LBL-15735
c.2

Lawrence Berkeley Laboratory
UNIVERSITY OF CALIFORNIA

RECEIVED
LAWRENCE
BERKELEY LABORATORY

APR 15 1983

LIBRARY AND
DOCUMENTS SECTION

STATISTICAL PROCEDURES FOR THE DESIGN AND ANALYSIS OF *in vitro* MUTAGENESIS ASSAYS

John Kaldor
(Ph.D. Thesis)

March 1983

TWO-WEEK LOAN COPY

This is a Library Circulating Copy which may be borrowed for two weeks. For a personal retention copy, call Tech. Info. Division, Ext. 6782.

Donner Laboratory

Biology & Medicine Division

LBL-15735

Prepared for the U.S. Department of Energy under Contract DE-AC03-76SF00098

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

STATISTICAL PROCEDURES FOR THE DESIGN AND ANALYSIS OF *in vitro* MUTAGENESIS ASSAYS

John Kaldor

Lawrence Berkeley Laboratory
University of California
Berkeley, CA 94720

This work was supported in part by the U.S. Department of Energy
under Contract No. DE-AC03-76SF00098,
and in part by the U.S. Environmental Protection Agency
under Contract IAG No. 80-D-X0953

Although the research described in this article has been funded wholly by the United States Environmental Protection Agency (EPA) through IAG No. 80-D-X0953 to the Department of Energy, it has not been subjected to EPA review and therefore does not necessarily reflect the views of EPA and no official endorsement should be inferred.

Contents

		Page
1.	Introduction	1
2.	The Assay Protocol	2
3.	Critical Review of Assumptions Underlying Stochastic Models of Cell Growth and Mutagenesis	4
4.	A Normal Approximation to the Distribution of M	12
5.	Statistical Inference from Mutagenesis Assay Data	13
6.	Simulation Study of Likelihood Procedures for Dose-Response Data	26
7.	Power and Design Considerations	29
8.	Data Analyses	40
9.	Discussion, Conclusions and Areas for Further Research	44
	Appendices	48
	Tables	57
	References	64
	Figures	66

Statistical Procedures for the Design and Analysis of *in vitro* Mutagenesis Assays

John Kaldor
University of California, Berkeley

ABSTRACT

In previous statistical treatments of a certain class of mutagenesis assays, stochastic models of mutation and cell growth have not been utilized. In this paper, we review the assumptions under which these models are derived, introduce some further assumptions, and propose ways to estimate and test hypotheses regarding the parameters of the models from assay data. It is shown via simulation and exact calculation that if the models are valid, the proposed statistical procedures provide very accurate Type I error rates for hypothesis tests, and coverage probabilities for confidence intervals. The cases of a linear dose response relationship for mutagenesis, and a comparison of a set of treated cell cultures with a set of control cultures are treated in detail.

Approximate power functions for hypothesis tests of interest are then derived, and these are also shown to be satisfactorily close to the true power functions. The approximations are used to develop guidelines for planning aspects of a mutagenesis assay, including the number, spacing and range of dose levels employed.

Examples of applications of the procedures are provided, and the paper concludes with a discussion of future statistical work which may be carried out in the area of mutagenesis assays.

1. Introduction

With the increasing use of *in vitro* mutagenesis assays as "short-term" tests for screening potential chemical carcinogens has come the need for associated statistical procedures, to determine whether an assay result is positive or negative, and to estimate some measure of a chemical's potency. Stochastic models for the mutation process have been in existence for some time (e.g. Luria and Delbruck, 1943; Armitage, 1952; Crump and Hoel, 1974). However, they have not been utilized by authors who have proposed estimation and testing procedures for mutagenesis assay data (e.g. Hsie, *et al.*, 1975; Amacher, *et al.*, 1980; Snee and Irr, 1981), with the exception of the treatment by Margolin, *et al.* (1981) of the *Salmonella* (Ames) assay. The major reason has probably been that model development was generally confined to derivations of characteristic functions and moments of the distributions of variables, rather than the distributions themselves. Recently, Tan (1980) elaborated upon previous models to allow for the possibility of cell death, and provided a means of explicitly calculating the appropriate density functions via a recursive formula. Tan (1981) made use of these results to derive the distribution of the number of mutant colonies counted in the CHO/HGPRTase mutagenesis assay.

Although these developments have opened up the possibility of using asymptotic procedures for parameter estimation and hypothesis testing, evaluation of the likelihood is computationally rather time-consuming. Furthermore, the complexity of the distribution does not allow closed-form calculation of such quantities as the information matrix, which is essential for asymptotic power calculations.

In this paper, we introduce a normal approximation to the distribution of the number of mutant colonies, which does not have these problems. We then suggest methods for interval estimation and hypothesis testing of parameters, and examine their validity through Monte Carlo simulation and exact calculation. We also critically review the assumptions under which stochastic models of mutagenesis, including the normal approximation, are derived. Finally, aspects of assay design are considered from the point of view of maximizing the power of specific hypothesis tests.

2. The Assay Protocol

In vitro mutagenesis assays are based on the fact that it is possible to set up experimental conditions in a cell culture under which only cells that have mutations at a specific genetic locus can divide and form colonies. The ability of a chemical to induce mutations in cellular DNA can then be examined by comparing the fraction of cells with a specific mutation in a cell population which has been exposed to the chemical, with the fraction of mutants in a cell population which has not been exposed. A number of different cell types and loci are currently used for mutagenesis testing. The experimental procedures employed vary somewhat among cell lines and loci, and even among laboratories using the same cell type and locus. However, many of the protocols have in common a sequence of essential steps, which are summarized here. The summary includes standard mammalian cell mutagenesis assay protocols, such as those described by O'Neill *et al.* (1977); Clive *et al.* (1979); and Fox (1975). It does not include certain specific mutagenesis assay protocols such as the "Ames test" (Ames *et al.*, 1975), and the fluctuation test (Green *et al.*, 1976), and *in situ* protocols for V79 cells (Huberman and Sachs, 1974).

