CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

022561Orig1s000

STATISTICAL REVIEW(S)

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

CARCINOGENICITY STUDY

NDA Number:	22,561 / Serial 0000
Drug Name:	Cladribine, oral formulation
Indication:	Treatment of multiple sclerosis
Applicant:	EMD Serono, Inc. Rockland, Massachusetts
	Testing Facilities: (b) (4)
Date:	Submitted 27 May 2010 To reviewer 3 August 2010
Review Priority:	Standard
Biometrics Division:	Division 6
Statistical Reviewer:	Steve Thomson
Concurring Reviewer:	Team Leader: Karl Lin, Ph. D.
Medical Division:	Neurology Products
Toxicologist:	Melissa Banks, Ph.D.
Project Manager:	Hamet Touré, PharmD, MPH, LT
Keywords:	Carcinogenicity, Cox regresson, Kaplan-Meier product limit, Survival analysis, Trend test, Nonparametric, Bayesian

Table of Contents

1. EXECUTIVE SUMMARY

The Sponsor's report indicate that the objective of the 26-week study was to "assess the potential carcinogenicity of Cladribine (dissolved tablets or drug substance) in CByB6F1-Tg(HRAS)2Jic Hemizygous mice and to evaluate the toxicokinetic behavior in CByB6F1–Tg(HRAS)2Jic wild type mice following daily oral exposure with a cyclic regimen for 26 weeks." (page 2 of report) Similarly the objective of the earlier 22 month DS95011 study "was to test the oncogenic potential of … cladribine (an antimetabolite) when administered subcutaneously to … mice intermittently (7 days of treatment followed by 21 drug-free days) for approximately 2 years." (page 6 of report).

1.1. Conclusions and Recommendations

This submission summarizes the results of a 26 week oral gavage study in Tg(HRAS)2 mice and a 22 month study in standard laboratory mice. Gross aspects of the study designs are summarized in the following tables, for each gender in each species:

Table 1. Study (b) (4) 281.02A: Design of 26-Week Mice Study (Dose Volume 10 mL/kg)

Treatment Groups	Description	Route of Administration	Dose Level (mg/kg)	# animals per gender	
				Main Study	TK
1. Vehicle[1]	Control	Gavage	0	25	9
2. Low	Tablets	Gavage	5	25	24/22[5]
3. Medium	Tablets	Gavage	15	25	24/22[5]
4. High	Tablets	Gavage	30	25	24/22[5]
5. Substance[2]	Drug	Gavage	15	25	24/22[5]
6. Excipient[3]	Main	Gavage	431	25	9
7. MNU[4]	Active Cntrl	IP Injection	75	25	0

[1] De-ionized water
[2] Cladribine Drug Substance
[3] 2-hydroxypropyl-β-cyclodextrin (It is arguable that this should considered as the vehicle)
[4] N-methyl-N-nitrosourea, positive control
[5] The Sponsor's report says that there were 24 animals per gender in these TK groups but the provided data set has 24 animals/group in male mice and 22 animals/group in females.

In the 26-Week Study, (b) (4) 281.02A, Groups 1 through 6, were dosed by oral gavage in cycles of 5 days of consecutive dosing followed by 23 days of non-dosing. The active control MNU Group 7 animals were injected with a single intraperitoneal injection on Day 1. The Sponsor labeled vehicle was deionized water. Note that 2-hydroxypropyl-β-cyclodextrin (HPBCD) in Group 6 was classified as an excipient in the tablet. In case it is not completely inert it would seem to make more sense, particularly in the carcinogenicity analysis, to use HPBCD in water, group 6, as the vehicle rather than deionized water as specified by the Sponsor.

In Study DS95011, animals were injected subcutaneously for 7 days of consecutive dosing followed by 21 days of non-dosing. The study report states that the "mouse (and not the rat was selected for this study due to similarities in plasma deoxycytidine levels and a target organ toxicity similar to primates." (page 9 of report)

Table 2. Study DS95011: Design of Study 22-Month Mice Study

Treatment Groups	# animals per study per gender	Dosage (mg/kg)	Concentration (mg/mL)	Dose Volume (mL/kg)
1. Vehicle	65	0	0	5
2. Untreated	65	0	0	0
3. Low	65	0.1	0.02	5
4. Medium	65	1.0	0.2	5
5. High	65	10.0	2.0	5

Somewhat more detailed descriptions of the studies are provided in Sections 3.2.1 and 3.2.2, below. The vehicle (whether water or HPBCD), low, medium, and high dose groups, groups 1 or 6 and 2 through 4 in the Tg(HRAS)2 study and groups 1, 3 to 5 in the earlier mouse study, are referred to as "dose groups" or "treated groups" and are used for tests of trend.

Table 3 below indicates the number of animals in the main study and the TK substudy that died before the final time of sacrifice. With the exception of the MNU control group 7, there are so few events that so there does not seem to be any point in testing such survival differences. Although thus of limited interest, figures A.1.1 and A.1.2 in Appendix 1 display survival curves for the pooled main study and TK animals. In these curves early sacrifices in the TK group are treated as censored observations, not as natural deaths.

Table 3. Study (b) (4) 281.02A: Summary of Survival

	Group 1 Vehicle	Group 2 Low	Group 3 Medium	Group 4 High	Group 5 Substance	Group 6 Excip.	Group 7 MNU
Male Main	2/25	1/25	0/25	0/25	0/25	0/25	21/25
TK	1/9	0/24	0/24	0/24	0/24	0/9	
Female Main	0/25	0/25	0 or 1/25[1]	1/25	2/25	0/25	12/25
TK	0/9	1/22	0/22	0/22	0/22	0/9	

[1] Although the death of this study is near the end of the study the Sponsor's report labels this as an early death but it is not flagged as such in the submitted data set.

The following Table 4 summarizes the results of tests comparing survival profiles across dose groups in the tumorigenicity data set for the 22-month (two year) study.

Table 4. Study DS95011: Statistical Significance of Tests of Homogeneity and Trend in Survival

Hypothesis Tested	Males		Females	
	Log rank	Wilcoxon	Log rank	Wilcoxon
Homogeneity over Groups 1-5	0.1033	0.1111	0.7778	0.8381
Homogeneity over Groups 1, 3-5	0.1426	0.1919	0.6231	0.7309
No trend over Groups 1, 3-5	0.1808	0.1649	0.5616	0.6142
No Difference Between Groups 1 vs 5	0.0676	0.0683	0.3824	0.4037
No Difference Between Groups 2 vs 5	0.0727	0.0465	0.6203	0.5484
No Difference Between Groups 1 vs 2	0.9244	0.9876	0.7168	0.8250

Figures A.1.3 through A.1.4, in Appendix 1, provide survival curves for each gender combinations in the DS95011 study. These results are somewhat more interesting than the results from the 26-Week Tg(HRAS)2 study. From Figure A.1.3, in male mice there appears to be a general, though not uniform, trend in decreasing survival over increasing dose. However, whether dealing with all five dose groups or with just the four dose groups remaining after deleting the no treatment group, there is no statistically significant evidence of differences in homogeneity at the usual 0.05 level (all four $p \geq 0.1033$). Despite the apparent trend in male mice in Figure A.1.3, the test that there is no trend is not statistically significant (Logrank $p = 0.1808$, Wilcoxon $p = 0.1649$), suggesting no real evidence of a trend. However the tests between the high dose and vehicle are somewhat close to the usual significance (Logrank $p = 0.0676$, Wilcoxon $p = 0.0683$). The early difference in male mice between the no treatment group and the high dose group leads to Wilcoxon test that is barely statistically significant ($p = 0.0465$). However the corresponding log rank test is not statistically significant ($p = 0.0727$), though close. In Figure A.1.4 in female mice the survival curves are all closely intertwined with no consistent patterns of dominance (all 12 $p \geq 0.3824$), indicating no differences in survival.

An experimental Bayesian nonparametric analysis of survival for Study DS95011 is given in Appendix 2. A Bayesian analysis postulates that probability is a useful measure of uncertainty about the parameters of a statistical model. The analysis will typically assess the probability that the parameters satisfy certain conditions. This analysis also suggests there are no strong dose related trends or differences.

The significance levels of the tests of tumorigenicity in the FDA analysis are based on poly-k tests applied to the data sets provided by the Sponsor and adapted by this reviewer. The poly-k test modifies the original Cochran-Armitage test of dose related trend in the occurrence of an event to adjust for differences in mortality (please see Bailer & Portier, 1988, Bieler & Williams, 1993). One problem with any such tumorgenicity analyses is that for each tumor-organ-gender-study combination there is one test of significance for each comparison of an actual treatment group to controls plus a test of overall trend. This implies a large number of tests, decreasing test specificity, and thus necessitating a multiplicity adjustment. For two species, two gender per species studies the so-called Haseman-Lin-Rahman rules adjust for the

multiplicity of tests of tumorigenicity by modifying the interpretation of the usual significance level (i.e. "p-value"). These specify that for tests of trend at a roughly overall 0.10 (10%) false positive error rate, one might claim statistical significance if the observed significance level is 0.025 for rare tumors (with a historical control incidence less than 1%) and 0.005 (incidence at or greater than 1%) for common tumors. Tests comparing the high dose group to controls would be considered statistically significant if the observed significance level is 0.05 for rare tumors and 0.01 for common tumors. This adjustment for multiplicity is discussed in Section 1.3.1.5 below. A number of other comparisons are also arguably relevant, but would also increase overall Type I error rate and thus decrease specificity of results. Rather surprisingly, none of the numerous tests in the Tg(HRAS)2 study were statistically significant at a 0.05 level, let alone at one of the multiplicity adjusted levels.

Tables 5 and 6, below, display for those comparisons for the DS95011 study, the tumor incidences and the results of tests of no differences between treatments for those neoplasms that had at least one test that achieved at least a nominal 0.05 level of statistical significance. Note that there is a clear difference in the rate of neoplasms in the vehicle group versus the no treatment group. The test of trend summarized below is over the vehicle, low, medium, and high dose groups. Significance levels of pairwise tests between the high, medium, and low dose group are on the same line. Underneath these pairwise tests to vehicle in the table are the corresponding pairwise tests to the no treatment group. Note that a test of trend using the no treatment group as the baseline would not be appropriate. Complete tables of incidences and the significance levels of statistical tests in both studies are provided in Appendix 3.

Table 5. Study DS95011: Tumor Incidences and Statistical Significance of Tests of Trend and Pairwise Differences in Male Standard Lab Mice with at least One Significant Test

Organ / Tumor	Incidence					Significance Levels				
		No					Vehicle/No Treat versus			
	Veh	Trt	Low	Med	Hi	Trend	High	Medium	Low	
LACRIMAL/HARDIRIAN GLANDS										
ADENOMA	6	11	3	8	28	0.0000	0.0000	0.3448	0.9055	Veh
							0.0002	0.7973	0.9940	No
Adenoma/Adenocarcinoma	6	11	4	8	29	0.0000	0.0000	0.3448	0.8162	Veh
							0.0001	0.7973	0.9827	No
LIVER										
ADENOMA, HEPATOCELLULAR	19	0	8	12	8	0.8928	0.9896	0.9281	0.9934	Veh
							0.0018	0.0001	0.0023	No
Adenoma/Carcinoma Hepato.	20	0	11	14	9	0.9183	0.9870	0.8832	0.9746	Veh
							0.0007	0.0000	0.0002	No
CARCINOMA, HEPATOCELLULAR	4	0	3	5	1	0.8938	0.9639	0.4759	0.7655	Veh
							0.4712	0.0244	0.1113	No
LUNG										
ADENOMA, BRONCHIOLO-ALVEOLAR	12	0	10	20	9	0.8196	0.7529	0.0554	0.7155	Veh
							0.0008	0.0000	0.0005	No
Adenoma/Carcinoma Bronch.-Alv.	15	0	15	24	13	0.7955	0.6445	0.0410	0.5145	Veh
							0.0000	0.0000	0.0000	No
CARCINOMA, BRONCHIOLO-ALVEOLAR	3	0	5	4	4	0.4150	0.4352	0.4576	0.3268	Veh
							0.0461	0.0503	0.0244	No
LYMPHORETICULAR SYSTEM										
LYMPHOMA	4	1	6	3	4	0.5370	0.5648	0.7319	0.3245	Veh
							0.1426	0.2673	0.0457	No
Lymphoma/Leukemia, Myleogenous	5	1	6	3	4	0.6071	0.6845	0.8250	0.4475	Veh
							0.1426	0.2673	0.0457	No

Table 5 (cont.) Study DS95011: Tumor Incidences and Statistical Significance of Tests of Trend and Pairwise Differences in Male Standard Lab Mice with at least One Significant Test

Organ / Tumor	Incidence (No Trt)					Significance Levels				
	Veh	Trt	Low	Med	Hi	Trend	Vehicle/No Treat versus High	Medium	Low	
SPLEEN										
HEMANGIOSARCOMA	1	0	3	5	1	0.7467	0.7276	0.0900	0.2947	Veh
							0.4712	0.0232	0.1113	No
Systemic										
HEMANGIOSARCOMA	4	0	5	11	3	0.8398	0.7310	0.0393	0.4642	Veh
							0.1012	0.0002	0.0244	No
Hemangioma/Hemangiosarcoma	5	0	9	14	3	0.9627	0.8243	0.0186	0.1754	Veh
							0.1012	0.0000	0.0011	No

In Table 5 above, in male mice, following the adjustment for multiplicity to get an overall rough 10% error rate and using the incidence in the no treatment group to decide if a tumor is rare or not, we would conclude, that the test of trend in both adenomas and pooled adenoma and adenocarcinoma of the lacrimal/Harderian glands was statistically significant (both $p < 0.00005 < 0.005$), as were the pairwise comparions between the high dose group and controls (both $p < 0.00005 < 0.01$). Applying the Haseman-Lin-Rahman rules to the comparisons of the high dose with the no treatment group can be expected to increase the overall Type I error to some value above the 10% level. Performing such tests would lead to statistically significant differences between the high dose and the no treatment group in adenoma and pooled adenoma/adenocarcinoma of the lacrimal harderian gland ($p = 0.0002$, $0.0001 < 0.05$, respectively). Similar comparisons in hepatocellular adenoma, and pooled adenoma and carcinoma of the liver would also be statistically significant ($p = 0.0018$, $0.0007 < 0.05$), as would be the tests of bronchiolar-alveolar adenoma, carcinoma, and pooled adnoma and carcinoma (0.0008, 0.0461, $0.00005 <$ all 0.05). Adding comparisons with other dose groups will only increase the Type I error rate even further. More detailed transcriptions for these comparisons are given in Section 3.2.2.1 below and in Appendix 3.

Table 6 below summarizes similar results in female mice. In Table 6, for female mice, the tests of trend in adenoma and adenocarcinoma of the Lacrimal/Hardirian glands would be considered as statistically significant ($p = 0.0001$ and $p < 0.00005$, both < 0.005). In these same glands the corresponding pairwise comparisons between the high dose and vehicle would classified as statistically significant ($p = 0.0030$ and $p = 0.0008$, both < 0.01).). Note that in the lung, the test of trend in bronchiolar-alveolar carcinoma and pooled adenoma and carcinoma would be close to significant ($p = 0.0058$, $0.0071 > 0.005$), while pairwise tests between the high dose group and vehicle in adenoma and pooled adenoma would be statistically significant ($p = 0.010 = 0.01$ and $p = 0.0013 < 0.01$). Adding comparisons of the high dose group with the no treatment group would be expected to increase overall Type I error. Accepting this inflation we would conclude the pairwise test between the high dose group and the no treatment group in pooled systemic hemangioma and hemangiosarcomas would also be statistically significant ($p = 0.0037 < 0.01$), as would be the test in pooled mammary gland tumors ($p = 0.0328 < 0.05$). Using the Haseman-Lin-Rahman rules, no other

tests of trend or pairwise comparison of the high dose to vehicle or no treatment group are statistically significant.

Table 6. Study DS95011: Tumor Incidences and Statistical Significance of Tests of Trend and Pairwise Differences in Female Standard Lab Mice with at least One Significant Test

Organ / Tumor	Incidence No Veh	Trt Low	Med	Hi	Trend	Significance Levels Vehicle/No Treat versus High	Medium	Low	
LACRIMAL/HARDIRIAN GLANDS									
ADENOMA	1	4	2	2	11 0.0001	0.0030	0.5230	0.5230	Veh
						0.0586	0.9067	0.9067	No
Adenoma/Adenocarcinoma	1	4	2	2	13 0.0000	0.0008	0.5230	0.5230	Veh
						0.0240	0.9067	0.9067	No
LIVER									
ADENOMA, HEPATOCELLULAR	1	0	2	5	1 0.7268	0.7675	0.1167	0.5230	Veh
						0.5100	0.0312	0.2576	No
Adenoma/Carcinoma Hepato.	1	0	2	6	1 0.7760	0.7675	0.0658	0.5230	Veh
						0.5100	0.0151	0.2576	No
LUNG									
ADENOMA, BRONCHIOLO-ALVEOLAR	2	9	9	15	11 0.1847	0.0100	0.0008	0.0300	Veh
						0.4403	0.1580	0.6202	No
Adenoma/Carcinoma Bronch.-Alv.	5	14	10	17	20 0.0071	0.0013	0.0057	0.1501	Veh
						0.2197	0.3821	0.8889	No
CARCINOMA, BRONCHIOLO-ALVEOLAR	3	5	1	4	9 0.0058	0.0858	0.5226	0.9460	Veh
						0.2534	0.7762	0.9883	No
LYMPHORETICULAR SYSTEM									
LYMPHOMA	17	10	12	22	21 0.1831	0.3744	0.3045	0.9055	Veh
						0.0437	0.0290	0.4673	No
MAMMARY GLAND/REGION									
Adenoma/Adenocarc./Fibroadenoma	1	0	3	3	5 0.0826	0.1165	0.3315	0.3240	Veh
						0.0328	0.1326	0.1288	No
OVARIES									
Leiomyoma/Lute-/Thec-/G-cell Tmr	2	0	5	3	2 0.7473	0.7140	0.5290	0.2438	Veh
						0.2576	0.1288	0.0312	No
Systemic									
HEMANGIOSARCOMA	6	0	7	3	4 0.7739	0.8563	0.9289	0.5454	Veh
						0.0663	0.1326	0.0083	No
Hemangioma/Hemangiosarcoma	8	0	9	5	8 0.4744	0.6397	0.8972	0.5528	Veh
						0.0037	0.0328	0.0019	No
UTERUS									
ADENOCARCINOMA, ENDOMETRIAL	0	0	2	0	4 0.0267	0.0664	.	0.2628	Veh
						0.0637	.	0.2576	No
ENDOMETRIAL STROMAL POLYP	3	0	5	7	3 0.7305	0.6705	0.1673	0.3688	Veh
						0.1288	0.0072	0.0312	No
LEIOMYOMA	2	0	5	2	1 0.8949	0.8897	0.7140	0.2525	Veh
						0.5100	0.2576	0.0328	No
Leiomyoma/-sarcoma/stromal sarc.	2	0	6	5	1 0.9294	0.8897	0.2525	0.1691	Veh
						0.5100	0.0328	0.0170	No

Other comparions to the vehicle and no treatment group are summarized in Section 3.2.2.1. and in Appendix 3.

1.2. Brief Overview of the Studies

This submission had a transgenic mouse study:

Study (b) (4) 281.02: A 26-Week Oral Dose Carcinogenicity Study of Cladribine in CByB6F1-Tg(HRAS)2Jic Hemizygous Mice and Toxicokinetic Study in CByB6F1-Tg(HRAS)2Jic Wild Mice,

plus an older study using standard laboratory mice:

Study DS95011: Two year Subcutaneous Carcinogenicity Study of Cladribine in Mice Using an Intermittent Dosing Schedule,

to assess the carcinogenic potential of Cladribine in mice. Each gender in the Tg(HRAS)2 study each study involved seven treatment groups with 25 animals per group in the main study. Study DS95011 involved five treatment groups including a vehicle control and a no treatment group, with 65 animals per group.

1.3. Statistical Issues and Findings

1.3.1. Statistical Issues

In this section, several issues, typical of statistical analyses of these studies, are considered. These issues include details on the survival analyses, tests on tumorigenicity, multiplicity of tests on neoplasms, and the validity of the designs.

1.3.1.1. Survival Analysis:

In the Tg(HRAS)2 study, there are too few events in the main study groups to detect a difference in survival, so survival analysis is performed only for the 22-month study. The survival analyses presented in this review are based on both the log rank test and the Wilcoxon test comparing survival curves. The log rank tests tend to put higher weight on later events, while the Wilcoxon test tends to weight events more equally, and thus is more sensitive to earlier differences in survival. The log rank test is most powerful when the survival curves track each other, and thus the hazards, i.e., the conditional probability of the event in the next infinitesimal interval, would be roughly proportional. Both tests were used to test both homogeneity of survival among the treatment groups and the effect of dose on trend in survival. Appendix 1 reviews the specific animal survival analyses in more detail. The results of the Sponsor's analyses are summarized in Sections 3.2.1.1 and 3.2.2.1.

An experimental Bayesian nonparametric analysis of survival is given in Appendix 2. A Bayesian analysis takes describes lack of knowledge about the parameter of interest in terms of a prior probability distribution, and then uses the data to modify these probability estimates. In this context it can be used to answer questions about the behavior of these parameters. In

particular it can be used to assess the probability that the parameter satisfies any criteria of interest. Note that this is a nonparametric procedure so the actual distribution is one of the parameters.

1.3.1.2. Multiplicity of Tests on Survival:

Using the logrank and Wilcoxon tests in the 22-Month study, there are 12 tests of survival for each gender combination. If we were to assume the tests are independent across comparisons, which clearly they are not, and assume that there is absolutely no difference in survival, the probability of at least one statistically significant result at the usual 0.05 level, is about 0.46. Such is the possible price paid for the multiplicity of hypothesis tests in the frequentist paradigm. Note that the Bayesian test is based on a hierarchical model and this provides an inherent adjustment for the multiplicity of tests.

1.3.1.3. Tests on Neoplasms:

The Sponsor's reports indicate that both studies were analyzed using Peto techniques, where fatal and mortality independent (i.e. observable) tumors were analyzed using life table techniques while incidental tumors were analyzed using a stratified Cochrane-Mantel type tests. The results from both tests are assumed to be independent and pooled, weighting by standard errors. Note that these tests require accurate determination of whether a tumor is fatal or incidental, which has been a major criticism of these techniques. In addition, from a statistical point of view, fixed time intervals for incidental tumors can cause some distributional problems in the Peto statistics, including loss of information, plus the independence assumption noted above is debatable.

Appendix 3 presents the results from the FDA poly-k analysis on tumor incidence in Tg(HRAS)2 mice and standard laboratory mice. The poly-k test is a modification of the original Cochran-Armitage test of trend in response to dose, adjusted for differences in mortality (please see Bailer & Portier, 1988, Bieler & Williams, 1993). It was noted in the report of the Society of Toxicological Pathology "town hall" meeting in June 2001 that the poly-k modification of the Cochran-Armitage tests of trend has been recommended over the corresponding Peto tests.

1.3.1.4. Multiplicity of Tests on Neoplasms:

Testing the various neoplasms necessitates a number of statistical tests, which in turn requires an adjustment in experiment-wise Type I error (i.e., the probability of rejecting a true null hypothesis). Based on his extensive experience with such carcinogenicity analyses in standard laboratory rodents, 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. Similarly, Lin and Rahman (1998) showed that tests of trend should be tested at a 0.025 (2.5%) level for rare tumors and 0.005 (0.5%) for common tumors. This approach is intended to balance both Type I error and Type II error (i.e., the error of concluding there is no evidence of a relation to tumorgenicity when there actually is such a

relation). Although this procedure is designed for combining standard rat and mouse studies it seems reasonably applicable to the two different mouse studies analyzed here.

