CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

022399Orig1s000

STATISTICAL REVIEW(S)

Statistical Review and Evaluation Addendum
(Carcinogenicity Studies)

Report #2
(An Addendum)

NDA Number: 22-399

Drug Name: (b) (4)™; ((±-1-([(α-isobutanoyloxyethoxy)carbonyl]-aminomethyl)-1-cyclohexane acetic acid

Sponsor: GlaxoSmithKline

Pharm/tox Reviewer: Terry S. Peters, D.V.M.
Division of Neurology Products

Project Manager: Beverly Conner, Pharm.D.
Division of Neurology Products

Statistical Reviewers: Steve Thomson
Karl K. Lin, Ph.D.
Division of Biometrics 6
Office of Biostatistics

Document Reviewed: Statistical Review and Evaluation (Carcinogenicity Studies) prepared by Dr. Karl Lin, Ph.D. 8/3/09

Introduction

Again, both the original review and this addendum were done under a very tight time schedule. Because of this restriction, the original review did not provide survival adjusted tests of tumorigenicity. At the request of the ECAC, the purpose of this addendum is to add such an analysis to this submission.

Survival Analysis

First, the statistical significances of the tests of differences in survival across treatment groups are given in Table A.1 below. Tests of homogeneity over all groups, dose related trend and the pairwise differences between the high dose group and the vehicle control were performed. Two main test statistics are provided, the log rank test and the so-called Wilcoxon test. The log rank tests puts equal weight on all events being assessed, while the Wilcoxon test weights them by the square of their rank in time, and thus places more weight on later events than does the log rank test. So the Wilcoxon test will generally be more sensitive to later separation of mortality than will be the log rank test. Kaplan Meier survival curves for survival as a function of dose were provided in the original report. In both genders in rats, the tests of no overall homogeneity, no trend over dose, and no difference between the high dose group and control were all highly statistically significant (all $p \leq 0.0001$). From the survival curves, in male rats there is a generally increasing trend in mortality in dose. In female rats, the high dose group and the medium dose group have survival curves that are generally intertwined, as were the curves of the low dose group and control, but there was still a generally increasing trend in mortality in dose. Again from the survival curves, in male mice the high dose group has a generally higher mortality than the remaining dose groups. The medium and low dose groups generally intertwined, but with generally higher mortality than the control group. Again all the tests noted above in male mice were statistically significant (all $p \leq 0.0047$). In female mice the survival curves are generally intertwined with no particular evidence of differences in survival (all $p \geq 0.2987$). Absence of proof is not proof of absence, but, as noted in the report, the lack of evidence for differences should indicate that the non-mortality adjusted tests in female mice are appropriate.

Table A.1 Statistical Significances of Tests of Homogeneity and Trend in Survival

Rats	Males		Females	
	Log Rank	Wilcoxon	Log Rank	Wilcoxon
Homogeneity over Groups 1-4	<0.0001	<0.0001	<0.0001	<0.0001
Trend over Groups 1-4	<0.0001	<0.0001	<0.0001	0.0001
Comparison of High and Low	<0.0001	<0.0001	<0.0001	<0.0001

Mice	Males		Females	
	Log Rank	Wilcoxon	Log Rank	Wilcoxon
Homogeneity over Groups 1-4	0.0027	0.0047	0.7492	0.7931
Trend over Groups 1-4	0.0003	0.0008	0.3727	0.4526
Comparison of High and Low	0.0002	0.0003	0.2987	0.3523

Tumorigenicity

Again, the purpose of this addendum is to provide mortality adjusted tests of carcinogenicity. The consensus of the Society of Toxicological Pathology town hall meeting in June 2001 seemed to be that the poly-k modification of the Cochran-Armitage test of trend for tumor incidence should be used. That is the analysis provided in this addendum. Note that because of the software used, analysis over all tumors was only slightly more demanding than the analysis of subsets of the tumors. The results of tests of trend, and the pairwise comparisons of each treatment group to the controls are presented in the tables below. As discussed in the report, if one is determined to control statistical error, the results of the test of trend are recommended. Even if one includes tests of differences between the high dose and control, note that including the results of the pairwise tests between the medium and low dose groups can be expected to inflate Type I error, perhaps considerably. Nonetheless, in case they are of actual interest they are included below.

Tables A.2 and A.3 below display the results of any test that is potentially statistically significant at a nominal 0.05 level. Complete results are presented in the appendix to this addendum. Note that testing carcinogens involves a large number of tests. Based on his extensive experience with such analyses, for pairwise tests between the high dose group and controls in two species, Haseman (1983) claimed that for a roughly 0.10 (10%) overall false positive error rate, rare tumors should be tested at a 0.05 (5%) level, and common tumors (with a historical control incidence greater than 1%) at a 0.01 level. For a standard chronic study in two species (i.e., mice and mice) study, based on simulations and their experience, Lin & Rahman (1998) proposed a further p-value adjustment for tests of trend. That is, for a roughly 0.10 (10%) overall false positive error rate in tests of trend, rare tumors should be tested at a 0.025 (2.5%) level and common tumors at a 0.005 (0.5%) level. In this analysis, we will use the observed incidence in the vehicle control group to decide if a tumor is rare or common when applying these rules for multiplicity adjustment. Note the discussion in the report justifies emphasis on the tests of trend.

In Table A.1 below, using incidence in the control group to determine whether the tumor is rare or common, in rats both acinar cell benign adenoma and combined adenoma and carcinoma would be classified as common in male rats and rare in female rats. In both genders, tests of trend in acinar cell adenoma would be statistically significant (Males: $p = 0.0009 < 0.005$, Females: $p = 0.0105 < 0.025$). Similarly tests of no trend in pooled acinar cell adenoma and carcinoma are also statistically significant in both genders (Males: $p = 0.0002 < 0.005$, Females: $p = 0.0022 < 0.025$). In male rats, tests of comparisons between the high dose group and controls of both acinar cell benign adenoma and combined adenoma and carcinoma were statistically significant ($p=0.0029$ and $p=0.001$ both < 0.01). In female rats only the test of no differences between the high dose group and controls of combined adenoma and carcinoma was statistically significant ($p=0.0317 < 0.05$), although the comparison in adenoma was close to significance. Similarly, in female rats, the test of trend and differences between the high dose and controls in benign granular tumors of the uterus were quite close to these somewhat arbitrary bounds to determine statistical significance (i.e., $p = 0.0051 > 0.005$ and

p=0.0106 > 0.01, respectively). No other comparison achieved statistical significance when using the Haseman-Lin-Rahman adjustments for multiplicity cited above.

Table A.2 Potentially Significant Results (p≤ 0.05) in Rats

organ / tumor	Cntrl	Low	Med	High	Trend	CvsL	CvsM	CvsHi
Male Rats								
pancreas								
Acinar adenoma+carcinoma	2	4	5	9	0.0002	0.3629	0.1535	0.0010
adenoma, acinar cell, benign	2	4	4	8	0.0009	0.3629	0.2498	0.0029
adenoma, islet cell, benign	3	5	1	0	0.0365	0.3707	0.3835	0.2697
small intestine, jejunum								
adenocarcinoma, malignant	0	1	0	2	0.0465	0.5067	1.0000	0.1248
Female Rats								
mammary gland								
adenocarcinoma, malignant	5	1	10	2	0.4807	0.1276	0.0438	0.3558
fibroadenoma, benign	18	17	17	23	0.0140	0.5219	0.2973	0.0339
pancreas								
Acinar adenoma+carcinoma	0	0	0	4	0.0022	1.0000	1.0000	0.0317
adenoma, acinar cell, benign	0	0	0	3	0.0105	1.0000	1.0000	0.0769
skin, subcutis								
fibroma, benign	4	3	0	0	0.0091	0.5470	0.1085	0.0988
uterus with cervix								
granular cell tumor, benign	1	3	3	7	0.0051	0.2776	0.1959	0.0106
polyp, stromal, benign	11	8	2	3	0.0233	0.3782	0.0317	0.0698
vagina								
granular cell tumor, benign	2	2	2	5	0.0409	0.6593	0.5593	0.1206

Using the Haseman-Lin-Rahman rules, in mice only malignant fibrosarcoma of subcutis skin in female mice is even close to statistical significance (p=0.0257>0.025). No other tests achieved multiplicity adjusted statistical significance.

Table A.3 Potentially Significant Results (p≤ 0.05) in Mice

organ / tumor	Cntrl	Low	Med	High	Trend	CvsL	CvsM	CvsHi
Male Mice								
harderian glands								
adenoma, benign	5	0	1	1	0.1328	0.0308	0.1313	0.1154
Female Mice								
skeletal muscle, biceps femoris								
sarcoma, undifferentiated, malignant	2	0	0	0	0.0491	0.1880	0.2177	0.2353
skin								
sarcoma, undifferentiated, malignant	2	0	0	0	0.0460	0.1792	0.2081	0.2253
skin, subcutis								
fibrosarcoma, malignant	0	0	3	3	0.0257	1.0000	0.1637	0.1405

Note that complete incidence tables are included in the appendix below.

Appendix AA.1 Complete Tumor Incidence Tables

Tables AA.1-AA.4, below, present complete incidence tables and survival adjusted analyses for all tumors given in the Sponsor's data sets. Due to time constraints these reflect the exact breakdowns of organs and tumors as provided by the sponsor. As noted above, very fine breakdowns of organs or tumors, e.g. nose levels a to d, result in very few tumors for each such breakdown. This, in turn, makes it difficult to determine if any treatment differences are statistically significant, whether one uses p-values or uses posterior probabilities.

Table AA.1 Tumor Incidence and Tests in Male Rats

organ / tumor	Cntrl	Low	Med	High	Trend	CvsL	CvsM	CvsHi
adrenal glands								
adenoma, cortical, benign	1	2	2	0	0.3415	0.5101	0.4273	0.6515
pheochromocytoma, benign	7	3	5	2	0.3540	0.1436	0.5638	0.3256
pheochromocytoma, complex, benign	1	0	0	0	0.2945	0.4933	0.5616	0.6515
bone, mandible								
odontoma, benign	1	0	0	0	0.2945	0.4933	0.5556	0.6515
brain								
astrocytoma, benign	0	1	1	1	0.1669	0.5067	0.4521	0.3485
granular cell tumor, benign	1	0	0	0	0.2945	0.4933	0.5556	0.6515
meningioma, malignant	1	0	0	0	0.2945	0.4933	0.5616	0.6515
cavity, abdominal								
fibrosarcoma, malignant	0	0	1	0	0.4110	.	0.4521	.
cavity, oral								
carcinoma, squamous cell,malignant	1	1	2	0	0.4263	0.7400	0.4273	0.6515
papilloma, squamous cell, benign	1	0	0	0	0.2945	0.4933	0.5556	0.6515
coagulating glands								
carcinoma, squamous cell,malignant	0	0	1	0	0.4110	.	0.4444	.
epididymides								
mesothelioma, malignant	0	1	0	0	0.5890	0.5067	.	.
eyes								
leiomyoma, benign	1	0	0	0	0.2945	0.4933	0.5616	0.6515
schwannoma, malignant	1	0	0	0	0.2945	0.4933	0.5556	0.6515
harderian glands								
adenoma, benign	1	0	2	0	0.5240	0.4933	0.4273	0.6515
heart								
neuroendocrine tumor, benign	0	1	0	0	0.5890	0.5132	.	.
schwannoma, benign	0	1	0	0	0.5890	0.5132	.	.
kidneys								
carcinoma, tubular cell, malignant	0	1	0	1	0.1977	0.5067	.	0.3485
lipoma, benign	1	0	0	1	0.3008	0.4933	0.5556	0.5916
liposarcoma, malignant	1	0	0	0	0.2945	0.4933	0.5556	0.6515
lacrimal glands, exorbital								
carcinoma, zymbals gland,malignant	0	0	1	1	0.1043	.	0.4521	0.3485
large intestine, rectum								
adenocarcinoma, malignant	0	0	0	1	0.1633	.	.	0.3582
liver								
adenoma, hepatocellular, benign	0	0	1	1	0.1043	.	0.4444	0.3485
carcinoma, hepatocellular,malig.	0	0	0	1	0.1575	.	.	0.3485
lymph node, mesenteric								
schwannoma, malignant	1	0	0	0	0.2945	0.4933	0.5616	0.6515
mammary gland								
fibroadenoma, benign	0	1	2	1	0.1631	0.5067	0.2009	0.3485
multicentric neoplasm								
hemangioma, benign	3	1	0	1	0.3811	0.2973	0.1657	0.5651
hemangiosarcoma, malignant	2	4	1	1	0.3396	0.3501	0.5842	0.7236
lymphoma, malignant	0	2	3	0	0.5256	0.2600	0.0877	.
sarcoma, histiocytic, malignant	1	1	1	0	0.4018	0.7400	0.7032	0.6515

Table AA.1 (cont.) Tumor Incidence and Tests in Male Rats

organ / tumor	Cntrl	Low	Med	High	Trend	CvsL	CvsM	CvsHi
nose, level a								
carcinoma, squamous cell, malign.	0	1	0	0	0.5890	0.5132	.	.
sarcoma, undifferentiated, malign.	1	0	0	0	0.2945	0.4933	0.5556	0.6515
nose, level b								
carcinoma, squamous cell, malign.	0	1	0	0	0.5890	0.5132	.	.
sarcoma, undifferentiated, malign.	1	0	0	0	0.2945	0.4933	0.5556	0.6515
schwannoma, malignant	1	0	0	0	0.2945	0.4933	0.5556	0.6515
nose, level c								
carcinoma, squamous cell, malign.	0	1	2	0	0.4760	0.5132	0.2009	.
chondroma, benign	0	1	0	0	0.5890	0.5132	.	.
fibrosarcoma, malignant	1	0	0	0	0.2945	0.4933	0.5616	0.6515
schwannoma, malignant	1	0	0	0	0.2945	0.4933	0.5556	0.6515
nose, level d								
carcinoma, squamous cell, malig.	1	1	2	0	0.4263	0.7400	0.4273	0.6515
fibrosarcoma, malignant	1	0	0	0	0.2945	0.4933	0.5616	0.6515
schwannoma, malignant	1	0	0	0	0.2945	0.4933	0.5556	0.6515
pancreas								
Acinar adenoma+carcinoma	2	4	5	9	0.0002	0.3629	0.1535	0.0010
adenoma, acinar cell, benign	2	4	4	8	0.0009	0.3629	0.2498	0.0029
adenoma, islet cell, benign	3	5	1	0	0.0365	0.3707	0.3835	0.2697
carcinoma, acinar cell, malignant	0	0	1	1	0.1043	.	0.4521	0.3485
carcinoma, islet cell, malignant	1	0	0	0	0.2945	0.4933	0.5556	0.6515
parathyroid glands								
adenoma, benign	0	2	1	0	0.4676	0.2533	0.4521	.
pituitary gland								
adenoma, pars distalis, benign	31	30	29	20	0.4516	0.5000	0.5408	0.5468
adenoma, pars intermedia, benign	1	2	0	0	0.1788	0.5200	0.5556	0.6515
prostate gland								
adenoma, benign	1	0	1	0	0.4956	0.4933	0.7032	0.6515
seminal vesicles								
carcinoma, squamous cell, malig.	0	0	2	0	0.3541	.	0.1941	.
skin								
adenoma, basal cell, benign	0	0	0	1	0.1633	.	.	0.3582
adenoma, sebaceous cell, benign	0	0	0	1	0.1575	.	.	0.3485
carcinoma, squamous cell, malign.	1	2	0	0	0.1771	0.5200	0.5616	0.6515
keratoacanthoma, benign	3	2	0	1	0.3027	0.4875	0.1714	0.5651
papilloma, squamous cell, benign	1	1	1	0	0.4018	0.7467	0.6948	0.6515
schwannoma, malignant	1	0	0	0	0.2945	0.4933	0.5616	0.6515
skin, subcutis								
fibroma, benign	3	5	2	0	0.0633	0.3859	0.6034	0.2697
fibrosarcoma, malignant	2	1	1	0	0.2211	0.5000	0.5819	0.4210
schwannoma, malignant	2	1	0	0	0.1004	0.4800	0.3120	0.4210
small intestine, jejunum								
adenocarcinoma, malignant	0	1	0	2	0.0465	0.5067	.	0.1248
testes								
adenoma, interstitial cell, benign	8	6	10	3	0.4724	0.3637	0.2558	0.4183
mesothelioma, malignant	0	1	0	0	0.5890	0.5067	.	.
thymus gland								
thymoma, malignant	0	0	0	1	0.1575	.	.	0.3485
thyroid gland								
adenoma, c-cell, benign	6	6	6	0	0.0519	0.5840	0.4782	0.0671
adenoma, follicular cell, benign	3	2	2	1	0.4058	0.4747	0.6177	0.5474
carcinoma, follicular cell, malig.	1	0	0	0	0.2945	0.4933	0.5556	0.6515
tongue								
carcinoma, squamous cell, malig.	0	0	1	0	0.4110	.	0.4521	.
papilloma, squamous cell, benign	0	1	0	0	0.5890	0.5132	.	.
zymbal`s gland								
carcinoma, zymbals gland, malign.	1	0	1	2	0.0708	0.4933	0.7032	0.2900

Table AA.2 Tumor Incidence and Tests in Female Rats

organ / tumor	Cntrl	Low	Med	High	Trend	CvsL	CvsM	CvsHi
adrenal glands								
adenocarcinoma, malignant	0	0	1	0	0.4386	.	0.4253	.
adenoma, cortical, benign	0	1	1	2	0.0832	0.4792	0.4186	0.1836
carcinoma (primary site unknown), malignant	0	0	1	0	0.4386	.	0.4253	.
carcinoma, cortical, malignant	1	0	0	1	0.3981	0.5208	0.5814	0.6800
pheochromocytoma, benign	3	3	3	0	0.0905	0.6206	0.4954	0.1786
brain								
astrocytoma, benign	0	1	3	1	0.2484	0.4845	0.0698	0.4382
granular cell tumor, benign	1	0	2	0	0.4212	0.5208	0.3875	0.5682
meningioma, benign	1	0	0	0	0.2941	0.5208	0.5814	0.5682
oligodendroglioma, benign	1	1	0	0	0.2454	0.7314	0.5814	0.5682
cavity, abdominal								
carcinoma (primary site unknown), malignant	0	0	1	0	0.4386	.	0.4253	.
mesothelioma, malignant	0	1	0	0	0.5647	0.4792	.	.
cavity, oral								
carcinoma, squamous cell, malig.	0	0	1	0	0.4386	.	0.4253	.
cavity, thoracic								
liposarcoma, malignant	0	1	0	0	0.5647	0.4792	.	.
mesothelioma, malignant	0	1	0	0	0.5647	0.4792	.	.
harderian glands								
carcinoma, squamous cell, malign.	0	0	1	0	0.4386	.	0.4253	.
heart								
adenocarcinoma, malignant	0	0	1	0	0.4386	.	0.4253	.
kidneys								
adenocarcinoma, malignant	0	0	1	1	0.1451	.	0.4253	0.4318
carcinoma (primary site unknown), malignant	0	0	1	0	0.4386	.	0.4253	.
nephroblastoma, malignant	0	0	1	0	0.4386	.	0.4253	.
liver								
adenoma, hepatocellular, benign	1	0	0	0	0.2941	0.5208	0.5814	0.5682
lung								
adenocarcinoma, malignant	0	0	2	1	0.1554	.	0.1780	0.4318
carcinoma, squamous cell, malign.	1	0	0	0	0.2941	0.5208	0.5814	0.5682
lymph node, mesenteric								
adenocarcinoma, malignant	0	0	1	0	0.4386	.	0.4253	.
carcinoma (primary site unknown), malignant	0	0	1	0	0.4386	.	0.4253	.
mammary gland								
adenocarcinoma, malignant	5	1	10	2	0.4807	0.1276	0.0438	0.3558
adenoma, benign	2	2	1	3	0.2191	0.6593	0.6125	0.3827
fibroadenoma, benign	18	17	17	23	0.0140	0.5219	0.2973	0.0339
mesentery/peritoneum								
adenocarcinoma, malignant	0	0	1	0	0.4386	.	0.4253	.
multicentric neoplasm								
hemangioma, benign	0	0	0	1	0.2235	.	.	0.4318
hemangiosarcoma, malignant	1	0	1	0	0.4427	0.5208	0.6648	0.5682
lymphoma, malignant	2	2	2	2	0.3904	0.6593	0.5818	0.5818
sarcoma, histiocytic, malignant	0	0	1	0	0.4386	.	0.4253	.
nose, level b								
carcinoma, squamous cell, malign.	0	0	1	0	0.4386	.	0.4253	.
nose, level c								
carcinoma, squamous cell, malign.	0	0	1	0	0.4386	.	0.4253	.
nose, level d								
carcinoma, squamous cell, malign.	0	0	1	0	0.4386	.	0.4253	.

Table AA.2 (cont.) Tumor Incidence and Tests in Female Rats

organ / tumor	Cntrl	Low	Med	High	Trend	CvsL	CvsM	CvsHi
ovaries								
adenocarcinoma, malignant	0	0	1	1	0.1451	.	0.4253	0.4318
adenoma, tubulostromal, benign	0	1	0	0	0.5673	0.4845	.	.
granulosa cell tumor, benign	1	0	0	0	0.2941	0.5208	0.5814	0.5682
nephroblastoma, malignant	0	0	1	0	0.4386	.	0.4253	.
sertoli cell tumor, benign	0	1	0	0	0.5647	0.4792	.	.
sex-cord/stromal tumor, malignant	0	1	0	0	0.5647	0.4792	.	.
pancreas								
Acinar adenoma+carcinoma	0	0	0	4	0.0022	.	.	0.0317
adenocarcinoma, malignant	0	0	1	0	0.4386	.	0.4253	.
adenoma, acinar cell, benign	0	0	0	3	0.0105	.	.	0.0769
adenoma, islet cell, benign	0	2	0	1	0.3766	0.2270	.	0.4318
carcinoma (primary site unknown), malignant	0	0	1	0	0.4386	.	0.4253	.
carcinoma, acinar cell, malignant	0	0	0	1	0.2235	.	.	0.4318
carcinoma, islet cell, malignant	0	1	0	0	0.5673	0.4845	.	.
nephroblastoma, malignant	0	0	1	0	0.4386	.	0.4253	.
parathyroid glands								
adenoma, benign	1	0	1	2	0.1209	0.5208	0.6648	0.3972
pituitary gland								
adenoma, pars distalis, benign	41	39	37	35	0.4118	0.4989	0.1247	0.5138
carcinoma, pars distalis, malig.	0	1	0	0	0.5647	0.4792	.	.
salivary gland, parotid								
adenoma, benign	0	0	1	0	0.4353	.	0.4186	.
skeletal muscle, diaphragm								
carcinoma, squamous cell, malig.	1	0	0	0	0.2941	0.5208	0.5814	0.5682
skin, subcutis								
fibroma, benign	4	3	0	0	0.0091	0.5470	0.1085	0.0988
fibrosarcoma, malignant	0	0	2	0	0.4419	.	0.1780	.
schwannoma, malignant	1	0	0	0	0.2941	0.5208	0.5814	0.5682
small intestine, duodenum								
adenocarcinoma, malignant	0	0	1	0	0.4386	.	0.4253	.
adenoma, benign	1	0	0	0	0.2941	0.5208	0.5814	0.5682
carcinoma (primary site unknown), malignant	0	0	1	0	0.4386	.	0.4253	.
small intestine, ileum								
adenocarcinoma, malignant	0	0	1	0	0.4386	.	0.4253	.
leiomyosarcoma, malignant	0	0	0	1	0.2235	.	.	0.4318
small intestine, jejunum								
adenocarcinoma, malignant	1	0	0	0	0.2941	0.5208	0.5814	0.5682
spinal cord, cervical								
astrocytoma, malignant	0	0	1	0	0.4353	.	0.4186	.
spleen								
adenocarcinoma, malignant	0	0	1	1	0.1451	.	0.4253	0.4318
carcinoma (primary site unknown), malignant	0	0	1	0	0.4386	.	0.4253	.
carcinoma, squamous cell,malignant	1	0	0	0	0.2941	0.5208	0.5814	0.5682
stomach, glandular								
adenocarcinoma, malignant	0	0	1	0	0.4386	.	0.4253	.
stomach, nonglandular								
papilloma, squamous cell, benign	1	0	0	0	0.2941	0.5208	0.5814	0.5682
thymus gland								
adenocarcinoma, malignant	0	0	1	0	0.4386	.	0.4253	.
thymoma, malignant	0	0	1	0	0.4386	.	0.4253	.
thyroid gland								
adenoma, c-cell, benign	8	13	5	6	0.2430	0.1141	0.4975	0.6083
adenoma, follicular cell, benign	1	2	0	2	0.2883	0.4684	0.5814	0.3972
carcinoma, follicular cell, malig.	2	2	1	0	0.1323	0.6593	0.6224	0.3200

Table AA.2 (cont.) Tumor Incidence and Tests in Female Rats

organ / tumor	Cntrl	Low	Med	High	Trend	CvsL	CvsM	CvsHi
urinary bladder								
adenocarcinoma, malignant	0	0	1	1	0.1451	.	0.4253	0.4318
carcinoma (primary site unknown), malignant	0	0	1	0	0.4386	.	0.4253	.
sarcoma, stromal, malignant	0	0	1	0	0.4353	.	0.4186	.
uterus with cervix								
adenocarcinoma, malignant	3	0	1	2	0.3478	0.1372	0.4293	0.6292
adenoma, benign	0	1	0	0	0.5673	0.4845	.	.
carcinoma, squamous cell,malignant	1	0	0	0	0.2941	0.5208	0.5814	0.5682
granular cell tumor, benign	1	3	3	7	0.0051	0.2776	0.1959	0.0106
leiomyosarcoma, malignant	0	1	1	0	0.5580	0.4792	0.4186	.
nephroblastoma, malignant	0	0	1	0	0.4386	.	0.4253	.
polyp, stromal, benign	11	8	2	3	0.0233	0.3782	0.0317	0.0698
sarcoma, stromal, malignant	1	1	2	0	0.3182	0.7314	0.3776	0.5682
schwannoma, malignant	0	1	0	0	0.5647	0.4792	.	.
vagina								
granular cell tumor, benign	2	2	2	5	0.0409	0.6593	0.5593	0.1206
polyp, benign	0	1	0	0	0.5647	0.4792	.	.
sarcoma, stromal, malignant	0	0	1	0	0.4353	.	0.4186	.
schwannoma, malignant	0	1	0	0	0.5647	0.4792	.	.
zymbal`s gland								
adenoma, benign	1	0	0	0	0.2941	0.5208	0.5814	0.5682
carcinoma, zymbals gland,malignant	0	0	1	0	0.4386	.	0.4253	.

Table AA.3 Tumor Incidence and Tests in Male Mice

organ / tumor	Cntrl	Low	Med	High	Trend	CvsL	CvsM	CvsHi
harderian glands								
adenoma, benign	5	0	1	1	0.1328	0.0308	0.1313	0.1154
liver								
adenoma, hepatocellular, benign	0	3	1	3	0.1451	0.0952	0.4571	0.1050
carcinoma, hepatocellular, malign.	6	5	6	9	0.1213	0.5882	0.4946	0.2538
lung								
adenoma, bronchiolar alveolar,ben.	3	6	2	3	0.2888	0.1679	0.5850	0.6432
carcinoma, bronchiolar alveolar, malignant	0	0	0	1	0.2464	.	.	0.4722
carcinoma, hepatocellular, malign.	1	1	1	1	0.4947	0.7286	0.7126	0.7286
fibrosarcoma, malignant	0	0	2	0	0.4857	.	0.2159	.
lymph node, axillary								
fibrosarcoma, malignant	0	0	1	0	0.4783	.	0.4571	.
lymph node, mandibular								
fibrosarcoma, malignant	0	0	1	0	0.4783	.	0.4571	.
multicentric neoplasm								
hemangiosarcoma, malignant	1	0	1	3	0.0608	0.5278	0.7126	0.2811
lymphoma, malignant	4	3	3	1	0.1300	0.5627	0.6315	0.2245
sarcoma, histiocytic, malignant	3	0	3	1	0.3934	0.1357	0.6140	0.3455
pituitary gland								
adenoma, pars distalis, benign	0	1	0	0	0.5217	0.4722	.	.
adenoma, pars intermedia, benign	0	0	1	0	0.4783	.	0.4571	.
skin, subcutis								
fibrosarcoma, malignant	0	1	2	1	0.3201	0.4722	0.2159	0.4865
stomach, nonglandular								
papilloma, squamous cell, benign	0	0	0	1	0.2464	.	.	0.4722
testes								
adenoma, interstitial cell, benign	1	0	0	0	0.2754	0.5278	0.5429	0.5278
thyroid gland								
adenoma, follicular cell, benign	0	1	0	0	0.5217	0.4722	.	.

