NPSOR-92-012

NAVAL POSTGRADUATE SCHOOL

Monterey, California

MODELING AND STATISTICAL ANALYSIS OF MEDAKA BIOASSAY DATA

Donald P. Gaver and Patricia A. Jacobs

May 1992

Prepared for:
Army Biomedical Research and Development Laboratory
Ft. Detrick MD

**NAVAL POSTGRADUATE SCHOOL,
MONTEREY, CALIFORNIA**

Rear Admiral R. W. West, Jr.
Superintendent

Harrison Shull
Provost

This report was funded by the Army Biomedical Research and Development Laboratory, Ft. Detrick, MD.

This report was prepared by:

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) NPSOR-92-012	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Naval Postgraduate School 6b. OFFICE SYMBOL (*If applicable*) OR	7a. NAME OF MONITORING ORGANIZATION Naval Postgraduate School
6c. ADDRESS (*City, State, and ZIP Code*) Monterey, CA 93943	7b. ADDRESS (*City, State, and ZIP Code*) Monterey, CA 93943-5006
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Army Biomedical Research and Development Laboratory 8b. OFFICE SYMBOL (*If applicable*)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (*City, State, and ZIP Code*) Ft. Detrick, MD	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO / WORK UNIT ACCESSION NO.

11. TITLE (*Include Security Classification*)
Modeling and Statistical Analysis of Medaka Bioassay Data

12. PERSONAL AUTHOR(S)
Donald P. Gaver and Patricia A. Jacobs

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (*Year, month day*)	15. PAGE COUNT
Technical Report	FROM_____TO_____	May, 1992	

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (*Continue on reverse if necessary and identify by block number*)
FIELD	GROUP	SUB-GROUP	Binomial Distribution; censored data; generalized linear model; bootstrap

19. ABSTRACT (*Continue on reverse if necessary and identify by block number*)

A histopathologic examination of tissues from *Oryzias latipes* (Japanese medaka fish) was performed to evaluate the carcinogenic potential of tricholoroethylene (TCE) in groundwater. The data were reported by Experimental Pathology Laboratories, Inc., in a report dated Jan. 19, 1990, submitted to the Army Biomedical Research and Development Laboratory, Ft. Detrick, MD.

This paper provides a brief statistical analysis of some aspects of those data. The analysis does not reveal a strong positive relationship between TCE concentration over the range considered and probability (risk or hazard) of incurring at least one end point manifestation (here cystic degeneration or liver neoplasm) in a fish. Uncertainties in the point estimates are assessed by bootstrapping. Both non-parametric (weak statistical assumptions) and parametric (stronger statistical assumptions) analyses give similar inconclusive dose-response indications.

A brief discussion is included of a biologically-based mathematical model that is likely to form an appropriate basis for more sophisticated data analysis.

One contribution of this paper is to discuss and illustrate techniques for quantitative analysis of other similar data. The methods can also be used to assist in choosing an experimental design.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSICIATION	
[x] UNCLASSIFIED/UNLIMITED [ ] SAME AS RPT. [ ] DTIC USERS	UNCLASSIFIED	
22a. NAME OF RESPONSIBLE INDIVIDUAL D. Gaver	22b. TELEPHONE (*Include Area Code*) (408) 646-2605	2c. OFFICE SYMBOL OR/Gv

MODELING AND STATISTICAL ANALYSIS OF MEDAKA BIOASSAY DATA

Donald P. Gaver*, Patricia A. Jacobs*
**Department of Operations Research, Naval Postgraduate School, Monterey, CA 93943*

ABSTRACT

A histopathologic examination of tissues from *Oryzias latipes* (Japanese medaka fish) was performed to evaluate the carcinogenic potential of tricholoroethylene (TCE) in groundwater. The data were reported by Experimental Pathology Laboratories, Inc., in a report dated Jan. 19, 1990, submitted to the Army Biomedical Research and Development Laboratory, Ft. Detrick, MD.

This paper provides a brief statistical analysis of some aspects of those data. The analysis does not reveal a strong positive relationship between TCE concentration over the range considered and probability (risk or hazard) of incurring at least one end point manifestation (here cystic degeneration or liver neoplasm) in a fish. Uncertainties in the point estimates are assessed by bootstrapping. Both non-parametric (weak statistical assumptions) and parametric (stronger statistical assumptions) analyses give similar inconclusive dose-response indications.

A brief discussion is included of a biologically-based mathematical model that is likely to form an appropriate basis for more sophisticated data analysis.

One contribution of this paper is to discuss and illustrate techniques for quantitative analysis of other similar data. The methods can also be used to assist in choosing an experimental design.

INTRODUCTION

The Japanese medaka fish (*Oryzias latipes*) has come to be of great interest as an indicator of groundwater toxicity; see Van Beneden *et al.* (1990), Gardner, *et al.* (1990). The Research Model Branch, Health Effects Research Division of the Army Biomedical Research and Development Laboratory, Ft. Detrick, MD, has conducted extensive experimentation with medaka so as to test its response to various known or suspected toxic agents or carcinogens. This paper provides a statistical analysis of data from such an experiment. Analysis provides a quantitative and focussed perspective on the message of the data that usefully supplements the more usual simple qualitative observations.

Design of Experiment

The experiment whose data is analyzed was planned and conducted as follows. Eight (8) groups of medaka were treated as shown in Table 1. Those groups treated

with DEN received pretreatment with 10 mg/l. of diethylnitrosamine for 48 hours at 17 days after hatch. The groups that received TCE received various concentrations of trichloroethylene (100%, 50%, 25% and 0%) on a biologically-motivated scale: 100% refers to undiluted groundwater containing TCE, and 50% and 25% refer to corresponding with pure water.

TABLE 1
TREATMENT COMBINATIONS AND CD RESPONSES

				# fish with symptom/# fish killed Sacrifice Time	
Group	DEN	no DEN	%TCE	3 months	6 months
1		X	0	6/25	4/15
3		X	25	4/25	5/13
5		X	50	2/25	4/14
7		X	100	3/25	3/14
2	X		0	11/25	6/12
4	X		25	4/25	8/13
6	X		50	6/25	5/12
8	X		100	7/25	3/8

The individual fish were assigned to tanks of water, presumably maintained at standard temperature, also presumably in a random manner. There do not appear to have been replicate tanks. After three months an interim sacrifice was made of 25 fish in each group; the number of fish showing cystic degeneration (CD), after three months/six months appear to the left of the slash (/) in the table. Thus Group 1 contained six fish out of 25 with CD after three months, and four out of 15 after six months; in the latter case 15 fish were exposed to the original concentration for the entire six months; this is referred to as the *chronic group*. Another group, the so-called *recovery group*, was placed in pure water for the second three month period. This group's response is not analyzed in this paper. Table 2 reports the incidence of liver neoplasms for the same fish in groups 2, 4, 6, 8 that were pretreated with DEN.

TABLE 2
INCIDENCE OF LIVER NEOPLASMS IN MEDAKA FOR GROUPS PRETREATED WITH DEN

		# fish with symptom/# fish killed Sacrifice time	
Group	% TCE	3 Months	6 Months
2	0	0/25	1/12
4	25	2/25	3/13
6	50	0/25	1/12
8	100	2/25	4/8

Model-based Analysis

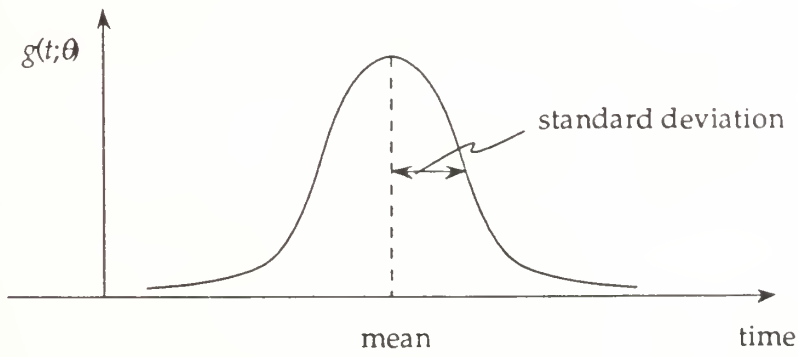
The experimental outcomes are viewed from the following perspective. Each individual fish subjected to a particular DEN-TCE treatment (e.g., DEN = 0, TCE = 50%) is initially thought of as a member of a population of similar fish. For various reasons, including that of genetic diversity, the individual fish will exhibit particular symptoms, i.e., reach specified biological endpoints such as cystic degeneration or neoplasm within specified organs, at widely different times. In addition some fish may die before any such symptoms manifest themselves. Consequently it is reasonable and natural to think of the occurrence of a particular endpoint as a probabilistic (or random, or stochastic) phenomenon, much as a coinflip or dice throw outcome is thought of, or in the same way that actuarial scientists regard human life durations when proposing life insurance contracts. That is, let T, the time to occurrence of a particular endpoint,such as cystic degeneration, be a *random variable*, a quantity whose value (for a particular fish) is determined by sampling from a population with a fixed distribution function that in turn depends upon the treatment of interest (DEN and TCE), but also upon water temperature and presence of other elements in the fish tank, and also individual fish traits. This distribution function is

$$\begin{array}{l}\text{Population Fraction of Fish with Symptom} \\ \text{Times, } T\text{, less than } t (t = 3 \text{ months, } 6 \text{ months})\end{array} = G(t,\theta)$$

The population parameter values that identify the particular distribution are called $\theta = (\theta_1, \theta_2, \ldots, \theta_p)$. For instance, θ_1 might be the population mean, and θ_2 the population standard deviation. Occasionally the specific distribution used to model the variability in a population of times is *normal*, i.e. its density function

$$g(t;\theta) = dG/dt = e^{-\frac{1}{2}(t-\theta_1)^2/\theta_2^2} \Big/ \sqrt{2\pi}\theta_2 ,$$

the familiar bell-shaped curve:



However, more frequently that data variability is better described as *log-normal*: logarithm of $T = X$ has a bell-shaped density for the raw T-values tend to straggle off to the right, i.e. the density of T is (possibly) "positively skewed." A simpler form that may be appropriate is the *exponential*:

$$G(t;\lambda) = 1 - e^{-\lambda t}.$$

In this model $\lambda = 1/$(Mean time to exhibit symptoms, e.g., CD). We will later use this exponential model in an illustrative analysis. Still another form that may be appropriate for describing the time to the onset of a cancerous growth is the *Weibull* which describes an increasing, or decraesing, time of exposure effect, depending upon data requirements. Later we shall describe a distribution that arises from plausible biological assumptions, particularly when cancer is considered—the Moolgavkar family of clonal expansion models; Moolgavkar *et al.* (1979). Use of the latter "biologically-based" family requires that at least four parameter values be estimated from data. In the light of the current experimental design such models are somewhat difficult to identify. The exponential distribution will be used in this report to illustrate a parametric analysis (one using a specific assumed form for $G(t;\theta)$).

Use of the Conceptual Model

Since TCE is a toxic substance it might be anticipated that (a) the mean or average fraction of fish exposed to x% of TCE that exhibit symptoms after $t = 3$ months (the first three months) would increase with x (the TCE concentration); likewise for six months of (chronic) exposure; and (b) that the mean fraction of the fish that survive the first three months that exhibit symptoms in the second three months (between three months and six months of exposure) *might* increase over the mean fraction in the first three months. The latter behavior would imply an *increasing hazard* property attributable to dosage with TCE. The increasing hazard property is consistent with the idea that prolonged exposure (to TCE here) increases the chances that particular endpoints will occur as time goes on. However, evidence for such behavior from the current data is not strong in the light of the uncertainties associated with sampling errors as assessed by bootstrapping.

