CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER: 20697

STATISTICAL REVIEW(S)

APR 16 1997

Statistical Review and Evaluation

DATE:

NDA#: 20-697

APPLICANT: Hoffmann-La Roche Inc.

NAME OF DRUG: TASMAR (tolcapone) Tablets

DOCUMENTS REVIEWED: Volumes 52 and 59 containing the Study Reports of the Mouse and Rat Studies and One Volume, Dated June 27, 1996 Containing the Data Diskettes and Data Listings for these Studies.

I. Background

Dr. Thomas Steele (HFD-120) requested from the Division of Biometrics I a statistical review of the rat and mouse studies data as well as an evaluation of the sponsor's findings.

II. The Rat Study

II.a. Design

The product was studied for 104 weeks in male and female Hannover-derived Wistar rats. There were two control and three treated groups of 60 animals/sex each. Ten animals per group and sex were used for plasma studies. On these, histopathology was performed only if gross lesions were apparent. The compound was administered orally in the diet at 50, 250, and 450 mg/kg/day. The two control groups were combined in the analyses. Terminal sacrifice was performed after 104 weeks of treatment.

II.b. Sponsor's Analyses of the Rat Study

Survival Analysis

The sponsor reported that survival data of each sex were analyzed by calculating Kaplan-Meier estimates, censoring on unnatural causes of death (death during blood sampling procedures, etc). They concluded that the compound had no effect on survival.

Tumor Data Analysis

The sponsor listed the following microscopic findings as being considered treatment-related, but no results of any statistical analyses were given. The findings are apparently based on all animals, i.e. including those used in the plasma study: Tubular cell carcinomas of the kidneys in males and females, spontaneous adenocarcionomas of the uterus in females, and squamous cell carcinoma and squamous cell papillomas (combined) in males. They also listed some tumors where the incidence decreased with increasing dose.

II.c. Reviewer's Analyses

The sponsor's comment on censoring only on unnatural causes of death does not make it clear whether animals terminally sacrificed were censored. This reviewer did not have information on animals lost due to trauma or blood sampling procedures and therefore independently performed analyses on the survival. For survival analysis the methods described in papers of Cox (Regression models and life tables, Journal of the Royal Statistical Society B 34, 187-220, 1972), and of Gehan (A generalized Wilcoxon test for comparing arbitrarily singly censored

samples, Biometrika 52, 203-223, 1965) were used. The corresponding computer program was written by Thomas, Breslow, and Gart (Trend and homogeneity analyses of proportions and life table data, Computers and Biomedical Research 10, 373-381, 1977, Version 2.1).

The sponsor lists several tumor findings as being associated with the treatment, but apparently did not analyze these data statistically. Statistical methods are only listed for body weight, food consumption, organ weights, and clinical laboratory data. The tumor data were therefore analyzed by this reviewer using the methods described in the paper of Peto et al. (Guidelines for simple sensitive significance test for carcinogenic effects in long-term animal experiments, Long term and short term screening assays for carcinogens: A critical appraisal, International Agency for Research against Cancer Monographs, Annex to Supplement, WHO, Geneva, 311-426, 1980) and the method of the exact permutation trend test developed by the Division of Biometrics. The following criteria for the levels of significance ensure a false positive rate of about ten percent for the trend tests of the usual two-species two-sexes studies: Tumors with less than 1.00% occurrence in the control group are considered rare and a positive trend test is statistically significant when it reaches a p-value of $\leq .025$ (one-sided). Higher tumor occurrences in the control group are considered common for these animals and a positive trend is statistically significant when its p-value is less than .005 (one-sided). An approximate permutation trend test is used when fatal and incidental tumors of the same kind are combined and have overlapping time intervals. All tests are survival adjusted and treatment groups are weighted by the actual dose levels. Trends of tumor incidence rates were computed two times, once excluding the animals used in the plasma study and once including them. These animals were histopathologically examined only in the presence of gross lesions and the finding of tumors is biased in two ways: small tumors may be missed, on the other hand examining only tissues with gross lesions raises the probability of detecting a tumor.

Survival Analysis

None of the trend tests nor the tests for departure from trend reached statistical significance for either the male or the female animals (Table 1, Figures 1 and 2). Also, none of the pairwise comparisons among the control and treated groups were statistically significant.

Tumor Data Analysis

This reviewer constructed tumor incidence tables for each recorded tissue for any tumor, treating fatal, incidental, and undetermined separately. Possible positive tumor trends with dose were then statistically analyzed adjusting for mortality despite the overall nonsignificant difference in survival. This reviewer did not analyze any negative trends observed in the tumor occurrences. The statistically significant tumor findings are summarized in Table 2. When all animals were used in the trend tests, the sponsor's figures were used which gave how many of the animals in the plasma study had a particular organ examined.

With this approach it was found that incidental, and fatal and incidental combined tubular cell carcinoma of the kidneys showed a statistically significant trend with dose (p=.0097 and p= 0102,

respectively) among the male rats when all animals were included. When excluding the animals used in the plasma study the corresponding p-values increased to .0449 and .0378, respectively. These latter p-value for the fatal and incidental combined tumors were not small enough to be called significant when adjusting for multiplicity of testing. Again among male rats, squamous cell papilloma of the stomach when combined with squamous cell carcinoma showed trends with p=.0110 when all animals were used in the statistic, and at p= .0093 when the plasma animals were excluded. These p-values are considered significant as these tumors did not occur among the control animals.

