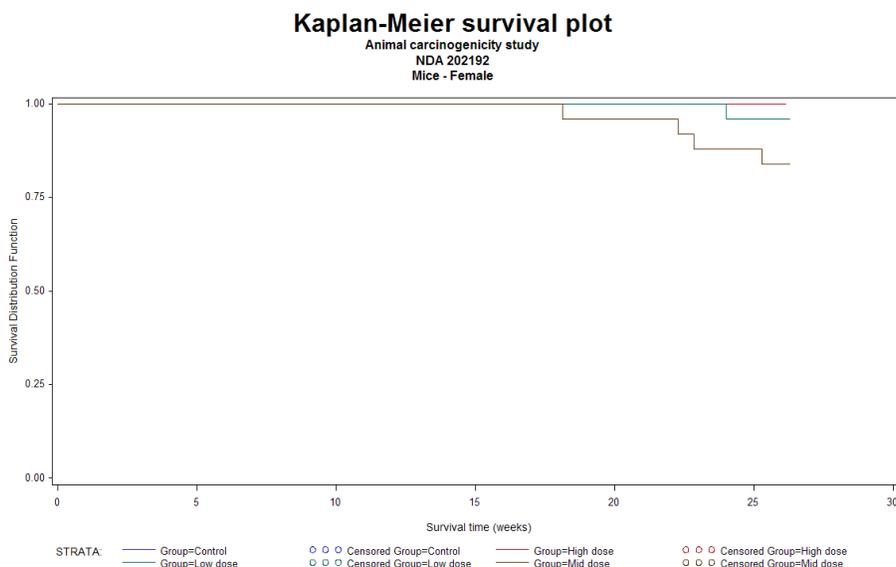
Table 1.2: Results of log-rank tests of survival

Species and Sex	Number of Groups	Test of homogeneity			Test of trend	
		χ^2 statistic	df	p -value	Two tailed p -value	One tailed p -value
Mice — Female	4	9.3181	3	0.0253	0.7364	0.6318
Mice — Male	4	6.2422	3	0.1004	0.7134	0.6433

Table 1.3: Log-rank tests of survival between treated groups and control

Species and Sex		Low dose	Mid dose	High dose
Mice — Female	χ^2 test statistic	0.4018	6.8189	0.0000
	p -value of comparison with control	0.5262	0.0090	1.0000
Mice — Male	χ^2 test statistic	0.5181	4.6488	0.0000
	p -value of comparison with control	0.4716	0.0311	1.0000

Figure 1.1: Survival curves for female mice



1.2.2 Tumor analysis

Theoretical underpinnings

The tumor data were analyzed for dose response relationships and pairwise comparisons of tumor incidence in each of the treated groups versus the vehicle 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

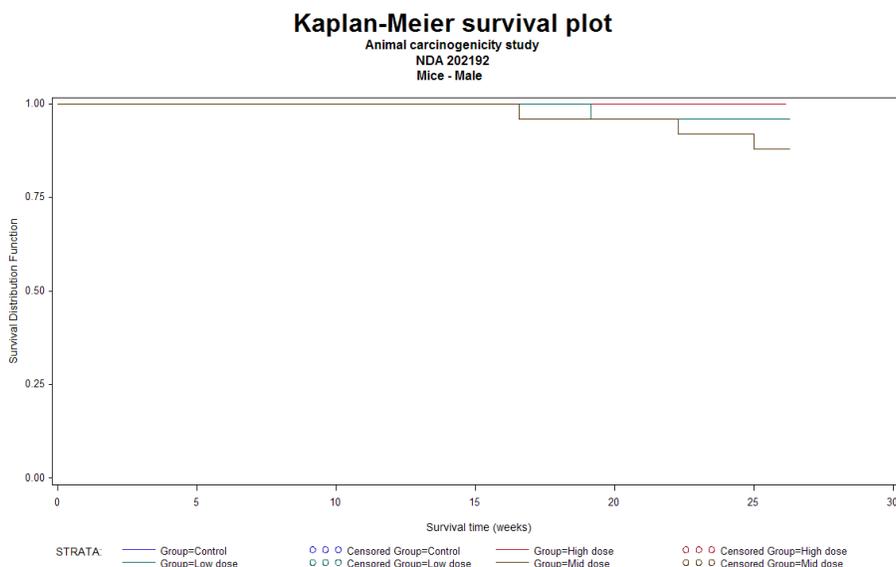
$$s_h = \left(\frac{w_h}{w_m} \right)^k < 1.$$

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. There is no consensus for the correct value to use for studies of transgenic mice. In the absence of such a consensus, this review uses the value $k = 1$, a value which is consistent with the assumption of constant hazard over the twenty six week period of the study. In any event, when there is little premature mortality (as is the case with this study — see section 1.2.1), the analyses are not very sensitive to variations in the value of k .

For the calculation of p -values, the exact permutation method was used. The tumor rates and the p -values of the tested tumor types are listed in tables 1.6 and 1.7. Corresponding tables for combination endpoints are presented in tables 1.8 and 1.9.

Figure 1.2: Survival curves for male mice



Under normal circumstances, since so many end points are being tested, it is appropriate to make some sort of multiplicity adjustment in order to control type I error. However, in the case of transgenic mice there is no guidance specifying how this should be done. Furthermore, in light of the fact that exact tests tend to be very conservative when considering rare events, the fact that there are only twenty five animals in each group, and the fact that tumorigenesis is very rare over the twenty six weeks that transgenic mouse studies typically run, it seems reasonable to consider test as having yielded positive findings whenever the p -value is below 0.05.