NON-PARAMETRIC ANALYSIS

Suppose the above general sampling model prevails. Then an estimate of the probability that a fish exhibits a particular symptom, e.g., CD, within time t (= 3 months) is

$$\begin{aligned}\text{Estimate of Probability} \\ \text{That } T \le t \text{ (= 3 months)} \end{aligned} \equiv \hat{G}(3;\theta)$$
$$= \frac{\text{Number of Fish Sacrificed at } t(=3) \text{ that Exhibit Symptom (e.g., CD)}}{\text{Number of Fish Exposed (for } t = 3)}.$$

This estimate is easily calculated for all treatments; however sometimes it is zero. This suggests that the number of fish exposed is too small to be detectably influenced by the dosage; in general we might expect *some* response. Since $G(3;\theta)$ is the so-called *hazard* associated with the appearance of the particular symptom during the first three months of exposure, $\hat{G}(3;\theta)$ is an estimate thereof on the basis of only 25 exposed fish; for a different 25 fish treated equivalently we generally anticipate a

different numerical value of $\hat{G}$. By *re-sampling* ("bootstrapping) it is possible to appraise the sample variation in the estimate $\hat{G}$: sample from a binomial distribution with $\hat{G}$ being the probability of "success" = symptom occurrence within t = 3, and N(= 25), the number exposed, to obtain a pseudo or bootstrapped sample number of fish exhibiting the symptom, and from this, divided by N, a possible sample value of hazard, i.e., $\hat{G}_1$. Repeat to get $\hat{G}_2$, again to obtain $\hat{G}_3, \ldots, \hat{G}_B$, where B is "large." The sampling has been repeated B = 500 times. Then compute

$$\text{Variance } \hat{G} = \frac{1}{B}\sum_{b=1}^{B}\left(\hat{G}_b - \bar{G}_B\right)^2$$

and the standard error of the original estimate, $\hat{G}$, is $\text{SE}[\hat{G}] = \sqrt{\text{variance } \hat{G}}$ where $\bar{G}_B$ is the sample mean of the bootstrap estimates; $\bar{G}_B = \sum_{b=1}^{B}\hat{G}_b / B$. For the first three-month data the above standard error can actually be calculated directly (no re-sampling necessary): $\text{SE}[\hat{G}] = \sqrt{\frac{\hat{G}(1-\hat{G})}{N}}$, but this formula approach is not so easy for the second three-month period. Roughly speaking, the true value of $G(3)$ lies within $\hat{G}$–2SE[$\hat{G}$] and $\hat{G}$ + 2SE[$\hat{G}$]. So an estimate, and an error estimate, for initial three months hazard is obtained. See Table 3 for quoted estimates and standard errors (in parentheses).

To compare to the second three months' hazard compute

$$\begin{array}{l}\text{Estimate of Probability}\\ \text{That } T \le t = 6 \text{ Months,}\\ \text{Given that } T > t = 3 \text{ Months}\end{array} \quad = \frac{\hat{G}(6,\theta) - \hat{G}(3,\theta)}{1-\hat{G}(3;\theta)} \equiv \hat{G}(6,3;\theta))$$

= Estimate of Second 3-Months' Hazard.

Notice that the estimates $\hat{G}(6;\theta)$ and $\hat{G}(3;\theta)$ must be obtained from different sets of counts, just as was done earlier, and consequently that there is no guarantee that $\hat{G}(6;\theta)$ is greater than $\hat{G}(3;\theta)$. Although no case of such reversal occurs in the present data, a few reversals have occurred when resampling or bootstrapping is done; in such cases the hazard value is set equal to zero. Standard errors of the second 3-months' hazards are calculated by bootstrapping $\hat{G}(6,2;\theta)$ by resampling for each component, $\hat{G}(6;\theta)$ and $\hat{G}(3;\theta)$, and combining as in the formula above.

Tables 3 and 4 summarize the results of the point estimates and their standard errors. Table 3 refers to CD, while Table 4 addresses neoplasms. Figures 1 through 5 graphically display the actual hazard sampling variations as assessed by bootstrapping. Figures 1-3 present boxplots of the bootstrap sample hazards $\hat{G}_b(3;\theta)$ and $\hat{G}_b(6,3;\theta)$.

The following description of the boxplot is taken from the documentation of GRAFSTAT, a developmental product of IBM which the Naval Postgraduate School is using under a test agreement with IBM. "The box portion of the plot extends from the lower quartile of the sample to the upper quartile. (The lower quartile is the point for which one quarter of the sample lies below and three quarters above. The upper quartile is analogous.) The line across the center of the box marks the median. The circle in the box represents the mean.

The distance from the lower to the upper quartile is called the **interquartile distance**, and it will be represented by Q. The points at the ends of the two lines (called **whiskers**) are the smallest and largest points, respectively, within 1.5Q of the quantiles. The points beyond the whiskers are outlying values."

Figures 1 and 2 present the boxplots for the CD hazards. Figure 3 presents boxplots for the neoplasm. The boxplots are grouped by level of TCE which is indicated at the bottom of the figure. The left boxplot in each grouping is for the 3 month hazard. The right boxplot in each group is for the 6 month hazard. Comparison of the boxplots for the 3 and 6 month hazards in Figures 1 and 2 suggests that the 3 and 6 month hazards are roughly the same. Comparing Figures 1 and 2 suggests that the pretreatment with DEN tends to increase the hazard. Comparison of the 3 and 6 month boxplots in Figure 3 suggests the respective hazards are the same except for the 6 month hazard at the 100% TCE level, which appears to be somewhat higher than the others for neoplasms.

Figure 4 presents the histograms of the CD hazard bootstrap samples. Once again the major effect seen is the increase in hazard for the fish pretreated with DEN.

Figure 5 presents the histograms of the neoplasm hazard bootstrap samples. Once again the only histogram that appears different is the histogram for the 6 month hazard at 100% TCE.

Conclusions

The general conclusion from the above analysis is that there is only a weak effect from TCE treatment change, regardless of whether DEN is used. The effect of DEN is noticeable: the second 3 months' hazard is always somewhat larger when DEN is used than is the case with no DEN. This is anticipated, but the quantitative degree of enhancement may be of interest.

TABLE 3
NONPARAMETRIC HAZARD FOR CYSTIC DEGENERATION
(BOOTSTRAP STANDARD ERROR)

				Estimated Hazard (Standard Error) $[\sqrt{\hat{p}\hat{q}/25}]$ Sacrifice Time	
Group	DEN	No DEN	% TCE	3 Months	6 Months
1		X	0	0.24 (0.09) [0.09]	0.04 (0.11)
3		X	25	0.16 (0.07) [0.07]	0.27 (0.16)
5		X	50	0.08 (0.05) [0.05]	0.22 (0.14)
7		X	100	0.12 (0.07) [0.06]	0.11 (0.11)
2	X		0	0.44 (0.10) [0.10]	0.11 (0.20)
4	X		25	0.16 (0.07) [0.07]	0.54 (0.17)
6	X		50	0.24 (0.09) [0.09]	0.23 (0.18)
8	X		100	0.28 (0.09) [0.09]	0.13 (0.19)

TABLE 4
NONPARAMETRIC HAZARD FOR NEOPLASMS
(BOOTSTRAP STANDARD ERROR)

		Estimated Hazard (Standard Error) $[\sqrt{\hat{p}\hat{q}/25}]$ Sacrifice Time	
Group	% TCE	3 Months	6 Months
2	0	0 (0) [0]	0.08 (0.07)
4	25	0.08 (0.05) [0.05]	0.16 (0.12)
6	50	0 (0) [0]	0.08 (0.07)
8	100	0.08 (0.06) [0.05]	0.46 (0.16)

PARAMETRIC ANALYSIS

In the present context a parametric analysis of data means that a particular mathematical form is adopted for the distribution of T, the time to symptom occurrence. It is desirable that such a form have a plausible biological origin, i.e. that it can be derived from suitable biological considerations, and that it adequately represent the data. The Moolgavkar *et al.* models (1973), (1979), (1983) seem to satisfy the former requirement, but involve at least four parameters, which is too many to attempt to fit using data from the present design. Instead, the simple exponential distribution,

$$G(t;\lambda) = 1 - e^{-\lambda t}$$

has been adopted for illustration. Note that the single parameter, λ, is actually interpretable as *the inverse of the mean of* T (time to symptom occurrence) *in the population*. If this model agrees reasonably well with the data then 1/estimated $\lambda = 1/\hat{\lambda}$ is easily understood and interpreted. The exponential model also implies that the theoretical first and second 3 month hazards are the same. Notice that since no actual times to symptom appearance are ever observed such a quantity is not available from non-parametric methodology. The parameter λ (actually it is best to estimate $\gamma = \log \lambda$) must be estimated from the counts at three months and six months. The method used here is that of *maximum likelihood;* details are provided in an appendix.

Tables 5 and 6 exhibit the results of the analysis.

These results seem surprising, since mean time to exhibit the CD symptom appears to *increase* with TCE dosage; as anticipated the effect of DEN is to reduce the time to symptom appearance; these results are in rough qualitative agreement with the non-parametric results. See also Figures 6-7, which indicate the uncertainty associated with the above numerical values. These results were obtained by bootstrapping.

Figure 6 displays boxplots of the values of $-\gamma$, the log mean time to CD for the bootstrap samples. Once again the boxplots are grouped in pairs by level of TCE which is indicated on the bottom of the figure. The leftmost, (respectively rightmost), boxplot in a group is for the fish not pretreated with DEN, (respectively pretreated with DEN). Once again the major effect is a decrease in mean time to occurrence of CD with pretreatment with DEN. The variability of the estimate makes other conclusions suspect. Figure 7 displays the histograms of the bootstrap estimate values of the $-\gamma$, the log mean time to CD.

Figures 8-9 present boxplots comparing the bootstrap estimates of the probability that CD occurs before 3 months obtained from the parametric exponential model and the nonparametric analysis. The estimate for the probability using the exponential model is

TABLE 5
MAXIMUM LIKELIHOOD ESTIMATES OF MEAN TIME TO EXHIBIT CD
(BOOTSTRAP STANDARD ERROR)

Group	DEN	no DEN	%TCE	Log Mean Time to CD	$P\{$CD occurs before 3 months$\}$
				(std error)	(std error)
1		X	0	2.66 (0.32)	0.19 (0.05)
3		X	25	2.68 (0.34)	0.19 (0.05)
5		X	50	3.17 (0.45)	0.12 (0.04)
7		X	100	3.19 (0.60)	0.12 (0.04)
2	X		0	1.85 (0.26)	0.38 (0.07)
4	X		25	2.3 (0.31)	0.26 (0.07)
6	X		50	2.40 (0.34)	0.24 (0.07)
8	X		100	2.33 (0.33)	0.25 (0.07)

TABLE 6
MAXIMUM LIKELIHOOD ESTIMATES OF LOG MEAN TIME TO EXHIBIT NEOPLASMS
(BOOTSTRAP STANDARD ERROR)

EXPONENTIAL MODEL
PRETREATMENT WITH DEN

Group	% TCE	Log Mean Time to Neoplasms	$P\{$Neoplasms occurs before 3 months$\}$
2	0	4.97 (4.47)	0.02 (0.02)
4	25	3.34 (0.99)	0.10 (0.04)
6	50	4.97 (4.60)	0.02 (0.02)
8	100	2.88 (0.39)	0.15 (0.05)

$$\hat{p}_e = 1 - \exp\left\{-e^{\hat{\gamma}} 3\right\}.$$

The estimate of the probability using the nonparametric hazard is the average of the first and second 3 month hazard. The boxplots are grouped by level of TCE. The left (respectively right) one in each group are the bootstrap estimates for the parametric exponential model (respectively the nonparametric hazard). The figures suggest that the two procedures yield roughly the same estimate.

Figure 10 presents similar boxplots for the bootstrap estimates of the probability the neoplasms occur before 3 months. Note that the exponential model estimates suggest that there is no effect at 100% TCE. Note that only the chronic data is being examined.

Figure 11 present histograms for a *simulation* experiment to illustrate the effect of using more fish in the experiments. Our experiment is extreme in that 200 fish are used in each group; 100 are sacrificed at 3 months and 100 sacrificed at 6 months. The nonparametric estimates of $G(3;\theta)$ and $G(6;\theta)$ for each group of the CD data are used as the true probabilities of CD occurring at 3 and 6 months respectively. For each simulation replication 2 random numbers are drawn; one from a binomial distribution with 100 trials and probability the estimate of $G(3;\theta)$ and the other from a binomial distribution with 100 trials and probability the estimate of $G(6;\theta)$. For each group 500 simulation replications are done and the two 3 month hazards are computed for each replication as before. A comparison of the histograms in Figures 4 and 11 shows the amount of decrease in the variability of the estimates that can be achieved by increasing the number of fish used in the experiment.

BIOLOGICALLY-BASED MODEL DESCRIPTION

It is widely believed that pre-cancerous conditions in an organ (the liver) occur as a result of cell clonal expansion, followed by a promotion (to tumor) event. Specific models for this have been proposed and developed by Moolgavkar and co-workers. More recent work is by C. J. Portier and co-workers. References appear later.