Among the female rats, when one combined tubular cell carcinoma of the kidney with tubular cell adenoma of the kidney for the core animals only, the trend test was just over the cut-off of statistical significance, namely p was .0258. When all animals were used in the trend test or when either tumor was tested for trend alone, statistical significance was not reached. Adenocarcinoma of the uterus showed significant trends when analyzing the fatal tumors, the incidental tumors, and the combined (fatal and incidental) tumors for the core animals only (p=.0067, p=.0164, and p=.0004, respectively). When including the plasma animals, the p-values corresponding to the incidental tumors and the combined tumors were not small enough to be statistically significant, in particular because two of the ten necropsied plasma animals of the control group had this tumor finding which resulted in the more stringent criterion of significance for common tumors. If this criterion is applied to adenocarcinoma of the uterus in general, then only the trend tests based on the combined fatal and incidental tumors would be called statistically significant.

This reviewer failed to reproduce the sponsor's p-values despite the given tumor incidence rates. In addition, it was not correct for the sponsor to use 50 animals as denominator of each dose group, as the denominators vary depending on the time interval of the study and whether or not the tumor was fatal or incidental. Some of these discrepancies may lead to different conclusions as to whether or not a trend statistic is considered statistically significant when taking into account the context of observation as well as the incidence rates among the concurrent controls.

III. The Mouse Study

III.a. Design

This study was conducted in Hanlbm:NMRI (SPF) mice. For each sex there were 50 animals per group. The two control groups were combined in the analyses. The dosed animals received 100, 300, and 800 mg/kg/day in the diet. Due to high mortality the females were terminated in week 80 and the males in week 95. Statistically significant differences were discussed by the sponsor only when they were considered to be biologically important.

III.b. Sponsor's Analyses of the Mouse Study

Survival Analysis

All animals were terminated early when mortality reached > 50 percent in any one of the

treatment groups. This happened after week 80 in the low dose females and after week 95 in the low dose males. Though the survival was less in the treated animals than in the controls, the sponsor did not observe any obvious treatment-related trend in the mortality experience of either the male or female mice.

Tumor Data Analysis

No neoplastic findings were considered to be treatment-related.

III.c. Reviewer's Analyses

The same statistical methods and approaches discussed for the rat study were applied to the mouse data.

Survival Analysis

The cummulative mortality is shown in Table 3 and Figures 3 and 4. None of the trend or departure from trend tests reached statistical significance. For the male mice no pairwise comparisons reached statistical significance either. Among the female mice, the pairwise comparison of controls versus low dose and controls versus high dose occasionally reached statistical significance depending on the conservatism of each test statistic ($.02 \leq p \leq .14$) (Table 4).

The intercurrent mortality tables of the sponsor and this reviewer differ slightly presumably because this reviewer counted animals dying naturally during the time of terminal sacrifice as part of the sacrificed animals, whereas the sponsor may have treated them as natural deaths. No difference in conclusions results from this discrepancy.

Tumor Data Analysis

None of the tumor findings exhibited a statistically significant positive trend with dose.

III.d. Validity of the Mouse Study

Before concluding that the mouse study showed no tumorigenic effect of tolcapone, the validity of the study needs to be determed. For this, two questions need to be answered (Haseman, Statistical Issues in the Design, Analysis and Interpretation of Animal Carcinogenicity Studies, Environmental Health Perspectives, Vol 58, pp 385-392, 1984):

(i) Were enough animals exposed for a sufficient length of time to allow for late developing tumors?

(ii) Were the dose levels high enough to pose a reasonable tumor challenge in the animals?

The following are some rules of thumb as suggested by experts in the field Haseman (Issues in

Carcinogenicity Testing: Dose Selection, Fundamental and Applied Toxicology, Vol 5, pp 66-78, 1985) had found that on the average, approximately 50 % of the animals in the high dose group survived the two-year study. In a personal communication with Dr. Karl Lin of HFD-715, he suggested that 50 % survival of the usual 50 initial animals in the high dose group between weeks 80-90 would be considered as a sufficient number and adequate exposure. Chu, Cueto, and Ward (Factors in the Evaluation of 200 National Cancer Institute Carcinogen Bioassays, Journal of Toxicology and Environmental Health, Vol 8, pp 251-280, 1981) proposed that "To be considered adequate, an experiment that has not shown a chemical to be carcinogenic should have groups of animals with greater than 50 % survival at one year". From these sources, it appears that the proportions of survival at weeks 52, 80-90, and at two years are of interest in determining the adequacy of exposure and number of animals at risk.

In determining the adequacy of the chosen dose levels, it is generally accepted that the high dose should be close to the MTD. Chu, Cueto, and Ward (1981) suggest:

(i) "A dose is considered adequate if there is a detectable weight loss of up to 10 % in a dosed group relative to the controls."

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

(iii) "In addition, doses are considered adequate if the dosed animals show a slightly increased mortality compared to the controls."

In another paper, Bart, Chu, and Tarone (Statistical Issues in Interpretation of Chronic Bioassay Tests for Carcinogenicity, Journal of the National Cancer Institute 62, 957-974, 1979), stated that the mean body weight curves over the entire study period should be taken into consideration with the survival curves, when adequacy of dose levels is to be examined. In particular, "Usually, the comparison should be limited to the early weeks of a study when no or little mortality has yet occurred in any of the groups. Here a depression of the mean weight in the treated groups is a indication that the treatment has been tested on levels at or approaching the MTD."

