

**CENTER FOR DRUG EVALUATION AND
RESEARCH**

APPLICATION NUMBER:
22-304

STATISTICAL REVIEW(S)



STATISTICAL REVIEW AND EVALUATION

Biometrics Division: VI (HFD-705)

NDA No.:	22-304
SERIAL No.:	S_000
DATE RECEIVED BY THE CENTER:	March 19, 2008
DRUG NAME:	Tapentadol Hydrochloride (CG5503)
DOSAGE FORM:	Oral
INDICATION:	Moderate to severe acute pain
SPONSOR:	Ortho-McNeil-Janssen Pharmaceuticals, Inc.
DOCUMENTS REVIEWED:	Electronic Copy Dated March 19, 2008
NAME OF STATISTICAL REVIEWER:	Meiyu Shen, Ph.D. Atiar Rahman, Ph.D.
NAME OF PHARM TOX REVIEWER:	Kathleen Young

Meiyu Shen, Mathematical Statistician
Atiar Rahman, Mathematical Statistician

Concur:

Karl K Lin, Ph.D.
Team Leader, DBVI

Distribution: NDA 22-304
HFD-705/Karl K Lin, Ph.D.
OND/ODEII/DMEP, Kathleen A. Young

TABLE OF CONTENTS

- 1. *Background* 3
- 2. *Mouse study (Study TP2518)* 3
 - 2.1 *Introduction* 3
 - 2.2 *Sponsor's analysis* 4
 - 2.3 *Data Analyzed and Sources* 4
 - 2.4 *Reviewer's analysis (I) - Including All Treatment Groups* 5
 - 2.4.1 *Survival analysis* 5
 - 2.4.2 *Tumor data analysis* 8
 - 2.4.3 *Conclusion* 8
 - 2.5 *Reviewer's Analysis (II) – Excluding Dose-Escalation Group* 11
 - 2.5.1 *Survival analysis* 11
 - 2.5.2 *Tumor data analysis* 15
 - 2.5.3 *Conclusion* 15
- 3. *Rat study (Study TP2418)* 19
 - 3.1 *Introduction* 19
 - 3.2 *Sponsor's analysis* 20
 - 3.3 *Data Analyzed and Sources* 20
 - 3.4 *Reviewers' analyses* 20
 - 3.4.1 *Survival analysis* 20
 - 3.4.2 *Tumor data analyses Using Peto Method* 23
 - 3.4.3 *Tumor Data Analysis Using Poly-3 Method* 27
 - 3.4.4 *Conclusion* 32

Statistical Review of NDA22304

1. BACKGROUND

CG5503 is for treatment of moderate to severe pain. In this submission (NDA 22304), the sponsor submitted one 2-year oral carcinogenicity study of CG5503 in mouse (Study TP2518) and one 2-year oral carcinogenicity study of CG5503 in rats (Study TP2418).

Dr. Meiyu Shen is the original statistical reviewer of the carcinogenicity studies of this NDA submission. She performed the original analysis of the survival and tumor data of both the mouse study and the rat study using data of all six treatment groups in the two studies, and completed a draft statistical review and evaluation report before she took a vacation during the summer.

While Dr. Shen was on vacation and the draft report was waiting for the secondary statistical reviewer Karl Lin, Ph.D. to concur, the Pharm/Tox Statistics Team of the Office of Biostatistics received an urgent request from Adam M. Wasserman, Ph.D., Supervisory Pharmacologist/Toxicologist, Division of Anesthesia, Analgesia and Rheumatology Products, to re-analyze the survival and tumor data of the mouse excluding the dose-escalation group. Dr. Wasserman and Dr. Kathleen Young, the primary pharm/tox reviewer of this submission, also informed the OB Pharm/Tox Statistics Team that they were going to present their reviews at the scheduled ECAC meeting on 8/26/2008, and would like to have statistical review results from the Team before the meeting.

Because urgent requests from the medical division, Dr. Lin asked Dr. Atiar Rahman of the OB Pharm/Tox Statistics Team to perform the re-analysis of the survival and tumor data of the mouse study instead of waiting for Dr. Shen to do that after her vacation. Dr. Rahman re-analyzed the survival and tumor data of the mouse study excluding the dose-escalation group as requested by Dr. Wasserman. Dr. Rahman also re-analyzed the survival and tumor data of the rat study merely to double check the results of the original analysis.

Dr. Atiar Rahman used the poly-3 test in his re-analysis versus the Peto test used in Dr. Shen's original analysis. Both tests are survival-adjusted tests. The dose-escalation group of the mouse study was excluded (for both males and females) in the re-analysis.

Only the results of the re-analysis of the mouse study and the rat study are added to the original analysis in this report. It is noted that the re-analysis also includes pairwise comparisons between the combined control and individual treated groups.

2. MOUSE STUDY (STUDY TP2518)

2.1 Introduction

The objective of this study was to determine the effects of CG5503 on the incidence and morphology of tumors following oral administration to the mouse for 104 weeks. Three treatment groups of 51 male and 51 female mice/group were administered the test article at respective dose levels of 50, 100, 200 mg free base/kg/day. One treatment group of 60 male and 60 female mice/group were administered the test article at respective dose levels of 300 mg free base/kg/day. Two control groups of 51 male and 51 female mice/group were administered the placebo.

Statistical Review of NDA22304

Group	Group description	Dose level (mg/kg/day)	Animal numbers			
			Main study		Satellite study	
			Male	Female	Male	Female
1	Control 1	0	1-51	307-357	613-624	724-735
2	Low	50	52-102	358-408	625-645	736-756
3	Intermediate (I)	100	103-153	409-459	646-666	757-777
4	Intermediate (II)	200*	154-204	460-510	667-690	778-801
5	High	200 (Wks 1-13)	205-255	511-561	691-714	802-825
		300 (Wks 14-28)	835-843#	844-852#		
		200 (Wks 29-91)**				
6	Control 2	0	256-306	562-612	715-723	826-834

dose escalation animals were administered the increased dose levels at Week 14 and 27 in advance of the main study and satellite animals

* as the number of surviving Group 4 males fell to 20 in Week 100, the remaining Group 4 males were retained off-dose for the remainder of the study

** All surviving Group 5 animals were subject to early termination from the study in Week 92, since Group 5 reached a level considered insufficient for the continued viability of this group.

At Week 14, 9 males and 9 females from the high dose group (Group 5, 200 mg/kg/day; designated dose escalation animals) were administered CG5503 at 300 mg/kg/day. After 7 days dosing, animals were considered to have tolerated this increase in dose level and all Group 5 animals were dosed at 300 mg/kg/day from Week 15.

Dosing of Group 5 was reduced to 200 mg/kg/day from Week 29 although it was scheduled to increase to 400 mg/kg/day at the start of Week 27.

2.2 Sponsor's analysis

The sponsor presented the mean survival estimate in Figures 1 and 2 of the sponsor's study report and the summary of survival estimates in Table 4 of the sponsor's study report. However, the sponsor did not present any statistical analysis for the mice mortality data. The sponsor noted that over the full duration of the study, mortality in males dosed at 200 and 200/300/200 mg/kg/day and females dosed at 200/300/200 mg/kg/day was significantly higher than in combined control groups.

Body weight was unaffected by treatment at doses up to 200 mg/kg/day. There was a slight reduction in body weight gain for Group 5 males between weeks 13-28 when administered 300 mg/kg/day.

Food consumption was unaffected by treatment.

There was a statistically significant increase at the 5% level for liver hepatocellular tumors, in males (for overall dose response).

2.3 Data Analyzed and Sources

The sponsor submitted the data in electronic format on March 19, 2008. The data are located in the EDR at the following link: \\cdsesub1\n22304\S_00.

2.4 Reviewer's analysis (I) - Including All Treatment Groups

Dr. Meiyu Shen independently analyzed the survival data for males and females, separately. She also independently analyzed the mice tumor data for males and females, separately using Peto's method.

2.4.1 Survival analysis

The summaries of the mortality data are given in Table 1 for males. The time intervals 0-52, 53-78, 79-90 and 91-105 weeks were chosen for males. The Kaplan-Meier curves for males are shown in Figure 1. Analysis of Dose-Mortality Trend for Male Mice is presented in Table 2. From Figure 1, we can see that the survival probability in control group was much higher than that in the dosed group over the full duration of the study. The highest dose group in males had the highest mortality rate. The analysis of Dose-Mortality trend for males in Table 2 showed a statistically significant dose-related trend among the control and the dosed groups because the p-value is 0.0005 (Cox method) and 0.0002 (Kruskal-Wallis tests), respectively, which is much smaller than 0.05.

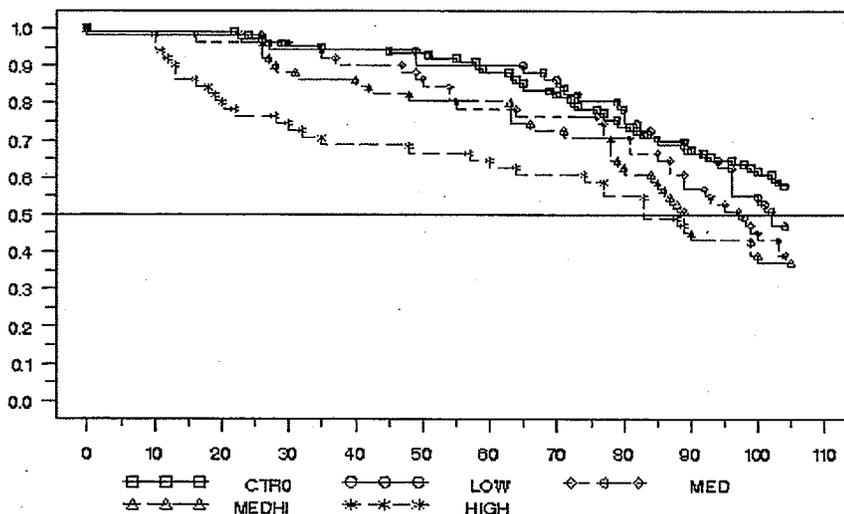


Figure 1 Kaplan-Meier Survival functions for Male Mice (including all treatment groups)

The summaries of the mortality data are given in Table 3 for females. The time intervals 0-52, 53-78, 79-90 and 91-105 weeks were chosen for females. The Kaplan-Meier curves for females are shown in Figure 2. Analysis of Dose-Mortality Trend for Female Mice is presented in Table 4. From Figure 2, we can see that the survival probability in control group is much higher than that in the highest dosed groups. The analysis of Dose-Mortality trend for females in Table 4 shows a statistically significant dose-related trend among the control and the dosed groups because the p-value is 0.0487 (Kruskal-Wallis tests), which is smaller than 0.05.

Statistical Review of NDA22304

Table 1 Analysis of mortality data for male mice (including all treatment groups)

Analysis of Mortality		No. Risk	No. Died	No. Alive	Pct Survival	Pct Mortality
CTR0	0-52	102	7	95	93.1	6.9
	53-78	95	16	79	77.5	22.5
	79-90	79	10	69	67.6	32.4
	FINALKILL 91-105	69	69	0		
LOW	0-52	51	3	48	94.1	5.9
	53-78	48	6	42	82.4	17.6
	79-90	42	7	35	68.6	31.4
	FINALKILL 91-105	35	35	0		
MED	0-52	51	7	44	86.3	13.7
	53-78	44	6	38	74.5	25.5
	79-90	38	7	31	60.8	39.2
	FINALKILL 91-105	31	31	0		
MEDHI	0-52	51	9	42	82.4	17.6
	53-78	42	6	36	70.6	29.4
	79-90	36	13	23	45.1	54.9
	FINALKILL 91-105	23	23	0		
HIGH	0-52	51	16	35	68.6	31.4
	53-78	35	5	30	58.8	41.2
	79-90	30	6	24	47.1	52.9
	FINALKILL 91-105	24	24	0		

Table 2 Analysis of Dose-Mortality Trend for Male Mice (including all treatment groups)

	Method			
	Cox		Kruskal-Wallis	
	Statistics	P-Value	Statistics	P-Value
Time-Adjusted Trend Test	1.6026	0.6588	2.0521	0.5617
Depart from Trend				
Dose-Mortality Trend	11.9583	0.0005	13.8353	0.0002
Homogeneity	13.5609	0.0088	15.8874	0.0032

Note: This test is run using Trend and Homogeneity Analyses of Proportions and Life Table Data Version 2.1, by Donald G. Thomas, National Cancer Institute

From Tables 1 and 3, it is seen that the females' mortality rate in the control group was 20% higher than the males' mortality rate in the control group because the females' mortality rates in the control group by Week 90 were 53.9% and the males' mortality rate by Week 90 were 32.4% for the control. The females' mortality rate in the control is almost higher than the males' mortality rate in any dosed group by Week 90.

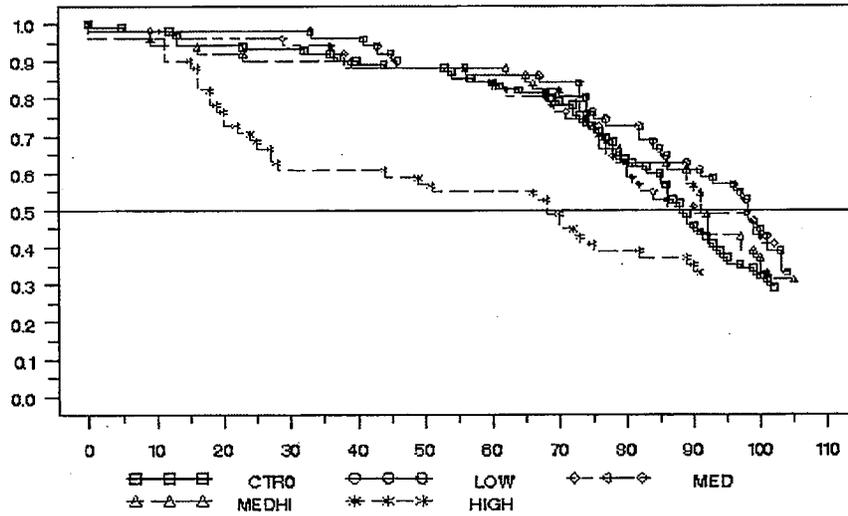


Figure 2 Kaplan-Meier Survival functions for Female Mice (including all treatment groups)

Table 3 Analysis of mortality data for female mice (including all treatment groups)

Analysis of Mortality		No. Risk	No. Died	No. Alive	Pct Survival	Pct Mortality
CTR0	0-52	102	11	91	89.2	10.8
	53-78	91	21	70	68.6	31.4
	79-90	70	23	47	46.1	53.9
	FINALKILL 91-105	47	47	0		
LOW	0-52	51	5	46	90.2	9.8
	53-78	46	8	38	74.5	25.5
	79-90	38	6	32	62.7	37.3
	FINALKILL 91-105	32	32	0		
MED	0-52	51	4	47	92.2	7.8
	53-78	47	14	33	64.7	35.3
	79-90	33	7	26	51.0	49.0
	FINALKILL 91-105	26	26	0		
MEDHI	0-52	51	5	46	90.2	9.8
	53-78	46	10	36	70.6	29.4
	79-90	36	7	29	56.9	43.1
	FINALKILL 91-105	29	29	0		
HIGH	0-52	51	22	29	56.9	43.1
	53-78	29	8	21	41.2	58.8
	79-90	21	3	18	35.3	64.7
	FINALKILL 91-105	18	18	0		

Statistical Review of NDA22304

Table 4 Analysis of Dose-Mortality Trend for Female Mice (including all treatment groups)

	Method			
	Cox		Kruskal-Wallis	
	Statistics	P-Value	Statistics	P-Value
Time-Adjusted Trend Test	12.9435	0.0048	18.1966	0.0004
Depart from Trend				
Dose-Mortality Trend	1.3698	0.2419	3.8862	0.0487
Homogeneity	14.3133	0.0064	22.0829	0.0002

2.4.2 Tumor data analysis

The dose response analyses in incidental tumors and fatal tumors were performed using the Peto prevalence method and the Peto death-rate method, respectively. The actual dose levels of treatment groups were used as the weights for the trend analysis. The number of tumor bearing animals of each tumor type and its p-values for many organs were presented in Tables 5 and 6 for males and females, respectively. Multiplicity for the trend testing was adjusted using a significance level of 0.025 for rare tumors, and 0.005 for common tumors because two species were studied. A tumor type with a background rate of 1 percent or less is classified as rare by Haseman; more frequent tumors are classified as common.

It is also well known that the approximation results may not be stable and reliable, and tend to underestimate the exact p-values when the total numbers of tumor occurrence across treatment groups are small. In this situation, the exact permutation trend test should be used to test for the positive trend. The exact permutation trend test is a generalization of the Fisher's exact test.

From Table 5, it is seen that the p-value (asymptotic method) of the trend test for the rare tumor SK (Sarcoma-NOS) in skin subcutis for males is 0.0214 (<0.025). From Table 6, it is seen that the p-value of the trend test for the rare tumor OV in ovary (benign granulose cell tumor) for females is 0.0203 (<0.025). The p-value of the trend test for the rare tumor OV in ovary (benign luteoma tumor) for females is 0.0154 (<0.025). The p-value of the trend test for the rare tumor SM in sternum+marrow (haenangioma tumor) for females is 0.0074 (<0.025).

2.4.3 Conclusion

The analysis of Dose-Mortality trend for males in Table 2 showed a statistically significant dose-related trend among the control and the dosed groups because the p-value is 0.0005 (Cox method) and 0.0002 (Kruskal-Wallis tests), respectively, which is much smaller than 0.05.

The analysis of Dose-Mortality trend for females in Table 4 shows a statistically significant dose-related trend among the control and the dosed groups because the p-value is 0.0487 (Kruskal-Wallis tests), which is smaller than 0.05.

The mortality rate in the 200/300/200 mg/kg/day group is much higher than the control for both females and males although the 300 mg/kg/day was dosed for 15 weeks (from Week 14 to 28) and switched back to 200 mg/kg/day for the rest 64 weeks (from Week 29 to 92).

From Tables 1 and 3, it is seen that the females' mortality rate in the control group was 20% higher than the males' mortality rate in the control group because the females' mortality rates in the control

Statistical Review of NDA22304

group by Week 90 were 53.9% and the males' mortality rate by Week 90 were 32.4% for the control.

From Table 5, it is seen that the p-value (asymptotic method) of the trend test for the rare tumor SK (Sarcoma-NOS) in skin subcutis for males is 0.0214 (<0.025). From Table 6, it is seen that the p-value of the trend test for the rare tumor OV in ovary (benign granulose cell tumor) for females is 0.0203 (<0.025). The p-value of the trend test for the rare tumor OV in ovary (benign luteoma tumor) for females is 0.0154 (<0.025). The p-value of the trend test for the rare tumor SM in sternum+marrow (haemangioma tumor) for females is 0.0074 (<0.025).