Analysis of tumor data

The electronic dataset used for this review does not list negative findings; organ-level records are present when neoplasms are found and when examinations conducted in some animals are not conducted in others. It follows that organs in which no tumors are found in any animals are not included in the submitted dataset. With this in mind, table 1.4 lists the organs which can definitely be inferred from the submitted dataset to have been examined in all or most animals. In the case, since no organs have been reported as being unexamined, we can actually infer that these organs underwent microscopic examination in all animals.

In addition, a number of combination endpoints were also tested at the pharmacology reviewer's request. These are listed in table 1.5.

The results of the statistical analyses of individual tumor types are reported in tables 1.6 and 1.7. Results of analyses of combination tumor types are presented in tables 1.8 and 1.9.

In addition to these standard analyses, and at the request of the reviewing pharmacologist, these analyses have been repeated with all three treated groups combined. The main advantage of combining the treated groups in this way is that when the a tumorigenic effect is believed to be sensitive to relatively low levels of the suspect tumorigenic agent. The results of these analyses are in the Appendix A. Note that in these tables, the p -values of the tests of trend are the same as the comparisons between the control and amalgamated treated group. This is an automatic consequence of the fact that the trend test is being taken across just two groups (control and treated).

No statistical tests, of either a single tumor type or combination yielded a p -value below 0.05. The study is therefore a negative study.

Table 1.4: Organs reported as being analyzed in most animals

Female mouse study		
adrenal glands	aorta	bone marrow, femur
bone marrow, sternum	bone, femur	bone, sternum
brain	cavity, nasal	esophagus
eyes	gall bladder	harderian glands
heart	intestine, cecum	intestine, colon
intestine, duodenum	intestine, ileum	intestine, jejunum
intestine, rectum	kidneys	liver
lungs with bronchi	lymph node, mandibular	lymph node, mediastinal
lymph node, mesenteric	mammary gland	multicentric
nerve, sciatic	ovaries	pancreas
parathyroid glands	pituitary gland	salivary glands
skeletal muscle (thigh)	skin	spinal cord, cervical
spinal cord, lumbar	spinal cord, thoracic	spleen
stomach	thymus	thyroid glands
trachea	urinary bladder	uterus
vagina		
Male mouse study		
adrenal glands	aorta	bone marrow, femur
bone marrow, sternum	bone, femur	bone, sternum
brain	cavity, nasal	epididymides
esophagus	eyes	gall bladder
harderian glands	heart	intestine, cecum
intestine, colon	intestine, duodenum	intestine, ileum
intestine, jejunum	intestine, rectum	kidneys
liver	lungs with bronchi	lymph node, mandibular
lymph node, mediastinal	lymph node, mesenteric	mammary gland
nerve, sciatic	pancreas	parathyroid glands
pituitary gland	prostate gland	salivary glands
seminal vesicles	skeletal muscle (thigh)	skin
spinal cord, cervical	spinal cord, lumbar	spinal cord, thoracic
spleen	stomach	testes
thymus	thyroid glands	trachea
urinary bladder		

Table 1.5: Combination endpoints tested

All hemangiosarcomas, regardless of site
All skin tumors

Table 1.6: Reported neoplastic tumors in study of female mice

Organ	Tumor type	Quantity	Control	Low dose	Mid dose	High dose
cavity, nasal	adenocarcinoma	<i>p</i> -value of test of trend or comparison	.9375	.4844	1	1
		Number of animals reported with tumor	1	2	0	0
		Poly-1 adjusted incidence rate	4.0%	8.0%	0.0%	0.0%
		95% CI for Poly-1 adjusted incidence rate (%)	(0.1,20.4)	(0.98,27.0)	(0,14.2)	(0,13.7)
		Poly-1 adjusted number of animals at risk	25.0	24.9	24.4	25.0
		<i>p</i> -value of test of trend or comparison	.2822	1	1	.6954
		Number of animals reported with tumor	2	0	0	2
		Poly-1 adjusted incidence rate	8.0%	0.0%	0.0%	8.0%
		95% CI for Poly-1 adjusted incidence rate (%)	(0.98,26.0)	(0,14.2)	(0,14.2)	(0.98,26.0)
		Poly-1 adjusted number of animals at risk	25.0	24.9	24.4	25.0
lungs with bronchi	alveolar–bronchiolar adenoma	<i>p</i> -value of test of trend or comparison	.8929	.6465	.9403	.9451
		Number of animals reported with tumor	3	3	1	1
		Poly-1 adjusted incidence rate	12%	12%	4.1%	4.0%
		95% CI for Poly-1 adjusted incidence rate (%)	(2.55,31.2)	(2.55,32.4)	(0.1,21.1)	(0.1,20.4)
		Poly-1 adjusted number of animals at risk	25.0	24.9	24.4	25.0
		<i>p</i> -value of test of trend or comparison	.7449	.4898		
		Number of animals reported with tumor	0	1	0	0
		Poly-1 adjusted incidence rate	0.0%	4.0%	0.0%	0.0%
		95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(0.1,21.1)	(0,14.2)	(0,13.7)
		Poly-1 adjusted number of animals at risk	25.0	24.9	24.4	25.0
multicentric	sarcoma	<i>p</i> -value of test of trend or comparison	.5000		.4898	
		Number of animals reported with tumor	0	0	1	0
		Poly-1 adjusted incidence rate	0.0%	0.0%	4.1%	0.0%
		95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(0,14.2)	(0.1,21.1)	(0,13.7)
		Poly-1 adjusted number of animals at risk	25.0	24.9	24.5	25.0
		<i>p</i> -value of test of trend or comparison	.7475	.5000		
		Number of animals reported with tumor	0	1	0	0
		Poly-1 adjusted incidence rate	0.0%	0.0%	4.1%	0.0%
		95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(0,14.2)	(0.1,21.1)	(0,13.7)
		Poly-1 adjusted number of animals at risk	25.0	24.9	24.5	25.0
skin	hemangiosarcoma	<i>p</i> -value of test of trend or comparison	.7475	.5000		
		Number of animals reported with tumor	0	1	0	0
		Poly-1 adjusted incidence rate	0.0%	4.0%	0.0%	0.0%
		95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(0.1,20.4)	(0,14.2)	(0,13.7)
		Poly-1 adjusted number of animals at risk	25.0	25.0	24.4	25.0
		<i>p</i> -value of test of trend or comparison	.7502	.2347	.2347	
		Number of animals reported with tumor	0	2	2	0
		Poly-1 adjusted incidence rate	0.0%	8.0%	8.1%	0.0%
		95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(0.98,27.0)	(0.98,27.0)	(0,13.7)
		Poly-1 adjusted number of animals at risk	25.0	24.9	24.4	25.0
spleen	hemangiosarcoma	<i>p</i> -value of test of trend or comparison	.7502	.2347	.2347	
		Number of animals reported with tumor	0	2	2	0
		Poly-1 adjusted incidence rate	0.0%	8.0%	8.1%	0.0%
		95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(0.98,27.0)	(0.98,27.0)	(0,13.7)
		Poly-1 adjusted number of animals at risk	25.0	24.9	24.4	25.0
		<i>p</i> -value of test of trend or comparison	.7502	.2347	.2347	
		Number of animals reported with tumor	0	2	2	0
		Poly-1 adjusted incidence rate	0.0%	8.0%	8.1%	0.0%
		95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(0.98,27.0)	(0.98,27.0)	(0,13.7)
		Poly-1 adjusted number of animals at risk	25.0	24.9	24.4	25.0