However, including the results of other tests besides overall trend and the pairwise tests between the chosen control and the high dose group can be expected to increase the overall type I error rate. So even if one uses the Haseman-Lin-Rahman rules, the overall type I error associated with including these other tests may be considerably larger than the rough 10% when these rules are restricted to the test of trend and pairwise differences between the high dose and vehicle. Although this reviewer has some concerns about the large number of statistical tests conducted in each study, it is hoped that the increase in sensitivity may justify the negative impact on specificity.

1.3.1.5. Validity of the Designs:

When determining the validity of designs there are two key points:

1) adequate drug exposure,
2) tumor challenge to the tested animals.

1) is related to whether or not sufficient animals survived long enough to be at risk of forming late-developing tumors and 2) is related to the Maximum Tolerated Dose (MTD), designed to achieve the greatest likelihood of tumorigenicity.

Lin and Ali (2006), quoting work by Haseman, have suggested that in standard laboratory rodent species, a survival rate of about 25 animals, out of 50 or more animals, between weeks 80-90 of a two-year study may be considered a sufficient number of survivors as well as one measure of adequate exposure. Note that as a percentage of animals that survived to week 91, this criterion is met in all dose groups in both genders in the DS95011 (Please see tables 14 and 15 on page 20). The application of this criterion to the transgenic mouse study is not clear.

Chu, Ceuto, and Ward (1981), citing earlier work by Sontag *et al* (1976) recommend that the MTD "is taken as 'the highest dose that causes no more than a 10% weight decrement as compared to the appropriate control groups, and does not produce mortality, clinical signs of toxicity, or pathologic lesions (other than those that may be related to a neoplastic response) that would be predicted to shorten the animal's natural life span.' " The mean weight values in the following Table 7 were taken from the Sponsor report for (b) (4) 281.02A (Tables 4, pages 64-75). The change from baseline is the simple difference between means and is not mortality adjusted. Note the only dose group that exceeds this decrement in the Tg(HRAS)2 animals is the high dose group.

Table 7. Mean Weights for Mice in Tg(HRAS)2 26-Week Study

Tg.RASH2 Mice Dose Group	Males				Females			
	Week		Change from baseline	% change relative to vehicle	Week		Change from baseline	% change relative to vehicle
	1	26			1	26		
1. Vehicle	22.9	27.5	4.6		19.0	23.6	4.6	
2. Low	23.5	30.7	7.2	157%	19.3	24.4	5.1	111%
3. Medium	23.6	29.8	6.2	135%	19.1	23.7	4.6	100%
4. High	22.3	24.5	2.2	48%	18.9	23.1	4.2	91%
5. Substance[1]	22.4	30.0	7.6	165%	18.4	24.0	5.6	122%
6. Excipient[2]	22.2	29.8	7.6	165%	18.7	24.8	6.1	133%
7. MNU[3]	22.5	27.3	4.8	104%	18.5	24.3	5.8	126%

Similarly, for Study DS95011, Table 8 the mean weight values were taken from the report (Table T3, pages 51-62), as are the values for the change from week -1 (Table T4, pages 63-74).

Table 8. Mean Weights for Mice in DS95011 22-Month Mice Study

Mice Dose Group	Males				Females			
	Week		Change from baseline	% change relative to vehicle	Week		Change from baseline	% change relative to vehicle
	1	96			1	96		
1.Vehicle	34.58	46.65	12.82		25.65	40.69	14.66	
2.Untreated	34.66	47.53	13.95	109%	26.00	41.07	14.92	102%
3.Low[1]	34.37	47.64	14.15	110%	25.54	40.05	14.46	99%
4.Medium	34.33	46.31	12.55	98%	25.84	40.98	14.81	101%
5.High	34.01	44.67	11.19	87%	25.82	38.16	11.94	81%

The high dose groups in both genders clearly exceed the Chu, Ceuto, and Ward criterion, although the value for males only slightly exceeds this value. This may be evidence that the MTD was exceeded in these dose groups.

The Sponsor's report notes that for the Tg(HRAS)2 study, Study (b) (4) 281.02A, indicates that "No test article-related changes in food consumption were observed. Sporadic incidences of statistically significant differences from control were observed in all groups, but did not have a consistent dose-response or temporal pattern and were not considered an indication of a test article effect." (page 31 of report) The report for Study DS95011 similarly claims that while there were some statistically significant differences in mean food consumption, these were "incidental and not cladribine related."

Although it is not clear if the usual comments about the MTD apply with the unusual forms of dosing used in these studies, if one assumes they are applicable then from 2) above, excess mortality not associated with any tumor or sacrifice in the higher dose groups might suggest that the MTD was exceeded. If dosing is close to the MTD one would expect slightly

higher mortality due to toxicity, but not so much that it largely reduces the number of animals exposed to the drug. As can be seen in Table 3 above, in the Tg.RASH2 study, except for the MNU group, there clearly were too few non-sacrifice deaths to effectively assess early deaths due to toxicity. In the 22-Month Mice Study, DS95011, in male mice there is slight evidence of higher mortality over increasing dose. A related way to assess whether or not the MTD was achieved is to measure mortality not associated with any identified tumor. Table 9, below, indicates that the number of animals in each dose group in the DS95011 Study that died of a natural death or moribund sacrifice, but did not show any tumors (i.e., the "Event"):

Table 9. DS95011 22-Month Study Natural Death with No Identified Tumor

		1. Vehicle 0 mg/kg	2. Untreated 0 mg/kg	3. Low 0.1 mg/kg	4.Medium 1.0 mg/kg	5. High 10 mg/kg
Male	Event	8	25	14	13	15
	No event	57	40	51	52	50
Female	Event	12	21	9	5	7
	No event	53	44	56	60	58

Using survival models for the event, there are statistically significant differences among dose groups in females, and differences that are close to significance in males. However, these do not seem to be associated with increasing dose and are primarily due to the high incidence of the events (where the "event" is death with no tumor) in the untreated group. Again, while, like the other observations above, these require the expertise of the toxicologist, since we might expect a difference in the number of events in the high dose group, this may be interpreted as evidence that the MTD was not achieved.

1.3.2. Statistical Findings

Please see Section 1.1 above.

2. INTRODUCTION

2.1. Overview

This submission summarizes the results of a 26-week study was in Tg(HRAS)2 mice and a two-year (actually 22 month study terminated in weeks 98-99). to "assess the potential carcinogenicity of Cladribine (dissolved tablets or drug substance) in in CByB6F1-Tg(HRAS)2Jic Hemizygous mice and to evaluate the toxicokinetic behavior in CByB6F1–Tg(HRAS)2Jic wild type mice following daily oral exposure with a cyclic regimen for 26 weeks." (page 2 of report) Similarly the objective of the older 22 month study "was to test the oncogenic potential of … cladribine (an antimetabolite) when administered subcutaneously to … mice intermittently (7 days of treatment followed by 21 drug-free days) for approximately 2 years." (page 6 of report).

2.2. Data Sources

Each hidden within several levels of DARRTS, the Sponsor provided a SAS transport data sets for both studies labeled tumor.sas7bdat. Both data sets required extensive preparation for analysis, since numerous null records were included, and variable names and characteristics differed from those requested by the FDA Office of Biostatistics. The data for the Tg(HRAS)2 study did not distinguish between main study animals and toxicokinetic animals. To make treatment groups more comparable, the animals identified as toxicokinetic animals were excluded from the tumor analyses. Also, this reviewer corrected several records in each study so that codes and labels were consistent.

3. STATISTICAL EVALUATION

3.1. Evaluation of Efficacy

NA

3.2. Evaluation of Safety

3.2.1 Study (b) (4) 281.02: A 26-Week Oral Dose Carcinogenicity Study of Cladribine in CByB6F1-Tg(HRAS)2Jic Hemizygous Mice and Toxicokinetic Study in CByB6F1-Tg(HRAS)2Jic Wild Mice

STUDY DURATION: 26 Weeks
EXPERIMENTAL START DATE: Males: 25 September 2008 (dose initiation)
Females: 2 October 2008
EXPERIMENTAL END DATE: Males: 17 March 2009 - 20 March 2009 (terminal necropsy)
Females: 24 March 2009 - 27 March 2009
RAT STRAIN: CByB6F1-Tg(HRAS)2Jic Hemizygous Mice
ROUTE: Oral gavage

In the 26-Week Study in Tg(HRAS)2 mice, (b) (4) 281.02A, Groups 1 through 6 were dosed by oral gavage in cycles of 5 days of consecutive dosing followed by 23 days of non-dosing. The active control MNU Group 7 animals were injected with a single intraperitoneal injection on Day 1. Animals were housed individually in stainless steel cages. Food and water were available *ad libitum*, except at the time of necropsy. Other gross aspects of the study design are summarized in the following table, for each gender:

Table 10. Study (b) (4) 281.02A: Design of 26-Week Mice Study (Dose Volume 10 mL/kg)

Treatment Groups	Description	Route of Administration	Dose Level (mg/kg)	# animals per gender	
				Main Study	TK
1. Vehicle[1]	Control	Gavage	0	25	9
2. Low	Tablets	Gavage	5	25	24/22[5]
3. Medium	Tablets	Gavage	15	25	24/22[5]
4. High	Tablets	Gavage	30	25	24/22[5]
5. Substance[2]	Drug	Gavage	15	25	24/22[5]
6. Excipient[3]	Main	Gavage	431	25	9
7. MNU[4]	Active Cntrl	IP Injection	75	25	0

[1] Deionized water
[2] Cladribine Drug Substance
[3] 2-hydroxypropyl-β-cyclodextrin (It is possible this should be the vehicle)
[4] N-methyl-N-nitrosourea, positive control
[5] The Sponsor's report says that there were 24 animals per gender in these TK groups but the provided data set has 24 animals/group in male mice and 22 animals/group in females.

The Sponsor justifies dosing levels as follows: "Dose levels were selected taking into account the results obtained from a previous dose range-finding study in the same species and strain ((b) (4) 281.01, RE7990 – IMP28853). In the previous study, mice received daily oral administration given on two five consecutive day cycles with a 23-day interval between cycles. At the end of the second dosing cycle period most animals given 60 mg/kg during the second cycle showed moderate to severe renal tubular degeneration/regeneration and some animals given 30 mg/kg for both 5 day cycles showed similar changes of slight to moderate severity. No changes were found at 20 mg/kg. Therefore, 30 mg/kg was considered to be the maximum tolerated dose (MTD). The lowest dose selected for this study corresponds to about 7 times the anticipated human therapeutic exposure." (page 25 of report)

3.2.1.1. Sponsor's Results and Conclusions

This section will present a summary of the Sponsor's analysis on survivability and tumorigenicity in Tg(HRAS)2 mice.

Survival analysis:

According to the Sponsor: "Treatment-related mortality was observed only in the positive control group." (page 33 of report) This is consistent with the results in the FDA analysis.

Tumorigenicity analysis:

The Sponsor's Statistical report indicates that tests on neoplasms were Peto tests of trend (please see Section 1.3.1.3). The report summarizes their interpretation of results as follows: "With the exception of positive control (Group 7) animals, no test article-related neoplastic findings were present. None of the neoplastic findings in Groups 2 – 6 occurred at an incidence that was statistically different from Group 1 animals.

”Male Group 7 (positive control) animals had a 52% incidence of lymphosarcoma, in agreement with published values (Takaoka M, et al, 2003), indicating that the positive control animals were responding appropriately. The incidence of lymphosarcoma in female animals was 32% but these animals received a lower dose of MNU as compared to the males . . . . The incidence of squamous cell papilloma/carcinoma of the forestomach, expected in positive control animals, was 68% in males and 64% in females; this was somewhat lower in both genders than previously-published data (Takaoka M, et al, 2003). Several other tumors with sporadic incidence were also observed in the positive control group. The incidence of many of the tumor types seen in Group 7 animals was statistically significantly greater than that of Group 1 . . .. ” (page 34 of report)

3.2.1.2. FDA Reviewer's Results

This section will present the Agency findings on survival and tumorigenicity in male and female transgenic mice.

Survival analysis:

Table 11 below indicates the number of animals in the main study and the TK substudy that died before the final time of sacrifice.

Table 11. Study (b) (4) 281.02A: Summary of Survival

	Group 1 Veh	Group 2 Low	Group 3 Med	Group 4 High	Group 5 Substance	Group 6 Excip.	Group 7 MNU
Male Main	2/25	1/25	0/25	0/25	0/25	0/25	21/25
TK	1/9	0/24	0/24	0/24	0/24	0/9	
Female Main	0/25	0/25	0 or 1/25[1]	1/25	2/25	0/25	12/25
TK	0/9	1/22	0/22	0/22	0/22	0/9	

[1] Although the death of this study is near the end of the study the Sponsor’s report labels this as an early death but it is not flagged as such in the submitted data set.

With the exception of the MNU active control group 7, as noted by the Sponsor, there are so few events that so there is no point in testing survival differences. Although thus of limited interest, figures A.1.1 and A.1.2 in Appendix 1 display survival curves for the pooled main study and TK animals. In these curves early sacrifices in the TK group are treated as censored observations, not as natural deaths.

Tumorigenicity analysis:

As with the Sponsor’s analysis there was no statistically significant evidence of any differences in carcinogenicity among the first six treatment groups. This included tests of trend over the first four dose groups, a test of trend with groups 2, 3, or 4 using the excipient as the baseline, corresponding pairwise comparisons with the vehicle or the excipient, pairwise tests between the drug substance and the excipient with the Sponsor’s water vehicle, and finally a pairwise comparison the drug substance with the medium dose group. Complete tables of incidence are given in Appendix 3.

3.2.2. Study DS95011: Two year Subcutaneous Carcinogenicity Study of Cladribine in Mice Using an Intermittent Dosing Schedule

STUDY DURATION: 22 Months
EXPERIMENTAL START DATE : 16 November 1995 (initiation of dosing)
EXPERIMENTAL END DATE: 8 October 1997 (last necropsy)
MOUSE STRAIN: (b) (4) Crl:CD1®(ICR) BR, VAF/PLUS® Mice
ROUTE: Intermittent Subcutaneous Injection

In Study DS95011, animals were injected subcutaneously for of 7 days of consecutive dosing followed by 21 days of non-dosing. The study report for states that the "mouse (and not the rat was selected for this study due to similarities in plasma deoxycytidine levels and a target organ toxicity similar to primates." (page 9 of report)

Table 12. Study DS95011: Design of Study 22-Month Mice Study

Treatment Groups	# animals per study per gender	Dosage (mg/kg)	Concentration (mg/mL)	Dose Volume (mL/kg)
1. Vehicle	65	0	0	5
2. Untreated	65	0	0	0
3. Low	65	0.1	0.02	5
4. Medium	65	1.0	0.2	5
5. High	65	10.0	2.0	5

The vehicle contains propylenel glycol, m-cresol, benzyl alcohol , di- and mono-basic sodium phosphate in a sterile water solution. The Sponsor states that animals were randomly allocated to treatment, stratified by weight. Animals were housed individually in stainless steel cages, rotated within the study room approximately every month. Food and water were available *ad libitum*.

The Sponsor justifies dosing as follows: "Exploratory subcutaneous 3-month range-finding studies were conducted in the mouse to determine the appropriate dosing regimen for this study. Two dosing regimens were used; daily dosing for 3 months at dosages randing from 1 to 30 mg/kg/day ([Study] DS94125) and an intermittent dosing schedule with each cycle consisting of 5 to 7 days of cladribine administration at dosages ranging from 10 to 80 mg/kg/day followed by a 23- to 27-day drug free period ([Study] DS94415)."

"Daily administration of 30 mg/kg/day of cladribine for 3 months resulted in the deaths of four male mice, a decrease in mean body weight gain, and had significant effects on the testes (atrophy/degeneration of seminiferous tubules, hypospermia/aspermia, and hyperplasia of interstitial cells), spleen (depletion of lymphocytes), and bone marrow (fatty deposition). The spleen and bone marrow of mice in the 1- and 10-mg/kg/day dosage groups were similarly affected. These findings and the observed lymphopenia, leucopenia, and thrombocytopenia were expected effects of this cytoxic agent on rapidly dividing cells. It appeared unlikely,

based on the severity of these findings after 3 months of cladribine administration, that mice would even tolerate a daily dosage as low as 1.0 mg/kg/day over the 2-year period of a carcinogenicity study.

"Three cycles of the intermittent dosing regimen resulted in lethality at dosages >60 mg/kg/day and dosage-dependent toxicity at all dosage levels. Dosages of 30, 60, and 80 mg/kg/day resulted in dosage-related effects on the bone marrow, lymphoid tissues, and testes, similar to those encountered with daily dosing at 30 mg/kg/day, as well as causing villous atrophy in the duodenum. Effects observed at 10 mg/kg/day were minimal and limited to the bone marrow and testes. With the exception of testicular atrophy which persisted in all 30-, 60-, and 80-mg/kg/day dosed mice and two mice from the 10-mg/kg/day group, these effects were no longer seen after 21-day drug-free period which followed the third dosing period.

"An intermittent dosing regimen was selected for this study because it closely resembles the treatment regimen being employed in humans, and the drug-free period should allow the mice to recover to some degree from the toxic effects of cladribine. This would include cellular regeneration. Survival over two years would also be likely with an intermittent schedule. Dosages selected for the intermittent dosing regimen were 0.1, 1.0, and 10.0 mg/kg/day, approximately 1.4, 14, and 140 times the clinical dosage of 0.07 mg/kg/day currently under investigation in the treatment of multiple sclerosis. The high-dosage 10 mg/kg/day also yields a Plasma AUC in mice approximately 30 times that seen in humans at this clinical dosage. In addition, a single dose of 10 mg/kg/day was shown to induce micronuclei in the bone marrow of mice (DS94317)." (pages 16 and 17 of report)

3.2.1.1. Sponsor's Results and Conclusions

This section will present a summary of the Sponsor's analysis on survivability and tumorigenicity in mice in the DS95011 study.

Survival analysis:

According to the Sponsor: "There were no statistically significant mortality differences" (page 18 of report) Note that the FDA analysis below did find some possibly significant results. However, if one adjusted for multiplicity these would no longer be close to statistical significance.

Tumorigenicity analysis:

The Sponsor's report indicates that tests on neoplasms were Peto tests of trend as noted in Section 1.3.1.3 above. The report summarizes their interpretation of results as follows: "An increased incidence of Harderian gland tumors, predominantly adenomas, was observed in the 10 mg/kg/day male and female mice . . . These tumors were well-differentiated and usually were of the papillary or cystic papillary type. An incidence of one or two adenocarcinoma per affected group was observed, but there was no indication of a progression of the benign tumors to malignant tumors. Histomorphologically, the Harderian gland tumors were similar among the control and cladribine-treated mice." (page of report) The following table is copied from the Sponsor's report.

Table 13: Incidence of Harderian Gland Tumors in Mice

Group Dosage No. Examined	1 Vehicle Control 65	2 Untreated Control 65	3 0.1 mg/kg/day 65	4 1.0 mg/kg/day 65	5 10 mg/kg/day 65
Adenomas -Male	6	11	3	8	28
Adenomas -Female	1	4	2	2	11
Adenocarcinomas -Male	0	0	1	0	1
Adenocarcinomas -Female	0	0	0	0	2

Continuing, "There were changes observed in the injection sites, but these changes were considered to have been the result of the repeated injection procedure and not the result of a local irritating effect of the vehicle or test article.

"All other microscopic changes, non-neoplastic and neoplastic, were considered to be spontaneous lesions that occur in laboratory mice of this age and strain and their type or incidence were not considered to be compound-related." (page 18 of report)

3.2.1.2. FDA Reviewer's Results

This section will present the Agency findings on survival and tumorigenicity in male and female mice in Study .

Survival analysis:

The following tables (Table 14 for male mice, Table 15 for female mice) summarize the mortality results for the dose groups. The data were grouped for the specified time period, and present the number of deaths during the time interval over the number at risk at the beginning of the interval. The percentage cited is the percent that survived at the end of the interval. The Kaplan-Meier survival plots in Appendix 1 provide a more detailed picture of mortality losses.

Table 14. Summary of Male Mice Survival (dosed at mg/dose/day)

Period (Weeks)	Vehicle 0	No treat-Ment 0	Low 0.1	Medium 1	High 10
1-52	1/65[1] 98.5%[2]	1/65 98.5%	4/65 93.8%	2/65 96.9%	4/65 93.8%
53-78	11/64 81.5%	8/64 86.2%	7/61 83.1%	11/63 80.0%	15/61 70.8%
79-91	6/69 72.3%	10/69 70.8%	13/69 63.1%	16/69 55.4%	8/69 58.5%
92-99	9/66 58.5%	9/66 56.9%	7/66 52.3%	10/66 40.0 %	10/66 40.1%
Terminal 99	38	37	34	26	28

[1] number of deaths / number at risk
[2] overall per cent survival to end of period.

In these tables the terminal period only includes those animals were sacrificed. Animals that died of other causes during the terminal period are included in the preceding, but overlapping time period.

Table 15. Summary of Female Mice Survival (dosed at mg/dose/day)

Period (Weeks)	Vehicle 0	No treat-Ment 0	Low 0.1	Medium 1	High 10
1-52	5/65[1] 92.3%[2]	4/65 93.8%	3/65 95.4%	5/65 92.3%	4/65 93.8%
53-78	10/64 76.9%	13/64 73.8%	11/61 78.6%	10/63 76.9%	11/61 76.9%
79-91	14/69 55.4%	13/69 53.8%	14/69 56.4%	8/69 64.6%	10/69 61.5%
92- 98	9/66 41.5%	5/66 6.2%	9/66 43.1%	8/66 52.3 %	8/66 49.2%
Terminal 98	27	30	28	34	32

[1] number of deaths / number at risk
[2] overall per cent survival to end of period.

Table 16 below provides the significance levels of the tests of homogeneity and trend over dose groups as proposed in Section 1.3.1.1, above.

Table 16. Study DS95011: Statistical Significance of Tests of Homogeneity and Trend in Survival

Hypothesis Tested	Males		Females	
	Log rank	Wilcoxon	Log rank	Wilcoxon
Homogeneity over Groups 1-5	0.1033	0.1111	0.7778	0.8381
Homogeneity over Groups 1, 3-5	0.1426	0.1919	0.6231	0.7309
No trend over Groups 1, 3-5	0.1808	0.1649	0.5616	0.6142
No Difference Between Groups 1 vs 5	0.0676	0.0683	0.3824	0.4037
No Difference Between Groups 2 vs 5	0.0727	0.0465	0.6203	0.5484
No Difference Between Groups 1 vs 2	0.9244	0.9876	0.7168	0.8250

Figures A.1.3 through A.1.4, in Appendix 1, provide survival curves for each gender combinations in the DS95011 study. These results are somewhat more interesting than the results from the 26-Week Tg(HRAS)2 study. From Figure A.1.3, in male mice there appears to be a general, though not uniform, trend in decreasing survival over increasing dose. However, whether dealing with all five dose groups or with just the four dose groups remaining after deleting the no treatment group, there is no statistically significant evidence of differences in homogeneity at the usual 0.05 level (all four $p \geq 0.1033$). Despite the apparent trend in male mice in Figure A.1.3, the test that there is no trend is not statistically significant (Logrank $p = 0.1808$, Wilcoxon $p = 0.1649$), suggesting no real evidence of a trend. However the tests

between the high dose and vehicle are somewhat close to the usual significance (Logrank p = 0.0676, Wilcoxon p = 0.0683). The early difference in male mice between the no treatment group and the high dose group leads to Wilcoxon test that is barely statistically significant (p = 0.0465). However the corresponding log rank test is not statistically significant (p = 0.0727), though close. In Figure A.1.4 in female mice the survival curves are all closely intertwined with no consistent patterns of dominance (all 12 $p \geq 0.3824$), indicating no differences in survival.