Table 5 Report on Test for Positive Linear Dose-Tumor Trends in Male Mice (including all treatment groups)

Organ Code	Organ Name	Tumor Code	Tumor Name	CTR0	LOW	MED	MEDHI	HIGH	P-Value (Exact Method)	P-Value (Asymptotic Method)
AD	ADRENAL	409	SUBCAPSULAR CELL ADENOMA	5	2	2	1	0	0.6919	0.6661
AD	ADRENAL	453	BENIGN PHAEOCHROMOCYTOMA	0	0	1	0	0	0.3911	0.3358
BR	BRAIN	434	MALIGNANT MENINGIOMA	1	0	0	0	0	1.0000	0.8250
CT	CONNECTIVE TISSUE	515	HISTIOCYTIC SARCOMA	0	1	0	0	0	0.5592	0.5591
DU	DUODENUM	184	OSTEOSARCOMA	0	0	1	0	0	0.3695	0.3199
FO	FOOT/LEG	326	SQUAMOUS CELL PAPILLOMA	1	0	0	0	0	1.0000	0.9229
HE	HAEMOLYMPHORETICULAR	152	GRANULOCYTIC LEUKAEMIA	2	0	0	0	0	1.0000	0.9128
HE	HAEMOLYMPHORETICULAR	177	MALIGNANT LYMPHOMA - PLEOMORPH	8	7	4	0	1	0.8428	0.8247
HE	HAEMOLYMPHORETICULAR	370	MALIGNANT LYMPHOMA - LYMPHOCYT	0	1	0	1	0	0.1688	0.1179
HE	HAEMOLYMPHORETICULAR	53	MALIGNANT LYMPHOMA - LYMPHOBLAST	0	0	2	0	0	0.4600	0.3367
KI	KIDNEY	561	TUBULAR CELL ADENOMA	0	0	1	0	0	0.3418	0.2826
LI	LIVER	302	HAEMANGIOSARCOMA	1	1	0	1	0	0.4069	0.3603
LI	LIVER	322	HISTIOCYTIC SARCOMA	1	0	0	1	0	0.3738	0.2614
LI	LIVER	383	HAEMANGIOMA	4	0	2	1	0	0.5721	0.5426
LI	LIVER	394	HEPATOCELLULAR ADENOMA	9	4	7	5	2	0.1419	0.1256
LI	LIVER	399	HEPATOCELLULAR CARCINOMA	2	0	0	4	0	0.0304	0.0167
LU	LUNG	198	BRONCHIOLO-ALVEOLAR ADENOMA	20	11	7	12	1	0.2762	0.2601
LU	LUNG	211	BRONCHIOLO-ALVEOLAR CARCINOMA	7	2	2	2	1	0.5342	0.5125
MU	MUSCLE	606	RHABDOMYOSARCOMA	0	0	1	0	0	0.3282	0.2786
PI	PITUITARY	318	ADENOMA	1	0	0	0	0	1.0000	0.8096
PR	PROSTATE	641	ADENOCARCINOMA	1	0	0	0	0	1.0000	0.8089
SI	SPINAL COLUMN	636	OSTEOSARCOMA	0	1	0	0	0	1.0000	0.8430
SK	SKIN + SUBCUTIS	105	SARCOMA - NOS	1	2	1	3	1	0.0341	0.0214 (!)
SK	SKIN + SUBCUTIS	153	OSTEOSARCOMA	1	0	0	0	0	1.0000	0.8365
SK	SKIN + SUBCUTIS	563	SQUAMOUS CELL PAPILLOMA	0	0	1	0	0	0.3418	0.2826
SK	SKIN + SUBCUTIS	640	MALIGNANT FIBROUS HISTIOCYTOMA	1	0	0	0	0	1.0000	0.8089
SM	STERNUM + MARROW	491	HAEMANGIOMA	2	0	0	0	0	1.0000	0.9449
SP	SPLEEN	554	HAEMANGIOMA	0	1	0	0	0	0.5605	0.5578
ST	STOMACH	246	ADENOMA	0	1	0	0	0	0.5605	0.5548
ST	STOMACH	564	SQUAMOUS CELL	0	0	1	0	0	0.3376	0.2762

Statistical Review of NDA22304

			PAPILLOMA							
TA	TAIL	483	SARCOMA - NOS	0	1	0	0	0	0.5641	0.5827
TA	TAIL	574	HAEMANGIOSARCOMA	0	0	0	1	0	0.2424	0.0515
TE	TESTIS	550	INTERSTITIAL CELL ADENOMA	1	1	0	2	0	0.0969	0.0620
TE	TESTIS	573	RETE TESTIS ADENOMA	0	0	1	0	0	0.3418	0.2826
TO	TONGUE	540	SQUAMOUS CELL PAPILLOMA	1	0	0	0	0	1.0000	0.8089
TY	THYROID	448	FOLLICULAR CELL ADENOMA	2	0	0	0	0	1.0000	0.9447
UB	URINARY BLADDER	642	TRANSITIONAL CELL PAPILLOMA	0	0	0	1	0	0.1484	0.0219

Note: The check mark (Ⓢ) indicates statistically significant test results, based on the decision rule of FDA.CDER.Divisions of Biometrics.

Table 6 Report on Test for Positive Linear Dose-Tumor Trends in Female Mice (including all treatment groups)

Organ Code	Organ Name	Tumor Code	Tumor Name	CTR0	LOW	MED	MEDHI	HIGH	P-Value (Exact Method)	P-Value (Asymptotic Method)
AD	ADRENAL	409	SUBCAPSULAR CELL ADENOMA	2	0	1	0	0	0.8105	0.7872
FE	FEMUR + MARROW	366	OSTEOMA	1	0	0	0	0	1.0000	0.8605
FO	FOOT/LEG	474	INFLAMM MYOFIBROBLAST. TUM	0	1	0	0	0	0.6000	0.5568
FO	FOOT/LEG	507	FIBROMA	0	0	0	1	0	0.2500	0.0473
HE	HAEMOLYMPHORETICULAR	152	GRANULOCYTIC LEUKAEMIA	1	0	1	0	0	0.6932	0.6727
HE	HAEMOLYMPHORETICULAR	177	MALIGNANT LYMPHOMA - PLEOMORPH	18	11	7	7	0	0.8984	0.8914
HE	HAEMOLYMPHORETICULAR	353	IMMUNOBLASTIC LYMPHOMA	0	0	0	0	1	0.0457	0.0412
HE	HAEMOLYMPHORETICULAR	370	MALIGNANT LYMPHOMA - LYMPHOCYT	1	2	2	0	0	0.6911	0.6592
HE	HAEMOLYMPHORETICULAR	405	HISTIOCYTIC SARCOMA	2	0	0	1	0	0.6201	0.5674
HE	HAEMOLYMPHORETICULAR	53	MALIGNANT LYMPHOMA-LYMPHOBLAST	1	0	0	0	0	1.0000	0.8375
LI	LIVER	322	HISTIOCYTIC SARCOMA	2	0	1	0	0	0.8741	0.8320
LI	LIVER	383	HAEMANGIOMA	1	0	0	0	0	1.0000	0.8336
LI	LIVER	394	HEPATOCELLULAR ADENOMA	0	1	0	0	0	0.6519	0.6352
LU	LUNG	198	BRONCHIOLO-ALVEOLAR ADENOMA	10	7	6	7	2	0.3453	0.3409
LU	LUNG	211	BRONCHIOLO-ALVEOLAR CARCINOMA	4	0	4	1	1	0.3769	0.3955
MA	MAMMARY GLAND	154	ADENOCARCINOMA	4	3	1	1	0	0.8524	0.8295
OV	OVARY	379	BENIGN GRANULOSA CELL TUMOUR	0	0	0	0	1	0.0203 (Ⓢ)	0.0323
OV	OVARY	457	CYSTADENOMA	4	2	0	1	0	0.8723	0.8452
OV	OVARY	510	BENIGN LUTEOMA	0	0	1	1	1	0.0154 (Ⓢ)	0.0122
OV	OVARY	513	BENIGN THECOMA	1	0	0	0	0	1.0000	0.7946
PA	PANCREAS	529	ISLET CELL ADENOMA	1	0	1	1	0	0.3373	0.2788
PI	PITUITARY	318	ADENOMA	5	4	2	2	0	0.7511	0.7335
SK	SKIN + SUBCUTIS	439	BENIGN HISTIOCYTOMA	0	0	1	0	0	0.3696	0.3411

Statistical Review of NDA22304

SK	SKIN + SUBCUTIS	594	SQUAMOUS CELL CARCINOMA	1	0	0	0	0	1.0000	0.8429
SK	SKIN + SUBCUTIS	619	BENIGN HAIR FOLLICLE TUMOUR	0	1	0	0	0	0.6519	0.6352
SM	STERNUM + MARROW	491	HAEMANGIOMA	0	0	0	0	1	0.0074 (ⓘ)	0.0324
SP	SPLEEN	554	HAEMANGIOMA	0	1	1	0	0	0.5171	0.4991
ST	STOMACH	246	ADENOMA	1	1	0	0	0	0.8772	0.8459
ST	STOMACH	254	ADENOCARCINOMA	0	0	1	0	0	0.4364	0.3902
ST	STOMACH	564	SQUAMOUS CELL PAPILLOMA	0	1	0	0	0	0.6493	0.6358
TA	TAIL	483	SARCOMA - NOS	0	0	0	1	0	0.1786	0.0352
UT	UTERUS	244	STROMAL POLYP	6	4	3	1	1	0.8554	0.8521
UT	UTERUS	290	HISTIOCYTIC SARCOMA	6	0	3	2	0	0.7274	0.7064
UT	UTERUS	317	LEIOMYOMA	4	0	3	1	0	0.7013	0.6797
UT	UTERUS	355	HAEMANGIOMA	2	0	1	2	0	0.2956	0.2582
UT	UTERUS	465	STROMAL SARCOMA	0	0	1	1	0	0.1328	0.0778
UT	UTERUS	476	LEIOMYOSARCOMA	1	1	0	0	0	0.8710	0.8235
UT	UTERUS	508	ADENOCARCINOMA	1	0	0	1	0	0.4323	0.3258
UT	UTERUS	575	HISTIOCYTOMA	1	0	0	0	0	1.0000	0.8429
UT	UTERUS	579	HAEMANGIOSARCOMA	0	1	0	0	0	0.6727	0.6374
VA	VAGINA	294	HAEMANGIOSARCOMA	0	0	0	1	0	0.2434	0.0634
VA	VAGINA	369	STROMAL POLYP	1	2	0	0	0	0.8600	0.8229
VA	VAGINA	505	STROMAL SARCOMA	0	0	1	0	0	0.4143	0.3749

Note: The check mark ⓘ indicates statistically significant test results, based on the decision rule of FDA.CDER.Divisions of Biometrics.

2.5 Reviewer's Analysis (II) – Excluding Dose-Escalation Group

Dr. Atiar Rahman independently analyzed the survival data for males and females, separately. He also independently analyzed the mice tumor data for males and females, separately using the poly-3 method excluding the dose-escalation group.

2.5.1 Survival analysis

The summaries of the mortality data are given in Table 1R for males. The time intervals 0-52, 53-78, 79-91, 92-104, and 105-105 weeks were chosen for males. The Kaplan-Meier curves for males are shown in Figure 1R. Analysis of Dose-Mortality Trend for Male Mice is presented in Table 2R. The analysis of Dose-Mortality trend for males in Table 2R shows that the dose-related trend is statistically significant.

APPEARS THIS WAY
ON ORIGINAL

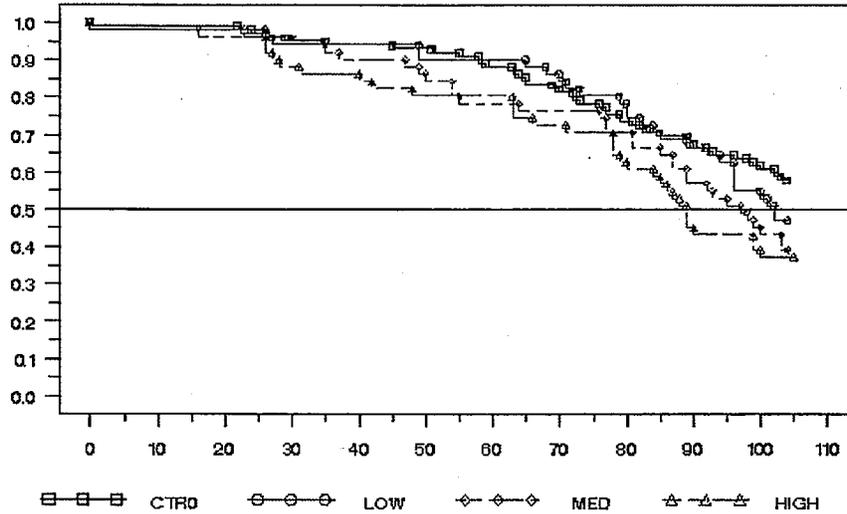


Figure 1R Kaplan-Meier Survival functions for Male Mice (Excluding Dose-Escalation Group)

Table 1R Analysis of mortality data for male mice (excluding dose-escalation group)

Analysis of Mortality Data for Male Mice by Treatment and Time

	Analysis of Mortality	No. Risk	No. Died	No. Alive	Pct Survival	Pct Mortality
CTR0	0-52	102	7	95	93.1	6.9
	53-78	95	16	79	77.5	22.5
	79-91	79	10	69	67.6	32.4
	92-104	69	10	59	57.8	42.2
	FINALKILL105-105	59	59	0		
LOW	0-52	51	3	48	94.1	5.9
	53-78	48	6	42	82.4	17.6
	79-91	42	7	35	68.6	31.4
	92-104	35	11	24	47.1	52.9
	FINALKILL105-105	24	24	0		
MED	0-52	51	7	44	86.3	13.7
	53-78	44	6	38	74.5	25.5
	79-91	38	7	31	60.8	39.2
	92-104	31	11	20	39.2	60.8
FINALKILL105-105	20	20	0			
HIGH	0-52	51	9	42	82.4	17.6

Statistical Review of NDA22304

Analysis of Mortality	No. Risk	No. Died	No. Alive	Pct Survival	Pct Mortality
53-78	42	6	36	70.6	29.4
79-91	36	13	23	45.1	54.9
92-104	23	3	20	39.2	60.8
FINALKILL105-105	20	20	0		

Table 2R Analysis of Dose-Mortality Trend for Male Mice (Excluding Dose-Escalation Group)

Analysis of Dose-Mortality Trend for Male Mice

	Method			
	Cox		Kruskal-Wallis	
	Statistics	P-Value	Statistics	P-Value
Time-Adjusted Trend Test				
 Depart from Trend	0.5467	0.7608	0.3088	0.8569
 Dose-Mortality Trend	6.6580	0.0099	6.7061	0.0096
 Homogeneity	7.2046	0.0657	7.0149	0.0714

Note: This test is run using Trend and Homogeneity Analyses of Proportions and Life Table Data Version 2.1, by Donald G. Thomas, National Cancer Institute

The summaries of the mortality data are given in Table 3R for females. The time intervals 0-52, 53-78, 79-91, 92-104, and 105-105 weeks were chosen for females. The Kaplan-Meier curves for females are shown in Figure 2R. Analysis of Dose-Mortality Trend for Female Mice is presented in Table 4R. The analysis of Dose-Mortality trend for females in Table 4R shows that the dose-related trend in survival is not statistically significant.

**APPEARS THIS WAY
ON ORIGINAL**

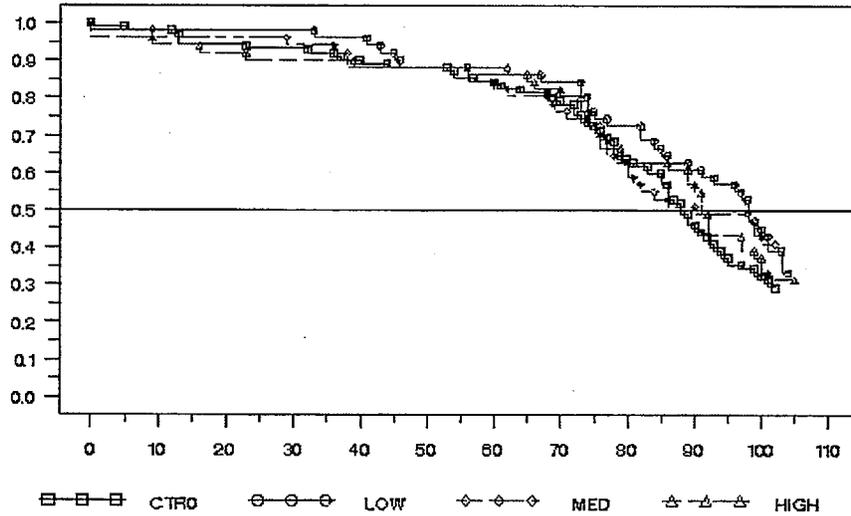


Figure 2R Kaplan-Meier Survival functions for Female Mice (Excluding Dose-Escalation Group)

Table 3R Analysis of mortality data for female mice (excluding dose-escalation group)

Analysis of Mortality Data for Female Mice by Treatment and Time						
	Analysis of Mortality	No. Risk	No. Died	No. Alive	Pct Survival	Pct Mortality
CTR0	0-52	102	11	91	89.2	10.8
	53-78	91	21	70	68.6	31.4
	79-91	70	25	45	44.1	55.9
	92-104	45	15	30	29.4	70.6
	FINALKILL105-105	30	30	0		
LOW	0-52	51	5	46	90.2	9.8
	53-78	46	8	38	74.5	25.5
	79-91	38	7	31	60.8	39.2
	92-104	31	14	17	33.3	66.7
	FINALKILL105-105	17	17	0		
MED	0-52	51	4	47	92.2	7.8
	53-78	47	14	33	64.7	35.3
	79-91	33	7	26	51.0	49.0
	92-104	26	5	21	41.2	58.8
	FINALKILL105-105	21	21	0		
HIGH	0-52	51	5	46	90.2	9.8
	53-78	46	10	36	70.6	29.4
	79-91	36	8	28	54.9	45.1

Statistical Review of NDA22304

Analysis of Mortality	No. Risk	No. Died	No. Alive	Pct Survival	Pct Mortality
92-104	28	11	17	33.3	66.7
FINALKILL105-105	17	17	0		

Table 4R Analysis of Dose-Mortality Trend for Female Mice (Excluding Dose-Escalation Group)

	Method			
	Cox		Kruskal-Wallis	
	Statistics	P-Value	Statistics	P-Value
Time-Adjusted Trend Test	1.4793	0.4773	1.7024	0.4269
Depart from Trend				
Dose-Mortality Trend	0.5821	0.4455	0.4733	0.4915
Homogeneity	2.0614	0.5598	2.1757	0.5367

2.5.2 Tumor data analysis

The dose response analyses and pairwise comparisons between the combined control group and individual treated groups were performed using the poly-3 method. The actual dose levels of treatment groups were used as the weights for the trend analysis. The number of tumor bearing animals of each tumor type and its p-values for many organs were presented in Tables 5R and 6R for males and females, respectively. Multiplicity for the trend testing was adjusted using a significance level of 0.025 for rare tumors, and 0.005 for common tumors because two species were studied. A tumor type with a background rate of 1 percent or less is classified as rare by Haseman; more frequent tumors are classified as common.

Results from Tables 5R and 6R show that there is no statistically significant positive trend or pairwise increase in incidence in all tumors tested in males and females of mice and rats.

2.5.3 Conclusion

When all treatment groups were included, the Dose-Mortality trends in both male and female mice were statistically significant.

When the dose-escalation group was excluded from the survival data analysis, the dose-mortality trend in male mice was statistically significant, but the dose-mortality trend in female mice was not statistically significant.

Results of analysis of tumor data of the mouse study show that there is no statistically significant positive trend and pairwise increase between the combined control and individual treated groups in incidence in all tumor types tested in male and females mice.

Statistical Review of NDA22304

Table 5R Report on Test for Positive Linear Dose-Tumor Trends in Male Mice (Excluding Dose-Escalation Group)

NDA 22_304
Dose Response Relationship Test and Pairwise Comparisons
Using Poly-3 test
Male Mice

Organ Name	Tumor Name	0 mg	50 mg	100 mg	200 mg	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
		Cont N=102	Low N=51	Med N=51	High N=51				
ADRENAL	BENIGN PHAEOCHROMOCYTOMA	0	0	1	0	0.366	.	0.313	.
	SUBCAPSULAR CELL ADENOMA	5	2	2	1	0.695	0.436	0.380	0.571
BONE	OSTEOMA	0	0	1	0	0.366	.	0.313	.
BRAIN	MALIGNANT MENINGIOMA	1	0	0	0	0.576	0.333	0.310	0.291
CONNECTIVE TISS	HISTIOCYTIC SARCOMA	0	1	0	0	0.366	0.336	.	.
DUODENUM	OSTEOSARCOMA	0	0	1	0	0.370	.	0.319	.
FOOT/LEG	SQUAMOUS CELL PAPILLOMA	1	0	0	0	0.576	0.333	0.310	0.291
HAEMOLYPHORETI	GRANULOCYTIC LEUKAEMIA	2	0	0	0	0.819	0.554	0.522	0.495
	MALIGNANT LYMPHOMA - LYMP	0	1	0	1	0.172	0.336	.	0.294
	MALIGNANT LYMPHOMA - PLEO	8	7	4	0	0.937	0.195	0.553	0.941
	MALIGNANT LYMPHOMA-LYMPHO	0	0	2	0	0.356	.	0.103	.
HARDERIAN GLAND	ADENOMA	0	2	0	0	0.438	0.111	.	.
KIDNEY	TUBULAR CELL ADENOMA	0	0	1	0	0.366	.	0.313	.
LIVER	HAEMANGIOMA	4	0	2	1	0.537	0.811	0.612	0.460
	HAEMANGIOSARCOMA	1	1	0	1	0.377	0.561	0.313	0.503
	HEPATOCELLULAR ADENOMA	9	4	7	5	0.215	0.448	0.182	0.382
	HEPATOCELLULAR CARCINOMA	2	0	0	4	0.021	0.558	0.525	0.063
	HISTIOCYTIC SARCOMA	1	0	0	1	0.354	0.333	0.310	0.499
LUNG	BRONCHIOLO-ALVEOLAR ADENO	20	11	7	12	0.232	0.466	0.657	0.212
	BRONCHIOLO-ALVEOLAR CARCI	7	2	2	2	0.659	0.630	0.569	0.517
MUSCLE	RHABDOMYOSARCOMA	0	0	1	0	0.366	.	0.313	.