The basic mechanism is treated as random or probabilistic: an **initiating event**, e.g., caused by contact with toxin, affects a cell within an organ in accordance with a simple Poisson process with rate parameter λ. That is, the chance of an uninitiated cell being initiated in time interval $(t, t+h)$ is approximately λh. *If* a cell is initiated during exposure time it clones itself into other cells at rate β; the original cells and its clones die randomly at rate δ. All cells in the organ perform thus independently, according to the model. Depending upon the values of β and δ (birth and death rates respectively) a colony of initiated cells (pre-cancerous, presumably) either tends to grow exponentially, or to die off to zero (also exponentially fast). The fates of colonies characterized by the same values of birth rate and death rate may actually be entirely different, as befits experience with variability characteristic of the real biological world. This behavior is *roughly* analogous to that of the flipping of the same coin: on one occasion 10 flips may well result in an excess of 5 Heads (7 Heads

and 3 Tails), analogous to more births (Heads) than deaths (Tails); on another sequence of 10 flips with the same coin the result may be exactly reversed (7 Tails, 3 Heads). Processes analogous to coin flipping or dice rolling can describe much, but possibly not all interesting biological variability pertinent to risk analysis. Other options are suggested later.

The values of β and δ describe clone colony properties in a precise probabilistic manner *if* the model is correct. It is certainly only approximate, but may still provide a useful tool for quantifying risk of tumor formation. The second step in the malignant cell development process is postulated to be *promotion*. A model for this is that at rate μ, i.e. with probability μh in time $(t, t+h)$, a promotion event occurs that affects one of the clone colony members in proportion to the current size of the colony; such events are assumed to occur in accordance with a Poisson process with rate proportional to instantaneous clone population size. At the instant that the *first* such promotion event occurs, the clone colony (if one exists, i.e. has been initiated) will be said to have developed a tumor, at least in informal layman's terms. Note that all original cells in an organ are assumed to be independently exposed to initiation and, thereafter, to promotion. Therefore all organ cells and subsequent clones, if any, must survive from initiation to the end of the observation period without being promoted in order for the organ to survive throughout.

The probabilistic mechanism described has been used to obtain a formula for the survival probability for an organ for any observation time t. See Appendix B for the formula and its derivation. Similar formulas have been derived also by Moolgavkar and others. Our formula provides the basis for statistically estimating from pathology data, (combinations of) the parameters: λ, the initiation rate; μ, the promotion rate; and β and δ, the clonal birth and death rates. Such estimates can, in turn, be used to estimate the probability of cell, and organ, survival for any time period. Appendix A contains a discussion of maximum likelihood estimation from data so as to specify parameters of a preliminary model. Further work is required to obtain additional statistical models and procedures to analyze other experimental data.

Extensions to the Model: Extra-variation of Parameters

The above model, and the consequences thereof in the form of a survival probability function, are appealing since they have a plausible biological basis. Organ-to-organ outcomes (tumor occurrence or not) vary randomly, but according to precisely the same mechanism in each organ; i.e. the same values of λ, μ, β and δ are assumed to hold for each organ. Note that this ignores likely variability between organs in different subjects (e.g., fish). Different, but superficially identical, biological entities, be they fish, rats, or humans, can be expected to have some differences; these can be said to be the result of genetic diversity. Specifically these differences may cause the effective parameters λ, μ β and δ to differ substantially across animals. *If* the above are estimated from data without recognizing the possibility of extra-variation, biased results will be obtained. See Harris (1990) for biological explanations of inter-organ (subject) variability.

There are several possible simple and preliminary ways of dealing with the above problem. One is by attempting to "explain" parameter variation by representing it as a function of some causal variable, such as the age, sex, weight, *etc.*, of the host subject. The technique is a variation of ordinary regression analysis; methods of McCullagh and Nelder (1983) suggest themselves. A description of a preliminary computational procedure to estimate model parameters is described in Appendix A. This procedure is used to estimate model parameters for a particular data set. A second approach is to assume that the variability between individual host organs can be represented by treating some or all of the parameters as random variables with their own distributions. A typical survival function is then obtained by *mixing*: the parametric survival function of Appendix B is "simply" randomized according to the (joint) distribution of the parameters. In principle it is desirable to recognize both sources of variability between individuals, adjusting for known sources of variation by a regression technique where possible, but recognizing the "unexplainable" variation by use of a mixing technique. The latter has been carried out to a limited extent, see Gaver and Jacobs (1992).

CONCLUSIONS AND SUMMARY

This report covers an initial short piece of research conducted under the sponsorship of the Army Biomedical Research and Development Laboratory. Its main contribution is to propose and illustrate quantitative assessments of treatment (here groundwater concentration) effects upon medaka. Those quantifications include the estimation of statistical sampling errors by the re-sampling or bootstrapping technique.

The somewhat inconclusive dose-response relationships revealed seem to imply the need for more sensitive experiments. Possibly such sensitivity can be achieved by working with more genetically homogeneous animals (medaka). Possibly, larger numbers of animal subjects will be helpful as well. Control and measurement of experimental conditions (e.g., tank temperature) and adjustments for their variations can play a useful part in the investigation.

It is hoped that the mathematical and statistical approaches illustrated here will help to promote an interest in the further use of such ideas among biologists and toxicologists.

APPENDIX A

MODEL FITTING METHODS FOR QUANTIFYING BIOASSAYS

PRELIMINARY STATISTICAL MODELS AND METHODS FOR ANALYZING BIOASSAY DATA

Suppose N organisms (for example fish) are used in an experiment. Groups of these organisms may be exposed to different treatments. Let $\mathbf{T}_i$ be the random time until organism i develops a particular symptom, e.g., cystic degeneration. Let $\mathbf{X}_i = (X_{i1}, X_{i2}, \ldots, X_{ip})$ be covariates which (possibly) influence $\mathbf{T}_i$; the $\mathbf{X}_i$ could be levels of substances having possible toxic effects to which the organisms are exposed. Let $G(t; \mathbf{x}_i) = P\left\{\mathbf{T}_i \le t \,|\, \mathbf{X}_i = \mathbf{x}_i\right\}$. We will assume that the organisms develop symptoms independently of each other. In this initial model, the symptom is either present or not.

Suppose that n_k organisms are sacrificed at time t_k with $t_1 < t_2 < \ldots < t_K$. We will label the organisms so that organisms 1 through n_1 are sacrificed at time t_1; organisms $n_1+1, \ldots, n_1+n_2$ are sacrificed at time t_2; etc. Let $s_i = 1$ if organism i exhibits the symptom when it is examined. Under the assumption of independence, the likelihood function is

$$L = \prod_{k=1}^{K} \prod_{i=1}^{n_k} G\left(t_k; \mathbf{x}_{n_{k-1}+i}\right)^{s_{n_{k-1}+i}} \left[\bar{G}\left(t_k; \mathbf{x}_{n_{k-1}+i}\right)^{\left(1-s_{n_{k-1}+i}\right)}\right] \qquad \text{(A.1)}$$

where $n_0 = 0$ and $\bar{G}(t; \mathbf{x}) = 1-G(t; \mathbf{x})$. The likelihood functions form the basis for estimation of parameters in the distributions that model survival times, i.e. G.

Example (Simple Binomial Model). If there are no covariates, then (A.1) becomes

$$L = \prod_{k=1}^{K} \binom{n_k}{f_k} G(t_k)^{f_k} \bar{G}(t_k)^{n_k - f_k} \qquad \text{(A.2)}$$

where f_k is the number of the n_k organisms exhibiting the symptom.

A procedure to estimate the parameters of the distribution G for the simple binomial model is as follows.

MAXIMUM LIKELIHOOD ESTIMATION IN THE SIMPLE BINOMIAL MODEL

(a) Likelihood and Parameter Estimation Formulas

Assume the distribution of the time to appearance of a symptom, G, is a function of the parameters θ_j, $j = 1, \ldots, J$. In this section we discuss maximum likelihood estimation of θ_j for the simple binomial model. Presumably the n_k subjects examined at time t_k, $k = 1, 2, \ldots, K$ have all been subjected to a common dosage of a potential toxin. The purpose of the present analysis is to predict survival probabilities as they depend on such dosage. The log-likelihood function for the simple binomial model is

$$l = \sum_{k=1}^{K} \ln\binom{n_k}{f_k} + f_k \ln G(t_k;\theta) + (n_k - f_k)\ln \bar{G}(t_k;\theta) \tag{A.3}$$

where $\theta = (\theta_1, \ldots, \theta_J)$. Differentiating, we obtain

$$\frac{\partial}{\partial \theta_j} l = \sum_{k=1}^{K} \frac{f_k}{G(t_k;\theta)} \frac{\partial}{\partial \theta_j} G(t_k;\theta) + \frac{(n_k - f_k)}{\bar{G}(t_k;\theta)} \left[-\frac{\partial}{\partial \theta_j} G(t_k;\theta) \right]$$

$$= \sum_{k=1}^{K} \frac{f_k[1 - G(t_k;\theta)] - (n_k - f_k)G(t_k;\theta)}{G(t_k;\theta)\bar{G}(t_k;\theta)} \left[\frac{\partial}{\partial \theta_j} G(t_k;\theta) \right]$$

$$= \sum_{k=1}^{K} \left[\frac{f_k - n_k G(t_k;\theta)}{G(t_k;\theta)\bar{G}(t_k;\theta)} \right] \frac{\partial}{\partial \theta_j} G(t_k;\theta). \tag{A.4}$$

Since $E[f_k] = n_k G(t_k;\theta)$

$$E\left[\frac{\partial^2}{\partial \theta_j \partial \theta_m} l \right] = -\sum_{k=1}^{K} n_k \frac{\frac{\partial}{\partial \theta_j} G(t_k;\theta) \frac{\partial}{\partial \theta_m} G(t_k;\theta)}{G(t_k;\theta)\bar{G}(t_k;\theta)}. \tag{A.5}$$

Thus a Newton procedure for finding the maximum likelihood estimates of $\{\theta_j; j = 1, \ldots, J\}$ would iteratively solve the system of linear equations

$$0 = \frac{\partial}{\partial \theta_j} l(\theta^0) + \sum_{m=1}^{P} E\left[\frac{\partial^2}{\partial \theta_j \partial \theta_m} l \right] (\theta_m - \theta_m^0) \tag{A.6}$$

where $\theta^0 = \left(\theta_1^0, \ldots, \theta_J^0\right)$. Such iterative procedures can be programmed for a digital computer, and the resulting parameter values can be used to compute predictions for survival probabilities, or risk, as the latter depend upon the parameters of such models as described in Appendix B.

APPENDIX B. TWO-STAGE CLONAL-EXPANSION MODEL

In this appendix we present a birth-death model for the distribution of time until a normal cell becomes promoted to a tumor.

We first develop an expression for the distribution of random time, S, until an initiated cell or one of its descendants becomes malignant.