Table 1.6: Reported neoplastic tumors in study of female mice (continued)

Organ	Tumor type	Quantity	Control	Low dose	Mid dose	High dose
stomach	Poly-1 adjusted number of animals at risk		25.0	24.9	24.5	25.0
	<i>p</i> -value of test of trend or comparison		1	1	1	1
	Number of animals reported with tumor		1	0	0	0
	Poly-1 adjusted incidence rate		4.0%	0.0%	0.0%	0.0%
thymus	95% CI for Poly-1 adjusted incidence rate (%)		(0.1,20.4)	(0,14.2)	(0,14.2)	(0,13.7)
	Poly-1 adjusted number of animals at risk		25.0	24.9	24.4	25.0
	<i>p</i> -value of test of trend or comparison		.5551		.1099	
	Number of animals reported with tumor		0	0	3	0
	Poly-1 adjusted incidence rate		0.0%	0.0%	12%	0.0%
	95% CI for Poly-1 adjusted incidence rate (%)		(0,13.7)	(0,14.2)	(2.55,32.4)	(0,13.7)
	Poly-1 adjusted number of animals at risk		25.0	24.9	24.4	25.0

Table 1.7: Reported neoplastic tumors in study of male mice

Organ	Tumor type	Quantity	Control	Low dose	Mid dose	High dose
cavity, nasal	adenocarcinoma	<i>p</i> -value of test of trend or comparison	.7449	.4898		
		Number of animals reported with tumor	0	1	0	0
		Poly-1 adjusted incidence rate	0.0%	4.0%	0.0%	0.0%
		95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(0,1,21.1)	(0,14.2)	(0,13.7)
		Poly-1 adjusted number of animals at risk	25.0	24.7	24.5	25.0
		<i>p</i> -value of test of trend or comparison	.6044	1	1	.8827
		Number of animals reported with tumor	2	0	0	1
liver	hepatocellular adenoma	Poly-1 adjusted incidence rate	8.0%	0.0%	0.0%	4.0%
		95% CI for Poly-1 adjusted incidence rate (%)	(0.98,26.0)	(0,14.2)	(0,14.2)	(0.1,20.4)
		Poly-1 adjusted number of animals at risk	25.0	24.7	24.5	25.0
		<i>p</i> -value of test of trend or comparison	.5000		.4898	
		Number of animals reported with tumor	0	0	1	0
		Poly-1 adjusted incidence rate	0.0%	0.0%	4.1%	0.0%
		95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(0,14.2)	(0.1,21.1)	(0,13.7)
lungs with bronchi	alveolar–bronchiolar adenoma	Poly-1 adjusted number of animals at risk	25.0	24.7	24.5	25.0
		<i>p</i> -value of test of trend or comparison	.5138	.6798	.4800	.6954
		Number of animals reported with tumor	2	2	3	2
		Poly-1 adjusted incidence rate	8.0%	8.1%	12%	8.0%
		95% CI for Poly-1 adjusted incidence rate (%)	(0.98,26.0)	(0.98,27.0)	(2.55,32.4)	(0.98,26.0)
		Poly-1 adjusted number of animals at risk	25.0	24.7	24.5	25.0

Table 1.7: Reported neoplastic tumors in study of male mice (continued)