Table AA.4 Tumor Incidence and Tests in Female Mice

organ / tumor	Cntrl	Low	Med	High	Trend	CvsL	CvsM	CvsHi
adipose tissue, white, umbilical region								
lipoma, benign	0	1	0	0	0.5091	0.5714	.	.
adrenal glands								
carcinoma, squamous cell,malignant	0	1	0	0	0.5091	0.5714	.	.
pheochromocytoma, benign	0	0	0	1	0.2364	.	.	0.5200
sarcoma, undifferentiated, malign.	1	0	0	0	0.2182	0.4286	0.4615	0.4800
aorta								
carcinoma, basal cell, malig.	0	1	0	0	0.5091	0.5714	.	.
bone marrow, femur								
carcinoma, basal cell, malig.	0	1	0	0	0.5091	0.5714	.	.
bone, femur								
osteosarcoma, malignant	1	0	0	0	0.2182	0.4286	0.4615	0.4800
sarcoma, undifferentiated, malig.	1	0	0	0	0.2182	0.4286	0.4615	0.4800
clitoral glands								
sarcoma, undifferentiated, malig.	1	0	0	0	0.2182	0.4286	0.4615	0.4800
gallbladder								
carcinoma, squamous cell, malign.	0	1	0	0	0.5091	0.5714	.	.
harderian glands								
adenoma, benign	4	1	1	2	0.3714	0.1001	0.1222	0.2953
heart								
carcinoma, squamous cell,malignant	0	1	0	0	0.5091	0.5714	.	.
schwannoma, benign	1	0	0	0	0.2182	0.4286	0.4615	0.4800
kidneys								
carcinoma, squamous cell,malignant	0	1	0	0	0.5091	0.5714	.	.
large intestine, rectum								
sarcoma, undifferentiated, malign.	1	0	0	0	0.2182	0.4286	0.4615	0.4800
larynx								
carcinoma, basal cell, malignant	0	1	0	0	0.5091	0.5714	.	.
liver								
adenoma, hepatocellular, benign	1	2	2	4	0.0766	0.6084	0.5587	0.2004
carcinoma, basal cell, malignant	0	1	0	0	0.5091	0.5714	.	.
carcinoma, hepatocellular, malign.	1	3	0	0	0.0574	0.4220	0.4615	0.4800
ito cell tumor, malignant	0	1	0	0	0.5091	0.5714	.	.
lung								
adenoma, bronchiolar alveolar,ben.	0	3	3	0	0.2428	0.1789	0.1560	.
carcinoma, basal cell, malignant	0	1	0	0	0.5091	0.5714	.	.
carcinoma,bronch. alveolar,malign.	0	2	1	1	0.4128	0.3221	0.5385	0.5200
carcinoma, hepatocellular, malign.	0	2	0	0	0.2569	0.3221	.	.
carcinoma, squamous cell, malign.	0	1	0	0	0.5091	0.5714	.	.
fibrosarcoma, malignant	0	0	1	1	0.1767	.	0.5472	0.5200
sarcoma, undifferentiated, malign.	1	0	0	0	0.2182	0.4286	0.4615	0.4800
lymph node, inguinal								
sarcoma, undifferentiated, malign.	1	0	0	0	0.2182	0.4286	0.4615	0.4800
lymph node, mandibular								
carcinoma, basal cell, malignant	0	1	0	0	0.5091	0.5714	.	.
lymph node, mesenteric								
carcinoma, squamous cell, malign.	0	1	0	0	0.5091	0.5714	.	.
schwannoma, malignant	0	0	1	0	0.4909	.	0.5385	.
mammary gland								
adenocarcinoma, malignant	0	0	0	2	0.0542	.	.	0.2653
multicentric neoplasm								
hemangioma, benign	2	5	1	1	0.1005	0.3479	0.4413	0.4550
hemangiosarcoma, malignant	2	1	4	1	0.4461	0.3916	0.4304	0.4694
lymphoma, malignant	5	12	11	10	0.2435	0.1850	0.1277	0.1917
sarcoma, histiocytic, malignant	3	3	0	2	0.3163	0.5595	0.1048	0.5000
nerve, sciatic								
sarcoma, undifferentiated, malign.	1	0	0	0	0.2182	0.4286	0.4615	0.4800

Table AA.4 (cont.) Tumor Incidence and Tests in Female Mice

organ / tumor	Cntrl	Low	Med	High	Trend	CvsL	CvsM	CvsHi
ovaries								
carcinoma, squamous cell,malignant	0	1	0	0	0.5091	0.5714	.	.
cystadenoma, benign	3	3	0	1	0.1076	0.5390	0.0982	0.2899
granulosa cell tumor, benign	0	0	0	1	0.2364	.	.	0.5200
schwannoma, malignant	0	0	1	0	0.4909	.	0.5385	.
pancreas								
carcinoma, squamous cell,malignant	0	1	0	0	0.5091	0.5714	.	.
schwannoma, malignant	0	0	1	0	0.4909	.	0.5385	.
parathyroid glands								
adenoma, benign	0	0	1	0	0.4909	.	0.5385	.
pituitary gland								
adenoma, pars distalis, benign	3	3	7	5	0.1678	0.5176	0.2170	0.3987
adenoma, pars intermedia, benign	1	0	0	0	0.2182	0.4286	0.4615	0.4800
skeletal muscle, biceps femoris								
sarcoma, undifferentiated, malig.	2	0	0	0	0.0491	0.1880	0.2177	0.2353
skin								
sarcoma, undifferentiated, malign.	2	0	0	0	0.0460	0.1792	0.2081	0.2253
skin, subcutis								
adenocarcinoma, malignant	0	0	0	1	0.2364	.	.	0.5200
carcinoma, basal cell, malignant	0	1	0	0	0.5091	0.5714	.	.
fibrosarcoma, malignant	0	0	3	3	0.0257	.	0.1637	0.1405
fibrous histiocytoma, malignant	1	0	0	0	0.2252	0.4386	0.4717	0.4902
sarcoma, undifferentiated, malign.	1	0	0	0	0.2182	0.4286	0.4615	0.4800
schwannoma, malignant	0	1	1	0	0.5229	0.5789	0.5385	.
small intestine, duodenum								
adenocarcinoma, malignant	0	1	0	0	0.5091	0.5714	.	.
schwannoma, malignant	0	0	1	0	0.4909	.	0.5385	.
small intestine, ileum								
adenocarcinoma, malignant	0	1	0	0	0.5091	0.5714	.	.
small intestine, jejunum								
adenocarcinoma, malignant	0	1	0	0	0.5091	0.5714	.	.
schwannoma, malignant	0	0	1	0	0.4909	.	0.5385	.
stomach, glandular								
carcinoma, basal cell, malignant	0	1	0	0	0.5091	0.5714	.	.
carcinoma, squamous cell,malignant	0	1	0	0	0.5091	0.5714	.	.
stomach, nonglandular								
carcinoma, squamous cell, malig.	0	1	0	0	0.5091	0.5714	.	.
schwannoma, malignant	0	0	1	0	0.4909	.	0.5385	.
thyroid gland								
adenoma, follicular cell, benign	1	2	0	1	0.4150	0.6084	0.4615	0.7347
ureters								
carcinoma, squamous cell, malign.	0	1	0	0	0.5091	0.5714	.	.
schwannoma, malignant	0	0	1	0	0.4909	.	0.5385	.
urinary bladder								
sarcoma, undifferentiated, malign.	1	0	0	0	0.2182	0.4286	0.4615	0.4800
uterus with cervix								
adenocarcinoma, malignant	0	1	0	1	0.3144	0.5714	.	0.5200
polyp, stromal, benign	3	4	1	2	0.2472	0.6618	0.2486	0.4609

Application Type/Number	Submission Type/Number	Submitter Name	Product Name
NDA-22399	ORIG-1	GLAXO GROUP LTD DBA GLAXOSMITHKLINE	SOLZIRA

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

/s/

STEVEN F THOMSON
03/02/2010

KARL K LIN
03/03/2010
Concur with review

Statistical Review and Evaluation
(Carcinogenicity Studies)

Report #1

(MODIFIED VERSION OF ORIGINAL REPORT ISSED 1/20/2010)

NDA Number: 22-399

Drug Name: (b) (4)™; ((±-1-([(α-isobutanoyloxyethoxy)carbonyl]-aminomethyl)-1-cyclohexane acetic acid

Sponsor: GlaxoSmithKline

Pharm/tox Reviewer: Terry S. Peters, D.V.M.
Division of Neurology Products

Project Manager: Beverly Conner, Pharm.D.
Division of Neurology Products

Statistical Reviewers: Karl K. Lin, Ph.D.
Division of Biometrics 6
Office of Biostatistics

Document Reviewed: Carcinogenicity Assessment Committee (Cac/Cac-Ec) Report And FDA-CDER Rodent Carcinogenicity Database Factsheet, prepared by P/T reviewer Terry S. Peters, D.V.M., 6/16/09

Reason for This Modified Version of the Original Report

The statistical review and evaluation, #1, of the carcinogenicity studies of this NDA submission was finalized and put into DARRTS 1/20/2010 by this statistical reviewer. Shortly after the official finalization of the statistical review report of the submission, I was informed by my supervisors that they had the objection to some wordings used in the subsection "On Decision Rules in the Interpretation of Statistically Significant Results" of Section 4 "Reviewer's Comments on the Analysis Results of the Rat and Mouse Studies" in the original report. The modifications have been made in that particular subsection only following their suggestions. Specifically, the name of the particular CDER body being criticized in the original report has been replaced by a less specific group of people referred as "some pharm/tox reviewers". Also few words considered inflammatory have also been dropped from the original report.

Summary

Because of urgent need to meet a very tight time schedule, an abbreviated statistical review using survival-unadjusted analysis instead of a full review using survival-adjusted analysis was performed on tumor data of the carcinogenicity studies included in this submission. Results of the survival-unadjusted analysis show that the positive dose responses in tumor incidence in pancreas acinar adenoma and pancreas acinar adenoma+carcinoma in female rats were statistically significant at 0.025 significance level when the tumor types were classified as rare based on the background rates (0 % < 1%) of the control group. The positive dose-responses in the above two tumor types in male rats were considered as not statistically significant at 0.005 significance level when the tumor types were classified as common based on the background rates of 2/60= 3.3% > 1%) of the control group. Results of the survival-unadjusted analysis show that no positive dose-responses in the selected tumor types in both male and female mice were statistically significant at 0.005 level of significance. All the selected tumor types were considered as common based on the background rates of those tumor types in the control group.

If survival-adjusted analysis is performed on the tumor data, some of the non-statistically significant results from the survival-unadjusted analysis may become statistically significant because there will be more tumor-bearing animals in the groups with higher mortalities if the animal did not die early. The study in male rats showed the strongest results of positive dose-response in mortality among the four species-gender studies. For pancreas acinar adenoma and acinar adenoma+carcinoma in male rats, if the animals in the high group and in the medium group had the same survivals as the control and low groups, then the tumor-bearing animals for those tumor types would had been higher than 4 and 8 in acinar adenoma, and 5 and 9 in acinar adenoma+carcinoma, and the positive dose-responses in those tumor types would had been statistically significant. The above argument may also hold for the tumor types of uterus granular cell tumor and vagina granular cell tumor in female rats.

For a multi-group study (e.g., 3 doses and placebo), trend tests are more powerful (i.e., more likely to detect a true effect) than pairwise comparisons. Tests for trend instead of pairwise comparison tests between control and high-dose groups should therefore be the primary tests in the evaluation of drug related increases in tumor rate. The statistically significant finding in the test for positive dose-response in tumor incidence alone should be considered as real drug effect instead of as not a real effect using the requirement followed by some pharm/tox reviewers. The pharm/tox reviewers require that the statistical test for positive dose-response and the statistical test for pairwise positive difference between the control and the high groups have to be statistically significant simultaneously in order to consider a significant finding in the positive dose-response test as a real effect.

1. Introduction

Two chronic carcinogenicity studies, one in rats and one in mice were included in this submission. These studies were intended to assess the carcinogenic potential of [redacted]™; ((±-1-([(α-isobutanoyloxyethoxy)carbonyl]-aminomethyl)-1-cyclohexane acetic acid when administered at appropriate dose levels by oral gavage for a planned duration of 104 weeks. Terry S. Peters, D.V.M. of the Division of Neurology Products, the pharm/tox reviewer of this NDA submission, requested through her project manager Beverly Conner a statistical review and evaluation of the carcinogenicity studies in January 2009. However, for unknown reasons, the request has never reached the leader of the Pharm/Tox Statistics Team of the Office of Biostatistics (OB). Because of the unfortunate thing happened in the consultation request process, this review has never been entered the workload report of the Pharm/Tox Statistics Team, and assigned a statistical reviewer until 7/27/2009 when Dr. Peters checked with this reviewer about the status of this consultation request and informed this reviewer that the NDA submission was going to be discussed at the ECAC meeting scheduled on August 4, 2009.

This reviewer informed Dr. Peters that there was no way a full statistical review can be done just in few days before the scheduled ECAC meeting, that the discussion of the results of this NDA submission should be rescheduled to allow the statistical reviewer enough time to complete the review. However, Dr. Peters did not want to reschedule the discussion because her medical division would like to complete the review of this NDA submission as originally scheduled. Dr. Peters also indicated that she had reviewed the carcinogenicity studies and did not find major issues in those studies. To help Dr. Peters discuss the results of the studies at the 8/4/09 ECAC meeting, this reviewer offered her the quick option of conducting an abbreviated statistical review for this NDA submission. It was proposed to perform the survival-unadjusted test for dose-response and the survival-unadjusted pairwise comparison test in incidence to the data of tumor types in the two studies that may appear to have significant positive trends or differences in incidence. Dr. Peters agreed the reviewer's above proposal. This statistical review was done based on the proposal and on the draft Carcinogenicity Assessment Committee (Cac/Cac-Ec) Report And FDA-CDER Rodent Carcinogenicity Database Factsheet, prepared by P/T reviewer Terry S. Peters, D.V.M., 6/16/09.

2. Rat Carcinogenicity Study

Study Design

Rat study duration (weeks): 104 weeks
Study starting date: 6/21/05
Study ending date: 6/22/07
Rat strain: Crl: WI rats
Route: Oral gavage

Dosing comments: Dosed in 0.1% v/v Tween®80 and 0.5% w/v methylcellulose at 20 mL/kg

Number Of rats:

- Control (C1): 60
- Low Dose (LD): 60
- Middle Dose (MD): 60
- High Dose (HD1): 60

Rat dose levels (mg/kg/day):

- Low dose: 500
- Middle dose: 2000
- High dose:5000

Reviewer's Tumor Data Analysis

The survival-unadjusted permutation test for positive dose-response and the survival-unadjusted pairwise test for positive difference in tumor incidence between the control and each of the treated groups were used to analyze the data of some selected tumor types that may show statistically significant positive trends and/or positive differences. The actual doses, 0, 500, 2000, and 5000 mg/kg/day were used as the weights in the analysis. Results of the survival-unadjusted analysis are present in Table 1 below.

The survival-unadjusted analysis results show that the positive dose responses in tumor incidence in pancreas acinar adenoma and pancreas acinar adenoma+carcinoma in female rats were statistically significant at 0.025 significance level when the tumor types were classified as rare based on the background rates (0 % < 1%) of the control group. The positive dose-responses in the above two tumor types but in male rats were considered as not statistically significant at 0.005 significance level when the tumor types were classified as common based on the background rates of 2/60= 3.3% > 1%) of the control group. It is noted that the rate of pancreas acinar adenoma+carcinoma was just the simple summation of the rates of acinar adenoma and acinar carcinoma by the assumption that no animal developed both tumor types. This assumption was used because the reviewer did not have time to actually look the raw tumor dataset.

Table 1: Results of Survival-Unadjusted Analysis of Tumor Data of the Rat Study

Neoplastic Lesion	Tumor Incidences				P-Values[2]			
	C	L	M	H	Trend	C vs. L	C vs. M	C vs. H
Males								
Pancreas/ Acinar adenoma	2/60	4/60	4/60	8/60	0.026	0.340	0.340	0.047
Pancreas/ Acinar adenoma+carcinoma[1]	2/60	4/60	5/60	9/60	0.013	0.340	0.220	0.027
Females								
Pancreas/ Acinar adenoma	0/60	0/60	0/60	3/60	0.015*	----	----	0.122
Pancreas/ Acinar adenoma+carcinoma[1]	0/60	0/60	0/60	4/60	0.004*	----	----	0.059
Uterus/ Granular cell tumor	1/60	3/60	3/60	7/60	0.015	0.309	0.309	0.031
Vagina/ Granular cell tumor	2/60	2/60	2/60	5/60	0.081	0.691	0.691	0.220

1: The rate of pancreas/ acinar adenoma+carcinoma is just the simple sum of the rates of acinar adenoma and acinar carcinoma by the assumption that no animal develops both tumor types.

2: P-values are calculated by the exact methods.

3. * Statistically significant at 0.025 level of significance when the tumor type is classified as rare based on the background rate of the control group.

3. Mouse Carcinogenicity Study

Study Design

Mouse study duration (weeks): 104 weeks
Study starting date: 6/15/05
Study ending date: 2/28/08
Mouse strain: B6C3F1/Crl mice
R Oral gavage
Dosing comments: Dosed in 0.1% v/v Tween®80 and 0.5% w/v methylcellulose at 20 mL/kg

Number of mice:

- Control (C1): 60
- Low Dose (LD): 60
- Middle Dose (MD): 60
- High Dose (HD1): 60

Mouse dose levels* (mg/kg/day):

- Low dose: 500 mg/kg/d
- Middle dose: 2000 mg/kg/d

- High dose: 5000 mg/kg/d
(*Dose adjusted during study)

Reviewer's Tumor Data Analysis

The same survival-unadjusted methods used and described in the reviewer's tumor data analysis of the rat study were used to analyze the tumor data of the mouse study. Results of the survival-unadjusted analysis of the mouse study are presented in Table 2. The results show that no positive dose-responses in the selected tumor types were statistically significant at 0.005 level of significance. All the selected tumor types were considered as common based on the background rates of those tumor types in the control group.

Table 2: Results of Survival-Unadjusted Analysis of Tumor Data of the Mouse Study

Neoplastic Lesion	Tumor Incidences				P-Values[1]			
	C	L	M	H	Trend	C vs. L	C vs. M	C vs. H
Males								
Liver/Hepatocellular carcinoma	6/60	5/60	6/60	9/60	0.134	0.736	0.619	0.291
Females								
Liver/Hepatocellular adenoma	1/60	2/60	2/60	4/60	0.093	0.500	0.500	0.182
Multicentric neoplasm/Lymphoma	5/11	12/20	11/16	10/13	0.074	0.344	0.209	0.122
Pituitary gland/Adenoma	3/58	3/59	7/59	5/59	0.212	0.669	0.168	0.368

1: P-values are calculated by the exact methods.

4. Reviewer's Comments on the Analysis Results of the Rat and Mouse Studies

On Results from Survival-Unadjusted Analysis

Like human beings, older rodents have a many fold higher probability of developing or dying of tumors than those of a younger age. Therefore, in the analysis of tumor data, it is essential to identify and adjust for possible differences in intercurrent mortality among treatment groups to eliminate or reduce biases caused by these differences. Intercurrent mortality refers to all deaths other than those resulting from a tumor being analyzed for evidence of carcinogenicity. It has been pointed out that the effects of differences in longevity on numbers of tumor-bearing animals can be very substantial, and so, whether or not they (the effects) appear to be, they should routinely be corrected when presenting experimental results.

As this reviewer pointed out to Dr. Peters, the survival-unadjusted analysis is quick but the analysis results may not be valid unless the mortalities of the treatment groups are similar. If there is a positive dose response in mortality among the treatment groups, then

the survival-unadjusted analysis may yield non-statistically significant results in tumor incidence when there are true carcinogenic effects. Therefore, it is important to keep this point in mind in the final interpretation of study results and in the determination of the carcinogenicity of the drug. This especially true in the studies in male and female rats, and in male mice. It seems, as shown in Figures 1 – 4, that there are significant positive dose responses in mortality in those three studies. If survival-adjusted analysis is performed on the tumor data, some of the non-statistically significant results from the survival-unadjusted analysis may become statistically significant because there will be more tumor-bearing animals in the groups with higher mortalities if the animal did not die early.

The study in male rats showed the strongest results of positive dose-response in mortality among the four species-gender studies. For pancreas acinar adenoma and acinar adenoma+carcinoma in male rats, if the animals in the high group and in the medium group had the same survivals as the control and low groups, then the tumor-bearing animals for those tumor types would had been higher than 4 and 8 in acinar adenoma, and 5 and 9 in acinar adenoma+carcinoma, and the positive dose-responses in those tumor types would had been statistically significant. The above argument may also hold for the tumor types of uterus granular cell tumor and vagina granular cell tumor in female rats.

Since the mortalities of the female mice were fairly similar across all the treatment groups. The results of no statistically significant findings in the survival-unadjusted analysis should be valid.

Figure 1: Mean Survival Estimate Curves of Male Rats

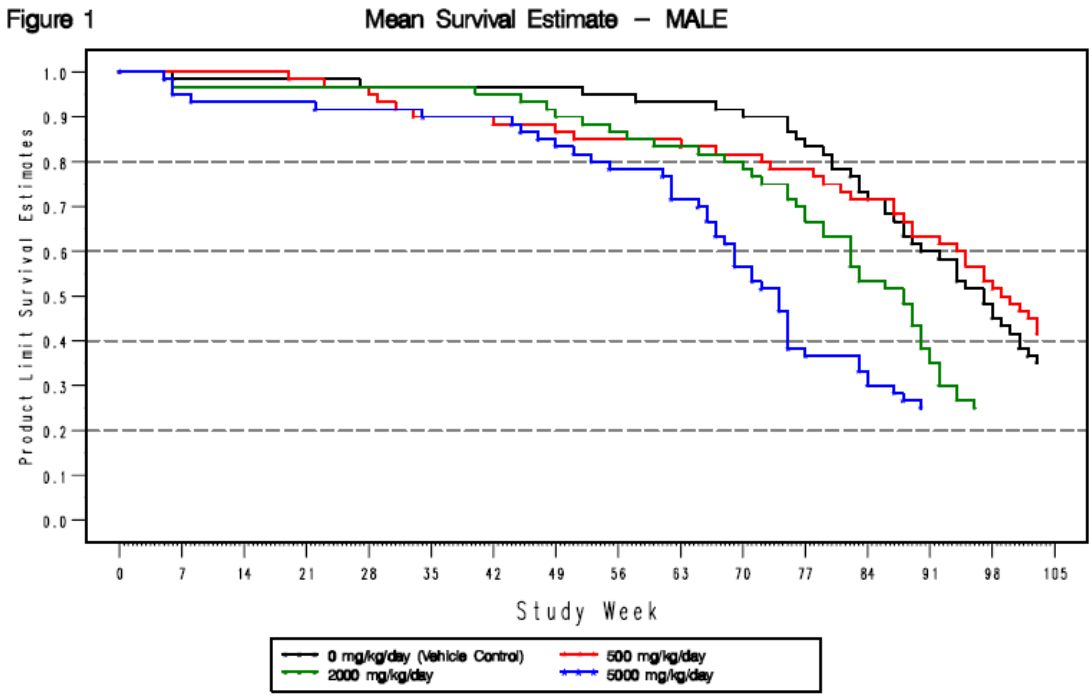


Figure 2: Mean Survival Estimate Curves of Female Rats

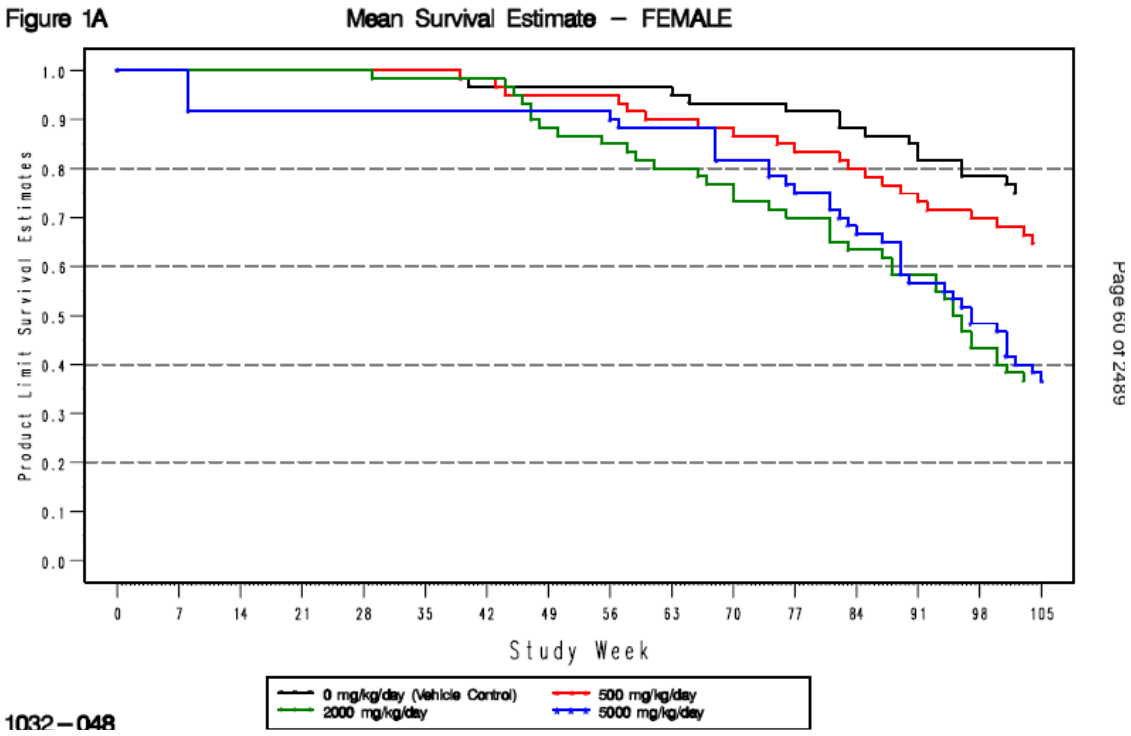


Figure 3: Mean Survival Estimate Curves of Male Mice

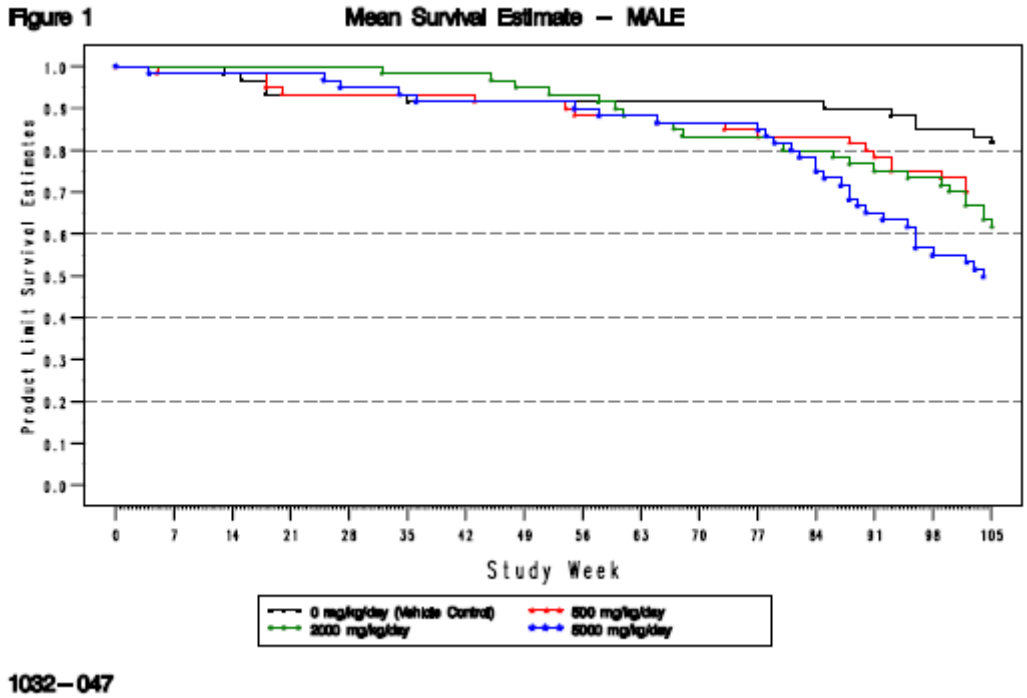
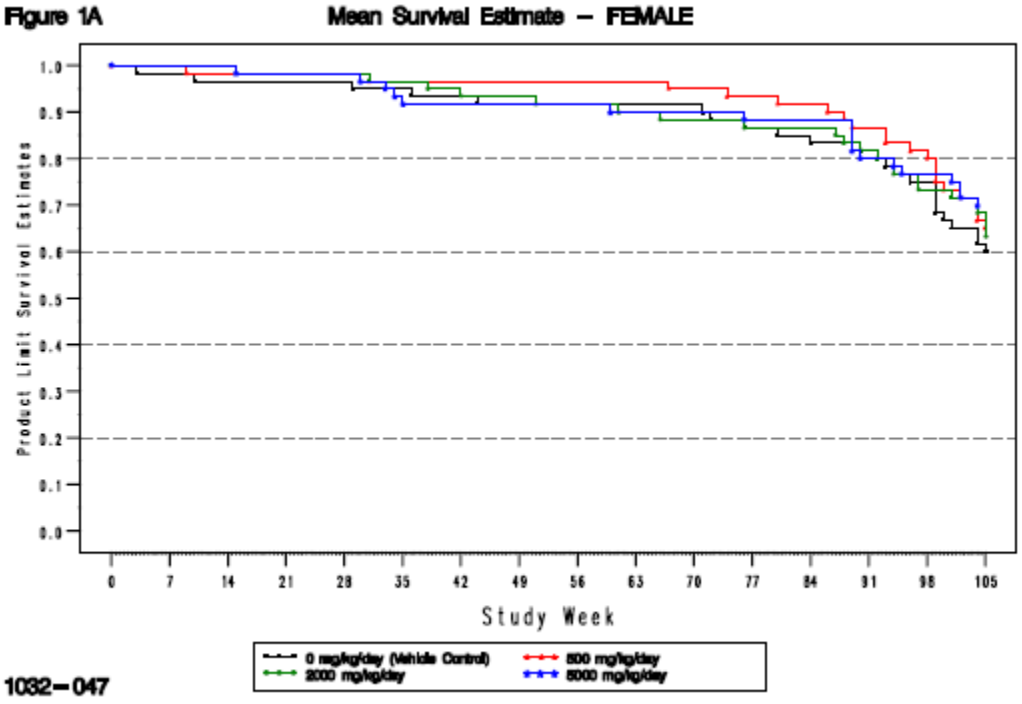


Figure 4: Mean Survival Estimate Curves of Female Mice



On Decision Rules in the Interpretation of Statistically Significant Results

It is well known in regulatory statistical literature that, for a multi-group study (e.g., 3 doses and placebo), trend tests are more powerful (i.e., more likely to detect a true effect) than pairwise comparisons. Tests for trend instead of pairwise comparison tests between control and high-dose groups should therefore be the primary tests in the evaluation of drug related increases in tumor rate although there are exceptional situations, however, in which pairwise comparisons between control and individual treated groups may be more appropriate than trend tests because trend tests assume that a carcinogenic effect is related to doses or systemic exposure weights, or ranks.