Assume that there is one initiated cell at time 0. Such cells divide at an exponential rate β, and die at an exponential rate δ. Any initiated cell turns malignant at an exponential rate μ; i.e. μ is the **promotion rate.**

(a) Time to Promotion of an Initiated Cell

Let S be the random time at which some initiated cell or its descendent turns malignant; note that S may actually be infinite if the population of initiated cell and its descendents dies out. Put

$$z(t) = P\{S > t\}.$$

The following probability argument provides an equation for $z(t)$: the event that $S > t+\Delta$ $(\Delta > 0)$ occurs if (i) neither birth (cloning), death, or promotion occurs in $(0, \Delta)$ *and* promotion does not occur in $(\Delta, t+\Delta)$; the probability of this is $[1 - (\beta+\delta+\mu)\Delta + o(\Delta)]z(t)$; *or* (ii) birth/cloning occurs in $(0,\Delta)$ and no promotion occurs in $(\Delta, t+\Delta)$; the probability of this event is $[\beta\Delta + o(\Delta)]z^2(t)$, where the square recognizes that at time Δ there are now two independent clonal families to be considered; *or* (iii) the original initiated cell dies in $(0, \Delta)$, the probability of which is $\delta\Delta + o(\Delta)$. Sum these three terms to obtain the probability that $S > t+\Delta$:

$$z(t + \Delta) = (1-(\beta + \delta + \mu)\Delta)z(t) + \beta\Delta z^2(t) + \delta\Delta.$$

Now subtract $z(t)$ from each side, divide by Δ and let $\Delta \rightarrow 0$. The result is the differential equation

$$\frac{dz(t)}{dt} = -(\beta + \delta + \mu)z(t) + \beta z^2(t) + \delta \qquad \text{(B.1)}$$

Hence $z(t)$ satisfies a Riccati equation with initial condition

$$z(0) \equiv 1 \qquad \text{(B.2)}$$

The solution to (B.1) with initial condition (B.2) is

$$z(t) = \frac{\rho_1(1-\rho_2) - \rho_2(1-\rho_1)e^{\beta(\rho_1-\rho_2)t}}{1-\rho_2-(1-\rho_1)e^{\beta(\rho_1-\rho_2)t}} \tag{B.3}$$

where $\rho_{1,2}$ are the solutions to the quadratic equation

$$x^2 - \left(1+\frac{\delta}{\beta}+\frac{\mu}{\beta}\right)x + \frac{\delta}{\beta} = 0; \tag{B.4}$$

$$\rho_{1,2} = \frac{1}{2}\left[\left(1+\frac{\delta}{\beta}+\frac{\mu}{\beta}\right) \pm \left[\left(1+\frac{\delta}{\beta}+\frac{\mu}{\beta}\right)^2 - 4\frac{\delta}{\beta}\right]^{\frac{1}{2}}\right]. \tag{B.5}$$

Since $\left[\left(1+\frac{\delta}{\beta}+\frac{\mu}{\beta}\right)^2 - 4\frac{\delta}{\beta}\right]^{1/2} \leq \left(1+\frac{\delta}{\beta}+\frac{\mu}{\beta}\right)$, both ρ_1 and ρ_2 are positive. Further $\rho_2 \leq 1$ and

$$\rho_1 - \rho_2 = \left[\left(1+\frac{\delta}{\beta}+\frac{\mu}{\beta}\right)^2 - 4\frac{\delta}{\beta}\right]^{\frac{1}{2}} > 0.$$

Hence,

$$\lim_{t\to\infty} z(t) = \rho_2. \tag{B.6}$$

If the death rate $\delta = 0$, then $\rho_2 = 0$ and $\lim_{t\to\infty} P\{T > t\} = 0$; if $\delta = 0$, then there is no death of initiated cells and thus an initiated cell will transition to a malignant cell in a finite time with probability 1. If $\delta > 0$, then the initiating cells can die, thus preventing a transition to malignancy and hence $\lim_{t\to\infty} P\{T > t\} = \rho_2 > 0$.

(b) Model for the Time until a Normal Cell becomes Malignant (is Promoted to Tumor)

Assume that each normal cell is initiated at an exponential rate λ_0. Let N be the total number of normal cells in an organ. Let T denote the first time a normal cell transitions to a malignant cell.

$$P\{T \geq t\} = \left[e^{-\lambda_0 t} + \int_0^t \lambda_0 e^{-\lambda_0 s} z(t-s)ds \right]^N \tag{B.7}$$

where z is a given in (B.3). Assume λ_0 is small and put $\lambda = \lambda_0 N$, a constant. Then

$$P\{T > t\} \approx \exp\left\{ N \ln\left[1 - \frac{\lambda}{N} t + \frac{\lambda}{N} \int_0^t z(s)ds \right] \right\} \tag{B.8}$$

$$\approx \exp\left\{ -\lambda t + \lambda \int_0^t z(s)ds \right\} \tag{B.9}$$

$$= \exp\left\{ \lambda(\rho_1 - 1)t - \lambda \frac{1}{\beta} \ln\left[\frac{1 - \rho_2 + (\rho_1 - 1)e^{\beta(\rho_1 - \rho_2)t}}{\rho_1 - \rho_2} \right] \right\}. \tag{B.10}$$

BIBLIOGRAPHY AND REFERENCES

Efron, B. and R. Tibshirani, "Bootstrap Methods for Standard Errors, Confidence Intervals, and other Measures of Statistical Accuracy," *Statistical Science*, Vol. 1, (1986), pp. 54-77.

Gardner, H.S., van der Schalie, M.J. Wolfe and R.A. Finch (1990). "New methods for on-site biological monitoring of effluent water quality." *In Situ Evaluations of Biological Hazards of Environmental Pollutants*, ed. by S. S. Sandhu *et al.*, Plenum Press, 1990; pp. 61-69.

Gaver, D. P. and P. A. Jacobs, *Susceptibility and Exposure Variation in the Two-stage Clonal Expansion Model*, in progress.

Harris, C. C., "Interindividual Variation in Human Chemical Carcinogenesis: Implications for Risk Assessment," *Scientific Issues in Quantitative Cancer Risk Assessment* (Abbreviate SIQCRA) ed. by S. H. Moolgavkar. Birkhauser. Boston, 1990. pp. 235-251.

Karlin, S., and H. M. Taylor, *A Second Course in Stochastic Processes*, Academic Press, New York, 1981.

Kopp-Schneider, A., C. J. Portier and F. Rippman, "The Application of a Multistage Model that Incorporates DNA Damage and Repair to the Analysis of Initiation/Promotion Experiments," *Math. Biosciences*, Vol. 105, 1991, pp. 139-166.

McCullagh, P., and J. A. Nelder, *Generalized Linear Models*, Chapman and Hall, New York, 1983.

Moolgavkar, S. H., and D. J. Venzon, "Two-event Models for Carcinogenesis: Incidence Curves for Childhood and Adult Tumors," *Math. Biosciences*, Vol. 47, 1979, pp. 55-77.

Moolgavkar, S. H., "Model for Human Carcinogenesis: Action of Environmental Agents," *Environmental Health Perspectives*, Vol. 50, 1973, pp. 285-291.

Moolgavkar, S. H., N. E. Day and R. G. Stevens, "Two-stage Model for Carcinogenesis: Epidemiology of Breast Cancer in Females," *JNCI*, Vol. 65, 1983, pp. 559-569.

Portier, C. J., "Two-stage Models of Tumor Incidence for Historical Control Animals in the National Toxicology Program's Carcinogenicity Experiment," *J. of Toxicology and Environmental Health*, Vol. 27, 1989, pp. 21-45.

Portier, C. J., "Utilizing Biologically Based Models to Estimate Carcinogenic Risk," *SIQCRA*, pp. 252-266.

Van Beneden, R. J., K. W. Henderson, D. G. Blair, T. S. Papas, and H. S. Gardner, "Oncogenes in hematopoietic and hepatic fish neoplasms." *Cancer Research (Suppl.)*, Vol 50, pp. 5671s-5674s.

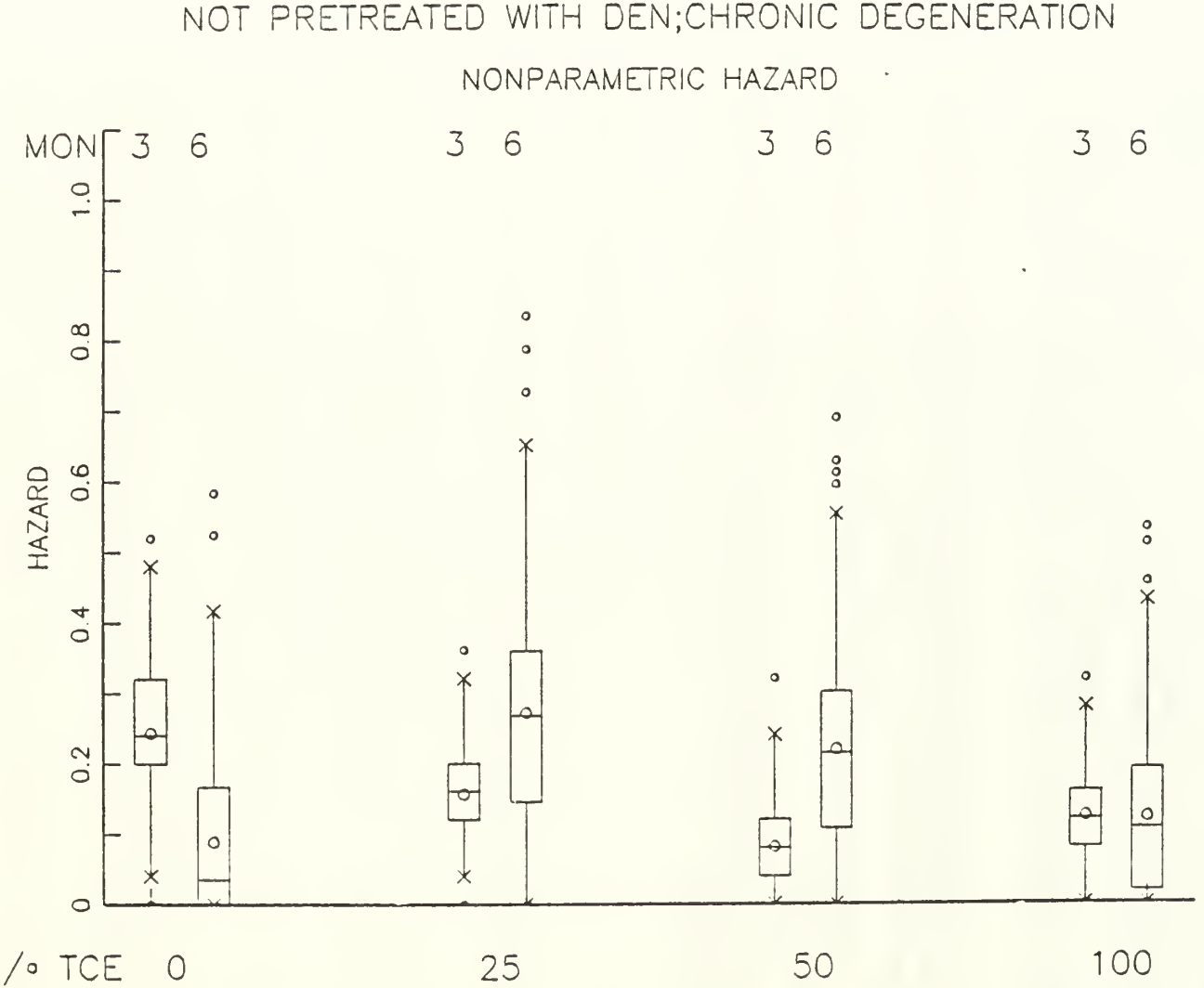

NOT PRETREATED WITH DEN;CHRONIC DEGENERATION
NONPARAMETRIC HAZARD
MON 3 6 3 6 3 6 3 6
HAZARD
0
0.2
0.4
0.6
0.8
1.0
% TCE 0 25 50 100