An experimental Bayesian nonparametric analysis of survival for Study DS95011 is given in Appendix 2. A Bayesian analysis postulates that probability is a useful measure of uncertainty about the parameters of a statistical model. The analysis will typically assess the probability that the parameters satisfy certain conditions. This analysis also suggests there are no strong dose related trends or differences.

Tumorigenicity analysis:

In Table 17, below, in male mice, following the Haseman-Lin-Rahman adjustment for multiplicity to get an overall rough 10% error rate and using the incidence in the no treatment group to decide if a tumor is rare or not, we would conclude, that the test of trend in both adenomas and pooled adenoma and adenocarcinoma of the lacrimal/Harderian glands was statistically significant (both $p < 0.00005 < 0.005$), as were the pairwise comparions between the high dose group and controls (both $p < 0.00005 < 0.01$). Similarly, the pairwise comparisons between the high dose group and the no treatment groups would be statistically significant (each p=0.0002, 0.0001, both <0.01). Applying the Haseman-Lin-Rahman rules to the comparisons of the high dose with the no treatment group can be expected to increase the overall Type I error to some value above the 10% level. Performing such tests would lead to statistically significant differences between the high dose and the no treatment group in adenoma and pooled adenoma/adenocarcinoma of the lacrimal harderian gland (p = 0.0002, $0.0001 < 0.05$, respectively). Similar comparisons in hepatocellular adenoma, and pooled adenoma and carcinoma of the liver would also be statistically significant (p = 0.0018, 0.0007, both < 0.05), as would be the tests of bronchiolar-alveolar adenoma, carcinoma, and pooled adenoma and carcinoma (p = 0.0008, 0.0461, 0.00005, all < 0.05).

Table 17. Study DS95011: Tumor Incidences and Statistical Significance of Tests of Trend and Pairwise Differences in Male Standard Lab Mice with at least One Significant Test

Organ / Tumor	Incidence					Significance Levels				
		No					Vehicle/No Treat versus			
	Veh	Trt	Low	Med	Hi	Trend	High	Medium	Low	
LACRIMAL/HARDIRIAN GLANDS										
ADENOMA	6	11	3	8	28	0.0000	0.0000	0.3448	0.9055	Veh
							0.0002	0.7973	0.9940	No
Adenoma/Adenocarcinoma	6	11	4	8	29	0.0000	0.0000	0.3448	0.8162	Veh
							0.0001	0.7973	0.9827	No
LIVER										
ADENOMA, HEPATOCELLULAR	19	0	8	12	8	0.8928	0.9896	0.9281	0.9934	Veh
							0.0018	0.0001	0.0023	No
Adenoma/Carcinoma Hepato.	20	0	11	14	9	0.9183	0.9870	0.8832	0.9746	Veh
							0.0007	0.0000	0.0002	No
CARCINOMA, HEPATOCELLULAR	4	0	3	5	1	0.8938	0.9639	0.4759	0.7655	Veh
							0.4712	0.0244	0.1113	No
LUNG										
ADENOMA, BRONCHIOLO-ALVEOLAR	12	0	10	20	9	0.8196	0.7529	0.0554	0.7155	Veh
							0.0008	0.0000	0.0005	No
Adenoma/Carcinoma Bronch.-Alv.	15	0	15	24	13	0.7955	0.6445	0.0410	0.5145	Veh
							0.0000	0.0000	0.0000	No
CARCINOMA, BRONCHIOLO-ALVEOLAR	3	0	5	4	4	0.4150	0.4352	0.4576	0.3268	Veh
							0.0461	0.0503	0.0244	No
LYMPHORETICULAR SYSTEM										
LYMPHOMA	4	1	6	3	4	0.5370	0.5648	0.7319	0.3245	Veh
							0.1426	0.2673	0.0457	No
Lymphoma/Leukemia, Myleogenous	5	1	6	3	4	0.6071	0.6845	0.8250	0.4475	Veh
							0.1426	0.2673	0.0457	No
SPLEEN										
HEMANGIOSARCOMA	1	0	3	5	1	0.7467	0.7276	0.0900	0.2947	Veh
							0.4712	0.0232	0.1113	No
Systemic										
HEMANGIOSARCOMA	4	0	5	11	3	0.8398	0.7310	0.0393	0.4642	Veh
							0.1012	0.0002	0.0244	No
Hemangioma/Hemangiosarcoma	5	0	9	14	3	0.9627	0.8243	0.0186	0.1754	Veh
							0.1012	0.0000	0.0011	No

Adding comparisons with other dose groups will only increase the expected Type I error rate even further (i.e., further decrease specificity). Accepting this effect, one would conclude that comparisons between the medium dose group and the vehicle in systemic hemangioma and pooled hemangioma and hemangiosarcoma would be statistically significant ($p = 0.0393$, 0.0186, both < 0.05). Corresponding results from the comparison to the no treatment group are more extreme ($p = 0.0002$, $p < 0.00005$, both < 0.05). Further, we could conclude that comparisons between the medium dose group and the no treatment group in hepatocellular adenoma, carcinoma, and pooled adenoma and carcinoma of the liver would also be statistically significant ($p = 0.0001$, $p=0.0244$ and $p <0.00005$, all < 0.05), as would be the corresponding tests in bronchiolar-alveolar adenoma and pooled adenoma and carcinoma in the lung (both $p < 0.00005 < 0.01$). Results are only slightly less extreme for the comparisons of the low dose group to the no treatment group. Comparisons between the low dose group and the vehicle in systemic hemangioma and pooled hemangioma and hemangiosarcoma would also be statistically significant ($p = 0.0244$, 0.0011, both < 0.05). Comparisons between the low dose group and the no treatment group in hepatocellular adenoma and pooled adenoma and carcinoma of the liver would also be statistically significant ($p = 0.0023$, 0.0002, both < 0.05), as would be the corresponding tests in bronchiolar-alveolar adenoma, carcinoma, and pooled

adenoma and carcinoma in the lung (p = 0.0005, p=0.0244, p< 0.00005, all < 0.05). Finally, the difference between medium dose group and the no treatment group in hemangiosarcoma is statistically significant (p = 0.0232 < 0.05).

Table 18 below summarizes similar results in female mice.

Table 18. Study DS95011: Tumor Incidences and Statistical Significance of Tests of Trend and Pairwise Differences in Female Standard Lab Mice with at least One Significant Test

Organ / Tumor	Incidence No Veh	Trt Low	Med	Hi	Trend	Significance Levels Vehicle/No Treat versus High	Medium	Low	
LACRIMAL/HARDIRIAN GLANDS									
ADENOMA	1	4	2	2	11	0.0001	0.0030	0.5230	0.5230 Veh
							0.0586	0.9067	0.9067 No
Adenoma/Adenocarcinoma	1	4	2	2	13	0.0000	0.0008	0.5230	0.5230 Veh
							0.0240	0.9067	0.9067 No
LIVER									
ADENOMA, HEPATOCELLULAR	1	0	2	5	1	0.7268	0.7675	0.1167	0.5230 Veh
							0.5100	0.0312	0.2576 No
Adenoma/Carcinoma Hepato.	1	0	2	6	1	0.7760	0.7675	0.0658	0.5230 Veh
							0.5100	0.0151	0.2576 No
LUNG									
ADENOMA, BRONCHIOLO-ALVEOLAR	2	9	9	15	11	0.1847	0.0100	0.0008	0.0300 Veh
							0.4403	0.1580	0.6202 No
Adenoma/Carcinoma Bronch.-Alv.	5	14	10	17	20	0.0071	0.0013	0.0057	0.1501 Veh
							0.2197	0.3821	0.8889 No
CARCINOMA, BRONCHIOLO-ALVEOLAR	3	5	1	4	9	0.0058	0.0858	0.5226	0.9460 Veh
							0.2534	0.7762	0.9883 No
LYMPHORETICULAR SYSTEM									
LYMPHOMA	17	10	12	22	21	0.1831	0.3744	0.3045	0.9055 Veh
							0.0437	0.0290	0.4673 No
MAMMARY GLAND/REGION									
Adenoma/Adenocarc./Fibroadenoma	1	0	3	3	5	0.0826	0.1165	0.3315	0.3240 Veh
							0.0328	0.1326	0.1288 No
OVARIES									
Leiomyoma/Lute-/Thec-/G-cell Tmr	2	0	5	3	2	0.7473	0.7140	0.5290	0.2438 Veh
							0.2576	0.1288	0.0312 No
Systemic									
HEMANGIOSARCOMA	6	0	7	3	4	0.7739	0.8563	0.9289	0.5454 Veh
							0.0663	0.1326	0.0083 No
Hemangioma/Hemangiosarcoma	8	0	9	5	8	0.4744	0.6397	0.8972	0.5528 Veh
							0.0037	0.0328	0.0019 No
UTERUS									
ADENOCARCINOMA, ENDOMETRIAL	0	0	2	0	4	0.0267	0.0664	.	0.2628 Veh
							0.0637	.	0.2576 No
ENDOMETRIAL STROMAL POLYP	3	0	5	7	3	0.7305	0.6705	0.1673	0.3688 Veh
							0.1288	0.0072	0.0312 No
LEIOMYOMA	2	0	5	2	1	0.8949	0.8897	0.7140	0.2525 Veh
							0.5100	0.2576	0.0328 No
Leiomyoma/-sarcoma/stromal sarc.	2	0	6	5	1	0.9294	0.8897	0.2525	0.1691 Veh
							0.5100	0.0328	0.0170 No

In female mice, the tests of trend in adenoma and adenocarcinoma of the Lacrimal/Hardirian glands would be considered as statistically significant (p = 0.0001 and p < 0.00005, both < 0.005). In these same glands the corresponding pairwise comparisons between the high dose and vehicle would classified as statistically significant (p = 0.0030 and p = 0.0008, both < 0.01).). Note that in the lung, the test of trend in bronchiolar-alveolar carcinoma and pooled adenoma and carcinoma would be close to significant (p = 0.0058,

0.0071 > 0.005), while pairwise tests between the high dose group and vehicle in adenoma and pooled adenoma would be statistically significant (p = 0.010 = 0.01 and p = 0.0013 < 0.01). Adding comparisons of the high dose group with the no treatment group would be expected to increase overall Type I error. Accepting this inflation we would conclude the pairwise test between the high dose group and the no treatment group in pooled systemic hemangioma and hemangiosarcomas would also be statistically significant (p = 0.0037 < 0.01), as would be the test in pooled mammary gland tumors (p = 0.0328 < 0.05). No other tests of trend or pairwise comparison of the high dose to vehicle or no treatment group are statistically significant.

Again, comparisons with other dose groups will only further increase the expected Type I error rate and thus decrease specificity. Accepting this cost, one would conclude that for bronchiolar-alveolar carcinoma and pooled adenoma and carcinoma, the tests of differences between the medium dose group and vehicle would statistically significant (p = 0.0008, 0.0057 < 0.01 Further, comparisons between the high, medium, and low dose groups with the no treatment group in systemic pooled hemangioma and hemangiosarcoma would be also statistically significant (p = 0.0037, 0.0328, 0.0019, all < 0.05), as would be the comparison for hemangioma in the low dose group (p = 0.0083 < 0.05). Comparisons between the high and medium dose group with the no treatment group in adenoma and pooled adenoma and carcinoma of the liver would also be statistically significant (p = 0.0312, 0.0151 both < 0.05), Comparisons between the medium dose group with the no treatment group in adenoma and pooled adenoma and carcinoma of the liver would also be statistically significant (p = 0.0312, 0.0151 both < 0.05), as would be the test between the low dose group and the no treatment group for various pooled tumors in the ovaries (p = 0. 0312 < 0.05). The comparisons between the medium and low dose group and the no treatment group in endrometrial stromal polyp of the uterus would be statistically significant (p = 0.0072, 0.0312, both < 0.05), as would be the comparisons for pooled leiomyoma, leiomysarcoma, and stromal sarcoma emangioma in the low dose group (p = 0.0328, 0.0170, both < 0.05)). Finally the test between the low dose and the no treatment group in leiomyoma was statistically significant (p = 0.0328 < 0.05).

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

NA

5. SUMMARY AND CONCLUSIONS

5.1. Statistical Issues and Collective Evidence

Please see Section 1.3 above.

5.2. Conclusions and Recommendations

Please see Section 1.1 above

APPENDICES:

Appendix 1. FDA Survival Analysis

Simple summary tabulations of the unscheduled deaths in the 26-Week Tg(HRAS)2 Mice Study are presented Table 1 and the identical Table 7 in the main body of the report. Except for the active MNU control group, there is virtually no evidence of a treatment related effect on survival in this study. Figures A.1.1 and A.1.2, below, present Kaplan-Meier estimated survival curves across dose groups for each gender for the 26-week Tg(HRAS)2 study for the pooled main study and TK groups. Note that TK animals that are sacrificed are treated as censored.

Figure A.1.1. Kaplan-Meier Survival Curves for 26-Week Tg(HRAS)2 Study Male Mice

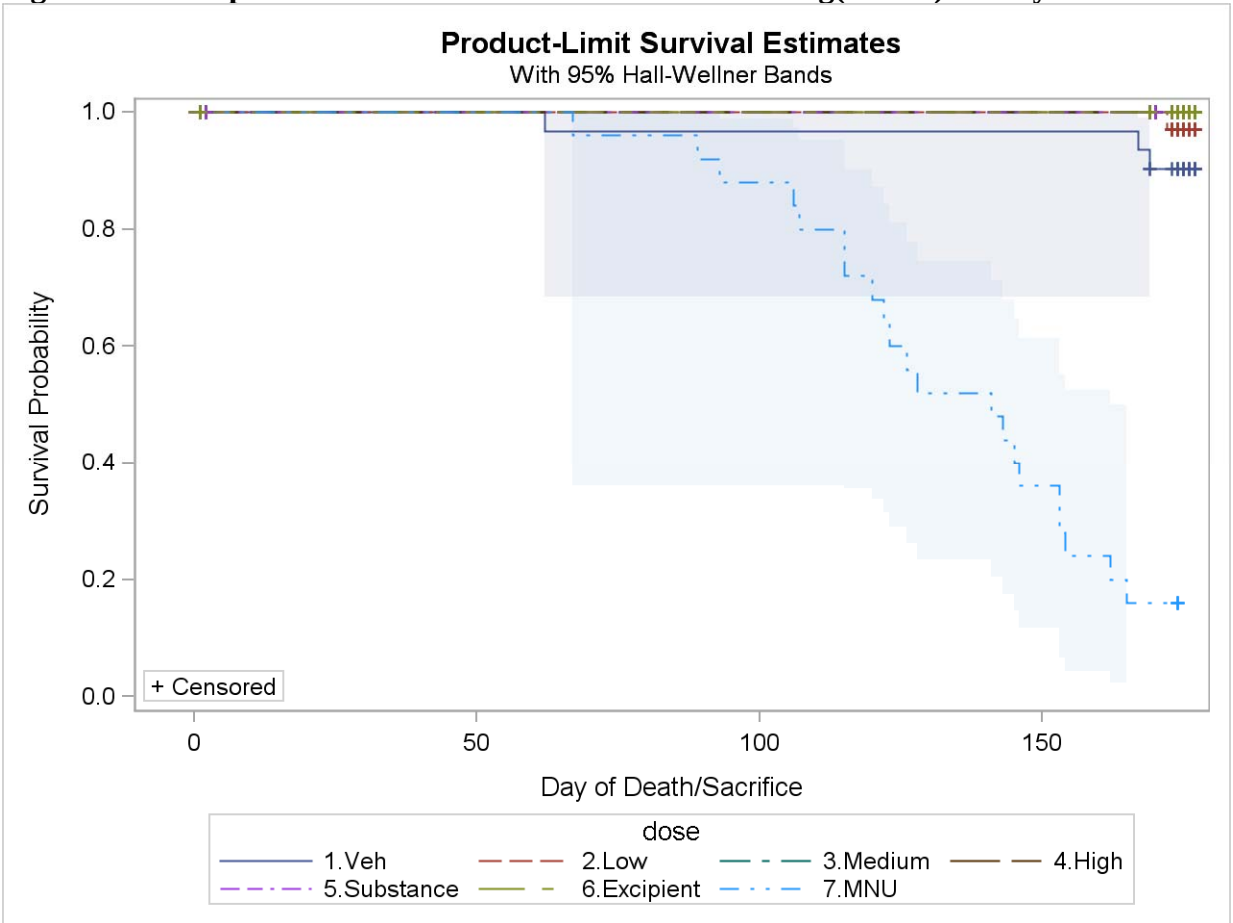
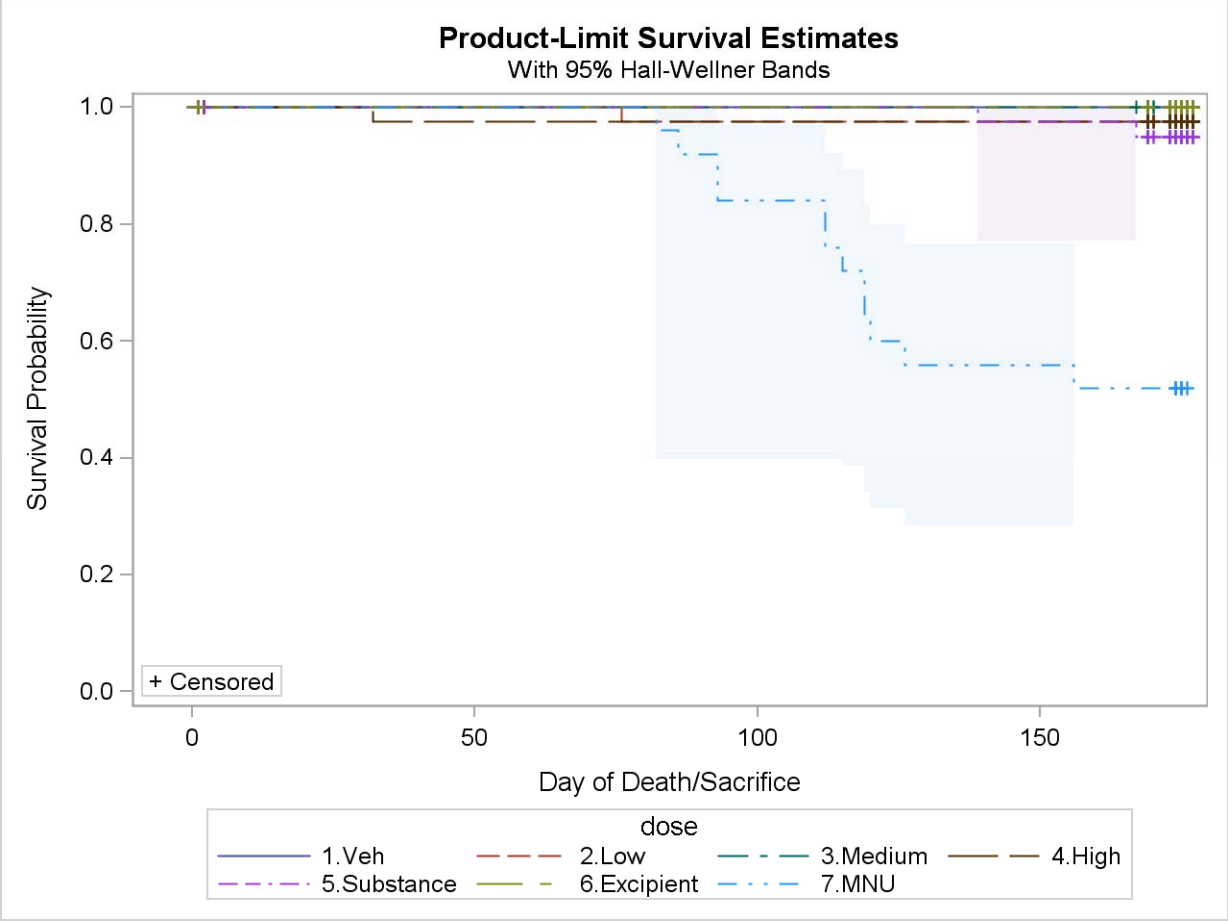


Figure A.1.2. Kaplan-Meier Survival Curves for 26-Week Tg(HRAS)2 Study Female Mice



The following Table A.1.1 summarizes the results of tests comparing survival profiles across dose groups in the tumorigenicity data set for the 22-month study DS95011.

Table A.1.1. Study DS95011: Statistical Significance of Tests of Homogeneity and Trend in Survival

Hypothesis Tested	Males		Females	
	Log rank	Wilcoxon	Log rank	Wilcoxon
Homogeneity over Groups 1-5	0.1033	0.1111	0.7778	0.8381
Homogeneity over Groups 1,3-5	0.1426	0.1919	0.6231	0.7309
No trend over Groups 1,3-5	0.1808	0.1649	0.5616	0.6142
No Difference Between Groups 1 vs 5	0.0676	0.0683	0.3824	0.4037
No Difference Between Groups 2 vs 5	0.0727	0.0465	0.6203	0.5484
No Difference Between Groups 1 vs 2	0.9244	0.9876	0.7168	0.8250

Figures A.1.3. through A.1.4. below, provide survival curves for each gender in the DS95011 study. Note these results are somewhat more interesting than the results from the 26-Week Tg(HRAS)2 study. From Figure A.1.3, in male mice there appears to be a general, though not uniform, trend in decreasing survival over increasing dose. However, whether dealing with all five dose groups or with just the four dose groups remaining after deleting the no treatment group, there is no statistically significant evidence of differences in homogeneity at the usual 0.05 level (all four $p \geq 0.1033$). Despite the apparent trend in male mice, the test that there is no trend is not statistically significant (Logrank $p = 0.1808$, Wilcoxon $p = 0.1649$), suggesting no real evidence of a trend. However the tests between the high dose and vehicle are somewhat close to the usual significance (Logrank $p = 0.0676$, Wilcoxon $p = 0.0683$). The early difference in male mice between the no treatment group and the high dose group leads to Wilcoxon test that is barely statistically significant ($p = 0.0465$). However the corresponding log rank test is not statistically significant ($p = 0.0727$), though close. In Figure A.1.4 in female mice the survival curves are all closely intertwined with no consistent patterns of dominance (all 12 $p \geq 0.3824$), indicating no differences in survival.