(i). **Exposure:** A measured quantity of predominantly non-mutant cells (usually $1-6\times10^6$) growing in culture is exposed to the test chemical for a fixed period, usually 4-8 hours, under conditions which allow cellular growth. Cell death may also occur.

(ii). **Expression:** The test chemical is removed and the cells are washed and allowed to grow for a further time period, usually 2-7 days. During this period any mutations which have been induced by the test chemical at the DNA level will presumably be expressed at the phenotypic level, thereby enabling their subsequent detection. New mutations may also arise spontaneously, and there may be residual effects of the chemical treatment on cellular growth, death and mutation rates.

At various stages during the expression period, the cells may be recultured. This involves sampling a fraction of the cells, which is then diluted with growth medium and allowed to continue growing. The remaining cells are discarded. The purpose of this process, which may be carried out several times during expression, is to maintain a cell density at which exponential growth can continue. The sampling fraction is chosen to achieve this density, which is generally specified by the assay protocol. Cell density is measured by a Coulter counter or haemocytometer.

At the end of expression, most of the cell culture is carried forward to the next stage, mutant selection, and a small fraction is employed for viable cell determination.

(iii). **Mutant Selection:** The cells are placed in growth medium on a Petri plate under conditions which allow only phenotypically mutant cells to multiply and form colonies. After an incubation period of a few weeks, these colonies are counted.

(iv). **Viable Cell Count:** This step is carried out concurrently with mutant selection. Its purpose is to estimate the number of viable cells at the end of the expression period; that is, the number of cells in the culture which are capable of initiating colony growth when placed on a Petri dish. A small known fraction of the cell population is sampled and plated under conditions which allow both normal and mutant cells to multiply and form colonies. The incubation period is the same as that used for mutant selection. The total number of viable cells at the end of expression is then estimated by multiplying the number of colonies counted after incubation by the inverse of the fraction sampled.

In a typical assay, these steps are carried out at 3-10 dose levels of test chemical including a zero dose, with possible replication at each dose. The data on which statistical inference from the assay is based are the counts of mutant and total colonies at the end of the mutant selection and viable cell count steps, respectively. The ratio of mutant colonies to the estimated viable cells in the culture sampled for mutant selection is called the *mutant fraction,* and is used as the summary measure of response for each replicate.

3. Critical Review of Assumptions Underlying Stochastic Models of Cell Growth and Mutagenesis

In order to carry out parametric inference using mutagenesis assay data, we require the distributions, or at least the moments, of the observed variables as functions of the unknown parameters. In this section we examine models under which distributions or moments of these variables can be derived, principally from the point of view of the biological and mathematical assumptions which are employed. In so doing, we set the stage for the normal approximation which is introduced in Section 4.

For the purpose of this discussion, it is useful to consider the protocol of a mutagenesis assay as being made up of a series of stages. The first stage is exposure, the next R stages are the reculturing periods during expression (assuming the cells are recultured $R-1$ times) and the $(R+2)^{th}$ and final stage is selection. The viable cell count, which is carried out in parallel with selection, is auxiliary to this sequence and is not defined as a stage. Denote the number of normal, mutant and total cells at time t by N_t, M_t and $T_t = N_t + M_t$ respectively, where t is measured from the start of the assay. We employ a notation similar to that used by Tan (1981) in defining the time interval occupied by the ith stage as (t_i, t'_{i+1}); and denoting N_{t_i}, M_{t_i} and T_{t_i} by N_i, M_i and T_i, respectively; and $N_{t'_i}$, $M_{t'_i}$, and $T_{t'_i}$ by N'_i, M'_i and T'_i. Furthermore, let M and T respectively represent the number of colonies counted on the mutant cell selection and viable cell count plates. Finally, we suppose that a fraction θ_i of the culture is sampled after the end of the ith stage, and carried forward to the $(i+1)^{th}$ stage, for $i = 1,\ldots,R+1$, and that a fraction γ of the culture is sampled for survival determination at the end of the expression period. Figure 1 provides a schematic of the protocol, and indicates where the above variables occur.

3.1. *Models for a Single Stage*

Of principal interest is the joint distribution (or moments) of T and M, the observed variables from a culture which has undergone the assay protocol. However, we initially restrict attention to stochastic models which describe normal and mutant cell growth during a single one of the first $R+1$ stages. We measure t from the beginning of the stage. The most general model for the process $\{(N_t, M_t), t \geqslant 0\}$ is a two-type branching process in which each cell, normal and mutant, faces the following competing risks:

R1. It may divide, thereby "dying" itself, and "giving birth" to two new cells of its own type;

R2. It may divide and give birth to one cell of its own type and one cell of the other type;

R3. It may divide and give birth to two cells of the other type;

R4. It may be transformed into a cell of the other type;

R5. It may die without issue.

We may think of the times to occurrence of each of these risks in the absence of the other four as failure times, with associated distributions. The fate of an individual cell then depends on which of these failure times is the smallest. The risks R2 - R4 result in a mutation; that is, one or two cells of one type arise from the "death" of a cell of the other type. The risks R1 - R3 collectively constitute the risk of division. In the context of mutagenesis, R3 and R4 have not been previously considered, and only Tan (1980) allows for R5.

As it stands, this formulation is too general to allow the derivation of the moments of (N_t, M_t), although asymptotic results are available for large t under fairly weak assumptions (e.g. see Jagers, 1975). Previous authors have adopted various simplifications to enable the exact derivation of these moments, or the probability generating function (p.g.f.) of (N_t, M_t). We follow all of them in assuming

A1. Each cell undergoes the risks R1 - R5 independently of all other cells.