Figure 1

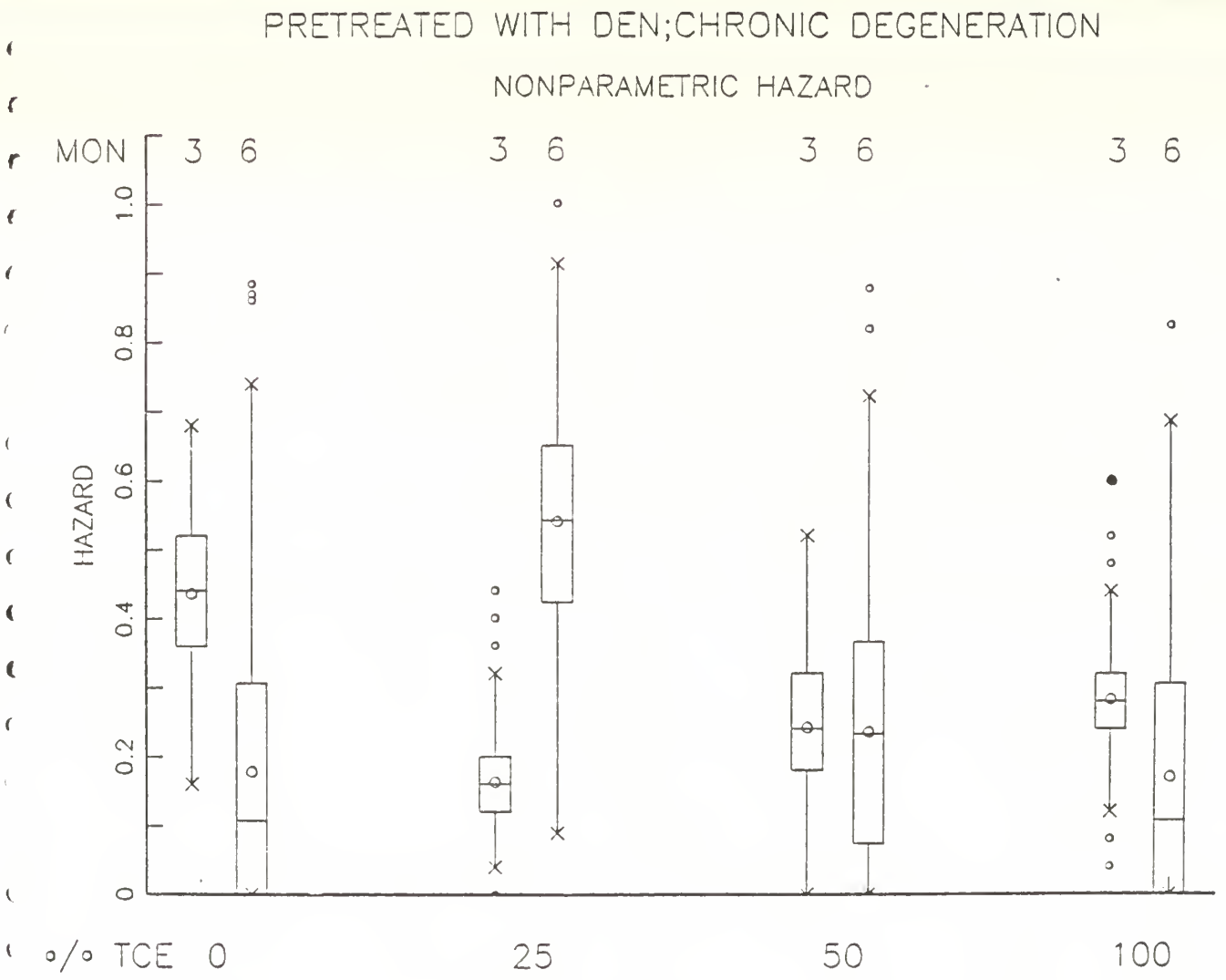

PRETREATED WITH DEN;CHRONIC DEGENERATION
NONPARAMETRIC HAZARD
MON
3 6 3 6 3 6 3 6
HAZARD
0
0.2
0.4
0.6
0.8
1.0
% TCE 0 25 50 100

Figure 2

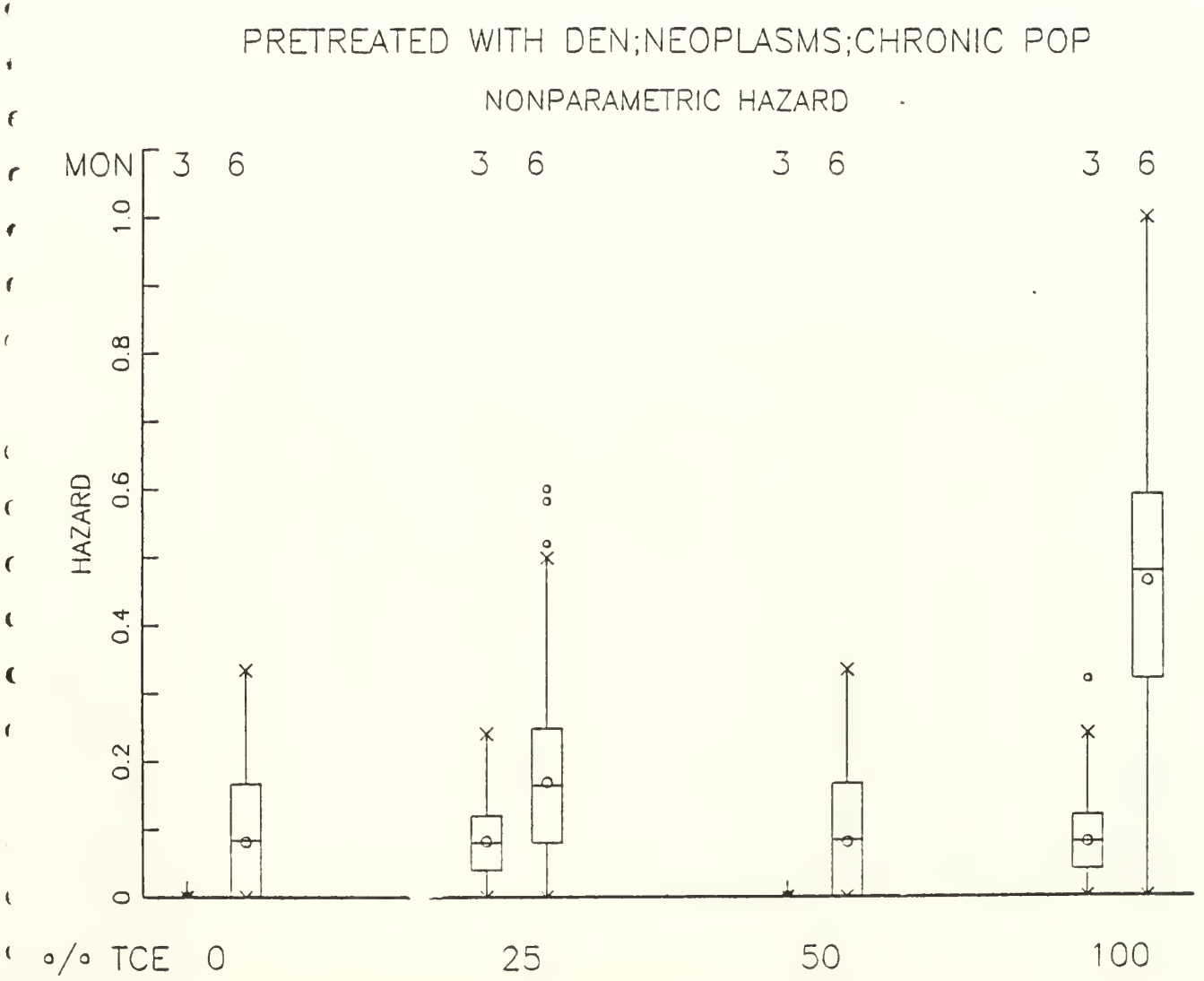

PRETREATED WITH DEN;NEOPLASMS;CHRONIC POP
NONPARAMETRIC HAZARD
MON
3 6 3 6 3 6 3 6
HAZARD
0
0.2
0.4
0.6
0.8
1.0
% TCE 0 25 50 100

Figure 3

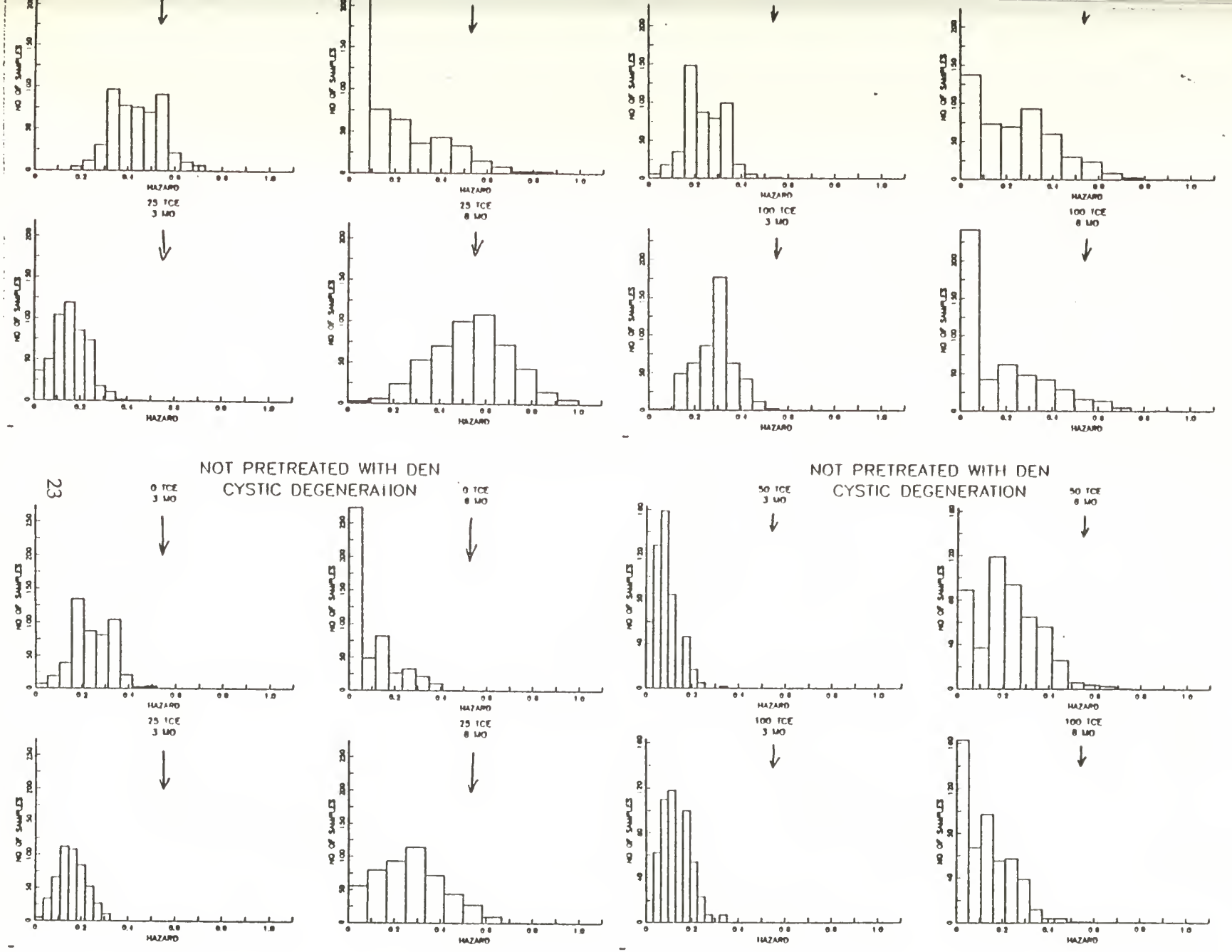
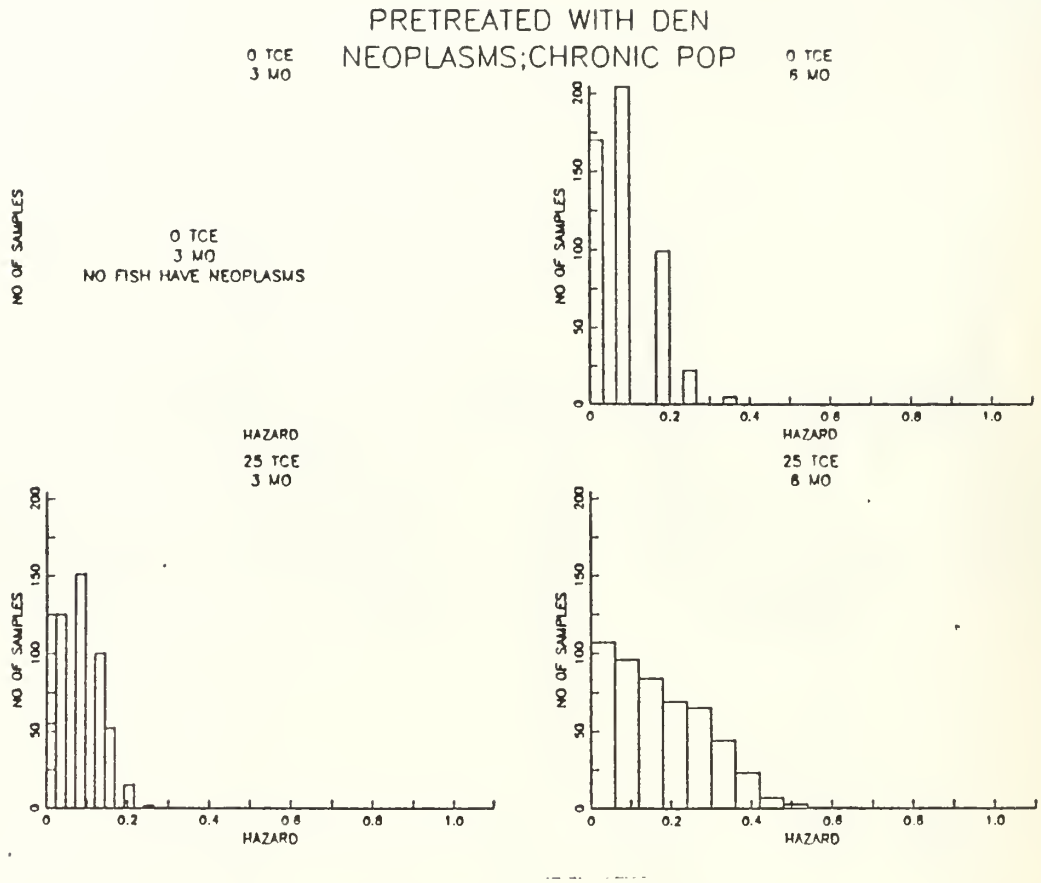