However, over the years, some pharm/tox reviewers have incorrectly applied the decision rules recommended in CDER statisticians in their efforts to reduce the false positive rate that measures the producer's risk (the well being of the sponsor) in toxicology studies without paying attention to the inflation in the false negative rate that measures the consumer's risk (the well being of the American public) in toxicology studies caused by its effort. To reduce the false positive rate, the reviewers set up their own requirement to consider a statistically significant finding as a true effect. They require that the results of both the trend test and the pairwise comparison test between the control and the high groups have to be statistically significant simultaneously at the levels of significance

recommended in the guidance for industry document in order to consider a statistically significant finding in the trend test as a true effect.

The statistically significant positive dose-responses in incidence in pancreas acinar adenoma and pancreas acinar adenoma+carcinoma in female rats will not be considered as real effects since the pairwise comparisons between the control and the high dose groups were not statistically significant in those two tumor types. This reviewer has strong objections to the requirement and practice based on the following sound scientific principles:

A. Because we make a decision about the true state of a population, such as a drug is carcinogenic or not in a population of mice and rats, based on limited information available to us from the data of an experiment (or a sample) with limited numbers of animals per treatment group, we will always commit two types of error called Type I and Type II errors in statistical inference. Type I error also called false positive error is the probability of concluding that there is a drug effect but in truth there is no drug effect. Type II error also called false negative error is the probability of concluding that there is no drug effect but in truth there is a drug effect. Type I error measures the producer's risk in toxicology studies, and measures the consumer's risk in clinical trials. Type II error measures the consumer's risk in toxicology studies and measures producer's risk in clinical trials. The false positive rate and the false negative rate run in opposite direction in the test of a statistical hypothesis. Trying to reduce one error rate, one will have to pay the price of increasing the other error rate as shown in Figure 5. Both false rates are bad. Decision-makers need to strike a balance in selecting the levels of risk between these two evils in their final interpretation about the carcinogenicity of a new drug. It is considered that the consumer's risk, not the producer's risk, should be the primary concern of regulatory authorities and agencies.

   In statistical analysis of carcinogenicity study data, the known false positive rate of an individual trend test that the pharm/tox reviewers try to reduce further is 0.005 (0.5%) or 0.025 (2.5%) in a two-species study, and 0.01 (1%) or 0.05 (5%) in a one-species study for a common and a rare tumor, respective. However, the magnitudes of the less familiar false negative rate of the trend test that the pharm/tox reviewers fail to consider can be very large, 100 or 200 times of the above known false positive rate or up to 0.7 (70%) to close to 1.0 (100%), for tumor types with low incidence rates in standard studies using 50-70 animals per treatment group. It will be difficult for the reviewers to defend their position, as regulators with the obligation to protect the well being of the consumer, for paying so much attention to reduce the already very low producer's risks and for ignoring the consumer's risks some of which can be extremely huge as shown in Figure 6.

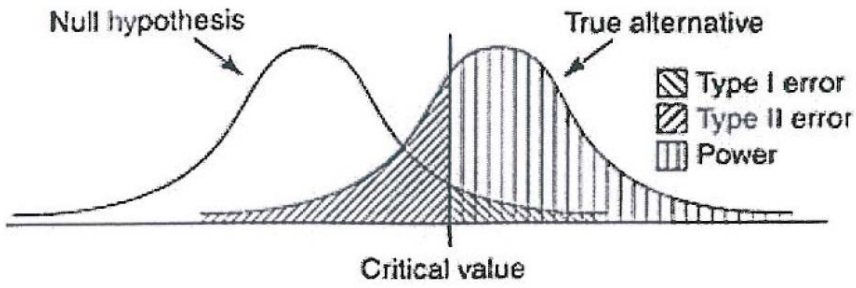
B. Results of an OB simulation study specifically conducted to address the important issue show that, as was expected and shown in Table 3, the false negative rates resulting from the pharm/tox reviewers' requirement of statistically significant results in both the trend test and pairwise comparison test simultaneously are higher than those from the procedure recommended in the guidance document that requires only a

statistically significant result in the trend test alone. The magnitude of inflation of false negative rate depends on the combination of the following factors simulated: (1) low or high tumor background rate, (2) tumors appearing early or late, (3) none, small, or large effect on tumor prevalence, and (4) none, small, or large effect on mortality. The third factor of the effect of the dose on tumor prevalence rate has the largest impact on the inflation of the false negative rate when both the trend test and the pairwise comparison tests are required to be statistically significant simultaneously in order to conclude that the effect is real. The inflations are the most serious in the situations in which the dose has the large effect on tumor prevalence. The inflation can be as high as 153.3% (i.e., more than double) of the false negative rate when the trend test alone is required to be statistically significant. This is the most alarming finding among those from the OB simulation study. When the dose of the test new drug has large effects on tumor prevalence, it is a clear indication that the drug is carcinogenic. Exactly in these most important situations the phrm/tox reviewers' practice causes the most serious inflation of the false negative (or the most serious reduction in statistical power to detect the true carcinogenic effect). The net result of this alarming finding is that the practice can be up to two and half times more likely to fail to detect a true carcinogenic effect than procedure based on the result of the trend test alone.

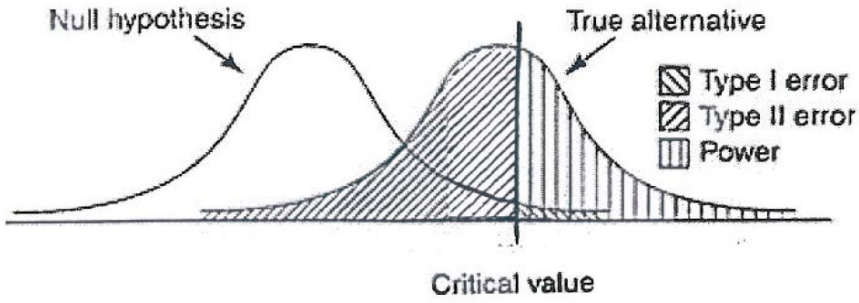
C. It is the main point that, with the group sizes (50-70 animals/group) used in regular chronic carcinogenicity studies as a surrogate of a big population of mice or rats with low tumor incidence rate endpoints, the false negative rate is already inherently big, the Agency should try to assume larger overall false positive rates, such as 0.1 (10%), in a study than those used in other types of drug development studies, such as clinical trials, to reduce the large false negative rate (or to increase the low power of detecting a true effect) inherited from the study design instead of trying to cut down the false positive rate further beyond that was estimated and considered as most appropriate in a regulatory environment. An overall false positive rate of 10% results in the multiplicity adjusted false positive rate for an individual trend test of 0.005 (0.5%) or 0.25 (2.5%) for common and rare tumors, respectively, in a two-species study or of 0.01 (1%) or 0.5 (5%), respectively, in a one-species study. It is important for the FDA to consider the producer's risk to make sure that a significantly positive result is not false positive since we all are benefited by not wasting the precious resources of the society. However, it is equally or even more important for FDA, as a regulator, to consider the consumer's risk to make sure that those non-significantly positive results are not false negative in order to provide an adequate protection for the health of American consumers.

Figure 5: Graphical Presentation of the Theoretical Relationship between Type I Error and Type II Error (1-Power) in Statistical Hypothesis Testing

Part A: Levels of Type II Error and Power under a Given Level of Type I Error



Part B: New Levels of Type II Error and Power When Type I error is reduced



Note of Figure 2: The slightly lighter shaded portion of the Type II error region in Part B of the figure is the increase of the Type II error (or the decrease of the power) when the Type I error is reduced from the level shown in Type I error region in Part A to that shown in Part B.

Figure 6: Type I Error (False Positive Rate: Producer's Risk) and Type II Error (False Negative Rate: Consumer's Risk) in a Statistical Test for Positive Trend in Tumor Incidence of a Given Tumor in a Carcinogenicity Study

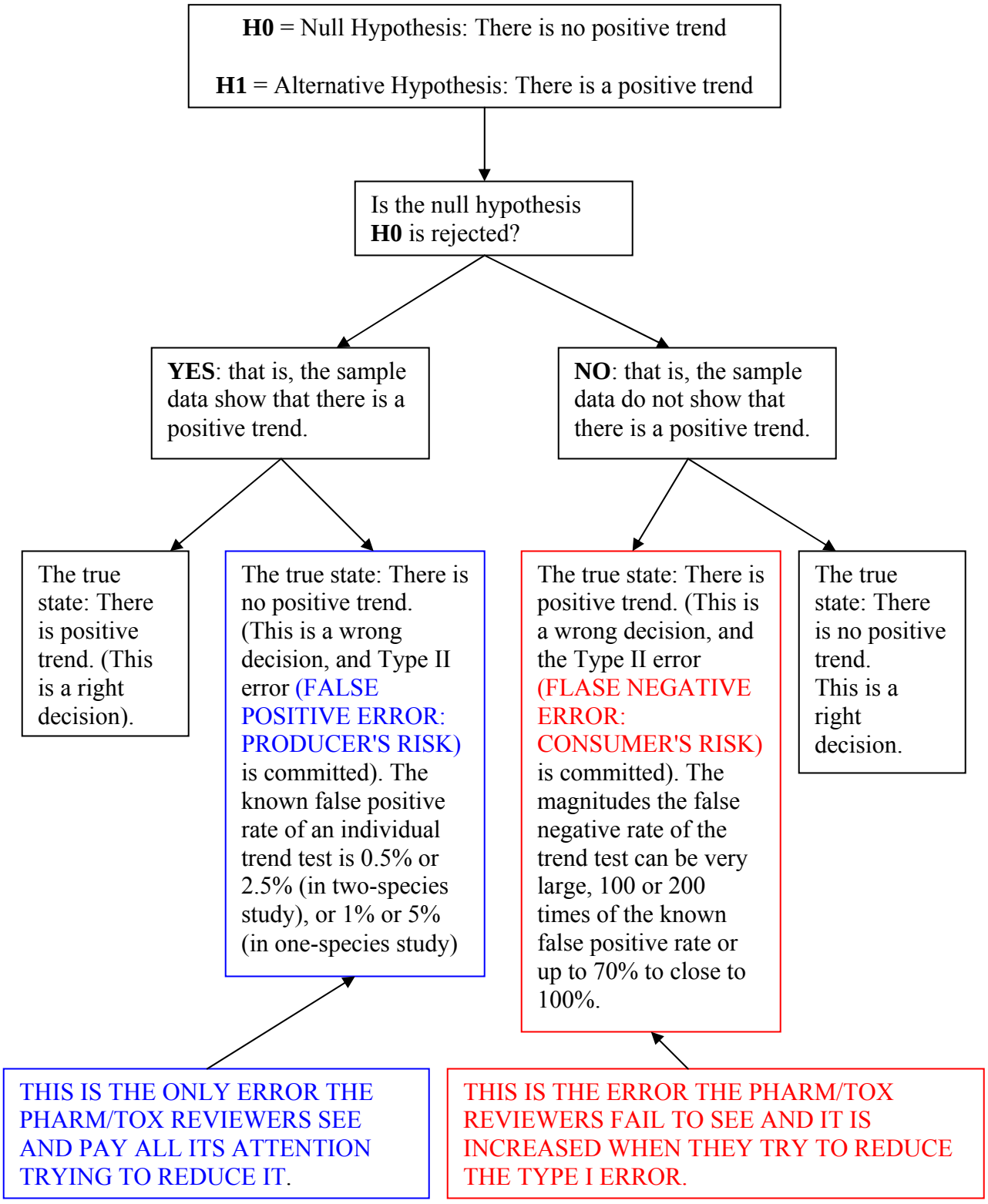


Table 3: Estimated False Negative Rates of Trend Test Alone and Trend Test Along with Pairwise Comparisons

Simulation Condition Number	Dose Effect on Death	Tumor Appearance	Dose Effect on Tumor Rate	Back Ground Tumor Rate	False Negative Rate			
					Trend	Trend and High	Trend and Any	Percent Change Tr-High	Percent Change Tr-Any
1	No	Early	No	.0500	.9840	.9934	.9919	.9553	.8028
2	No	Early	Small	.0500	.6283	.7084	.6957	12.75	10.73
3	No	Early	Large	.0500	.1313	.1780	.1595	35.57	21.48
4	No	Late	No	.0501	.9827	.9927	.9915	1.018	.8955
5	No	Late	Small	.0501	.6314	.7208	.7076	14.16	12.07
6	No	Late	Large	.0501	.1408	.2018	.1811	43.32	28.62
7	No	Early	No	.2000	.9953	.9979	.9974	.2612	.2110
8	No	Early	Small	.2000	.8377	.8805	.8715	5.109	4.035
9	No	Early	Large	.2000	.3424	.4270	.3980	24.71	16.24
10	No	Late	No	.2001	.9952	.9972	.9972	.2010	.2010
11	No	Late	Small	.2001	.8399	.8869	.8772	5.596	4.441
12	No	Late	Large	.2001	.3754	.4864	.4565	29.57	21.60
13	Small	Early	No	.0500	.9855	.9985	.9978	1.319	1.248
14	Small	Early	Small	.0500	.6967	.8465	.8324	21.50	19.48
15	Small	Early	Large	.0500	.2152	.4112	.3574	91.08	66.08
16	Small	Late	No	.0501	.9819	.9991	.9977	1.752	1.609
17	Small	Late	Small	.0501	.7220	.9161	.8903	26.88	23.31
18	Small	Late	Large	.0501	.2682	.6794	.6021	153.3	124.5
19	Small	Early	No	.2000	.9948	.9996	.9995	.4825	.4725
20	Small	Early	Small	.2000	.8753	.9694	.9606	10.75	9.745
21	Small	Early	Large	.2000	.4649	.7564	.7110	62.70	52.94
22	Small	Late	No	.2001	.9961	.9999	.9996	.3815	.3514
23	Small	Late	Small	.2001	.8935	.9939	.9885	11.24	10.63
24	Small	Late	Large	.2001	.5380	.9455	.9095	75.74	69.05
25	Large	Early	No	.0500	.9856	.9994	.9989	1.400	1.349
26	Large	Early	Small	.0500	.8381	.9587	.9480	14.39	13.11
27	Large	Early	Large	.0500	.5358	.8133	.7796	51.79	45.50
28	Large	Late	No	.0501	.9828	1.000	1.000	1.750	1.750
29	Large	Late	Small	.0501	.8675	.9960	.9886	14.81	13.96
30	Large	Late	Large	.0501	.6447	.9807	.9428	52.12	46.24
31	Large	Early	No	.2000	.9940	1.000	1.000	.6036	.6036
32	Large	Early	Small	.2000	.9414	.9994	.9985	6.161	6.065
33	Large	Early	Large	.2000	.7445	.9823	.9700	31.94	30.29
34	Large	Late	No	.2001	.9956	1.000	1.000	.4419	.4419
35	Large	Late	Small	.2001	.9585	1.000	.9999	4.330	4.319
36	Large	Late	Large	.2001	.8350	.9998	.9989	19.74	19.63

Table 3: False Negative Rates of Trend Test Alone and Trend Test Along with Pairwise Comparisons (Continued)

Notes of Table 8: (1) Columns under (a) "Trend", (b) "Trend and High", and (c) "Trend and Any" list the false negative rates, respectively, from requiring statistically significant results of the trend test alone, of the trend test and C-H pairwise comparison test simultaneously, and of the trend test and any of the three (C-L, C-M, C-H) pairwise comparison tests. (2) The last two columns list the percent changes of false negative rate of (b) over (a) and (c) over (a), respectively. (3) The estimated false negative rates under simulation numbers 1, 4, 7, 10, 13, 16, 19, 22, 25, 28, 31, and 34 are actually the estimated false positive rates because the assumption of no dose effect on tumor prevalence rate is used in those simulations.

Application Type/Number	Submission Type/Number	Submitter Name	Product Name
NDA-22399	ORIG-1	GLAXO GROUP LTD DBA GLAXOSMITHKLINE	SOLZIRA

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

/s/

KARL K LIN

02/18/2010

This is a modified version of the original statistical review and evauation report #1 issued 1/20/2010

Statistical Review and Evaluation Addendum
(Carcinogenicity Studies)

Report #2
(An Addendum)

NDA Number: 22-399

Drug Name: (b) (4)™; ((±-1-([(α-isobutanoyloxyethoxy)carbonyl]-aminomethyl)-1-cyclohexane acetic acid

Sponsor: GlaxoSmithKline

Pharm/tox Reviewer: Terry S. Peters, D.V.M.
Division of Neurology Products

Project Manager: Beverly Conner, Pharm.D.
Division of Neurology Products

Statistical Reviewers: Steve Thomson
Karl K. Lin, Ph.D.
Division of Biometrics 6
Office of Biostatistics

Document Reviewed: Statistical Review and Evaluation (Carcinogenicity Studies) prepared by Dr. Karl Lin, Ph.D. 8/3/09

Introduction

Again, both the original review and this addendum were done under a very tight time schedule. Because of this restriction, the original review did not provide survival adjusted tests of tumorigenicity. At the request of the ECAC, the purpose of this addendum is to add such an analysis to this submission.

Survival Analysis

First, the statistical significances of the tests of differences in survival across treatment groups are given in Table A.1 below. Tests of homogeneity over all groups, dose related trend and the pairwise differences between the high dose group and the vehicle control were performed. Two main test statistics are provided, the log rank test and the so-called Wilcoxon test. The log rank tests puts equal weight on all events being assessed, while the Wilcoxon test weights them by the square of their rank in time, and thus places more weight on later events than does the log rank test. So the Wilcoxon test will generally be more sensitive to later separation of mortality than will be the log rank test. Kaplan Meier survival curves for survival as a function of dose were provided in the original report. In both genders in rats, the tests of no overall homogeneity, no trend over dose, and no difference between the high dose group and control were all highly statistically significant (all $p \leq 0.0001$). From the survival curves, in male rats there is a generally increasing trend in mortality in dose. In female rats, the high dose group and the medium dose group have survival curves that are generally intertwined, as were the curves of the low dose group and control, but there was still a generally increasing trend in mortality in dose. Again from the survival curves, in male mice the high dose group has a generally higher mortality than the remaining dose groups. The medium and low dose groups generally intertwined, but with generally higher mortality than the control group. Again all the tests noted above in male mice were statistically significant (all $p \leq 0.0047$). In female mice the survival curves are generally intertwined with no particular evidence of differences in survival (all $p \geq 0.2987$). Absence of proof is not proof of absence, but, as noted in the report, the lack of evidence for differences should indicate that the non-mortality adjusted tests in female mice are appropriate.

Table A.1 Statistical Significances of Tests of Homogeneity and Trend in Survival

Rats	Males		Females	
	Log Rank	Wilcoxon	Log Rank	Wilcoxon
Homogeneity over Groups 1-4	<0.0001	<0.0001	<0.0001	<0.0001
Trend over Groups 1-4	<0.0001	<0.0001	<0.0001	0.0001
Comparison of High and Low	<0.0001	<0.0001	<0.0001	<0.0001

Mice	Males		Females	
	Log Rank	Wilcoxon	Log Rank	Wilcoxon
Homogeneity over Groups 1-4	0.0027	0.0047	0.7492	0.7931
Trend over Groups 1-4	0.0003	0.0008	0.3727	0.4526
Comparison of High and Low	0.0002	0.0003	0.2987	0.3523

Tumorigenicity

Again, the purpose of this addendum is to provide mortality adjusted tests of carcinogenicity. The consensus of the Society of Toxicological Pathology town hall meeting in June 2001 seemed to be that the poly-k modification of the Cochran-Armitage test of trend for tumor incidence should be used. That is the analysis provided in this addendum. Note that because of the software used, analysis over all tumors was only slightly more demanding than the analysis of subsets of the tumors. The results of tests of trend, and the pairwise comparisons of each treatment group to the controls are presented in the tables below. As discussed in the report, if one is determined to control statistical error, the results of the test of trend are recommended. Even if one includes tests of differences between the high dose and control, note that including the results of the pairwise tests between the medium and low dose groups can be expected to inflate Type I error, perhaps considerably. Nonetheless, in case they are of actual interest they are included below.

Tables A.2 and A.3 below display the results of any test that is potentially statistically significant at a nominal 0.05 level. Complete results are presented in the appendix to this addendum. Note that testing carcinogens involves a large number of tests. Based on his extensive experience with such analyses, for pairwise tests between the high dose group and controls in two species, Haseman (1983) claimed that for a roughly 0.10 (10%) overall false positive error rate, rare tumors should be tested at a 0.05 (5%) level, and common tumors (with a historical control incidence greater than 1%) at a 0.01 level. For a standard chronic study in two species (i.e., mice and mice) study, based on simulations and their experience, Lin & Rahman (1998) proposed a further p-value adjustment for tests of trend. That is, for a roughly 0.10 (10%) overall false positive error rate in tests of trend, rare tumors should be tested at a 0.025 (2.5%) level and common tumors at a 0.005 (0.5%) level. In this analysis, we will use the observed incidence in the vehicle control group to decide if a tumor is rare or common when applying these rules for multiplicity adjustment. Note the discussion in the report justifies emphasis on the tests of trend.

In Table A.1 below, using incidence in the control group to determine whether the tumor is rare or common, in rats both acinar cell benign adenoma and combined adenoma and carcinoma would be classified as common in male rats and rare in female rats. In both genders, tests of trend in acinar cell adenoma would be statistically significant (Males: $p = 0.0009 < 0.005$, Females: $p = 0.0105 < 0.025$). Similarly tests of no trend in pooled acinar cell adenoma and carcinoma are also statistically significant in both genders (Males: $p = 0.0002 < 0.005$, Females: $p = 0.0022 < 0.025$). In male rats, tests of comparisons between the high dose group and controls of both acinar cell benign adenoma and combined adenoma and carcinoma were statistically significant ($p=0.0029$ and $p=0.001$ both < 0.01). In female rats only the test of no differences between the high dose group and controls of combined adenoma and carcinoma was statistically significant ($p=0.0317 < 0.05$), although the comparison in adenoma was close to significance. Similarly, in female rats, the test of trend and differences between the high dose and controls in benign granular tumors of the uterus were quite close to these somewhat arbitrary bounds to determine statistical significance (i.e., $p = 0.0051 > 0.005$ and

p=0.0106 > 0.01, respectively). No other comparison achieved statistical significance when using the Haseman-Lin-Rahman adjustments for multiplicity cited above.

Table A.2 Potentially Significant Results (p≤ 0.05) in Rats

organ / tumor	Cntrl	Low	Med	High	Trend	CvsL	CvsM	CvsHi
Male Rats								
pancreas								
Acinar adenoma+carcinoma	2	4	5	9	0.0002	0.3629	0.1535	0.0010
adenoma, acinar cell, benign	2	4	4	8	0.0009	0.3629	0.2498	0.0029
adenoma, islet cell, benign	3	5	1	0	0.0365	0.3707	0.3835	0.2697
small intestine, jejunum								
adenocarcinoma, malignant	0	1	0	2	0.0465	0.5067	1.0000	0.1248
Female Rats								
mammary gland								
adenocarcinoma, malignant	5	1	10	2	0.4807	0.1276	0.0438	0.3558
fibroadenoma, benign	18	17	17	23	0.0140	0.5219	0.2973	0.0339
pancreas								
Acinar adenoma+carcinoma	0	0	0	4	0.0022	1.0000	1.0000	0.0317
adenoma, acinar cell, benign	0	0	0	3	0.0105	1.0000	1.0000	0.0769
skin, subcutis								
fibroma, benign	4	3	0	0	0.0091	0.5470	0.1085	0.0988
uterus with cervix								
granular cell tumor, benign	1	3	3	7	0.0051	0.2776	0.1959	0.0106
polyp, stromal, benign	11	8	2	3	0.0233	0.3782	0.0317	0.0698
vagina								
granular cell tumor, benign	2	2	2	5	0.0409	0.6593	0.5593	0.1206

Using the Haseman-Lin-Rahman rules, in mice only malignant fibrosarcoma of subcutis skin in female mice is even close to statistical significance (p=0.0257>0.025). No other tests achieved multiplicity adjusted statistical significance.