Figure A.1.3. Kaplan-Meier Survival Curves for Male Mice in 22-Month Study

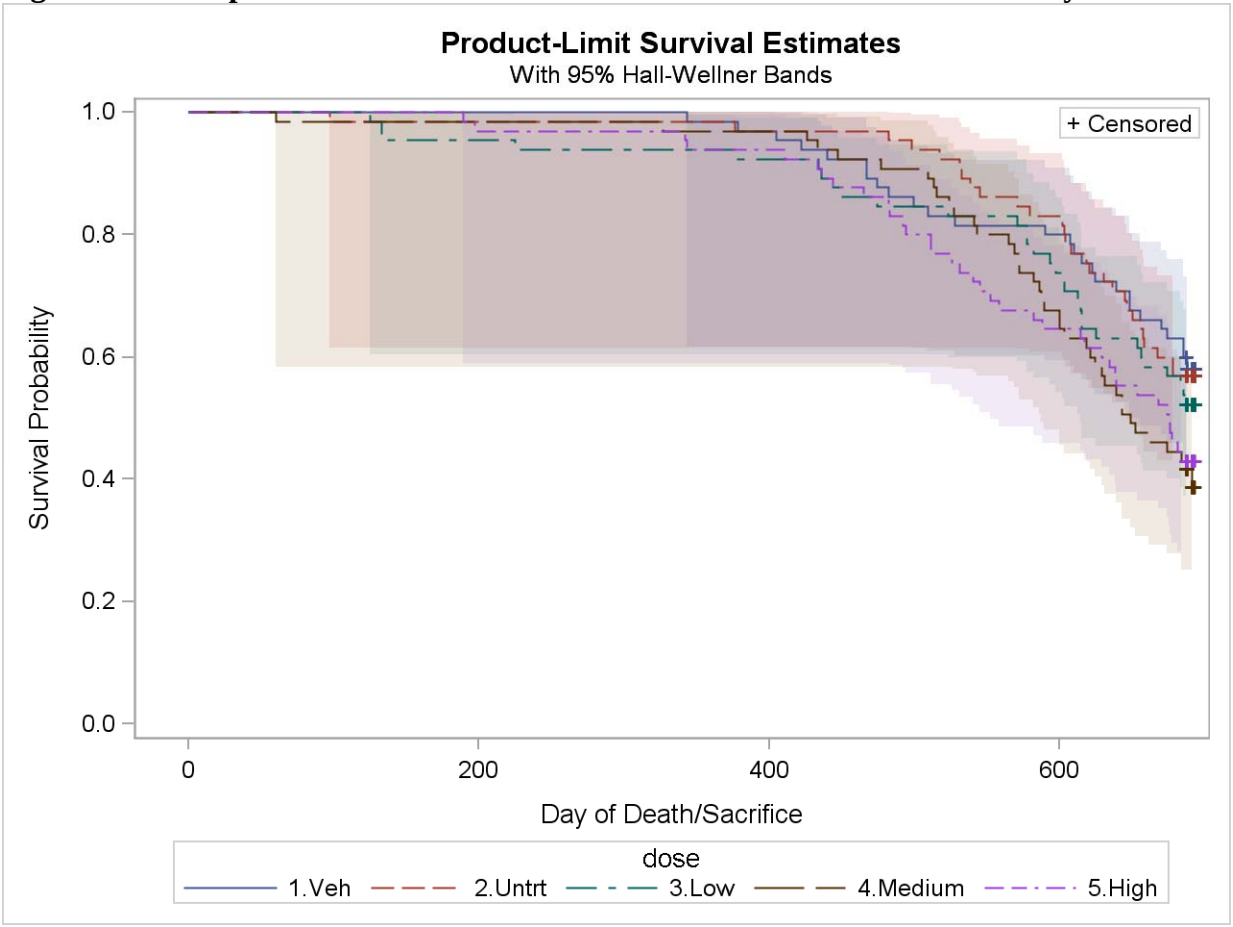
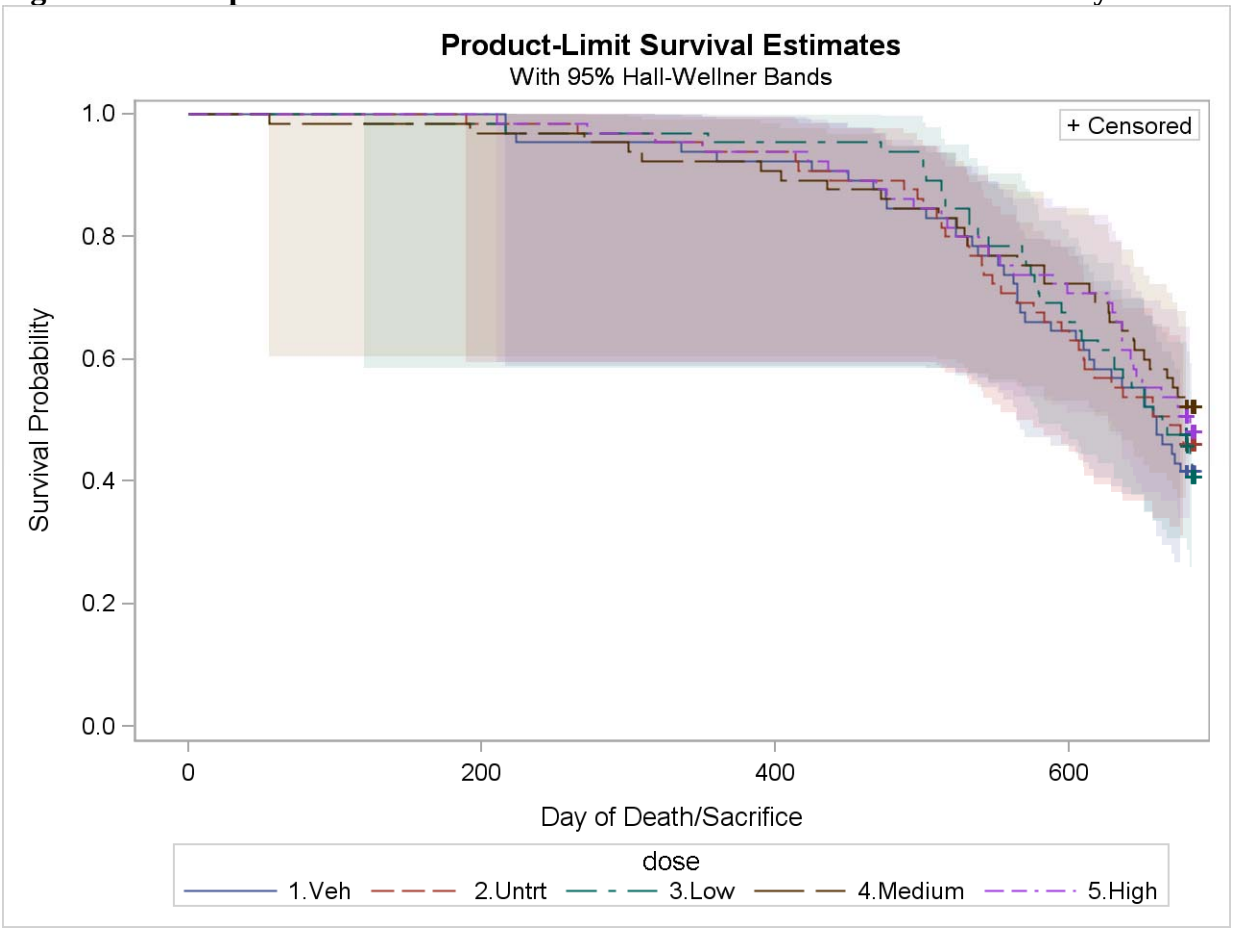


Figure A.1.4. Kaplan-Meier Survival Curves for Female Mice in 22-Month Study



Appendix 2. FDA Nonparametric Bayesian Survival Analysis for Study DS95011

The probability of a subject surviving past time t is given by the survival function, i.e., for random survival time T, S(t) = P(T > t). Statistical inference on survival is based on proposing a probability model for S(t) or one of its derivations. The probability model is defined so that hypotheses to be investigated are specified as parameters in the model. A frequentist analysis takes parameters as fixed and assesses the likelihood of the observed data. A Bayesian analysis starts by noting that parameters are not known, and assumes that a so-called prior probability distribution is a natural measure of this lack of exact knowledge. Then the Bayesian analysis assesses the impact of the actual observed data on this prior. In a nonparametric Bayesian analysis at least one of these parameters is an infinite dimensional space of probability distributions. The nonparametric analysis used here is based upon using a so-called Dependent Dirichlet Process (DDP) as the prior on this space of probability distributions. Note that there is little to no need to model the survival in the 26-Week Tg(HRAS)2 study, and thus it is not addressed here.

Specifically, let T_i denote a random variable representing the survival time of the ith animal. For time until natural death time t_i we write $T_i = t_i$, but if the animal is sacrified at time a_i, all we know is that the time until natural death is greater than a_i, written as $T_i \in (a_i, \infty)$, i.e. T_i is in the time interval (a_i, ∞). Note that animals whose death is in this interval are said to be censored. One useful probability model is to model the logarithm of T_i with a normal distribution, i.e., the T_i are modeled using a lognormal distribution. For this analysis, we model the distribution of $\log(T_i)$ as a mixture of normal distributions weighted by a Dirichlet process on the normal parameters. The prior is defined as a Dirichlet process where the baseline distribution models the linear parameters as a normal distribution on the linear mean parameters and the variance parameters with a Gamma distribution. The prior of the precision parameter of the Dirichlet process is specified as a gamma distribution. The priors for the other hyperparameters are conjugate distributions. Mathematically we can write:

$$\log(T_i) = t_i \,|\, \mathrm{f}_{X_i} \sim \mathrm{f}_{X_i}$$
$$\mathrm{f}_{X_i} = \int N(X_i\beta, \sigma^2) G(d\beta\, d\sigma^2)$$
$$G \,|\, \alpha, G_0 \sim DP(\alpha G_0)$$

The distributions of the hyperparameters above are specified as follows:

$$G_0 = N(\beta \,|\, \mu_b, s_b)\Gamma(\sigma^2 \,|\, \tau_1/2, \tau_2/2)$$
$$\alpha \,|\, a_0, b_0 \sim Gamma(a_0, b_0)$$
$$\mu_b \,|\, m_0, S_0 \sim N(m_0, S_0)$$
$$s_b \,|\, \nu, \Psi \sim InvWishart(\nu, \Psi)$$
$$\tau_2 \,|\, \tau_{s1}, \tau_{s2} \sim Gamma(\tau_{s1}, \tau_{s2})$$

This is an experimental procedure using the DPpackage (Jara, 2007) in R (R Development Core Team, 2009), and, results only should be considered as supporting. The basic reference is de Iorio, et al (2009). The parameterization used to indicate doses was so-called dummy coding, which, in analogy with linear models as discussed in de Iorio et al (2004), implies that effect parameters for treatment doses correspond to the difference with the vehicle controls. That is, the means mubd2, mubd3, and mubd4 in the tables below indicate the differences in posterior location between the vehicle and the low, medium, and high dose groups, respectively. The HPD interval is the estimated highest posterior density interval for these differences. Conditional on the data, the probability that the indicated difference is in the HPD interval is approximately 0.95.

Male Mice

Dummy Parameters	Mean	Median	Std. Dev.	95%HPD-Low	95%HPD-Upper
mub(Intercept)	7.02683	6.99443	0.64871	5.74177	8.36471
mubd2	-0.36669	-0.35430	0.83098	-2.26791	1.05052
mubd3	-0.43648	-0.37362	0.83471	-2.05702	1.20271
mubd4	-0.38031	-0.35299	0.82665	-2.05798	1.17606

Since 0 is near the middle of each HPD interval, there is little evidence of strong differences between the actual treatment groups and the vehicle group.

Slope Parameters	Mean	Median	Std. Dev.	95%HPD-Low	95%HPD-Upper
mub(Intercept)	6.474e+00	6.470e+00	2.338e-01	6.026e+00	6.941e+00
mubd	2.003e-03	2.076e-03	4.486e-03	-5.705e-03	1.013e-02

Again, the HPD interval is near 0, here with all values close to 0. So again, there is little evidence that the slope parameter is non-zero.

Female Mice

Dummy Parameters	Mean	Median	Std. Dev.	95%HPD-Low	95%HPD-Upp
mub(Intercept)	6.506316	6.524730	0.555718	5.525074	7.594967
mubd2	0.198031	0.179197	0.726624	-1.135382	1.677173
mubd3	-0.165499	-0.174304	0.671440	-1.461326	1.295996
mubd4	0.384021	0.366278	0.767093	-1.246231	1.703910

Again, since 0 is near the middle of each HPD interval, there is no evidence of strong differences between the actual treatment groups and the vehicle groups.

Slope Parameters	Mean	Median	Std. Dev.	95%HPD-Low	95%HPD-Upper
mub(Intercept)	6.336e+00	6.339e+00	2.843e-01	5.804e+00	6.894e+00
mubd	9.114e-03	8.015e-03	9.123e-03	-4.224e-03	2.813e-02

Once again, the HPD interval is near 0, with all values close to 0. So again, there is little evidence that the slope parameter is non-zero.

Appendix 3. FDA Poly-k Tumorigenicity Analysis

The poly-k test, here with k=3, modifies the original Cochran-Armitage test to adjust for differences in mortality (please see Bailer & Portier, 1988, Bieler & Williams, 1993). The tests used here are small sample exact permutation tests of tumor incidence. These do assume all marginal totals are fixed, a debatable assumption. This assumption implies that in the pairwise tests when one dose group has no tumors of the specific type and the other does, there is only one permutation of this pattern. Since that means that the only permutation of the data is the one observed, that means that all possible permutations are as extreme as the pattern observed, and thus the significance level of the observed pattern can be logically expressed as 1.0. One could use the same sort of argument when there were no tumors of the specific type being analyzed in either column of the 2x2 table corresponding to a pairwise comparison. Then an argument could be made that the p-value for this test should also be 1.0. However, largely for readability, in the tables below these p-values are considered as missing (i.e., corresponding to a null test), denoted by ".". Note that StatXact adjusts for the variance, which would be 0. Then the significance levels of the test statistics are based on the result of a division by 0, i.e., undefined, and hence StatXact codes these p-values as missing.

Up until fairly recently, the Biometrics Division has usually emphasized so-called Peto carcinogenicity tests, asm used by the Sponsor. However this methodology has been replaced by the poly-k modification of the Cochran-Armitage tests of trend for reasons discussed in Section 1.3.1.3.

To adjust for the multiplicity of tests the so-called Haseman-Lin-Rahman rules discussed in Section 1.3.1.4 are often applied. That is, when testing for trend over dose and the difference between the highest dose group with a control group, to control the overall Type I error rate to roughly 10% for a standard two species, two sex study, one compares the unadjusted significance level of the trend test to 0.005 for common tumors (incidence > 1%) and 0.025 for rare tumors, and the pairwise test to 0.01 for common tumors and 0.05 for rare tumors. Even assuming they are applicable to the Tg(HRAS)2 study using these adjustments for other tests, like the trend over the vehicle, low, and medium dose groups and the pairwise comparisons between the vehicle and the medium and low dose groups and the no treatment and excipient groups, etc., can be expected to increase the overall type I error rate to some value above the nominal rough 10% level, possibly considerably higher than the nominal 10% rate.

Except for the active control, group 7, there were few tumors in the Tg.(HRAS)2 study, and thus, not surprisingly, no statistically significant differences or tests of trend. The only statistically significant results were observed in Study DS95011. These are presented in Tables A.3.1 and A.3.2 below. The test of trend in this latter study, summarized below, is over the vehicle, low, medium, and high dose groups. Significance levels of pairwise tests between the high, medium, and low dose group are on the same line. Underneath the p-values of these pairwise tests to vehicle are the results of the corresponding pairwise tests to the no treatment

group. Note that there is a clear difference in the rate of neoplasms in the vehicle group versus the no treatment group. In male mice, following the Haseman-Lin-Rahman adjustment for multiplicity to get an overall rough 10% error rate and using the incidence in the no treatment group to decide if a tumor is rare or not, we would conclude, that the test of trend in both adenomas and pooled adenoma and adenocarcinoma of the lacrimal/Harderian glands was statistically significant (both $p < 0.00005 < 0.005$), as were the pairwise comparions between the high dose group and controls (both $p < 0.00005 < 0.01$). Similarly, the pairwise comparisons between the high dose group and the no treatment groups would be statistically significant (each p=0.0002, 0.0001, both <0.01). Similarly, the pairwise comparisons between the high dose group and the no treatment groups would be statistically significant (each p=0.0002, 0.0001, both <0.01). Applying the Haseman-Lin-Rahman rules to the comparisons of the high dose with the no treatment group can be expected to increase the overall Type I error to some value above the 10% level. Performing such tests would lead to statistically significant differences between the high dose and the no treatment group in adenoma and pooled adenoma/adenocarcinoma of the lacrimal harderian gland ($p = 0.0002, 0.0001 < 0.05$, respectively). Similar comparisons in hepatocellular adenoma, and pooled adenoma and carcinoma of the liver would also be statistically significant ($p = 0.0018, 0.0007$, both < 0.05), as would be the tests of bronchiolar-alveolar adenoma, carcinoma, and pooled adenoma and carcinoma ($p = 0.0008, 0.0461, 0.00005$, all < 0.05).

Table A.3.1. Study DS95011: Tumor Incidences and Statistical Significance of Tests of Trend and Pairwise Differences in Male Standard Lab Mice with at least One Significant Test

Organ / Tumor	Incidence (No Veh)	(Trt Low)	(Med)	(Hi)	Trend	Significance Levels (Vehicle/No Treat versus High)	(Medium)	(Low)	
LACRIMAL/HARDIRIAN GLANDS									
ADENOMA	6	11	3	8	28	0.0000	0.0000	0.3448	0.9055 Veh
							0.0002	0.7973	0.9940 No
Adenoma/Adenocarcinoma	6	11	4	8	29	0.0000	0.0000	0.3448	0.8162 Veh
							0.0001	0.7973	0.9827 No
LIVER									
ADENOMA, HEPATOCELLULAR	19	0	8	12	8	0.8928	0.9896	0.9281	0.9934 Veh
							0.0018	0.0001	0.0023 No
Adenoma/Carcinoma Hepato.	20	0	11	14	9	0.9183	0.9870	0.8832	0.9746 Veh
							0.0007	0.0000	0.0002 No
CARCINOMA, HEPATOCELLULAR	4	0	3	5	1	0.8938	0.9639	0.4759	0.7655 Veh
							0.4712	0.0244	0.1113 No
LUNG									
ADENOMA, BRONCHIOLO-ALVEOLAR	12	0	10	20	9	0.8196	0.7529	0.0554	0.7155 Veh
							0.0008	0.0000	0.0005 No
Adenoma/Carcinoma Bronch.-Alv.	15	0	15	24	13	0.7955	0.6445	0.0410	0.5145 Veh
							0.0000	0.0000	0.0000 No
CARCINOMA, BRONCHIOLO-ALVEOLAR	3	0	5	4	4	0.4150	0.4352	0.4576	0.3268 Veh
							0.0461	0.0503	0.0244 No
LYMPHORETICULAR SYSTEM									
LYMPHOMA	4	1	6	3	4	0.5370	0.5648	0.7319	0.3245 Veh
							0.1426	0.2673	0.0457 No
Lymphoma/Leukemia, Myleogenous	5	1	6	3	4	0.6071	0.6845	0.8250	0.4475 Veh
							0.1426	0.2673	0.0457 No
SPLEEN									
HEMANGIOSARCOMA	1	0	3	5	1	0.7467	0.7276	0.0900	0.2947 Veh
							0.4712	0.0232	0.1113 No

Table A.3.1. (cont.) Study DS95011: Tumor Incidences and Statistical Significance of Tests of Trend and Pairwise Differences in Male Standard Lab Mice with at least One Significant Test

Organ / Tumor	Incidence No Veh	Trt	Low	Med	Hi	Trend	Significance Levels Vehicle/No Treat versus High	Medium	Low	
Systemic										
HEMANGIOSARCOMA	4	0	5	11	3	0.8398	0.7310	0.0393	0.4642	Veh
							0.1012	0.0002	0.0244	No
Hemangioma/Hemangiosarcoma	5	0	9	14	3	0.9627	0.8243	0.0186	0.1754	Veh
							0.1012	0.0000	0.0011	No

Adding comparisons with other dose groups will only increase the expected Type I error rate even further (i.e., further decrease specificity). Accepting this effect, one would conclude that comparisons between the medium dose group and the vehicle in systemic hemangioma and pooled hemangioma and hemangiosarcoma would be statistically significant ($p = 0.0393$, 0.0186, both < 0.05). Corresponding results from the comparison to the no treatment group are more extreme ($p = 0.0002$, $p < 0.00005$, both < 0.05). Further, we could conclude that comparisons between the medium dose group and the no treatmen group in hepatocellular adenoma, carcinoma, and pooled adenoma and carcinoma of the liver would also be statistically significant ($p = 0.0001$, $p=0.0244$ and $p <0.00005$, all < 0.05), as would be the corresponding tests in bronchiolar-alveolar adenoma and pooled adenoma and carcinoma in the lung (both $p < 0.00005 < 0.01$). Results are only slightly less extreme for the comparisons of the low dose group to the no treatment group. Comparisons between the low dose group and the vehicle in systemic hemangioma and pooled hemangioma and hemangiosarcoma would also be statistically significant ($p = 0.0244$, 0.0011, both < 0.05). Comparisons between the low dose group and the no treatment group in hepatocellular adenoma and pooled adenoma and carcinoma of the liver would also be statistically significant ($p = 0.0023$, 0.0002, both < 0.05), as would be the corresponding tests in bronchiolar-alveolar adenoma, carcinoma, and pooled adenoma and carcinoma in the lung ($p = 0.0005$, $p=0.0244$, $p< 0.00005$, all < 0.05). Finally, the difference between medium dose group and the no treatment group in hemangiosarcoma is statistically significant ($p = 0.0232 < 0.05$).

Table A.3.2 below summarizes similar results in female mice:.

Table A.3.2. Study DS95011: Tumor Incidences and Statistical Significance of Tests of Trend and Pairwise Differences in Female Standard Lab Mice with at least One Significant Test

Organ / Tumor	Incidence (No Trt)					Significance Levels				
	Veh	Low	Med	Hi		Trend	Vehicle/No Treat versus High	Medium	Low	
LACRIMAL/HARDIRIAN GLANDS										
ADENOMA	1	4	2	2	11	0.0001	0.0030	0.5230	0.5230	Veh
							0.0586	0.9067	0.9067	No
Adenoma/Adenocarcinoma	1	4	2	2	13	0.0000	0.0008	0.5230	0.5230	Veh
							0.0240	0.9067	0.9067	No
LIVER										
ADENOMA, HEPATOCELLULAR	1	0	2	5	1	0.7268	0.7675	0.1167	0.5230	Veh
							0.5100	0.0312	0.2576	No
Adenoma/Carcinoma Hepato.	1	0	2	6	1	0.7760	0.7675	0.0658	0.5230	Veh
							0.5100	0.0151	0.2576	No
LUNG										
ADENOMA, BRONCHIOLO-ALVEOLAR	2	9	9	15	11	0.1847	0.0100	0.0008	0.0300	Veh
							0.4403	0.1580	0.6202	No
Adenoma/Carcinoma Bronch.-Alv.	5	14	10	17	20	0.0071	0.0013	0.0057	0.1501	Veh
							0.2197	0.3821	0.8889	No
CARCINOMA, BRONCHIOLO-ALVEOLAR	3	5	1	4	9	0.0058	0.0858	0.5226	0.9460	Veh
							0.2534	0.7762	0.9883	No
LYMPHORETICULAR SYSTEM										
LYMPHOMA	17	10	12	22	21	0.1831	0.3744	0.3045	0.9055	Veh
							0.0437	0.0290	0.4673	No
MAMMARY GLAND/REGION										
Adenoma/Adenocarc./Fibroadenoma	1	0	3	3	5	0.0826	0.1165	0.3315	0.3240	Veh
							0.0328	0.1326	0.1288	No
OVARIES										
Leiomyoma/Lute-/Thec-/G-cell Tmr	2	0	5	3	2	0.7473	0.7140	0.5290	0.2438	Veh
							0.2576	0.1288	0.0312	No
Systemic										
HEMANGIOSARCOMA	6	0	7	3	4	0.7739	0.8563	0.9289	0.5454	Veh
							0.0663	0.1326	0.0083	No
Hemangioma/Hemangiosarcoma	8	0	9	5	8	0.4744	0.6397	0.8972	0.5528	Veh
							0.0037	0.0328	0.0019	No
UTERUS										
ADENOCARCINOMA, ENDOMETRIAL	0	0	2	0	4	0.0267	0.0664	.	0.2628	Veh
							0.0637	.	0.2576	No
ENDOMETRIAL STROMAL POLYP	3	0	5	7	3	0.7305	0.6705	0.1673	0.3688	Veh
							0.1288	0.0072	0.0312	No
LEIOMYOMA	2	0	5	2	1	0.8949	0.8897	0.7140	0.2525	Veh
							0.5100	0.2576	0.0328	No
Leiomyoma/-sarcoma/stromal sarc.	2	0	6	5	1	0.9294	0.8897	0.2525	0.1691	Veh
							0.5100	0.0328	0.0170	No

In female mice, the tests of trend in adenoma and adenocarcinoma of the Lacrimal/Hardirian glands would be considered as statistically significant ($p = 0.0001$ and $p < 0.00005$, both < 0.005). In these same glands the corresponding pairwise comparisons between the high dose and vehicle would classified as statistically significant ($p = 0.0030$ and $p = 0.0008$, both < 0.01).). Note that in the lung, the test of trend in bronchiolar-alveolar carcinoma and pooled adenoma and carcinoma would be close to significant ($p = 0.0058$, $0.0071 > 0.005$), while pairwise tests between the high dose group and vehicle in adenoma and pooled adenoma would be statistically significant ($p = 0.010 = 0.01$ and $p = 0.0013 < 0.01$). Adding comparisons of the high dose group with the no treatment group would be expected to increase overall Type I error. Accepting this inflation we would conclude the pairwise test

between the high dose group and the no treatment group in pooled systemic hemangioma and hemangiosarcomas would also be statistically significant ($p = 0.0037 < 0.01$), as would be the test in pooled mammary gland tumors ($p = 0.0328 < 0.05$). No other tests of trend or pairwise comparison of the high dose to vehicle or no treatment group are statistically significant.