A2. The process is time-homogeneous; that is, the time when the cell is born has no effect on the probabilities of succumbing to R1 - R5.

There is some evidence that division times of cells from a common progenitor are correlated (Kubitschek, 1967). However, without A1 further analytic treatment appears to be impossible. The assumption A2 is less crucial from a mathematical standpoint, and is not generally valid biologically. For example, during exposure, the concentration of a test chemical could decay and alter the probabilities of R1 - R5. During expression, the division time distribution is likely to decrease with time as the cells recover from the treatment and return to normal growth. Also, cell density affects growth rates even if the chemical has no effect.

Some authors (the Markov models of Armitage, 1952; the D/M model of Crump and Hoel, 1974; Tan, 1980) have assumed that the time to occurrence of various of the risks R1 - R5 is independent of the cell's age for mutant cells, where age refers to the time since the cell's birth. This assumption, which in conjunction with A1 and A2, implies that the failure times are exponentially distributed, and hence have a coefficient of variation (c.v.) of unity, may be acceptable for the risks R4 and R5 provided their mechanisms are "single-hit". However, it is almost certainly not valid for cellular division, which is known to be a multi-stage process with c.v. substantially less than one. Unfortunately, the other mathematically tractable alternative, a fixed division time, errs in the other direction, since the c.v. is then zero. The truth would appear to lie somewhere in between the two extremes, with c.v.'s for mammalian cell division times somewhere in the range .1 - .7 (see Jagers and Norrby, 1974). The choice of a distribution for the time to division or death is not really crucial for normal cells. As pointed out by previous authors, their population size is very large and hence is effectively deterministic unless cell killing is extreme. We adopt the assumption

A3. The probability of the risks R1 - R5 occurring is independent of the cell's age.

Under the assumptions A1 - A3, standard techniques can be used to derive differential equations from whose solutions the p.g.f. and moments of (N_t, M_t) conditional on (N_0, M_0) can be straightforwardly obtained (see Karlin and Taylor, 1975). Although the equations which yield the moments can be solved directly, the equations from which the p.g.f. is obtained are Riccati equations which cannot be solved in closed form without further simplification of assumptions.

We now introduce the parameters κ_{kl} for $k = 1,2$ and $l = 1,...,5$, where $\kappa_{kl}dt + o(dt)$ is the probability of the lth risk occurring to a cell of type k (for normal cells $k = 1$; for mutants, $k = 2$) in the time $(t, t + dt)$. Equivalently, $1/\kappa_{kl}$ is the mean of the time to occurrence distribution of the lth risk for a cell of type k, in the absence of the other four risks. We again follow all previous authors, who have explicitly assumed

A4. $\kappa_{22} = \kappa_{23} = \kappa_{24} = 0$;

and implicitly assumed

A5. $\kappa_{13} = \kappa_{14} = 0$.

The first of these assumptions, which states that mutation from mutant to normal cell cannot occur, is biologically reasonable for assays which measure forward mutation. There are generally many more ways for a chemical to interact with cellular DNA in such a way that a mutant is formed from a wild type cell than vice versa. The assumption can also be justified simply on probabilistic grounds by the small number of mutant cells and the low probability of mutation (Tan, 1980). The validity of assumption A5 depends on a number of factors. Its implication is that a phenotypically mutant cell only arises at division, and then in only one daughter cell. For this to be true, the mutation would have to arise only in one strand of the replicated chromosomes in the parent cell, and the mutant allele would have to be recessive. While the latter condition holds for most loci used in mutagenesis assays there is generally insufficient biological evidence to decide whether the former is valid.

Finally, we make the assumption

A6. $\kappa_{11} + \kappa_{12} = \kappa_{21} = \lambda$ and $\kappa_{15} = \kappa_{25} = \mu$,

which is based on the fact that the enzyme product of the locus at which mutation is being measured is generally dispensable *in vitro,* so that normal and mutant cells should be at equal risk of division and death (Abbondandolo, 1977).

Bartlett (1978) describes how the p.g.f. of M_t may be obtained under A1 - A6 with $\mu = 0$. However, this result is of limited use because toxicity is an important factor in mutagenesis assays. Tan (1980) derives the p.g.f. of M_t for large N_0 under the assumptions A1 - A5, and the condition

$$p = p^* N_0^{-1} + o(N_0^{-1}) \tag{3.1}$$

where $p = \kappa_{12}$ and $p^* \geqslant 0$. The condition (3.1) states that the initial number of normal cells is large compared to the probability of mutation in one unit of time, but that their product is non-negligible. It is thus analogous to the condition under which the Poisson approximation to the binomial distribution is obtained. It is in fact a very realistic assumption, as the initial number of normal cells is generally of the order of 10^6, while p is typically in the range $10^{-4} - 10^{-8}$, where time is measured in hours. If A6 with $\mu = 0$ is applied to the p.g.f. derived by Tan (1980), the resulting p.g.f. is equal to that derived by previous authors (Crump and Hoel, 1974; Bartlett, 1978) under the assumption of deterministic growth of normal cells and age-independent division of mutants. However, Tan (1980) provided the very useful development of a recursive formula in k to calculate the probability that $M_t = k$. Table 1 displays the first two moments and the coefficient of skewness of this distribution, as compared to the corresponding "exact" quantities obtained by solving the appropriate differential equations under A1 - A6 (see Appendix 1) for a range of values of N_0, t, λ, μ and p which might occur in practice. The table also displays the mean and c.v. of N_t, and the correlation between N_t and M_t. It is clear that the approximation (3.1) is very good for these parameter values. We also note that the c.v. of N_t is very small, and that N_t and M_t are approximately uncorrelated.