Table A.3 Potentially Significant Results (p≤ 0.05) in Mice

organ / tumor	Cntrl	Low	Med	High	Trend	CvsL	CvsM	CvsHi
Male Mice								
harderian glands								
adenoma, benign	5	0	1	1	0.1328	0.0308	0.1313	0.1154
Female Mice								
skeletal muscle, biceps femoris								
sarcoma, undifferentiated, malignant	2	0	0	0	0.0491	0.1880	0.2177	0.2353
skin								
sarcoma, undifferentiated, malignant	2	0	0	0	0.0460	0.1792	0.2081	0.2253
skin, subcutis								
fibrosarcoma, malignant	0	0	3	3	0.0257	1.0000	0.1637	0.1405

Note that complete incidence tables are included in the appendix below.

Appendix AA.1 Complete Tumor Incidence Tables

Tables AA.1-AA.4, below, present complete incidence tables and survival adjusted analyses for all tumors given in the Sponsor's data sets. Due to time constraints these reflect the exact breakdowns of organs and tumors as provided by the sponsor. As noted above, very fine breakdowns of organs or tumors, e.g. nose levels a to d, result in very few tumors for each such breakdown. This, in turn, makes it difficult to determine if any treatment differences are statistically significant, whether one uses p-values or uses posterior probabilities.

Table AA.1 Tumor Incidence and Tests in Male Rats

organ / tumor	Cntrl	Low	Med	High	Trend	CvsL	CvsM	CvsHi
adrenal glands								
adenoma, cortical, benign	1	2	2	0	0.3415	0.5101	0.4273	0.6515
pheochromocytoma, benign	7	3	5	2	0.3540	0.1436	0.5638	0.3256
pheochromocytoma, complex, benign	1	0	0	0	0.2945	0.4933	0.5616	0.6515
bone, mandible								
odontoma, benign	1	0	0	0	0.2945	0.4933	0.5556	0.6515
brain								
astrocytoma, benign	0	1	1	1	0.1669	0.5067	0.4521	0.3485
granular cell tumor, benign	1	0	0	0	0.2945	0.4933	0.5556	0.6515
meningioma, malignant	1	0	0	0	0.2945	0.4933	0.5616	0.6515
cavity, abdominal								
fibrosarcoma, malignant	0	0	1	0	0.4110	.	0.4521	.
cavity, oral								
carcinoma, squamous cell,malignant	1	1	2	0	0.4263	0.7400	0.4273	0.6515
papilloma, squamous cell, benign	1	0	0	0	0.2945	0.4933	0.5556	0.6515
coagulating glands								
carcinoma, squamous cell,malignant	0	0	1	0	0.4110	.	0.4444	.
epididymides								
mesothelioma, malignant	0	1	0	0	0.5890	0.5067	.	.
eyes								
leiomyoma, benign	1	0	0	0	0.2945	0.4933	0.5616	0.6515
schwannoma, malignant	1	0	0	0	0.2945	0.4933	0.5556	0.6515
harderian glands								
adenoma, benign	1	0	2	0	0.5240	0.4933	0.4273	0.6515
heart								
neuroendocrine tumor, benign	0	1	0	0	0.5890	0.5132	.	.
schwannoma, benign	0	1	0	0	0.5890	0.5132	.	.
kidneys								
carcinoma, tubular cell, malignant	0	1	0	1	0.1977	0.5067	.	0.3485
lipoma, benign	1	0	0	1	0.3008	0.4933	0.5556	0.5916
liposarcoma, malignant	1	0	0	0	0.2945	0.4933	0.5556	0.6515
lacrimal glands, exorbital								
carcinoma, zymbals gland,malignant	0	0	1	1	0.1043	.	0.4521	0.3485
large intestine, rectum								
adenocarcinoma, malignant	0	0	0	1	0.1633	.	.	0.3582
liver								
adenoma, hepatocellular, benign	0	0	1	1	0.1043	.	0.4444	0.3485
carcinoma, hepatocellular,malig.	0	0	0	1	0.1575	.	.	0.3485
lymph node, mesenteric								
schwannoma, malignant	1	0	0	0	0.2945	0.4933	0.5616	0.6515
mammary gland								
fibroadenoma, benign	0	1	2	1	0.1631	0.5067	0.2009	0.3485
multicentric neoplasm								
hemangioma, benign	3	1	0	1	0.3811	0.2973	0.1657	0.5651
hemangiosarcoma, malignant	2	4	1	1	0.3396	0.3501	0.5842	0.7236
lymphoma, malignant	0	2	3	0	0.5256	0.2600	0.0877	.
sarcoma, histiocytic, malignant	1	1	1	0	0.4018	0.7400	0.7032	0.6515

Table AA.1 (cont.) Tumor Incidence and Tests in Male Rats

organ / tumor	Cntrl	Low	Med	High	Trend	CvsL	CvsM	CvsHi
nose, level a								
carcinoma, squamous cell, malign.	0	1	0	0	0.5890	0.5132	.	.
sarcoma, undifferentiated, malign.	1	0	0	0	0.2945	0.4933	0.5556	0.6515
nose, level b								
carcinoma, squamous cell, malign.	0	1	0	0	0.5890	0.5132	.	.
sarcoma, undifferentiated, malign.	1	0	0	0	0.2945	0.4933	0.5556	0.6515
schwannoma, malignant	1	0	0	0	0.2945	0.4933	0.5556	0.6515
nose, level c								
carcinoma, squamous cell, malign.	0	1	2	0	0.4760	0.5132	0.2009	.
chondroma, benign	0	1	0	0	0.5890	0.5132	.	.
fibrosarcoma, malignant	1	0	0	0	0.2945	0.4933	0.5616	0.6515
schwannoma, malignant	1	0	0	0	0.2945	0.4933	0.5556	0.6515
nose, level d								
carcinoma, squamous cell, malig.	1	1	2	0	0.4263	0.7400	0.4273	0.6515
fibrosarcoma, malignant	1	0	0	0	0.2945	0.4933	0.5616	0.6515
schwannoma, malignant	1	0	0	0	0.2945	0.4933	0.5556	0.6515
pancreas								
Acinar adenoma+carcinoma	2	4	5	9	0.0002	0.3629	0.1535	0.0010
adenoma, acinar cell, benign	2	4	4	8	0.0009	0.3629	0.2498	0.0029
adenoma, islet cell, benign	3	5	1	0	0.0365	0.3707	0.3835	0.2697
carcinoma, acinar cell, malignant	0	0	1	1	0.1043	.	0.4521	0.3485
carcinoma, islet cell, malignant	1	0	0	0	0.2945	0.4933	0.5556	0.6515
parathyroid glands								
adenoma, benign	0	2	1	0	0.4676	0.2533	0.4521	.
pituitary gland								
adenoma, pars distalis, benign	31	30	29	20	0.4516	0.5000	0.5408	0.5468
adenoma, pars intermedia, benign	1	2	0	0	0.1788	0.5200	0.5556	0.6515
prostate gland								
adenoma, benign	1	0	1	0	0.4956	0.4933	0.7032	0.6515
seminal vesicles								
carcinoma, squamous cell, malig.	0	0	2	0	0.3541	.	0.1941	.
skin								
adenoma, basal cell, benign	0	0	0	1	0.1633	.	.	0.3582
adenoma, sebaceous cell, benign	0	0	0	1	0.1575	.	.	0.3485
carcinoma, squamous cell, malign.	1	2	0	0	0.1771	0.5200	0.5616	0.6515
keratoacanthoma, benign	3	2	0	1	0.3027	0.4875	0.1714	0.5651
papilloma, squamous cell, benign	1	1	1	0	0.4018	0.7467	0.6948	0.6515
schwannoma, malignant	1	0	0	0	0.2945	0.4933	0.5616	0.6515
skin, subcutis								
fibroma, benign	3	5	2	0	0.0633	0.3859	0.6034	0.2697
fibrosarcoma, malignant	2	1	1	0	0.2211	0.5000	0.5819	0.4210
schwannoma, malignant	2	1	0	0	0.1004	0.4800	0.3120	0.4210
small intestine, jejunum								
adenocarcinoma, malignant	0	1	0	2	0.0465	0.5067	.	0.1248
testes								
adenoma, interstitial cell, benign	8	6	10	3	0.4724	0.3637	0.2558	0.4183
mesothelioma, malignant	0	1	0	0	0.5890	0.5067	.	.
thymus gland								
thymoma, malignant	0	0	0	1	0.1575	.	.	0.3485
thyroid gland								
adenoma, c-cell, benign	6	6	6	0	0.0519	0.5840	0.4782	0.0671
adenoma, follicular cell, benign	3	2	2	1	0.4058	0.4747	0.6177	0.5474
carcinoma, follicular cell, malig.	1	0	0	0	0.2945	0.4933	0.5556	0.6515
tongue								
carcinoma, squamous cell, malig.	0	0	1	0	0.4110	.	0.4521	.
papilloma, squamous cell, benign	0	1	0	0	0.5890	0.5132	.	.
zymbal`s gland								
carcinoma, zymbals gland, malign.	1	0	1	2	0.0708	0.4933	0.7032	0.2900

Table AA.2 Tumor Incidence and Tests in Female Rats

organ / tumor	Cntrl	Low	Med	High	Trend	CvsL	CvsM	CvsHi
adrenal glands								
adenocarcinoma, malignant	0	0	1	0	0.4386	.	0.4253	.
adenoma, cortical, benign	0	1	1	2	0.0832	0.4792	0.4186	0.1836
carcinoma (primary site unknown), malignant	0	0	1	0	0.4386	.	0.4253	.
carcinoma, cortical, malignant	1	0	0	1	0.3981	0.5208	0.5814	0.6800
pheochromocytoma, benign	3	3	3	0	0.0905	0.6206	0.4954	0.1786
brain								
astrocytoma, benign	0	1	3	1	0.2484	0.4845	0.0698	0.4382
granular cell tumor, benign	1	0	2	0	0.4212	0.5208	0.3875	0.5682
meningioma, benign	1	0	0	0	0.2941	0.5208	0.5814	0.5682
oligodendroglioma, benign	1	1	0	0	0.2454	0.7314	0.5814	0.5682
cavity, abdominal								
carcinoma (primary site unknown), malignant	0	0	1	0	0.4386	.	0.4253	.
mesothelioma, malignant	0	1	0	0	0.5647	0.4792	.	.
cavity, oral								
carcinoma, squamous cell, malig.	0	0	1	0	0.4386	.	0.4253	.
cavity, thoracic								
liposarcoma, malignant	0	1	0	0	0.5647	0.4792	.	.
mesothelioma, malignant	0	1	0	0	0.5647	0.4792	.	.
harderian glands								
carcinoma, squamous cell, malign.	0	0	1	0	0.4386	.	0.4253	.
heart								
adenocarcinoma, malignant	0	0	1	0	0.4386	.	0.4253	.
kidneys								
adenocarcinoma, malignant	0	0	1	1	0.1451	.	0.4253	0.4318
carcinoma (primary site unknown), malignant	0	0	1	0	0.4386	.	0.4253	.
nephroblastoma, malignant	0	0	1	0	0.4386	.	0.4253	.
liver								
adenoma, hepatocellular, benign	1	0	0	0	0.2941	0.5208	0.5814	0.5682
lung								
adenocarcinoma, malignant	0	0	2	1	0.1554	.	0.1780	0.4318
carcinoma, squamous cell, malign.	1	0	0	0	0.2941	0.5208	0.5814	0.5682
lymph node, mesenteric								
adenocarcinoma, malignant	0	0	1	0	0.4386	.	0.4253	.
carcinoma (primary site unknown), malignant	0	0	1	0	0.4386	.	0.4253	.
mammary gland								
adenocarcinoma, malignant	5	1	10	2	0.4807	0.1276	0.0438	0.3558
adenoma, benign	2	2	1	3	0.2191	0.6593	0.6125	0.3827
fibroadenoma, benign	18	17	17	23	0.0140	0.5219	0.2973	0.0339
mesentery/peritoneum								
adenocarcinoma, malignant	0	0	1	0	0.4386	.	0.4253	.
multicentric neoplasm								
hemangioma, benign	0	0	0	1	0.2235	.	.	0.4318
hemangiosarcoma, malignant	1	0	1	0	0.4427	0.5208	0.6648	0.5682
lymphoma, malignant	2	2	2	2	0.3904	0.6593	0.5818	0.5818
sarcoma, histiocytic, malignant	0	0	1	0	0.4386	.	0.4253	.
nose, level b								
carcinoma, squamous cell, malign.	0	0	1	0	0.4386	.	0.4253	.
nose, level c								
carcinoma, squamous cell, malign.	0	0	1	0	0.4386	.	0.4253	.
nose, level d								
carcinoma, squamous cell, malign.	0	0	1	0	0.4386	.	0.4253	.

Table AA.2 (cont.) Tumor Incidence and Tests in Female Rats

organ / tumor	Cntrl	Low	Med	High	Trend	CvsL	CvsM	CvsHi
ovaries								
adenocarcinoma, malignant	0	0	1	1	0.1451	.	0.4253	0.4318
adenoma, tubulostromal, benign	0	1	0	0	0.5673	0.4845	.	.
granulosa cell tumor, benign	1	0	0	0	0.2941	0.5208	0.5814	0.5682
nephroblastoma, malignant	0	0	1	0	0.4386	.	0.4253	.
sertoli cell tumor, benign	0	1	0	0	0.5647	0.4792	.	.
sex-cord/stromal tumor, malignant	0	1	0	0	0.5647	0.4792	.	.
pancreas								
Acinar adenoma+carcinoma	0	0	0	4	0.0022	.	.	0.0317
adenocarcinoma, malignant	0	0	1	0	0.4386	.	0.4253	.
adenoma, acinar cell, benign	0	0	0	3	0.0105	.	.	0.0769
adenoma, islet cell, benign	0	2	0	1	0.3766	0.2270	.	0.4318
carcinoma (primary site unknown), malignant	0	0	1	0	0.4386	.	0.4253	.
carcinoma, acinar cell, malignant	0	0	0	1	0.2235	.	.	0.4318
carcinoma, islet cell, malignant	0	1	0	0	0.5673	0.4845	.	.
nephroblastoma, malignant	0	0	1	0	0.4386	.	0.4253	.
parathyroid glands								
adenoma, benign	1	0	1	2	0.1209	0.5208	0.6648	0.3972
pituitary gland								
adenoma, pars distalis, benign	41	39	37	35	0.4118	0.4989	0.1247	0.5138
carcinoma, pars distalis, malig.	0	1	0	0	0.5647	0.4792	.	.
salivary gland, parotid								
adenoma, benign	0	0	1	0	0.4353	.	0.4186	.
skeletal muscle, diaphragm								
carcinoma, squamous cell, malig.	1	0	0	0	0.2941	0.5208	0.5814	0.5682
skin, subcutis								
fibroma, benign	4	3	0	0	0.0091	0.5470	0.1085	0.0988
fibrosarcoma, malignant	0	0	2	0	0.4419	.	0.1780	.
schwannoma, malignant	1	0	0	0	0.2941	0.5208	0.5814	0.5682
small intestine, duodenum								
adenocarcinoma, malignant	0	0	1	0	0.4386	.	0.4253	.
adenoma, benign	1	0	0	0	0.2941	0.5208	0.5814	0.5682
carcinoma (primary site unknown), malignant	0	0	1	0	0.4386	.	0.4253	.
small intestine, ileum								
adenocarcinoma, malignant	0	0	1	0	0.4386	.	0.4253	.
leiomyosarcoma, malignant	0	0	0	1	0.2235	.	.	0.4318
small intestine, jejunum								
adenocarcinoma, malignant	1	0	0	0	0.2941	0.5208	0.5814	0.5682
spinal cord, cervical								
astrocytoma, malignant	0	0	1	0	0.4353	.	0.4186	.
spleen								
adenocarcinoma, malignant	0	0	1	1	0.1451	.	0.4253	0.4318
carcinoma (primary site unknown), malignant	0	0	1	0	0.4386	.	0.4253	.
carcinoma, squamous cell,malignant	1	0	0	0	0.2941	0.5208	0.5814	0.5682
stomach, glandular								
adenocarcinoma, malignant	0	0	1	0	0.4386	.	0.4253	.
stomach, nonglandular								
papilloma, squamous cell, benign	1	0	0	0	0.2941	0.5208	0.5814	0.5682
thymus gland								
adenocarcinoma, malignant	0	0	1	0	0.4386	.	0.4253	.
thymoma, malignant	0	0	1	0	0.4386	.	0.4253	.
thyroid gland								
adenoma, c-cell, benign	8	13	5	6	0.2430	0.1141	0.4975	0.6083
adenoma, follicular cell, benign	1	2	0	2	0.2883	0.4684	0.5814	0.3972
carcinoma, follicular cell, malig.	2	2	1	0	0.1323	0.6593	0.6224	0.3200

Table AA.2 (cont.) Tumor Incidence and Tests in Female Rats

organ / tumor	Cntrl	Low	Med	High	Trend	CvsL	CvsM	CvsHi
urinary bladder								
adenocarcinoma, malignant	0	0	1	1	0.1451	.	0.4253	0.4318
carcinoma (primary site unknown), malignant	0	0	1	0	0.4386	.	0.4253	.
sarcoma, stromal, malignant	0	0	1	0	0.4353	.	0.4186	.
uterus with cervix								
adenocarcinoma, malignant	3	0	1	2	0.3478	0.1372	0.4293	0.6292
adenoma, benign	0	1	0	0	0.5673	0.4845	.	.
carcinoma, squamous cell,malignant	1	0	0	0	0.2941	0.5208	0.5814	0.5682
granular cell tumor, benign	1	3	3	7	0.0051	0.2776	0.1959	0.0106
leiomyosarcoma, malignant	0	1	1	0	0.5580	0.4792	0.4186	.
nephroblastoma, malignant	0	0	1	0	0.4386	.	0.4253	.
polyp, stromal, benign	11	8	2	3	0.0233	0.3782	0.0317	0.0698
sarcoma, stromal, malignant	1	1	2	0	0.3182	0.7314	0.3776	0.5682
schwannoma, malignant	0	1	0	0	0.5647	0.4792	.	.
vagina								
granular cell tumor, benign	2	2	2	5	0.0409	0.6593	0.5593	0.1206
polyp, benign	0	1	0	0	0.5647	0.4792	.	.
sarcoma, stromal, malignant	0	0	1	0	0.4353	.	0.4186	.
schwannoma, malignant	0	1	0	0	0.5647	0.4792	.	.
zymbal`s gland								
adenoma, benign	1	0	0	0	0.2941	0.5208	0.5814	0.5682
carcinoma, zymbals gland,malignant	0	0	1	0	0.4386	.	0.4253	.

Table AA.3 Tumor Incidence and Tests in Male Mice

organ / tumor	Cntrl	Low	Med	High	Trend	CvsL	CvsM	CvsHi
harderian glands								
adenoma, benign	5	0	1	1	0.1328	0.0308	0.1313	0.1154
liver								
adenoma, hepatocellular, benign	0	3	1	3	0.1451	0.0952	0.4571	0.1050
carcinoma, hepatocellular, malign.	6	5	6	9	0.1213	0.5882	0.4946	0.2538
lung								
adenoma, bronchiolar alveolar,ben.	3	6	2	3	0.2888	0.1679	0.5850	0.6432
carcinoma, bronchiolar alveolar, malignant	0	0	0	1	0.2464	.	.	0.4722
carcinoma, hepatocellular, malign.	1	1	1	1	0.4947	0.7286	0.7126	0.7286
fibrosarcoma, malignant	0	0	2	0	0.4857	.	0.2159	.
lymph node, axillary								
fibrosarcoma, malignant	0	0	1	0	0.4783	.	0.4571	.
lymph node, mandibular								
fibrosarcoma, malignant	0	0	1	0	0.4783	.	0.4571	.
multicentric neoplasm								
hemangiosarcoma, malignant	1	0	1	3	0.0608	0.5278	0.7126	0.2811
lymphoma, malignant	4	3	3	1	0.1300	0.5627	0.6315	0.2245
sarcoma, histiocytic, malignant	3	0	3	1	0.3934	0.1357	0.6140	0.3455
pituitary gland								
adenoma, pars distalis, benign	0	1	0	0	0.5217	0.4722	.	.
adenoma, pars intermedia, benign	0	0	1	0	0.4783	.	0.4571	.
skin, subcutis								
fibrosarcoma, malignant	0	1	2	1	0.3201	0.4722	0.2159	0.4865
stomach, nonglandular								
papilloma, squamous cell, benign	0	0	0	1	0.2464	.	.	0.4722
testes								
adenoma, interstitial cell, benign	1	0	0	0	0.2754	0.5278	0.5429	0.5278
thyroid gland								
adenoma, follicular cell, benign	0	1	0	0	0.5217	0.4722	.	.

Table AA.4 Tumor Incidence and Tests in Female Mice

organ / tumor	Cntrl	Low	Med	High	Trend	CvsL	CvsM	CvsHi
adipose tissue, white, umbilical region								
lipoma, benign	0	1	0	0	0.5091	0.5714	.	.
adrenal glands								
carcinoma, squamous cell,malignant	0	1	0	0	0.5091	0.5714	.	.
pheochromocytoma, benign	0	0	0	1	0.2364	.	.	0.5200
sarcoma, undifferentiated, malign.	1	0	0	0	0.2182	0.4286	0.4615	0.4800
aorta								
carcinoma, basal cell, malig.	0	1	0	0	0.5091	0.5714	.	.
bone marrow, femur								
carcinoma, basal cell, malig.	0	1	0	0	0.5091	0.5714	.	.
bone, femur								
osteosarcoma, malignant	1	0	0	0	0.2182	0.4286	0.4615	0.4800
sarcoma, undifferentiated, malig.	1	0	0	0	0.2182	0.4286	0.4615	0.4800
clitoral glands								
sarcoma, undifferentiated, malig.	1	0	0	0	0.2182	0.4286	0.4615	0.4800
gallbladder								
carcinoma, squamous cell, malign.	0	1	0	0	0.5091	0.5714	.	.
harderian glands								
adenoma, benign	4	1	1	2	0.3714	0.1001	0.1222	0.2953
heart								
carcinoma, squamous cell,malignant	0	1	0	0	0.5091	0.5714	.	.
schwannoma, benign	1	0	0	0	0.2182	0.4286	0.4615	0.4800
kidneys								
carcinoma, squamous cell,malignant	0	1	0	0	0.5091	0.5714	.	.
large intestine, rectum								
sarcoma, undifferentiated, malign.	1	0	0	0	0.2182	0.4286	0.4615	0.4800
larynx								
carcinoma, basal cell, malignant	0	1	0	0	0.5091	0.5714	.	.
liver								
adenoma, hepatocellular, benign	1	2	2	4	0.0766	0.6084	0.5587	0.2004
carcinoma, basal cell, malignant	0	1	0	0	0.5091	0.5714	.	.
carcinoma, hepatocellular, malign.	1	3	0	0	0.0574	0.4220	0.4615	0.4800
ito cell tumor, malignant	0	1	0	0	0.5091	0.5714	.	.
lung								
adenoma, bronchiolar alveolar,ben.	0	3	3	0	0.2428	0.1789	0.1560	.
carcinoma, basal cell, malignant	0	1	0	0	0.5091	0.5714	.	.
carcinoma,bronch. alveolar,malign.	0	2	1	1	0.4128	0.3221	0.5385	0.5200
carcinoma, hepatocellular, malign.	0	2	0	0	0.2569	0.3221	.	.
carcinoma, squamous cell, malign.	0	1	0	0	0.5091	0.5714	.	.
fibrosarcoma, malignant	0	0	1	1	0.1767	.	0.5472	0.5200
sarcoma, undifferentiated, malign.	1	0	0	0	0.2182	0.4286	0.4615	0.4800
lymph node, inguinal								
sarcoma, undifferentiated, malign.	1	0	0	0	0.2182	0.4286	0.4615	0.4800
lymph node, mandibular								
carcinoma, basal cell, malignant	0	1	0	0	0.5091	0.5714	.	.
lymph node, mesenteric								
carcinoma, squamous cell, malign.	0	1	0	0	0.5091	0.5714	.	.
schwannoma, malignant	0	0	1	0	0.4909	.	0.5385	.
mammary gland								
adenocarcinoma, malignant	0	0	0	2	0.0542	.	.	0.2653
multicentric neoplasm								
hemangioma, benign	2	5	1	1	0.1005	0.3479	0.4413	0.4550
hemangiosarcoma, malignant	2	1	4	1	0.4461	0.3916	0.4304	0.4694
lymphoma, malignant	5	12	11	10	0.2435	0.1850	0.1277	0.1917
sarcoma, histiocytic, malignant	3	3	0	2	0.3163	0.5595	0.1048	0.5000
nerve, sciatic								
sarcoma, undifferentiated, malign.	1	0	0	0	0.2182	0.4286	0.4615	0.4800

Table AA.4 (cont.) Tumor Incidence and Tests in Female Mice

organ / tumor	Cntrl	Low	Med	High	Trend	CvsL	CvsM	CvsHi
ovaries								
carcinoma, squamous cell,malignant	0	1	0	0	0.5091	0.5714	.	.
cystadenoma, benign	3	3	0	1	0.1076	0.5390	0.0982	0.2899
granulosa cell tumor, benign	0	0	0	1	0.2364	.	.	0.5200
schwannoma, malignant	0	0	1	0	0.4909	.	0.5385	.
pancreas								
carcinoma, squamous cell,malignant	0	1	0	0	0.5091	0.5714	.	.
schwannoma, malignant	0	0	1	0	0.4909	.	0.5385	.
parathyroid glands								
adenoma, benign	0	0	1	0	0.4909	.	0.5385	.
pituitary gland								
adenoma, pars distalis, benign	3	3	7	5	0.1678	0.5176	0.2170	0.3987
adenoma, pars intermedia, benign	1	0	0	0	0.2182	0.4286	0.4615	0.4800
skeletal muscle, biceps femoris								
sarcoma, undifferentiated, malig.	2	0	0	0	0.0491	0.1880	0.2177	0.2353
skin								
sarcoma, undifferentiated, malign.	2	0	0	0	0.0460	0.1792	0.2081	0.2253
skin, subcutis								
adenocarcinoma, malignant	0	0	0	1	0.2364	.	.	0.5200
carcinoma, basal cell, malignant	0	1	0	0	0.5091	0.5714	.	.
fibrosarcoma, malignant	0	0	3	3	0.0257	.	0.1637	0.1405
fibrous histiocytoma, malignant	1	0	0	0	0.2252	0.4386	0.4717	0.4902
sarcoma, undifferentiated, malign.	1	0	0	0	0.2182	0.4286	0.4615	0.4800
schwannoma, malignant	0	1	1	0	0.5229	0.5789	0.5385	.
small intestine, duodenum								
adenocarcinoma, malignant	0	1	0	0	0.5091	0.5714	.	.
schwannoma, malignant	0	0	1	0	0.4909	.	0.5385	.
small intestine, ileum								
adenocarcinoma, malignant	0	1	0	0	0.5091	0.5714	.	.
small intestine, jejunum								
adenocarcinoma, malignant	0	1	0	0	0.5091	0.5714	.	.
schwannoma, malignant	0	0	1	0	0.4909	.	0.5385	.
stomach, glandular								
carcinoma, basal cell, malignant	0	1	0	0	0.5091	0.5714	.	.
carcinoma, squamous cell,malignant	0	1	0	0	0.5091	0.5714	.	.
stomach, nonglandular								
carcinoma, squamous cell, malig.	0	1	0	0	0.5091	0.5714	.	.
schwannoma, malignant	0	0	1	0	0.4909	.	0.5385	.
thyroid gland								
adenoma, follicular cell, benign	1	2	0	1	0.4150	0.6084	0.4615	0.7347
ureters								
carcinoma, squamous cell, malign.	0	1	0	0	0.5091	0.5714	.	.
schwannoma, malignant	0	0	1	0	0.4909	.	0.5385	.
urinary bladder								
sarcoma, undifferentiated, malign.	1	0	0	0	0.2182	0.4286	0.4615	0.4800
uterus with cervix								
adenocarcinoma, malignant	0	1	0	1	0.3144	0.5714	.	0.5200
polyp, stromal, benign	3	4	1	2	0.2472	0.6618	0.2486	0.4609

Application Type/Number	Submission Type/Number	Submitter Name	Product Name
NDA-22399	ORIG-1	GLAXO GROUP LTD DBA GLAXOSMITHKLINE	SOLZIRA

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

/s/

KARL K LIN
01/20/2010
Statistical review of carcinogenicity studies of NDA 22-399, report #2 (survival-adjusted analysis)

Statistical Review and Evaluation
(Carcinogenicity Studies)

Report #1

NDA Number: 22-399

Drug Name: [redacted]™; ((±-1-([(α-isobutanoyloxyethoxy)carbonyl]-aminomethyl)-1-cyclohexane acetic acid

Sponsor: GlaxoSmithKline

Pharm/tox Reviewer: Terry S. Peters, D.V.M.
Division of Neurology Products

Project Manager: Beverly Conner, Pharm.D.
Division of Neurology Products

Statistical Reviewers: Karl K. Lin, Ph.D.
Division of Biometrics 6
Office of Biostatistics

Document Reviewed: Carcinogenicity Assessment Committee (Cac/Cac-Ec) Report And FDA-CDER Rodent Carcinogenicity Database Factsheet, prepared by P/T reviewer Terry S. Peters, D.V.M., 6/16/09

Summary

Because of urgent need to meet a very tight time schedule, an abbreviated statistical review using survival-unadjusted analysis instead of a full review using survival-adjusted analysis was performed on tumor data of the carcinogenicity studies included in this submission. Results of the survival-unadjusted analysis show that the positive dose responses in tumor incidence in pancreas acinar adenoma and pancreas acinar adenoma+carcinoma in female rats were statistically significant at 0.025 significance level when the tumor types were classified as rare based on the background rates (0 % < 1%) of the control group. The positive dose-responses in the above two tumor types in male rats were considered as not statistically significant at 0.005 significance level when the tumor types were classified as common based on the background rates of 2/60= 3.3% > 1%) of the control group. Results of the survival-unadjusted analysis show that no positive dose-responses in the selected tumor types in both male and female mice were statistically significant at 0.005 level of significance. All the selected tumor types were considered as common based on the background rates of those tumor types in the control group.