Again, comparisons with other dose groups will only further increase the expected Type I error rate and thus decrease specificity. Accepting this cost, one would conclude that for bronchiolar-alveolar carcinoma and pooled adenoma and carcinoma, the tests of differences between the medium dose group and vehicle would statistically significant ($p = 0.0008, 0.0057 < 0.01$ Further, comparisons between the high, medium, and low dose groups with the no treatment group in systemic pooled hemangioma and hemangiosarcoma would be also statistically significant ($p = 0.0037, 0.0328, 0.0019$, all < 0.05), as would be the comparison for hemangioma in the low dose group ($p = 0.0083 < 0.05$). Comparisons between the high and medium dose group with the no treatment group in adenoma and pooled adenoma and carcinoma of the liver would also be statistically significant ($p = 0.0312, 0.0151$ both < 0.05), Comparisons between the medium dose group with the no treatment group in adenoma and pooled adenoma and carcinoma of the liver would also be statistically significant ($p = 0.0312, 0.0151$ both < 0.05), as would be the test between the low dose group and the no treatment group for various pooled tumors in the ovaries ($p = 0. 0312 < 0.05$). The comparisons between the medium and low dose group and the no treatment group in endrometrial stromal polyp of the uterus would be statistically significant ($p = 0.0072, 0.0312$, both < 0.05), as would be the comparisons for pooled leiomyoma, leiomysarcoma, and stromal sarcoma emangioma in the low dose group ($p = 0.0328, 0.0170$, both < 0.05)). Finally the test between the low dose and the no treatment group in leiomyoma was statistically significant ($p = 0.0328 < 0.05$).

The remaining tables give complete incidences of all labeled neoplasms. The table for the Tg(HRAS)2 study is complicated. For each organ-tumor combination the first line provides tumor incidence over the Sponsor labeled vehicle (i.e., water), low, medium, high, substance, excipient, and MNU groups (i.e., groups 1-7) respectively. The first line continues with significance level of the test of trend over the first four groups, and the results of the pairwise tests of the high, medium, and low dose groups against vehicle. Note that it may make more sense to use the excipient group 6 as the vehicle control. The corresponding tests of trend and pairwise differences using this as the vehicle group are presented on the second line for the organ-tumor combination, after the vertical line "|". Preceding these are the significance levels of the pairwise tests of the excipient versus the Sponsor's vehicle, i.e. water, the drug substance versus water, and finally the medium dose group versus the drug substance (at the same nominal dose).

Table A.3.3. Tg(HRAS)2 Study: Tumor Incidences and Statistical Significance of Tests of Trends and Pairwise Differences in Male Mice

Organ / Tumor	Incidence/Signif Levels Veh	Low	Med	Hi	Sub	Ex	MNU	Significance Levels Vehicle/Excipient versus Trend	High	Medium	Low	
	(Ex vs| Veh	Sub vs| Veh	Med vs) Sub								
Brain												
Lymphosarcoma	0	0	0	1	0	0	5	0.2500	0.5000	.	.	Veh
	.		.				|	0.2500	0.5000	.	.	Ex
Harderian gls.												
Adenoma	0	0	0	0	1	0	0	.	.	.	.	Veh
	.		0.5000	0.5000			|	.	.	.	.	Ex
Heart												
Lymphosarcoma	0	0	0	1	0	0	11	0.2500	0.5000	.	.	Veh
	.		.				|	0.2500	0.5000	.	.	Ex
Joints,knee												
Lymphosarcoma	0	0	0	1	0	0	7	0.2500	0.5000	.	.	Veh
	.		.				|	0.2500	0.5000	.	.	Ex
Kidneys												
Lymphosarcoma	0	0	0	1	0	0	10	0.2500	0.5000	.	.	Veh
	.		.				|	0.2500	0.5000	.	.	Ex
Liver												
Lymphosarcoma	0	0	0	1	0	0	11	0.2500	0.5000	.	.	Veh
	.		.				|	0.2500	0.5000	.	.	Ex
Lungs/Bronchi												
A/B adenoma	0	1	0	0	1	0	1	0.7500	.	.	0.5000	Veh
	.		0.5000	0.5000			|	0.7500	.	.	0.5000	Ex
Bronch alveolar carcinoma	0	0	0	0	1	0	0	.	.	.	.	Veh
	.		0.5000	0.5000			|	.	.	.	.	Ex
Bronch. Alv.Adenoma/Carc.	0	1	0	0	2	0	1	0.7500	.	.	0.5000	Veh
	.		0.2449	0.2449			|	0.7500	.	.	0.5000	Ex
Lymphosarcoma	0	0	0	1	0	0	11	0.2500	0.5000	.	.	Veh
	.		.				|	0.2500	0.5000	.	.	Ex
Lymph ns.,mand.												
Lymphosarcoma	0	0	0	1	0	0	8	0.2500	0.5000	.	.	Veh
	.		.				|	0.2500	0.5000	.	.	Ex
Lymphoreticular												
Lymphosarcoma	0	0	0	1	0	0	13	0.2500	0.5000	.	.	Veh
	.		.				|	0.2500	0.5000	.	.	Ex
Pituitary												
Lymphosarcoma	0	0	0	1	0	0	4	0.2500	0.5000	.	.	Veh
	.		.				|	0.2500	0.5000	.	.	Ex
Reticuloendothel												
Lymphosarcoma	0	0	0	1	0	0	13	0.2500	0.5000	.	.	Veh
	.		.				|	0.2500	0.5000	.	.	Ex
Skin mammary												
Lymphosarcoma	0	0	0	1	0	0	0	0.2500	0.5000	.	.	Veh
	.		.				|	0.2500	0.5000	.	.	Ex
Spleen												
Hemangiosarcoma	2	3	0	3	1	3	1	0.4552	0.5000	1.0000	0.5000	Veh
	0.5000		0.8827	0.5000			|	0.6071	0.6664	1.0000	0.6664	Ex
Stomach												
Forestomach papilloma	0	0	0	0	1	1	13	.	.	.	.	Veh
	0.5000		0.5000	0.5000			|	1.0000	1.0000	1.0000	1.0000	Ex
Lymphosarcoma	0	0	0	1	0	0	0	0.2500	0.5000	.	.	Veh
	.		.				|	0.2500	0.5000	.	.	Ex
Systemic												
Hemangiosarcoma	2	3	0	3	1	3	2	0.4552	0.5000	1.0000	0.5000	Veh
	0.5000		0.8827	0.5000			|	0.6071	0.6664	1.0000	0.6664	Ex
Lymphosarcoma	0	0	0	1	0	0	13	0.2500	0.5000	.	.	Veh
	.		.				|	0.2500	0.5000	.	.	Ex

Table A.3.3. (cont.) Tg(HRAS)2 Study: Tumor Incidences and Statistical Significance of Tests of Trends and Pairwise Differences in Male Mice

Organ / Tumor	Incidence/Signif Levels Veh Low Med Hi Sub Ex MNU (Ex vs| Sub vs|Med vs) Veh Veh Sub							Significance Levels Vehicle/Excipient versus Trend	High	Medium	Low	
Thymus												
Lymphosarcoma	0	0	0	1	0	0	12	0.2500	0.5000	.	.	Veh
	.	.	.				|	0.2500	0.5000	.	.	Ex
Thymoma	0	0	0	0	0	1	0	.	.	.	.	Veh
	0.5000	.	.				|	1.0000	1.0000	1.0000	1.0000	Ex
Thyrs./paraths.												
Lymphosarcoma	0	0	0	1	0	0	1	0.2500	0.5000	.	.	Veh
	.	.	.				|	0.2500	0.5000	.	.	Ex
Trachea												
Lymphosarcoma	0	0	0	1	0	0	2	0.2500	0.5000	.	.	Veh
	.	.	.				|	0.2500	0.5000	.	.	Ex

Table A.3.4. Tg(HRAS)2 Study: Tumor Incidences and Statistical Significance of Tests of Trends and Pairwise Differences in Female Mice

Organ / Tumor	Incidence/Signif Levels Veh Low Med Hi Sub Ex MNU (Ex vs| Sub vs|Med vs) Veh Veh Sub							Significance Levels Vehicle/Excipient versus Trend	High	Medium	Low	
Adrenals												
Lymphosarcoma	0	0	1	0	0	1	3	0.5000	.	0.5102	.	Veh
	0.5102	.	1.0000				|	.	.	.	.	Ex
Aorta,thoracic												
Lymphosarcoma	0	0	1	0	0	1	4	0.5000	.	0.5102	.	Veh
	0.5102	.	1.0000				|	.	.	.	.	Ex
Bone,sternum												
Lymphosarcoma	0	0	1	0	0	1	6	0.5000	.	0.5102	.	Veh
	0.5102	.	1.0000				|	.	.	.	.	Ex
Brain												
Lymphosarcoma	0	0	1	0	0	0	1	0.5000	.	0.5102	.	Veh
	.	.	1.0000				|	.	.	.	.	Ex
Clitoral gls.												
Lymphosarcoma	0	0	1	0	0	1	1	0.5000	.	0.5102	.	Veh
	0.5102	.	1.0000				|	.	.	.	.	Ex
Gallbladder												
Lymphosarcoma	0	0	0	0	0	1	2	.	.	.	.	Veh
	0.5102	.	.				|	.	.	.	.	Ex
Harderian gls.												
Lymphosarcoma	0	0	1	0	0	1	2	0.5000	.	0.5102	.	Veh
	0.5102	.	1.0000				|	.	.	.	.	Ex
Heart												
Lymphosarcoma	0	0	1	0	0	1	4	0.5000	.	0.5102	.	Veh
	0.5102	.	1.0000				|	.	.	.	.	Ex
Joints,knee												
Lymphosarcoma	0	0	1	0	0	1	6	0.5000	.	0.5102	.	Veh
	0.5102	.	1.0000				|	.	.	.	.	Ex
Kidneys												
Lymphosarcoma	0	0	1	0	0	1	6	0.5000	.	0.5102	.	Veh
	0.5102	.	1.0000				|	.	.	.	.	Ex
Liver												
Lymphosarcoma	0	0	1	0	0	1	6	0.5000	.	0.5102	.	Veh
	0.5102	.	1.0000				|	.	.	.	.	Ex

Table A.3.4. (cont.) Tg(HRAS)2 Study: Tumor Incidences and Statistical Significance of Tests of Trends and Pairwise Differences in Female Mice

Organ / Tumor	Incidence/Signif Levels Veh Low Med Hi Sub Ex MNU (Ex vs| Sub vs|Med vs) Veh Veh Sub							Significance Levels Vehicle/Excipient versus Trend	High	Medium	Low	
Lungs/Bronchi												
A/B adenoma	1	0	1	1	1	0	1	0.4034	0.7553	0.7653	1.0000	Veh
	1.0000	0.7653	0.7551					.	.	.	.	Ex
Bronch. Alv.Adenoma/Carc.	1	0	1	1	1	0	3	0.4034	0.7553	0.7653	1.0000	Veh
	1.0000	0.7653	0.7551					.	.	.	.	Ex
Lymphosarcoma	0	0	1	0	0	1	6	0.5000	.	0.5102	.	Veh
	0.5102	.	1.0000					.	.	.	.	Ex
Lymph ns.,mand.												
Lymphosarcoma	0	0	1	0	0	1	5	0.5000	.	0.5102	.	Veh
	0.5102	.	1.0000					.	.	.	.	Ex
Lymph ns.,mes.												
Lymphosarcoma	0	0	1	0	0	0	3	0.5000	.	0.5102	.	Veh
	.	.	1.0000					.	.	.	.	Ex
Lymphoreticular												
Lymphosarcoma	0	0	1	0	0	1	9	0.5000	.	0.5102	.	Veh
	0.5102	.	1.0000					.	.	.	.	Ex
Mammary gl., F												
Lymphosarcoma	0	0	1	0	0	1	0	0.5000	.	0.5102	.	Veh
	0.5102	.	1.0000					.	.	.	.	Ex
Ovaries												
Lymphosarcoma	0	0	1	0	0	1	4	0.5000	.	0.5102	.	Veh
	0.5102	.	1.0000					.	.	.	.	Ex
Oviducts												
Lymphosarcoma	0	0	1	0	0	1	3	0.5000	.	0.5102	.	Veh
	0.5102	.	1.0000					.	.	.	.	Ex
Pancreas												
Lymphosarcoma	0	0	1	0	0	1	1	0.5000	.	0.5102	.	Veh
	0.5102	.	1.0000					.	.	.	.	Ex
Reticuloendothel												
Lymphosarcoma	0	0	1	0	0	1	8	0.5000	.	0.5102	.	Veh
	0.5102	.	1.0000					.	.	.	.	Ex
Salivary gls.												
Lymphosarcoma	0	0	0	0	0	1	1	.	.	.	.	Veh
	0.5102	.	.					.	.	.	.	Ex
Skin												
Hemangiosarcoma	0	0	0	0	1	0	0	.	.	.	.	Veh
	.	0.5102	0.5000					.	.	.	.	Ex
Skin mass(es)												
papilloma	1	1	0	0	1	0	11	0.9382	1.0000	1.0000	0.7551	Veh
	1.0000	0.7449	0.4898					.	.	.	.	Ex
Spleen												
Hemangiosarcoma	0	1	2	1	1	1	4	0.2948	0.5000	0.2551	0.5102	Veh
	0.5102	0.5000	0.8752					.	.	.	.	Ex
Lymphosarcoma	0	0	1	0	0	1	7	0.5000	.	0.5102	.	Veh
	0.5102	.	1.0000					.	.	.	.	Ex
Stomach												
Forestomach papilloma	0	0	0	1	0	0	16	0.2449	0.5000	.	.	Veh
	.	.	.					.	.	.	.	Ex
Hemangiosarcoma	0	0	0	0	0	1	0	.	.	.	.	Veh
	0.5102	.	.					.	.	.	.	Ex
Lymphosarcoma	0	0	0	0	0	1	0	.	.	.	.	Veh
	0.5102	.	.					.	.	.	.	Ex

Table A.3.4. (cont.) Tg(HRAS)2 Study: Tumor Incidences and Statistical Significance of Tests of Trends and Pairwise Differences in Female Mice

Organ / Tumor	Incidence/Signif Levels Veh	Low	Med	Hi	Sub	Ex	MNU	Significance Levels (Ex vs| Sub vs|Med vs) Vehicle/Excipient versus Trend	High	Medium	Low	
Systemic												
Hemangioma	1	0	0	0	0	0	0	1.0000	1.0000	1.0000	1.0000	Veh
	1.0000	1.0000	.					.	.	.	.	Ex
Hemangioma/-sarcoma	1	1	2	1	2	2	6	0.4831	0.7449	0.5000	0.7551	Veh
	0.5000	0.5000	0.6954					.	.	.	.	Ex
Hemangiosarcoma	0	1	2	1	2	2	6	0.2948	0.5000	0.2551	0.5102	Veh
	0.2551	0.2551	0.6954					.	.	.	.	Ex
Lymphosarcoma	0	0	1	0	0	1	9	0.5000	.	0.5102	.	Veh
	0.5102	.	1.0000					.	.	.	.	Ex
Thymus												
Lymphosarcoma	0	0	1	0	0	1	8	0.5000	.	0.5102	.	Veh
	0.5102	.	1.0000					.	.	.	.	Ex
Thyrs./paraths.												
Lymphosarcoma	0	0	0	0	0	1	1	.	.	.	.	Veh
	0.5102	.	.					.	.	.	.	Ex
Urinary bladder												
Lymphosarcoma	0	0	1	0	0	1	2	0.5000	.	0.5102	.	Veh
	0.5102	.	1.0000					.	.	.	.	Ex
Uterus												
Hemangioma	1	0	0	0	0	0	0	1.0000	1.0000	1.0000	1.0000	Veh
	1.0000	1.0000	.					.	.	.	.	Ex
Hemangioma/-sarcoma	1	0	0	0	0	0	1	1.0000	1.0000	1.0000	1.0000	Veh
	1.0000	1.0000	.					.	.	.	.	Ex
Lymphosarcoma	0	0	1	0	0	1	4	0.5000	.	0.5102	.	Veh
	0.5102	.	1.0000					.	.	.	.	Ex
Vagina												
Lymphosarcoma	0	0	1	0	0	1	2	0.5000	.	0.5102	.	Veh
	0.5102	.	1.0000					.	.	.	.	Ex

The following table provides incidences for those organ-tumor combinations that occurred only in the active control MNU group.

Table A.3.5. Tg(HRAS)2 Study: Organ -Tumor Combinations Observed in Group 7 (MNU) Only (i.e. 0 incidence in other groups)

Organ	Tumor	Group 7 Incidence
Male Animals		
Adrenals	Lymphosarcoma	6
	Squamous cell carcinoma	1
Aorta,thoracic	Lymphosarcoma	8
Bone,sternum	Lymphosarcoma	9
Colon	Squamous cell carcinoma	1
Duodenum	Lymphosarcoma	1
Esophagus,thor.	Lymphosarcoma	3
GI System	Adenomas/Adenocarcinomas	1
Gallbladder	Lymphosarcoma	2
Harderian gls.	Lymphosarcoma	2
Jejunum	Adenocarcinoma	1
Lymph ns.,mes.	Lymphosarcoma	4
Lymph ns.,renal	Lymphosarcoma	1
Mouth	Papilloma	1
Oral Cavity/Fore	Sq.Cell Papillomas/carcinonoma	5
Pancreas	Lymphosarcoma	1
	Squamous cell carcinoma	1
Prostate	Transitional cell papillom	1

Table A.3.5. (cont.) Tg(HRAS)2 Study: Organ -Tumor Combinations Observed in Group 7 (MNU) Only (i.e. 0 incidence in other groups)

Organ	Tumor	Group 7 Incidence
Male Animals (cont.)		
Salivary gls.	Squamous cell carcinoma	1
Skin	Sq.Cell Neoplasms/Keratoca	2
	Sq.Cell Papillomas/carcinonoma	1
	Squamous Papilloma	1
Skin mass(es)	keratoacanthoma	1
	papilloma	1
Sp.Cord,lumb	Hemangiosarcoma	1
	Lymphosarcoma	1
Spleen	Lymphosarcoma	10
Stomach	Forestomach SCC	3
	Squamous cell carcinoma	1
Urinary bladder	Papilloma	1
Female Animals		
Cecum	Lymphosarcoma	2
Duodenum	Lymphosarcoma	2
Ear	Papilloma	1
Esophagus,thor.	Papilloma	1
GI System	Adenomas/Adenocarcinomas	1
Jejunum	Adenocarcinoma	1
	Lymphosarcoma	2
Lungs/Bronchi	Bronch alveolar carcinoma	2
Lymph node(s)	Lymphosarcoma	1
Oral Cavity/Fore	Sq.Cell Papillomas/carcinonoma	2
Salivary gls.	Squamous cell carcinoma	1
Skin	Sq.Cell Neoplasms/Keratoca	3
	Sq.Cell Papillomas/carcinonoma	3
	Squamous Papilloma	2
Skin mass(es)	Hemangiosarcoma	1
	Squamous cell carcinoma	1
Stomach	Squamous cell carcinoma	1
Trachea	Lymphosarcoma	1
	Papilloma	1
Uterus	Hemangiosarcoma	1
Vagina	Papilloma	2

Table A.3.6. Study DS95011: Tumor Incidences and Statistical Significance of Tests of Trend and Pairwise Differences in Male Standard Lab Mice

Organ / Tumor	Incidence No Veh	Trt Low	Med	Hi	Significance Levels Vehicle/No Treat versus Trend	High	Medium	Low
ADRENAL GLANDS								
ADENOMA, CORTICAL	2	0 1	2	0	0.8653	1.0000	0.6542	0.8678
						.	0.2244	0.4860
ADENOMA, SPINDLE-CELL	0	0 1	0	0	0.7366	.	.	0.4906
						.	.	0.4860
PHEOCHROMOCYTOMA (BENIGN)	0	0 0	1	0	0.4854	.	0.4857	.
						.	0.4811	.
PHEOCHROMOCYTOMA (MALIGNANT)	1	0 0	0	1	0.4218	0.7276	1.0000	1.0000
						0.4712	.	.
Pheochromocytoma B+M	1	0 0	1	1	0.3694	0.7276	0.7379	1.0000
						0.4712	0.4811	.
BONE (STIFLE)								
HEMANGIOSARCOMA	0	0 0	0	1	0.2390	0.4757	.	.
						0.4712	.	.