3.2. *Models for the Full Protocol*

Tan (1981) is the only previous author to have applied stochastic models of mutagenesis to a particular assay protocol of the type under consideration here. In order to do this, it is necessary to make an assumption about the relationship between the number of cells of both types present at the end of one stage, and the number present at the beginning of the next. Specifically, he assumes

A7. The number of cells in a fractional sample of a culture is distributed binomially, conditional on the number of cells in the sampled culture.

This assumption relies on uniform mixing of the sampled culture, and exact fractional sampling. Chase and Hoel (1972) discuss a model in which the fraction itself is random. Although this is probably a more accurate representation of reality, we follow A7, which implies in particular that for $i = 2,\ldots,R+1$, N_i and M_i have the binomial distribution conditional on N_i' and M_i', with parameters N_i' and θ_{i-1}, and M_i' and θ_{i-1} respectively, and that T and M are distributed binomially with parameters T_{R+2}' and $\gamma\pi$, and M_{R+2}' and $\theta_{R+1}\pi$ respectively, conditional on T_{R+2}' and M_{R+2}'. Here π is the *cloning efficiency,* or the probability that a plated cell will form a visible colony by the end of the incubation period for selection or viable cell count.

We note at this point that although θ_i and γ are assumed to be known exactly, they can vary among replicate cultures. Under A1 - A7, this variation is partly due to randomness in total cell density, but mostly due to the randomness associated with *measuring* the cell density. (Recall that θ_i is chosen to achieve a specified cell density; γ is usually chosen to achieve a specified number of cells.)

Tan (1981) utilized his earlier assumptions and results (Tan, 1980) to derive the p.g.f. and density function of M conditional on M_1 and N_i, $i = 1,\ldots,R+1$ by making use of the above-mentioned conditional relationship between M_i' and M_i, for $i = 2,\ldots,R+1$ and between M_{R+2}' and M. The condition (3.1) was applied to all stages, through the assumption that

$$p_i = p^*_i N_i^{-1} + o(N_i^{-1}),\ i = 1,\ldots,R+1 \tag{3.2}$$

where $p_i \geqslant 0$, and N_i is large for $i = 1,\ldots,R+1$.

It is possible to obtain the moments of M and T conditional on (N_1, M_1) without relying on (3.2). Under the assumptions A1 - A7, we can derive difference equations in i for the moments of (N_i, M_i), $i = 2,\ldots,R+1$, which can be easily solved, and use the solutions to obtain the moments of (N_t, M_t) for any $t \geqslant 0$ (see Appendix 1), and for $t = t'_{R+2}$ in particular. Then by making use of the fact that conditional on M'_{R+2} and N'_{R+2}, M and T have binomial distributions, their moments conditional on (N_1, M_1) can be derived.

For most protocols and parameter values which arise in practice, the characteristics of the distribution of (N'_{R+2}, M'_{R+2}) are very similar to those of the distribution of (N_t, M_t) evidenced by Table 1. In particular, the moments of M'_{R+2} obtained under A1 - A7 described above are very close to those derived under the approximation (3.2); the c.v. of N'_{R+2} is very small; and M'_{R+2} and N'_{R+2} are approximately uncorrelated. Therefore, since $M_{R+2} << N_{R+2}$, most of the variability in T arises from the fact that γ, the fraction of cells sampled from T'_{R+2} for viable cell count, is very small. Hence, the distribution of T conditional on (N_1, M_1) is very close to its distribution conditional on T'_{R+2}; namely, binomial with parameters T'_{R+2} and $\gamma\pi$, where T'_{R+2} is large and $\gamma\pi$ is small. We formalize these observations in the approximation

$$T \text{ is distributed as Poisson with mean } \gamma\pi E(T'_{R+2}), \text{ and is independent of } M. \tag{3.3}$$

In fact, since the number of mutants in the culture up to time t'_{R+2} is always very small compared with the number of normal cells, which is large, we could employ the more general approximation that $T_t = N_t$, and that $Var(N_t) = 0$, for $0 \leqslant t \leqslant t'_{R+2}$.

Under A1 - A7,

$$E(T'_{R+2}) = (N_1 + M_1)\left[\prod_{i=1}^{R} \theta_i\right] e^{\sum_{i=1}^{R+1} d_i r_i},$$

where $d_i = \lambda_i - \mu_i$, and $r_i = t'_{i+1} - t_i$.

Let λ_i, μ_i and p_i be respectively the division, death and mutation parameters for the ith stage, for $i = 1,\ldots,R+1$. Under the assumptions A1 - A7, the approximations (3.2) and (3.3), and conditional on M_1 and N_i, $i = 1,\ldots,R+1$

$$E(M) = \left(\prod_{i=1}^{R+1} \theta_i^*\right) e^{\sum_{i=1}^{R+1} d_i r_i} \cdot \left\{N_1 \sum_{i=1}^{R+1} p_i r_i + M_1\right\} \tag{3.4}$$

$$Var(M) = \sum_{i=1}^{R+1} \left(\prod_{m=i+1}^{R+1} \theta_m^* e^{d_m r_m}\right)^2 e^{d_i r_i} \theta_i^* N_i$$

$$\cdot \left\{\left[\bar{\theta}_i^* + \theta_i^* (\lambda_i + \mu_i)\ (e^{d_i r_i} - 1) \cdot \frac{1}{d_i}\right] \left[\sum_{m=1}^{i-1} p_m r_m + \frac{M_1}{N_1}\right]\right.$$

$$\left. + \left[\bar{\theta}_i^* r_i + \theta_i^* \left\{\frac{2\lambda_i}{d_i} (e^{d_i r_i} - 1) - (\lambda_i + \mu_i)\ r_i\right\} \frac{1}{d_i}\right] p_i\right\}, \tag{3.5}$$

where $\theta_i^* = \theta_i$ for $i = 1,\ldots,R$ and $\theta_{R+1}^* = \theta_{R+1}\pi$; and $\bar{\theta}_i^* = 1 - \theta_i^*$. We define $\sum_{m=1}^{0} p_m r_m = 0$, and $\prod_{m=R+2}^{R+1} \theta_m^* e^{d_m r_m} = 1$.