The study was terminated early but 50 percent survival was seen at weeks 52 and at weeks 80 for the female mice and 95 for the male mice. One can therefore conclude that there were a sufficient number of animals at risk for a sufficient length of time to allow for the manifestation of any late developing tumors.

From Figures 5 and 6, the sponsor's group mean body weight plots, one can see that the dosed animals experienced lower body weight than the controls.(It also appears that around week 34 the low dose animals, both male and female, were not properly fed. The dip in their bodyweight curves may warrant further investigation.) As the mortality of all treatment groups was strongly affected by the compound and as some non-neoplastic microscopic findings (increased incidence

of hepatocellular hypertrophy, Kupffer cell proliferation, granulocytosis, and single cell necrosis) were observed for both the medium and high dose groups, it is not clear whether the high dose was close to the MTD. The MTD may actually have been exceeded.

By all but the strictest criteria one can conclude that a sufficient number of animals were exposed to the compound for a sufficient length of time to permit the development of late developing tumors. However, the evaluation of whether the high dose was close to the MTD proved more difficult. The markers outlined by Chu, Cueto, and Ward (1981) are all exceeded and may indicate that the high dose actually exceeded the MTD.

IV. Summary and Conclusion

The two year rat study seemed to be a well executed study in which the doses administered had no apparent effect on survival. Findings in tumor incidence rates were somewhat dependent on whether or not the animals used for plasma level determinations were included. In general, tubular cell carcinoma of the kidney showed a statistically significant trends among the male rats and borderline among the female rats if only the core animals were used in the test. The combined tumors of squamouse cell papilloma and carcinoma of the stomach showed statistically significant trend amoung the male rats but did not occur among the female rats. The female rats also experienced increases in adenocarcinoma of the uterus with dose. If this tumor is common among these animals then only the combined fatal and incidental tumors reached statistical significance

The low dose animals in the mouse study experienced early mortality and the study was terminated when this group reached over 50 percent mortality. At this point, the remaining dose groups had also reached about 50 percent mortality. As the termination happened at week 80 for the female mice and at week 95 for the male mice, enough animals were exposed for a sufficient length of time for the study to be valid from this point of view. In assessing whether the high dose was close to the MTD, this reviewer observed that the markers suggested in the literature were generally exceeded in this study and that the high dose may actually have exceeded the MTD. Therefore, the validity of this study could not be concluded and the lack of observed trends in tumor incidence rates may not reflect lack of carcinogenic effect of this compound.

A final point is, that the sponsor's statistical methods applied to the mortality data and to the tumor data were not clearly described and their results not presented. Therefore, a complete statistical review had to be performed. In addition, the sponsor's p-values given for tumor trend tests could not be reproduced by this reviewer despite the tumor incidence rates provided.

Roswitha E. Kelly
Mathematical Statistician

Concur:

Todd Sahlroot, Ph. D.
Team Leader

4/16/97

George Chi, Ph.D.
Director, DB I

cc: Archival NDA 20-697, Tasmar (tolcapone) Tablets, Hoffmann-La Roche Inc.
HFD-120/Division File
HFD-120/Dr. Steele
HFD-120/Dr. Fitzgerald
HFD-710/Chron.
HFD-710/Dr. Chi
HFD-710/Dr. Sahlroot
HFD-710/Ms. Kelly
HFD-710/RKELLY/04/07/97/wp-tasmar1.rev

Table 1
INTERCURRENT MORTALITY RATES

Weeks	MALE RATS mg/kg/day 0	50	250	450
0- 52	0/100 (0%)	1/50 (2%)	0/50 (0%)	0/50 (0%)
53- 78	3/100 (3%)	3/49 (6%)	3/50 (6%)	1/50 (2%)
79- 91	12/97 (15%)	5/46 (18%)	0/47 (6%)	0/49 (2%)
92-104	7/85 (22%)	0/41 (18%)	8/47 (22%)	6/49 (14%)
Term. Sac.	78/100 (78%)	41/50 (82%)	39/50 (78%)	43/50 (86%)

Weeks	FEMALE RATS mg/kg/day 0	50	250	450
0- 52	0/100 (0%)	1/50 (2%)	0/50 (0%)	0/50 (0%)
53- 78	4/100 (4%)	3/49 (8%)	1/50 (2%)	1/50 (2%)
79- 91	8/96 (12%)	4/46 (16%)	3/49 (8%)	2/49 (6%)
92-104	9/88 (21%)	5/42 (26%)	5/46 (18%)	6/47 (18%)
Term. Sac.	79/100 (79%)	37/50 (74%)	41/50 (82%)	41/50 (82%)

Note: Except for Terminal Sacrifice, an entry of this table represents the number of animals dying or being sacrificed during the time interval divided by the number of animals entering the time interval. The entry in parenthesis is the cumulative mortality percent, i.e. the cumulative percent of animals dying up to the end of the time interval. The entry for Terminal Sacrifice represents the number of animals surviving till the end of the study divided by the initial number of animals. The entry in parentheses for this row represents the number of animals surviving to terminal sacrifice.