Table A.3.6. (cont.) Study DS95011: Tumor Incidences and Statistical Significance of Tests of Trend and Pairwise Differences in Male Standard Lab Mice

Organ / Tumor	Incidence No Veh	Trt Low	Med	Hi	Significance Levels Vehicle/No Treat versus Trend	High	Medium	Low
BONE MARROW (STIFLE)								
HEMANGIOMA	0	0 1	0	0	0.7366	.	.	0.4906
						.	.	0.4860
BRAIN								
ASTROCYTOMA	0	0 0	1	0	0.4829	.	0.4808	.
						.	0.4762	.
COAGULATING GLAND								
ADENOMA	0	0 1	0	0	0.7366	.	.	0.4906
						.	.	0.4860
INJECTION SITE								
FIBROSARCOMA	0	0 0	1	1	0.1734	0.4757	0.4808	.
						0.4712	0.4762	.
HEMANGIOSARCOMA	0	0 0	1	0	0.4854	.	0.4857	.
						.	0.4811	.
MALIGNANT FIBROUS HISTIOCYTOMA	1	0 0	1	0	0.7364	1.0000	0.7379	1.0000
						.	0.4811	.
Injection site/gross skin								
Fibrosarc./Sarc./Fib. Histio.	2	0 0	3	1	0.5064	0.8559	0.4640	1.0000
						0.4712	0.1079	.
KIDNEYS								
CARCINOMA, TUBULAR	0	0 1	0	1	0.2953	0.4757	.	0.4906
						0.4712	.	0.4860
LACRIMAL/HARDIRIAN GLANDS								
ADENOCARCINOMA	0	0 1	0	1	0.2953	0.4757	.	0.4906
						0.4712	.	0.4860
ADENOMA	6	11 3	8	28	0.0000	0.0000	0.3448	0.9055
						0.0002	0.7973	0.9940
Adenoma/Adenocarcinoma	6	11 4	8	29	0.0000	0.0000	0.3448	0.8162
						0.0001	0.7973	0.9827
LIVER								
ADENOMA, HEPATOCELLULAR	19	0 8	12	8	0.8928	0.9896	0.9281	0.9934
						0.0018	0.0001	0.0023
Adenoma/Carcinoma Hepato.	20	0 11	14	9	0.9183	0.9870	0.8832	0.9746
						0.0007	0.0000	0.0002
CARCINOMA, HEPATOCELLULAR	4	0 3	5	1	0.8938	0.9639	0.4759	0.7655
						0.4712	0.0244	0.1113
HEMANGIOMA	0	0 2	0	0	0.7973	.	.	0.2383
						.	.	0.2338
HEMANGIOSARCOMA	3	0 1	4	1	0.7260	0.9258	0.4685	0.9339
						0.4712	0.0525	0.4860
Hemangioma/Hemangiosarcoma	3	0 3	4	1	0.8377	0.9258	0.4685	0.6336
						0.4712	0.0525	0.1113
LUNG								
ADENOMA, BRONCHIOLO-ALVEOLAR	12	0 10	20	9	0.8196	0.7529	0.0554	0.7155
						0.0008	0.0000	0.0005
Adenoma/Carcinoma Bronch.-Alv.	15	0 15	24	13	0.7955	0.6445	0.0410	0.5145
						0.0000	0.0000	0.0000
CARCINOMA, BRONCHIOLO-ALVEOLAR	3	0 5	4	4	0.4150	0.4352	0.4576	0.3268
						0.0461	0.0503	0.0244
OSTEOSARCOMA (PRIMARY UNDETERMINED	0	0 0	0	1	0.2390	0.4757	.	.
						0.4712	.	.

Table A.3.6. (cont.) Study DS95011: Tumor Incidences and Statistical Significance of Tests of Trend and Pairwise Differences in Male Standard Lab Mice

Organ Tumor	Incidence No Veh	Trt	Low	Med	Hi	Significance Levels Vehicle/No Treat versus Trend	High	Medium	Low
LYMPHORETICULAR SYSTEM									
LEUKEMIA, MYELOGENOUS	1	0	0	0	0	1.0000	1.0000	1.0000	1.0000
							.	.	.
LYMPHOMA	4	1	6	3	4	0.5370	0.5648	0.7319	0.3245
							0.1426	0.2673	0.0457
Lymphoma/Leukemia, Myleogenous	5	1	6	3	4	0.6071	0.6845	0.8250	0.4475
							0.1426	0.2673	0.0457
SARCOMA, HISTIOCYTIC	0	0	3	1	1	0.5191	0.4757	0.4857	0.1146
							0.4712	0.4811	0.1113
PANCREAS									
ADENOMA, ISLET-CELL	0	0	1	0	0	0.7366	.	.	0.4906
							.	.	0.4860
PITUITARY									
ADENOMA, PARS DISTALIS	0	0	1	1	0	0.6060	.	0.4857	0.4906
							.	0.4811	0.4860
PREPUTIAL GLAND									
ADENOMA	0	0	0	0	1	0.2390	0.4757	.	.
							0.4712	.	.
PROSTATE									
ADENOMA	1	0	0	0	0	1.0000	1.0000	1.0000	1.0000
							.	.	.
SKIN (GROSS LESION)									
ADENOCARCINOMA, NOS	0	0	1	0	0	0.7366	.	.	0.4906
							.	.	0.4860
HEMANGIOMA	1	0	0	1	0	0.7339	1.0000	0.7328	1.0000
							.	0.4762	.
MALIGNANT FIBROUS HISTIOCYTOMA	0	0	0	1	0	0.4854	.	0.4857	.
							.	0.4811	.
SARCOMA, UNDIFFERENTIATED	1	0	0	0	0	1.0000	1.0000	1.0000	1.0000
							.	.	.
SKIN (ROUTINE SECTION)									
HEMANGIOMA	0	0	0	1	0	0.4829	.	0.4808	.
							.	0.4762	.
TUMOR, BASAL-CELL	0	0	0	0	1	0.2390	0.4757	.	.
							0.4712	.	.
SPLEEN									
HEMANGIOSARCOMA	1	0	3	5	1	0.7467	0.7276	0.0900	0.2947
							0.4712	0.0232	0.1113
STOMACH									
ADENOMA	2	0	0	0	0	1.0000	1.0000	1.0000	1.0000
							.	.	.
Adenoma/Sq. Cell Carcinoma	3	0	0	0	0	1.0000	1.0000	1.0000	1.0000
							.	.	.
CARCINOMA, SQUAMOUS-CELL	1	0	0	0	0	1.0000	1.0000	1.0000	1.0000
							.	.	.
PAPILLOMA	1	0	0	0	1	0.4218	0.7276	1.0000	1.0000
							0.4712	.	.
Skin(gross lesion/section									
HEMANGIOMA	1	0	0	2	0	0.6670	1.0000	0.4709	1.0000
							.	0.2244	.
Systemic									
HEMANGIOMA	1	0	4	3	0	0.9219	1.0000	0.2803	0.1699
							.	0.1046	0.0525
HEMANGIOSARCOMA	4	0	5	11	3	0.8398	0.7310	0.0393	0.4642
							0.1012	0.0002	0.0244
Hemangioma/Hemangiosarcoma	5	0	9	14	3	0.9627	0.8243	0.0186	0.1754
							0.1012	0.0000	0.0011

Table A.3.6. (cont.) Study DS95011: Tumor Incidences and Statistical Significance of Tests of Trend and Pairwise Differences in Male Standard Lab Mice

Organ / Tumor	Incidence (No Veh Trt Low Med Hi)	Trend	Significance Levels Vehicle/No Treat versus High	Medium	Low
TAIL					
HEMANGIOMA	0 0 0 1 0	0.4829	.	0.4808	.
			.	0.4762	.
HEMANGIOSARCOMA	0 0 1 0 0	0.7366	.	.	0.4906
			.	.	0.4860
Hemangioma/Hemangiosarcoma	0 0 1 1 0	0.6047	.	0.4808	0.4906
			.	0.4762	0.4860
KERATOACANTHOMA	1 0 0 0 0	1.0000	1.0000	1.0000	1.0000
			.	.	.
NEUROFIBROMA	1 0 0 0 0	1.0000	1.0000	1.0000	1.0000
			.	.	.
OSTEOMA	0 0 0 1 0	0.4854	.	0.4857	.
			.	0.4811	.
TESTES					
HEMANGIOMA	0 0 1 0 0	0.7366	.	.	0.4906
			.	.	0.4860
HEMANGIOSARCOMA	0 0 0 1 0	0.4829	.	0.4808	.
			.	0.4762	.
Hemangioma/Hemangiosarcoma	0 0 1 1 0	0.6047	.	0.4808	0.4906
			.	0.4762	0.4860
TUMOR, INTERSTITIAL-CELL	1 0 1 2 1	0.4696	0.7276	0.4709	0.7429
			0.4712	0.2244	0.4860
THYROID					
ADENOMA, FOLLICULAR	1 0 0 0 0	1.0000	1.0000	1.0000	1.0000
			.	.	.
URINARY BLADDER					
TUMOR, MESENCHYMAL	1 0 0 0 2	0.1427	0.4633	1.0000	1.0000
			0.2196	.	.

Table A.3.7. Study DS95011: Tumor Incidences and Statistical Significance of Tests of Trend and Pairwise Differences in Female Standard Lab Mice

Organ / Tumor	Incidence (No Veh Trt Low Med Hi)	Trend	Significance Levels Vehicle/No Treat versus High	Medium	Low
ADRENAL GLANDS					
ADENOMA, CORTICAL	1 0 0 0 0	1.0000	1.0000	1.0000	1.0000
			.	.	.
ADENOMA, SPINDLE-CELL	3 0 1 3 3	0.3354	0.6801	0.6801	0.9460
			0.1288	0.1288	0.5100
PHEOCHROMOCYTOMA (BENIGN)	0 0 1 0 0	0.7612	.	.	0.5152
			.	.	0.5100
BONE (STIFLE)					
CHONDROSARCOMA	0 0 0 1 0	0.5075	.	0.5152	.
			.	0.5100	.
BONE MARROW (STIFLE)					
HEMANGIOMA	0 0 1 0 0	0.7612	.	.	0.5152
			.	.	0.5100
BRAIN					
MENINGIOMA	0 0 0 1 0	0.5075	.	0.5152	.
			.	0.5100	.

Table A.3.7. (cont.) Study DS95011: Tumor Incidences and Statistical Significance of Tests of Trend and Pairwise Differences in Female Standard Lab Mice

Organ / Tumor	Incidence No Veh	Trt Low	Med	Hi	Trend	Significance Levels Vehicle/No Treat versus High	Medium	Low
CERVIX								
ADENOCARCINOMA	0	0	1	0	0.5075	.	0.5152	.
						.	0.5100	.
ADENOMA	0	0	1	0	0.7612	.	.	0.5152
						.	.	0.5100
Adenoma/-carinoma/Sq.Cell Carc.	0	0	2	1	0.7456	.	0.5152	0.2628
						.	0.5100	0.2576
CARCINOMA, SQUAMOUS-CELL	0	0	1	0	0.7612	.	.	0.5152
						.	.	0.5100
LEIOMYOMA	0	0	1	2	0.6509	.	0.2628	0.5152
						.	0.2576	0.5100
LEIOMYOSARCOMA	0	0	0	1	0.5099	.	0.5200	.
						.	0.5149	.
Leiomyoma/Leiomyosarcoma	0	0	1	3	0.7117	.	0.1367	0.5152
						.	0.1326	0.5100
SCHWANNOMA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
						.	.	.
INJECTION SITE								
FIBROSARCOMA	0	0	0	0	0.2537	0.5152	.	.
						0.5100	.	.
MASTOCYTOMA	0	0	0	1	0.5075	.	0.5152	.
						.	0.5100	.
Injection site/gross skin								
Fibrosarc./Mastocytoma/Rhabdo.	0	0	0	2	0.2078	0.5152	0.2628	.
						0.5100	0.2576	.
KIDNEYS								
ADENOMA, TUBULAR	0	0	1	0	0.7612	.	.	0.5152
						.	.	0.5100
FIBROMA	1	0	0	0	1.0000	1.0000	1.0000	1.0000
						.	.	.
LACRIMAL/HARDIRIAN GLANDS								
ADENOCARCINOMA	0	0	0	0	0.0653	0.2679	.	.
						0.2626	.	.
ADENOMA	1	4	2	2	0.0001	0.0030	0.5230	0.5230
						0.0586	0.9067	0.9067
Adenoma/Adenocarcinoma	1	4	2	2	0.0000	0.0008	0.5230	0.5230
						0.0240	0.9067	0.9067
LIVER								
ADENOMA, HEPATOCELLULAR	1	0	2	5	0.7268	0.7675	0.1167	0.5230
						0.5100	0.0312	0.2576
Adenoma/Carcinoma Hepato.	1	0	2	6	0.7760	0.7675	0.0658	0.5230
						0.5100	0.0151	0.2576
CARCINOMA, HEPATOCELLULAR	0	0	0	1	0.5075	.	0.5152	.
						.	0.5100	.
HEMANGIOSARCOMA	4	0	3	1	0.7506	0.9067	0.9747	0.7982
						0.2576	0.5100	0.1288
LUNG								
ADENOMA, BRONCHIOLO-ALVEOLAR	2	9	9	15	0.1847	0.0100	0.0008	0.0300
						0.4403	0.1580	0.6202
Adenoma/Carcinoma Bronch.-Alv.	5	14	10	17	0.0071	0.0013	0.0057	0.1501
						0.2197	0.3821	0.8889
CARCINOMA, BRONCHIOLO-ALVEOLAR	3	5	1	4	0.0058	0.0858	0.5226	0.9460
						0.2534	0.7762	0.9883
LYMPH NODE,MANDIBULAR								
HEMANGIOSARCOMA	0	0	1	0	0.7612	.	.	0.5152
						.	.	0.5100

Table A.3.7. (cont.) Study DS95011: Tumor Incidences and Statistical Significance of Tests of Trend and Pairwise Differences in Female Standard Lab Mice

Organ / Tumor	Incidence No Veh	Trt Low	Med	Hi	Significance Levels Vehicle/No Trt versus Trend	High	Medium	Low
LYMPHORETICULAR SYSTEM								
LYMPHOMA	17	10	12	22	21	0.1831	0.3744	0.3045	0.9055
							0.0437	0.0290	0.4673
SARCOMA, HISTIOCYTIC	5	3	6	4	3	0.8346	0.8845	0.7949	0.5569
							0.6705	0.5221	0.2638
MAMMARY GLAND/REGION									
ADENOACANTHOMA	0	0	1	1	0	0.6369	.	0.5152	0.5152
							.	0.5100	0.5100
ADENOCARCINOMA	1	0	3	2	4	0.1473	0.2003	0.5152	0.3240
							0.0663	0.2576	0.1288
ADENOMA	0	0	0	1	0	0.5075	.	0.5152	.
							.	0.5100	.
Adenoma/Adenocarc./Fibroadenoma	1	0	3	3	5	0.0826	0.1165	0.3315	0.3240
							0.0328	0.1326	0.1288
FIBROADENOMA	0	0	0	0	1	0.2537	0.5152	.	.
							0.5100	.	.
MESENTERY									
HEMANGIOMA	0	0	1	0	0	0.7612	.	.	0.5152
							.	.	0.5100
OVARIES									
ADENOMA, PAPILLARY	0	0	0	1	0	0.5075	.	0.5152	.
							.	0.5100	.
ADENOMA, TUBULAR	0	0	0	1	0	0.5075	.	0.5152	.
							.	0.5100	.
ADENOMA, TUBULO-STROMAL	1	0	0	0	0	1.0000	1.0000	1.0000	1.0000
							.	.	.
Adenoma,Pap./Cystadenoma	0	0	0	1	1	0.1928	0.5152	0.5152	.
							0.5100	0.5100	.
Adenoma,Tub. or Tub.-Stromal	1	0	0	1	0	0.7562	1.0000	0.7624	1.0000
							.	0.5100	.
CYSTADENOMA	0	0	0	0	1	0.2537	0.5152	.	.
							0.5100	.	.
HEMANGIOMA	1	0	0	2	0	0.6977	1.0000	0.5303	1.0000
							.	0.2626	.
HEMANGIOSARCOMA	0	0	1	1	0	0.6369	.	0.5152	0.5152
							.	0.5100	0.5100
Hemangioma/Hemangiosarcoma	1	0	1	3	0	0.8124	1.0000	0.3396	0.7675
							.	0.1326	0.5100
LEIOMYOMA	0	0	1	0	0	0.7612	.	.	0.5152
							.	.	0.5100
LUTEOMA	0	0	2	2	2	0.2315	0.2628	0.2628	0.2628
							0.2576	0.2576	0.2576
Leiomyoma/Lute-/Thec-/G-cell Tmr	2	0	5	3	2	0.7473	0.7140	0.5290	0.2438
							0.2576	0.1288	0.0312
THECOMA	1	0	0	0	0	1.0000	1.0000	1.0000	1.0000
							.	.	.
TUMOR, GRANULAR-CELL	0	0	1	0	0	0.7612	.	.	0.5152
							.	.	0.5100
TUMOR, GRANULOSA-CELL	1	0	1	1	0	0.8392	1.0000	0.7675	0.7675
							.	0.5100	0.5100
PANCREAS									
ADENOMA, ISLET-CELL	1	0	1	0	0	0.9439	1.0000	1.0000	0.7675
							.	.	0.5100
PITUITARY									
ADENOMA, PARS DISTALIS	2	0	4	4	2	0.6978	0.7140	0.3683	0.3683
							0.2576	0.0637	0.0637

Table A.3.7. (cont.) Study DS95011: Tumor Incidences and Statistical Significance of Tests of Trend and Pairwise Differences in Female Standard Lab Mice

Organ Tumor	Incidence No Veh	Trt	Low	Med	Hi	Trend	Significance Levels Vehicle/No Treat versus High	Medium	Low
SKIN (GROSS LESION)									
CARCINOMA, BASAL-CELL	1	0	0	0	0	1.0000	1.0000	1.0000	1.0000
							.	.	.
PAPILLOMA (INVERTED)	1	0	0	0	0	1.0000	1.0000	1.0000	1.0000
							.	.	.
RHABDOMYOSARCOMA	0	0	0	1	0	0.5075	.	0.5152	.
							.	0.5100	.
SPLEEN									
HEMANGIOMA	0	0	0	0	2	0.0634	0.2628	.	.
							0.2576	.	.
HEMANGIOSARCOMA	2	0	1	0	1	0.6770	0.8861	1.0000	0.8861
							0.5100	.	0.5100
Hemangioma/Hemangiosarcoma	2	0	1	0	3	0.1660	0.5191	1.0000	0.8861
							0.1288	.	0.5100
STOMACH									
CARCINOMA, SQUAMOUS-CELL	0	0	0	0	1	0.2537	0.5152	.	.
							0.5100	.	.
PAPILLOMA (SQUAMOUS)	0	0	1	0	0	0.7612	.	.	0.5152
							.	.	0.5100
Systemic									
HEMANGIOMA	2	0	2	2	4	0.1592	0.3582	0.7135	0.7063
							0.0637	0.2626	0.2576
HEMANGIOSARCOMA	6	0	7	3	4	0.7739	0.8563	0.9289	0.5454
							0.0663	0.1326	0.0083
Hemangioma/Hemangiosarcoma	8	0	9	5	8	0.4744	0.6397	0.8972	0.5528
							0.0037	0.0328	0.0019
TAIL									
HEMANGIOSARCOMA	0	0	0	1	0	0.5075	.	0.5152	.
							.	0.5100	.
THYROID									
ADENOMA, FOLLICULAR	0	0	0	2	0	0.5075	.	0.2628	.
							.	0.2576	.
URINARY BLADDER									
TUMOR, MESENCHYMAL	0	0	0	1	0	0.5075	.	0.5152	.
							.	0.5100	.
UTERUS									
ADENOCARCINOMA, ENDOMETRIAL	0	0	2	0	4	0.0267	0.0664	.	0.2628
							0.0637	.	0.2576
DECIDUOMA	1	0	1	0	0	0.9439	1.0000	1.0000	0.7675
							.	.	0.5100
ENDOMETRIAL STROMAL POLYP	3	0	5	7	3	0.7305	0.6705	0.1673	0.3688
							0.1288	0.0072	0.0312
HEMANGIOMA	1	0	0	1	2	0.1733	0.5230	0.7675	1.0000
							0.2576	0.5100	.
HEMANGIOSARCOMA	0	0	1	0	1	0.3222	0.5152	.	0.5152
							0.5100	.	0.5100
Hemangioma/Hemangiosarcoma	1	0	1	1	3	0.1243	0.3396	0.7675	0.7675
							0.1326	0.5100	0.5100
LEIOMYOMA	2	0	5	2	1	0.8949	0.8897	0.7140	0.2525
							0.5100	0.2576	0.0328
LEIOMYOSARCOMA	0	0	0	2	0	0.5075	.	0.2679	.
							.	0.2626	.
Leiomyoma/-sarcoma/stromal sarc.	2	0	6	5	1	0.9294	0.8897	0.2525	0.1691
							0.5100	0.0328	0.0170
SARCOMA, STROMAL	0	0	1	2	0	0.6509	.	0.2628	0.5152
							.	0.2576	0.5100
TUMOR, GRANULAR-CELL	0	0	0	0	1	0.2537	0.5152	.	.
							0.5100	.	.

Table A.3.7. (cont.) Study DS95011: Tumor Incidences and Statistical Significance of Tests of Trend and Pairwise Differences in Female Standard Lab Mice

Organ / Tumor	Incidence No Veh	Trt Low	Med	Hi	Significance Levels Vehicle/No Treat versus Trend	High	Medium	Low
VAGINA								
POLYP(S)	1	0 0	0	0	1.0000	1.0000	1.0000	1.0000
						.	.	.
SARCOMA, NOS	1	0 0	0	0	1.0000	1.0000	1.0000	1.0000
						.	.	.
Sarcoma NOS/Polyps	2	0 0	0	0	1.0000	1.0000	1.0000	1.0000
						.	.	.

Appendix 4. References

Bailer, A. and Portier, C. (1988), "Effects of Treatment-Induced Mortality on Tests for Carcinogenicity in Small Samples", *Biometrics*, **44**, 4, 417-431.

Bieler, G.S., and Williams, R.L. (1993), "Ratio Estimates, the Delta Method, and Quantal Response Tests for Increased Carcinogenicity", *Biometrics*, **49**, 4, 793-801.

Chu, K.C., Ceuto, C., and Ward, J.M. (1981), "Factors in the Evaluation of 200 National Cancer Institute Carcinogen Bioassays", *Journal of Toxicology and Environmental Health,* 8, 251-280.

De Iorio, M., Muller, P., Rosner, R.L., and MacEachern, S. (2004), "An ANOVA Model for Dependent Random Measures", Bayesian Nonparametric Nonproportional Hazards Survival Modeling", *Journal of the American Statistical Association*. **99**, 465, 205-215.

De Iorio, M., Johnson, W.O., Mueller, P., and Rosner, R.L. (2009), "Bayesian Nonparametric Nonproportional Hazards Survival Modeling", *Biometrics*, 65, 3, 762-771.

Haseman, J. K. (1983), "A Reexamination of False-positive Rates for Carcinogenicity Studies", *Fundamental and Applied Toxicology,* 3, 334-339.

Jara, A. (2007), "Applied Bayesian Non- and Semi-parametric Inference using DPpackage", *Rnews***, 7**, 3, 17-26.

Jara, A., Hanson, T.E., Quintana, F.A., Mueller, P., and Rosner, G.L. (To be published), "DPpackage: Bayesian Non- and Semi-parametric Modelling in R," *Journal of Statistical Software.*

Lin, K. K. and Ali, M.W. (2006), "Statistical Review and Evaluation of Animal Tumorigenicity Studies", *Statistics in the Pharmaceutical Industry,Third Edition,* edited by C.R. Buncher and J.Y. Tsay, Marcel Dekker, Inc. New York.

Lin, K. K. and Rahman, M.A. (1998), "Overall False Positive Rates in Tests for Linear Trend in Tumor Incidence in Animal Carcinogenicity Studies of New Drugs", *Journal of Biopharmaceutical Statistics*. **8,** 1, 1-15.

McConnell, E.E., Solleveld, H.A., Swenberg, J.A., and Boorman, G.A. (1986), "Guidelines for Combining Neoplasms for Evaluation of Rodent Carcinogenesis Studies", *Journal of the National Cancer Institute*. **76**, 283-289.