3.3. *Additional Sources of Variability*

In developing the distribution functions to be used as a basis for inference, we have only incorporated variation arising through the random processes of cell division, death, and mutation, and the sampling of cultures between stages. A number of previous authors (Grafe and Vollmar, 1979; Margolin, *et al.*, 1981) have noted in the context of the *Salmonella* (Ames) mutagenesis assay, that the variability of counts is greater than that of the Poisson distribution which would be predicted on comparable theoretical grounds. Margolin, *et al.* (1981) have suggested that this increased variability is due to the inability of experimentalists to "maintain near constancy across replicates of the [cells'] environment." For the protocols considered

above, this phenomenon would be manifested by variation in such quantities as the initial number of exposed cells, amount of nutrients supplied to the cells and dose of test chemical, within dose levels. These quantities, which have been implicitly regarded as constant, would then be stochastic. In order to investigate the importance of these additional sources of variability, we require data from experiments in which the assay protocol was applied to a large number of replicate cell cultures at a fixed dose of test chemical. Unfortunately, the availability of data of this kind for the protocols discussed here is at present very limited.

We have also not discussed *phenotypic delay* in this section, although its incorporation into models of mutagenesis has been a major concern of previous authors. Briefly, phenotypic delay is the phenomenon whereby there is a lag between the time at which a mutation occurs at the level of the gene, and the time at which the mutant cell can actually be identified. This occurs, for example, if selection of the mutant requires a much reduced level of the enzyme product of the normal gene, and a period of time after mutation is necessary for residual normal enzyme to be diluted out of the cell. We have ignored phenotypic delay as a factor in the derivation of the moments and distribution of M because the expression period is generally long enough to allow phenotypic expression of all mutations induced by the test chemical; only the relatively few spontaneous mutants which arise during the latter stages of expression may not be sufficiently expressed to permit their selection.

In what follows, we adopt the assumptions A1 - A7 and the approximations (3.2) and (3.3), except where specified.

4. A Normal Approximation to the Distribution of M

As discussed in Section 3, the density function derived by Tan (1981) should provide a very good approximation for the density of M under A1 - A6 over the range of mutation parameter values which occurs in practice. However, because of its analytic and computational complexity, it is natural to seek alternative approximations for use as a basis for statistical inference regarding the unknown parameters of the model. One strategy is to employ a simpler

distribution, which agrees with the distribution of the number of mutant colonies up to a specified number of moments. The normal distribution suggests itself as a possibility, by the following argument:

Let X^k denote the number of mutant colonies arising by time t'_{R+3} (the end of selection) from the kth normal cell at the start of exposure. Then if we make the assumption

A8. $M_1 = 0$,

we may express M as $M = \sum_{k=1}^{N_1} X^k$. Since the random variables X^k are independently and identically distributed with finite mean and variance (see Appendix 1) straightforward application of the Central Limit Theorem suggests that M is approximately normally distributed for large N_1. This approximation is analogous to the normal approximation to the binomial distribution, just as the approximations (3.1) and (3.2) used by Tan (1980; 1981) are analogous to the Poisson approximation. We do not claim that the normal approximation is as accurate as that obtained under the Tan approximation; the true distribution is obviously skewed and only defined over the positive integers, properties it shares with the latter approximation. However, the normal approximation may be quite adequate for larger $E[M]$, and it certainly offers the advantage of analytical and computational convenience.

5. Statistical Inference from Mutagenesis Assay Data

In Sections 3 and 4, we discussed the density derived by Tan (1981) for M, the number of mutant colonies counted at the end of selection; the normal approximation to this density; and an approximation to the distribution of T, the number of colonies counted at the end of the viable cell determination. We now consider procedures for estimation and testing of the parameters of the model using dose-response data from an assay, when the assumptions A1 - A8, and the approximations (3.2) and (3.3) hold. Although the procedures will be generally applicable to protocols which conform to the outline in Section 2, we will base numerical simulations and computations on a specific protocol (Clive, *et al.*, 1979) for measuring chemically induced

mutation at the $TK^{+/-}$ locus in mouse lymphoma cells. The values of relevant protocol variables are:

$$R = 2; \tag{5.1a}$$

$$r_1 = 4 \text{ hours}, \; r_2 = r_3 = 24 \text{ hours}; \tag{5.1b}$$

$$T_1 = 6\times10^6, \; T_2 = T_2', \; T_3 = \min(T_3', 3\times10^6), \; T_4 = \min(T_4', 3\times10^6); \tag{5.1c}$$

600 cells are sampled for survival determination. (5.1d)

Suppose C dose levels of a chemical plus a control are tested in a mutagenesis assay, with D_c replicates at the cth dose level. Let $M^{c,d}$ and $T^{c,d}$ respectively denote the number of mutant colonies at the end of selection and the number of colonies at the end of viable cell count, in the dth replicate at the cth dose. We indicate the control by $c = 0$, and let x_c denote the dose of test chemical at the cth level. Under the assumptions A1 - A8, there are $3(R+1) + 1$ parameters unknown at each dose level, namely, λ_i^c, μ_i^c, and p_i^c for $i = 1,...,R+1$ and the cloning efficiency π^c, where the superscripts reflect the possible dependence of the parameters on the dose of chemical at the cth dose level. We will sometimes, where appropriate, write $\lambda_i(x)$, $\mu_i(x)$, $p_i(x)$ and $\pi(x)$ to express the functional dependence on dose x for *all* $x \geqslant 0$. Since in most assays $D_c \leqslant 3$, further assumptions are required before we can carry out estimation and testing of the parameters of the model using only the data $(M^{c,d}, T^{c,d})$, $c = 0,1,...,C$ and $d = 1,...,D_c$. Accordingly, we introduce the following simplifications. in Section 5.3 we discuss these assumptions, and indicate ways they may be weakened if various auxiliary measurements are utilized.