If survival-adjusted analysis is performed on the tumor data, some of the non-statistically significant results from the survival-unadjusted analysis may become statistically significant because there will be more tumor-bearing animals in the groups with higher mortalities if the animal did not die early. The study in male rats showed the strongest results of positive dose-response in mortality among the four species-gender studies. For pancreas acinar adenoma and acinar adenoma+carcinoma in male rats, if the animals in the high group and in the medium group had the same survivals as the control and low groups, then the tumor-bearing animals for those tumor types would had been higher than 4 and 8 in acinar adenoma, and 5 and 9 in acinar adenoma+carcinoma, and the positive dose-responses in those tumor types would had been statistically significant. The above argument may also hold for the tumor types of uterus granular cell tumor and vagina granular cell tumor in female rats.

For a multi-group study (e.g., 3 doses and placebo), trend tests are more powerful (i.e., more likely to detect a true effect) than pairwise comparisons. Tests for trend instead of pairwise comparison tests between control and high-dose groups should therefore be the primary tests in the evaluation of drug related increases in tumor rate. The statistically significant finding in the test for positive dose-response in tumor incidence alone should be considered as real drug effect instead of the CAC requirement that the statistical test for positive dose-response and the statistical test for pairwise positive difference between the control and the high groups have to be statistically significant simultaneously in order to consider a significant finding in the positive dose-response test as a real effect.

1. Introduction

Two chronic carcinogenicity studies, one in rats and one in mice were included in this submission. These studies were intended to assess the carcinogenic potential of [redacted]™; ((±-1-([(α-isobutanoyloxyethoxy)carbonyl]-aminomethyl)-1-cyclohexane acetic acid when administered at appropriate dose levels by oral gavage for a planned duration of 104 weeks. Terry S. Peters, D.V.M. of the Division of Neurology Products, the pharm/tox reviewer of this NDA submission, requested through her project manager Beverly Conner a statistical review and evaluation of the carcinogenicity studies in January 2009. However, for unknown reasons, the request has never reached the leader of the Pharm/Tox Statistics Team of the Office of Biostatistics (OB). Because of the unfortunate thing happened in the consultation request process, this review has never been entered the workload report of the Pharm/Tox Statistics Team, and assigned a statistical reviewer until 7/27/2009 when Dr. Peters checked with this reviewer about the status of this consultation request and informed this reviewer that the NDA submission was going to be discussed at the ECAC meeting scheduled on August 4, 2009.

This reviewer informed Dr. Peters that there was no way a full statistical review can be done just in few days before the scheduled ECAC meeting, that the discussion of the results of this NDA submission should be rescheduled to allow the statistical reviewer enough time to complete the review. However, Dr. Peters did not want to reschedule the discussion because her medical division would like to complete the review of this NDA submission as originally scheduled. Dr. Peters also indicated that she had reviewed the carcinogenicity studies and did not find major issues in those studies. To help Dr. Peters discuss the results of the studies at the 8/4/09 ECAC meeting, this reviewer offered her the quick option of conducting an abbreviated statistical review for this NDA submission. It was proposed to perform the survival-unadjusted test for dose-response and the survival-unadjusted pairwise comparison test in incidence to the data of tumor types in the two studies that may appear to have significant positive trends or differences in incidence. Dr. Peters agreed the reviewer's above proposal. This statistical review was done based on the proposal and on the draft Carcinogenicity Assessment Committee (Cac/Cac-Ec) Report And FDA-CDER Rodent Carcinogenicity Database Factsheet, prepared by P/T reviewer Terry S. Peters, D.V.M., 6/16/09.

2. Rat Carcinogenicity Study

Study Design

Rat study duration (weeks): 104 weeks
Study starting date: 6/21/05
Study ending date: 6/22/07
Rat strain: Crl: WI rats
Route: Oral gavage
Dosing comments: Dosed in 0.1% v/v Tween®80 and 0.5% w/v methylcellulose at 20 mL/kg

Number Of rats:

- Control (C1): 60
- Low Dose (LD): 60
- Middle Dose (MD): 60
- High Dose (HD1): 60

Rat dose levels (mg/kg/day):

- Low dose: 500
- Middle dose: 2000
- High dose:5000

Reviewer's Tumor Data Analysis

The survival-unadjusted permutation test for positive dose-response and the survival-unadjusted pairwise test for positive difference in tumor incidence between the control and each of the treated groups were used to analyze the data of some selected tumor types that may show statistically significant positive trends and/or positive differences. The actual doses, 0, 500, 2000, and 5000 mg/kg/day were used as the weights in the analysis. Results of the survival-unadjusted analysis are present in Table 1 below.

The survival-unadjusted analysis results show that the positive dose responses in tumor incidence in pancreas acinar adenoma and pancreas acinar adenoma+carcinoma in female rats were statistically significant at 0.025 significance level when the tumor types were classified as rare based on the background rates (0 % < 1%) of the control group. The positive dose-responses in the above two tumor types but in male rats were considered as not statistically significant at 0.005 significance level when the tumor types were classified as common based on the background rates of 2/60= 3.3% > 1%) of the control group. It is noted that the rate of pancreas acinar adenoma+carcinoma was just the simple summation of the rates of acinar adenoma and acinar carcinoma by the assumption that no animal developed both tumor types. This assumption was used because the reviewer did not have time to actually look the raw tumor dataset.

Table 1: Results of Survival-Unadjusted Analysis of Tumor Data of the Rat Study

Neoplastic Lesion	Tumor Incidences				P-Values[2]			
	C	L	M	H	Trend	C vs. L	C vs. M	C vs. H
Males								
Pancreas/ Acinar adenoma	2/60	4/60	4/60	8/60	0.026	0.340	0.340	0.047
Pancreas/ Acinar adenoma+carcinoma[1]	2/60	4/60	5/60	9/60	0.013	0.340	0.220	0.027
Females								
Pancreas/ Acinar adenoma	0/60	0/60	0/60	3/60	0.015*	----	----	0.122
Pancreas/ Acinar adenoma+carcinoma[1]	0/60	0/60	0/60	4/60	0.004*	----	----	0.059
Uterus/ Granular cell tumor	1/60	3/60	3/60	7/60	0.015	0.309	0.309	0.031
Vagina/ Granular cell tumor	2/60	2/60	2/60	5/60	0.081	0.691	0.691	0.220

1: The rate of pancreas/ acinar adenoma+carcinoma is just the simple sum of the rates of acinar adenoma and acinar carcinoma by the assumption that no animal develops both tumor types.
2: P-values are calculated by the exact methods.
3. * Statistically significant at 0.025 level of significance when the tumor type is classified as rare based on the background rate of the control group.

3. Mouse Carcinogenicity Study

Study Design

Mouse study duration (weeks): 104 weeks
Study starting date: 6/15/05
Study ending date: 2/28/08
Mouse strain: B6C3F1/Crl mice
R Oral gavage
Dosing comments: Dosed in 0.1% v/v Tween®80 and 0.5% w/v methylcellulose at 20 mL/kg

Number of mice:

- Control (C1): 60
- Low Dose (LD): 60
- Middle Dose (MD): 60
- High Dose (HD1): 60

Mouse dose levels* (mg/kg/day):

- Low dose: 500 mg/kg/d
- Middle dose: 2000 mg/kg/d
- High dose: 5000 mg/kg/d

(*Dose adjusted during study)

Reviewer's Tumor Data Analysis

The same survival-unadjusted methods used and described in the reviewer's tumor data analysis of the rat study were used to analyze the tumor data of the mouse study. Results of the survival-unadjusted analysis of the mouse study are presented in Table 2. The results show that no positive dose-responses in the selected tumor types were statistically significant at 0.005 level of significance. All the selected tumor types were considered as common based on the background rates of those tumor types in the control group.

Table 2: Results of Survival-Unadjusted Analysis of Tumor Data of the Mouse Study

Neoplastic Lesion	Tumor Incidences				P-Values[1]			
	C	L	M	H	Trend	C vs. L	C vs. M	C vs. H
Males								
Liver/Hepatocellular carcinoma	6/60	5/60	6/60	9/60	0.134	0.736	0.619	0.291
Females								
Liver/Hepatocellular adenoma	1/60	2/60	2/60	4/60	0.093	0.500	0.500	0.182
Multicentric neoplasm/Lymphoma	5/11	12/20	11/16	10/13	0.074	0.344	0.209	0.122
Pituitary gland/Adenoma	3/58	3/59	7/59	5/59	0.212	0.669	0.168	0.368

1: P-values are calculated by the exact methods.

4. Reviewer's Comments on the Analysis Results of the Rat and Mouse Studies

On Results from Survival-Unadjusted Analysis

Like human beings, older rodents have a many fold higher probability of developing or dying of tumors than those of a younger age. Therefore, in the analysis of tumor data, it is essential to identify and adjust for possible differences in intercurrent mortality among treatment groups to eliminate or reduce biases caused by these differences. Intercurrent mortality refers to all deaths other than those resulting from a tumor being analyzed for evidence of carcinogenicity. It has been pointed out that the effects of differences in longevity on numbers of tumor-bearing animals can be very substantial, and so, whether or not they (the effects) appear to be, they should routinely be corrected when presenting experimental results.

As this reviewer pointed out to Dr. Peters, the survival-unadjusted analysis is quick but the analysis results may not be valid unless the mortalities of the treatment groups are similar. If there is a positive dose response in mortality among the treatment groups, then the survival-unadjusted analysis may yield non-statistically significant results in tumor incidence when there are true carcinogenic effects. Therefore, it is important to keep this

point in mind in the final interpretation of study results and in the determination of the carcinogenicity of the drug. This especially true in the studies in male and female rats, and in male mice. It seems, as shown in Figures 1 – 4, that there are significant positive dose responses in mortality in those three studies. If survival-adjusted analysis is performed on the tumor data, some of the non-statistically significant results from the survival-unadjusted analysis may become statistically significant because there will be more tumor-bearing animals in the groups with higher mortalities if the animal did not die early.

The study in male rats showed the strongest results of positive dose-response in mortality among the four species-gender studies. For pancreas acinar adenoma and acinar adenoma+carcinoma in male rats, if the animals in the high group and in the medium group had the same survivals as the control and low groups, then the tumor-bearing animals for those tumor types would had been higher than 4 and 8 in acinar adenoma, and 5 and 9 in acinar adenoma+carcinoma, and the positive dose-responses in those tumor types would had been statistically significant. The above argument may also hold for the tumor types of uterus granular cell tumor and vagina granular cell tumor in female rats.

Since the mortalities of the female mice were fairly similar across all the treatment groups. The results of no statistically significant findings in the survival-unadjusted analysis should be valid.

Figure 1: Mean Survival Estimate Curves of Male Rats

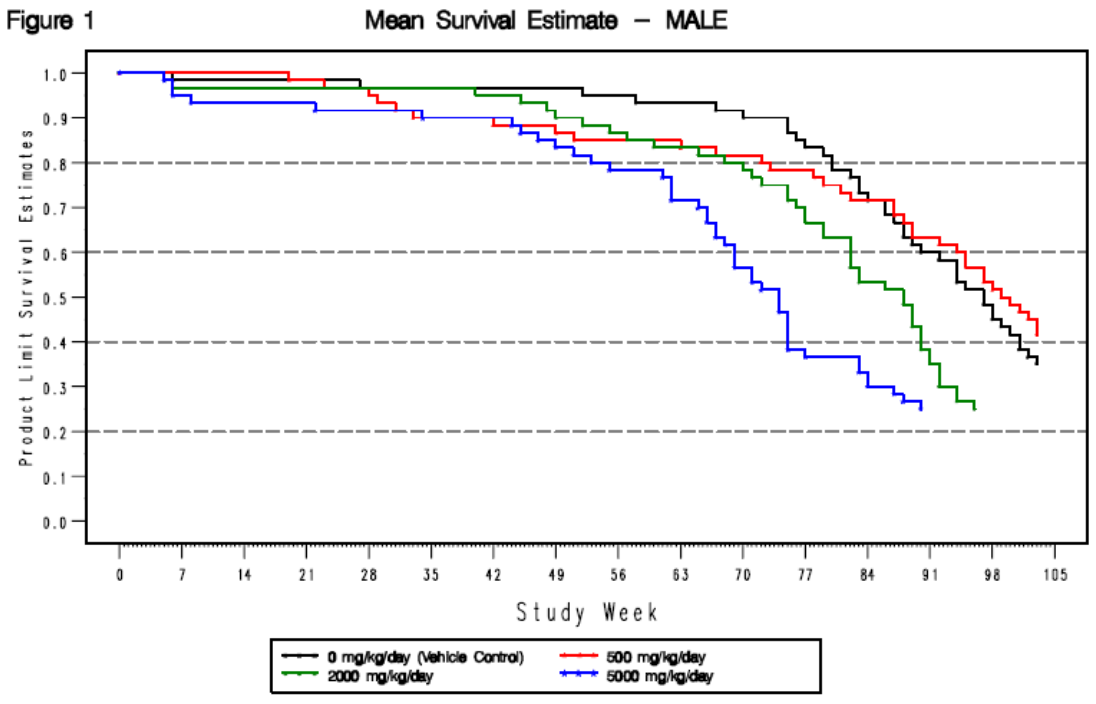


Figure 2: Mean Survival Estimate Curves of Female Rats

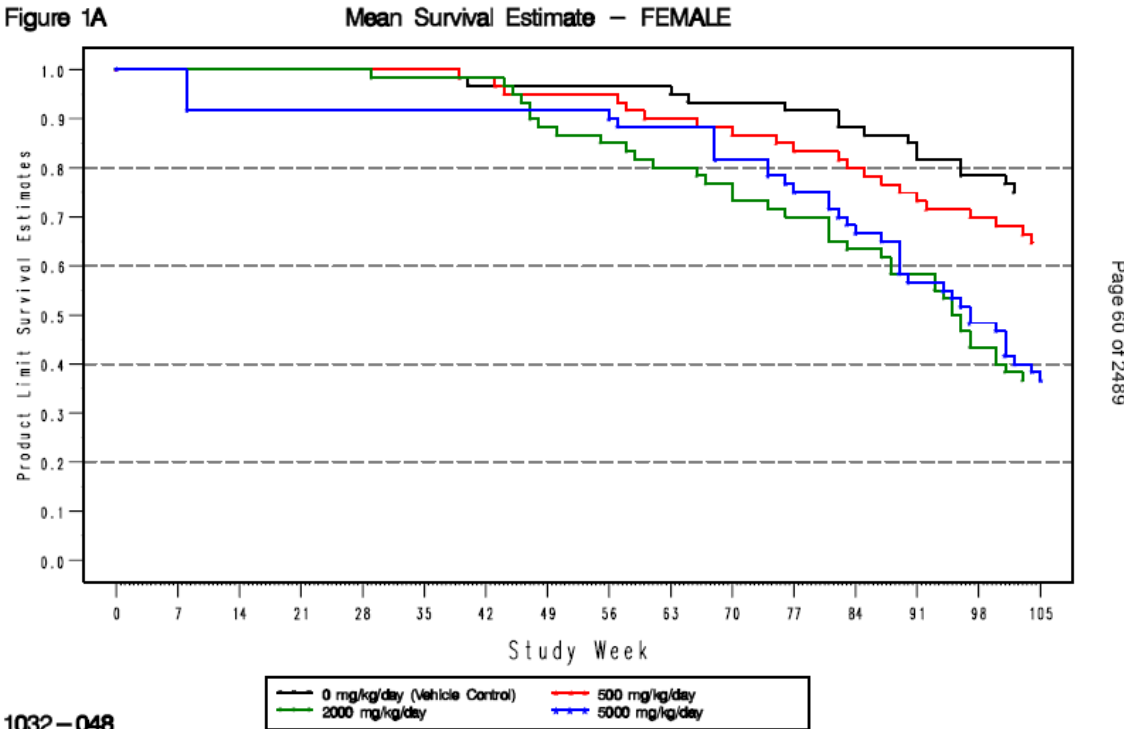


Figure 3: Mean Survival Estimate Curves of Male Mice

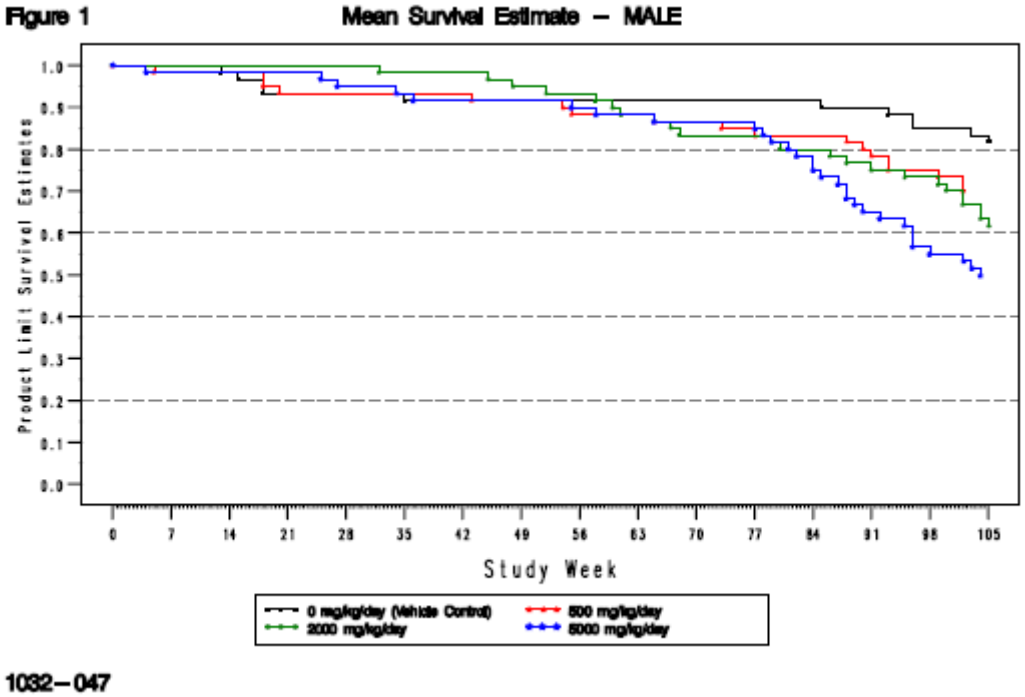
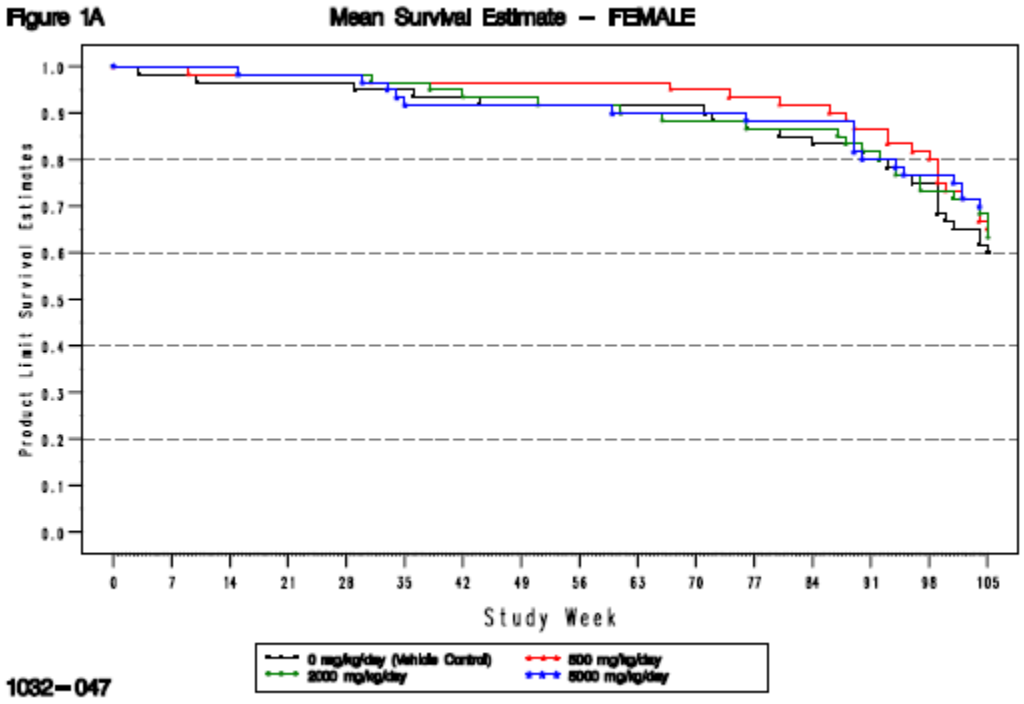


Figure 4: Mean Survival Estimate Curves of Female Mice



On Decision Rules in the Interpretation of Statistically Significant Results

It is well known in regulatory statistical literature that, for a multi-group study (e.g., 3 doses and placebo), trend tests are more powerful (i.e., more likely to detect a true effect) than pairwise comparisons. Tests for trend instead of pairwise comparison tests between control and high-dose groups should therefore be the primary tests in the evaluation of drug related increases in tumor rate although there are exceptional situations, however, in which pairwise comparisons between control and individual treated groups may be more appropriate than trend tests because trend tests assume that a carcinogenic effect is related to doses or systemic exposure weights, or ranks.

However, over the years, the Carcinogenesis Assessment Committee (CAC) of the Center for Drug Evaluation and Research (CDER), the decision-making body in the CDER review of pharmacology/toxicology studies of new drugs, has incorrectly applied the decision rules recommended in CDER statisticians in its great one-sided effort to reduce the false positive rate that measures the producer's risk (the well being of the sponsor) in toxicology studies without paying any attention to the inflation in the false negative rate that measures the consumer's risk (the well being of the American public) in toxicology studies caused by its effort. To reduce the false positive rate, the CAC set up its own requirement to consider a statistically significant finding as a true effect. The CAC requires that the results of both the trend test and the pairwise comparison test between

the control and the high groups have to be statistically significant simultaneously at the levels of significance recommended in the guidance for industry document in order to consider a statistically significant finding in the trend test as a true effect.

The statistically significant positive dose-responses in incidence in pancreas acinar adenoma and pancreas acinar adenoma+carcinoma in female rats will not be considered as real effects since the pairwise comparisons between the control and the high dose groups were not statistically significant in those two tumor types. This reviewer has strong objections to the CAC requirement and practice based on the following sound scientific principles:

A. Because we make a decision about the true state of a population, such as a drug is carcinogenic or not in a population of mice and rats, based on limited information available to us from the data of an experiment (or a sample) with limited numbers of animals per treatment group, we will always commit two types of error called Type I and Type II errors in statistical inference. Type I error also called false positive error is the probability of concluding that there is a drug effect but in truth there is no drug effect. Type II error also called false negative error is the probability of concluding that there is no drug effect but in truth there is a drug effect. Type I error measures the producer's risk in toxicology studies, and measures the consumer's risk in clinical trials. Type II error measures the consumer's risk in toxicology studies and measures producer's risk in clinical trials. The false positive rate and the false negative rate run in opposite direction in the test of a statistical hypothesis. Trying to reduce one error rate, one will have to pay the price of increasing the other error rate as shown in Figure 5. Both false rates are bad. Decision-makers need to strike a balance in selecting the levels of risk between these two evils in their final interpretation about the carcinogenicity of a new drug. It is considered that the consumer's risk, not the producer's risk, should be the primary concern of regulatory authorities and agencies.

   In statistical analysis of carcinogenicity study data, the known false positive rate of an individual trend test that the CAC tries to reduce further is 0.005 (0.5%) or 0.025 (2.5%) in a two-species study, and 0.01 (1%) or 0.05 (5%) in a one-species study for a common and a rare tumor, respective. However, the magnitudes of the less familiar false negative rate of the trend test that the CAC fails to consider can be very large, 100 or 200 times of the above known false positive rate or up to 0.7 (70%) to close to 1.0 (100%), for tumor types with low incidence rates in standard studies using 50-70 animals per treatment group. It will be difficult for the CAC to defend its position, as a regulator body with obligation to protect the well being of the consumer, for paying so much attention to reduce the already very low producer's risks and for ignoring the consumer's risks some of which can be extremely huge as shown in Figure 6.

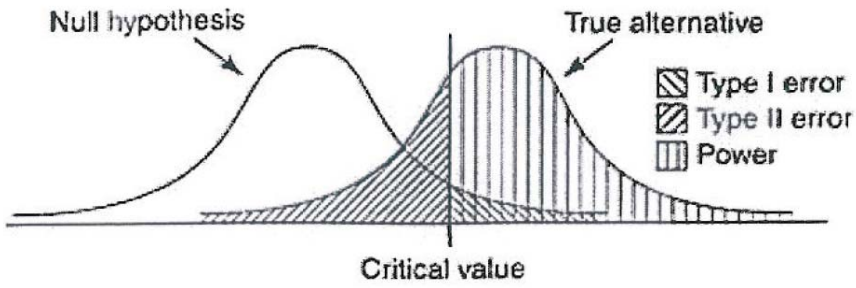
B. Results of an OB simulation study specifically conducted to address the important issue show that, as was expected and shown in Table 3, the false negative rates resulting from the CAC requirement of statistically significant results in both the trend test and pairwise comparison test simultaneously are higher than those from the procedure recommended in the guidance document that requires only a statistically

significant result in the trend test alone. The magnitude of inflation of false negative rate depends on the combination of the following factors simulated: (1) low or high tumor background rate, (2) tumors appearing early or late, (3) none, small, or large effect on tumor prevalence, and (4) none, small, or large effect on mortality. The third factor of the effect of the dose on tumor prevalence rate has the largest impact on the inflation of the false negative rate when both the trend test and the pairwise comparison tests are required to be statistically significant simultaneously in order to conclude that the effect is real. The inflations are the most serious in the situations in which the dose has the large effect on tumor prevalence. The inflation can be as high as 153.3% (i.e., more than double) of the false negative rate when the trend test alone is required to be statistically significant. This is the most alarming finding among those from the OB simulation study. When the dose of the test new drug has large effects on tumor prevalence, it is a clear indication that the drug is carcinogenic. Exactly in these most important situations the CAC practice causes the most serious inflation of the false negative (or the most serious reduction in statistical power to detect the true carcinogenic effect). The net result of this alarming finding is that the CAC practice can be up to two and half times more likely to fail to detect a true carcinogenic effect than procedure based on the result of the trend test alone.