Organ	Tumor type	Quantity	Control	Low dose	Mid dose	High dose
skeletal muscle (thigh)	hemangioma	<i>p</i> -value of test of trend or comparison	.2551			.5000
		Number of animals reported with tumor	0	0	0	1
		Poly-1 adjusted incidence rate	0.0%	0.0%	0.0%	4.0%
		95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(0,14.2)	(0,14.2)	(0,1,20.4)
		Poly-1 adjusted number of animals at risk	25.0	24.7	24.5	25.0
		<i>p</i> -value of test of trend or comparison	.7475	.5000		
skin	sarcoma	Number of animals reported with tumor	0	1	0	0
		Poly-1 adjusted incidence rate	0.0%	4.0%	0.0%	0.0%
		95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(0,1,20.4)	(0,14.2)	(0,13.7)
		Poly-1 adjusted number of animals at risk	25.0	25.0	24.5	25.0
		<i>p</i> -value of test of trend or comparison	.5000		.4898	
		Number of animals reported with tumor	0	0	1	0
spleen	sarcoma	Poly-1 adjusted incidence rate	0.0%	0.0%	4.1%	0.0%
		95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(0,14.2)	(0,1,21.1)	(0,13.7)
		Poly-1 adjusted number of animals at risk	25.0	24.7	24.5	25.0
		<i>p</i> -value of test of trend or comparison	.5000		.4898	
		Number of animals reported with tumor	0	0	1	0
		Poly-1 adjusted incidence rate	0.0%	0.0%	4.0%	0.0%
testes	hemangioma	95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(0,14.2)	(0,1,21.1)	(0,13.7)
		Poly-1 adjusted number of animals at risk	25.0	24.7	24.8	25.0
		<i>p</i> -value of test of trend or comparison	.1894		.4898	.5000
		Number of animals reported with tumor	0	0	1	1
		Poly-1 adjusted incidence rate	0.0%	0.0%	4.1%	4.0%
		95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(0,14.2)	(0,1,21.1)	(0,1,20.4)
thyroid glands	adenoma	Poly-1 adjusted number of animals at risk	25.0	24.7	24.5	25.0
		<i>p</i> -value of test of trend or comparison	.7449	.4898		
		Number of animals reported with tumor	0	1	0	0
		Poly-1 adjusted incidence rate	0.0%	4.0%	0.0%	0.0%
		95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(0,1,21.1)	(0,14.2)	(0,13.7)
		Poly-1 adjusted number of animals at risk	25.0	24.7	24.5	25.0

Table 1.8: Reported combinations in study of female mice

Endpoint	Quantity	Control	Low dose	Mid dose	High dose
All hemangiosarcomas	<i>p</i> -value of test of trend or comparison	.8219	.1173	.2347	
	Number of animals reported with tumor	0	3	2	0
	Poly-1 adjusted incidence rate	0.0%	12%	8.1%	0.0%
	95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(2.55,31.2)	(0.98,27.0)	(0,13.7)
	Poly-1 adjusted number of animals at risk	25.0	25.0	24.5	25.0
All skin tumors	<i>p</i> -value of test of trend or comparison	.7475	.5000		
	Number of animals reported with tumor	0	1	0	0
	Poly-1 adjusted incidence rate	0.0%	4.0%	0.0%	0.0%
	95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(0.1,20.4)	(0,14.2)	(0,13.7)
	Poly-1 adjusted number of animals at risk	25.0	25.0	24.4	25.0

Table 1.9: Reported combinations in study of male mice

Endpoint	Quantity	Control	Low dose	Mid dose	High dose
All hemangiosarcomas	<i>p</i> -value of test of trend or comparison	.3758	1	.4844	.7551
	Number of animals reported with tumor	1	0	2	1
	Poly-1 adjusted incidence rate	4.0%	0.0%	8.2%	4.0%
	95% CI for Poly-1 adjusted incidence rate (%)	(0.1,20.4)	(0,14.2)	(0.98,27.0)	(0.1,20.4)
	Poly-1 adjusted number of animals at risk	25.0	24.7	24.5	25.0
All skin tumors	<i>p</i> -value of test of trend or comparison	.5052		.2347	
	Number of animals reported with tumor	0	0	2	0
	Poly-1 adjusted incidence rate	0.0%	0.0%	8.0%	0.0%
	95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(0,14.2)	(0.98,27.0)	(0,13.7)
	Poly-1 adjusted number of animals at risk	25.0	24.7	24.9	25.0

1.3 Unexamined and Autolytic organs

1.3.1 Missing animals

There are no animals reported as being completely unexamined.

1.3.2 Unexamined organs

With the exception of the mediastinum, which as reported as being unexamined in every animal, no organs are reported as being unexamined.

1.3.3 Autolytic organs

No organs are reported as being autolyzed to the extent that a usable sample could not be obtained.

Chapter 2

Urethane study

2.1 Sponsor's analyses

2.1.1 Survival analysis

Kaplan-Meier survival estimates were constructed for each group within each sex, and the generalized Wilcoxon test was used to test for a difference in survival between the urethane groups and the controls. In both sexes, the animals treated with urethane experienced a significant reduction in survival compared with the control group. No statistics or p -values were reported.

2.1.2 Tumor data analysis

The sponsor tested each reported tumor type, in each sex, separately, using Peto's [5] mortality prevalence method. For each tumor type, the urethane group was compared with the control group. Exact versions of the tests were used for tumor types with lower prevalence (how low is not specified), and tests were conducted at both the 1% and 5% significance levels.

Significant results are reported for alveolar-bronchiolar adenomas, carcinomas, and for adenomas and carcinomas combined, for splenic hemangiosarcomas, and for the combination endpoints of hemangioma and hemangiosarcomas at all sites, and mesenchymal tumors at all sites. This level of specificity is slightly misleading however, as the only tumors reported in the urethane group were bronciolo-alveolar adenomas and carcinomas, and hemangiosarcomas.