NOT PRETREATED WITH DEN
CYSTIC DEGENERATION
NOT PRETREATED WITH DEN
CYSTIC DEGENERATION
0 TCE
3 MO
0 TCE
8 MO
25 TCE
3 MO
25 TCE
8 MO
50 TCE
3 MO
50 TCE
8 MO
100 TCE
3 MO
100 TCE
8 MO
NO OF SAMPLES
HAZARD

Figure 4


PRETREATED WITH DEN
NEOPLASMS;CHRONIC POP
0 TCE
3 MO
0 TCE
6 MO
0 TCE
3 MO
NO FISH HAVE NEOPLASMS
NO OF SAMPLES
HAZARD
25 TCE
3 MO
25 TCE
6 MO

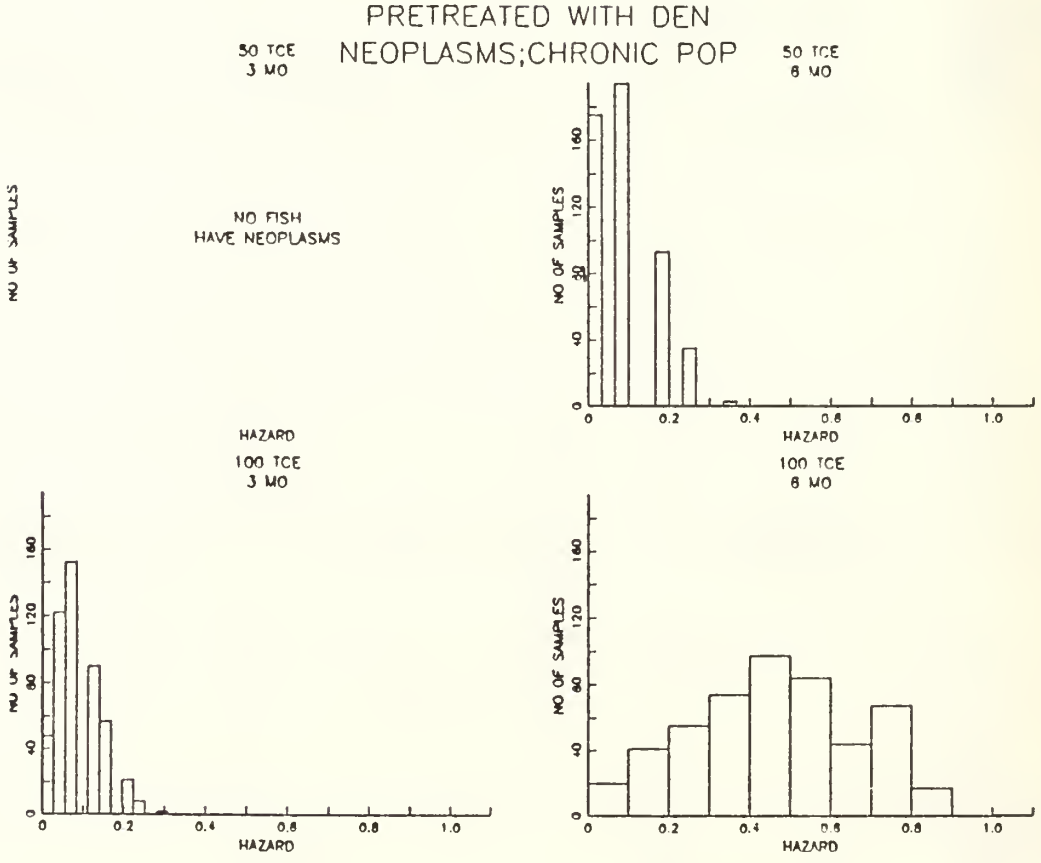

PRETREATED WITH DEN
NEOPLASMS;CHRONIC POP
50 TCE
3 MO
50 TCE
6 MO
NO FISH
HAVE NEOPLASMS
NO OF SAMPLES
HAZARD
100 TCE
3 MO
100 TCE
6 MO

Figure 5

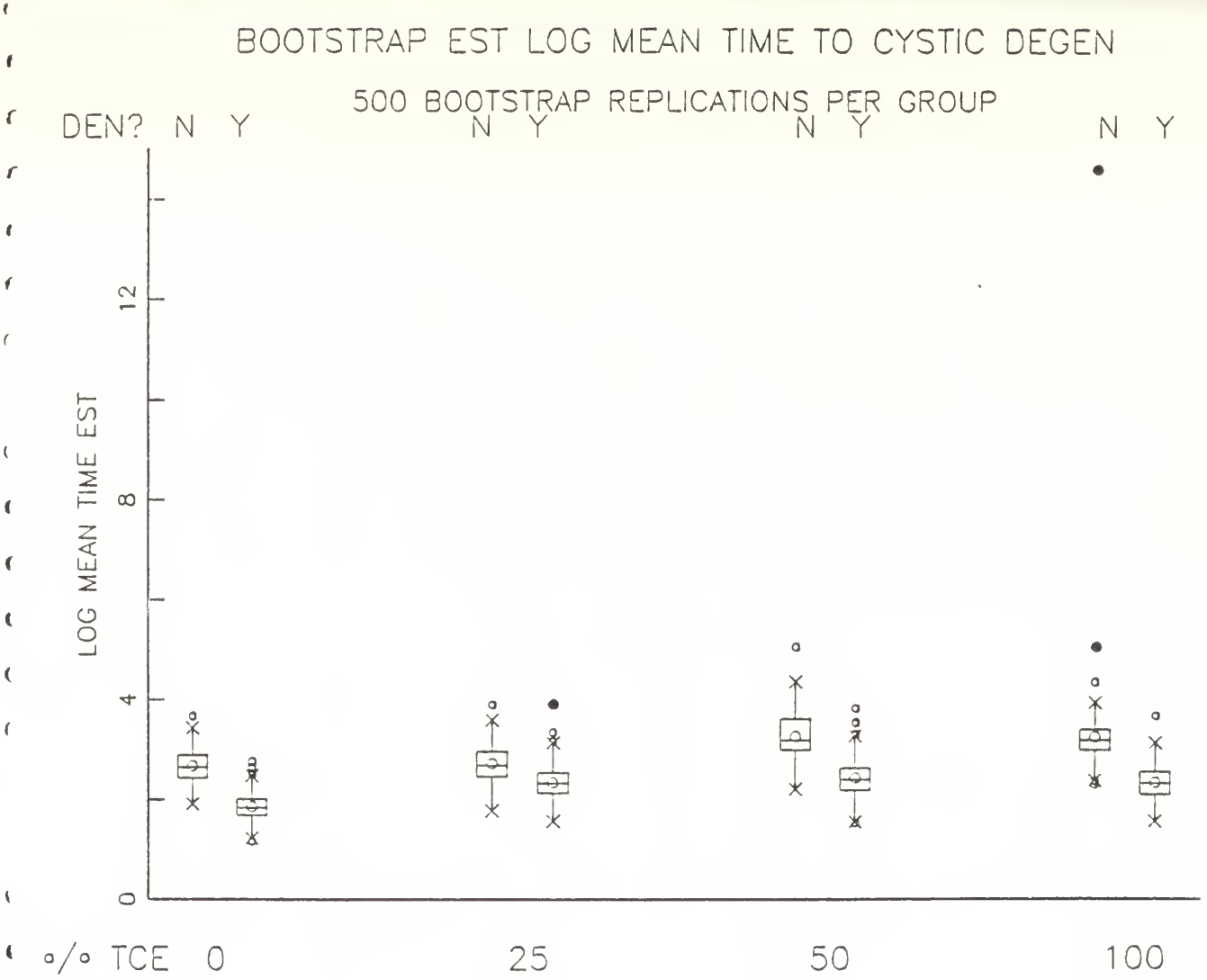
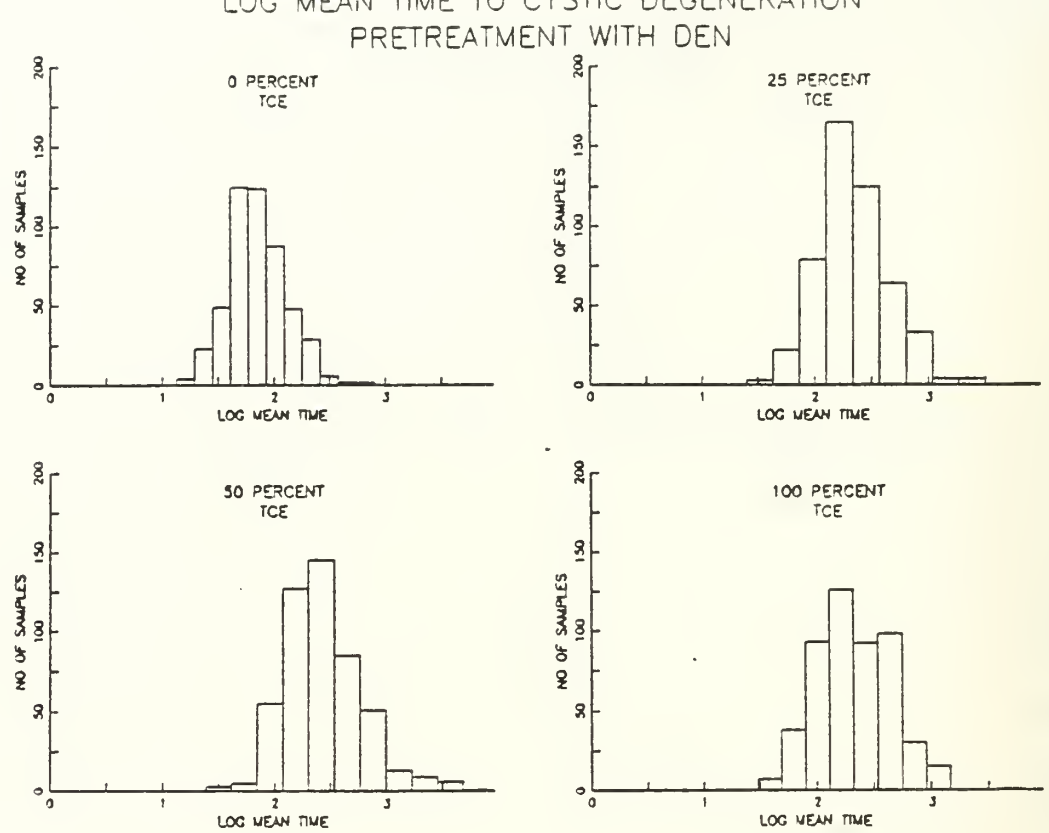

BOOTSTRAP EST LOG MEAN TIME TO CYSTIC DEGEN
500 BOOTSTRAP REPLICATIONS PER GROUP
DEN? N Y N Y N Y N Y
LOG MEAN TIME EST
0
4
8
12
% TCE 0 25 50 100

Figure 6


LOG MEAN TIME TO CYSTIC DEGENERATION
PRETREATMENT WITH DEN
0 PERCENT TCE
25 PERCENT TCE
50 PERCENT TCE
100 PERCENT TCE
NO OF SAMPLES
LOG MEAN TIME