A9. $\lambda_i^c = \lambda_1^0$, $\mu_i^c = \mu_1^0$ and $p_i^c = p_1^0$ for $i = 2,...,R+1$ and all c, indicating that "control" rates of division, death and mutation return after exposure at all doses.

A10. The division rate and cloning efficiency are unaffected by the test chemical, so that $\lambda_1^c = \lambda$ and $\pi^c = \pi$ for all c.

A11. $\mu_1^0 = 0$, so that no cell death occurs during exposure in the controls.

These assumptions reduce the number of unknown parameters to λ, π, p_1^c for $c = 0,1,\ldots,C$ and μ_1^c for $c = 1,\ldots,C$, representing one unknown parameter for the control dose, two for the other doses, and growth and cloning efficiency parameters which apply to all doses. In illustrative computations and simulations involving the protocol specified by (5.1), we will use the values

$$\lambda = .0578 \tag{5.1e}$$

$$\pi = 1 \tag{5.1f}$$

The value of λ was chosen to correspond to the normal mouse lymphoma cellular division time of about 12 hours. The choice of π represents a perfect cloning efficiency which would not apply in practice. However, for the purpose of such calculations, we can consider π, which is generally in the range .5 - .95, as being absorbed into the fraction sampled for mutant selection and viable cell count (i.e. θ_{R+1} and γ respectively).

Since we would expect that increased dose of test chemical could only increase the risk of mutation and cell death during exposure, there is a natural ordering within the mutation and toxicity parameters; namely

$$p_1^0 \leqslant p_1^1 \leqslant \cdots \leqslant p_1^C \tag{5.2}$$

$$\mu_1^0 \leqslant \mu_1^1 \leqslant \cdots \leqslant \mu_1^C \tag{5.3}$$

Of principal interest is p_1^c / λ, the proportion of mutants produced per cell division during the period of exposure to the test chemical, and the relationship between this proportion and the dose of test chemical. The other unknown parameters are in effect nuisance parameters. We are typically concerned with one or more of the following objectives:

(i) Estimate and provide confidence intervals for p_1^c / λ, for $c = 0,1,\ldots C$.

(ii) Test the hypothesis H_{0c}: $p_1^c = p_1^0$ against the one-sided alternative H_{1c}: $p_1^c > p_1^0$, for $c = 1,\ldots,C$.

(iii) Test the hypothesis H_E: $p_1^0 = p_1^1 = \cdots = p_1^C$ of equal mutation rates against the alternative of increasing trend in rates, namely H_T: $p_1^0 \leqslant p_1^1 \leqslant \cdots \leqslant p_1^C$, with at least one inequality strict. This generalizes (ii) above.

(iv) Estimate, and provide confidence intervals for, the parameters of a functional relationship between x, the dose of test chemical at the cth level, and $p_1(x) / \lambda$. For example, if the chemical's mutagenic activity followed "single-hit" kinetics, we would be interested in the parameters α and β in the model

$$p_1(x) / \lambda = 1 - \exp(-\alpha - \beta x) \approx \alpha + \beta x \tag{5.4}$$

when p_1 is small, as is typically the case.

(v) Test hypotheses concerning the parameters of a functional relationship between x and p_1 / λ. For example, in (5.4) we we may wish to test H_0: $\beta = 0$ against the one-sided alternative H_1: $\beta > 0$.

5.1. *Inference When Only the Mutation Parameters are Unknown*

First consider estimation and hypothesis tests for the mutation parameters $\{p_1^c,\ c = 0,1,\ldots,C\}$ when λ, and the nuisance parameters π and $\{\mu_1^c,\ c = 1,\ldots,C\}$ are known. Since the distribution of $T^{c,d}$ is independent of the mutation parameters (see (3.3)) we can only utilize the mutant colony counts $M^{c,d}$. Of the above objectives, (i), (ii), (iv) and (v) can be accomplished using standard likelihood methods, based on either the iteratively-defined density derived by Tan (1981) for $M^{c,d}$ under the assumptions A1 - A11, or its normal approximation, as discussed in Section 4. These methods include estimation of parameters by maximizing the joint density of $M^{c,d}$, $c = 0,1,\ldots,C$ and $d = 1,\ldots,D_c$ with respect to the unknown parameters; and estimating confidence intervals and testing hypotheses either by assuming that the maximum likelihood

estimates of parameters are normally distributed, or by assuming that the likelihood ratio has a central χ^2 distribution under the null hypothesis in question.

It may be important to make explicit the conditions under which the "standard likelihood methods" referred to above are valid. We can either consider the limit as C, the number of dose levels, becomes large, or the limit for large D_c, the number of replicates per dose. In Appendix 2, we prove general asymptotic results, and show that particular cases of interest satisfy the assumptions. However, because typical assays rarely test more than 2 replicates at 10 doses, it is more important to know how well the large-sample approximations perform in what are in fact rather small samples. The exact calculations in this section and the Monte Carlo simulations carried out in Section 6 are designed to answer this question for some of these large-sample methods.

In order to maximize the likelihood over unknown parameters, some non-linear functional optimization procedure is required. We employ Newton-Raphson iteration, with starting values as in Appendix 4. The complex form of the Tan (1981) density requires special treatment to obtain the required first and second derivatives of the likelihood with respect to the unknown parameters. Moreover, evaluation of the probability of the observations for each new parameter value in the iterative process requires a re-application of the iterative formula for the density.