2.2 Reviewer's analysis

To verify the sponsor's analyses and to perform additional analysis suggested by the reviewing pharmacologist, independent survival and tumor data analyses were performed. Data used in these analyses were provided by the sponsor electronically.

2.2.1 Survival analysis

Intercurrent mortality data are presented in table 2.1 (although note that the termination date for the urethane animals was 119 days (female animals) or 121 days (male animals), whereas the termination date for the control group was 182 days). The results of the log-rank test of survival between the urethane group and the control group are presented in table 2.2. Kaplan-Meier survival plots are displayed as figures 2.1 and 2.2.

It is abundantly clear that the urethane treatment is associated with sharply increased mortality.

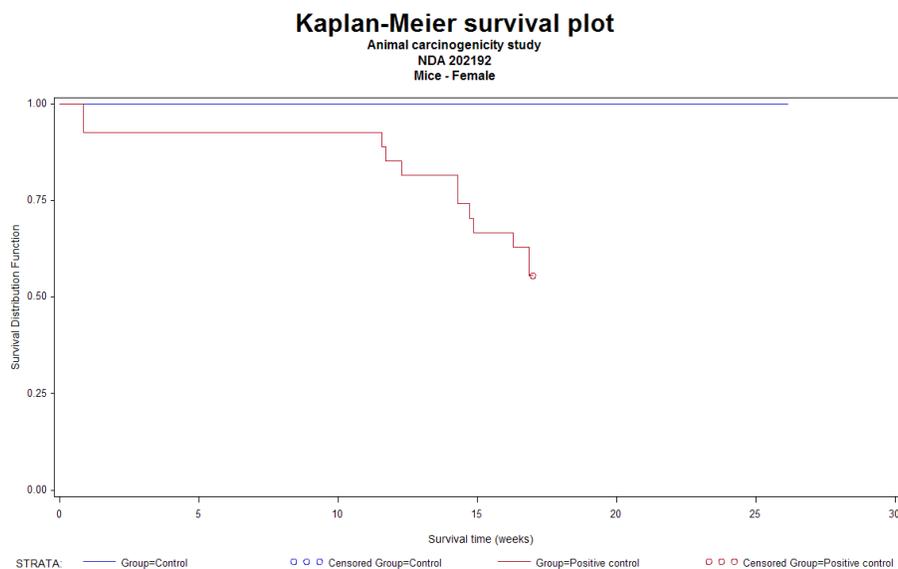
Table 2.1: Survival rates at key times (urethane study)

Species and Sex	Dose Group	Dose (mg per kg)	Number and percentage alive			
			Start	12 weeks (%)	Termination	(%)
Mice — Female	Control	0	25	25 100%	25	100%
	Positive control	1000	27	23 85%	15	60.0%
Mice — Male	Control	0	25	25 100%	25	100%
	Positive control	1000	25	24 96%	13	52.0%

Table 2.2: Log-rank tests of survival between urethane group and control (urethane study)

Species and Sex		Low dose	Mid dose	High dose
Mice — Female	χ^2 test statistic	0.4018	6.8189	0.0000
	p -value of comparison with control	0.5262	0.0090	1.0000
Mice — Male	χ^2 test statistic	0.5181	4.6488	0.0000
	p -value of comparison with control	0.4716	0.0311	1.0000

Figure 2.1: Survival curves for female mice (urethane study)



2.2.2 Tumor analysis

Analysis of tumor data The animals in the urethane group only had their lungs, with bronchi, and spleens examined microscopically. There were no exceptions to this; every urethane treated animal had these two organs examined, and none had any additional organs examined. Accordingly, comparison with the control group is only possible for tumors in these sites.

As can be seen, administration of urethane is strongly associated with an increased incidence of alveolar-bronchiolar adenomas and carcinomas and splenic hemangiosarcomas in both male and female mice, and of hemangiosarcomas in the lungs and bronchi in male mice.

Figure 2.2: Survival curves for male mice (urethane study)

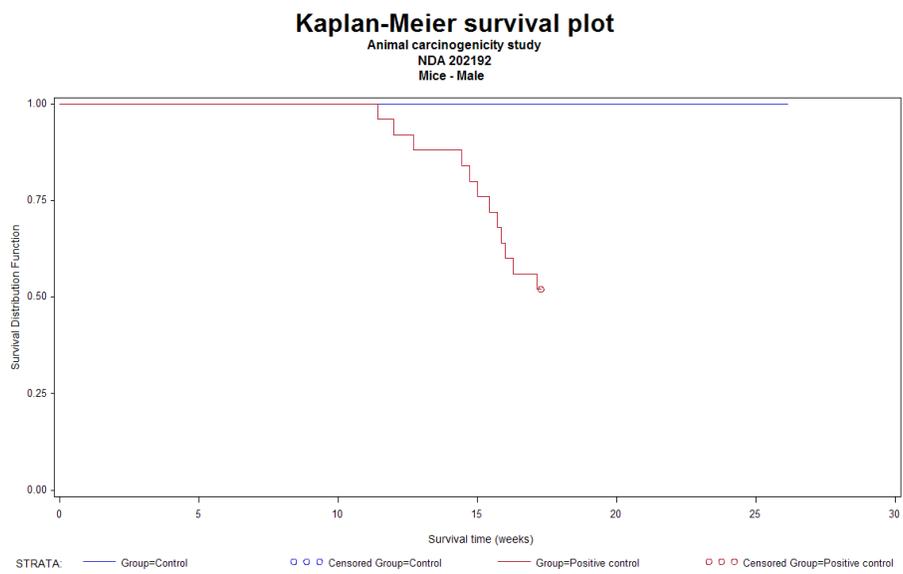


Table 2.3: Reported neoplastic tumors in study of female mice (urethane study)