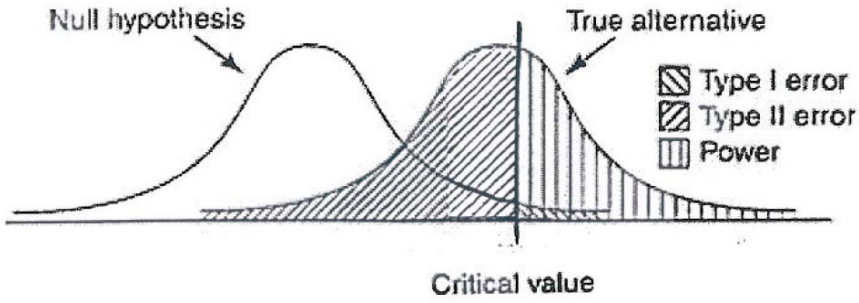
C. It is the main point that, with the group sizes (50-70 animals/group) used in regular chronic carcinogenicity studies as a surrogate of a big population of mice or rats with low tumor incidence rate endpoints, the false negative rate is already inherently big, the Agency should try to assume larger overall false positive rates, such as 0.1 (10%), in a study than those used in other types of drug development studies, such as clinical trials, to reduce the large false negative rate (or to increase the low power of detecting a true effect) inherited from the study design instead of trying to cut down the false positive rate further beyond that was estimated and considered as most appropriate in a regulatory environment. An overall false positive rate of 10% results in the multiplicity adjusted false positive rate for an individual trend test of 0.005 (0.5%) or 0.25 (2.5%) for common and rare tumors, respectively, in a two-species study or of 0.01 (1%) or 0.5 (5%), respectively, in a one-species study. It is important for the FDA to consider the producer's risk to make sure that a significantly positive result is not false positive since we all are benefited by not wasting the precious resources of the society. However, it is equally or even more important for FDA, as a regulator, to consider the consumer's risk to make sure that those non-significantly positive results are not false negative in order to provide an adequate protection for the health of American consumers.

Figure 5: Graphical Presentation of the Theoretical Relationship between Type I Error and Type II Error (1-Power) in Statistical Hypothesis Testing

Part A: Levels of Type II Error and Power under a Given Level of Type I Error



Part B: New Levels of Type II Error and Power When Type I error is reduced



Note of Figure 2: The slightly lighter shaded portion of the Type II error region in Part B of the figure is the increase of the Type II error (or the decrease of the power) when the Type I error is reduced from the level shown in Type I error region in Part A to that shown in Part B.

Figure 6: Type I Error (False Positive Rate: Producer's Risk) and Type II Error (False Negative Rate: Consumer's Risk) in a Statistical Test for Positive Trend in Tumor Incidence of a Given Tumor in a Carcinogenicity Study

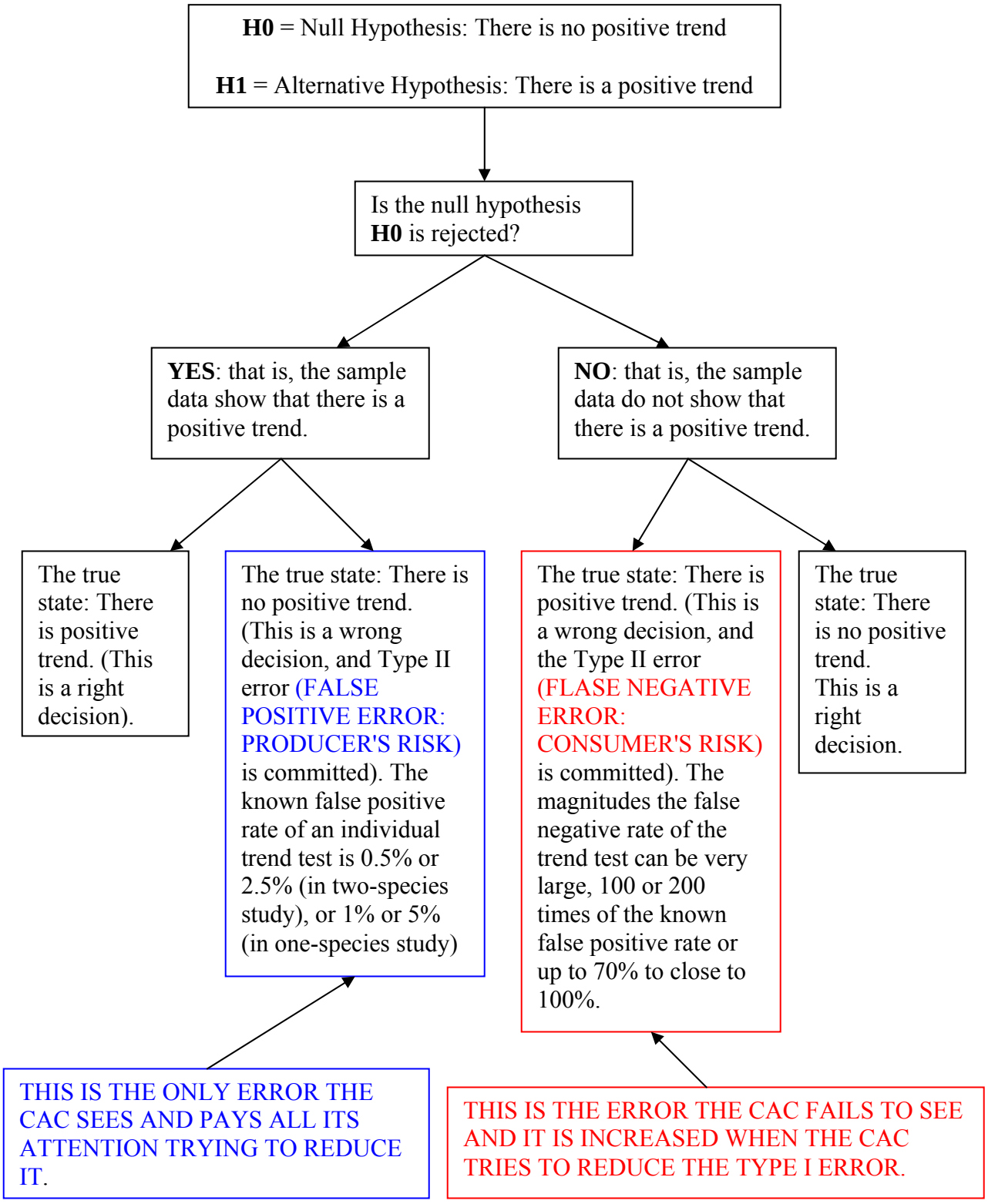


Table 3: Estimated False Negative Rates of Trend Test Alone and Trend Test Along with Pairwise Comparisons

Simulation Condition Number	Dose Effect on Death	Tumor Appearance	Dose Effect on Tumor Rate	Back Ground Tumor Rate	False Negative Rate			
					Trend	Trend and High	Trend and Any	Percent Change Tr-High	Percent Change Tr-Any
1	No	Early	No	.0500	.9840	.9934	.9919	.9553	.8028
2	No	Early	Small	.0500	.6283	.7084	.6957	12.75	10.73
3	No	Early	Large	.0500	.1313	.1780	.1595	35.57	21.48
4	No	Late	No	.0501	.9827	.9927	.9915	1.018	.8955
5	No	Late	Small	.0501	.6314	.7208	.7076	14.16	12.07
6	No	Late	Large	.0501	.1408	.2018	.1811	43.32	28.62
7	No	Early	No	.2000	.9953	.9979	.9974	.2612	.2110
8	No	Early	Small	.2000	.8377	.8805	.8715	5.109	4.035
9	No	Early	Large	.2000	.3424	.4270	.3980	24.71	16.24
10	No	Late	No	.2001	.9952	.9972	.9972	.2010	.2010
11	No	Late	Small	.2001	.8399	.8869	.8772	5.596	4.441
12	No	Late	Large	.2001	.3754	.4864	.4565	29.57	21.60
13	Small	Early	No	.0500	.9855	.9985	.9978	1.319	1.248
14	Small	Early	Small	.0500	.6967	.8465	.8324	21.50	19.48
15	Small	Early	Large	.0500	.2152	.4112	.3574	91.08	66.08
16	Small	Late	No	.0501	.9819	.9991	.9977	1.752	1.609
17	Small	Late	Small	.0501	.7220	.9161	.8903	26.88	23.31
18	Small	Late	Large	.0501	.2682	.6794	.6021	153.3	124.5
19	Small	Early	No	.2000	.9948	.9996	.9995	.4825	.4725
20	Small	Early	Small	.2000	.8753	.9694	.9606	10.75	9.745
21	Small	Early	Large	.2000	.4649	.7564	.7110	62.70	52.94
22	Small	Late	No	.2001	.9961	.9999	.9996	.3815	.3514
23	Small	Late	Small	.2001	.8935	.9939	.9885	11.24	10.63
24	Small	Late	Large	.2001	.5380	.9455	.9095	75.74	69.05
25	Large	Early	No	.0500	.9856	.9994	.9989	1.400	1.349
26	Large	Early	Small	.0500	.8381	.9587	.9480	14.39	13.11
27	Large	Early	Large	.0500	.5358	.8133	.7796	51.79	45.50
28	Large	Late	No	.0501	.9828	1.000	1.000	1.750	1.750
29	Large	Late	Small	.0501	.8675	.9960	.9886	14.81	13.96
30	Large	Late	Large	.0501	.6447	.9807	.9428	52.12	46.24
31	Large	Early	No	.2000	.9940	1.000	1.000	.6036	.6036
32	Large	Early	Small	.2000	.9414	.9994	.9985	6.161	6.065
33	Large	Early	Large	.2000	.7445	.9823	.9700	31.94	30.29
34	Large	Late	No	.2001	.9956	1.000	1.000	.4419	.4419
35	Large	Late	Small	.2001	.9585	1.000	.9999	4.330	4.319
36	Large	Late	Large	.2001	.8350	.9998	.9989	19.74	19.63

Table 3: False Negative Rates of Trend Test Alone and Trend Test Along with Pairwise Comparisons (Continued)

Notes of Table 8: (1) Columns under (a) "Trend", (b) "Trend and High", and (c) "Trend and Any" list the false negative rates, respectively, from requiring statistically significant results of the trend test alone, of the trend test and C-H pairwise comparison test simultaneously, and of the trend test and any of the three (C-L, C-M, C-H) pairwise comparison tests. (2) The last two columns list the percent changes of false negative rate of (b) over (a) and (c) over (a), respectively. (3) The estimated false negative rates under simulation numbers 1, 4, 7, 10, 13, 16, 19, 22, 25, 28, 31, and 34 are actually the estimated false positive rates because the assumption of no dose effect on tumor prevalence rate is used in those simulations.

Application Type/Number	Submission Type/Number	Submitter Name	Product Name
NDA-22399	ORIG-1	GLAXO GROUP LTD DBA GLAXOSMITHKLINE	SOLZIRA

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

/s/

KARL K LIN

01/20/2010

This is report #1 of the statistical review and evaluation (survival-unadjusted) of the carcinogenicity studies of this submission.

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

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/Serial Number:	22,399
Drug Name:	Solzira™ (Gabapentin Enacarbil) Extended Release Tablet
Indication(s):	Restless Legs Syndrome
Applicant:	GlaxoSmithKline
Date(s):	Document Date: January 9, 2009
	PDUFA Date: November 9, 2009
Review Priority:	Standard Review
Biometrics Division:	Division I
Statistical Reviewer:	Sharon Yan, Ph.D.
Concurring Reviewers:	Kun Jin, Ph.D., Team Leader
	James Hung, Ph.D., Division Director
Medical Division:	Division of Neurological Drug Products
Clinical Team:	Susanne Goldstein, M.D., Clinical Reviewer
	Gerald Podskalny, M.D. Clinical Acting Team Leader
	Eric Bastings, M.D., Deputy Director
Project Manager:	Beverly Conner, Pharm.D.

Table of Contents

STATISTICAL REVIEW AND EVALUATION **1**

1. EXECUTIVE SUMMARY **3**

1.1 CONCLUSIONS AND RECOMMENDATIONS 3
1.2 BRIEF OVERVIEW OF CLINICAL STUDIES 3
1.3 STATISTICAL ISSUES AND FINDINGS 4

2. INTRODUCTION **5**

2.1 OVERVIEW 5
2.2 DATA SOURCES 5

3. STATISTICAL EVALUATION **6**

3.1 EVALUATION OF EFFICACY 6
3.1.1 Studies XP052 and XP053 ..... *6*
3.1.2 Study XP060 ..... *13*
3.1.3 Study XP081 ..... *16*
3.1.4 Study XP083 ..... *20*
3.2 EVALUATION OF SAFETY 28

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS **29**

4.1 GENDER, RACE AND AGE 29
4.2 OTHER SPECIAL/SUBGROUP POPULATIONS 30

5. SUMMARY AND CONCLUSIONS **31**

5.1 STATISTICAL ISSUES AND COLLECTIVE EVIDENCE 31
5.2 CONCLUSIONS AND RECOMMENDATIONS 31

1. EXECUTIVE SUMMARY

The statistical review of this NDA covers two pivotal studies (XP052 and XP053), a long-term maintenance study (XP060), a phase II dose-exposure/response PK study (XP081), and a simulated driving performance study (XP083).

1.1 Conclusions and Recommendations

The efficacy results obtained from the analyses of the two pivotal studies XP052 and XP053 across two co-primary endpoints support the conclusion that XP13512 is effective in treating patients with restless legs syndrome (RLS). The effectiveness of XP13512 is also supported by Study XP060 and XP081. All doses studied (XP13512 600 mg, 1200 mg, 1800 mg, and 2400 mg) appear to be effective. XP13512 600 mg appears to be as efficacious as higher doses, and XP13512 2400 mg, the highest dose studied, does not appear to be more effective than the lower doses studied.

1.2 Brief Overview of Clinical Studies

XP13512 is a prodrug of gabapentin designed to overcome the pharmacokinetic limitations of gabapentin. The XP13512 clinical development program for RLS consists of 4 Phase II studies (XP021, XP045, XP081, and XP083) and 4 Phase III studies (XP052, XP053, XP060, and XP055). Studies XP052, XP053, and XP060 constitute the principal efficacy studies for the treatment of primary RLS with XP13512.

Studies XP052 and XP053 are pivotal, Phase III, 12-week, multicenter, randomized, double-blind, placebo-controlled, parallel-group studies in subjects with moderate-to-severe primary RLS. The co-primary efficacy endpoints are: (i) the change from baseline in IRLS Rating Scale total score and (ii) the proportion of subjects who were rated as responders (“much improved” or “very much improved”) on the investigator-rated CGI-I.

Study XP060 is a multicenter, randomized withdrawal study in subjects with moderate-to-severe primary RLS. Eligible subjects were initially enrolled in a 24-week single blind treatment period during which they received XP13512. Subjects who completed the initial treatment period and met the responder criteria were then randomized to receive either XP13512 or placebo during the 12-week double-blind treatment period. The primary study objective is to assess the maintenance of efficacy of XP13512 1200 mg in the long-term treatment of subjects with moderate-to-severe primary RLS. The primary efficacy variable iss the proportion of RLS subjects who relapsed during the double-blind treatment period.

Study XP081 is a phase II dose-exposure/response PK study, which is reviewed to provide supporting evidence for effectiveness of 600 mg XP13512.

In addition to the above studies, a 2-week Simulated Driving Performance Study XP083 is reviewed to evaluate the safety of driving under the treatment of XP13512.

1.3 Statistical Issues and Findings

In Studies XP052 and XP053, the change from baseline in the IRLS Rating Scale total score at the end of treatment (Week 12) and the proportion of responders on the investigator-rated CGI-I at Week 12 were co-primary efficacy endpoints. The results on both co-primary endpoints in both of the 12-week placebo-controlled efficacy studies provided sufficient evidence that XP13512 is effective in the treatment of RLS.

In Study XP052, the mean change from baseline to Week 12 for the IRLS Rating Scale total score was -13.2 in the XP13512 1200 mg group and -8.8 in the placebo group. The difference was statistically significant (p=0.0003). The proportion of responders on the investigator-rated CGI-I Scale at Week 12 was 76.1% in the XP13512 1200 mg group compared with 38.9% in the placebo group, and the estimated odds of improvement for XP13512 1200 mg relative to placebo were 5.1 (p<0.0001).

In Study XP053, the mean change from baseline to Week 12 for the IRLS Rating Scale total score was -13.0 in the XP13512 1200 mg group, -13.8 in the XP13512 600 mg group, and -9.8 in the placebo group (1200 mg vs. placebo: p=0.0017; 600 mg vs. placebo: p<0.0001). The proportion of responders on the investigator-rated CGI-I Scale at Week 12 LOCF was 77.5% in the XP13512 1200 mg group, 72.8% in the XP13512 600 mg group, compared with 44.8% in the placebo group. The odds of being a responder were 4.29 times that in the placebo group in the XP13512 1200 mg group (p<.0001) and 3.32 time that in the placebo group in the XP13512 600 mg group (p < .0001).

In Study XP060, the treatment difference between XP13512 and placebo was statistically significant in the primary efficacy endpoint of relapse rate: 9.4% of subjects in the XP13512 group and 22.7% of subjects in the placebo group relapsed by the end of the double-blind treatment period. The odds ratio for experiencing a relapse was 0.353 (XP13512 vs. placebo; p=0.0158).

Results from study XP081 are similar to the ones from XP052 and XP053.

In the two pivotal studies XP052 and XP053, the dropout rate is about 15%. Statistical significance between the treatment groups was reached beginning from Week 1 in observed case analysis, and the efficacy results appear to be consistent across study period in both studies.

No major statistical issues were identified.

2. INTRODUCTION

2.1 Overview

XP13512 is a prodrug of gabapentin designed to overcome the pharmacokinetic limitations of gabapentin. Studies XP052, XP053, and XP060 constitute the principal efficacy studies for the treatment of primary RLS with XP13512 in the clinical development program. A phase II dose response study XP081 is also included in this review in order to provide supportive evidence for the efficacy of 600 mg XP13512, which is not studied in XP052 and XP060.

Studies XP052 and XP053 were pivotal, Phase III, 12-week, multicenter, randomized, double-blind, placebo-controlled, parallel-group studies in subjects with moderate-to-severe primary RLS. The co-primary efficacy endpoints were: (i) the change from baseline in IRLS Rating Scale total score and (ii) the proportion of subjects who were rated as responders ("much improved" or "very much improved") on the investigator-rated CGI-I. A total of 222 subjects were randomized in 22 centers in Study XP052, and 325 subjects were randomized in 27 centers in Study XP053. Both studies were conducted in US.

Study XP060 was a multicenter, randomized withdrawal study in subjects with moderate-to-severe primary RLS. Eligible subjects were initially enrolled in a 24-week single blind treatment period during which they received XP13512. Subjects who completed the initial treatment period and met the responder criteria were then randomized to receive either XP13512 or placebo during the 12-week double-blind treatment period. The primary study objective was to assess the maintenance of efficacy of XP13512 1200 mg in the long-term treatment of subjects with moderate-to-severe primary RLS. The primary efficacy variable was the proportion of RLS subjects who relapsed during the double-blind treatment period. A total of 194 subjects were randomized into 26 study sites in US.

Study XP081 was conducted to measure XP13512 released gabapentin pharmacokinetics and to assess if there was a XP13512 dose/exposure-response relationship for the treatment of patients with RLS. This was a multicenter, randomized, double-blind, placebo-controlled, parallel-group, 12-week study, comparing 4 doses (600 mg, 1200 mg, 1800 mg, and 2400 mg) of XP13512 with placebo to subjects with RLS. There was no assignment of primary or secondary endpoints in this study as this study was conducted to measure XP13512 released gabapentin pharmacokinetics. IRLS Rating Scale total score and investigator-rated CGI-I were assessed, and thus efficacy of XP13512 doses other than 1200 mg, particularly 600 mg, can be evaluated from the available data.

Study XP083 was a 2-week Simulated Driving Performance Study to evaluate the safety of driving under the treatment of XP13512.

2.2 Data Sources

All document reviewed for this NDA submission are in electronic form. The path to CDER Electronic Document Room for the submission is listed below:

\\Cdsesub1\evsprod\NDA022399

3. STATISTICAL EVALUATION

3.1 Evaluation of Efficacy

3.1.1 Studies XP052 and XP053

3.1.1.1 Description of the Studies

The primary study objective of Studies XP052 and XP053 was to compare the efficacy of XP13512 1200 mg taken once daily versus placebo for the treatment of subjects suffering from Restless Legs Syndrome (RLS).

The two studies had almost identical design except that XP053 had an additional arm of lower dose group of 600 mg XP13512.

Both studies were multi-center, randomized, double-blind, placebo-controlled, parallel-group trials. Eligible subjects were randomized to receive XP13512 or placebo for the 12-week treatment period. Study XP052 had 2 treatment arms, and subjects were randomized to a once daily dose of 1200 mg XP13512 or matching placebo. Study XP053 had an additional dose arm of 600 mg XP13512. This lower dose arm was added to the Protocol Amendment 1 based on FDA's recommendation.

Each of the two studies planned to enroll 105 subjects per treatment arm.

3.1.1.2 Efficacy Evaluation

Efficacy Endpoints

The two studies have identical co-primary endpoints: (1) the mean change from Baseline to the end of treatment in IRLS Rating Scale total score; and (2) the proportion of subjects at the end of treatment who are "much improved" or "very much improved" on the Investigator-rated Clinical Global Impression (CGI) of Improvement.

The IRLS Rating Score is a 10-item self rating scale for assessing severity of RLS in scores 0 to 4 for each item. The maximum score for IRLS Rating Score is 40, which indicate extremely severe RLS symptoms. The investigator rated CGI-I is a 7-category measure from very much improved to very much worse.

Statistical Analysis Methods

The change from baseline in IRLS total score is to be analyzed by analysis of covariance (ANCOVA) including effects for pooled site, treatment, and the baseline value as a covariate. The treatment-by-pooled-site interaction is to be evaluated at 0.10 significance level and to be removed if not significant. The response to treatment from the Investigator-rated CGI of Improvement at the end of treatment is to be analyzed using a logistic regression model that included treatment and pooled site as explanatory factors.

The assumption of the normal distribution for the ANCOVA linear model for the primary analysis of IRLS Total Score is to be evaluated using a single Kolmogorov-Smirnov test at a 0.05 two-sided significance level. If the residual distribution is not consistent with a normal distribution, then the baseline adjusted and the baseline IRLS Total Score are both to be transformed to a rank score. An ANCOVA model is to be used for the primary analysis with rank transformed baseline adjusted IRLS Score as the outcome (dependent) variable and the baseline IRLS score as the covariate

The primary efficacy analysis is to be conducted on the modified ITT (MITT) population, which includes all patients in the Safety Population who also satisfies all of the following conditions: (i) completed the IRLS rating scale at baseline; and (ii) completed at least one on-treatment IRLS rating scale score during the treatment period. This population is to be analyzed as randomized.

The coprimary endpoints are each to be tested at the 0.05 significance level. Only if both tests are statistically significant, the study will be considered to have provided positive evidence of efficacy.

For study XP053, the primary comparison of interest is XP13512 1200 mg versus placebo. Comparison involving XP13512 600 mg vs. placebo is to be carried out as secondary analysis.

3.1.1.3 Study Results from XP052

3.1.1.3.1 Study Population Results

A sample size of 210 subjects (105 per treatment arm) was planned, and a total of 222 subjects were actually randomized: 114 subjects to XP13512 and 108 subjects to placebo. Two subjects, both randomized to XP13512 group, were excluded from the efficacy analyses. One of the subjects withdrew at the investigator's request before taking any study medication, and the other withdrew before any post-randomization assessments. Thus, the modified ITT patient population consisted of 112 subjects in XP13512 group and 108 in the placebo group. A summary of study completion status and primary reason for withdrawal for subjects in the Randomized Population is provided in Table 1. The study completion and the premature withdrawal rates were similar between treatment groups.

Overall, the primary reason for withdrawal was AEs (12 [5.4%] subjects). A total of 9 (7.9%) subjects in the XP13512 group withdrew due to AEs compared to 3 (2.8%) subjects in the placebo group. More subjects withdrew due to treatment failure in the placebo group (6 [5.6%] subjects) compared with the XP13512 group (0 subjects).

Table 1 Study Completion and Withdrawal – XP052 (Source: Table 7 of Sponsor's Study Report)

	Number (%) of Subjects		
	Placebo N=108	XP13512 N=114	Total N=222
Completion Status			
Completed	92 (85.2)	100 (87.7)	192 (86.5)
Prematurely Withdrawn	16 (14.8)	14 (12.3)	30 (13.5)
Primary Reason for Withdrawal			
Adverse event	3 (2.8)	9 (7.9)	12 (5.4)
Subject Withdrew Consent	3 (2.8)	4 (3.5)	7 (3.2)
Treatment Failure	6 (5.6)	0	6 (2.7)
Ineligibility (did not meet entry criteria)	2 (1.9)	0	2 (.9)
Termination of Study or Withdrawal of Subject by Sponsor[a]	0	1 (0.9)	1 (0.5)
Protocol Non-Compliance (after randomization)	1 (0.9)	0	1 (0.5)
Investigator Judgement[b]	1 (0.9)	0	1 (0.5)

Data Source: DSTable 1.1

Note: Disposition is calculated based on the number of randomized subjects.

a. Subject 140/2010 withdrew at the sponsor's request because of the subjects work schedule (shift work) which made them ineligible for the study, and the subject had not taken a dose of drug.

b. Subject 133/2005 was withdrawn at the request of the Investigator because the investigator judged the subject to be non-compliant with investigational product and was requesting to use a prohibited medication.

The two treatment groups were similar with regard to demographic characteristics. The mean age of subjects was 51.1 years (range 18-81 years). There was a predominance of females (59.7%), and the majority of subjects were White (96.8%).

Overall, the median duration of RLS symptoms was 9.9 years and the mean number of days RLS symptoms reported on the 7-Day Subject RLS Record prior to Baseline was 6 days. The majority (68.3%) of subjects had no previous RLS treatment.

3.1.1.3.2 Efficacy Results

Change from Baseline in IRLS Rating Scale

One of the co-primary efficacy endpoints was the change from Baseline in the IRLS Rating Scale total score at end of treatment. The IRLS Rating Scale mean total score at Baseline was similar in both treatment groups. To be eligible for the study, subjects were to have a total score of ≥15 on the IRLS Rating Scale total score at Baseline. At the end of the treatment period, both treatment groups had a large reduction in IRLS total scores with a mean reduction of 8.75 points for the placebo group and 13.23 points for the XP13512 1200 mg group. The adjusted treatment difference for the change from Baseline in the IRLS Rating Scale total score at Week 12 using

LOCF for the MITT Population was -4.0 (95% CI: -6.2, -1.9, p=0.0003) indicating a statistically significant difference in favor of XP13512. The following table provides mean change in IRLS total score with observed cases at each visit and at the end of the treatment period using LOCF.

Table 2 Change from Baseline in IRLS Total Score - XP052 (Source: Reviewer's Analysis)

	Change from IRLS Total Score									
	Base-Line	Visit3 Week1	Visit4 Week2	Visit5 Week3	Visit6 Week4	Visit7 Week6	Visit8 Week8	Visit9 Wk10	Visit10 Wk12	**Visit10 LOCF**
Placebo										
N	108	104	103	102	99	97	93	92	90	108
Mean	22.57	-4.61	-6.53	-7.15	-7.49	-8.00	-8.59	-9.33	-9.39	-8.75
SD	(4.91)	(7.30)	(6.64)	(7.19)	(7.97)	(7.38)	(7.62)	(8.50)	(8.10)	(8.63)
XP13512										
N	112	107	107	104	101	102	102	96	98	112
Mean	23.07	-11.19	-11.86	-12.25	-13.87	-12.91	-13.67	-14.75	-13.76	-13.23
SD	(4.86)	(7.84)	(8.14)	(8.59)	(7.94)	(8.78)	(7.49)	(8.50)	(8.67)	(9.21)
p-value		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	.0001	.0003

Proportion of Subjects who were Responders

The CGI-I Scale at the end of the treatment using LOCF is presented in Table 3. The proportions of responders using observed cases at each visit and LOCF at the end of the treatment period are provided in Table 4. At Week 12 using LOCF in the MITT Population, 76.1% of the subjects in the XP13512 group compared with 38.9% of the subjects in the placebo group were responders on the CGI-I Scale. The odds of having a score of much improved or very much improved in the XP13512 group was 5.1 times that in the placebo group (95% CI 2.8, 9.2). This difference was statistically significant in favor of XP13512 (p<.0001).