Peto, R., Pike, M.C., Day, N.E., Gray, R.G., Lee, P.N., Parrish, S., Peto, J., Richards, S., and Wahrendorf, J. (1980). "Guidelines for sample sensitive significance tests for carcinogenic effects in long-term animal experiments", *IARC Monographs on the Evaluation of Carcinogenic Risk of Chemicals to Humans, supplement 2: Long term and Short term Screening Assays for Carcinogens: A Critical Appraisal*, International Agency for Research Against Cancer, 311-426.

R Development Core Team (2009). *R: A language and environment for statistical computing. R Foundation for Statistical Computing*, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Rahman, M.A. and Lin, K.K. (2008), "A Comparison of False Positive Rates of Peto and Poly-3 Methods for Long Term Carcinogenicity Data Analysis Using Multiple Comparison Adjustment Method Suggested by Lin and Rahman", *Journal of Biopharmaceutical Statistics*. 18, 949-958.

STP Peto Working Group (2002), "Statistical Methods for Carcinogenicity Studies", *Toxicologic Pathology*. 30 (3), 403-414.

U.S. Department of Health and Human Services (2001), *Guidance for Industry Statistical Aspects of the Design, Analysis, and Interpretation of Chronic Rodent Carcinogenicity Studies of Pharmaceuticals* (DRAFT GUIDANCE), Center for Drug Evaluation and Research, Food and Drug Administration.

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
01/14/2011

KARL K LIN
01/18/2011
Concur with review

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:	22561/000-
Drug Name:	Cladribine Tablets
Indication(s):	Multiple Sclerosis
Applicant:	EMD Serono, Inc.
Date of Submission:	May 27, 2010
PDUFA Due Date:	November 27, 2010
Review Priority:	Priority Review
Biometrics Division:	Division I
Statistical Reviewer:	Sharon Yan
Concurring Reviewers:	Kun Jin, Team Leader
	James Hung, Director
Medical Division:	Neurology
Clinical Team:	Jody Green, Clinical Reviewer
	Evelyn Mentari, Safety Reviewer
	Billy Dunn, Clinical Team Leader
	Eric Bastings, Deputy Director
	Russell Katz, Director
Project Manager:	Hamet Toure

Table of Contents

1. EXECUTIVE SUMMARY

1.1 Conclusions and Recommendations

The efficacy data from the CLARITY study (protocol 25643) support the superiority of Cladribine over placebo in the reduction of relapse rate, active T1 Gd-enhanced lesions and active T2 lesions. No dose response was observed in the clinical endpoints, and Cladribine 5.25 mg/kg does not appear to provide additional benefits over cladribine 3.5 mg/kg in the clinical endpoint of relapse rate or disability progression.

1.2 Brief Overview of Clinical Studies

Cladribine is proposed to be indicated for the treatment of relapsing forms of multiple sclerosis (b) (6). Demonstration of efficacy of cladribine is based on the pivotal study CLARITY (protocol 25643).

The CLARITY trial is a Phase III randomized, double-blind, three-arm, placebo-controlled, multi-center clinical trial, evaluating the efficacy and safety of cladribine tablets over 96 weeks in subjects with relapsing-remitting MS (RRMS). The trial randomized 1326 subjects at 155 centers in 32 countries. Subjects were randomized 1:1:1 to receive cladribine tablets 3.5 mg/kg cumulative dose, cladribine tablets 5.25 mg/kg cumulative dose, and placebo.

1.3 Statistical Issues and Findings

Oral cladribine treatment, both the high dose and the low dose, were found to be statistically superior to placebo in the primary endpoint of relapse rate. The treatment difference yielded a p-value of less than 0.0001 in the comparison between each of the cladribine dose groups and the placebo group from the primary analysis.

Statistically significant treatment differences were also found in all key secondary MRI endpoints as well as in disability progression.

No major statistical issues were found.

2. INTRODUCTION

2.1 Overview

EMD Serono submitted two studies in supporting the efficacy claim of cladribine: pivotal Study CLARITY and supportive study Scripps-C.

The Scripps-C trial was a single-center, placebo-controlled, double-blind trial that equally randomized 49 subjects to receive s.c. parenteral cladribine 2.1 mg/kg cumulative dose (equivalent to cladribine tablets 5.25 mg/kg cumulative dose) or placebo.

At the meeting with EMD Serono held on 26 January 2010, the FDA requested source data verification of Scripps-C via a third-party audit and Scripps-C CRFs. After reviewing the audit report, the FDA concluded that Study Scripps-C was not acceptable to be used as supportive study for efficacy. Therefore, the establishment of efficacy for oral Cladribine will be based on CLARITY study only. Consistent and robust efficacy results with high level of significance in treatment difference are expected for demonstration of efficacy with a single study.

2.2 Data Sources

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

\\Cdsesub1\evsprod\NDA022561

3. STATISTICAL EVALUATION

3.1 Evaluation of Efficacy

3.1.1 Description of the Study

The primary objective of the trial was to evaluate the efficacy of cladribine versus placebo in the reduction of qualifying relapse rate during 96 weeks of treatment in subjects with RRMS.

The study was a randomized, double-blind, three-arm parallel, placebo-controlled, multi-center trial. The trial included a pre-trial evaluation period (up to 28 days prior to the start of treatment), an initial treatment period during Week 0-48, and a retreatment period during Week 48-96. The overall trial design is displayed in Figure 1.

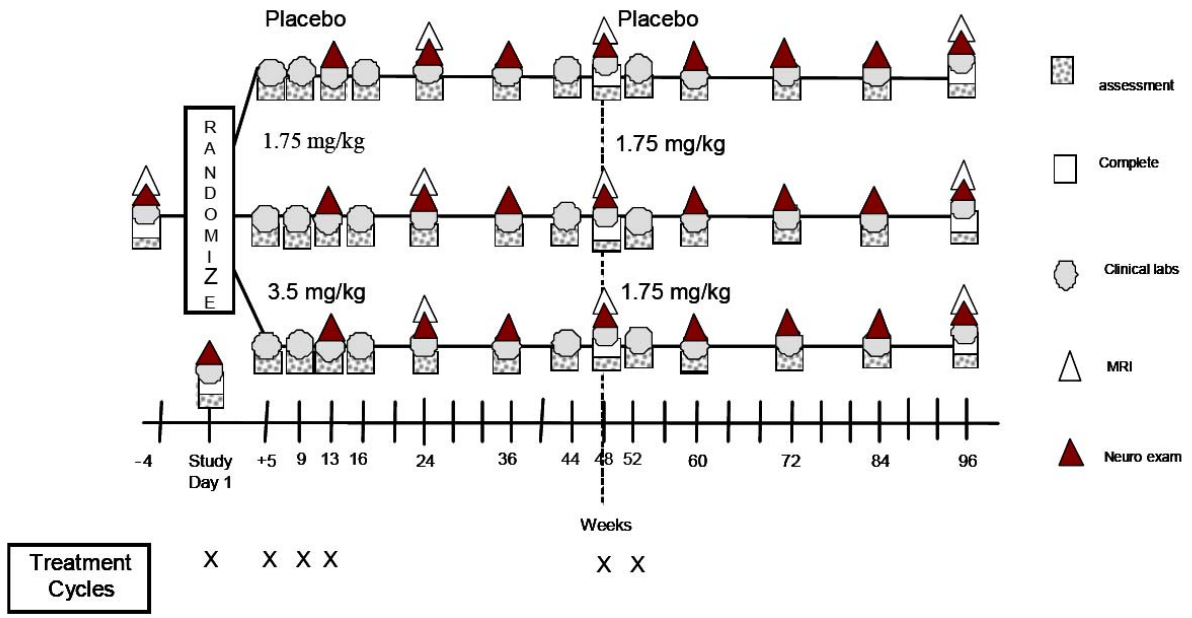


Figure 1 Trial Design (Source: Figure 25643-1 from sponsor's study report)

The study enrolled subjects with relapsing remitting MS who had one or more relapses within twelve months prior to Trial Day 1.

Eligible subjects were equally randomized by a central randomization system to receive:

- Cladribine tablets 3.5 mg/kg (administered p.o. as 0.875 mg/kg/course for two courses plus placebo p.o. for two courses during the first 48 weeks and 0.875 mg/kg/course for two courses during the second 48 weeks), or
- Cladribine tablets 5.25 mg/kg (administered p.o. as 0.875 mg/kg/course for four courses during the first 48 weeks and 0.875 mg/kg/course for two courses during the second 48 weeks), or
- matching placebo (administered p.o. for four courses during the first 48 weeks and two courses during the second 48 weeks).

Cladribine was administered orally in 10 mg tablets. The number of tablets administered was standardized based on weight. A course was defined as daily administration given consecutively over four to five days during a 28-day period. The courses administered for all treatment groups were initiated at Trial Day 1, Week 5, Week 9 and Week 13 (for the first 48-week period), and at Week 48 and Week 52 (for second 48-week period).

In addition to the usual safety assessments, all subjects were assessed at the Week 44 visit for their lymphocyte count prior to retreatment at Week 48. For all randomized subjects, a rescue option of treatment with Rebif® (44 mcg three times a week [tiw]) was available if the subject experienced more than one qualifying relapse, and/or experienced a sustained increase in their EDSS of $\geq$ 1 point, or $\geq$ 1.5 points if baseline EDSS was 0 (over a period of three months or greater), during a calendar year beginning at Week 24.

The blind was maintained by utilizing a Treating Physician who viewed clinical laboratory results and assessed adverse events (AEs) and safety information, and an independent blinded Evaluating Physician who performed neurological exams.

It was planned that 1290 subjects (430 subjects in each group) would provide 90% power to detect a 25% relative reduction in the primary efficacy endpoint of qualifying relapse rate during 96 weeks, when comparing each of the two cladribine dose groups to the placebo group.

3.1.2 Efficacy Variables

Primary Efficacy Endpoint

The primary endpoint of the study was the qualifying relapse rate at 96 weeks.

Qualifying relapses were defined as a two grade increase in one or more Kurtzke Functional Systems (KFS) or a one grade increase in two or more KFS, excluding changes in bowel/bladder or cognition, in the absence of fever, lasting for ≥ 24 hours, and preceded by ≥ 30 days of clinical stability or improvement.

Secondary Efficacy Endpoints

Secondary efficacy endpoints of the study included the following:

- Proportion of subjects qualifying relapse-free at 96 weeks
- Disability progression at 96 weeks (time to sustained change in Expanded Disability Status Score (EDSS) ≥ one point, or ≥ 1.5 points if baseline EDSS was 0, over a period of at least three months)
- Mean number of active T1 gadolinium-enhanced lesions per subject per scan at 96 Weeks
- Mean number of active T2 lesions per subject per scan at 96 weeks
- Mean number of CU lesions defined as 1) new T1 gadolinium-enhancing, or 2) new T2 nonenhancing or enlarging lesions, or 3) both, without double-counting (designated "combined unique lesions") per subject per scan at 96 weeks

Neurological examinations including the KFS exam, ambulation up to 500 meters and EDSS were obtained at the Pre-Study Evaluation, Study Day 1 and at Weeks 13, 24, 36, 48, 60, 72, 84 and 96. MRI scans were assessed at the pre-trial evaluation and at Weeks 24, 48 and 96.

3.1.3 Statistical Analysis Methods

The statistical analysis methods described in this section were prospectively specified in the Statistical Analysis Plan (SAP).

The intend-to-treat (ITT) population was the primary analysis populations for efficacy analysis. The Evaluable population was utilized as the supportive analysis population. The ITT population

included all subjects who were randomized into the trial. Subjects who completed treatment without a major protocol deviation with 96-week data were included in the Evaluable population.

If at least 10% of the subjects took Rebif® (or any disease modifying drug) as rescue medication, then the models used for the continuous efficacy parameters were to be augmented to include an indicator for intake of rescue medication and interaction between treatment and the indicator for intake of rescue medication.

Since less than 10% (actual 3.5%; 47 out of 1326) of the subjects took Rebif® (or any disease modifying drug) as rescue medication, the models used for the continuous efficacy parameters were not augmented to include an indicator for intake of rescue medication.

Analysis of Primary Efficacy Endpoint

The qualifying relapse rate was analyzed using a Poisson regression model with fixed effects for treatment group and region with log of time on study as an offset variable in the model. An approximate Chi-square test based on Wald statistics was used to compare treatment groups. In addition, the relative risk of developing a qualifying relapse and its associated 95% CI (97.5% CI) were estimated for each treatment group comparison. Annualized qualifying relapse rate and its associated 95% CI (97.5% CI) were estimated for each treatment group.

Hochberg′s step-up procedure was employed in order to compare the cladribine dose groups with the placebo group. The hypothesis testing was to proceed in the following manner. First, the cladribine dose group with the largest p-value when comparing its qualifying relapse rate with that of the placebo group was examined. If the p-value was ≤0.05, then both the low and high dose cladribine groups were considered to be significantly different from the placebo group in the qualifying relapse rate. If the p-value exceeded 0.05, then the other cladribine dose group was considered significantly different from the placebo group if the corresponding comparison p-value was ≤0.025. Otherwise, both the low and high dose cladribine groups were considered not significantly different from the placebo group in the qualifying relapse rate. If both cladribine dose groups were significantly different from the placebo group, then the appropriate 95% CI for inference were presented. If only one cladribine dose group was significant, then the appropriate 97.5% CI for inference was presented.

Analyses of Secondary Efficacy Endpoints

The three MRI parameters were analyzed using a non-parametric ANCOVA (analysis of covariance) model on ranked data with fixed effects for treatment group and region with adjustment for baseline T1 Gd+ lesion number, as there was no available data for baseline T2 or CU lesions.

If both cladribine doses were significant for the primary parameter, then the three MRI parameters were to be tested in the order of T_1, T_2, and CU for high dose cladribine versus placebo followed by T_1, T_2, and CU for low dose cladribine versus placebo in a hierarchical manner at the 0.05 level. If only one cladribine dose was significant for the primary parameter, then these MRI parameters were to be tested in the order of T_1, T_2, and CU for the significant

dose of cladribine versus placebo at the 0.025 level. However, the sequential testing of the parameters was only to be carried out if the test for the previous parameter was significant.

The proportion of qualifying relapse-free subjects at the end of 96 weeks was analyzed using a logistic regression model with fixed effects for treatment group and region. The odds ratio of being qualifying relapse-free in each of the cladribine groups versus the placebo group and the associated 95% (97.5%) CI was estimated.

Time to 3-month sustained change in EDSS score was analyzed using a Cox proportional hazards model with fixed effects for treatment group and region. An approximate Chi square test based on Wald statistic was used to compare treatment groups. The hazard ratio of time to 3-month sustained change in EDSS score in each of the cladribine groups versus the placebo group and the associated 95% (97.5%) CI was estimated. Kaplan-Meier plots of time to 3-month sustained change in EDSS score (survival function) were presented by treatment group. In addition, the Kaplan-Meier estimates of time to 3-month sustained change in EDSS score was presented for each treatment group.

For the analysis of continuous parameters, confirmation of the ANOVA model assumptions (normality) was conducted. Normality was assessed using the normal probability plot and the Shapiro-Wilk statistic. If the model assumptions were satisfied, then the ANOVA model with fixed effects for treatment and region was performed on the raw data; otherwise, the ANOVA model with fixed effects for treatment and region was performed on the ranked data.

3.1.4 Study Results

3.1.4.1 Disposition of Subjects

A total of 1641 subjects were screened for enrollment, and of these subjects, 1326 were randomized into the trial from 155 investigative sites across 32 countries worldwide. Of these 1326 subjects, in accordance with the 1:1:1 randomization scheme, 456 were randomized to the cladribine 5.25 mg/kg group, 433 were randomized to the cladribine 3.5 mg/kg group and 437 were randomized to the placebo group. The slight discrepancies between groups were attributable to the block size of 6 that was used in the randomization scheme, as the randomization was stratified by site.

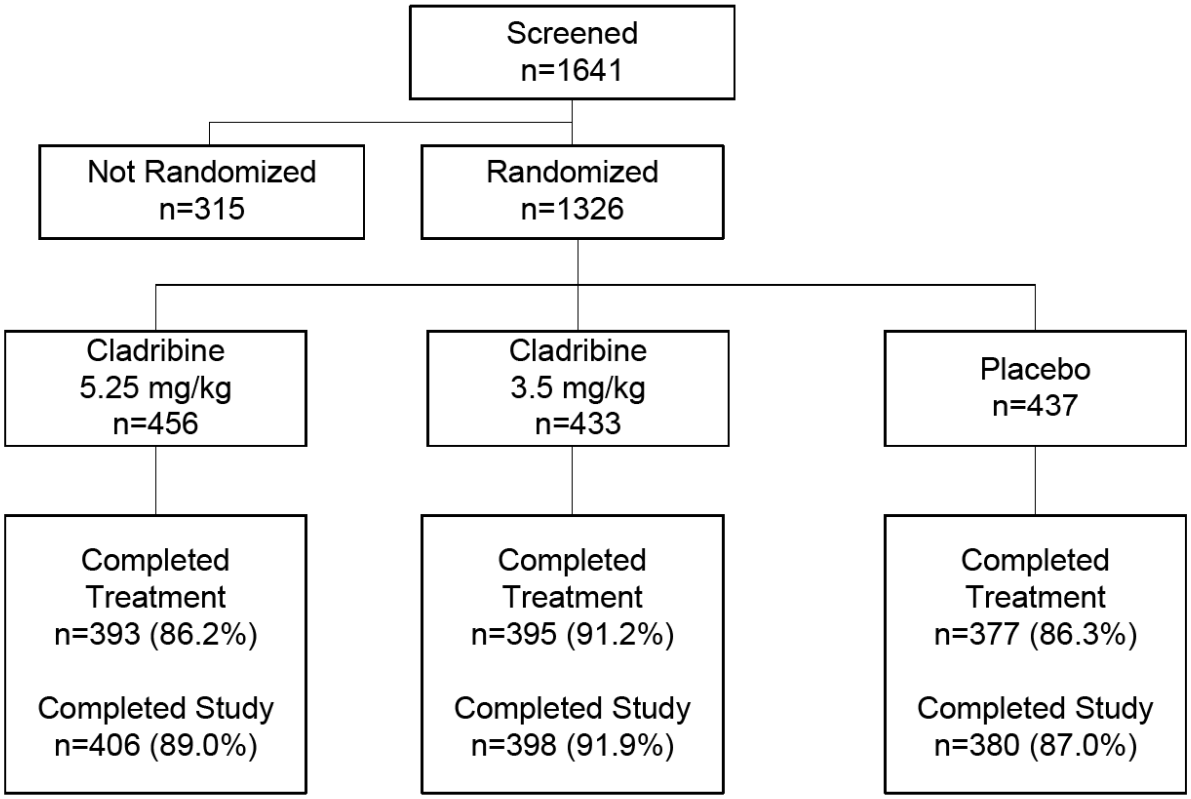


Figure 2 Disposition of Subjects (Source: Figure 25643-3 of sponsor's study report)

Of the 32 countries, the majority of subjects were enrolled from Russia (n=300), Bulgaria (n=190), the United States (n=94), Ukraine (n=80), Czech Republic (n=67), Italy (n=66), Tunisia (n=54), and France (n=47).

Table 1 displays the reasons for premature trial discontinuation. The percentages of subjects who completed the trial were similar across the cladribine 5.25 mg/kg, cladribine 3.5 mg/kg, and placebo groups (89.0%, 91.9% and 87.0%, respectively). Of the subjects who withdrew from the trial, one of the most common reason cited was the category "adverse event" in which the cladribine 5.25 mg/kg group displayed a higher rate (2.0%) compared to the cladribine 3.5 mg/kg (1.2%) and placebo (1.1%) groups. The largest proportion of subjects that withdrew from the trial due to disease progression was in the placebo group (4.8%) compared to cladribine 5.25 mg/kg (0.9%) and cladribine 3.5 mg/kg (1.2%). Four subjects died during the trial (cladribine 5.25 mg/kg, n=1; cladribine 3.5 mg/kg, n=1, and placebo, n=2).

Table 1 Study Termination by Treatment Group – ITT Population (Source: Table 25643-3 of sponsor's study report)

	Status	Cladribine 5.25 mg/kg (n=456) n (%)	Cladribine 3.5 mg/kg (n=433) n (%)	Placebo (n=437) n (%)	Total (n=1326) n (%)
Completed Study	Yes	406 (89.0)	398 (91.9)	380 (87.0)	1184 (89.3)
	No	50 (11.0)	35 (8.1)	57 (13.0)	142 (10.7)
Reasons for Withdrawing from Study Prematurely	Adverse event	9 (2.0)	5 (1.2)	5 (1.1)	19 (1.4)
	Lost to follow-up	11 (2.4)	8 (1.8)	4 (0.9)	23 (1.7)
	Protocol violation	4 (0.9)	4 (0.9)	10 (2.3)	18 (1.4)
	Death	1 (0.2)	1 (0.2)	2 (0.5)	4 (0.3)
	Disease progression	4 (0.9)	5 (1.2)	21 (4.8)	30 (2.3)
	Other	21 (4.6)	12 (2.8)	15 (3.4)	48 (3.6)

T:\BIOMET08\CLADRIBINE\25643\FINAL\TABLES\PGM\T_SD2B.SAS 10FEB2009

3.1.4.2 Demographic Characteristics

Demographic characteristics were well balanced among the treatment groups. No statistically significant differences were observed among any of the demographic characteristics. The mean age (years) was 39.1 for the cladribine 5.25 mg/kg group, 37.9 for the cladribine 3.5 mg/kg group and 38.7 for the placebo group. Most subjects were white (98%), and approximately 68% of subjects were female.

The baseline MRI and neurological assessments are presented in the following table. The mean (s.d.) number of T1 hypointense lesions was greater for the cladribine 5.25 mg/kg group compared to cladribine 3.5 mg/kg and placebo (8.5 [9.3], 7.1 [8.2] and 7.4 [8.0], respectively), and the mean T2 lesion volume (mm^3) was greater in the cladribine 5.25 mg/kg group compared to cladribine 3.5 mg/kg and placebo (17202.1 [17467.7], 14828.0 [16266.8], and 14287.6 [13104.8], respectively). These findings suggest that the cladribine 5.25 mg/kg group may at baseline have had an increased disease burden based on these MRI parameters.

Table 2 Baseline MRI and Neurological Assessment by Treatment Group - ITT Population (Source: Table 25643-9 of sponsor's study report)

Characteristic	Statistics	Cladribine 5.25 mg/kg (n=456)	Cladribine 3.5 mg/kg (n=433)	Placebo (n=437)	p-value
EDSS category, n (%)	n (missing)	456 (0)	433 (0)	437 (0)	0.149[a]
	0	11 (2.4)	12 (2.8)	13 (3.0)	
	1	80 (17.5)	75 (17.3)	70 (16.0)	
	2	119 (26.1)	133 (30.7)	127 (29.1)	
	3	108 (23.7)	108 (24.9)	96 (22.0)	
	4	84 (18.4)	71 (16.4)	83 (19.0)	
	>=5	54 (11.8)	34 (7.9)	48 (11.0)	
EDSS	Mean (SD)	3.0 (1.4)	2.8 (1.2)	2.9 (1.3)	
	Median	3.0	2.5	3.0	
	Min; Max	0.0; 5.5	0.0; 6.0	0.0; 5.5	
Number of T1 Gadolinium-enhanced Lesions	n (missing)	456 (0)	433 (0)	437 (0)	0.547[b]
	Mean (SD)	1.0 (2.3)	1.0 (2.7)	0.8 (2.1)	
	Median	0.0	0.0	0.0	
	Min; Max	0.0; 20.0	0.0; 32.0	0.0; 27.0	
Number of T1 Hypointense Lesions	n (missing)	456 (0)	433 (0)	437 (0)	0.058[b]
	Mean (SD)	8.5 (9.3)	7.1 (8.2)	7.4 (8.0)	
	Median	5.0	4.0	5.0	
	Min; Max	0.0; 57.0	0.0; 48.0	0.0; 44.0	
T2 Lesion Volume (mm^3)	n (missing)	456 (0)	433 (0)	437 (0)	0.058[b]
	Mean (SD)	17202.1 (17467.7)	14828.0 (16266.8)	14287.6 (13104.8)	
	Median	11106.0	9659.0	10140.5	
	Min; Max	236.0; 103645.0	106.0; 128747.0	150.0; 76770.0	

T:\BIOMET08\CLADRIBINE\25643\FINAL\TABLES\PGM\T_BCA.SAS 13JAN2009

[a] From Cochran-Mantel-Haenszel row means score test, adjusted for region.