Organ	Tumor type	Quantity	Control	Treated
lungs with bronchi	alveolar–bronchiolar adenoma	P–value of test of trend or comparison	< 0.0001	< 0.0001
		Number of animals reported with tumor	3	25
		Poly–1 adjusted incidence rate	12%	100%
		95% CI for poly–1 adjusted incidence rate (%)	(2.55,31.2)	(86.3,100)
		Poly–1 adjusted number of animals at risk	25.0	25.0
		P–value of test of trend or comparison	< 0.0001	< 0.0001
		Number of animals reported with tumor	0	18
		Poly–1 adjusted incidence rate	0.0%	82%
		95% CI for poly–1 adjusted incidence rate (%)	(0,13.7)	(59.7,97.0)
		Poly–1 adjusted number of animals at risk	25.0	21.9
spleen	hemangiosarcoma	P–value of test of trend or comparison	.3750	.3750
		Number of animals reported with tumor	0	1
		Poly–1 adjusted incidence rate	0.0%	6.4%
		95% CI for poly–1 adjusted incidence rate (%)	(0,13.7)	(0.16,31.9)
		Poly–1 adjusted number of animals at risk	25.0	15.7
		P–value of test of trend or comparison	< 0.0001	< 0.0001
		Number of animals reported with tumor	0	24
		Poly–1 adjusted incidence rate	0.0%	98%
		95% CI for poly–1 adjusted incidence rate (%)	(0,13.7)	(79.6,100)
		Poly–1 adjusted number of animals at risk	25.0	24.5

Table 2.4: Reported neoplastic tumors in study of male mice (urethane study)

Organ	Tumor type	Quantity	Control	Treated
lungs with bronchi	alveolar–bronchiolar adenoma	P–value of test of trend or comparison	< 0.0001	< 0.0001
		Number of animals reported with tumor	2	25
		Poly–1 adjusted incidence rate	8.0%	100%
		95% CI for poly–1 adjusted incidence rate (%)	(0.98,26.0)	(86.3,100)
		Poly–1 adjusted number of animals at risk	25.0	25.0
		P–value of test of trend or comparison	0.0003	0.0003
		Number of animals reported with tumor	0	8
		Poly–1 adjusted incidence rate	0.0%	44%
		95% CI for poly–1 adjusted incidence rate (%)	(0,13.7)	(20.3,69.2)
		Poly–1 adjusted number of animals at risk	25.0	18.2

Table 2.4: Reported neoplastic tumors in study of male mice (urethane study) (continued)

Organ	Tumor type	Quantity	Control	Treated
	hemangiosarcoma	P-value of test of trend or comparison	0.0180	0.0180
		Number of animals reported with tumor	0	4
		Poly-1 adjusted incidence rate	0.0%	24%
		95% CI for poly-1 adjusted incidence rate (%)	(0,13.7)	(6.81,52.4)
		Poly-1 adjusted number of animals at risk	25.0	17.0
spleen	hemangiosarcoma	P-value of test of trend or comparison	< 0.0001	< 0.0001
		Number of animals reported with tumor	0	24
		Poly-1 adjusted incidence rate	0.0%	98%
		95% CI for poly-1 adjusted incidence rate (%)	(0,13.7)	(79.6,100)
		Poly-1 adjusted number of animals at risk	25.0	24.4

Chapter 3

Evaluation of the validity of negative studies

3.1 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.

Criteria for the retrospective assessment of dose levels are better established for two year studies, but have not been studied systematically for 26 week studies. Nonetheless, the basic principles should still apply. It is therefore reasonable to assess the dose levels in a 26 week study by reasoning from the following concepts.

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?

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 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.

3.2 Assessment of the validity of the carcinogenicity study

There is clearly no reason to suppose that the dose levels were excessive. Table 3.1 shows the mean weight changes by group over the course of the study. It appears from these results that INCB018424 is associated with diminished weight gain in both male and female mice. The dose levels should therefore be considered appropriate.

Table 3.1: Weight changes by group

Species and Sex	Betamethasone dipropionate						
	Δ_C	Δ_L	$\frac{\Delta_L}{\Delta_C} - 1$	Δ_M	$\frac{\Delta_M}{\Delta_C} - 1$	Δ_H	$\frac{\Delta_H}{\Delta_C} - 1$
Mice — Female	4.66	4.34	+1.7%	4.01	-0.6%	3.98	-10.5%
Mice — Male	6.43	5.79	+1.0%	4.30	-2.0%	5.74	-9.5%

3.3 Assessment of the validity of the urethane study

The urethane study is clearly a positive study, with very strong indications that urethane is strongly associated with increased incidences of lung tumors and hemangiosarcomas. However, since the experimental procedure was different for the urethane studies and the control animals, with different histopathological analyses carried out and different dosing regimes, it is hard to see how the success of this trial adds weight to the study of INCB018424.

Chapter 4

Conclusion

4.1 Carcinogenicity study

4.1.1 Tumor findings

The study is a negative study. Few tumors of any kind were reported in the study, and there were no statistically significant results. While survival levels were very high across all groups, there is evidence of a dose related diminution of weight gain, and so the dose levels can be concluded to have been adequate.