**APPEARS THIS WAY
ON ORIGINAL**

Statistical Review of NDA22304

NDA 22_304
Dose Response Relationship Test and Pairwise Comparisons
Using Poly-3 test (Excluding Dose-Escalation Group)

Male Mice

Organ Name	Tumor Name	0 mg	50 mg	100 mg	200 mg	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
		Cont N=102	Low N=51	Med N=51	High N=51				
PITUITARY	ADENOMA	1	0	0	0	0.579	0.336	0.313	0.294
PROSTATE	ADENOCARCINOMA	1	0	0	0	0.579	0.336	0.313	0.294
SKIN + SUBCUTIS	MALIGNANT FIBROUS HISTIOC	1	0	0	0	0.579	0.336	0.313	0.294
	OSTEOSARCOMA	1	0	0	0	0.576	0.333	0.310	0.291
	SARCOMA - NOS	1	2	1	3	0.054	0.265	0.525	0.078
	SQUAMOUS CELL PAPILLOMA	0	0	1	0	0.366	.	0.313	.
SPINAL COLUMN	OSTEOSARCOMA	0	1	0	0	0.364	0.342	.	.
SPLEEN	HAEMANGIOMA	0	1	0	0	0.366	0.336	.	.
STERNUM + MARRO	HAEMANGIOMA	2	0	0	0	0.824	0.561	0.529	0.503
STOMACH	ADENOMA	0	1	0	0	0.366	0.336	.	.
	SQUAMOUS CELL PAPILLOMA	0	0	1	0	0.366	.	0.313	.
TAIL	HAEMANGIOSARCOMA	0	0	0	1	0.175	.	.	0.294
	SARCOMA - NOS	0	1	0	0	0.366	0.336	.	.
TESTIS	INTERSTITIAL CELL ADENOMA	1	1	0	2	0.131	0.561	0.313	0.206
	RETE TESTIS ADENOMA	0	0	1	0	0.366	.	0.313	.
THYROID	FOLLICULAR CELL ADENOMA	2	0	0	0	0.822	0.558	0.525	0.499
TONGUE	SQUAMOUS CELL PAPILLOMA	1	0	0	0	0.579	0.336	0.313	0.294
URINARY BLADDER	TRANSITIONAL CELL PAPILLO	0	0	0	1	0.175	.	.	0.294

APPEARS THIS WAY
ON ORIGINAL

Statistical Review of NDA22304

NDA 22_304
Dose Response Relationship Test and Pairwise Comparisons
Using poly-3 test (Excluding Dose-Escalation Group)

Male Mice

(Tumor types with Some P_values <=0.05)

Organ Name	Tumor Name	0 mg	50 mg	100 mg	200 mg	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
		Cont N=102	Low N=51	Med N=51	High N=51				
LIVER	HEPATOCELLULAR CARCINOMA	2	0	0	4	0.021	0.558	0.525	0.063

Table 6 Report on Test for Positive Linear Dose-Tumor Trends in Female Mice (Excluding Dose-Escalation Group)

NDA 22_304
Dose Response Relationship Test and Pairwise Comparisons
Using Poly-3 test
Female Mice

Organ Name	Tumor Name	0 mg	50 mg	100 mg	200 mg	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
		Cont N=102	Low N=51	Med N=51	High N=51				
ADRENAL	SUBCAPSULAR CELL ADENOMA	2	0	1	0	0.709	0.597	0.271	0.572
BONE	OSTEOMA	2	0	0	0	0.859	0.602	0.577	0.577
FEMUR + MARROW	OSTEOMA	1	0	0	0	0.622	0.367	0.347	0.347
FOOT/LEG	FIBROMA	0	0	0	1	0.206	.	.	0.354
	INFLAMM MYOFIBROBLAST. TU	0	1	0	0	0.402	0.367	.	.
HAEMOLYMPHORETI	GRANULOCYTIC LEUKAEMIA	1	0	1	0	0.493	0.367	0.585	0.347
	HISTIOCYTIC SARCOMA	2	0	0	1	0.436	0.597	0.572	0.271
	MALIGNANT LYMPHOMA - LYMP	1	2	2	0	0.552	0.313	0.285	0.347
	MALIGNANT LYMPHOMA - PLEO	18	11	7	7	0.816	0.477	0.681	0.708
	MALIGNANT LYMPHOMA-LYMPHO	1	0	0	0	0.618	0.364	0.344	0.344
HARDERIAN GLAND	ADENOCARCINOMA	0	0	0	1	0.206	.	.	0.354
LIVER	HAEMANGIOMA	1	0	0	0	0.622	0.367	0.347	0.347
	HEPATOCELLULAR ADENOMA	0	1	0	0	0.402	0.367	.	.
	HISTIOCYTIC SARCOMA	2	0	1	0	0.713	0.602	0.276	0.577
LUNG	BRONCHIOLO-ALVEOLAR ADENO	10	7	6	7	0.293	0.421	0.491	0.349
	BRONCHIOLO-ALVEOLAR CARCI	4	0	4	1	0.524	0.842	0.288	0.577
MAMMARY GLAND	ADENOCARCINOMA	4	3	1	1	0.774	0.501	0.563	0.577
ORAL CAVITY	OSTEOSARCOMA	0	0	1	0	0.406	.	0.354	.
OVARY	BENIGN LUTEOMA	0	0	1	1	0.121	.	0.347	0.347
	BENIGN THECOMA	1	0	0	0	0.622	0.367	0.347	0.347
	CYSTADENOMA	4	2	0	1	0.805	0.379	0.821	0.563

NDA 22_304
Dose Response Relationship Test and Pairwise Comparisons
Using Poly-3 test (Excluding Dose-Escalation Group)

Female Mice

Organ Name	Tumor Name	0 mg	50 mg	100 mg	200 mg	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
		Cont N=102	Low N=51	Med N=51	High N=51				
PANCREAS	ISLET CELL ADENOMA	1	0	1	1	0.312	0.367	0.577	0.577
PITUITARY	ADENOMA	5	4	2	2	0.653	0.424	0.470	0.470
SKIN + SUBCUTIS	BENIGN HAIR FOLLICLE TUMO	0	1	0	0	0.402	0.367	.	.
	BENIGN HISTIOCYTOMA	0	0	1	0	0.406	.	0.354	.
	SQUAMOUS CELL CARCINOMA	1	0	0	0	0.622	0.367	0.347	0.347
SPLEEN	HAEMANGIOMA	0	1	1	0	0.491	0.367	0.347	.
STOMACH	ADENOCARCINOMA	0	0	1	0	0.406	.	0.354	.
	ADENOMA	1	1	0	0	0.691	0.610	0.347	0.347
	SQUAMOUS CELL PAPILLOMA	0	1	0	0	0.402	0.367	.	.
TAIL	SARCOMA - NOS	0	0	0	1	0.201	.	.	0.347
UTERUS	ADENOCARCINOMA	1	0	0	1	0.410	0.367	0.347	0.585
	HAEMANGIOMA	2	0	1	2	0.246	0.602	0.285	0.445
	HAEMANGIOSARCOMA	0	1	0	0	0.402	0.367	.	.
	HISTIOCYTIC SARCOMA	6	0	3	2	0.594	0.939	0.390	0.579
	HISTIOCYTOMA	1	0	0	0	0.622	0.367	0.347	0.347
	LEIOMYOMA	4	0	3	1	0.567	0.842	0.470	0.563
	LEIOMYOSARCOMA	1	1	0	0	0.692	0.602	0.347	0.347
	STROMAL POLYP	6	4	3	1	0.834	0.534	0.363	0.759
STROMAL SARCOMA	0	0	1	1	0.124	.	0.347	0.354	
VAGINA	HAEMANGIOSARCOMA	0	0	0	1	0.206	.	.	0.354
	STROMAL POLYP	1	2	0	0	0.712	0.313	0.347	0.347
	STROMAL SARCOMA	0	0	1	0	0.402	.	0.347	.

3. RAT STUDY (STUDY TP2418)

3.1 Introduction

The objective of this study was to evaluate the carcinogenic potential of CG5503 following daily oral administration in SPF Wistar rats at dose levels of 10, 50, 125 or 250 mg/kg body weight/day for at least 104 consecutive weeks. Two control groups received identical feed devoid of test item. The groups comprised 50 animals per sex for oncogenicity evaluation and were sacrificed after 104 weeks of treatment (allocation A). Additional 10 rats per sex and group were used for the determination of plasma drug concentrations at weeks 4, 13 and 26 (allocation B). After the last blood sampling in week 26, these animals were killed and discarded without further examinations. Thus a total of 360 males and 360 females were assigned to this study.

Statistical Review of NDA22304

Group	Nominal dose* (mg/kg)	Achieved dose* (mg/kg)		Assigned number of animals per dose		
		M	F	Total number of animals	Number of allocation A animals Oncogenicity evaluation	Number of allocation B animals Plasma level determination
1	0	0	0	120	50 M + 50 F	10 M + 10 F
2	0	0	0	120	50 M + 50 F	10 M + 10 F
3	10	9.99	9.98	120	50 M + 50 F	10 M + 10 F
4	50	49.85	50.20	120	50 M + 50 F	10 M + 10 F
5	125	124.32	124.83	120	50 M + 50 F	10 M + 10 F
6	250	249.06	251.55	120	50 M + 50 F	10 M + 10 F

* = The dose levels refer to the hydrochloride salt of CG5503, M = male; F = female

3.2 Sponsor's analysis

The sponsor presented the mean survival estimate in Figures 1 and 2 of the sponsor's study report and the summary of survival estimates in Table 10 of the sponsor's study report. However, the sponsor did not present any statistical analysis for the rate mortality data.

The hepatocellular hypertrophy at 125 and 250 mg/kg/day in males and females were statistically significant ($p < 0.0005$). This finding was considered to be test item-related. The follicular cell hypertrophy and focal follicular cell hyperplasia in the thyroid gland in females at 250 mg/kg/day showed a positive trend with p-values of 0.0010 and 0.0007, respectively. These findings are considered to be caused by the liver cell hypertrophy and the consequently enhanced liver enzyme activities. The type, incidence, and severity of all other non-neoplastic changes in this study are considered to be incidental as they are common in rats of this strain and age.

3.3 Data Analyzed and Sources

The sponsor submitted the data in electronic format on March 18, 2008. The data are located in the EDR at the following link: [\\cdsesub1\n22304\S_00](#).

3.4 Reviewers' analyses

Dr. Meiyu Shen independently analyzed the survival data for males and females, separately. Dr. Shen and Dr. Rahman also independently analyzed the tumor data of the rat study for males and females, separately using the Peto method and the poly-3 method, respectively.

3.4.1 Survival analysis

The summaries of the mortality data are given in Table 7 for males. The time intervals 0-52, 53-78, 79-91 and 92-104 weeks were chosen for males. The Kaplan-Meier curves for males are shown in Figure 3. Analysis of Dose-Mortality Trend for Male Mice is presented in Table 8. From Figure 3, we can see that the mortality rate in the low and medium dosed groups is higher than that in the control group overall. The analysis of Dose-Mortality trend for males in Table 8 does not show a statistically significant dose-related trend among the control and the dosed groups because the p-value is 0.8217 (Cox method) and 0.6893 (Kruskal-Wallis tests), respectively, which is much larger than 0.05.

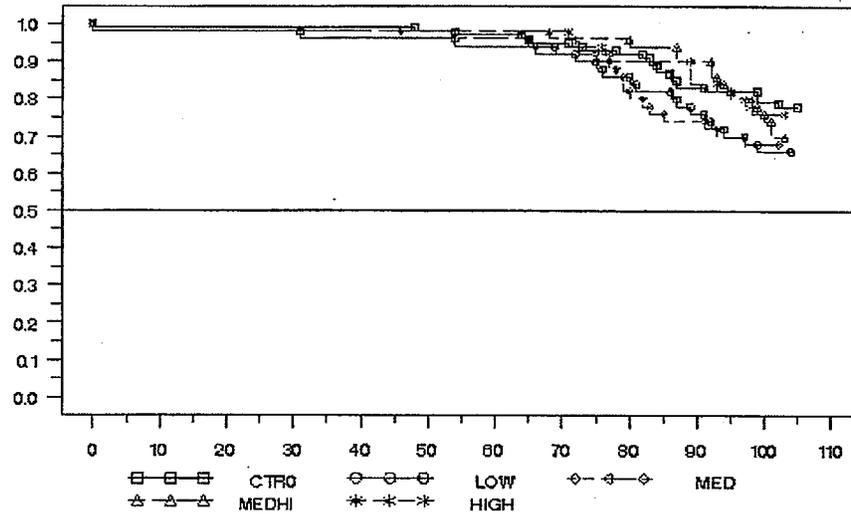


Figure 3 Kaplan-Meier Survival Functions for Male Rats

Table 7 Analysis of Mortality Data for Male Rats by Treatment and Time

Analysis of Mortality		No. Risk	No. Died	No. Alive	Pct Survival	Pct Mortality
CTR0	0-52	100	1	99	99.0	1.0
	53-78	99	6	93	93.0	7.0
	79-91	93	10	83	83.0	17.0
	92-104	83	4	79	79.0	21.0
	FINALKILL105-107	79	79	0		
LOW	0-52	50	1	49	98.0	2.0
	53-78	49	5	44	88.0	12.0
	79-91	44	6	38	76.0	24.0
	92-104	38	5	33	66.0	34.0
	FINALKILL105-107	33	33	0		
MED	0-52	50	1	49	98.0	2.0
	53-78	49	5	44	88.0	12.0
	79-91	44	7	37	74.0	26.0
	92-104	37	3	34	68.0	32.0
	FINALKILL105-107	34	34	0		
MEDHI	53-78	50	1	49	98.0	2.0
	79-91	49	2	47	94.0	6.0
	92-104	47	12	35	70.0	30.0
	FINALKILL105-107	35	35	0		
HIGH	53-78	50	4	46	92.0	8.0
	79-91	46	1	45	90.0	10.0
	92-104	45	7	38	76.0	24.0
	FINALKILL105-107	38	38	0		

Statistical Review of NDA22304

Table 8 Analysis of Dose-Mortality Trend for Male Rats

	Method			
	Cox		Kruskal-Wallis	
	Statistics	P-Value	Statistics	P-Value
Time-Adjusted Trend Test	4.3279	0.2282	4.4705	0.2149
Depart from Trend				
Dose-Mortality Trend	0.0508	0.8217	0.1598	0.6893
Homogeneity	4.3787	0.3572	4.6303	0.3274

Note: This test is run using Trend and Homogeneity Analyses of Proportions and Life Table Data Version 2.1, by Donald G. Thomas, National Cancer Institute

The summaries of the mortality data are given in Table 9 for females. The time intervals 0-52, 53-78, 79-91 and 92-104 weeks were chosen for females. The Kaplan-Meier curves for females are shown in Figure 4. Analysis of Dose-Mortality Trend for Female Mice is presented in Table 10. From Figure 4, we can see that there is not much difference in the survival probability between the control group and the dosed groups overall. The analysis of Dose-Mortality trend for females in Table 10 does not show a statistically significant dose-related trend among the control and the dosed groups at the 0.05 significance level because the p-value is 0.2726 (Cox method) and 0.2636 (Kruskal-Wallis tests), respectively.

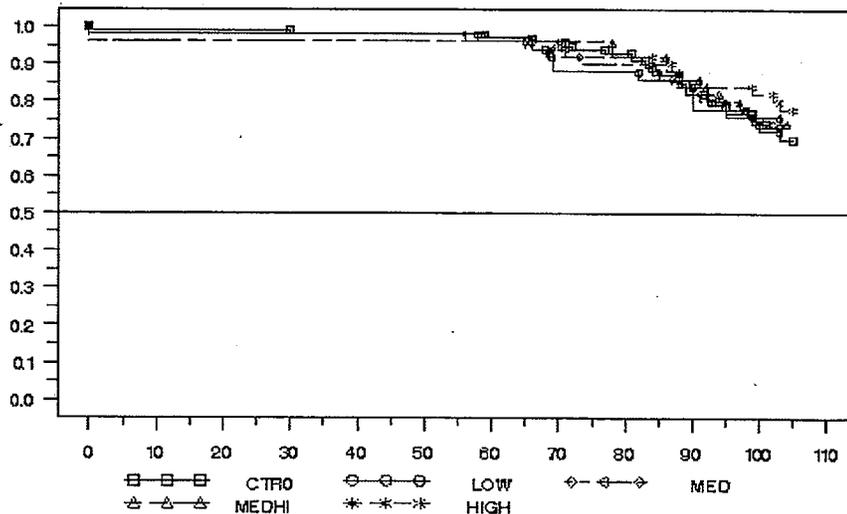


Figure 4 Kaplan-Meier Survival Functions for Female Rats

Table 9 Analysis of Mortality Data for Female Rats by Treatment and Time

Analysis of Mortality		No. Risk	No. Died	No. Alive	Pct Survival	Pct Mortality
CTR0	0-52	100	1	99	99.0	1.0
	53-78	99	5	94	94.0	6.0
	79-91	94	9	85	85.0	15.0
	92-104	85	12	73	73.0	27.0
	FINALKILL105-107	73	73	0		
LOW	53-78	50	4	46	92.0	8.0
	79-91	46	4	42	84.0	16.0
	92-104	42	6	36	72.0	28.0
	FINALKILL105-107	36	36	0		
MED	53-78	50	4	46	92.0	8.0
	79-91	46	5	41	82.0	18.0
	92-104	41	4	37	74.0	26.0
	FINALKILL105-107	37	37	0		
MEDHI	53-78	50	2	48	96.0	4.0
	79-91	48	5	43	86.0	14.0
	92-104	43	6	37	74.0	26.0
	FINALKILL105-107	37	37	0		
HIGH	53-78	50	3	47	94.0	6.0
	79-91	47	3	44	88.0	12.0
	92-104	44	4	40	80.0	20.0
	FINALKILL105-107	40	40	0		

Table 10 Analysis of Dose-Mortality Trend for Female Rats

	Method			
	Cox		Kruskal-Wallis	
	Statistics	P-Value	Statistics	P-Value
Time-Adjusted Trend Test	0.0494	0.9971	0.0278	0.9988
Depart from Trend				
Dose-Mortality Trend	1.2035	0.2726	1.2500	0.2636
Homogeneity	1.2529	0.8693	1.2778	0.8651

Note: This test is run using Trend and Homogeneity Analyses of Proportions and Life Table Data Version 2.1, by Donald G. Thomas, National Cancer Institute

3.4.2 Tumor data analyses Using Peto Method

In Dr. Shen's tumor data analysis, the dose response analyses in incidental tumors and fatal tumors were performed using the Peto prevalence method and the Peto death-rate method, respectively.

In Dr. Rahman's tumor data analysis, the dose-response analysis and pairwise comparisons between the combined control and individual treated groups were performed using the poly-3 method.

In Both Dr. Shen's and Dr. Rahman's analyses, the actual dose levels of treatment groups were used as the weights for the trend analysis. The numbers of tumor bearing animals and p-values of trend tests of individual tumor types from Dr. Shen's tumor data analysis were presented in Tables 11 and

Statistical Review of NDA22304

12 for males and females, respectively. The numbers of tumor bearing animals and p-values of trend tests and pairwise comparisons of individual tumor types from Dr. Shen's tumor data analysis were presented in Tables 11R and 12R for males and females, respectively.

Multiplicity for the trend testing was adjusted using a significance level of 0.025 for rare tumors, and 0.005 for common tumors because two species were studied. A tumor type with a background rate of 1 percent or less is classified as rare by Haseman; more frequent tumors are classified as common.

It is also well known that the approximation results may not be stable and reliable, and tend to underestimate the exact p-values when the total numbers of tumor occurrence across treatment groups are small. In this situation, the exact permutation trend test should be used to test for the positive trend. The exact permutation trend test is a generalization of the Fisher's exact test.

From Table 11, it is seen that there is no statistically significantly positive trend in both incidental tumors and fatal tumors for males. From Table 12, it is seen that there is a statistically significantly positive trend in rare tumor hepatocellular adenoma in females' liver because the p-value is 0.0149 (<0.025). There is a statistically significantly positive trend in rare tumor thymic lymphoma in females' thymus because the p-value is 0.0214 (<0.025).