Table 2: Rats, Significant Tumor Findings

SEX	ORGAN AND TUMOR TYPE	GROUPING	PLASMA ANIMALS	P-VALUE
M	KIDNEY: TUBULAR CELL CARCINOMA	INCIDENTAL	INCL.	.0097 TREND
		FATAL AND INCIDENTAL	INCL.	.0102 TREND
		INCIDENTAL	EXCL.	.0449 TREND
		FATAL	EXCL.	NS
		FATAL AND INCIDENTAL	EXCL.	.0378 TREND
M	STOMACH: SQUAMOUS CELL PAPILLOMA AND SQUAMOUS CELL CARCINOMA COMBINED	INCIDENTAL	INCL.	.0110 TREND
			EXCL.	.0093 TREND
F	KIDNEY: TUBULAR CELL CARCINOMA	INCIDENTAL	INCL.	NS
		INCIDENTAL	EXCL.	NS
F	KIDNEY: TUBULAR CELL ADENOMA	INCIDENTAL	INCL.	NS
		INCIDENTAL	EXCL.	NS
F	KIDNEY: TUBULAR CELL ADENOMA AND CARCINOMA COMBINED	INCIDENTAL	INCL.	.0322 TREND
		INCIDENTAL	EXCL.	..0258 TREND
F	UTERUS: ADENOCARCINOMA	INCIDENTAL	INCL.	.0762 TREND
		FATAL AND INCIDENTAL	INCL.	.0033 TREND
		FATAL	EXCL.	.0067 TREND
		INCIDENTAL	EXCL.	.0164 TREND
		FATAL AND INCIDENTAL	EXCL.	.0004 TREND

Table 3
INTERCURRENT MORTALITY RATES

Weeks	MALE MICE mg/kg/day 0	100	300	800
0- 52	7/100 (7%)	3/50 (6%)	2/50 (4%)	5/50 (10%)
53- 78	18/93 (25%)	10/47 (26%)	9/48 (22%)	5/45 (20%)
79- 94	17/75 (42%)	14/37 (54%)	12/39 (46%)	14/40 (48%)
Term. Sac.	58/100 (58%)	23/50 (46%)	27/50 (54%)	26/50 (52%)

Weeks	FEMALE MICE mg/kg/day 0	100	300	800
0- 52	9/100 (18%)	6/50 (12%)	6/50 (12%)	6/50 (12%)
53- 79	30/91 (39%)	22/44 (56%)	16/44 (44%)	23/44 (58%)
Term. Sac.	61/100 (61%)	22/50 (44%)	28/50 (56%)	21/50 (42%)

Note: Except for Terminal Sacrifice, an entry of this table represents the number of animals dying or being sacrificed during the time interval divided by the number of animals entering the time interval. The entry in parenthesis is the cumulative mortality percent, i.e. the cumulative percent of animals dying up to the end of the time interval. The entry for Terminal Sacrifice represents the number of animals surviving till the end of the study divided by the initial number of animals. The entry in parentheses for this row represents the number of animals surviving to terminal sacrifice.

Table 4

Results of Intercurrent Mortality Analyses

Male Mice: All pairwise comparison were non-significant

Female Mice

Groups Compared	Direction	Two-tailed P-Value of Test Cox	Kruskal/Wallis	
C,L,M,H	pos	.091	.129	
C,L	pos	.112	.142	x
C,M	pos	.658	.576	
C,H	pos	.057	.071	xx
L,M	neg	.387	.403	
L,H	pos	.877	.756	
M,H	pos	.277	.300	