There were no organs reported as unexamined or autolytic.

4.2 Urethane study

The group treated with urethane showed clear indications of increases incidence of lung tumors (bronchio-alveolar adenomas and carcinomas, splenic hemangiosarcomas, and, in the males, hemangiosarcomas of the lungs. In that sense, the study was a success, with clear (intended) positive findings. However, given that these animals did not undergo comparable analyses, the success of this trial is of little relevance when assessing the reliability of the INCB018424 study.

Appendix A

Tumor incidence calculations with all treated groups combined

Table A.1: Reported neoplastic tumors in study of female mice (treated groups combined)

Organ	Tumor type	Quantity	Control	Treated
cavity, nasal	adenocarcinoma	<i>p</i> -value of test of trend or comparison	.8438	.8438
		Number of animals reported with tumor	1	2
		Poly-1 adjusted incidence rate	4.0%	2.7%
		95% CI for Poly-1 adjusted incidence rate (%)	(0.1,20.4)	(0.32,9.4)
		Poly-1 adjusted number of animals at risk	25.0	74.3
harderian glands	adenoma	<i>p</i> -value of test of trend or comparison	.9514	.9514
		Number of animals reported with tumor	2	2
		Poly-1 adjusted incidence rate	8.0%	2.7%
		95% CI for Poly-1 adjusted incidence rate (%)	(0.98,26.0)	(0.32,9.4)
		Poly-1 adjusted number of animals at risk	25.0	74.3
lungs with bronchi	alveolar–bronchiolar adenoma	<i>p</i> -value of test of trend or comparison	.8919	.8919
		Number of animals reported with tumor	3	5
		Poly-1 adjusted incidence rate	12%	6.7%
		95% CI for Poly-1 adjusted incidence rate (%)	(2.55,31.2)	(2.2,15.1)
		Poly-1 adjusted number of animals at risk	25.0	74.3
	alveolar–bronchiolar carcinoma	<i>p</i> -value of test of trend or comparison	.7475	.7475
		Number of animals reported with tumor	0	1
		Poly-1 adjusted incidence rate	0.0%	1.3%
		95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(0.03,7.3)
		Poly-1 adjusted number of animals at risk	25.0	74.3
multicentric	sarcoma	<i>p</i> -value of test of trend or comparison	.7475	.7475
		Number of animals reported with tumor	0	1
		Poly-1 adjusted incidence rate	0.0%	1.3%
		95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(0.03,7.3)
		Poly-1 adjusted number of animals at risk	25.0	74.3
skin	hemangiosarcoma	<i>p</i> -value of test of trend or comparison	.7475	.7475
		Number of animals reported with tumor	0	1
		Poly-1 adjusted incidence rate	0.0%	1.3%
		95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(0.03,7.3)
		Poly-1 adjusted number of animals at risk	25.0	74.3
spleen	hemangiosarcoma	<i>p</i> -value of test of trend or comparison	.3057	.3057
		Number of animals reported with tumor	0	4
		Poly-1 adjusted incidence rate	0.0%	5.4%
		95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(1.47,13.3)
		Poly-1 adjusted number of animals at risk	25.0	74.3

Table A.1: Reported neoplastic tumors in study of female mice (treated groups combined) (continued)

Organ	Tumor type	Quantity	Control	Treated
stomach		Poly-1 adjusted number of animals at risk	25.0	74.5
	papilloma	<i>p</i> -value of test of trend or comparison	1	1
		Number of animals reported with tumor	1	0
		Poly-1 adjusted incidence rate	4.0%	0.0%
thymus		95% CI for Poly-1 adjusted incidence rate (%)	(0.1,20.4)	(0,4.9)
		Poly-1 adjusted number of animals at risk	25.0	74.3
	<i>p</i> -value of test of trend or comparison	.4133	.4133	
	Number of animals reported with tumor	0	3	
	Poly-1 adjusted incidence rate	0.0%	4.0%	
	95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(0.83,11.4)	
		Poly-1 adjusted number of animals at risk	25.0	74.3

Table A.2: Reported neoplastic tumors in study of male mice (treated groups combined)

Organ	Tumor type	Quantity	Control	Treated
cavity, nasal	adenocarcinoma	<i>p</i> -value of test of trend or comparison	.7475	.7475
		Number of animals reported with tumor	0	1
		Poly-1 adjusted incidence rate	0.0%	1.3%
		95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(0.03,7.3)
harderian glands	adenoma	Poly-1 adjusted number of animals at risk	25.0	74.2
		<i>p</i> -value of test of trend or comparison	.9853	.9853
		Number of animals reported with tumor	2	1
		Poly-1 adjusted incidence rate	8.0%	1.3%
liver	hepatocellular adenoma	95% CI for Poly-1 adjusted incidence rate (%)	(0.98,26.0)	(0.03,7.3)
		Poly-1 adjusted number of animals at risk	25.0	74.2
		<i>p</i> -value of test of trend or comparison	.7475	.7475
		Number of animals reported with tumor	0	1
lungs with bronchi	alveolar – bronchiolar adenoma	Poly-1 adjusted incidence rate	0.0%	1.3%
		95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(0.03,7.3)
		Poly-1 adjusted number of animals at risk	25.0	74.2
		<i>p</i> -value of test of trend or comparison	.5934	.5934
		Number of animals reported with tumor	2	7
		Poly-1 adjusted incidence rate	8.0%	9.4%
		95% CI for Poly-1 adjusted incidence rate (%)	(0.98,26.0)	(3.84,18.5)
		Poly-1 adjusted number of animals at risk	25.0	74.2