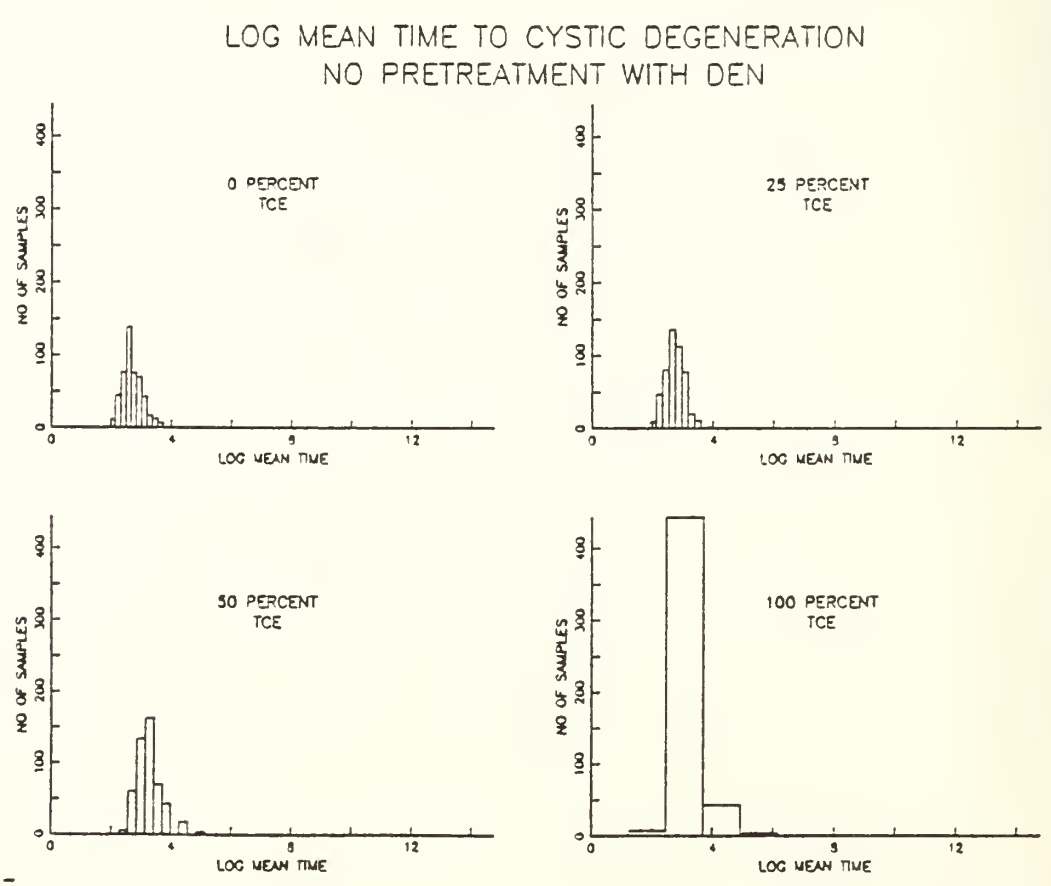

LOG MEAN TIME TO CYSTIC DEGENERATION
NO PRETREATMENT WITH DEN
0 PERCENT TCE
25 PERCENT TCE
50 PERCENT TCE
100 PERCENT TCE
NO OF SAMPLES
LOG MEAN TIME

Figure 7

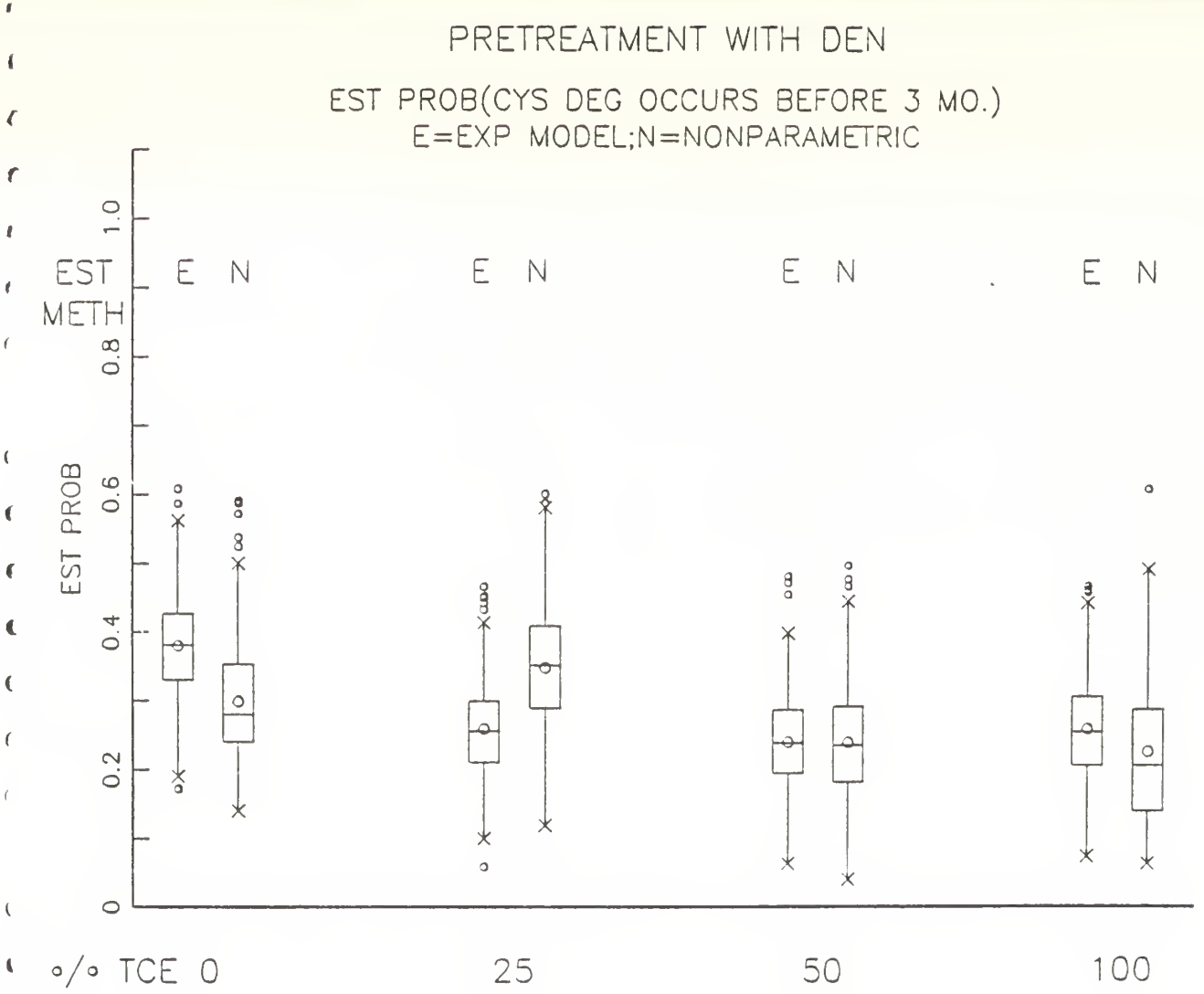

PRETREATMENT WITH DEN
EST PROB(CYS DEG OCCURS BEFORE 3 MO.)
E=EXP MODEL;N=NONPARAMETRIC
EST METH
E N
E N
E N
E N
EST PROB
0
0.2
0.4
0.6
0.8
1.0
% TCE 0
25
50
100

Figure 8

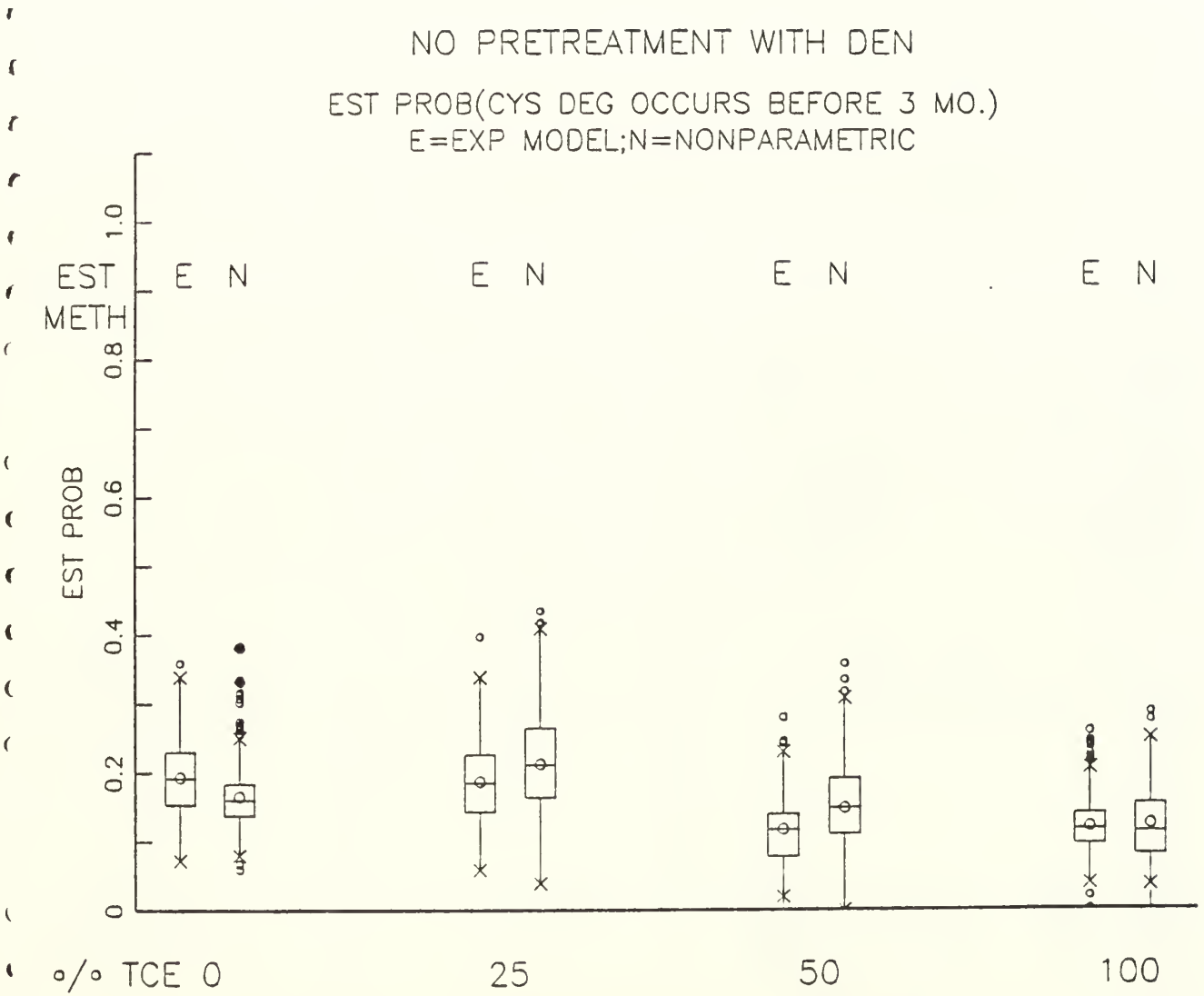

NO PRETREATMENT WITH DEN
EST PROB(CYS DEG OCCURS BEFORE 3 MO.)
E=EXP MODEL;N=NONPARAMETRIC
EST METH
E N
E N
E N
E N
EST PROB
0
0.2
0.4
0.6
0.8
1.0
% TCE 0
25
50
100