Table 11 Report on Test for Positive Linear Dose-Tumor for Male Using Peto Method

Organ Code	Organ Name	Tumor Code	Tumor Name	CTR0	LOW	MED	MEDHI	HIGH	P-Value (Exact Method)	P-Value (Asymptotic Method)
101	CEREBRUM	10102	Granular cell tumor	0	0	1	0	1	0.1943	0.1841
101	CEREBRUM	10104	Meningioma	0	0	0	0	1	0.1735	0.0273
102	CEREBELLUM	10202	Benign Astrocytoma	1	0	0	0	0	1.0000	0.7245
102	CEREBELLUM	10203	Meningioma	0	0	0	0	1	0.1735	0.0273
1500	STOMACH	150017	Leiomyosarcoma	1	0	0	0	0	1.0000	0.7874
1800	LIVER	180039	Hepatocellular adenoma	2	0	0	1	2	0.1712	0.1370
2000	PANCREAS	200015	Islet cell adenoma	2	3	2	0	1	0.7831	0.7929
2000	PANCREAS	200020	Mixed acinar-islet cell adenom	0	0	0	1	0	0.3719	0.3123
2100	KIDNEYS	210030	Tubular cell adenoma	1	2	1	1	0	0.8214	0.8399
2100	KIDNEYS	210031	Renal Lipoma	1	0	0	0	0	1.0000	0.7874
2100	KIDNEYS	210041	Tubular cell carcinoma	0	1	1	0	0	0.8036	0.8375
2100	KIDNEYS	210043	Mesenchymal tumor	1	1	0	1	0	0.6620	0.7097
2500	TESTES	250018	Benign Leydig cell tumor	2	1	0	0	1	0.5679	0.5701
2500	TESTES	250021	Benign mesothelioma	1	0	0	0	0	1.0000	0.7874
2600	EPIDIDYMIDES	260008	Benign mesothelioma	1	0	0	0	0	1.0000	0.7874
2800	SEMINAL VESICLES	280012	Leiomyosarcoma	0	0	0	1	0	0.3333	0.2865
400	HEART	40028	Malignant atriocaval mesotheli	0	1	0	0	0	0.6422	0.7546
4100	PITUITARY GLAND	410011	Adenoma of pars distalis	29	19	20	23	16	0.3233	0.3263
4100	PITUITARY GLAND	410016	Ganglioneuroma (pars nervosa)	0	0	1	0	0	0.4862	0.5980
4200	THYROID GLAND	420019	C-cell adenoma	5	5	2	4	4	0.3023	0.3082
4200	THYROID GLAND	420020	Follicular cell adenoma	3	0	0	2	3	0.0847	0.0692

Statistical Review of NDA22304

4200	THYROID GLAND	420022	Follicular cell carcinoma	0	0	0	0	1	0.2258	0.0615
4401	ADRENAL CORTICES	440112	Adenoma	1	0	0	1	0	0.7566	0.6901
4401	ADRENAL CORTICES	440118	Adenocarcinoma	0	0	0	0	1	0.1743	0.0276
4402	ADRENAL MEDULLAS	440204	Benign pheochromocytoma	3	1	0	0	3	0.1637	0.1503
4402	ADRENAL MEDULLAS	440207	Malignant pheochromocytoma	0	1	1	0	0	0.6480	0.7538
4500	HEMOLYMPHORET. SYS	450005	Malignant lymphoma	4	0	4	0	0	0.7191	0.5793
4500	HEMOLYMPHORET. SYS	450006	Histiocytic sarcoma	1	0	0	0	0	1.0000	0.8461
4600	SPLEEN	460010	Hemangiosarcoma	0	0	1	0	0	0.7097	0.7547
4600	SPLEEN	460019	Leiomyoma	1	1	0	0	0	0.8709	0.8521
5000	THYMUS	500014	Benign thymoma	1	1	0	0	0	0.8730	0.8531
5000	THYMUS	500015	Thymic lymphoma	2	1	1	1	0	0.7976	0.8115
5104	MESENT. LYMPH NODE	510416	Hemangiosarcoma	3	0	1	0	2	0.2487	0.2311
5104	MESENT. LYMPH NODE	510419	Hemangioma	3	1	1	3	0	0.8240	0.8213
5700	SKIN/SUBCUTIS	570031	Hair follicle tumor	0	0	1	0	0	0.4886	0.5992
5701	SKIN NON-ROUTINE S	570108	Fibrosarcoma	1	0	0	0	0	1.0000	0.7569
5701	SKIN NON-ROUTINE S	570109	Keratoacanthoma	4	2	4	2	1	0.6974	0.7072
5701	SKIN NON-ROUTINE S	570112	Osteosarcoma	1	0	0	0	0	1.0000	0.7569
5701	SKIN NON-ROUTINE S	570114	Lipoma	1	0	2	0	0	0.5711	0.6513
5701	SKIN NON-ROUTINE S	570115	Squamous cell papilloma	3	1	1	0	0	0.9263	0.9036
5701	SKIN NON-ROUTINE S	570118	Fibroma	1	2	2	0	1	0.5972	0.6210
5701	SKIN NON-ROUTINE S	570119	Squamous cell carcinoma	0	1	0	0	0	0.6364	0.7418
5701	SKIN NON-ROUTINE S	570120	Sebaceous cell adenoma	0	0	1	0	0	0.5000	0.5477

Note: The check mark (☑) indicates statistically significant test results, based on the decision rule of FDA.CDER.Divisions of Biometrics.

Table 12 Report on Test for Positive Linear Dose-Tumor for Females Using Peto Method

Organ Code	Organ Name	Tumor Code	Tumor Name	CTR0	LOW	MED	MEDH	HIGH	P-Value (Exact Method)	P-Value (Asymptotic Method)
101	CEREBRUM	10104	Meningioma	0	0	0	0	1	0.1794	0.0298
101	CEREBRUM	10105	Ependymoma	1	1	0	0	0	0.8696	0.8420
1702	COLON	170205	Leiomyoma	1	0	0	0	0	1.0000	0.7945
1800	LIVER	180022	Hemangioma	1	0	0	0	0	1.0000	0.7946
1800	LIVER	180039	Hepatocellular adenoma	0	0	0	1	2	0.0149 (☑)	0.0031
1800	LIVER	180040	Hepatocellular carcinoma	0	0	0	1	0	0.3453	0.2979
2000	PANCREAS	200015	Islet cell adenoma	0	0	0	3	0	0.2211	0.1764
2000	PANCREAS	200017	Islet cell carcinoma	1	0	0	1	0	0.5693	0.5752
2000	PANCREAS	200020	Mixed acinar-islet cell adenom	0	0	1	0	0	0.5112	0.6108
2100	KIDNEYS	210030	Tubular cell adenoma	0	0	1	0	1	0.1450	0.1137
2100	KIDNEYS	210031	Renal Lipoma	0	0	0	1	0	0.3453	0.2979

Statistical Review of NDA22304

2100	KIDNEYS	210041	Tubular cell carcinoma	0	0	1	0	1	0.1463	0.1227
3200	OVARIES	320020	Benign granulosa cell tumor	1	0	0	0	1	0.3541	0.2249
3200	OVARIES	320026	Sex cord stromal tumor	0	2	1	2	1	0.2658	0.2863
3200	OVARIES	320029	Benign thecoma	0	1	0	1	0	0.4081	0.4773
3400	UTERUS	340008	Glandular polyp	1	0	0	0	1	0.3557	0.2259
3400	UTERUS	340009	Endometrial-stromal polyp	13	3	4	2	9	0.1964	0.1959
3400	UTERUS	340011	Hemangioma	1	0	0	1	0	0.5270	0.5233
3400	UTERUS	340013	Adenoma	1	0	0	0	2	0.0992	0.0438
3400	UTERUS	340014	Leiomyosarcoma	1	0	0	0	1	0.3068	0.1863
3400	UTERUS	340015	Endometrial-stromal sarcoma	1	1	0	1	0	0.8649	0.7111
3400	UTERUS	340019	Adenocarcinoma	1	0	2	1	2	0.1310	0.1119
3400	UTERUS	340020	Leiomyoma	0	0	1	0	0	0.5090	0.6110
3400	UTERUS	340022	Granular cell tumor	0	0	1	0	1	0.1508	0.1282
3400	UTERUS	340025	Squamous cell carcinoma	0	0	0	1	0	0.3468	0.2987
3401	CERVIX	340101	Stromal cell sarcoma	1	0	0	0	1	0.3573	0.2276
3401	CERVIX	340104	Fibroma	3	0	2	0	0	0.9275	0.9139
3401	CERVIX	340106	Granular cell tumor	0	1	0	0	1	0.2114	0.2052
3401	CERVIX	340109	Leiomyoma	0	0	1	1	0	0.4121	0.4314
3500	VAGINA	350004	Leiomyoma	0	0	1	0	0	0.5113	0.6124
3500	VAGINA	350011	Adenomatous polyp	1	0	0	0	0	1.0000	0.7954
3500	VAGINA	350012	Granular cell tumor	0	0	1	0	0	0.5113	0.6124
4100	PITUITARY GLAND	410011	Adenoma of pars distalis	52	24	22	22	22	0.8661	0.8654
4200	THYROID GLAND	420019	C-cell adenoma	7	5	6	4	7	0.1685	0.1669
4200	THYROID GLAND	420020	Follicular cell adenoma	1	2	2	1	0	0.8191	0.8170
4200	THYROID GLAND	420021	C-cell carcinoma	1	0	0	1	0	0.5724	0.5805
4200	THYROID GLAND	420022	Follicular cell carcinoma	0	0	1	0	0	0.5112	0.6108
4300	PARATHYROID GLANDS	430002	Adenoma	1	0	0	0	0	1.0000	0.7896
4402	ADRENAL MEDULLAS	440204	Benign pheochromocytoma	2	0	2	0	0	0.8639	0.8619
4402	ADRENAL MEDULLAS	440207	Malignant pheochromocytoma	0	0	0	1	0	0.3453	0.2979
4500	HEMOLYMPHORET. SYS	450005	Malignant lymphoma	1	2	2	1	0	0.3810	0.4050
4500	HEMOLYMPHORET. SYS	450006	Histiocytic sarcoma	1	0	1	0	0	0.5000	0.3745
5000	THYMUS	500014	Benign thymoma	2	5	2	2	2	0.5949	0.6095
5000	THYMUS	500015	Thymic lymphoma	1	0	1	3	3	0.0214	0.0125
5104	MESENT. LYMPH NODE	510419	Hemangioma	1	0	1	1	0	0.5940	0.6210
5108	MANDIB. LYMPH NODES	510815	Lymphosarcoma	0	0	0	0	1	0.1794	0.0298
5333	MANDIBULAR GLANDS	533308	Mixed Malignant tumor	0	0	0	0	1	0.1364	0.0165
5600	MAMMARY GLAND AREA	560003	Fibroadenoma	20	8	8	7	9	0.6003	0.6062
5600	MAMMARY GLAND AREA	560004	Adenocarcinoma	4	2	7	1	3	0.4712	0.4828
5600	MAMMARY GLAND AREA	560012	Adenoma	2	3	1	4	2	0.2181	0.2206
5600	MAMMARY GLAND	560013	Mammary fibroma	5	1	1	4	1	0.5993	0.6139

Statistical Review of NDA22304

AREA										
5701	SKIN NON-ROUTINE S	570108	Fibrosarcoma	3	0	0	0	0	1.0000	0.9041
5701	SKIN NON-ROUTINE S	570109	Keratoacanthoma	1	0	0	0	0	1.0000	0.7871
5701	SKIN NON-ROUTINE S	570114	Lipoma	0	0	1	0	0	0.5333	0.5759
5701	SKIN NON-ROUTINE S	570118	Fibroma	0	0	0	1	0	0.2667	0.2371
5701	SKIN NON-ROUTINE S	570119	Squamous cell carcinoma	0	1	1	0	0	0.6333	0.7269
5701	SKIN NON-ROUTINE S	570123	Basal cell carcinoma	0	0	0	1	0	0.2917	0.2335
5701	SKIN NON-ROUTINE S	570124	Malignant Schwannoma	0	1	0	0	0	0.7333	0.7503
6500	EYES	650019	Malignant Schwannoma	1	0	0	0	0	1.0000	0.7910
6800	BODY CAVITIES	680005	Granular cell tumor	0	0	0	0	1	0.1818	0.0349
6800	BODY CAVITIES	680010	Hemangioma	0	0	1	0	0	0.5909	0.6466

Note: The check mark (✓) indicates statistically significant test results, based on the decision rule of FDA.CDER.Divisions of Biometrics.

3.4.3 Tumor Data Analysis Using Poly-3 Method

From Table 11R, it is seen that there is no statistically significantly positive trend in tumors tested for males. From Table 12R, it is seen that there is a statistically significantly positive trend in rare tumor hepatocellular adenoma in females' liver because the p-value is 0.020 (<0.025). There is a statistically significantly positive trend in rare tumor thymic lymphoma in females' thymus because the p-value is 0.019 (<0.025).

Table 11 R

NDA 22_304
Dose Response Relationship Test and Pairwise Comparisons
Using Poly-3 test
Male Rats

Organ Name	Tumor Name	0 mg	10 mg	50 mg	125 mg	250 mg	P_Value				
		Cont N=100	Low N=50	Med N=50	Midhi N=50	High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H	P_Value MH C vs. H
ADRENAL CORTICE	Adenocarcinoma	0	0	0	0	1	0.170	.	.	.	0.333
	Adenoma	1	0	0	1	0	0.451	0.318	0.318	0.564	0.333
ADRENAL MEDULLA	Benign pheochromocytoma	3	1	0	0	3	0.166	0.380	0.686	0.714	0.317
	Malignant pheochromocytom	0	1	1	0	0	0.593	0.323	0.318	.	.
CEREBELLUM	Benign Astrocytoma	1	0	0	0	0	0.660	0.318	0.318	0.338	0.333
	Meningioma	0	0	0	0	1	0.170	.	.	.	0.333
CEREBRUM	Granular cell tumor	0	0	1	0	1	0.142	.	0.318	.	0.333
	Meningioma	0	0	0	0	1	0.170	.	.	.	0.333
EPIDIDYMIDES	Benign mesothelioma	1	0	0	0	0	0.886	0.537	0.537	0.564	0.557
HEART	Malignant atriocaval meso	0	1	0	0	0	0.502	0.318	.	.	.
HEMOLYMPHORET.	Histiocytic sarcoma	1	0	0	0	0	0.660	0.318	0.318	0.338	0.333

Statistical Review of NDA22304

	Malignant lymphoma	4	0	4	0	0	0.940	0.779	0.229	0.804	0.798
ILIAC LYMPH NOD	Hemangioma	1	0	0	0	0	0.660	0.318	0.318	0.338	0.333
KIDNEYS	Mesenchymal tumor	1	1	0	1	0	0.604	0.544	0.318	0.564	0.333
	Renal Lipoma	1	0	0	0	0	0.660	0.318	0.318	0.338	0.333
	Tubular cell adenoma	1	2	1	1	0	0.754	0.095	0.537	0.264	0.333
	Tubular cell carcinoma	0	1	1	0	0	0.594	0.318	0.318		
LIVER	Hepatocellular adenoma	2	0	0	1	2	0.137	0.537	0.537	0.264	0.416
MESENT. LYMPH N	Hemangioma	3	1	1	3	0	0.697	0.380	0.380	0.327	0.707
	Hemangiosarcoma	3	0	1	0	2	0.323	0.686	0.380	0.714	0.542
PANCREAS	Islet cell adenoma	2	3	2	0	1	0.781	0.184	0.380	0.564	0.707

**APPEARS THIS WAY
ON ORIGINAL**

Statistical Review of NDA22304

NDA 22_304
Dose Response Relationship Test and Pairwise Comparisons
Using Poly-3 test
Male Rats

Organ Name	Tumor Name	0 mg	10 mg	50 mg	125 mg	250 mg	P_Value				
		Cont N=100	Low N=50	Med N=50	Midhi N=50	High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	C vs. MH	P_Value C vs. H
PANCREAS	Mixed acinar-islet cell a	0	0	0	1	0	0.343	.	.	0.338	.
PARANASAL SINUS	Benign Schwannoma	1	0	0	0	0	0.660	0.318	0.318	0.338	0.333
PITUITARY GLAND	Adenoma of pars distalis	29	19	20	23	16	0.412	0.164	0.080	0.033	0.468
	Ganglioneuroma (pars nerv	0	0	1	0	0	0.343	.	0.318	.	.
SEMINAL VESICLE	Leiomyosarcoma	0	0	0	1	0	0.343	.	.	0.338	.
SKIN NON-ROUTIN	Fibroma	1	2	2	0	1	0.603	0.238	0.238	0.338	0.557
	Fibrosarcoma	1	0	0	0	0	0.660	0.318	0.318	0.338	0.333
	Keratoacanthoma	4	2	4	2	1	0.765	0.623	0.222	0.327	0.551
	Lipoma	1	0	2	0	0	0.733	0.318	0.244	0.338	0.333
	Osteosarcoma	1	0	0	0	0	0.660	0.318	0.318	0.338	0.333
	Sebaceous cell adenoma	0	0	1	0	0	0.343	.	0.318	.	.
	Squamous cell carcinoma	0	1	0	0	0	0.502	0.318	.	.	.
SKIN/SUBCUTIS	Squamous cell papilloma	3	1	1	0	0	0.941	0.380	0.380	0.714	0.707
	Hair follicle tumor	0	0	1	0	0	0.343	.	0.318	.	.
SPLEEN	Hemangiosarcoma	0	0	1	0	0	0.343	.	0.318	.	.
	Leiomyoma	1	1	0	0	0	0.777	0.537	0.318	0.338	0.333
STOMACH	Leiomyosarcoma	1	0	0	0	0	0.660	0.318	0.318	0.338	0.333
TESTES	Benign Leydig cell tumor	2	1	0	0	1	0.517	0.686	0.537	0.564	0.707
	Benign mesothelioma	1	0	0	0	0	0.886	0.537	0.537	0.564	0.557
THYMUS	Benign thymoma	1	1	0	0	0	0.777	0.537	0.318	0.338	0.333
	Thymic lymphoma	2	1	1	1	0	0.776	0.686	0.686	0.264	0.557
THYROID GLAND	C-cell adenoma	5	5	2	4	4	0.320	0.175	0.394	0.359	0.347

**APPEARS THIS WAY
ON ORIGINAL**

Statistical Review of NDA22304

NDA 22_304
Dose Response Relationship Test and Pairwise Comparisons
Using Poly-3 test
Male Rats

Organ Name	Tumor Name	0 mg	10 mg	50 mg	125 mg	250 mg	P_Value				
		Cont N=100	Low N=50	Med N=50	Midhi N=50	High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	C vs. MH	P_Value C vs. H
THYROID GLAND	Follicular cell adenoma	3	0	0	2	3	0.071	0.686	0.686	0.551	0.327
	Follicular cell carcinoma	0	0	0	0	1	0.170				

Table 12R

DA 22_304
Dose Response Relationship Test and Pairwise Comparisons
Using Poly-3 test
Female Rats

Organ Name	Tumor Name	0 mg	10 mg	50 mg	125 mg	250 mg	P_Value				
		Cont N=100	Low N=50	Med N=50	Midhi N=50	High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	C vs. MH	P_Value C vs. H
ADRENAL MEDULLA	Benign pheochromocytoma	2	0	2	0	0	0.914	0.704	0.537	0.717	0.717
	Malignant pheochromocytom	0	0	0	1	0	0.342	.	.	0.341	.
BODY CAVITIES	Granular cell tumor	0	0	0	0	1	0.171	.	.	.	0.341
	Hemangioma	0	0	1	0	0	0.342	.	0.331	.	.
CEREBRUM	Ependymoma	1	1	0	0	0	0.783	0.554	0.331	0.341	0.341
	Meningioma	0	0	0	0	1	0.171	.	.	.	0.341
CERVIX	Fibroma	3	0	2	0	0	0.914	0.704	0.537	0.717	0.717
	Granular cell tumor	0	1	0	0	1	0.200	0.331	.	.	0.341
	Leiomyoma	0	0	1	1	0	0.398	.	0.331	0.341	.
	Stromal cell sarcoma	1	0	0	0	1	0.342	0.331	0.331	0.341	0.567
CLITORAL GLANDS	Adenoma	0	0	1	0	0	0.342	.	0.331	.	.
COLON	Leiomyoma	1	0	0	0	0	0.669	0.331	0.331	0.341	0.341
EYES	Malignant Schwannoma	1	0	0	0	0	0.667	0.328	0.328	0.338	0.338
HEMOLYMPHORET.	Histiocytic sarcoma	1	0	1	0	0	0.648	0.331	0.561	0.341	0.341
	Malignant lymphoma	1	2	2	1	0	0.804	0.268	0.261	0.567	0.341
KIDNEYS	Renal Lipoma	0	0	0	1	0	0.342	.	.	0.341	.
	Tubular cell adenoma	0	0	1	0	1	0.144	.	0.336	.	0.341
	Tubular cell carcinoma	0	0	1	0	1	0.144	.	0.336	.	0.341
LIVER	Hemangioma	1	0	0	0	0	0.669	0.331	0.331	0.341	0.341
	Hepatocellular adenoma	0	0	0	1	2	0.020	.	.	0.341	0.114
	Hepatocellular carcinoma	0	0	0	1	0	0.342	.	.	0.341	.
MAMMARY GLAND A	Adenocarcinoma	4	2	7	1	3	0.504	0.649	0.034	0.556	0.458

Statistical Review of NDA22304

NDA 22_304
Dose Response Relationship Test and Pairwise Comparisons
Using Poly-3 test
Female Rats