Estimates of the mutation parameters which are of a simpler form than the maximum likelihood estimates may be obtained by making use of the fact that under the additional assumptions A9 - A11, (3.4) reduces to

$$E(M^{c,d}) = \pi \left(\prod_{i=1}^{R+1} \theta_i^{c,d} \right) e^{\lambda \sum_{i=1}^{R+1} r_i - \mu_1^c r_1} N_1 \cdot \left\{ p_1^c \; r_1 + p_1^0 \sum_{i=2}^{R+1} r_i \right\}$$

$$= K^{c,d} \left\{ p_1^c \; r_1 + p_1^0 \sum_{i=2}^{R+1} r_i \right\},$$

say, where N_1, the number of cells initially exposed; $\theta_i^{c,d}$, the fraction sampled after the ith stage in the dth replicate at the cth dose; and r_i, the length of the ith stage are all known; we note that $K_{c,d}/\pi$ is in fact the expected number of normal cells in the culture at the start of selection. Then

$$\hat{p}_1^{0,d} = M^{0,d} \Big/ \left(K^{0,d} \sum_{i=1}^{R+1} r_i \right) \tag{5.5}$$

is an unbiased estimate of p_1^0 based on the dth observation of the control. An estimate based on all D_0 control observations can be obtained by simply averaging. The resulting estimate may then be used in

$$\hat{p}_1^{c,d} = \left\{ M^{c,d} \Big/ K^{c,d} - \hat{p}_1^0 \sum_{i=2}^{R+1} r_i \right\} \Big/ r_1, \tag{5.6}$$

an estimate of p_1^c based on the dth observation at the cth dose. Once again the average can be taken to provide an estimate based on all observations at the cth dose.

These method-of-moments type estimators for p_1^c also provide the basis for a simple means of carrying out objective (ii) above. We consider testing the hypothesis H_{0c}: $p_1^c = p_1^0 = p$, say, against the alternative H_{1c}: $p_1^c > p_1^0$. Let

$$\hat{p}^{c,d} = M^{c,d} \Big/ \left(K^{c,d} \sum_{i=1}^{R+1} r_i \right)$$

Then, under H_{0c}, A1 - A12, and the approximations (3.2) and (3.3), $\hat{p}^{c,d}$ is unbiased for p, with variance

$$Var(\hat{p}^{c,d}) = \frac{1}{\left(K^{c,d} \sum_{i=1}^{R+1} r_i \right)^2} Var(M^{c,d}),$$

where $Var(M^{c,d})$ is given by (3.5). Therefore, under H_{0c},

$$\Gamma = \frac{\frac{1}{D_c}\sum_{d=0}^{D_c}\hat{p}^{c,d} - \frac{1}{D_0}\sum_{d=0}^{D_0}\hat{p}^{0,d}}{\left\{\frac{1}{D_c^2}\sum_{d=0}^{D_c} Var(\hat{p}^{c,d}) + \frac{1}{D_0^2}\sum_{d=0}^{D_0} Var(\hat{p}^{0,d})\right\}^{1/2}} \tag{5.7}$$

is approximately a standard normal variate, for large D_c and D_0. The hypothesis H_{0c} is rejected for $\Gamma > Z_{(1-\nu)}$, where Z_Q is the value of a $N(0,1)$ random variable below which a fraction Q of the distribution lies, and ν is the nominal significance level of the test.

Of course, the variance expressions in the denominator of (5.7) involve the unknown p as well as the other parameters which we are at present assuming known. However, under H_{0c},

$$\hat{p} = \frac{1}{D_0+D_c}\left\{\sum_{d=1}^{D_0}\hat{p}^{0,d} + \sum_{d=1}^{D_c}\hat{p}^{c,d}\right\}$$

is an unbiased estimate of p, which utilizes all the observations at the cth and control exposure levels. This estimate is also consistent for large D_0 and D_c, since $Var(M^{c,d})$ is finite. Therefore, its use in place of p where required in (5.7) does not invalidate the asymptotic normality of Γ , and the test can thus be used as a simpler alternative to the likelihood ratio test for H_{0c}. Table 2A gives exact Type I error rates for both tests under the protocol of Clive, *et al.* (1979) (see (5.1)) for the case where $D_0 = D_c = 1$, for a number of values of the toxicity parameters μ_1^c, and 2 values of the common p. The likelihood ratio test statistic Λ is calculated using the normal approximation to the density of M. In the table, we use $S_\mu = e^{-\mu_1^c r_1}$ as the measure of toxicity. This is the number of cells which survive exposure in the treated culture, as a fraction of control, since by A10, $N_1 e^{\lambda r_1}$ cells survive exposure in the control culture, while $N_1 e^{(\lambda-\mu_1^c) r_1}$ survive in the treated culture. S_μ, the *surviving fraction,* is used by biologists as the measure of cell killing. The exact Type I error rates were obtained by employing the distributions derived by Tan (1981) to evaluate the probability of the outcomes which result in rejection of H_{0c} using one or the other of the two tests. It is clear that in both cases, the true Type

I error rates are very close to the nominal ones of the tests, even though the sample size in each case is only 2.