x: p(ChiSquare) = .036

xx: p(ChiSquare) = .021

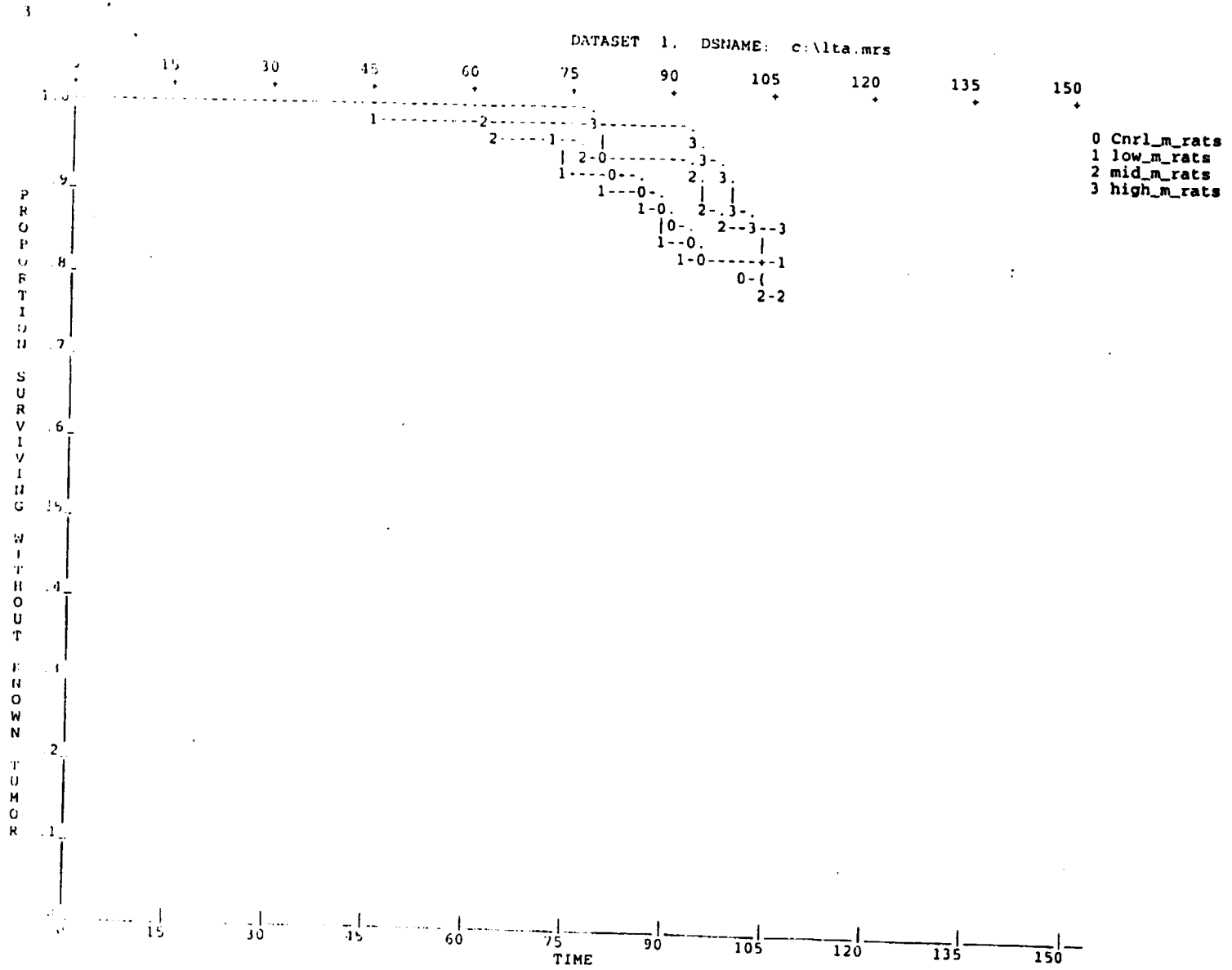

Figure 2

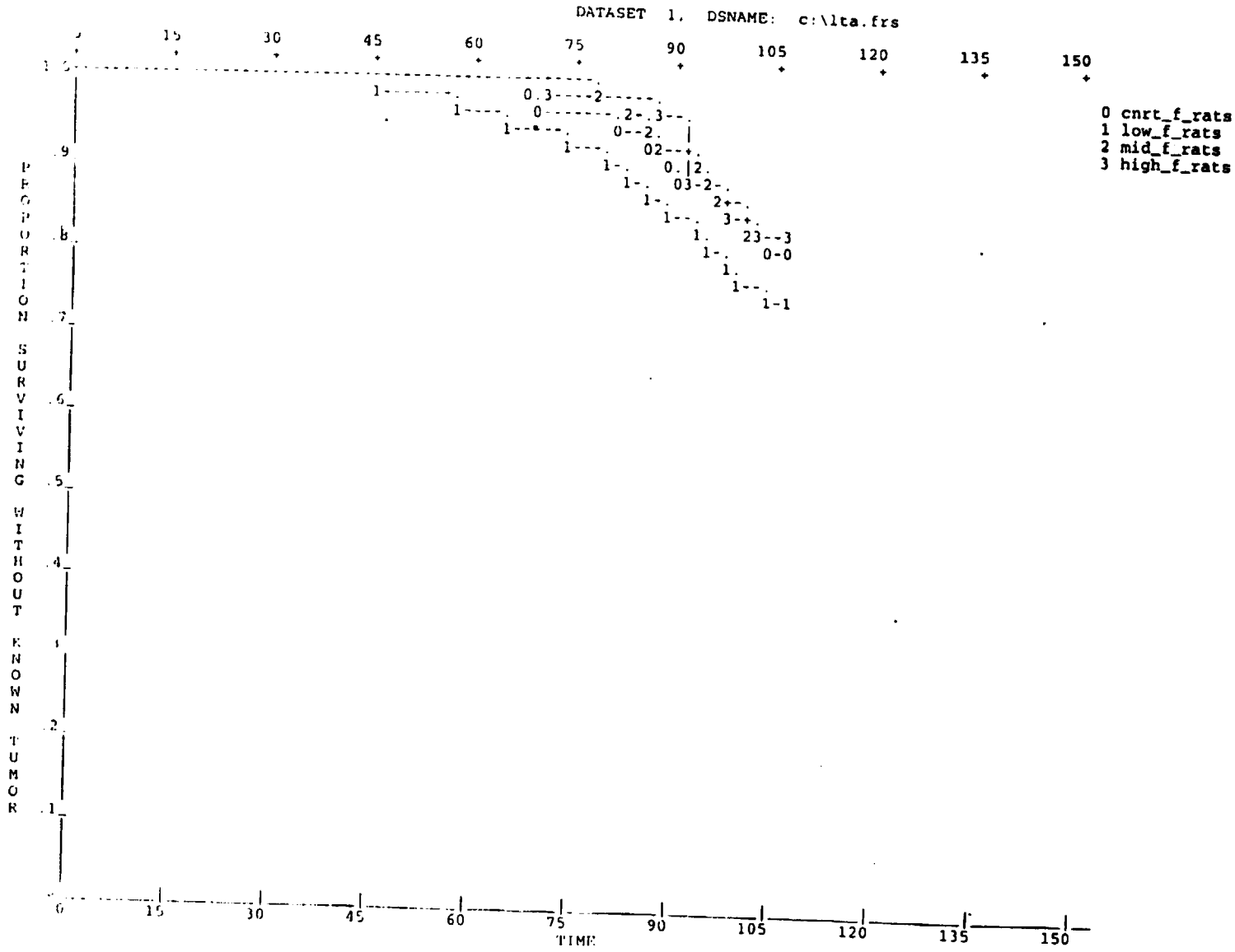