Table 3 CGI-I Score and Responder Rates at Week 12 - XP052 (Source: Table 15 of Sponsor's Study Report)

	Number (%) of Subjects				
	Placebo N=108	XP13512 1200 mg N=112	Odds ratio	95% CI	p-value
n	108	109			
Very much improved	20 (18.5)	55 (50.5)			
Much improved	22 (20.4)	28 (25.7)			
Minimally improved	21 (19.4)	10 (9.2)			
No change	39 (36.1)	15 (13.8)			
Minimally worse	6 (5.6)	0			
Much worse	0	1 (<1)			
Very much worse	0	0			
Total Responders	42 (38.9%)	83(76.1%)	5.1	2.8, 9.2	<0.0001

Table 4 Responder Rates at Each Visit – XP052 (Source: Reviewer's Analysis)

	CGI – XP052					
	Visit 3 Week 1	Visit 4 Week 2	Visit 6 Week 4	Visit 8 Week 8	Visit10 Week 12	**Visit10 LOCF**
Placebo						
N	105	103	99	93	90	108
# (%) Responders	26 (24.76%)	33 (32.04%)	43 (43.43%)	43 (46.24%)	39 (43.33%)	42 (38.89%)
XP13512 1200 mg						
N	107	106	100	102	95	109
# (%) Responders	62 (57.94%)	74 (69.81%)	78 (78.00%)	82 (80.39%)	75 (78.95%)	83 (76.15%)
p-value	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001

3.1.1.4 Study Results from XP053

3.1.1.4.1 Study Population Results

A total of 315 subjects were planned for the study and 325 subjects were actually randomized: 113 subjects to XP13512 1200 mg, 115 subjects to XP13512 600 mg, and 97 subjects to placebo. Three subjects were excluded from the Safety Population due to failed entry criteria or withdrawal of consent. All 3 subjects withdrew from the study prior to taking a dose of study drug. Another subject in the XP13512 600 mg group was excluded from the MITT Population. The subject reported an AE of somnolence that led to withdrawal from the study drug shortly following the Baseline Visit; no on-treatment IRLS Rating Scale scores were obtained for this subject. Thus, the MITT patient population comprises of 321 subjects: 96 subjects in placebo group, 114 subjects in XP13512 600 mg group, and 111 subjects in XP13512 1200 mg group.

A summary of study completion status and primary reason for withdrawal for subjects in the Randomized Population is provided in Table 5. Study completion rates were higher in the XP13512 groups (86.7% for XP13512 1200 mg, 90.4% for XP13512 600 mg) compared with the placebo group (79.4%). Overall, the most common reasons for withdrawal were AE (21 [6.5%] subjects) and “subject withdrew consent” (15 [4.6%] subjects). More subjects withdrew consent in the placebo group (8 [8.2%] subjects) compared with the XP13512 1200 mg (4 [3.5%] subjects) and XP13512 600 mg (3 [2.6%] subjects) groups.

Table 5 Summary of Subject Disposition - XP053 (Source: Table 7 of Sponsor’s Study Report)

	Number (%) of Subjects			
	Placebo N=97	XP13512 600 mg N=115	XP13512 1200 mg N=113	Total N=325
Completion Status				
Completed	77 (79.4)	104 (90.4)	98 (86.7)	279 (85.8)
Prematurely Withdrawn	20 (20.6)	11 (9.6)	15 (13.3)	46 (14.2)
Primary Reason for Withdrawal				
Adverse Event	6 (6.2)	7 (6.1)	8 (7.1)	21 (6.5)
Subject Withdrew Consent	8 (8.2)	3 (2.6)	4 (3.5)	15 (4.6)
Treatment Failure	3 (3.1)	0	0	3 (0.9)
Ineligibility (did not meet entry criteria)	0	0	2 (1.8)	2 (0.6)
Protocol Non-Compliance (after randomization)	1 (1.0)	0	1 (0.9)	2 (0.6)
Lost to Follow-Up	1 (1.0)	1 (0.9)	0	2 (0.6)
Termination of Study or Withdrawal of Subject by Sponsor[a]	1 (1.0)	0	0	1 (0.3)

Data Source: DSTable 6.1

Note: Disposition is calculated based on the number of randomized subjects.

a. Subject 197/3025 was withdrawn per the sponsor's request due to ineligibility (did not meet entrance criteria).

The 3 treatment groups were similar with regard to demographic characteristics. The mean age of subjects in the Safety Population was 48.9 years (range 21.0-77.0 years). There was a predominance of females (58.7%), and the majority of subjects were White (94.2%).

Overall, the median duration of RLS symptoms was 9.9 years, and the mean number of days RLS symptoms reported on the 7-Day Subject RLS Record prior to Baseline was 6.3 days. The majority (64.4%) of subjects had no previous RLS treatment.

3.1.1.4.2 Efficacy Results

Change from Baseline in IRLS Rating Scale

The IRLS Rating Scale mean total scores at Baseline were similar among the treatment groups. To be eligible for the study, subjects were to have a total score of ≥15 on the IRLS Rating Scale

at Screening and Baseline. At the end of the 12-week treatment period, all three treatment groups had a large reduction in mean IRLS Rating Scale total scores. The mean reductions in IRLS Rating Scale total score were 9.84, 13.82, and 12.95 points for placebo group, XP13512 600 mg group, and XP13512 1200 mg group, respectively. The primary comparison was between the XP13512 1200 mg and placebo groups at Week 12 (Visit 10) using LOCF. Both XP13512 dose groups showed statistically significant efficacy benefit over placebo group. The following table presents the mean change from baseline in IRLS Rating Scale total score at each visit (observed cases) and at the end of the treatment period (LOCF).

Table 6 Change from Baseline in IRLS Total Score – XP053 (Source: Reviewer's Analysis)

	Change from IRLS Total Score – XP053									
	Base-Line	Visit3 Week1	Visit4 Week2	Visit5 Week3	Visit6 Week4	Visit7 Week6	Visit8 Week8	Visit9 Wk10	Visit10 Wk12	**Visit10 LOCF**
Placebo										
N	96	88	91	87	84	83	81	74	74	96
Mean	23.81	-6.51	-7.80	-7.17	-8.62	-8.99	-8.09	-9.19	-10.97	-9.84
SD	(4.58)	(5.53)	(6.38)	(7.07)	(5.80)	(7.16)	(6.75)	(7.68)	(7.72)	(7.69)
600 mg										
N	114	110	110	105	104	102	102	103	101	114
Mean	23.11	-10.13	-11.13	-10.80	-11.44	-12.92	-12.64	-13.83	-14.17	-13.82
SD	(4.93)	(7.67)	(7.63)	(8.23)	(7.86)	(7.65)	(8.32)	(8.07)	(8.11)	(8.09)
p-value		<.0001	.0002	.0002	.0018	.0001	<.0001	<.0001	.0015	<.0001
1200 mg										
N	111	105	102	103	101	97	95	97	93	111
Mean	23.18	-9.25	-11.76	-12.36	-13.00	-12.69	-12.87	-13.02	-14.24	-12.95
SD	(5.32)	(8.03)	(8.78)	(8.99)	(9.22)	(9.85)	(8.50)	(9.49)	(8.74)	(9.12)
p-value		.0019	<.0001	<.0001	<.0001	.0012	<.0001	.0019	.0048	.0017

In this analysis of the IRLS total score, the treatment by pooled site interaction term was statistically significant. Further investigation indicated that the results from Site 149 were not consistent with results from other sites. The sponsor audited the site, and reported that it appeared to be in compliance with the protocol. Except for Site 149, effect is consistent across study sites.

Proportion of Subjects who were Responders

The other co-primary efficacy endpoint was the proportion of subjects with a score of much improved or very much improved (termed responders) on the investigator-rated CGI-I Scale at the end of treatment (Week 12 using LOCF). The proportion of responders was 77.48% in the XP13512 1200 mg group, 72.81% in the XP13512 600 mg group, compared with 44.79% of the subjects in the placebo group. The odds of having a score of much improved or very much improved in the XP13512 1200 mg group was 4.29 times that in the placebo group (95% CI: 2.338, 7.861; $p<0.0001$) and in the XP13512 600 mg group was 3.32 times that in the placebo group (95% CI: 1.841, 5.992; $p<0.0001$). The following table presents the proportion of subjects who were responders at each visit (observed cases) and at the end of the treatment period (LOCF).

Table 7 Responder Rate at Each Visit - XP053 (Source: Reviewer's Analysis)

	CGI – XP053					
	Visit 3 Week 1	Visit 4 Week 2	Visit 6 Week 4	Visit 8 Week 8	Visit 10 Week 12	**Visit10** **LOCF**
Placebo						
N	89	95	95	96	96	96
# (%) Responders	26 (29.21%)	36 (37.78%)	43 (45.26%)	41 (42.71%)	43 (44.79%)	43 (44.79%)
XP13512 600 mg						
N	108	112	113	114	114	114
# (%) Responders	54 (50%)	74 (66.07%)	71 (62.83%)	78 (68.42%)	83 (72.81%)	83 (72.81%)
p-value	.0030	<.0001	.0133	.0003	<.0001	<.0001
XP13512 1200 mg						
N	106	110	111	111	111	111
# (%) Responders	59 (55.66%)	74 (67.27%)	78 (70.27%)	77 (69.37%)	86 (77.48%)	86 (77.48%)
p-value	.0002	<.0001	.0004	.0001	<.0001	<.0001

3.1.2 Study XP060

3.1.2.1 Description of the Study

The primary objective for this study with withdrawal design was to assess the maintenance of efficacy of XP13512 1200 mg taken once daily in the long-term treatment of subjects with primary RLS.

This was a multicenter, maintenance of effect study in subjects with primary RLS. The study involved an initial 24-week single-blind (SB) period of treatment with 1200 mg/day XP13512. Subjects who completed the initial SB treatment period and met the responder criteria were then randomized to receive either XP13512 or placebo during the 12-week DB treatment period (Figure 1).

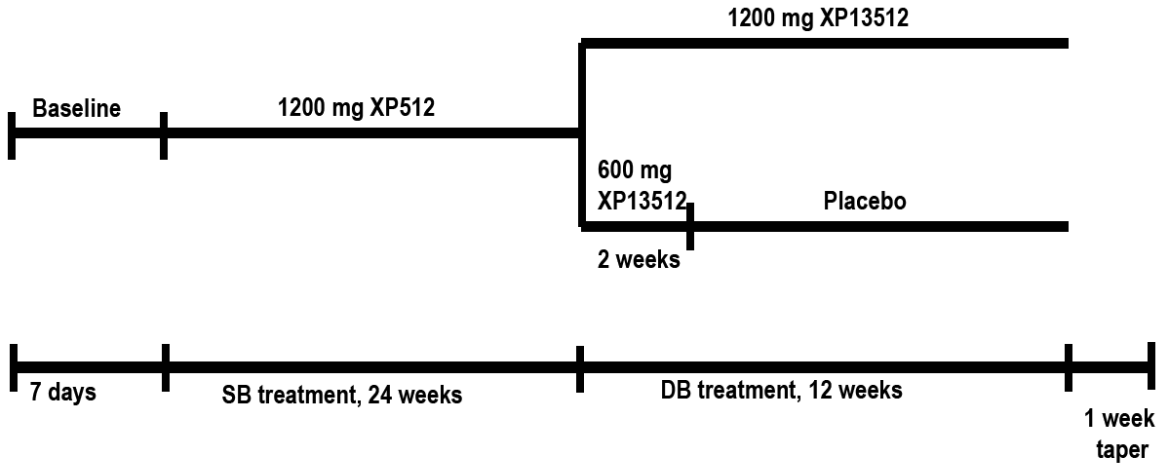


Figure 1 Study Design - XP060 (Source: Figure 1 of Sponsor's Study Report)

The responder criteria were as follows:

- total IRLS score decreased by 6 or more points relative to their Baseline score;
- total IRLS score decreased to less than 15;
- had an assessment of “much improved” or “very much improved” on the investigator-rated Clinical Global Impression of Improvement (CGI-I);
- stable on 1200 mg XP13512 dose for at least the month prior; and
- successfully completed the entire 24-week SB treatment period.

Subjects enrolled into the DB treatment period were randomized 1:1 to receive 1200 mg XP13512 or placebo.

A total of 180 subjects (90 subjects per arm) were planned to be randomized into DB period, and 194 subjects were actually randomized.

3.1.2.2 Efficacy Evaluation

The primary efficacy variable was the proportion of RLS subjects who relapsed, which was defined as worsening of RLS symptoms (defined below) or withdrawal due to lack of efficacy during the 12-week DB treatment period (the period from Randomization on Visit 14 [Week 24] through the end of treatment). Relapse was defined as:

- an increase (i.e., worsening) in the total IRLS score by at least 6 or more points relative to the subject's score at Randomization on Visit 14 (Week 24), achieving an IRLS score of at least 15 and an assessment of "much worse" or "very much worse" on the investigator-rated Clinical Global Impression of Change (CGI-C). In order for a subject to be defined as having achieved the endpoint of relapse, these criteria must have been met at 2 consecutive visits at least 1 week apart. The date of relapse was then defined as the first date at which the above criteria were met; or
- withdrawal due to lack of efficacy during the DB treatment period.

The primary analysis variable was to be analyzed by a logistic regression model, which included terms for treatment group, Visit 14 (Week 24) IRLS total score, and pooled study site.

3.1.2.3 Study Population Results

A total of 327 subjects were enrolled into the SB treatment period. Among the 221 subjects who completed the SB period of the study, 194 subjects met the responder criteria and were randomized into the DB treatment period: 96 subjects to XP13512 and 98 subjects to placebo. One subject withdrew without any efficacy assessment in the DB treatment period, and was therefore excluded from the ITT Populations.

In the XP13512 group, 4 (4.2%) subjects withdrew consent compared with 2 (2.0%) subjects in the placebo group. No one in the XP13512 group withdrew due to AE during the DB period. However, one subject in the XP13512 group withdrew during the taper period due to an AE which started in the DB period. Three (3.1%) subjects in the placebo group withdrew due to AEs, one of them had AE that started during the SB period.

Subjects who met the definition of relapse were not required to withdraw from the study. Four subjects in each treatment group withdrew due to lack of efficacy without meeting relapse criteria. An additional two subjects in the XP13512 treatment group met relapse criteria and withdrew due to lack of efficacy.

In the DB-ITT Population, the mean age of subjects was 52.2 years (range 23 to 82 years) in the placebo group and was 50.7 years (range 19 to 73 years) in the XP13512 group. The majority of subjects were female (53.6% in placebo and 64.6% in XP13512), and nearly all subjects (93.8% in placebo and 96.9% in XP13512) were white.

In the DB-ITT Population, the median duration of RLS symptoms in subjects randomized to the placebo group was 10.4 years and for subjects randomized to the XP13512 group was 7.1 years. The mean number of days in which RLS symptoms were reported on the 7-Day Subject RLS Record prior to Baseline was 6 days for both groups. The majority of subjects in the placebo and XP13512 groups had no previous RLS treatment (61.9% and 64.2%, respectively).

3.1.2.4 Efficacy Results

During the DB treatment period, 9.4% of subjects in the XP13512 group relapsed compared with 22.7% of subjects in the placebo group. The odds ratio for experiencing a relapse in RLS symptoms was 0.353 (XP13512 vs. placebo) (95% CI: 0.2, 0.8; p=0.0158). The difference in the proportion of subjects who experienced a relapse was significantly lower in the XP13512 treatment group compared to placebo.

Table 8 Proportion of Subjects who Experienced a Relapse During the Double-Blind Period – XP060 (Source: Table 15 of Sponsor's Study Report)

	Number (%) of Subjects				
	Placebo N=97	XP13512 N=96	Odds ratio[a]	95% CI	p-value
Subjects who Relapsed	22 (22.7)	9 (9.4)	0.353	(0.2, 0.8)	0.0158

Data Source: DS Table 7.1

a. From a logistic regression model including terms for treatment group, Visit 14 (Week 24) IRLS assessment, and pooled study site.

Of the 9 relapses in the XP13512 group, 5 (55.6%) met the IRLS/CGI criteria. The corresponding number for the placebo group was 18 (81.8%). In both XP13512 and placebo groups, 4 subjects withdrew due to lack of efficacy.

Multiple sensitivity analyses of efficacy were conducted to ensure that the outcome described by the primary analysis variable was a true measure of relapse. This sensitivity assessment is

important since there are multiple components to the relapse outcome, and the nature of relapse associated with RLS treatment withdrawal can be confused with a transient drug withdrawal or rebound effect, or a transient worsening of RLS symptoms.

Original data at all visits were examined thoroughly for possible sign of withdrawal effect. Among the placebo-treated subjects who met the relapse criteria, no one relapsed within 10 days after withdrawal of the drug. A total of 9 subjects relapsed about 2 weeks from the beginning of the double-blind treatment. Some of the 9 subjects stayed in the study. It appeared that those relapses were not likely to be from withdrawal effect.

The following table presents mean (median) of IRLS total score at the beginning of the double-blind period and IRLS total score and CGI-C rating scores at the last visit. Note that for subjects who relapsed, the scores represent the assessment at the visit for which they met the relapse criteria. Subjects must have had CGI-C score of 6 (much worse) or 7 (very much worse) to meet the criteria of relapse.

Table 9 IRLS Rating Scale and CGI-I during Double-Blind Period – XP060 (Source: Reviewer's Analysis)

	All Subjects		Relapsed Subjects	
	Placebo N=97	**XP13512 1200 mg N=96**	**Placebo N=22**	**XP13512 1200 mg N=9**
IRLS				
Baseline	5.30 (6.00)	5.10 (6.00)	5.32 (5.00)	7.88 (8.00)
Last Visit	9.72 (9.00)	7.40 (6.50)	18.59 (17.50)	20.44 (21.00)
Change	4.42 (2.00)	2.29 (0.00)	13.27 (13.50)	12.56 (13.00)
CGI-C	4.32 (4.00)	3.92 (4.00)	6.14 (6.00)	6.11 (6.00)

3.1.3 Study XP081

3.1.3.1 Description of the Study

This study was conducted to measure XP13512 released gabapentin pharmacokinetics and to assess if there was a XP13512 dose/exposure-response relationship for the treatment of patients with RLS. The objective of the study was to assess the relationship between the gabapentin exposure produced by four dose levels of XP13512 (600 mg, 1200 mg, 1800 mg, and 2400 mg) and the relief of symptoms in patients with Restless Legs Syndrome (RLS).

This was a multicenter, randomized, double-blind, placebo-controlled, parallel-group study, comparing 4 doses of XP13512 with placebo given once daily to subjects with RLS. Eligible subjects were randomized in equal numbers into 1 of 5 treatment groups (XP13512 600 mg, 1200 mg, 1800 mg, or 2400 mg or placebo). The Double-Blind Treatment began with a 9-day titration period and continued for 12 weeks. The randomization was stratified by study site and Baseline IRLS total score category ($\leq$22 versus >22).

3.1.3.2 Efficacy Evaluation

There was no assignment of primary or secondary endpoints in this study as this study was conducted to measure XP13512 released gabapentin pharmacokinetics. Multiple efficacy endpoints were evaluated in this study including:

- Change from Baseline in the IRLS Rating Scale total score after 1 week, 4 weeks, and at the end of treatment;
- Proportion of subjects responding to treatment where a "response" is a report of "very much improved" or "much improved" on the investigator-rated CGI-I after 1 week, 4 weeks, and at the end of treatment;

The planned analyses of efficacy variables were limited to the presentation of descriptive statistics by dose group.

3.1.3.3 Study Populations Results

A total of 217 subjects were randomized: 48 subjects to XP13512 600 mg, 45 subjects to XP13512 1200 mg, 38 subjects to XP13512 1800 mg, 45 subjects to XP13512 2400 mg, and 41 subjects to placebo.

A total of 6 subjects were excluded from the MITT Population because they did not have any post-Baseline assessments for the IRLS Rating Scale.

An overall 73.3% of subjects completed the study. The study completion and the premature withdrawal rates were similar among the treatment groups. More subjects withdrew due to an AE in the XP13512 treatment groups (7.9% to 13.3%) compared with placebo (2.4%). Conversely, more subjects withdrew from the study due to withdrawing consent in the placebo group (14.6%) compared with the XP13512 treatment groups (600 mg=10.4%; 1200 mg=8.9%; 1800 mg=2.6%; and 2400 mg=none). The number of subjects withdrawing due to other reasons was low and similar across treatment groups.

The mean age of the population ranged from 38 for the XP13512 1800 mg group to 48 for the XP13512 600 mg group. There were a greater percentage of males in the XP13512 1200 mg group (48.9%) compared to other groups (range 29.3% to 35.6%). The majority of subjects were White (95.4%).

Overall, the mean IRLS total score was similar between treatment groups (range 22.5 to 23.9). To be eligible for the study, subjects were to have a total score of $\geq$ 15 on the IRLS at baseline. The median duration of RLS symptoms was 7.9 years: the placebo group had the shortest median duration (4.8 years) and the XP13512 1200 mg group had the longest median duration (10.7 years). The mean number of days RLS symptoms was reported on the 7-Day Subject RLS Record prior to baseline was approximately 6.0 days in all treatment groups, which may be lower than the actual value because subjects were not required to complete all 7 days of the diary once they reached 4 days of RLS symptoms. The majority (65.4%) of subjects overall had no previous

RLS treatment, with the lowest proportion of such subjects in the XP13512 1800 mg group (54.1%) and the highest proportion in the XP13512 1200 mg group (72.1%).

3.1.3.4 Efficacy Results

There was no assignment of primary or secondary efficacy endpoints in this study. The reviewer has focused on the same efficacy endpoints as co-primary variables in the two pivotal studies for which the data are available, and performed the same analyses that were applied to the pivotal studies. It should be noted that this study was not designed to make statistical comparisons between individual dose groups and placebo, and the sample size of each treatment group is much smaller than the ones in the pivotal studies. The purpose of the analyses was to find supporting evidence for dose response, particularly for XP13512 600 mg, which appeared to be efficacious in Study XP053 and not studied in XP052.

Subjects in different XP13512 treatment groups reached their target doses on different study days based on the titration schedule. Subjects in the 600 mg group reached their target dose on Day 1. Subjects in the 1200 mg, 1800 mg, and 2400 mg groups reached their target doses on Days 4, 7, and 10, respectively. Therefore, subjects in the 2400 mg group had not reached their target dose at the Week 1 assessment.

Change from Baseline in IRLS Rating Scale

The following table presents the mean change of IRLS total score from baseline at each visit (observed cases) and at the end of the treatment (LOCF). The nominal p-values from pairwise comparisons of each XP13512 dose group versus placebo without multiplicity adjustment are provided for the endpoint at Week 12 LOCF.

Table 10 IRLS Total Scores - XP081 (Source: Reviewer's Analysis)

	Base line	Change from Baseline in IRLS Total Score								
		Visit3 Week1	Visit4 Week2	Visit5 Week3	Visit6 Week4	Visit7 Week6	Visit8 Week8	Visit9 Wk10	Visit10 Wk12	Visit10 LOCF
Placebo										
N	40	34	32	36	34	31	32	33	30	40
Mean	22.45	-5.62	-6.84	-8.06	-8.71	-7.52	-9.41	-9.09	-9.17	-9.28
SD	(5.32)	(7.30)	(8.85)	(8.28)	(7.76)	(9.65)	(9.79)	(9.63)	(8.37)	(8.13)
600 mg										
N	47	45	44	42	38	38	36	34	33	47
Mean	23.87	-8.91	-11.20	-10.81	-12.42	-11.87	-13.58	-13.00	-15.67	-13.81
SD	(5.33)	(7.69)	(8.29)	(9.48)	(9.00)	(9.32)	(9.85)	(8.70)	(8.00)	(9.48)
p-value										.0394
1200 mg										
N	43	41	39	39	39	32	31	32	27	43
Mean	23.91	-10.10	-11.45	-12.38	-13.13	-14.88	-13.06	-14.75	-16.22	-13.81
SD	(5.49)	(7.68)	(8.07)	(8.87)	(7.44)	(8.78)	(9.78)	(8.14)	(9.74)	(9.84)
p-value										.0445
1800 mg										
N	37	37	35	35	30	32	33	32	33	37
Mean	23.62	-10.59	-13.89	-14.23	-15.13	-16.59	-15.24	-14.91	-15.15	-13.95
SD	(4.25)	(8.42)	(8.05)	(8.28)	(8.67)	(7.82)	(7.89)	(8.85)	(8.13)	(8.70)

p-value										.0256
2400 mg										
N	44	42	43	39	37	34	34	35	31	44
Mean	23.34	-9.02	-12.84	-11.92	-13.38	-15.24	-14.41	-13.74	-15.35	-12.86
SD	(5.70)	(7.10)	(8.39)	(7.21)	(7.57)	(7.38)	(9.08)	(8.24)	(7.86)	(9.52)
p-value										.0895

At the baseline, all treatment groups had similar IRLS total scores. All treatment groups showed improvement in IRLS total scores from Week 1 assessment, and more improvement was gained and maintained throughout the treatment. At the end of the treatment period using LOCF, all XP13512 dose groups showed numerically larger mean reduction in IRLS total score than the placebo group. The magnitude of the improvement was similar in all XP13512 groups. The difference among all treatment groups did not reach statistical significance (p=.1581) in the overall statistical testing using the same ANCOVA model that applied in the two pivotal studies (XP052 and XP053). When all XP13512 dose groups were compared to placebo group using Dunnett's adjustment for multiplicity, none of the dose group reached statistical significance of 0.05 as well, though the pairwise comparison without multiplicity adjustment showed that all but XP13512 2400 mg dose groups were statistically significantly different from placebo group at significance level of 0.05. The nominal p-values were 0.0394, .0445, .0256, .0895 for XP 600 mg, 1200 mg, 1800 mg, and 2400 mg, respectively, compared to placebo group.

In this study, the baseline level of IRLS total score, the magnitude of change from baseline to the end of the treatment, and treatment difference were all similar to the levels found in the two pivotal studies. The sample size of each treatment group was about half of the sizes of the pivotal studies, which could be the reason of resulted insignificance of statistical testing.

Proportion of Subjects who were Responders

A summary of the proportions of responders (much improved or very much improved) in the investigator-rated CGI-I Scale at each visit (observed cases) and at Week 12 using LOCF is presented in Table 11. The proportion of responders (very much improved or much improved) on the CGI-I Scale at Week 12 using LOCF in the MITT Population was numerically greater in the XP13512 600 mg, 1200 mg, 1800 mg, and 2400 mg groups (63.8%, 65.1%, 73.0%, and 81.8%, respectively) compared with the placebo group (45.0%). The same logistic model that applied to the two pivotal studies was applied first. However, the validity of the logistic model fit was in question, although the statistical significance in an overall test among all treatment groups was reached. The CMH method was then applied in place of logistic model, and significance of the treatment differences was confirmed. The nominal p-values provided in the following table are from pairwise comparisons using Cochran-Mantel-Haenszel (CMH) test without multiplicity adjustment.