Organ Name	Tumor Name	0 mg	10 mg	50 mg	125 mg	250 mg	P_Value				
		Cont N=100	Low N=50	Med N=50	Midhi N=50	High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H	
MAMARY GLAND A	Adenoma	2	3	1	4	2	0.267	0.210	0.704	0.103	0.429
	Fibroadenoma	20	8	8	7	9	0.689	0.606	0.629	0.759	0.607
	Mammary fibroma	5	1	1	4	1	0.581	0.659	0.649	0.365	0.678
MANDIB.LYMPH NO	Lymphosarcoma	0	0	0	0	1	0.171	.	.	.	0.341
MANDIBULAR GLAN	Mixed Malignant tumor	0	0	0	0	1	0.030	.	.	.	0.118
MESENT. LYMPH N	Hemangioma	1	0	1	1	0	0.535	0.331	0.554	0.567	0.341
OVARIES	Benign granulosa cell tum	1	0	0	0	1	0.342	0.331	0.331	0.341	0.567
	Benign thecoma	0	1	0	1	0	0.409	0.331	.	0.114	.
	Sex cord stromal tumor	0	2	1	2	1	0.271	0.108	0.331	0.114	0.346
PANCREAS	Islet cell adenoma	0	0	0	3	0	0.213	.	.	0.038	.
	Islet cell carcinoma	1	0	0	1	0	0.454	0.331	0.331	0.567	0.341
	Mixed acinar-islet cell a	0	0	1	0	0	0.342	.	0.331	.	.
PARATHYROID GLA	Adenoma	1	0	0	0	0	0.669	0.331	0.331	0.341	0.341
PITUITARY GLAND	Adenoma of pars distalis	52	24	22	22	22	0.849	0.593	0.723	0.796	0.796
SKIN NON-ROUTIN	Basal cell carcinoma	0	0	0	1	0	0.342	.	.	0.341	.
	Fibroma	0	0	0	1	0	0.342	.	.	0.341	.
	Fibrosarcoma	3	0	0	0	0	0.965	0.704	0.704	0.717	0.717
	Keratoacanthoma	1	0	0	0	0	0.669	0.331	0.331	0.341	0.341
	Lipoma	0	0	1	0	0	0.342	.	0.331	.	.
	Malignant Schwannoma	0	1	0	0	0	0.506	0.331	.	.	.
	Squamous cell carcinoma	0	1	1	0	0	0.592	0.336	0.331	.	.
THYMUS	Benign thymoma	2	5	2	2	2	0.556	0.040	0.403	0.421	0.421
	Thymic lymphoma	1	0	1	3	3	0.019	0.331	0.554	0.114	0.114

APPEARS THIS WAY
ON ORIGINAL

NDA 22_304
Dose Response Relationship Test and Pairwise Comparisons
Using Poly-3 test
Female Rats

Organ Name	Tumor Name	0 mg	10 mg	50 mg	125 mg	250 mg	P_Value				
		Cont N=100	Low N=50	Med N=50	Midhi N=50	High N=50	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	C vs. H	P_Value MH C vs. H
////////////////////////////////////											
THYROID GLAND	C-cell adenoma	7	5	6	4	7	0.131	0.225	0.225	0.553	0.093
	C-cell carcinoma	1	0	0	1	0	0.454	0.331	0.331	0.567	0.341
	Follicular cell adenoma	1	2	2	1	0	0.809	0.254	0.254	0.567	0.341
	Follicular cell carcinoma	0	0	1	0	0	0.342	.	0.331	.	.
TOOTH/TEETH	Ameloblastic odontoma	0	0	1	0	0	0.342	.	0.331	.	.
UTERUS	Adenocarcinoma	1	0	2	1	2	0.121	0.331	0.261	0.567	0.268
	Adenoma	1	0	0	0	2	0.092	0.331	0.331	0.341	0.268
	Endometrial-stromal polyp	13	3	4	2	9	0.182	0.845	0.727	0.940	0.306
	Endometrial-stromal sarco	1	1	0	1	0	0.607	0.561	0.331	0.567	0.341
	Glandular polyp	1	0	0	0	1	0.342	0.331	0.331	0.341	0.567
	Granular cell tumor	0	0	1	0	1	0.144	.	0.331	.	0.341
	Hemangioma	1	0	0	1	0	0.454	0.331	0.331	0.567	0.341
	Leiomyoma	0	0	1	0	0	0.342	.	0.331	.	.
	Leiomyosarcoma	1	0	0	0	1	0.342	0.331	0.331	0.341	0.567
Squamous cell carcinoma	0	0	0	1	0	0.342	.	.	0.341	.	
VAGINA	Adenomatous polyp	1	0	0	0	0	0.669	0.331	0.331	0.341	0.341
	Granular cell tumor	0	0	1	0	0	0.342	.	0.331	.	.
	Leiomyoma	0	0	1	0	0	0.342	.	0.331	.	.

3.4.4 Conclusion

From Figure 3, we can see that the mortality rate in the low and medium dosed groups is higher than that in the control group overall. The analysis of Dose-Mortality trend for males in Table 8 does not show a statistically significant dose-related trend among the control and the dosed groups because the p-value is 0.8217 (Cox method) and 0.6893 (Kruskal-Wallis tests), respectively, which is much larger than 0.05.

From Figure 4, we can see that there is not much difference in the survival probability between the control group and the dosed groups overall. The analysis of Dose-Mortality trend for females in Table 10 does not show a statistically significant dose-related trend among the control and the dosed groups at the 0.05 significance level because the p-value is 0.2726 (Cox method) and 0.2636 (Kruskal-Wallis tests), respectively.

From Tables 11 and 11R, it is seen that there is no statistically significantly positive trend in the tumors tested either using the Peto Method or the ploy-3 method for males. From Tables 12 and 12R, it is seen that the positive trend in rare tumor hepatocellular adenoma liver in females and the positive trend in rare tumor thymic lymphoma in thymus in females are statistically significant in both Peto and poly-3 tests.

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

/s/

Karl Lin

9/10/2008 03:49:15 PM

BIOMETRICS

Drs. Meiyu Shen and Atiar Rahman are the primary
reviewers of the carcinogenicity studies of this submission.
Karl Lin put the review results from both
of them into this report.



**STATISTICAL REVIEW AND EVALUATION
(Addendum)**

Biometrics Division: VI (HFD-705)

NDA NO.:	22-304
SERIAL NO.:	S_000
DATE RECEIVED BY THE CENTER:	March 19, 2008
DRUG NAME:	Tapentadol Hydrochloride (CG5503)
DOSAGE FORM:	Oral
INDICATION:	Moderate to severe acute pain
SPONSOR:	Ortho-McNeil-Janssen Pharmaceuticals, Inc.
DOCUMENTS REVIEWED:	Electronic Copy Dated March 19, 2008
NAME OF STATISTICAL REVIEWER:	Meiyu Shen, Ph.D. Atiar Rahman, Ph.D.
NAME OF PHARM TOX REVIEWER:	Kathleen Young

Meiyu Shen, Mathematical Statistician
Atiar Rahman, Mathematical Statistician

Concur:

Karl K Lin, Ph.D.
Team Leader, DBVI

Distribution: NDA 22-304
HFD-705/Karl K Lin, Ph.D.
OND/ODEII/DMEP, Kathleen A. Young

Background

The original statistical review and evaluation of the carcinogenicity studies were done by Dr. Meiyu Shen and Dr. Atiar Rahman of the Pharm/Tox Statistics Team of the Office of Biostatistics before the ECAC meeting that discussed the final results of this submission. At the meeting ECAC members requested that the Pharm/Tox Statistics Team perform the following additional analyses:

Lymphoma (all sites)

Hemangioma (all sites)

Combination of liver adenoma and carcinoma for both genders of the two species

This addendum report contains the results of the additional analyses requested by the ECAC.

Results of Additional Analyses

The results of the additional analyses are presented in the table below.

**APPEARS THIS WAY
ON ORIGINAL**

(Results of additional analyses for some combined tumor types suggested by ECAC)

NDA 22_304
Dose Response Relationship Test and Pairwise Comparisons
Using Poly-3 test

Male Rats

Organ Name	Tumor Name	0 mg	10 mg	50 mg	125 mg	250 mg	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. MH	P_Value C vs. H
		Cont N=100	Low N=50	Med N=50	Midhi N=50	High N=50					
ALL_ORGANS	HEMANGIOMA LYMPHOMA	4 6	1 1	1 5	3 1	0 0	0.785 0.969	0.511 0.700	0.511 0.252	0.441 0.739	0.807 0.911
LIVER	ADENNA+CARCONOMA	2	0	0	1	2	0.137	0.537	0.537	0.264	0.416

Female Rats

Organ Name	Tumor Name	0 mg	10 mg	50 mg	125 mg	250 mg	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. MH	P_Value C vs. H
		Cont N=100	Low N=50	Med N=50	Midhi N=50	High N=50					
ALL_ORGANS	HEMANGIOMA LYMPHOMA	3 2	0 2	2 3	2 4	0 3	0.741 0.126	0.704 0.421	0.537 0.210	0.556 0.103	0.717 0.217
LIVER	ADENNA+CARCONOMA	0	0	0	2	2	0.022			0.114	0.114

Male Mice

Organ Name	Tumor Name	0 mg	50 mg	100 mg	200 mg	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
		Cont N=102	Low N=51	Med N=51	High N=51				
ALL_ORGANS	HEMANGIOMA LYMPHOMA	6 8	1 8	2 6	1 1	0.771 0.784	0.750 0.126	0.482 0.276	0.665 0.794
LIVER	ADENNA+CARCONOMA	11	4	7	9	0.039	0.596	0.288	0.082

Female Mice

Organ Name	Tumor Name	0 mg	50 mg	100 mg	200 mg	P_Value Dos Resp	P_Value C vs. L	P_Value C vs. M	P_Value C vs. H
		Cont N=102	Low N=51	Med N=51	High N=51				
ALL_ORGANS	HEMANGIOMA LYMPHOMA	3 20	1 13	2 9	2 7	0.392 0.862	0.462 0.379	0.577 0.591	0.577 0.782
LIVER	ADENNA+CARCONOMA	0	1	0	0	0.402	0.367		

APPEARS THIS WAY
ON ORIGINAL

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

/s/

Karl Lin

9/10/2008 04:10:23 PM

BIOMETRICS

This addendum to the original review and evaluation report
contains results of additional statistical analyses requested by
the ECAC.



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/Serial Number: 22-304/000

Drug Name: Tapentadol hydrochloride immediate release tablets

Indication(s): Moderate to severe acute pain

Applicant: Johnson & Johnson Pharmaceutical Research & Development, L.L.C.

Date(s): Letter 1/22/2008; PDUFA 11/23/2008

Review Priority: Standard

Biometrics Division: Division of Biometrics II

Statistical Reviewer: Jonathan Norton, Ph.D.

Concurring Reviewers: Dionne Price, Ph.D.
Thomas Permutt, Ph.D.

Medical Division: Division of Anesthesia, Analgesia, and Rheumatology Products

Clinical Team: Ellen Fields, M.D.

Project Manager: Matthew Sullivan

Keywords: active control/non-inferiority, analysis of covariance, multiple comparisons, subgroup analyses

Table of Contents

LIST OF TABLES.....	3
LIST OF FIGURES.....	4
1. EXECUTIVE SUMMARY	5
1.1 CONCLUSIONS AND RECOMMENDATIONS	5
1.2 BRIEF OVERVIEW OF CLINICAL STUDIES	5
1.3 STATISTICAL ISSUES AND FINDINGS	6
2. INTRODUCTION	6
2.1 OVERVIEW.....	6
2.2 DATA SOURCES	9
3. STATISTICAL EVALUATION	9
3.1 EVALUATION OF EFFICACY	9
3.2 EVALUATION OF SAFETY	30
4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS	30
4.1 GENDER, RACE AND AGE	30
4.2 OTHER SPECIAL/SUBGROUP POPULATIONS	34
5. SUMMARY AND CONCLUSIONS	35
5.1 STATISTICAL ISSUES AND COLLECTIVE EVIDENCE	35
5.2 CONCLUSIONS AND RECOMMENDATIONS	36
5.3 REVIEW OF THE PROPOSED LABEL	36

**APPEARS THIS WAY
ON ORIGINAL**

LIST OF TABLES

Table 1: Controlled Clinical Studies of Efficacy – Completed (Sources: Listing of Clinical Studies, Reviewer’s Guide, clinical study reports)..... 8

Table 2: Applicant's Disposition of Subjects in ITT/Safety Analysis Set..... 12

Table 3: Applicant's Demographic and Baseline Characteristics 13

Table 4: Applicant’s Descriptive Statistics and Pairwise Comparison of SPID at Hour 48 hours, LOCF Imputation..... 16

Table 5: Descriptive Statistics Pairwise Comparison of SPID at 48 hours, BOCF and WOCF Imputation 16

Table 6: Applicant's Onset of Pain Relief (Source: Table 22 in CSR) 19

Table 7: Pairwise Comparison of SPID at Hour 48 hours, LOCF Imputation, by Center 20

Table 8: Applicant's Disposition of Subjects in Safety Analysis Set 23

Table 9: Applicant's Demographic and Baseline Characteristics 26

Table 10: Applicant’s Descriptive Statistics and Pairwise Comparison of Day Five SPID, LOCF Imputation 28

Table 11: Descriptive Statistics and Pairwise Comparison of SPID at Day Five, BOCF and WOCF Imputation 28

Table 12: Descriptive Statistics for SPID-48 (LOCF), by Gender 30

Table 13: Descriptive Statistics for SPID-48 (LOCF), by Race..... 31

Table 14: Descriptive Statistics for SPID-48 (LOCF), by Age 32

Table 15: Descriptive Statistics for Five Day SPID (LOCF), by Gender..... 32

Table 16: Descriptive Statistics for Five Day SPID (LOCF), by Race 33

Table 17: Descriptive Statistics for Five Day SPID (LOCF), by Age..... 33

Table 18: Descriptive Statistics for SPID 48 by Reload (2nd dose < 3 hrs), LOCF Imputation.. 34

Table 19: Descriptive Statistics of Day Five SPID, Subjects with No Gap after Baseline 35

**APPEARS THIS WAY
ON ORIGINAL**

LIST OF FIGURES

Figure 1: Study Design from Applicant (Source: Clinical Study Report)..... 10
Figure 2: Cumulative Distribution of Responders, 48 Hours 17
Figure 3: Study Design (Source: Clinical Study Report)..... 22
Figure 4: Cumulative Distribution of Responders, Day Five 29

**APPEARS THIS WAY
ON ORIGINAL**

1. EXECUTIVE SUMMARY

1.1 Conclusions and Recommendations

The Applicant seeks to have Tapentadol hydrochloride immediate release tablets approved for the indication of moderate to severe acute pain. Tapentadol HCl is a new molecular entity which acts as both a mu-opioid receptor agonist and as a norepinephrine uptake inhibitor. The proposed dosages are 50 mg, 75 mg, and 100 mg, to be taken every 4 to 6 hours as needed.

Considering the totality of evidence, I find tapentadol IR to be effective for the proposed indication

b(4)

1.2 Brief Overview of Clinical Studies

The application is primarily based on two pivotal studies. KF5503/32, which I will refer to as the Bunionectomy Study, was a double-blind, placebo- and active-controlled study of IR tapentadol for treatment of acute pain following bunionectomy. It was conducted at five sites in the United States, and there were 603 randomized subjects. They were assigned to the following treatment groups in approximately equal numbers: placebo, tapentadol 50 mg, tapentadol 75 mg, tapentadol 100 mg, and oxycodone HCl 15 mg. Patients qualified for the Bunionectomy study by reaching a sufficient pain score within nine hours of surgery. They were then randomized into the double-blind treatment period, which lasted for 72 hours (barring early discontinuation). The treatment period was followed by a nine day open-label study.

The primary efficacy variable in the Bunionectomy Study was the 48-hour sum of pain intensity differences (SPID₄₈), using an 11-point numeric rating scale for pain. There was also an extensive list of secondary efficacy variables. These included time to first rescue medication, as well the proportion of subjects who needed it. The responder rate was analyzed for different threshold values for percent change in pain intensity. Pain intensity was also assessed using a variety of other measures at hours 12, 24, 48, and 72. Times to perceptible, confirmed perceptible, and meaningful pain relief were measured using the double stopwatch method. In addition, the Patient Global Impression of Change was reported.

KF5503/33, which I will call the End Stage Joint Disease (ESJD) Study, was similar in design to the Bunionectomy Study. The main differences were that it was not post-surgical and the duration was somewhat longer. The study was conducted at 81 sites in the United States, Canada, the United Kingdom, Australia, and New Zealand. There were 674 subjects who were randomized to the following treatment groups: placebo, tapentadol IR 50 mg, tapentadol IR 75 mg, and oxycodone IR 10 mg. Note that this study did not include a 100 mg dose of tapentadol as the Bunionectomy Study did, and it used a lower dose of oxycodone. After going through screening, subjects in the ESJD study entered the Run-In Period (whose definition is discussed more in Section 1.3), during which they recorded their pain scores and completed a daily bowel movement questionnaire. Those who qualified then began the Double-Blind Treatment period,

which lasted for ten days. During this period, they received study medication every four to six hours and reported their pain level twice a day. The efficacy analysis was similar to that for the Bunionectomy Study. The principal difference in the efficacy analysis was that the primary efficacy variable was the *Five Day* SPID, rather than the SPID₄₈.

1.3 Statistical Issues and Findings

In both pivotal studies, the primary efficacy analysis was an analysis of covariance on the SPID (48 hour or five day) with the factors of treatment and center, and baseline pain intensity as a covariate. Multiplicity was controlled across doses in each study using the Hochberg step-up procedure. The primary imputation method was last observation carried forward (LOCF). In both studies, all tested doses of tapentadol were found superior to placebo in the primary efficacy analysis. Consistent results were obtained when LOCF was replaced with more conservative imputation methods.

Although the primary efficacy analysis favored tapentadol in both studies, there are procedural and statistical issues that merit attention. The Division of Scientific Investigations (DSI) found irregularities at two of the clinical sites. DSI concluded that the data from the sites are still acceptable, however, and my own analysis showed that the results from these sites were consistent with those from other sites. Another problem was that the definition of the Run-In Period used in the ESJD study was apparently modified post-hoc. I ultimately determined that none of these problems ultimately undermined a claim of efficacy for tapentadol, based on superiority to placebo.

┌

b(4)

└

2. INTRODUCTION

2.1 Overview

The Applicant seeks to have Tapentadol hydrochloride immediate release tablets approved for the indication of moderate to severe acute pain. Tapentadol HCl is a new molecular entity which acts as both a mu-opioid receptor agonist and as a norepinephrine uptake inhibitor. The proposed dosages are 50 mg, 75 mg, and 100 mg, to be taken every 4 to 6 hours as needed.

The Division of Anesthesia, Analgesia, and Rheumatology Products (DAARP) had a meeting with the Applicant on December 16, 2005 regarding the acute indications for tapentadol (IND 61,345). Several key clinical and statistical issues regarding the pivotal acute studies were addressed. It was agreed that a flexible dosing interval was acceptable, provided that the timing of the doses was carefully captured and analyzed. In response to a question about whether the label could instruct patients to take a second "reloading" dose as soon as an hour after the first dose if it failed to provide sufficient analgesia, DAARP stated that this was acceptable provided that "it is studied in that manner and the data support the benefit and safety of such use." In regard to the proper duration for an efficacy study, DAARP recommended a minimum of three to five days, but acknowledged that this may be too long a period for a bunionectomy study. Patients awaiting knee surgery were discussed as an additional study population. In terms of a primary pain endpoint, DAARP stated that pain should be evaluated over at least 48 hours and that last observation carried forward (LOCF) imputation was not acceptable. For additional endpoints, DAARP suggested that onset of analgesia be assessed using the double stopwatch method, as well as to time to re-medication.

b(4)

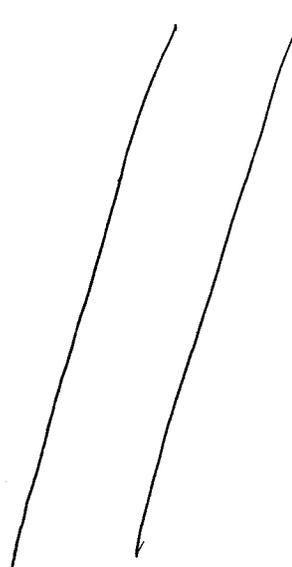
DAARP also had a pre-NDA meeting with the Applicant on June 5, 2007. In response to an Applicant question, DAARP stated that the Phase 3 bunionectomy (KF5503/32) and end-stage joint disease (KF5503/33) protocols appeared to be sufficient to support an application for the proposed acute indication. In regard to the Summary of Clinical Efficacy, DAARP requested that the Applicant perform a subset analysis of those patients in the bunionectomy study who received a second dose within three hours of the first dose ("reloaded"). DAARP also indicated that the proposed statistical analysis plan appeared adequate, with the caveat that LOCF imputation was not acceptable for the primary endpoint. The Applicant replied that they would retain LOCF for the primary endpoint, but that they understood that a study that failed with more conservative imputation methods would not support approval.

b(4)

Table 1 lists the controlled clinical studies for safety and efficacy reported by the Applicant. It should be noted I have not verified the results listed in the table, with the exceptions of those from KF5503/32 and KF5503/33. Although all of the listed studies were designed to test for efficacy, three of the ten studies failed to show an effect at the tested doses of tapentadol. All three of these studies, however, used an extended release formulation of tapentadol, rather than the immediate release formulation that is the subject of the present application. Moreover, two of

these trials used an approved analgesic (tramadol) as an active control and did not find it be significantly different from placebo, either.