Figure 2

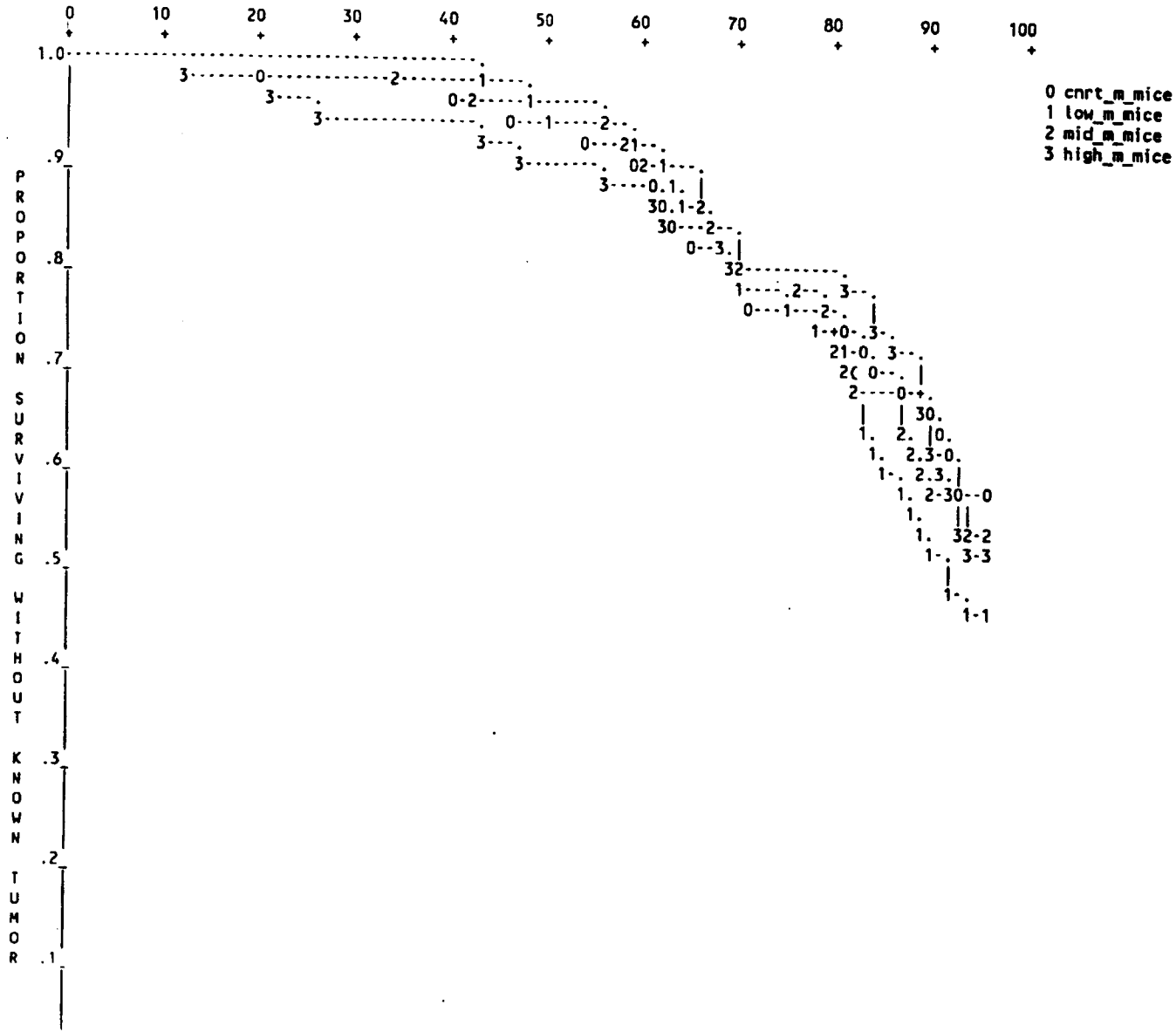

Figure 3

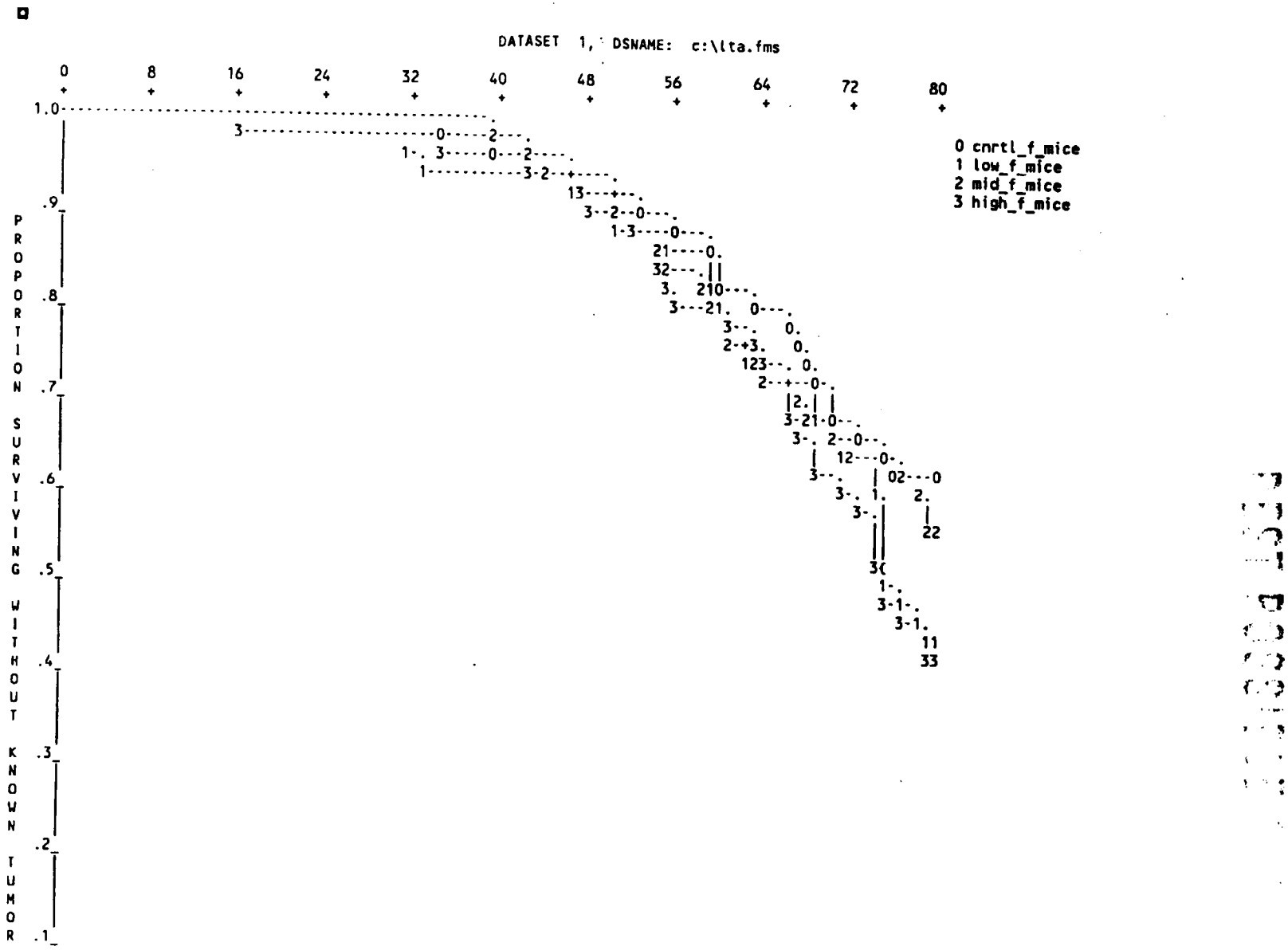
DATASET 1, DSNAME: c:\lta.fms
0 cnrtl_f_mice
1 low_f_mice
2 mid_f_mice
3 high_f_mice
PROPORTION SURVIVING WITHOUT KNOWN TUMOR