Table 11 Responder Rate - XP081 (Source: Reviewer's Analysis)

	CGI – XP081					
	Visit 3 Week 1	Visit 4 Week 2	Visit 6 Week 4	Visit 8 Week 8	Visit 10 Week 12	Visit10 LOCF
Placebo						
N	35	32	34	31	29	40
# (%) Responders	11 (31.43%)	10 (31.25%)	17 (50.00%)	15 (48.39%)	13 (44.83%)	18 (45.00%)
XP13512 600 mg						
N	46	43	37	36	33	47
# (%) Responders	23 (50.00%)	24 (55.81%)	23 (62.16%)	23 (63.89%)	24 (72.73%)	30 (63.83%)
Nominal p-value						.0801
XP13512 1200 mg						
N	40	39	39	31	26	43
# (%) Responders	23 (57.50%)	27 (69.23%)	27 (69.23%)	25 (80.65%)	20 (76.92%)	28 (65.12%)
Nominal p-value						.0671
XP13512 1800 mg						
N	36	35	30	33	31	37
# (%) Responders	23 (63.89%)	27 (77.14%)	20 (66.67%)	27 (81.82%)	25 (80.65%)	27 (72.97%)
Nominal p-value						.0134
XP13512 2400 mg						
N	42	43	36	34	31	44
# (%) Responders	21 (50.00%)	33 (76.74%)	28 (77.78%)	28 (82.35%)	28 (90.32%)	36 (81.82%)
Nominal p-value						.0005

A large number of missing values occurred in the above analysis. Among the 217 subjects, 6 did not have post-baseline value and were excluded. Another 6 subjects did not have post-baseline assessments (but did have assessments at Visit 1 and 2) were not excluded. In addition, 57 subjects had their values at week 12 for IRLS score carried from previous assessments. Therefore, LOCF may not present the best picture. Observed case analysis may reflect better actual means.

3.1.4 Study XP083

3.1.4.1 Description of the Study

XP083 was a randomized, double-blind, active- and placebo-controlled safety study to assess simulated driving performance after two weeks of treatment with XP13512 in patients with RLS. Eligible subjects were randomized in a 1:1:1:1 ratio to 1 of the following 4 treatment groups:

A) XP13512 Placebo + Diphenhydramine Placebo (Pbo)

B) XP13512 1200 mg/day + Diphenhydramine Placebo

C) XP13512 1800 mg/day + Diphenhydramine Placebo

D) XP13512 Placebo + 50 mg Diphenhydramine (Pbo/DPH)

Subjects in group D received 50 mg diphenhydramine on Day 16 to assess the effects of an agent known to have sedative properties. Subjects in groups A, B, and C received a diphenhydramine-matching placebo on Day 16.

After the Baseline period, subjects returned to the study site to complete 2 Baseline Visits: Day -1 in the evening and Day 1 in the morning. The purpose of conducting 2 Baseline Visits was to provide Baseline values for driving, alertness, cognition, and other efficacy assessments at a comparable time of day when assessments would be performed during the Treatment Period. Assessments at the Day -1 Baseline Visit were conducted in the evening to be used for comparison to subsequent assessments performed in the evening (Day 14 and Day 16), during the Treatment Period. The Day 1 Baseline Visit assessments occurred the following morning, to be used for comparison to subsequent assessments performed in the morning (Day 15), during the Treatment Period. A schematic diagram of the overall study design is presented in Figure 2.

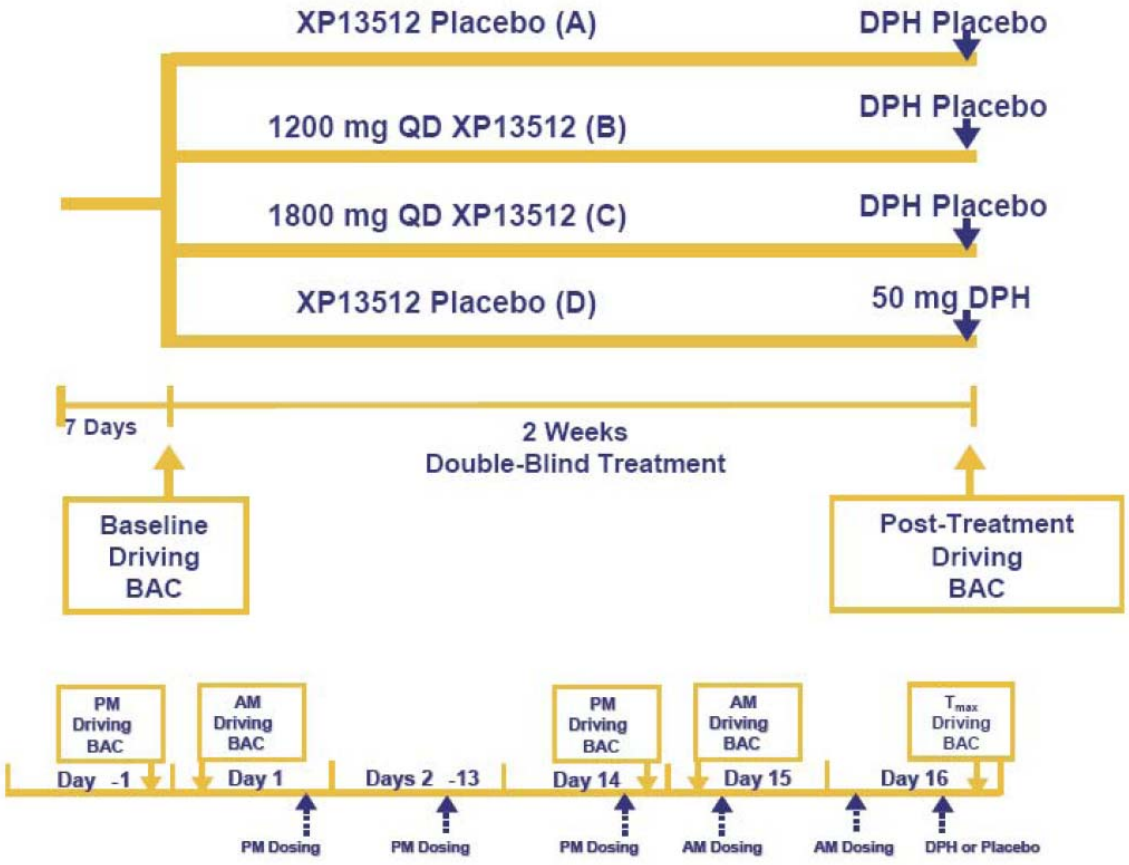


Figure 2 Study Design - XP083 (Source: Figure 1 of sponsor's Study Report)
*BAC stands for brief assessment of cognition.

After 14 days of treatment, the driving and cognitive assessments were conducted in the evening (7 PM to 9 PM), 2 to 4 hours after the 5 PM dosing, and the following morning (7 AM to 9 AM) to correlate with when subjects are likely to be ambulatory and engaged in driving. Two Baseline

driving assessments (1 in the evening [5 PM to 8 PM] and 1 the following morning [7 AM to 9 AM]) were conducted at the corresponding time points on Days -1 and 1 prior to the treatment.

Under the proposed prescribing condition (5 PM dosing), the time to maximal drug concentration (Tmax) of XP13512 would occur in the middle of the night, while most subjects would be sleeping (i.e., 12 AM to 1 AM). Therefore, in order to explore the effect of XP13512 on driving performance and alertness at the time of maximal concentration, the dosing time was adjusted to 10 AM to 11 AM on Days 15 and 16 to allow assessments at Tmax in the evening on Day 16.

The timing of the 3 post-dose driving assessments in relation to the dosing time would provide different gabapentin plasma concentrations, with the Day 14 evening assessment correlating with a relatively low concentration, the Day 16 estimated Tmax assessment with a high concentration, and the Day 15 morning assessment in between the 2 concentrations.

This study utilized an active control of diphenhydramine. Diphenhydramine is a non-prescription antihistamine commonly used to treat allergies. Use of diphenhydramine is associated with the side effect of drowsiness and its use has been shown to have an effect on driving and cognition.

3.1.4.2 Driving Simulator

In this study, STISIM Drive™, a fixed-platform PC -based driving simulation system was used. The simulator setup and placement of controls were similar to an actual car.

During simulated driving tests, subjects were advised to drive within the simulated environment according to local laws and to observe the posted speed limit of 55 mph. Simulator speakers provided audio feedback of characteristic road noise associated with acceleration, deceleration, rapid braking and crashes. In addition, subjects who exceeded 65 mph or fell below 45 mph received automated audio messages reminding them of the 55 mph speed limit. These prompts were repeated at 30-second intervals, if the subject remained outside of the established threshold values (i.e., < 45 mph or > 65 mph).

The simulated driving assessment included a 5-minute practice drive, a 2-minute brake reaction time test, and a 1-hour test drive. The 1-hour test drive consisted of a rural 2-lane highway with multiple gradual curves, occasional hills, and several oncoming vehicles approximately every 10 minutes. The data were collected in 6 sequential 10-minute time blocks (referred to as epochs) for the 1-hour test drive.

3.1.4.3 Study Population Results

A total of 130 subjects were randomized, and 122 subjects had at least one baseline and End of Study (Days 14 to 16) driving assessment to be included in MITT patients population: 33 subjects to placebo, 28 subjects to XP13512 1200 mg, 33 subjects to XP13512 1800 mg, and 28 subjects to Pbo/DPH.

The mean age of the population was 49.6, 46.8, 49.3 and 40.6 for the placebo group, XP13512 1200 mg group, XP13512 1800 mg group, and Diphenhydramine Placebo group, respectively.

There were a greater percentage of males in the XP13512 1800 mg group (50.0 %) compared to other groups (range 32.3% to 41.2%). Nearly all subjects were White (99.2%).

3.1.4.4 Simulated Driving Performance

Lane Position Variability

The primary assessment for this study was the change from Baseline (Day -1) in lane position variability (LPV) measured by simulated driving performance at Day 16 (estimated Tmax). Under normal prescribing conditions (5 pm dosing), Tmax would occur while most subjects were asleep. To assess the effect of XP13512 at its maximal concentration, the dosing regimen was adjusted such that subjects took study drug between 10 and 11 AM on Day 15 and Day 16 instead of 5 PM in the evening in the previous 14 days. This would allow assessment of driving performance at Tmax to occur in the early evening on Day 16.

Lane position variability at Baseline (Day -1) and Day 16 are presented in Table 12. Lane position variability (SD) was 1.40 (0.32), 1.46 (0.32), 1.37 (0.20), and 1.36 (0.25) ft for the placebo, XP13512 1200 mg, XP13512 1800 mg, and Pbo/DPH groups at Baseline (Day -1). At Day 16 (estimated Tmax) LPV was greater for the XP13512 1200 mg (1.61 ft), the XP13512 1800 mg (1.52 ft), and the Pbo/DPH groups (1.52 ft) compared with the placebo group (1.26 ft).

Table 12 Lane Position Variability Change for Day 16 - XP083 (source: Table 11 of sponsor's Study Report)

	Pbo	XP13512 1200 mg	XP13512 1800 mg	Pbo/DPH [a]		
	N=33	N=28	N=33	N=28		ANOVA [b]
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	95% CI for Mean	95% CI for LS Mean
Baseline (Day -1)	1.40 (0.32)	1.46 (0.32)	1.37 (0.20)	1.36 (0.25)		
Day 16	1.26 (0.31)	1.61 (0.48)	1.52 (0.37)	1.52 (0.50)		
Change from Baseline to Day 16						
Mean (SD)	–0.11 (0.17)	0.15 (0.38)	0.15 (0.27)	0.16 (0.40)		
LS Mean (SE)	-0.10 (0.06)	0.15 (0.06)	0.15 (0.06)	0.16 (0.06)		
XP13512 1200 mg – Pbo					0.10, 0.41	0.08, 0.42
XP13512 1800 mg – Pbo					0.14, 0.37	0.09, 0.41
Pbo/DPH - Pbo					0.10, 0.43	0.09, 0.42
XP13512 1200 mg – Pbo/DPH					-0.22, 0.20	
XP13512 1800 mg – Pbo/DPH					-0.18, 0.17	

Data Source: DSTable 8.4 and DSTable 9.4

a. Pbo/DPH group received diphenhydramine on Day 16 only.
b. Analysis was based on a repeated measures ANOVA model with fixed effects for treatment group, pooled site, visit, and treatment group by visit.

Lane position variability change for Day 14 and Day 15 assessments are summarized in Table 13. The Pbo/DPH group received placebo on Day 14 and Day 15. Compared to corresponding

baseline, XP13512 1200 mg group had large increase in LPV while other groups had little changes.

Table 13 Lane Position Variability on Day 14 and Day 15 – XP083 (Source: Table 12 of Sponsor's Study Report)

	Pbo	XP13512 1200 mg	XP13512 1800 mg	Pbo/DPH [a]	95% CI for Mean	ANOVA [b]
	N=33	N=28	N=33	N=28		
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)		95% CI for LS-Mean
Baseline (Day -1)	1.40 (0.32)	1.46 (0.32)	1.37 (0.20)	1.36 (0.25)		
Day 14	1.34 (0.38)	1.62 (0.62)	1.36 (0.38)	1.29 (0.26)		
Change from Baseline (Day -1) to Day 14						
Mean	-0.06 (0.17)	0.17 (0.43)	-0.01 (0.28)	-0.08 (0.15)		
LS Mean	-0.06 (0.05)	0.17 (0.05)	-0.01 (0.05)	-0.08 (0.05)		
XP13512 1200 mg – Pbo					0.06, 0.39	0.09, 0.37
XP13512 1800 mg – Pbo					-0.06, 0.17	-0.08, 0.19
Baseline (Day 1)	1.35 (0.28)	1.49 (0.36)	1.40 (0.29)	1.45 (0.35)		
Day 15	1.35 (0.31)	1.62 (0.45)	1.44 (0.46)	1.34 (0.28)		
Change from Baseline (Day 1) to Day 15						
Mean	-0.01 (0.14)	0.13 (0.40)	0.02 (0.32)	-0.10 (0.19)		
LS Mean	-0.01 (0.05)	0.13 (0.05)	0.02 (0.05)	-0.10 (0.05)		
XP13512 1200 mg – Pbo					-0.01, 0.29	-0.00, 0.28
XP13512 1800 mg – Pbo					-0.10, 0.15	-0.12, 0.16

Data Source: DSTable 8.4 and DSTable 9.4

a. Pbo/DPH group received diphenhydramine on Day 16 only.

b. Analysis was based on a repeated measures ANOVA model with fixed effects for treatment group, pooled site, visit, and treatment group by visit.

In this simulated driving performance study, means and modeling did not provide a good picture of whether or not and how subjects were affected by the study drug. A number of subjects who had extremely large values were driving the means in a significant degree. The standard deviation (SD), an important measure of subject-to-subject differences in LPV, provided some information of large discrepancies of subject-to-subject differences among the treatment group.

In further examination of the data, the reviewer found that most subjects had very little change in LPV at all assessments while about 10% to 20% of the XP13512 and Diphenhydramine - treated subjects had a quite substantial increase in LPV at the End of Study assessments. The following table shows that a majority of subjects in any group had a decrease in LPV or a mean increase of no more than 0.10 at any assessment. However, 20 subjects, none of them in the placebo group, had an increase of at least 0.30 in LPV at Day 16 assessment from their corresponding baseline value, 3 of them (1 in each of the XP13512 groups and 1 in Pbo/DPH group) exhibited extremely large increase in LPV.

Table 14 Number of Subjects who have Change of LPV in the Specified Range (Source: Reviewer's Analysis)

	Placebo	XP13512 1200 mg	XP13512 1800 mg	Pbo/DPH
Day 14 Change				
≤ -0.10	13	8	9	11
-0.10 ~ 0.10	17	10	17	13
0.10 ~ 0.30	1	2	5	2
0.30 ~ 0.50	1	3	1	0
0.50 ~ 1.00	1	4	0	0
> 1.00	0	1 (1.77 /1207001)	1 (1.24/1357005)	0
Day 15 Change				
≤ -0.10	8	6	11	13
-0.10 ~ 0.10	17	11	14	11
0.10 ~ 0.30	7	4	3	2
0.30 ~ 0.50	0	2	2	1
0.50 ~ 1.00	0	4	0	0
> 1.00	0	1 (1.66/1827001)	1 (1.35/1357005)	0
Day 16 Change				
≤ -0.10	12	7	3	5
-0.10 ~ 0.10	16	9	13	13
0.10 ~ 0.30	2	6	10	3
0.30 ~ 0.50	0	1	5	2
0.50 ~ 1.00	0	4	1	4
> 1.00	0	1 (1.23/1237004)	1 (1.16/2217007)	1 (1.69/2307005)

Although it occurred in a relatively small number of subjects, the pattern of such a large increase in LPV clearly indicated that it is drug related. In particular, the Pbo/DPH group had no subject who had an increase of more than 0.5 at any assessment when subjects were treated with placebo only. The group had 5 such subjects at Day 16 assessment when subjects took Diphenhydramine.

The reviewer did not find evidence that such large increase in LPV was related to age, gender, baseline LPV, baseline IRLS score, or caused by protocol violation. The data did not provide enough evidence of dose response either.

Simulated Crashes

A simulated crash was defined as a collision with an oncoming car or obstacle (e.g., tree), or when the distance to the center line was greater than 18 ft on either side of the road. The proportion of subjects with simulated crashes and the number of simulated crashes are summarized by treatment group in Table 15 and Table 16.

At each of the Baseline (Day -1 or Day 1) assessments, a greater proportion of subjects in the XP13512 1200 mg group experienced simulated crashes compared with the placebo, XP13512 1800 mg, and Pbo/DPH groups.

At the Day 14 [PM] assessment, the number or proportion of subjects who had simulated crashes was greater for the XP13512 1200 mg group (6 [21.4%]) when compared with the other 3 groups. Most subjects had 1 to 3 simulated crashes. Three subjects in the XP13512 1200 mg group each had 4, 5, and 13 crashes, respectively.

At the Day 15 [AM] assessment, a total of 10 subjects (35.7%) in the XP13512 1200 mg group experienced simulated crashes, an increase from 4 subjects (14.3%) at Baseline (Day 1). Seven of them had 1 to 2 simulated crashes, 2 subjects had 4 crashes, and 1 subject had 13 simulated crashes. The placebo and XP13512 1800 mg group each had 1 subject with 1 simulated crash. No subjects had simulated crashes in the Pbo/DPH group.

At the Day 16 (estimated Tmax) assessment, no subjects in the placebo group experienced simulated crashes, whereas all the active treatment groups had an increase from Baseline (Day -1) in the number of subjects with simulated crashes, with 8 (28.6%) in the XP13512 1200 mg group, 6 (18.2%) in the XP13512 1800 mg group, and 3 (10.7%) in the Pbo/DPH group. Most subjects had only 1 or 3 simulated crashes. One subject in the XP13512 1200 mg group and 1 subject in the Pbo/DPH group had 4 simulated crashes. One subject each in the XP13512 1200 mg and 1800 mg groups experienced 17 and 13 simulated crashes, respectively.

Table 15 Number of Subjects with Simulated Crashes – XP083 (Source: Table 14 of Sponsor's Study Report)

		Pbo	XP13512 1200 mg	XP13512 1800 mg	Pbo/DPH [a]
		N=33	N=28	N=33	N=28
Number of Subjects with Crashes, n (%)					
	Day -1	3 (9.1)	6 (21.4)	3 (9.1)	2 (7.1)
	Day 1	1 (3.1)	4 (14.3)	3 (9.4)	3 (11.1)
	Day 14	4 (12.1)	6 (21.4)	1 (3.0)	1 (3.6)
	Day 15	1 (3.0)	10 (35.7)	1 (3.2)	0 (0)
	Day 16	0 (0)	8 (28.6)	6 (18.2)	3 (10.7)

Data Source: DSTable 8.7.1

a. Pbo/DPH group received diphenhydramine on Day 16 only.

Table 16 Distribution of Subjects by Number of Simulated Crashes (Source: Table 15 of Sponsor's Study Report)

	Pbo N=33	XP13512 1200 mg N=28	XP13512 1800 mg N=33	Pbo/DPH [a] N=28
Day -1, Overall (0 to 60 minutes), n (%)				
N	33	28	33	28
1 Crash	2 (6.1)	4 (14.3)	3 (9.1)	2 (7.1)
2 Crashes	0	1 (3.6)	0	0
3 Crashes	1 (3.0)	1 (3.6)	0	0
Day 1, Overall (0 to 60 minutes), n (%)				
N	32	28	32	27
1 Crash	0	2 (7.1)	3 (9.4)	3 (11.1)
2 Crashes	1 (3.1)	0	0	0
3 Crashes	0	2 (7.1)	0	0
Day 14, Overall (0 to 60 minutes), n (%)				
N	33	28	33	28
1 Crash	2 (6.1)	1 (3.6)	0	1 (3.6)
2 Crashes	2 (6.1)	2 (7.1)	0	0
3 Crashes	0	0	1 (3.0)	0
4 Crashes	0	1 (3.6)	0	0
5 Crashes	0	1 (3.6)	0	0
13 Crashes	0	1 (3.6)	0	0
Day 15, Overall (0 to 60 minutes), n (%)				
N	33	28	31	28
1 Crash	1 (3.0)	4 (14.3)	1 (3.2)	0
2 Crashes	0	3 (10.7)	0	0
4 Crashes	0	2 (7.1)	0	0
13 Crashes	0	1 (3.6)	0	0
Day 16, Overall (0 to 60 minutes), n (%)				
N	30	28	33	28
1 Crash	0	5 (17.9)	3 (9.1)	2 (7.1)
3 Crashes	0	1 (3.6)	2 (6.1)	0
4 Crashes	0	1 (3.6)	0	0
5 Crashes	0	0	0	1 (3.6)
13 Crashes	0	0	1 (3.0)	0
17 Crashes	0	1 (3.6)	0	0

Data Source: DSTable 8.7.1

Note: Number of crashes over the indicated time interval, defined as a collision with an oncoming car or obstacle (e.g., tree), or when the vehicle deviated from the center line greater than 18 ft. on either side of the road.

a. Pbo/DPH group received diphenhydramine on Day 16 only.

Four subjects, 3 in the XP13512 1200 mg group and 1 in the XP13512 1800 mg group, experienced multiple simulated crashes (> 5). The sponsor reported that there were no notable findings in their medical history and physical or neurological examinations. These subjects did not take any concomitant medications. There were no AEs reported for these 4 subjects on any of the simulated driving assessment days. Subjects either reported no change or improvement in sleep from baseline to the nights prior to the simulated driving assessments based on the PghSD. Some demographic characteristics, efficacy, alertness measures, and plasma gabapentin levels for these 4 subjects who experienced multiple simulated crashes are summarized in Table 17.

Table 17 Characteristics of Subjects with Multiple Crashes (Source: Table 16 of Sponsor's Study Report)

		XP13512 1200 mg			XP13512 1800 mg
Subject ID		123/7004	182/7001	221/7003	221/7007
Age/Gender		56/M	49/M	45/F	57/F
Day with Multiple Crashes		16	15	14	16
Other Days with Crashes		14, 15	-1, 14, 16	-1, 1, 15	N/A
ESS	Day -1	24	12	3	19
	Day 14	8	15	2	8
IRLS Rating Scale Score	Day -1	20	22	27	22
	Day 14	22	27	11	18
VAS Alertness (Pre/Post Driving Score)	Day -1 or Day 1 Corresponding to Crash Day	94/88	59/76	100/46	77/52
	Day with Multiple Crashes	88/17	26/29	44/34	81/11
Plasma Gabapentin Level (ng/mL)	Day 14	380	1470	5800	2690
	Day 15	5740	2350	2330	5050
	Day 16	5260	3340	5130	13000

Data Source: DSListing 3, DSListing 17.1, DSListing 18, DSListing 20.2, DSListing 20.3, DSListing 25.2, DSListing 25.3, DSListing 25.4, and DSListing 26.

Speed variability was also examined. The reviewer found that the speed variability data was less reliable in drawing any conclusions about the drug effect because of large subject-to-subject variations exhibited at both baseline and End-of-Study simulated driving tests. The speed variability data is not presented in this review.

3.2 Evaluation of Safety

Refer to Clinical Review by Dr. Goldstein for Evaluation of Safety.

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Race and Age

Analyses of efficacy by gender and age group were performed for the two pivotal studies XP052 (Table 18) and XP053 (Table 19). The treatment effect appeared to be larger in female subjects than in male subjects, and there was little treatment difference in males in Study XP053. The baseline IRLS total scores for the males and females were similar (not shown).

In XP052, subjects with older age appeared to have larger treatment effect than the subject with younger age, but such difference is not observed in XP053.

Table 18 Mean Change in IRLS Total Score and Responder Rates by Gender and Age Group – XP052 (Source: Reviewer's Analysis)

	Placebo	XP13512 1200 mg	Nominal p-value
IRLS Total Score			
Gender			
Male			
N	43	46	
Mean (SD)	-6.91	-11.80	.0021
Female			
N	65	66	
Mean (SD)	-9.97 (9.35)	-14.23 (10.17)	.0236
Age (year)			
≤ 50			
N	54	50	
Mean (SD)	-8.65 (9.36)	-11.48 (9.54)	.1762
> 50			
N	54	62	
Mean (SD)	-8.85 (7.92)	-14.65 (8.76)	.0004
CGI			
Gender			
Male			
N	43	45	
# (%) Responders	9 (20.93%)	30 (66.67%)	<.0001
Female			
N	65	64	
# (%) Responders	33 (50.77%)	53 (82.81%)	.0004
Age			
≤ 50			
N	54	49	
# (%) Responders	23 (42.59%)	34 (69.39%)	.0083
> 50			
N	54	60	
# (%) Responders	19 (35.19%)	49 (81.67%)	<.0001

Table 19 Mean Change in IRLS Total Score and Responder Rate by Gender and Age Group – XP053 (Source: Reviewer's Analysis)

	Placebo	XP13512 600 mg	XP13512 1200 mg
IRLS Total Score			
Gender			
Male			
N	39	48	46
Mean (SD)	-10.45 (6.83)	-11.90 (8.06)	-10.43 (8.52)
p-value		.3061	.4557
Female	57		
N	-9.44 (8.27)	66	65
Mean (SD)		-15.23 (7.88)	-14.74 (9.17)
p-value		<.0001	.0012
Age (year)			
≤ 50			
N	56	60	56
Mean (SD)	-11.18 (7.53)	-13.55 (8.94)	-12.70 (9.07)
p-value		.1163	.1364
> 50			
N	40	54	55
Mean (SD)	-7.98 (7.62)	-14.13 (7.11)	-13.22 (9.24)
p-value		<.0001	.0053
CGI			
Gender			
Male			
N	39	48	46
# (%) Responders	17 (43.59%)	30 (62.50%)	34 (73.91%)
p-value		.1475	.0043
Female			
N	57	66	65
# (%) Responders	26 (45.61%)	53 (80.30%)	52 (80.00%)
p-value		.0001	.0001
Age			
≤ 50			
N	56	60	56
# (%) Responders	27 (48.21%)	43 (71.67%)	43 (76.79%)
p-value		.0627	.0050
> 50			
N	40	54	55
# (%) Responders	16 (40.00%)	40 (74.07%)	43 (78.18%)
p-value		.0019	.0002

4.2 Other Special/Subgroup Populations

No other analyses for special/subgroup populations were performed.

5. SUMMARY AND CONCLUSIONS

5.1 Statistical Issues and Collective Evidence

The two pivotal studies collectively provided sufficient evidence that XP13512 is effective in treating patients with restless legs syndrome (RLS). No major statistical issues were identified.

5.2 Conclusions and Recommendations

The efficacy results obtained from the analyses of the two pivotal studies XP052 and XP053 across two co-primary endpoints support the conclusion that XP13512 is effective in treating patients with restless legs syndrome (RLS). The effectiveness of XP13512 is also supported by Study XP060 and XP081. All doses studied (XP13512 600 mg, 1200 mg, 1800 mg, and 2400 mg) appear to be effective. XP13512 600 mg appears to be as efficacious as higher doses, and XP13512 2400 mg, the highest dose studied, does not appear to be more effective than the lower doses studied.

Application Type/Number	Submission Type/Number	Submitter Name	Product Name
NDA-22399	ORIG-1	GLAXO GROUP LTD DBA GLAXOSMITHKLINE	SOLZIRA

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

/s/

XIAORONG YAN
10/09/2009

KUN JIN
10/09/2009

HSIEN MING J J HUNG
10/10/2009