Table 1: Controlled Clinical Studies of Efficacy – Completed (Sources: Listing of Clinical Studies, Reviewer’s Guide, and clinical study reports)

Study ID and Phase	Study Design and Indication/Population	Total Treated Subjects (including Placebo)	Applicant-Reported Results for Primary Endpoint
KF5503/02 Phase 2	Comparison of IR tapentadol, tramadol, ibuprofen, and placebo for dental pain	400	All active treatments superior to placebo for 8-hour total pain relief
KF5503/05 Phase 2	Comparison of IR tapentadol, morphine, ibuprofen, and placebo for post-surgical podiatric pain	517	Tapentadol doses \geq 50 mg superior to placebo for 8-hour total pain relief
KF5503/09 Phase 2	Comparison of ER tapentadol and placebo for chronic hip or knee joint OA	381	
KF5503/10 Phase 2	Comparison of ER tapentadol, ER tramadol, and placebo for chronic lower back pain	444	
KF5503/19 Phase 2	Comparison of ER tapentadol, ER oxycodone and placebo for pain due to OA of the knee	665	
KF5503/20 Phase 2	Comparison of ER tapentadol, ER tramadol and placebo for lower back pain	693	
KF5503/21 Phase 2	Comparison of IR tapentadol, oxycodone and placebo for post-surgical podiatric pain	269	
KF5503/22 Phase 2	Comparison of IR tapentadol, oxycodone and placebo for pain following bunionectomy	480	
			All tapentadol doses superior to placebo in SPRID ₁₂

b(4)

KF5503/32 Phase 3	Comparison of IR tapentadol, oxycodone and placebo for acute pain following bunionectomy	602	All tapentadol doses superior to placebo in 48 hour SPID
KF5503/33 Phase 3	Comparison of tapentadol, oxycodone and placebo in subjects eligible for joint replacement due to chronic OA	666	Both tapentadol doses superior to placebo in 5-day SPID

Notes: ER = extended release, IR=immediate release, OA = osteoarthritis. All tapentadol doses refer to base form. Only studies for which the Applicant provided a report are listed. As of 5/22/2008, there were two additional studies which were listed as "Completed; Full report in progress".

The application's claim of efficacy is based on two pivotal studies, KF5503/32 and KF5503/33. These studies will be the focus of the remainder of the review.

2.2 Data Sources

The electronic version of this NDA can be found at \\cdsesub1\evsprod\NDA022304.

3. STATISTICAL EVALUATION

3.1 Evaluation of Efficacy

KF5503/32 (Bunionectomy Study)

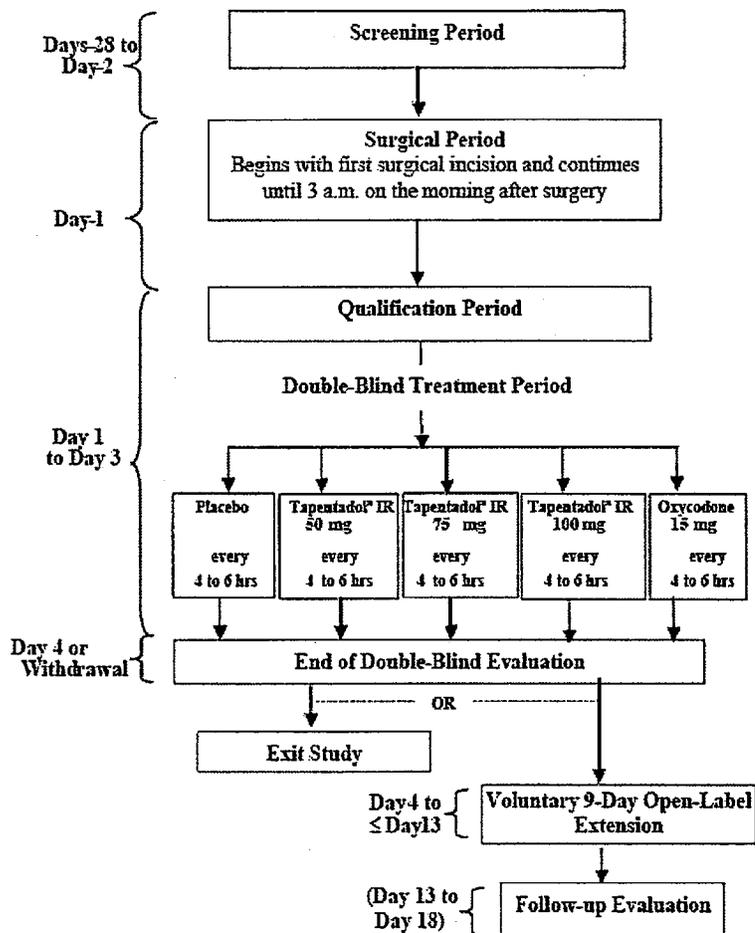
Study Design and Endpoints

KF5503/32, which will be referred to herein as the Bunionectomy Study, was a double-blind, placebo- and active-controlled study of the safety and efficacy of IR tapentadol for treatment of acute pain following bunionectomy. The study was conducted at five sites in the United States, and 603 subjects were randomized into five treatment groups in approximately equal numbers: placebo, tapentadol 50 mg, tapentadol 75 mg, tapentadol 100 mg, and oxycodone HCl 15 mg.

Figure 1, which was provided by the Applicant, illustrates the temporal sequence of the study. The Surgical Period began with the first surgical incision and continued until the end of popliteal sciatic block or systemic analgesia. The Qualification Period occurred for a maximum of nine hours after the end of the Surgical Period. During this time, subjects were monitored to see if they qualified for the next phase of the study by reaching a pain intensity score of 4 or greater in the required timeframe. If they reached the target pain level, then they were randomized and entered the Double-Blind Treatment Period. This period lasted for 72 hours barring earlier discontinuation. Patients who required rescue medication were discontinued for lack of efficacy. Subjects who completed the treatment period and who were considered medically stable were

eligible to enter the nine day open-label extension study. Consenting subjects returned to the clinic 13 to 18 days after randomization for a safety evaluation.

Figure 1: Study Design from Applicant (Source: Clinical Study Report)



The primary efficacy variable was the 48-hour sum of pain intensity differences (SPID₄₈), using an 11-point numeric rating scale. There was also an extensive list of secondary efficacy variables. Time to first rescue medication was defined to start from intake of first study medication, and subjects who did not take rescue medication or withdrew had their times censored at time of completion or discontinuation. The proportion of subjects who took rescue medication was also analyzed. Another secondary endpoint was the responder rate for different threshold values for percent change in pain intensity, assessed at 48 hours. The response rates at the 30% and 50% improvement levels were pre-specified for particular attention. Another set of secondary endpoints was total pain relief (TOTPAR), SPID (excluding 48 hours), and the sum of

total pain relief and sum of pain intensity difference (SPRID) at hours 12, 24, 48, and 72. Additional pain endpoints specified for exploratory purposes were pain relief (PAR), pain intensity difference (PID), sum of PAR and PID (PRID), and the sum of total pain relief and sum of pain intensity difference (SPRID) at various time points. Times to perceptible, confirmed perceptible, and meaningful pain relief were assessed using the double stopwatch method. The time to *confirmed perceptible* pain relief was defined as the time to perceptible pain relief (first stopwatch), provided that the subject experienced meaningful pain relief (second stopwatch). Finally, Patient Global Impression of Change was also assessed.

In addition to the previous endpoints which compared tapentadol HCl with placebo, oxycodone was also compared with placebo as a test of assay sensitivity. The endpoints assessed were SPID, TOTPAR, and SPRID at hours 12, 24, 48, and 72. SPID₄₈ was also the primary endpoint used in the non-inferiority comparison of Tapentadol IR 75 mg with oxycodone HCl IR 15 mg.

The intent-to-treat (ITT) analysis set was defined as “all randomized subjects who receive any amount of study medication(s) (i.e., at least one study medication intake following randomization) and have valid (non-missing) baseline pain assessment.” The baseline value was defined as the “last non-missing observation assessed prior to the first dose”.

Patient Disposition, Demographic and Baseline Characteristics

There were 918 subjects screened, of which 603 were randomized to the five treatment groups in approximately equal ratios. Of these subjects, 602 received the study drug and hence were included in the safety population, the exception being a subject who was found not to have met the pain criterion for inclusion. All of these subjects were also included in the ITT analysis set. The subsequent disposition of these subjects is shown in Table 2. As the table shows, there were a total of 143 subjects (24%) who withdrew during the Double-Blind Treatment Period. The majority (85%) of withdrawals were due to lack of efficacy. Withdrawal rates markedly decreased with dosage of tapentadol, going from 50% in the placebo group to 11% in the 100 mg group. The oxycodone group had a withdrawal rate of 14%. Among the 459 subjects who completed the Double-Blind Treatment Period, 428 entered the open-label study and all but one subject completed it.

APPEARS THIS WAY
ON ORIGINAL

Table 2: Applicant's Disposition of Subjects in ITT/Safety Analysis Set

(Source: Table 8 in Clinical Study Report)

	Placebo (N=120)	Tapentadol IR 50 mg (N=119)	Tapentadol IR 75 mg (N=120)	Tapentadol IR 100 mg (N=118)	Oxycodone IR 15 mg (N=125)	Total (N=602)
Reason for Withdrawal/termination	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
Completed	60 (50)	91 (76)	96 (80)	105 (89)	107 (86)	459 (76)
Withdrawn/Discontinued	60 (50)	28 (24)	24 (20)	13 (11)	18 (14)	143 (24)
Subject Choice ^a	0	1 (1)	1 (1)	1 (1)	3 (2)	6 (1)
Adverse Event	1 (1)	4 (3)	6 (5)	0	2 (2)	13 (2)
Lack of Efficacy	59 (49)	23 (19)	17 (14)	12 (10)	11 (9)	122 (20)
Other	0	0	0	0	2 (2)	2 (<1)

^a Subject withdrew consent

Percentages calculated with the number of subjects in each group as denominator.

Completion and discontinuation information was based on the study termination eCRF page.

Lack of Efficacy is defined as use of rescue medication during the double-blind period.

Table 3 shows the baseline characteristics of the subjects in the ITT set. The source is Table 9 in Clinical Study Report, but I confirmed the reported values.

APPEARS THIS WAY
ON ORIGINAL

Table 3: Applicant's Demographic and Baseline Characteristics

	(Study R331333-PAL-3003; KF5503/32: Intent-to-Treat Analysis Set)					
	Placebo (N=120)	Tapentadol IR 50 mg (N=119)	Tapentadol IR 75 mg (N=120)	Tapentadol IR 100 mg (N=118)	Oxycodone IR 15 mg (N=125)	Total (N=602)
Sex, n (%)						
N	120	119	120	118	125	602
Male	12 (10)	18 (15)	13 (11)	19 (16)	15 (12)	77 (13)
Female	108 (90)	101 (85)	107 (89)	99 (84)	110 (88)	525 (87)
Racial/ethnic Group, n (%)						
N	120	119	120	118	125	602
White	68 (57)	56 (47)	71 (59)	62 (53)	76 (61)	333 (55)
Black	23 (19)	27 (23)	19 (16)	24 (20)	25 (20)	118 (20)
Hispanic	26 (22)	32 (27)	24 (20)	30 (25)	23 (18)	135 (22)
Other	3 (3)	4 (3)	6 (5)	2 (2)	1 (1)	16 (3)
Age (Years)						
N	120	119	120	118	125	602
Category, n (%)						
<65	111 (93)	113 (95)	114 (95)	111 (94)	119 (95)	568 (94)
≥65	9 (8)	6 (5)	6 (5)	7 (6)	6 (5)	34 (6)
Mean (SD)	44.3 (14.45)	41.5 (13.27)	44.8 (13.61)	44.4 (13.68)	46.4 (13.02)	44.3 (13.66)
Median	45.0	42.0	47.5	46.5	49.0	46.0
Range	(18;77)	(18;75)	(19;72)	(18;74)	(18;73)	(18;77)
Weight (kg)						
N	120	119	120	118	125	602
Mean (SD)	75.6 (17.28)	76.4 (19.00)	74.3 (16.96)	78.2 (18.92)	77.9 (17.14)	76.5 (17.87)
Median	69.5	71.8	71.4	74.6	74.5	72.0
Range	(46;129)	(49;148)	(47;135)	(48;127)	(48;150)	(46;150)
Baseline Body Mass Index (kg/m²)						
N	120	119	120	118	125	602
Mean (SD)	27.8 (6.00)	28.1 (5.77)	27.6 (6.17)	28.5 (5.85)	28.9 (6.03)	28.2 (5.96)
Median	26.6	27.7	26.8	28.0	27.6	27.4
Range	(16;46)	(19;48)	(16;53)	(19;44)	(19;55)	(16;55)
Baseline Pain Intensity Score						
N	120	119	120	118	125	602
Category, n (%)						
Moderate	31 (26)	25 (21)	32 (27)	33 (28)	27 (22)	148 (25)
Severe	89 (74)	94 (79)	88 (73)	85 (72)	98 (78)	454 (75)

Protocol Deviations

The Division of Scientific Investigations (DSI) performed a directed inspection of sites 1004 and 1005 and issued a Form 483 on May 15, 2008, which included two observations. Observation 1 was "Failure to ensure proper monitoring of the study", specifically that the Applicant's clinical trials and management team (which was run by a CRO) failed to file documents associated with monitoring visits in a timely manner. Observation 2 was, "The sponsor failed to fully report to FDA all information relevant to safety and efficacy data discovered during monitored (*sic*) of the clinical sites as they are not in the clinical study report." Specifically, missed pain assessments from six subjects were not listed in the clinical report as deviations. The Applicant responded to

these observations on May 21, 2008. In response to Observation 1, the Applicant acknowledged the CRO was “not always timely” and described remedial measures. In regard to Observation 2, the Applicant acknowledged that 14 missing pain assessments were not reported as protocol deviations. The Applicant stated that their error in omitting these protocol deviations did not change the results of the primary or secondary analyses because the subjects were still included in the ITT set. DSI concluded that “The sponsor’s response is acceptable... and the data submitted by the sponsor may be used in support of the respective indication.” They did suggest, however, that the review division may want to exclude subjects 304078, 305069, 305147 and 305178.

Another inspection of site 1005 was performed by DSI from May 27th to May 29th 2008. The inspectors found that one subject (305140) took a prohibited medication (the antidepressant Effexor) during the study. They also found that the consent form “did not note the possibility that the FDA may inspect the records, as required by 21 CFR 50.25(a)(5)”. DSI concluded that “The data appear acceptable in support of the pending application”.

I assessed the impact of these irregularities by analyzing the primary endpoint by center, and also by re-running the primary analysis with the suggested subjects excluded. See *Possible Impact of Protocol Deviations*.

Statistical Methods

Methods for Primary Efficacy Variable

The primary efficacy analysis on the SPID₄₈ was an analysis of covariance with the factors of treatment and center, and baseline pain intensity as a covariate. The Hochberg step-up Bonferroni procedure (implemented in SAS PROC MULTTEST) was used to control the type I error when testing the three tapentadol dosage levels against placebo.

Values for missing pain assessments were imputed as follows: Intermittent missing measurements were imputed by linear interpolation. For the primary analysis, pain measurements after discontinuation were imputed using last observation carried forward. If there were no post-baseline values, then the baseline was carried forward. As a test of sensitivity, baseline observation carried forward (BOCF) and worst observation carried forward (WOCF) were used as alternative imputation methods after discontinuation. BOCF could not be used for PAR because there was no baseline.

Methods for Secondary Efficacy Variables

The planned statistical methods for the secondary endpoints were as follows. The “time to first rescue medication” was to be estimated using Kaplan-Meier and compared using a log-rank test, with center as a stratification factor. (The planned analysis was altered due to low use of rescue medication.) The proportion of subjects needing rescue medication was compared in each group using the Cochran-Mantel-Haenszel (CMH) test, stratifying by center.

The percent change from baseline pain intensity was computed at 48 hours. The distribution of responder rates, using different threshold values, was then compared between treatment groups using a Wilcoxon rank-sum test with stratification by center. (The Clinical Study Report and Statistical Analysis Plan refer to the "Gehan test", which is actually a generalization of the Wilcoxon test for censored survival data.) In addition, the proportion of subjects achieving a 30% and 50% response was compared across groups using the CMH test (controlling for center). The same analysis was performed for subjects who took an early second dose. TOTPAR, PID, and SPRID were computed at each analysis time point (Hour 12, 24, 48, and 72) and analyzed using ANCOVA models with treatment, center, and baseline pain intensity as factors. Descriptive statistics were also computed for PAR, PID, PRID (=PAR+PID), TOTPAR, SPID, and SPRID for each treatment group. In addition, the results for PAR, PID, and PRID were plotted over time.

The distributions of time to perceptible, meaningful, and confirmed perceptible pain relief were estimated using Kaplan-Meier and compared using log-rank statistics (stratifying for center). If subjects did not experience pain relief, then their times were censored at 12 hours from the first dose or the time of early discontinuation (whichever came first). If a subject did not have confirmed perceptible pain relief, then their time was censored at 12 hours or time of discontinuation. The PGIC scores were summarized and compared with placebo using the CMH test.

A sequential gatekeeping strategy was used to test selected secondary variables, some of which would not traditionally be considered *efficacy* endpoints in this setting. The Statistical Analysis Plan (SAP) describes the gatekeeping procedure as follows (p. 1159 in Clinical Study Report):

The following selected secondary hypotheses will be tested in a sequential manner if the null hypothesis for the primary endpoint is rejected. At each step, when the preceding null hypothesis fails to be rejected, further comparisons will not be performed.

1. For those CG5503 IR dose groups that are shown to be superior to the placebo group in SPID48, compare with the placebo group in terms of the "Time to First Rescue Medication" using the Hochberg's procedure at the 0.05 significance level;
2. If CG5503 base IR 75mg is shown to be superior to the placebo group in SPID48, test the superiority of CG5503 base IR 75mg against Oxycodone 15mg in terms of the composite event of nausea and vomiting rate at a two-sided 0.05 significance level.
3. If CG5503 base IR 75mg is shown to be superior to the oxycodone 15mg in terms of the composite event, compare the NI of CG5503 base IR 75mg against Oxycodone 15mg at a one-sided 0.025 significance level with SPID48. The NI margin will be 10% of the entire range of the primary endpoint (i.e. 48 out of 480 points).
4. Test the superiority of CG5503 base IR 75mg against Oxycodone 15mg in terms of the constipation adverse event rate at a two-sided 0.05 significance level.

The "composite event of nausea or vomiting" refers to a subject reporting either nausea or vomiting. The comparisons of the rates of nausea/vomiting and constipation for tapentadol and oxycodone were done using the CMH test, controlling for center. The Applicant's proposed label does *not* include a comparison of these adverse outcomes for tapentadol vs. oxycodone.

Results and Conclusions

Primary Efficacy Endpoint

I was able to replicate the Applicant's finding that all three doses of tapentadol were superior to placebo using LOCF imputation, as seen in Table 4. In a meeting and subsequent letter, DAARP had strongly suggested that the Applicant not use LOCF imputation for the primary analysis. As the p-values in Table 5 show, however, the finding of efficacy for all doses of tapentadol holds even when BOCF and WOCF are used to impute missing pain values from discontinuations. Considering the descriptive statistics, the large standard deviations (relative to the corresponding means) are noteworthy, particularly in the placebo group. These large values reflect the fact that many individual SPID values were negative, which happened when a subject's pain increased relative to the baseline value.

Table 4: Applicant's Descriptive Statistics and Pairwise Comparison of SPID at Hour 48 hours, LOCF Imputation

	Placebo (N=120)	Tapentadol IR 50 mg (N=119)	Tapentadol IR 75 mg (N=120)	Tapentadol IR 100 mg (N=118)	Oxycodone HCl IR 15 mg (N=125)
0-48 Hours					
Mean (SD)	24.5 (120.93)	119.1 (125.86)	139.1 (118.93)	167.2 (98.99)	172.3 (110.86)
Median	43.4	127.6	131.3	158.5	170.6
(Range)	(-278;274)	(-185;402)	(-199;462)	(-94;408)	(-190;431)
LS Means (diff from placebo)	--	88.2	113.5	141.4	142.4
95% CI	--	60.71 to 115.59	86.12 to 140.81	113.98 to 168.90	115.28 to 169.47
Adjusted p-value vs. placebo ^a	--	<0.001	<0.001	<0.001	

^a Based on analysis of covariance model with factors of treatment, center, and baseline pain intensity as a covariate. Adjusted p-values using Hochberg procedure. Oxycodone group is not included.