Figure 4

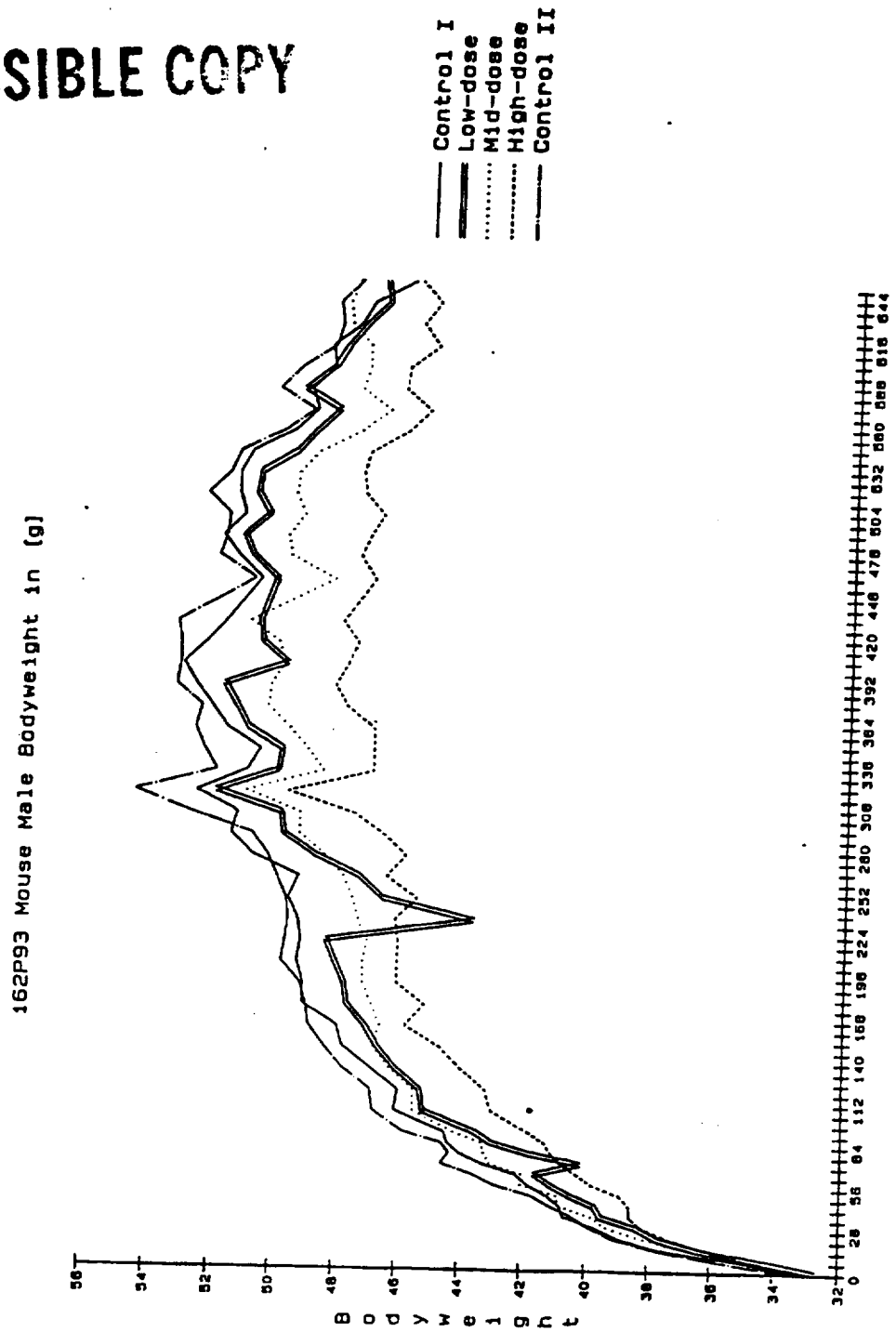

Figure 5

34

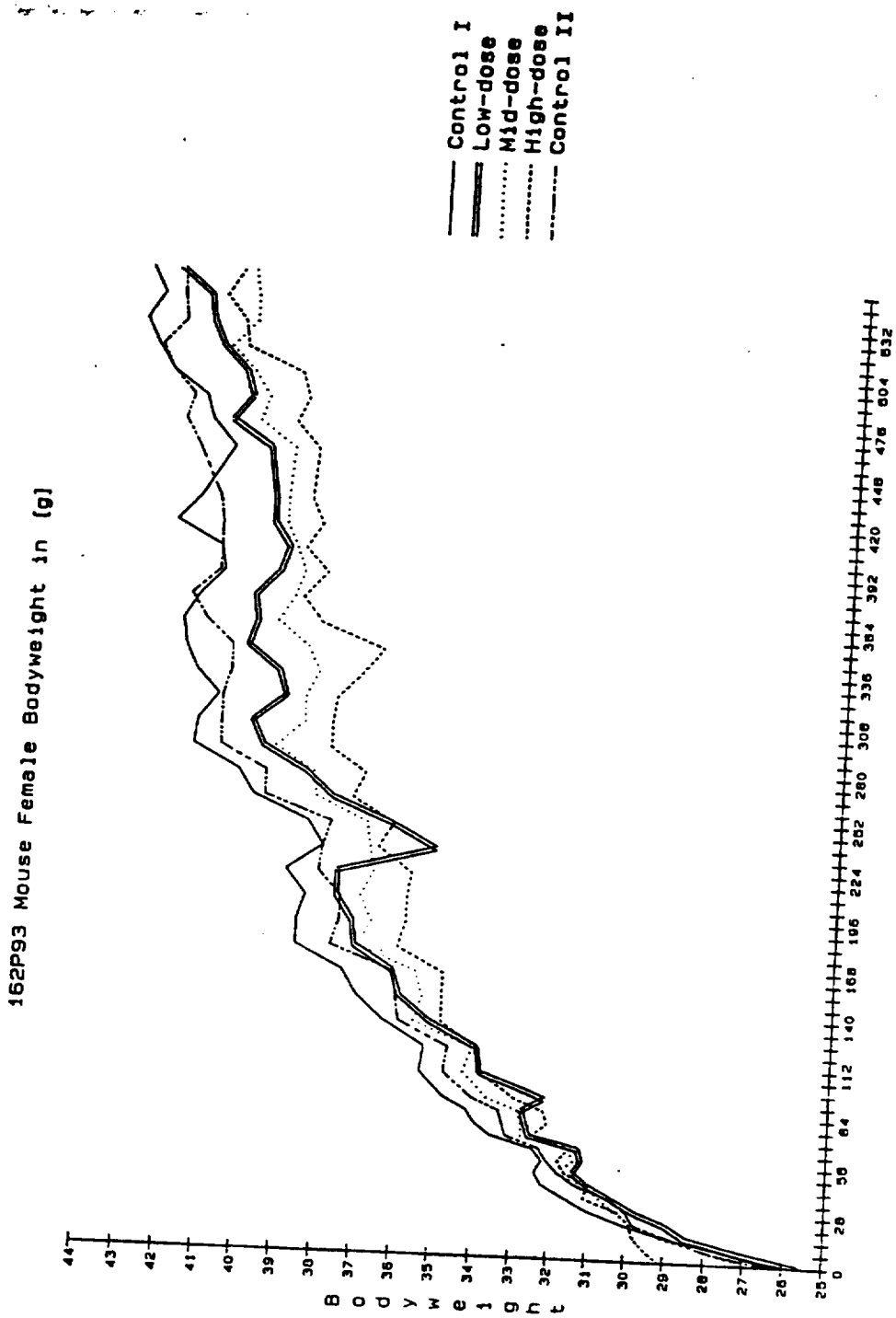
Control I
Low-dose
Mid-dose
High-dose
Control II
162P93 Mouse Female Bodyweight in [g]
Bodyweight
days
0 28 56 84 112 140 168 196 224 252 280 308 336 364 392 420 448 476 504 532
25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44

Figure 6