Figure 9

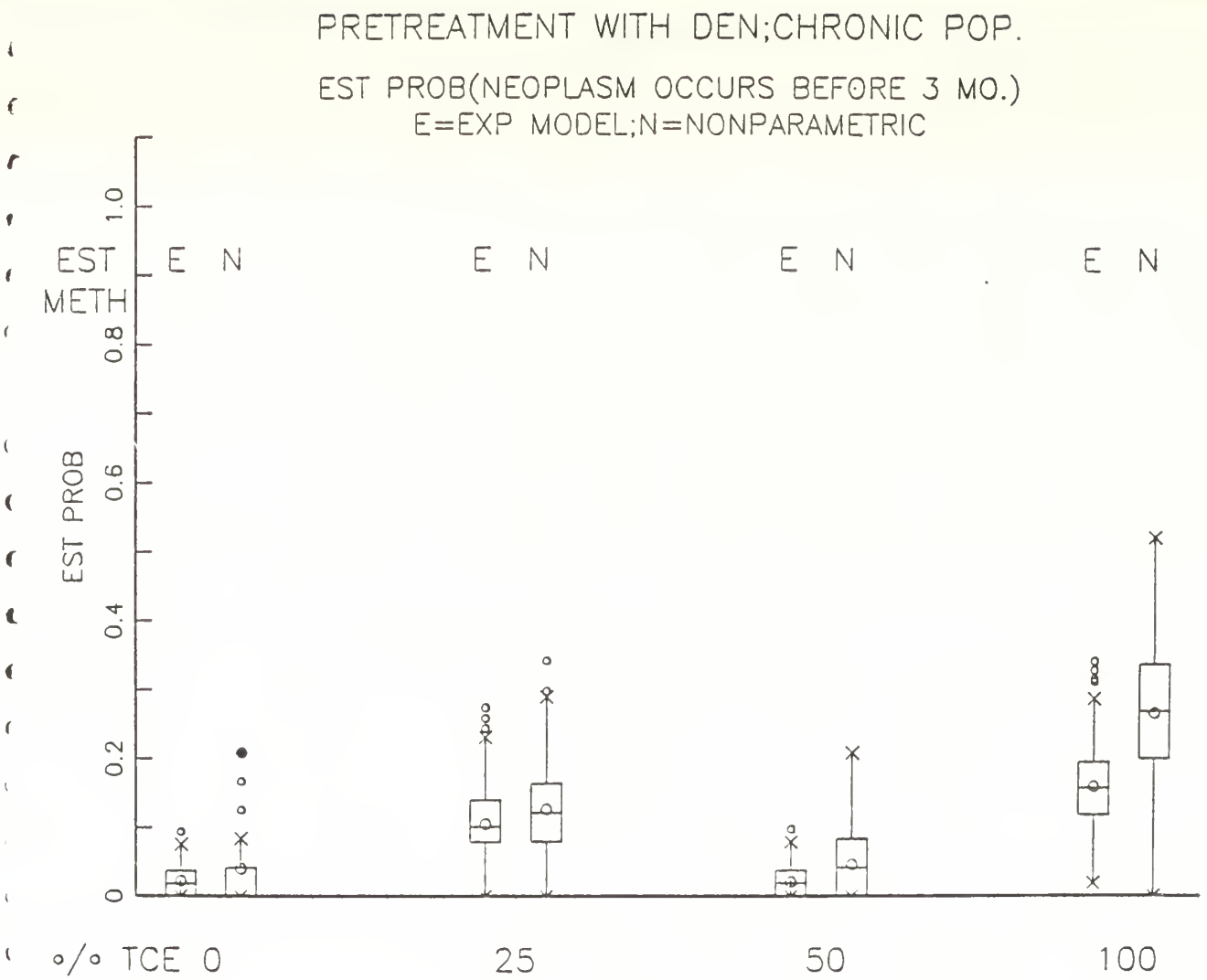
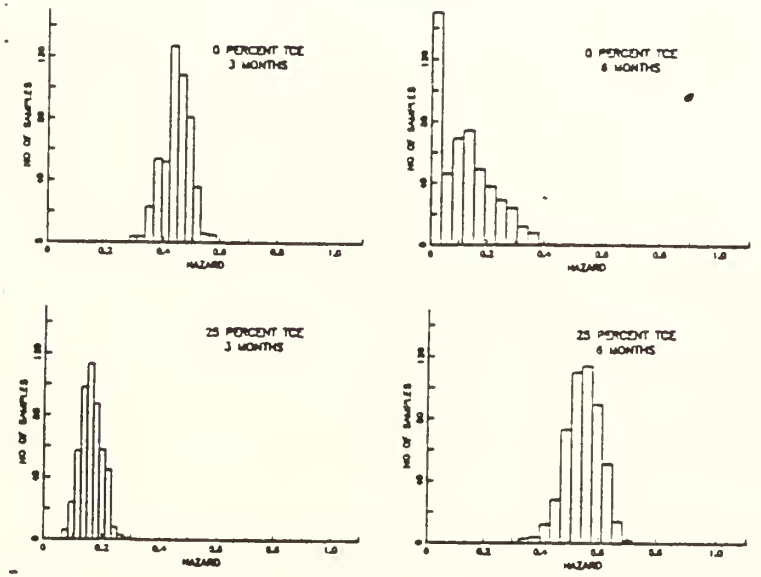
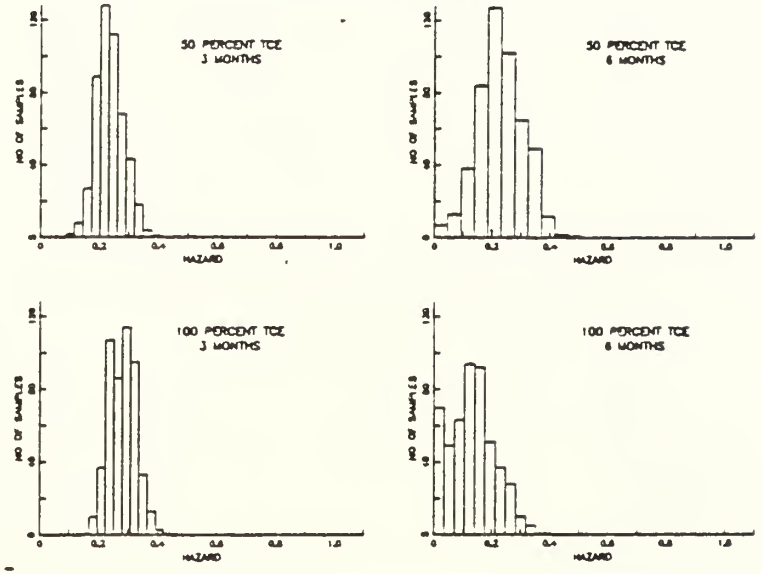

PRETREATMENT WITH DEN;CHRONIC POP.
EST PROB(NEOPLASM OCCURS BEFORE 3 MO.)
E=EXP MODEL;N=NONPARAMETRIC
EST METH
E N E N E N E N
EST PROB
0
0.2
0.4
0.6
0.8
1.0
% TCE 0
25
50
100

Figure 10

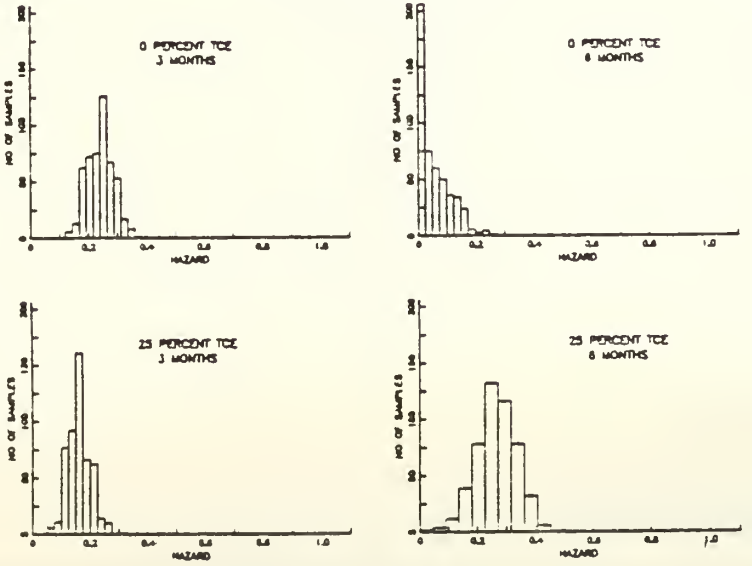
PRETREATED WITH DEN
CYSTIC DEGENERATION



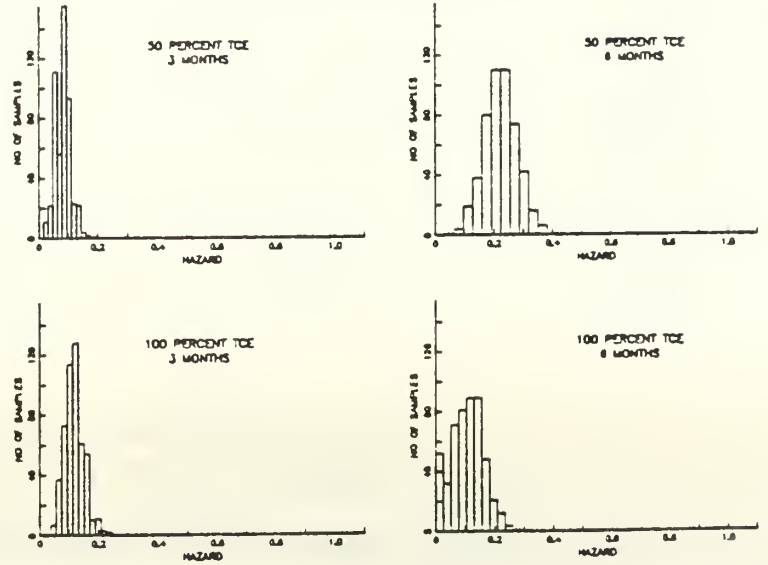
PRETREATED WITH DEN
CYSTIC DEGENERATION



NOT PRETREATED WITH DEN
CYSTIC DEGENERATION



NOT PRETREATED WITH DEN
CYSTIC DEGENERATION



INITIAL DISTRIBUTION LIST

1. Library (Code 52) 2
   Naval Postgraduate School
   Monterey, CA 93943-5000

2. Defense Technical Information Center 2
   Cameron Station
   Alexandria, VA 22314

3. Office of Research Administration (Code 81) 1
   Naval Postgraduate School
   Monterey, CA 93943-5000

4. Prof. Peter Purdue 1
   Code OR-Pd
   Naval Postgraduate School
   Monterey, CA 93943-5000

5. Department of Operations Research (Code 55) 1
   Naval Postgraduate School
   Monterey, CA 93943-5000

6. Prof. Donald Gaver, Code OR-Gv 5
   Naval Postgraduate School
   Monterey, CA 93943-5000

7. Prof. Patricia Jacobs 5
   Code OR/Jc
   Naval Postgraduate School
   Monterey, CA 93943-5000

8. Center for Naval Analyses 1
   4401 Ford Avenue
   Alexandria, VA 22302-0268

9. Dr. David Brillinger 1
   Statistics Department
   University of California
   Berkeley, CA 94720

10. Prof. W. R. Schucany 1
    Dept. of Statistics
    Southern Methodist University
    Dallas, TX 75222

11. Prof. H. Solomon 1
Department of Statistics
Sequoia Hall
Stanford University
Stanford, CA 94305

12. Dr. Ed Wegman 1
George Mason University
Fairfax, VA 22030

13. Dr. Neil Gerr 1
Office of Naval Research
Arlington, VA 22217

14. Dr. J. Abrahams, Code 1111, Room 607 1
Mathematical Sciences Division, Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

15. Prof. D. L. Iglehart 1
Dept. of Operations Research
Stanford University
Stanford, CA 94350

16. Prof. J. B. Kadane 1
Dept. of Statistics
Carnegie-Mellon University
Pittsburgh, PA 15213

17. Prof. J. Lehoczky 1
Department of Statistics
Carnegie-Mellon University
Pittsburgh, PA 15213

18. Prof. M. Mazumdar 1
Dept. of Industrial Engineering
University of Pittsburgh
Pittsburgh, PA 15235

19. Prof. Joseph R. Gani 1
Mathematics Department
University of California
Santa Barbara, CA 93106

20. Prof. Frank Samaniego 1
Statistics Department
University of California
Davis, CA 95616

21. Prof. Tom A. Louis 1
School of Public Health
University of Minnesota
Mayo Bldg. A460
Minneapolis, MN 55455

22. Prof. Brad Carlin 1
School of Public Health
University of Minnesota
Mayo Bldg. A460
Minneapolis, MN 55455

23. Dr. John Orav 1
Biostatistics Department
Harvard School of Public Health
677 Huntington Ave.
Boston, MA 02115

24. Dr. D. C. Hoaglin 1
Department of Statistics
Harvard University
1 Oxford Street
Cambridge, MA 02138

25. Prof. Carl N. Morris 1
Statistics Department
Harvard University
1 Oxford St.
Cambridge, MA 02138

26. Dr. John E. Rolph 1
RAND Corporation
1700 Main St.
Santa Monica, CA 90406

27. Prof. J. W. Tukey 1
Statistics Dept., Fine Hall
Princeton University
Princeton, NJ 08540

28. Dr. D. Pregibon 1
AT&T Bell Telephone Laboratories
Mountain Avenue
Murray Hill, NJ 07974

29. Dr. Jon Kettenring 1
Bellcore
445 South Street
Morris Township, NJ 07962-1910

30. Dr. D. F. Daley 1
Statistic Dept. (I.A.S.)
Australian National University
Canberra, A.C.T. 2606
AUSTRALIA

31. Koh Peng Kong 1
OA Branch, DSO
Ministry of Defense
Blk 29 Middlesex Road
SINGAPORE 1024

32. Professor Sir David Cox 1
Nuffield College
Oxford, OXI INF
ENGLAND

33. Dr. A. J. Lawrence 1
Dept. of Mathematics,
University of Birmingham
P. O. Box 363
Birmingham B15 2TT
ENGLAND

34. Dr. D. Vere-Jones 1
Dept. of Math, Victoria Univ. of Wellington
P. O. Box 196
Wellington
NEW ZEALAND

35. Dr. Robert Carpenter 1
Naval Medical Research Institute, Toxicology Unit
Wright-Patterson Air Force Base,
Dayton, OH 45433