Table A.2: Reported neoplastic tumors in study of male mice (treated groups combined) (continued)

Organ	Tumor type	Quantity	Control	Treated
skeletal muscle (thigh)	hemangioma	<i>p</i> -value of test of trend or comparison	.7475	.7475
		Number of animals reported with tumor	0	1
		Poly-1 adjusted incidence rate	0.0%	1.3%
		95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(0.03,7.3)
		Poly-1 adjusted number of animals at risk	25.0	74.2
		<i>p</i> -value of test of trend or comparison	.7475	.7475
		Number of animals reported with tumor	0	1
		Poly-1 adjusted incidence rate	0.0%	1.3%
		95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(0.03,7.3)
		Poly-1 adjusted number of animals at risk	25.0	74.5
skin	hemangiosarcoma	<i>p</i> -value of test of trend or comparison	.7475	.7475
		Number of animals reported with tumor	0	1
		Poly-1 adjusted incidence rate	0.0%	1.3%
		95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(0.03,7.3)
		Poly-1 adjusted number of animals at risk	25.0	74.2
		<i>p</i> -value of test of trend or comparison	.7475	.7475
		Number of animals reported with tumor	0	1
		Poly-1 adjusted incidence rate	0.0%	1.3%
		95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(0.03,7.3)
		Poly-1 adjusted number of animals at risk	25.0	74.2
spleen	hemangiosarcoma	<i>p</i> -value of test of trend or comparison	.5568	.5568
		Number of animals reported with tumor	0	2
		Poly-1 adjusted incidence rate	0.0%	2.7%
		95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(0.32,9.4)
		Poly-1 adjusted number of animals at risk	25.0	74.2
		<i>p</i> -value of test of trend or comparison	.7475	.7475
		Number of animals reported with tumor	0	1
		Poly-1 adjusted incidence rate	0.0%	1.3%
		95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(0.03,7.3)
		Poly-1 adjusted number of animals at risk	25.0	74.6
testes	hemangiosarcoma	<i>p</i> -value of test of trend or comparison	.7475	.7475
		Number of animals reported with tumor	0	1
		Poly-1 adjusted incidence rate	0.0%	1.3%
		95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(0.03,7.3)
		Poly-1 adjusted number of animals at risk	25.0	74.2
		<i>p</i> -value of test of trend or comparison	.7475	.7475
		Number of animals reported with tumor	0	1
		Poly-1 adjusted incidence rate	0.0%	1.3%
		95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(0.03,7.3)
		Poly-1 adjusted number of animals at risk	25.0	74.2
thyroid glands	adenoma	<i>p</i> -value of test of trend or comparison	.7475	.7475
		Number of animals reported with tumor	0	1
		Poly-1 adjusted incidence rate	0.0%	1.3%
		95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(0.03,7.3)
		Poly-1 adjusted number of animals at risk	25.0	74.2
		<i>p</i> -value of test of trend or comparison	.7475	.7475
		Number of animals reported with tumor	0	1
		Poly-1 adjusted incidence rate	0.0%	1.3%
		95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(0.03,7.3)
		Poly-1 adjusted number of animals at risk	25.0	74.2

Table A.3: Reported combinations in study of female mice (treated groups combined)

Endpoint	Quantity	Control	Treated
All hemangiosarcomas	<i>p</i> -value of test of trend or comparison	.2252	.2252
	Number of animals reported with tumor	0	5
	Poly-1 adjusted incidence rate	0.0%	6.7%
	95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(2.2,15.1)
All skin tumors	Poly-1 adjusted number of animals at risk	25.0	74.5
	<i>p</i> -value of test of trend or comparison	.7475	.7475
	Number of animals reported with tumor	0	1
	Poly-1 adjusted incidence rate	0.0%	1.3%
	95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(0.03,7.3)
	Poly-1 adjusted number of animals at risk	25.0	74.4

Table A.4: Reported combinations in study of male mice (treated groups combined)

Endpoint	Quantity	Control	Treated
All hemangiosarcomas	<i>p</i> -value of test of trend or comparison	.7362	.7362
	Number of animals reported with tumor	1	3
	Poly-1 adjusted incidence rate	4.0%	4.0%
	95% CI for Poly-1 adjusted incidence rate (%)	(0.1,20.4)	(0.83,11.4)
All skin tumors	Poly-1 adjusted number of animals at risk	25.0	74.2
	<i>p</i> -value of test of trend or comparison	.5568	.5568
	Number of animals reported with tumor	0	2
	Poly-1 adjusted incidence rate	0.0%	2.7%
	95% CI for Poly-1 adjusted incidence rate (%)	(0,13.7)	(0.32,9.4)
	Poly-1 adjusted number of animals at risk	25.0	74.6

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] 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.

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
09/16/2011

KARL K LIN
09/16/2011
Concur with review

STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

NDA Number: 202192

Applicant: Novartis

Stamp Date: 6/3/11

Drug Name: Ruxolitinib

NDA/BLA Type: NDA

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

Information requests:

- For Study 352, provide a one record per subject dataset including primary, secondary endpoints, patient demographics and disposition.

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.				To be checked
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.				To be checked
Investigation of effect of dropouts on statistical analyses as described by applicant appears adequate.	x			

STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

Reviewing Statistician _____ Date _____

Supervisor/Team Leader _____ Date _____

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

/s/

HONG LU
07/22/2011

MARK D ROTHMANN
07/22/2011