Table 5: Descriptive Statistics Pairwise Comparison of SPID at 48 hours, BOCF and WOCF Imputation

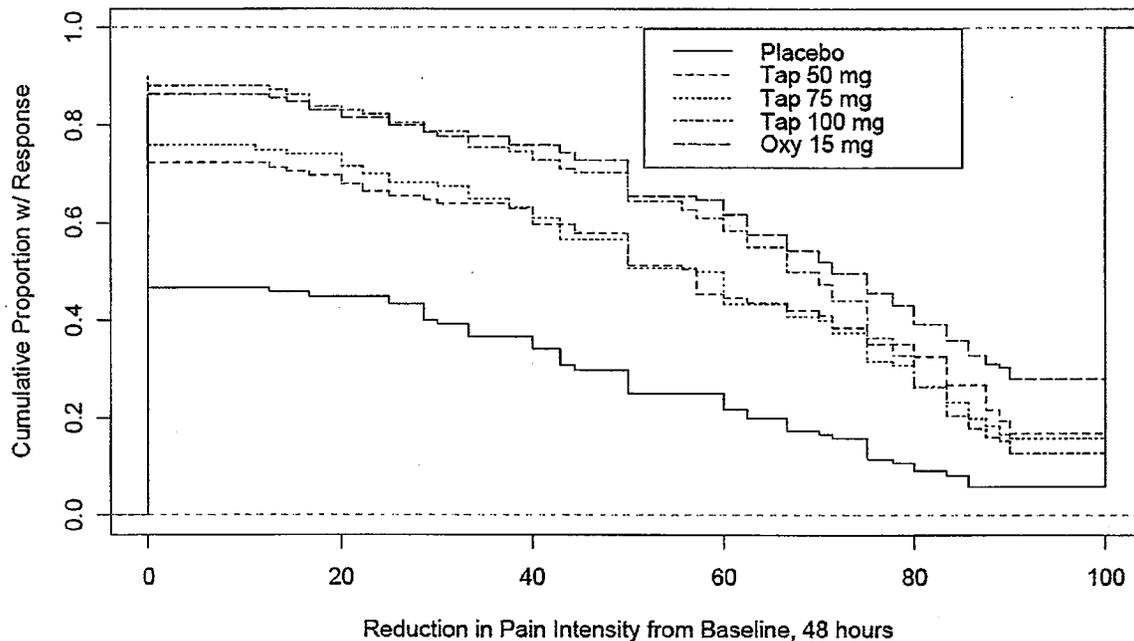
Imputation	Statistic	Placebo	Tap. 50 mg	Tap. 75 mg	Tap. 100 mg.	Oxy. 15 mg
BOCF	Mean	57.5	128.2	140.5	164.9	171.2
	Standard Deviation	79.6	105.6	111.1	99.9	102.5
	LS Mean (diff from placebo)	--	65.4	82.2	106.5	109.4
	p-value (unadj.)	--	< .0001	<.0001	<.0001	<.0001
WOCF	Mean	19.9	113.4	134.4	163.3	166.2
	Standard Deviation	122.5	129.1	120.9	102.0	113.1
	LS Mean (diff from placebo)	--	87.1	113.4	142.3	141.0
	p-value (unadj.)	--	< .0001	< .0001	< .0001	< .0001

(Note: p-values not adjusted for multiplicity)

Responder Analysis

Figure 2 shows the cumulative proportion of subjects who showed different percentages of reduction in pain intensity from baseline at 48 hours. It is my replicated version of the Applicant's Figure 4 in the Clinical Study Report. Subjects who discontinued prior to 48 hours or who had a worse score than baseline were given a value of zero (non-responder).

Figure 2: Cumulative Distribution of Responders, 48 Hours



I replicated the Applicant's finding that all active treatments had a different responder distribution than placebo, using the Wilcoxon rank-sum test ($p < .0001$ in all cases). This result supports the appearance of a strong separation from placebo in the figure. I also verified the findings that all three doses of tapentadol were more likely to achieve both 30% and 50% response than placebo (unadjusted $p < .0001$ in all cases).

Tapentadol vs. Oxycodone

I verified the results in this section, which were reported by the Applicant. A sequential testing strategy was used to compare tapentadol 75 mg to oxycodone 15 mg. Once it was verified that tapentadol 75 mg was superior to placebo on the SPID₄₈, the next step (as I read the procedure) was to compare the two treatments on time to rescue analgesia. A stratified log-rank test showed that tapentadol 75 mg differed from placebo in the distribution of time to rescue medication ($p <$

b(4)

.0001). Having established this, the next step was to compare tapentadol 75 mg to oxycodone 15 mg for the compound event of nausea/vomiting during the double-blind treatment period. This adverse event was experienced by 41% of subjects in the tapentadol IR group compared with 70% of subjects in the oxycodone HCl IR group, yielding $p < .0001$ from the CMH test (stratified for site). Since that comparison favored tapentadol, tapentadol 75 mg was then tested for non-inferiority against oxycodone 15 mg based on the SPID48. The lower bound for the 95% confidence interval (two-sided) for the difference between the two SPID scores (tapentadol minus oxycodone) was -56.0. Since this exceeded the prespecified non-inferiority margin, the Applicant failed to show that tapentadol 75 mg was non-inferior to oxycodone.

The Applicant also performed a "post-hoc" (as they termed it) comparison of the tapentadol 100 mg dose to oxycodone 15 mg, using the step-wise approach originally planned for the 75 mg dose. Since tapentadol 100 mg was superior to placebo on the SPID₄₈, the next step was to show that the two treatments differ in the distributions of time to rescue analgesia ($p < .0001$). In regard to nausea/vomiting, subjects in the 100 mg group experienced this event 53% of the time, compared to 70% in the oxycodone group. This difference was significant ($p = .007$). In regard to non-inferiority on the SPID48 outcome, the lower bound on the confidence interval was -28.1. If one accepts the non-inferiority margin of 48 as appropriate, then this result indicates that the 100 mg dose is non-inferior to oxycodone 15 mg. The previous steps having favored tapentadol, in the final test of the sequence the 100 mg dose was not shown to have a lower rate of constipation than oxycodone ($p = .24$).

Time to Onset of Analgesia

Table 6 shows the descriptive statistics for the times to perceptible relief, meaningful relief, and confirmed reliefs for the different treatment groups. It also shows the results of log-rank tests to compare the "survival" (time to onset) curves for each of the active treatment groups to placebo. The table was provided by the Applicant, but I confirmed the results for confirmed perceptible relief because ζ is in the proposed label.

b(4)

APPEARS THIS WAY
ON ORIGINAL

Table 6: Applicant's Onset of Pain Relief (Source: Table 22 in CSR)

	Placebo (N=120)	Tapentadol IR 50 mg (N=119)	Tapentadol IR 75 mg (N=120)	Tapentadol IR 100 mg (N=118)	Oxycodone HCl IR 15 mg (N=125)
Perceptible Relief					
Events (%)	96 (80.0)	105 (88.2)	116 (96.7)	115 (97.5)	121 (96.8)
Median	34.0	46.0	31.0	35.5	30.0
(95% CI) ^a	(27.0; 59.0)	(37.0; 58.0)	(28.0; 44.0)	(31.0; 42.0)	(28.0; 34.0)
Nominal p-value vs. placebo ^b		0.935	0.029	0.045	<0.001
Meaningful Relief					
Events (%)	65 (54.2)	94 (79.0)	101 (84.2)	103 (87.3)	107 (85.6)
Median	240.0	123.0	104.0	94.0	77.0
(95% CI) ^a	(155.0; 468.0)	(93.0; 164.0)	(71.0; 128.0)	(84.0; 118.0)	(60.0; 92.0)
Nominal p-value vs. placebo ^b		0.008	<0.001	<0.001	<0.001
Confirmed Perceptible Relief					
Events (%)	65 (54.2)	93 (78.2)	100 (83.3)	103 (87.3)	106 (84.8)
Median	100.0	46.0	32.0	37.0	31.0
(95% CI) ^a	(39.0;)	(37.0; 59.0)	(29.0; 46.0)	(32.0; 44.0)	(28.0; 36.0)
Nominal p-value vs. placebo ^b		0.005	<0.001	<0.001	<0.001

^a In minutes; based on Kaplan–Meier product limit estimates.

^b Pairwise comparison: Log rank test stratified with center.

Possible Impact of Protocol Deviations

Due to the protocol deviations observed in centers 1004 and 1005, I tested for robustness of the results for the primary endpoint across center. As Table 7 shows, the finding of efficacy for tapentadol was quite consistent across centers. The only apparent exception was a possible lack of efficacy of the 50 mg dose in center 1001. Since there is no reason to suspect center 1001 in particular, I attribute the low estimate of effect at this center/dose to statistical variation and multiplicity. (The standard error on the least-squares mean at this center/dose was 31.7.) I also re-ran the primary analysis with the four subjects noted by DSI excluded, and the results were virtually identical to those from the full ITT set.

**APPEARS THIS WAY
ON ORIGINAL**

Table 7: Pairwise Comparison of SPID at Hour 48 hours, LOCF Imputation, by Center
 (Note: p-values not adjusted for multiplicity.)

Site	N	Statistic	Tap. 50 mg	Tap. 75 mg	Tap 100 mg.
1001	102	LS Means (diff from placebo)	9.2	93.2	127.4
		p-value (unadj.)	.77	.004	.0002
1002	153	LS Means	91.5	94.5	155.4
		p-value	.0006	.0003	< .0001
1003	98	LS Means	124.9	175.8	154.4
		p-value	.0005	< .0001	< .0001
1004	116	LS Means	80.8	90.8	111.4
		p-value	.015	.0054	.0007
1005	133	LS Means	113.0	124.1	142.2
		p-value	.0008	.0002	< .0001

KF5503/33 (End-Stage Joint Disease Study)

Study Design and Endpoints

KF5503/33, which will also be referred to herein as the End-Stage Joint Disease Study, was a double-blind, placebo- and active-controlled study of the safety and efficacy of IR tapentadol for treatment of pain in patients eligible for primary joint replacement surgery for end-stage joint disease. The study was conducted at a total of 81 sites in the United States, Canada, the United Kingdom, Australia, and New Zealand. Of the 1101 subjects screened, a total of 674 were randomized to the following treatment groups: 172 subjects to placebo, 161 to tapentadol IR 50 mg, 169 to tapentadol IR 75 mg, and 172 to the oxycodone IR group.

Figure 3, which comes from the Applicant's Clinical Study Report, illustrates the design of the study. The study began with the Screening Period, in which subjects first visited the study center, gave informed consent, and received a diary to record their pain scores. That was followed by a Run-In Period during which subjects recorded their pain intensity scores twice daily, and also completed a daily bowel movement questionnaire. (The definition of the Run-in period is discussed more under *Patient Disposition, Demographic and Baseline Characteristics*.) The Double-Blind Treatment Period lasted for 10 days and involved two clinic visits. In the first visit (Day 1), qualified subjects were randomized to one of the four treatment groups. The second visit, which occurred between Day 6 and Day 8, was a mid-period evaluation. Subjects took their first dose of study medication after they arrived home on Day 1. Subsequent doses followed every four to six hours, during the subject's waking hours. The subjects completed pain assessments twice a day. In the Follow-Up Period, they returned to the clinic for a final evaluation. Subjects who withdrew early underwent a follow-up evaluation at the next scheduled visit or earlier, at the investigator's discretion.

Inclusion criteria included (in addition to other criteria) being diagnosed with non-inflammatory end-stage joint disease, either waiting for joint replacement therapy or having been offered it and

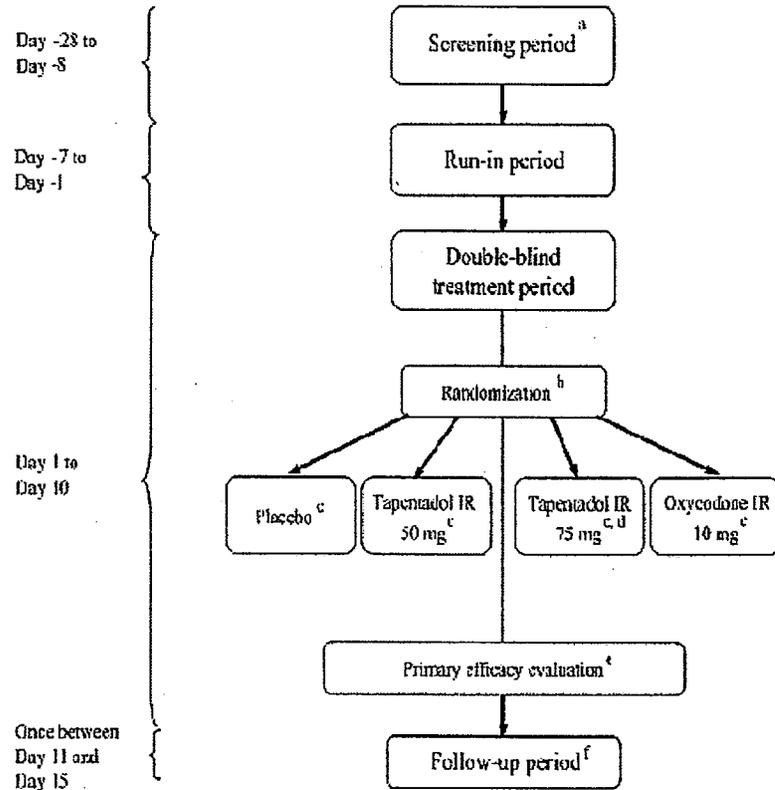
declined it, and having a minimum baseline pain score. The pain score requirement is discussed more in the next section.

The primary efficacy variable was a Five Day SPID, which was based on an 11-point numeric scale and assessed twice a day. There were also several secondary efficacy variables which largely overlapped those evaluated in the Bunionectomy Study. Time to first rescue medication was assessed in those subjects who received it (and who were then supposed to be withdrawn for lack of efficacy). A responder analysis was performed using change from baseline in average pain intensity on Day Five, and additionally on Days 2 and 10. Additional secondary pain endpoints included pain relief (PAR), total pain relief (TOTPAR), pain intensity difference (PID), and sum of TOTPAR and SPID (SPRID), all evaluated at Days 2, 5, and 10. Also included among the secondary endpoints were the times to perceptible, meaningful, and confirmed perceptible Pain Relief. An additional efficacy endpoint was patient global impression of change (PGIC) compared to placebo.

The assessments in the previous paragraph were used to compare the two doses of tapentadol to placebo. In addition, oxycodone was compared to placebo using the SPID, TOTPAR, and SPRID on Days 2, 5, and 10. Oxycodone was also compared to two doses of tapentadol using the Five Day SPID. Also used to compare oxycodone and tapentadol were the composite rate of nausea or vomiting and the rate of constipation.

**APPEARS THIS WAY
ON ORIGINAL**

Figure 3: Study Design (Source: Clinical Study Report)



^a First clinic visit.

^b Second clinic visit. All other baseline evaluations are to be collected before randomization.

^c The first dose will be taken in the evening of Day 1. Subsequent doses will be taken orally every 4 to 6 hours relative to the previous dose during waking hours.

^d Includes a titration step of 50 mg for Day 1 and 75 mg for Days 2 to 10.

^e Third clinic visit.

^f Fourth clinic visit to undergo the end-of-treatment evaluations. Those subjects who withdraw from the study will undergo their early-withdrawal evaluations during the next scheduled visit or earlier as determined by the investigator.

Patient Disposition, Demographic and Baseline Characteristics

There were 1101 subjects screened, of whom 674 were randomized to the four treatment groups. Eight of these randomized subjects did not received any study drug (3 in the placebo group, 4 in the tapentadol 50 mg group, 1 in the tapentadol 70 mg group) and were omitted from subsequent analyses. The reasons that these subjects were withdrawn included “withdrew consent” (3 subjects), “lost to follow-up” (3 subjects), and “other” (2 subjects). The two subjects in the “other” category included one who had an atrial fibrillation and another who was found to meet an exclusion criterion (SSRI regimen not stable for 28 days prior to screening).

Since they received a dose of study medication, the remaining 666 subjects were included in the safety analysis set. Their subsequent disposition is summarized in Table 8 which was provided by the Applicant. As can be seen from the table, the lowest rate of withdrawal was in the placebo group (10%) while the highest was in the oxycodone group. The withdrawal rates in the tapentadol 50 mg and 75 mg groups were 18% and 26%, respectively. Among all of the treatment groups, subjects in the oxycodone group were most likely to withdraw due to an adverse event (30%). Compared to the other groups, subjects in the placebo group were most likely to withdraw due to lack of efficacy (4%). Overall, only 2% of subjects withdraw due to lack of efficacy. In addition to the twelve listed in the table, there were six additional subjects who received rescue medication but were *not* treated as having withdrawn due to lack of efficacy. I verified that the data from these subjects did not change the primary efficacy outcome.

**Table 8: Applicant's Disposition of Subjects in Safety Analysis Set
(Source: Table 8 in Clinical Study Report)**

Completion Status Reason for Withdrawal	Placebo	Tapentadol IR 50 mg	Tapentadol IR 75 mg	Oxycodone HCl IR 10 mg	Total
	(N=169) n (%)	(N=157) n (%)	(N=168) n (%)	(N=172) n (%)	(N=666) n (%)
Completed ^a	152(90)	129(82)	125(74)	112(65)	518(78)
Withdrawn ^a	17(10)	28(18)	43(26)	60(35)	148(22)
Subject Choice (subject withdrew consent)	2(1)	1(1)	3(2)	2(1)	8(1)
Lost to Follow-up	0	0	1(1)	1(1)	2(<1)
Adverse Event	7(4)	21(13)	31(18)	52(30)	111(17)
Lack of Efficacy ^b	6(4)	2(1)	2(1)	2(1)	12(2)
Other	2(1)	4(3)	6(4)	3(2)	15(2)

Note: Percentages calculated with the number of subjects in each group as denominator.

^a Completion and discontinuation information was based on the study termination eCRF page.

^b Lack of efficacy is defined as use of rescue medication during the double-blind period.

Among the 666 subjects in the safety analysis set, seven were excluded from the ITT population. The Clinical Study Report did not provide an explanation for why these subjects were excluded. DAARP sent the Applicant an information request on 10 July 2008, and received a response which read in part:

The definition of the intent-to-treat analysis set was those subjects who were randomized, received at least one dose of study medication following randomization and had a valid baseline pain intensity score. The valid baseline pain intensity score was pre-defined in the statistical analysis plan (SAP). According to the SAP (section 2.1.2, page 8 and section 2.2.11.4.1, page 17), the baseline pain intensity score was defined as the average of the last 3 consecutive days that pain intensity was collected during the run-in period, given that at least 5 pain assessments were available during the last 3 days. If a pain assessment was recorded in the run-in diary in the morning of the day of randomization, that assessment was also taken into account for calculation of the baseline pain intensity score.

A window of up to 14 days between the last day of the run-in period and day of randomization was implemented to address variability in study visit schedules at study sites. Run-in diaries that were completed more than 14 days prior to the day of randomization were considered too remote to use in calculation of the baseline score. Therefore, if the last day of the run-in period was more than 14 days prior to the day of randomization, a baseline pain intensity score was not calculated based on those measurements.

There were 7 subjects who were randomized, received study medication but who were considered not having a valid baseline pain intensity score and were excluded from the ITT analysis set. Of these 7 subjects, 4 had insufficient run-in period measurements to calculate a baseline pain intensity score, and 3 had run-in period measurements that exceeded the 14-day window.

The Applicant goes on to describe the baseline assessments available from specific subjects.

There was some ambiguity in the SAP as to how the Run-in period was defined. On page 6 of the SAP (p. 1152 in CSR), the period was defined as “Day -7 through -1”. On page 7 of the SAP, under the heading “Relative Study Day”, the following was stated:

The first double-blind dose date will be used as the reference start date in computing relative study days.

- Relative days for event on or after the first dose = event date – first dose date + 1; and
- Relative days for event prior to the first dose = event date – first dose date.

As an illustration, consider a subject who received the first double-blind dose of medication on January 15. Based on the information on pages 6 and 7, the Run-in period would be January 8 through January 14. (This definition is also suggested by Figure 3.) On page 8, however, the following is stated,

At an eCRF meeting on 9/25/06, a decision was made by the team to assume the start of the PI diary as the start of the run-in period.

This contradicts the earlier definition of the Run-in period, making it a variable window that does not begin on a specific Study Day. The response from the Applicant introduces yet another definition of the Run-in period, requiring it to end between Day -14 and day -1. The 14 day rule was not in the Statistical Analysis Plan. Although the 14 day rule may have been introduced post-hoc, a mitigating circumstance is that it only affected three subjects, less than .5% of the safety set.

The fact that the Applicant did not use the fixed Run-in period had substantive implications for interpretation of the study. This can be made clearer by considering this inclusion criterion (p. 31 of CSR, emphasis added):

Before randomization on Day 1, *pain was not adequately controlled with the current stable analgesic regimen* based on the following criteria:

- Mean pain intensity was equal to or greater than 5 (after rounding 4.5 and above to an integer) on an 11-point (0 to 10) numerical rating scale (NRS) during the last 3 days of pain assessments during the run-in period.
- Minimum single-assessment pain intensity score was equal to or greater than 3 on an 11-point (0 to 10) NRS during the last 3 days of pain assessments during the run-in period.

The meaning of this criterion depends on how the run-in period is defined. If the period ends on Day -1, then it requires that the pain be out of control *in the 3 days before randomization*. With the floating window, on the other hand, it is possible that the out of control pain was observed more than a week before randomization. I performed a subgroup analysis (Section 4) to confirm that the efficacy finding did not depend on using a floating window.

Baseline characteristics of the ITT population are shown in Table 9. It was provided by the Applicant, but I verified their findings.