[b] From a two-way ANOVA model on ranked data with fixed effects for treatment group and region.

Table 3 summarizes the MS history for the cladribine and placebo treatment groups within the ITT population. MS history characteristics, in general, were well balanced across the treatment groups. Notably, there was a shorter median time from first attack prior to Trial Day 1 in the cladribine 3.5 mg/kg treatment group compared to the other two groups. The majority of subjects experienced one or two relapses within the 12 months prior to SD1 (cladribine 5.25 mg/kg, 95.6%; cladribine 3.5 mg/kg, 94.2%; placebo, 95.2%). Only one or two subjects in each group received treatment in the 3 months prior to SD1.

Table 3 MS History by Treatment Group - ITT Population (Source: Table 25643-10 of sponsor's study report)

Multiple Sclerosis Characteristic	Statistics	Cladribine 5.25 mg/kg (n=456)	Cladribine 3.5 mg/kg (n=433)	Placebo (n=437)	p-value
Time since first attack (years) prior to Study Day 1	n (missing)	456 (0)	433 (0)	437 (0)	0.005[a]
	Mean (SD)	9.3(7.6)	7.9(7.2)	8.9(7.4)	
	Median	7.2	5.8	7.1	
	Min; Max	0.4; 35.2	0.3; 42.3	0.4; 39.5	
Time since most recent relapse (months) prior to Study Day 1	n (missing)	456 (0)	433 (0)	437 (0)	0.352[a]
	Mean (SD)	5.3(3.0)	5.4(2.9)	5.4(2.7)	
	Median	4.3	4.8	5.0	
	Min; Max	0.9; 13.3	1.1; 15.2	0.9; 12.8	
Number of relapses within the past 12 months prior to Study Day 1, n (%)	n (missing)	456 (0)	433 (0)	437 (0)	0.667[b]
	0	2 (0.4)	0	0	
	1	323 (70.8)	303 (70.0)	306 (70.0)	
	2	113 (24.8)	105 (24.2)	110 (25.2)	
	3	14 (3.1)	22 (5.1)	19 (4.3)	
	>=4	4 (0.9)	3 (0.7)	2 (0.5)	
Subjects who received treatment during the last 3 months prior to Study Day 1	n (missing)	456 (0)	433 (0)	437 (0)	0.836[c]
	Yes	2 (0.4)	1 (0.2)	1 (0.2)	
	No	454 (99.6)	432 (99.8)	436 (99.8)	
Subjects with abnormalities related to MS on neurological examination	n (missing)	456 (0)	433 (0)	437 (0)	0.834[c]
	Yes	442 (96.9)	418 (96.5)	425 (97.3)	
	No	14 (3.1)	15 (3.5)	12 (2.7)	
Subjects who have signs and symptoms related to MS	n (missing)	456 (0)	433 (0)	437 (0)	0.333[c]
	Yes	428 (93.9)	416 (96.1)	416 (95.2)	
	No	28 (6.1)	17 (3.9)	21 (4.8)	

T:\BIOMET08\CLADRIBINE\25643\FINAL\TABLES\PGM\T_MH1A.SAS 21JAN2009

[a]From a two-way ANOVA model on ranked data with fixed effects for treatment group and region.

[b]From Cochran-Mantel-Haenszel row means score test, adjusted for region.

[c]From Cochran-Mantel-Haenszel general association test, adjusted for region.

Table 4 presents the MS therapy received by subjects prior to Trial Day 1 by treatment group. A greater proportion of subjects in the cladribine 3.5 mg/kg group were naïve to disease modifying drugs (DMD) compared to the cladribine 5.25 mg/kg and placebo groups (73.9%, 67.8%, and 67.5%, respectively). There was similar access across the spectrum of DMD agents previously used by subjects within the three treatment groups, with prior treatment with Copaxone® varying the most (8.3%, 4.4%, and 6.6%, for the cladribine 5.25 mg/kg group, cladribine 3.5 mg/kg and placebo groups, respectively).

Table 4 MS therapy taken by subjects prior to Study Day 1 - ITT population (Source: Table 25643-11 of sponsor's study report)

Disease Modifying Drug	Cladribine 5.25 mg/kg (n=456) n (%)	Cladribine 3.5 mg/kg (n=433) n (%)	Placebo (n=437) n (%)
Any Disease Modifying Drugs Taken	147 (32.2)	113 (26.1)	142 (32.5)
Avonex	59 (12.9)	44 (10.2)	46 (10.5)
Betaseron	42 (9.2)	42 (9.7)	56 (12.8)
Copaxone	38 (8.3)	19 (4.4)	29 (6.6)
Rebif	44 (9.6)	36 (8.3)	44 (10.1)
Tysabri	1 (0.2)	0	1 (0.2)
Other	14 (3.1)	7 (1.6)	17 (3.9)

3.1.4.3 Efficacy Results

Primary Efficacy Endpoint - Qualifying Relapse Rate

The efficacy results in relapse rate reported by the sponsor were confirmed by the reviewer. As displayed in Table 5, the annualized qualifying relapse rates were 0.15 for cladribine 5.25 mg/kg, 0.14 for cladribine 3.5 mg/kg, and 0.33 for placebo. Treatment with cladribine 5.25 mg/kg resulted in a 54.5% relative reduction in annualized qualifying relapse rate compared to placebo. Treatment with cladribine 3.5 mg/kg resulted in a 57.6% relative reduction in annualized qualifying relapse rate compared to placebo. Treatment with cladribine 5.25 mg/kg and cladribine 3.5 mg/kg compared to the placebo group resulted in a highly statistically significant difference for the qualifying relapse rate at 96 weeks ($p<0.001$ for both comparisons).

The treatment difference between the two cladribine dose groups is not statistically significant. The data suggests that treatment of cladribine 5.25 mg/kg does not provide additional benefit to the treatment of cladribine 3.5 mg/kg in reducing relapse rate.

Table 5 Relapse Rate at Week 96 - ITT Population (Source: reviewer's analysis)

		Cladribine 5.25 mg/kg (n=456)	Cladribine 3.5 mg/kg (n=433)	Placebo (n=437)
Annualized relapse rate (95% CI)	Qualifying (Primary)	0.15 (.12, .17)	0.14 (.12, .17)	0.33 (.29, .38)
	Relative Risk	0.43 (.35, .54)	.43 (.34, .54)	
	p-value	<.0001	<.0001	
	All relapses	0.28 (.24, .32)	.27 (.23, .31)	0.63 (.57, .69)
Number of Qualify Relapse	Mean (SD)	0.25 (0.58)	0.25 (0.59)	0.56 (0.88)
	Median	0	0	0
Number (%) with	0 relapse	368 (80.70%)	351 (81.06%)	276 (63.16%)
	1 relapse	69 (15.13%)	63 (14.55%)	99 (22.65%)
	2 relapses	13 (2.85%)	13 (3.00%)	44 (10.07%)
	3 relapses	5 (1.10%)	5 (1.15%)	15 (3.43%)
	4 or more relapses	1 (0.22%)	1 (0.23%)	3 (0.69%)

The results from analysis using ITT patient population were supported by results from analysis using Evaluable population.

Time to first qualifying relapse was analyzed by Cox's proportional hazard model with fixed effects of treatment and region. The following table presents the results. Kaplan-Meier curves of time to first qualifying relapse are plotted and presented in Figure 3.

Table 6 Time to first qualifying relapse - ITT population (Source: reviewer's analysis)

		Cladribine 5.25 mg/kg (n=456)	Cladribine 3.5 mg/kg (n=433)	Placebo (n=437)
Time to 1st qualifying relapse	N (events)	88 (19.3%)	82 (18.9%)	161 (36.8%)
	N (censored)	368 (80.7%)	351 (81.1%)	276 (63.2%)
	Hazard ratio (95% CI)	0.46 (0.36, 0.60)	0.44 (0.34, 0.58)	
	Nominal p-value	<.0001	<.0001	

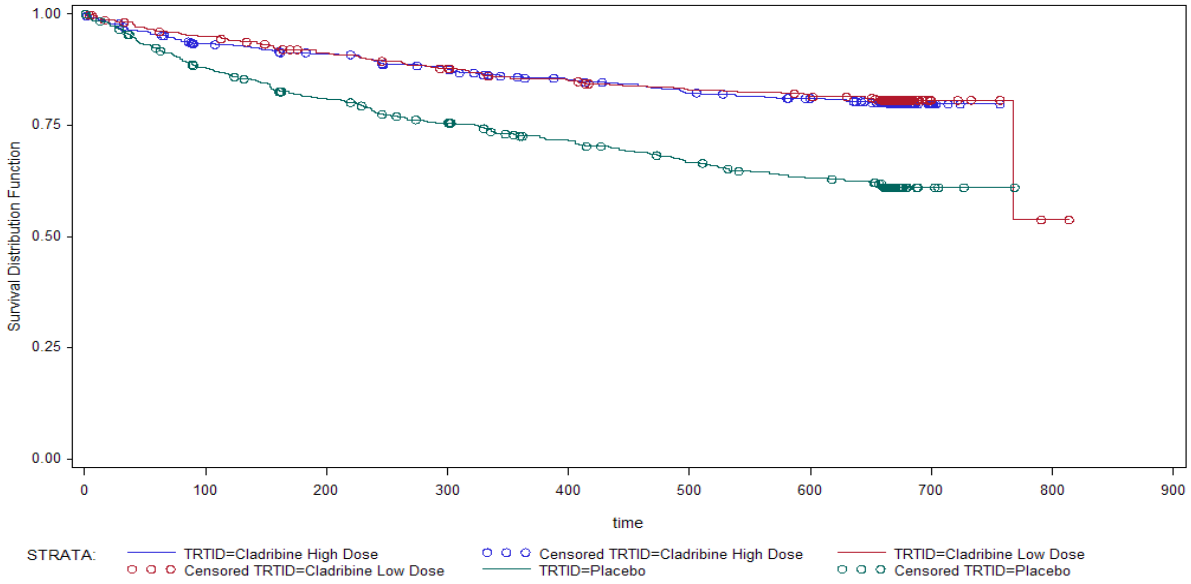


Figure 3 Kaplan-Meier curve of time to first relapse – ITT population (Source: reviewer's analysis)

Secondary Efficacy Endpoints – MRI Endpoints

It needs to be pointed out that although 5 variables are listed as secondary endpoints, only the three MRI endpoints are considered inferential. The three MRI endpoints are mean number of active T1 Gd-enhanced lesions, mean number of active T2 lesions, and mean number of CU lesions. The statistical testing with multiplicity adjustment is planned for the three MRI endpoints only. The following table presents the results from analysis of the three MRI endpoints.

Table 7 MRI results - ITT population (Source: reviewer's analysis)

		Cladribine 5.25 mg/kg (n=456)	Cladribine 3.5 mg/kg (n=433)	Placebo (n=437)
New T1 Gd-Enhanced Lesions	Mean (SD)	0.07 (0.37)	0.09 (0.30)	0.86 (1.78)
	Median	0.0	0.0	0.33
	LS Mean (SE)	0.11 (0.05)	0.12 (0.05)	0.91 (0.05)
	Caldribine-Placebo	-0.80 (0.07)	-0.78 (0.07)	
	95% CI	(-0.94, -0.66)	(-0.92, -0.65)	
	p-value	<0.0001	<0.0001	
	# Free of New T1	415 (91.01%)	376 (86.84%)	211 (48.28%)
	Nominal p-value	<0.0001	<0.0001	
Active T2 Lesions	Mean (SD)	0.29 (0.56)	0.35 (0.66)	1.38 (2.11)
	Median	0.0	0.0	0.67
	LS Mean (SE)	0.33 (0.06)	0.38 (0.07)	1.43 (0.06)
	Caldribine-Placebo	-1.10 (0.09)	-1.05 (0.09)	
	95% CI	(-1.27, -0.94)	(-1.22, -0.87)	
	p-value	<0.0001	<0.0001	
	# Free of Active T2	285 (62.50%)	267 (61.66%)	124 (28.38%)
	Nominal p-value	<0.0001	<0.0001	
CU Lesions	Mean (SD)	0.33 (0.64)	0.39 (0.71)	1.65 (2.55)
	Median	0	0	0.67
	LS Mean (SE)	0.38 (0.08)	0.43 (0.08)	1.72 (0.08)
	Caldribine-Placebo	-1.34 (0.10)	-1.28 (0.10)	
	95% CI	(-1.54, -1.14)	(-1.49, -1.08)	
	p-value	<0.0001	<0.0001	
	# Free of Active T2	277 (60.75%)	258 (59.58%)	114 (26.09%)
	Nominal p-value	<0.0001	<0.0001	

Among the 1326 patients, more than 1000 patients had 0 active T1 lesions. Only about 300 patients had one or more lesions during the study.

Similar to T1 lesions, the majority of patients did not have active T2 lesions. A small number of patients had more than 10 active T2 lesions, and a few had more than 20 active T2 lesions, almost all of them occurred in the placebo group.

The results of CU lesion counts are similar to those of T2 lesions, with majority had 0 lesion, a small number had larger than 10 lesions, and a few had larger than 20 lesions, almost all of them occurred in the placebo group.

Secondary Efficacy Endpoint - Disability Progression

As mentioned above, disability progression is considered as a secondary efficacy endpoint, but not for inferential purpose, and no multiplicity adjustment was specified. The sponsor reported

results of disability progression pre-rescue. The following table presents disability progression pre-rescue and during the entire study. The statistical analysis plan did not specify to use pre-rescue data, and the data from the entire study was assumed to be used, as understood by the reviewer. Nominal p-values from both analyses were below 0.05.

Table 8 Analysis of disability progression - ITT population (Source: reviewer's analysis)

Disability Progression		**Cladribine 5.25 mg/kg (n=456)**	**Cladribine 3.5 mg/kg (n=433)**	**Placebo (n=437)**
Pre-rescue	N (events)	62 (13.6%)	58	82
	N (censored)	394 (86.4%)	375	355
	Hazard ratio (95% CI)	0.69 (0.49, 0.96)	0.67 (0.48, 0.93)	
	Nominal p-value	0.026	0.018	
During study	N (events)	65 (14.3%)	59 (13.6%)	91 (20.8%)
	N (censored)	391 (85.7%)	374 (86.4%)	346 (79.2%)
	Hazard ratio (95% CI)	0.65 (0.47, 0.89)	0.61 (0.44, 0.84)	
	Nominal p-value	0.007	0.003	

The Kaplan-Meier curve for time to disability progression is presented in the following figure.

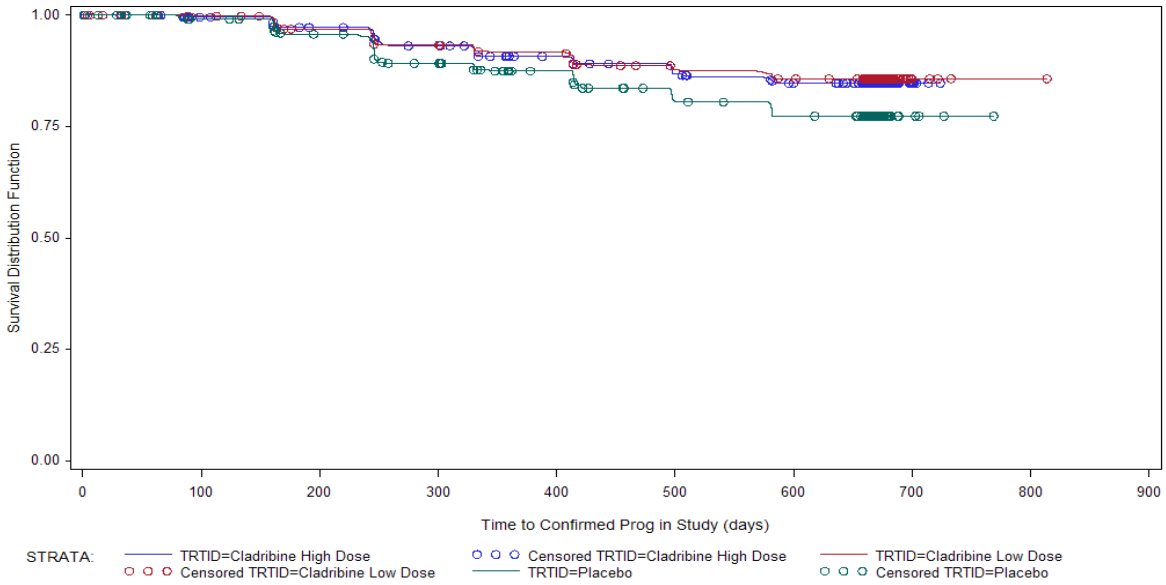


Figure 4 Kaplan-Meier curve of time to disability progression - ITT population (Source: reviewer's analysis)

3.2 Evaluation of Safety

Please refer to Clinical Evaluation by Dr. Jody Green and Evaluation of Safety by Dr. Evelyn Mentari.

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Race and Age

The analyses of relapse rate by gender and age group were performed, and the results are presented in Table 9. The median age of 39 years was used as cut-off value. Analysis of relapse rate by race was not performed since 98% of the patients were White.

Table 9 Relapse rate by gender and age group - ITT population (Source: reviewer's analysis)

		Cladribine 5.25 mg/kg (n=456)	Cladribine 3.5 mg/kg (n=433)	Placebo (n=437)
Male	N	144	135	149
	Relapse Rate	0.12	0.15	0.38
	Relative Risk	0.32	0.40	
	Nominal p-value	<.0001	<.0001	
Female	N	312	298	288
	Relapse Rate	0.16	0.14	0.31
	Relative Risk	0.51	0.44	
	Nominal p-value	<.0001	<.0001	
Age < 39 years	N	210	227	221
	Relapse Rate	0.15	0.15	0.39
	Relative Risk	0.39	0.39	
	Nominal p-value	<.0001	<.0001	
Age ≥ 39 years	N	246	206	216
	Relapse Rate	0.14	0.13	0.27
	Relative Risk	0.51	0.48	
	Nominal p-value	<.0001	<.0001	

4.2 Other Special/Subgroup Populations

Relapse rate with respect to region was analyzed, and the results are presented in the following table.

Table 10 Relapse rate by region – ITT population (Source: reviewer's analysis)

		Cladribine 5.25 mg/kg (n=456)	Cladribine 3.5 mg/kg (n=433)	Placebo (n=437)
Americas	N	47	40	44
	ARR (95%CI)	0.27 (0.15, 0.38)	0.16 (0.06, 0.26)	0.44 (0.29, 0.60)
	Relative Risk	0.60	0.35	
	p-value	0.069	0.004	
E. Europe	N	173	166	173
	ARR (95%CI)	0.16 (0.12, 0.20)	0.13 (0.09, 0.17)	0.31 (0.25, 0.37)
	Relative Risk	0.49	0.41	
	p-value	<0.001	<0.001	
W. Europe	N	91	85	80
	ARR (95%CI)	0.13 (0.07, 0.19)	0.15 (0.09, 0.21)	0.37 (0.27, 0.47)
	Relative Risk	0.36	0.40	
	p-value	<0.001	<0.001	
Rest of World	N	45	42	40
	ARR ((95%CI)	0.07 (0.01, 0.12)	0.18 (0.08, 0.28)	0.28 (0.16, 0.40)
	Relative Risk	0.23	0.63	
	p-value	0.004	0.197	
Russia	N	100	100	100
	ARR (95%CI)	0.12 (0.07, 0.17)	0.14 (0.09, 0.20)	0.32 (0.23, 0.40)
	Relative Risk	0.38	0.45	
	p-value	<0.001	0.001	

Region 4 (rest of the world) included countries of Israel, Lebanon, Turkey, Tunisia, Morocco, and Saudi Arabia. Only 5 relapses occurred among the 45 patients in the high dose group during the study in that region, resulting in a annualized relapse rate of 0.066, compared to 13 relapses among 42 patients in the low dose group (relapse rate=0.18) and 20 relapses among 40 patients in the placebo group (relapse rate=0.28).

A total of 94 subjects were enrolled in US sites. The percentages of study discontinuation and treatment discontinuation in US sites (19.1% and 22.3%, respectively) are higher than the average discontinuation rate (10.7% and 12.1%, respectively). The reviewer is particularly concerned of the outcome of two US sites: site 207 enrolled 7 subjects including 3 subjects discontinued treatment and one subject with unusually large decrease in EDSS score (from 4.5 to 0.0); site 209 enrolled 5 subjects including 2 subjects discontinued treatment and 3 subjects had unusually large increase in EDSS scores (EDSS change for the 5 subjects are 1.0 to 3.0, 2.0 to 2.5, 2.5 to 6.0, 3.0 to 6.5, 1.5 to 6.0. Because of the small numbers of subjects enrolled in each site, inspection of US sites was not performed.

Relapse rate with respect to whether subjects were previously treated by disease modifying drug was analyzed. The following table presents the results.

Table 11 Relapse rate by previous treatment of disease modifying drug - ITT population (Source: reviewer's analysis)

		Cladribine 5.25 mg/kg (n=456)	Cladribine 3.5 mg/kg (n=433)	Placebo (n=437)
Previously Treated	N	147	113	142
	Relapse Rate	0.18	0.22	0.40
	Relative Risk	0.45	0.55	
	Nominal p-value	<.0001	0.0013	
Naïve Subjects	N	309	320	295
	Relapse Rate	0.13	0.12	0.31
	Relative Risk	0.43	0.39	
	Nominal p-value	<.0001	<.0001	

5. SUMMARY AND CONCLUSIONS

5.1 Statistical Issues and Collective Evidence

The establishment of efficacy of oral cladribine is soly based on CLARITY study. Consistent and robust findings across primary and secondary efficacy endpoints in the entire ITT patient population and subgroup populations with respect to demographic and baseline characteristics are expected in order to grant the market approval for the oral cladribine.

Although concerns of data quality in some US sites remains, the impact of such data to the outcome of the study is minimal as the total number of subjects enrolled in these sites were small, and statistically significant treatment difference in favor of one or both oral cladribine doses in the relapse rate is demonstrated in all regions and in majority of countries with large enrollment.

5.2 Conclusions and Recommendations

The efficacy data from the CLARITY study (protocol 25643) support the superiority of Cladribine over placebo in reducing the relapse rate and in the reduction of MRI T1 Gs-enhanced lesions and active T2 lesions. Both doses of cladribine 5.25 mg/kg and cladribine 3.5 mg/kg are efficacious in reducing the relapse rate as well as reduction of active T1 and T2 lesions. The treatment of cladribine 5.25 mg/kg does not provide additional clinical benefit over the treatment of cladribine 3.5 mg/kg.

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
11/08/2010

KUN JIN
11/08/2010
I concur with this review.

HSIEN MING J J HUNG
11/08/2010