APPEARS THIS WAY
ON ORIGINAL

Table 9: Applicant's Demographic and Baseline Characteristics

(Source: Table 9 in Clinical Study Report)

	Placebo (N=169)	Tapentadol IR 50 mg (N=153)	Tapentadol IR 75 mg (N=166)	Oxycodone HCl IR 10 mg (N=171)	Total (N=639)
Sex, n (%)					
N	169	153	166	171	659
Male	80 (47)	79 (52)	88 (53)	88 (51)	335 (51)
Female	89 (53)	74 (48)	78 (47)	83 (49)	324 (49)
Racial/ethnic Group, n (%)					
N	169	153	166	171	659
White	158 (93)	138 (90)	148 (89)	156 (91)	600 (91)
Black	9 (5)	5 (3)	6 (4)	10 (6)	30 (5)
Hispanic	0	5 (3)	7 (4)	3 (2)	15 (2)
Other	2 (1)	5 (3)	5 (3)	2 (1)	14 (2)
Age (Years)					
N	169	153	166	171	659
Category, n (%)					
<65	104 (62)	91 (59)	103 (62)	101 (59)	399 (61)
≥65	65 (38)	62 (41)	63 (38)	70 (41)	260 (39)
Mean (SD)	61.3 (10.08)	60.6 (10.16)	60.8 (10.04)	62.1 (9.05)	61.2 (9.83)
Median	62.0	60.0	61.5	62.0	62.0
Range	(20;79)	(31;79)	(34;78)	(41;79)	(20;79)
Weight (kg)					
N	169	153	166	171	659
Mean (SD)	98.4 (24.96)	96.4 (25.02)	97.2 (22.03)	96.1 (22.95)	97.0 (23.71)
Median	93.8	92.5	95.1	93.0	93.4
Range	(48;175)	(54;200)	(54;181)	(54;181)	(48;200)
Body Mass Index (kg/m²)					
N	168	151	166	171	656
Mean (SD)	33.8 (7.71)	33.0 (8.02)	33.6 (7.79)	33.2 (6.86)	33.4 (7.58)
Median	32.1	31.2	32.8	32.2	32.0
Range	(19;60)	(21;76)	(21;64)	(20;52)	(19;76)
Pain Intensity Score					
N	169	153	166	171	659
Category, n (%)					
Moderate	48 (28)	43 (28)	52 (31)	60 (35)	203 (31)
Severe	121 (72)	110 (72)	114 (69)	111 (65)	456 (69)

Statistical Methodologies

As with the Bunionectomy Study, the primary efficacy analysis was an ANCOVA with the factors of treatment and pooled center, and baseline PI as a covariate. The primary imputation method was LOCF and the Hochberg step-up procedure was used to control for multiplicity. Aside from assessing the SPID over five days instead of 48 hours, the efficacy primary analysis was essentially the same one used in the Bunionectomy Study.

Analysis of the secondary efficacy variables also paralleled the methods used in the Bunionectomy Study. Major differences were that pain was assessed at different time points and that time to onset of analgesia was not assessed. In a deviation from the Statistical Analysis Plan, the Applicant used the log-rank test as well as the planned Wilcoxon test to compare responder curves.

As with the Bunionectomy Study, there was a sequential testing procedure for selected secondary variables. It is described as follows in the Statistical Analysis Plan (p. 1157 in Clinical Study Report):

The following selected secondary hypotheses will be tested in a sequential manner if the null hypothesis for the primary endpoint is rejected. At each step, when the preceding null hypothesis for a particular treatment fails to be rejected, further comparisons with that treatment will not be performed.

- a. Test the superiority of CG5503 base IR 50mg against oxycodone 10mg in terms of the composite event of nausea and vomiting rate at a two-sided 0.05 significance level.
- b. If CG5503 base IR 50mg is shown to be superior to the oxycodone group in composite event of nausea and vomiting rate, compare the non-inferiority of CG5503 base IR 75mg against oxycodone 10mg at a one-sided 0.025 significance level with 5-day SPID. The NI margin will be 10% of the entire range of the primary endpoint (i.e. 120 out of 1200 points).
- c. If CG5503 base IR 75mg is shown to be non-inferior to the oxycodone group in 5-day SPID, test the superiority of CG5503 base IR 75mg against oxycodone 10mg in terms of composite event of nausea and vomiting rate at a two-sided 0.05 significance level.
- d. If CG5503 base IR 75mg is shown to be superior to oxycodone 10mg in composite event of nausea and vomiting, compare the non-inferiority of CG5503 base IR 50mg against oxycodone 10mg at a one-sided 0.025 significance level with 5-day SPID. The NI margin will be 10% of the entire range of the primary endpoint.
- e. If CG5503 base IR 50mg is shown to be non-inferior to the oxycodone group in 5-day SPID, test the superiority of CG5503 base IR 50mg to oxycodone 10mg in the treatment emergent adverse event rate of constipation at a two-sided 0.05 significance level.
- f. If CG5503 base IR 50mg is shown to be superior to oxycodone 10mg in adverse event rate of constipation, test the superiority of CG5503 base IR 75mg to oxycodone 10mg in the treatment emergent adverse event rate of constipation at a two-sided 0.05 significance level;
- g. If CG5503 base IR 75mg is shown to be superior to oxycodone 10mg in adverse event rate of constipation, compare the CG5503 IR dose groups with the placebo group in terms of the "Time to First Rescue Medication" using the Hochberg's procedure at the 0.05 significance level.

The comparisons between tapentadol and oxycodone on nausea, vomiting, and constipation are not in the label. Time to rescue analgesia is not mentioned in the label, either.

Results and Conclusions

Primary Efficacy Endpoint

I replicated the Applicant's finding that both doses of tapentadol were superior to placebo on the Five Day SPID using LOCF imputation, as seen in Table 10. As Table 11 shows, I found also found superiority for both tapentadol doses when BOCF and WOCF imputation were used.

Table 10: Applicant's Descriptive Statistics and Pairwise Comparison of Day Five SPID, LOCF Imputation

	Placebo (N=169) n (%)	Tapentadol IR 50 mg (N=153) n (%)	Tapentadol IR 75 mg (N=166) n (%)	Oxycodone HCl IR 10 mg (N=171) n (%)
Day 1-5				
N	169	153	166	171
Mean (SD)	130.6 (182.77)	229.2 (228.92)	223.8 (217.76)	236.5 (222.82)
Median	86.6	164.1	210.2	206.7
(Range)	(-358;695)	(-480;881)	(-308;823)	(-268;884)
LS Means (diff from placebo)	--	101.2	97.5	111.9
95% CI	--	54.58 to 147.89	51.81 to 143.26	66.49 to 157.38
Raw p-value	--	<0.001	<0.001	<0.001
Adjusted p-value using Hochberg	--	<0.001	<0.001	--

The summary and analysis are based on the LOCF imputation method. Higher value in SPID indicates greater pain relief.

Table 11: Descriptive Statistics and Pairwise Comparison of SPID at Day Five, BOCF and WOCF Imputation

(Note: p-values not adjusted for multiplicity.)

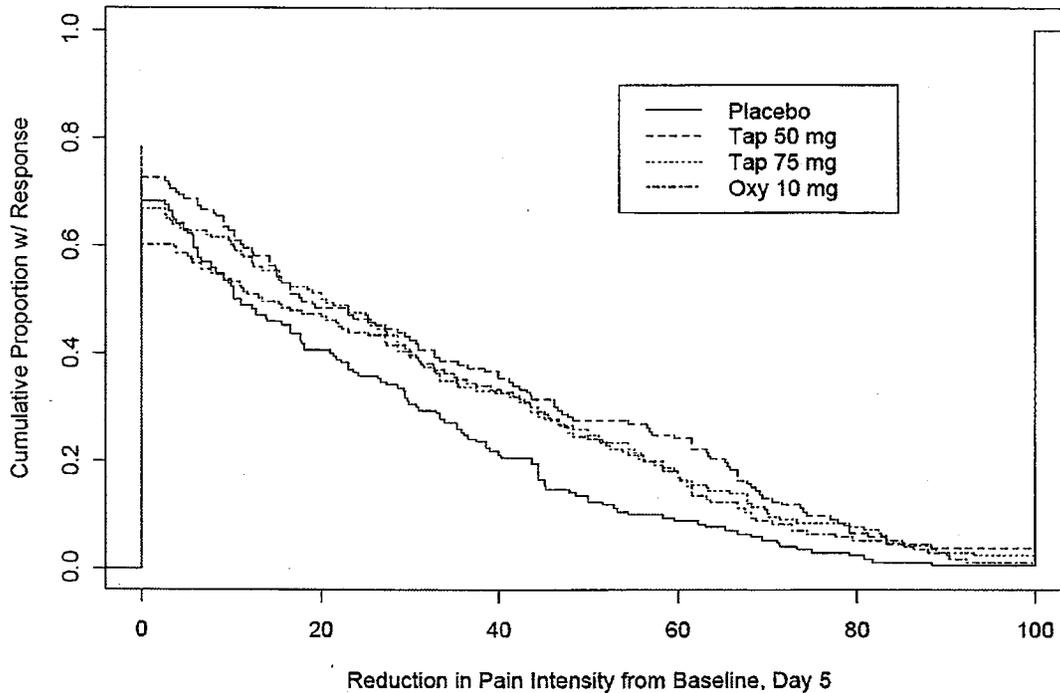
Imputation	Statistic	Placebo	Tap. 50 mg	Tap. 75 mg	Oxy. 10 mg
BOCF	Mean	131.1	219.5	207.2	202.3
	Standard Deviation	180.1	221.6	206.2	204.3
	LS Mean (diff from placebo)	--	92.1	80.1	78.9
	p-value (unadj.)	--	<.0001	.0003	.0003
WOCF	Mean	128.2	214.0	201.3	194.8
	Standard Deviation	184.1	230.2	214.5	214.0
	LS Mean (diff from placebo)	--	89.9	77.7	74.6
	p-value (unadj.)	--	.0001	.0007	.0010

Responder Analysis

Figure 4 shows the cumulative proportion of subjects who showed different percentages of reduction in pain intensity from baseline at Day Five. It is my replication of the Applicant's Figure 2 in the proposed label. Subjects who used rescue medication or withdrew from the study by Day Five were given a value of zero. Comparing the responder curve from this trial to that from the bunionectomy study (Figure 2 in this report), there is a less clear separation from placebo. Using the pre-specified Wilcoxon test, the tapentadol 50 mg dose was the only active treatment whose responder curve was significantly different from placebo. The log-rank test was proposed by the Applicant post-hoc, however, and it shows all active treatments different from

placebo. I also verified the Applicant's findings that both doses of tapentadol were more likely to achieve both 30% and 50% response than placebo (unadjusted $p < .05$ in all cases).

Figure 4: Cumulative Distribution of Responders, Day Five



Tapentadol vs. Oxycodone

I replicated the Applicant's results from the sequential testing on page 26 through step d.

... Tapentadol 50 mg had a lower rate of nausea/vomiting than oxycodone 10 mg ($p < .001$). The next step was to compare tapentadol 75 mg with oxycodone for non-inferiority on the 5-day SPID. The 95% lower bound for the difference of SPIDs was -59.9, within the pre-specified margin of 120. Having shown non-inferiority, the next step was to show that tapentadol 75 mg was superior to oxycodone 10 mg on nausea/vomiting ($p < .001$). Finally comparing tapentadol 50 mg with oxycodone on the five day SPID, the lower bound for the difference was -57.3, again within the NI margin.

b(4)

The Sponsor's non-inferiority margin of 120 was excessively wide, however. This can be seen by comparing the NI margin to the observed effects of treatment on the primary outcome. For example, the effect of oxycodone 10 mg (an approved analgesic) on the Five Day SPID was 111.9, within the NI margin. (This was the estimated effect using LOCF imputation; the other imputation methods yielded a smaller effect.) In fact, the point estimates of the effects from *all* of the active treatments in this study were within the Sponsor's margin.

3.2 Evaluation of Safety

The safety of tapentadol was reviewed by Ellen Fields, M.D.

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Race and Age

KF5503/32 (Bunionectomy Study)

Tables 12-14 show descriptive statistics for the primary efficacy outcome in ITT population by gender, race, and age group. The imputation method used in these tables and in the Clinical Study Report is LOCF, but I obtained qualitatively similar results using BOCF. My LOCF results match the Applicant's. The study population was predominantly non-elderly women, so it is difficult to draw conclusions about the impact of gender and geriatric status. There is little indication that black, white, and Hispanic patients had substantially different outcomes.

Table 12: Descriptive Statistics for SPID-48 (LOCF), by Gender

Gender	Treatment	N Obs	Mean	Std Dev
FEMALE	Placebo	108	24.9	120.7
	Tapentadol 50 mg	101	129.1	122.1
	Tapentadol 75 mg	107	139.1	115.5
	Tapentadol 100 mg	99	163.7	100.2
	Oxycodone 15 mg	110	185.7	102.3
MALE	Placebo	12	21.3	128.8
	Tapentadol 50 mg	18	62.7	135.4
	Tapentadol 75 mg	13	139.1	150.0
	Tapentadol 100 mg	19	185.6	92.7
	Oxycodone 15 mg	15	74.1	125.2

APPEARS THIS WAY
ON ORIGINAL

Table 13: Descriptive Statistics for SPID-48 (LOCF), by Race

Race	Treatment	N Obs	Mean	Std Dev
White	Placebo	68	22.1	117.2
	Tapentadol 50 mg	56	104.6	133.3
	Tapentadol 75 mg	71	133.4	120.0
	Tapentadol 100 mg	62	153.8	96.0
	Oxycodone 15 mg	76	163.3	105.7
Black	Placebo	23	15.3	140.2
	Tapentadol 50 mg	27	127.1	109.4
	Tapentadol 75 mg	19	145.6	122.2
	Tapentadol 100 mg	24	192.6	107.2
	Oxycodone 15 mg	25	169.9	134.8
Hispanic	Placebo	26	33.4	122.0
	Tapentadol 50 mg	32	130.1	129.3
	Tapentadol 75 mg	24	181.9	100.7
	Tapentadol 100 mg	30	164.8	91.7
	Oxycodone 15 mg	23	200.3	99.2
Other	Placebo	3	72.7	47.6
	Tapentadol 50 mg	4	180.5	101.5
	Tapentadol 75 mg	6	15.2	81.7
	Tapentadol 100 mg	2	315.2	73.6
	Oxycodone 15 mg	1	265.0	.

**APPEARS THIS WAY
ON ORIGINAL**

Table 14: Descriptive Statistics for SPID-48 (LOCF), by Age

Age Group (years)	Treatment	N Obs	Mean	Std Dev
<65	Placebo	111	20.5	122.4
	Tapentadol 50 mg	113	118.5	128.6
	Tapentadol 75 mg	114	134.8	118.1
	Tapentadol 100 mg	111	165.8	99.3
	Oxycodone 15 mg	119	170.1	111.9
≥65	Placebo	9	73.8	92.5
	Tapentadol 50 mg	6	130.5	55.1
	Tapentadol 75 mg	6	221.8	112.1
	Tapentadol 100 mg	7	189.2	98.1
	Oxycodone 15 mg	6	215.4	84.0

KF5503/33 (End-Stage Joint Disease Study)

Tables 15-17 show descriptive statistics for the primary efficacy outcome in the ITT population by gender, race, and age group. LOCF imputation is again used in the tables, following what was reported by the Applicant, but I obtained qualitatively similar results using BOCF. My LOCF results match the Applicant's. There is some indication of a gender difference at the Tapentadol 50 mg dose, but no apparent pattern across doses. No age effect is apparent, and there are inadequate sample sizes to assess racial differences.

Table 15: Descriptive Statistics for Five Day SPID (LOCF), by Gender

Gender	Treatment	N Obs	Mean	Std Dev
FEMALE	Placebo	89	138.9	175.2
	Tapentadol 50 mg	74	265.3	247.3
	Tapentadol 75 mg	78	227.2	218.5
	Oxycodone 10 mg	83	223.1	242.7
MALE	Placebo	80	121.4	191.5
	Tapentadol 50 mg	79	195.3	206.2
	Tapentadol 75 mg	88	220.7	218.3
	Oxycodone 10 mg	88	249.1	202.9

Table 16: Descriptive Statistics for Five Day SPID (LOCF), by Race

Racial/Ethnic Group Code	Treatment Group Code	N Obs	Mean	Std Dev
White	Placebo	158	129.4	176.7
	Tapentadol 50 mg	138	237.3	232.6
	Tapentadol 75 mg	148	227.2	221.4
	Oxycodone 10 mg	156	236.2	224.6
Black	Placebo	9	162.5	291.9
	Tapentadol 50 mg	5	91.0	105.3
	Tapentadol 75 mg	6	191.2	191.1
	Oxycodone 10 mg	10	260.9	191.3
Hispanic	Tapentadol 50 mg	5	166.9	235.3
	Tapentadol 75 mg	7	270.6	190.8
	Oxycodone 10 mg	3	294.6	281.1
Other	Placebo	2	86.0	121.6
	Tapentadol 50 mg	5	205.3	201.8
	Tapentadol 75 mg	5	96.3	170.1
	Oxycodone 10 mg	2	50.4	217.2

Table 17: Descriptive Statistics for Five Day SPID (LOCF), by Age

Age Group (years)	Treatment	N Obs	Mean	Std Dev
<65	Placebo	104	131.4	189.5
	Tapentadol 50 mg	91	240.1	215.8
	Tapentadol 75 mg	103	222.2	215.5
	Oxycodone 10 mg	101	256.6	231.7
≥65	Placebo	65	129.3	172.9
	Tapentadol 50 mg	62	213.2	247.9
	Tapentadol 75 mg	63	226.3	223.1
	Oxycodone 10 mg	70	207.6	207.6

4.2 Other Special/Subgroup Populations

KF5503/32 (Bunionectomy Study)

In response to a request from DAARP, the Applicant assessed the outcomes for the subgroup of subjects who “reloaded”, *i.e.*, took the second dose of study medication less than three hours after the first dose. Table 18 shows the descriptive statistics for the primary efficacy variable stratified by reloading. With the exception of the 100 mg group, the reloaders tended to show worse pain outcomes (lower SPIDs). The largest difference can be found in the placebo group, where subjects who reloaded had a mean SPID₄₈ of -2 compared to a value of 67.2 for those who didn’t reload. These results make intuitive sense. One would expect the reloaders to generally be those patients who are getting the least effect from their first dose of study medication, with those in the placebo group getting no biological effect at all.

DAARP had requested evidence from the Sponsor that the “reload” dose was effective, and later (more specifically) a subgroup analysis for “reloaders”. While the Case Study Report does not include any inferential statistics to this purpose, the Sponsor’s descriptive statistics (which are replicated in the table) do suggest overall efficacy in the “reload” group. As a caveat, this variable does not address the efficacy of the “reload” dose specifically, but that of the whole 48 hours of treatment including the “reload” dose.

Table 18: Descriptive Statistics for SPID 48 by Reload (2nd dose < 3 hrs), LOCF Imputation

Reloaded?	Statistic	Placebo	Tap. 50 mg	Tap. 75 mg	Tap. 100 mg	Oxy. 15 mg
Yes	N	74	60	56	55	54
	Mean	-2.0	109.4	120.6	171.3	147.2
	SD	122.8	131.9	124.0	109.2	115.0
No	N	46	59	64	63	71
	Mean	67.2	129.0	155.3	163.7	191.3
	SD	105.9	119.7	112.8	89.9	104.4

KF5503/33 (End-Stage Joint Disease Study)

It was noted on page 23 that subjects in the End-Stage Joint Disease Study could have had their baseline pain level assessed as long as 14 days before receiving their first dose of study drug. In principle, a subject’s pain level could have decreased in the intervening time period. Since the SPID is computed relative to the baseline pain level, any such decrease in pain would increase the SPID and could potentially show a false effect of treatment.

In order to assess whether the results were susceptible to this artifact, I analyzed the primary efficacy variable for the 601 subjects who had their last baseline pain measurement within a day of randomization.

Table 19 shows the descriptive statistics of the five day SPID for these subjects, computed using LOCF imputation. I also replicated the primary efficacy analysis in this group, finding that all three active treatments were superior to placebo (unadjusted $p < .001$).

Table 19: Descriptive Statistics of Day Five SPID, Subjects with No Gap after Baseline

Statistic	Placebo	Tap. 50 mg	Tap. 75 mg	Oxy. 15 mg
N	152	141	152	156
Mean (SD)	136.9 (182.7)	235.1 (232.2)	217.5 (212.3)	239.0 (227.0)
Median	101.5	188.5	202.2	209.0

5. SUMMARY AND CONCLUSIONS

5.1 Statistical Issues and Collective Evidence

The application presented several procedural and statistical issues which were relevant to the efficacy of the product. First of all, DSI found irregularities at the two of clinical sites used in the Bunionectomy Study. DSI concluded that the data from the sites are still acceptable, however, and my own analysis showed that they results for the primary endpoint were fairly consistent across sites (Table 7). A second concern was that the definition of the Run-In Period used in the ESJD study was apparently modified post-hoc. Using a strict definition that required the period to end immediately before randomization, however, did not alter the overall finding of efficacy. A third issue is that Applicant used LOCF imputation for the primary efficacy analysis, rather than a more conservative method, but the results were ultimately not dependent on the imputation method used.

┌

b(4)

└

4 Page(s) Withheld

 Trade Secret / Confidential (b4)

✓ Draft Labeling (b4)

 Draft Labeling (b5)

 Deliberative Process (b5)

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

/s/

Jonathan D Norton
10/3/2008 02:20:38 PM
BIOMETRICS

Dionne Price
10/3/2008 02:26:16 PM
BIOMETRICS
concur

Thomas Permutt
10/3/2008 02:28:13 PM
BIOMETRICS
concur