This argument can be extended, with some modifications, to provide a means of carrying out the objective (iii) above using the $\bar{\chi}^2$ isotonic regression test (Barlow, *et al.*, 1972). The ordering (5.2) for the parameters p_i^c is equivalent to the ordering

$$q^0 \leqslant q^1 \leqslant \cdots \leqslant q^C,$$

where $q^c = E\,(M^{c,d})\,/\,K^{c,d} = p_1^c\, r_1 + p_1^0 \sum_{k=2}^{R+1} r_k$. If $\theta_i^{c,d}$ does not depend on d for fixed i and c, so that $K^{c,d} = K^c$, say, the variables $M^{c,d}/K^c$ satisfy the requirements of a theorem stated in Appendix 3, which is a simple generalization of a theorem in the Appendix of Collings, *et al.* (1981). If we let

$$\bar{q} = \frac{1}{D} \sum_{c=0}^{C} \frac{1}{K^c} \sum_{d=1}^{D_c} M^{c,d}$$

where $D = \sum_{c=0}^{D} D_c$, and $\dot{q}^c$ be the isotonic regression estimate of

$$\frac{1}{D_c\, K^c} \sum_{d=1}^{D_c} M^{c,d}$$

with respect to weights $D_c(K^c)^2\,/\,Var(M^{c,d})$, then

$$\Theta = \sum_{c=0}^{C} D_c(K^c)^2\,(\dot{q}^c - \bar{q})^2\,/\;Var\;(\hat{M}^{c,d})$$

has asymptotically the $\bar{\chi}^2$ distribution on $C+1$ degrees of freedom, where $Var\;(M^{c,d})$ is obtained from (3.5), and under H_E, its estimate $Var\;(\hat{M}^{c,d})$ is the same expression with

$$\bar{p} = \bar{q}\,/\left(\sum_{k=1}^{R+1} r_k\right)$$

in place of p_i^c.

The isotonic regression test, while being intuitively appealing, has some distinct drawbacks. First, the isotonic regression used in the test statistic has to be calculated by an iterative algorithm such as the Pool-the-Adjacent-Violators procedure (see Barlow, *et al.*, 1972). More importantly, the critical values of the statistic, which is asymptotically a mixture of χ^2 random variables under H_E, are only available for $C+1 \leqslant 6$, in the unequal weights case which (possibly) arises here even under equal sample sizes if μ_1, the toxicity, increases with dose. Collings, *et al.* (1981) have recently shown that when the outcome variables are binomially distributed, the isotonic regression test has no better power to detect increasing trend alternatives than the simpler Cochran-Armitage trend test. It is possible that a statistic analogous to the Cochran-Armitage statistic could be constructed to test H_E; it will be a subject of further research to compare the isotonic regression test with a test based on such a statistic. In Section 6, we introduce a likelihood-ratio statistic for testing that the slope of a linear relationship between x, the dose of test chemical, and $p_1(x)$ is equal to zero. This test could also be compared with the two trend tests mentioned above, in terms of power to detect specific increasing trend alternatives.

5.2. *Inference When the Mutation and Toxicity Parameters are Unknown*

The assumption that the division rate and cloning efficiency are known is reasonable for most varieties of cells used in mutagenesis assays, provided A10 is valid. However, it is more likely that the toxicity parameters μ_1^c are unknown, although some chemicals have well established toxicity curves in specific cell lines. We now consider procedures for estimating μ_1^c.

The distributions of both $M^{c,d}$ and $T^{c,d}$ both depend on μ_1^c, so that it would be theoretically possible to estimate the toxicity parameters simultaneously with the mutation parameters, by maximizing the likelihood based on the joint density of $(M^{c,d}, T^{c,d})$, for $c = 0,1,\ldots,C$ and $d = 1,\ldots,D_c$ over $\{\mu_1^c,\ c = 1,\ldots,C\}$ and $\{p_1^c,\ c = 0,1,\ldots,C\}$. However, we take a more pragmatic approach, and employ maximum likelihood estimation based only on the viable cell counts $T^{c,d}$. This approach is justifiable on a number of grounds. First the mutation counts $M^{c,d}$, $d = 1,\ldots,D$ would be expected to contain much less information about the parameter μ_1^c than

the survival counts. Second, the approach is consistent with assay protocol, in that the viable cell counts are made specifically to provide information about cellular growth. Finally, it provides a computationally simpler estimation algorithm, since the toxicity parameters need not be re-estimated for different hypotheses concerning the mutation parameters.

Now under the approximation (3.3), $T^{c,d}$ is distributed as Poisson with mean $J^{c,d}\gamma^{c,d}$, where $J^{c,d} = K^{c,d} / \theta_{R+1}^{c,d}$ and $\gamma^{c,d}$ is the fraction sampled for viable cell count in the dth replicate at the cth dose. Therefore the maximum likelihood estimate (m.l.e.) of μ_1^c based on the viable cell counts at the cth dose only is

$$\hat{\mu}_1^c = -\frac{1}{r_1} \log \left[\frac{\sum_{d=1}^{D_c} T^{c,d}}{N_1 \pi \sum_{d=1}^{D_c} \left(\gamma^{c,d} \prod_{i=1}^{R} \theta_i^{c,d} \right)} \cdot e^{-\lambda \sum_{i=1}^{R+1} r_i} \right] \tag{5.8}$$

The control counts $T^{0,d}$, for $d = 1,\ldots,D$ can be used in a similar way to estimate λ, if it is in fact unknown. An estimate of μ_1^c with better characteristics can possibly be obtained by utilizing the constraint (5.3) in the maximization of the likelihood, although the extra computational effort may not be justified.

The methods discussed above for estimation and hypothesis testing of the mutation parameters when $\{\mu_1^c,\ c = 1,\ldots,C\}$ are known can be used with estimates in place of the true values. The large sample approximations should still be valid, because the mutant counts $M^{c,d}$ contain negligible information about the toxicity parameters compared to the viable cell counts $T^{c,d}$. Table 2B gives estimated Type I error rates for the test of H_{0c} using Γ, (see (5.7)), and the likelihood ratio test statistic Λ, when the level of toxicity during exposure in the treated culture is estimated as in (5.8), and $D_0 = D_c = 1$. The error rates were estimated using 5,000 simulated "experiments" with $S_\mu = .1$ and $p = 7.5\times10^{-7}$, and the assay protocol (5.1). The density derived by Tan (1981) was used to simulate values of $M^{c,d}$ and the Poisson density with parameters as in (3.3) provided values of $T^{c,d}$. Once again, both statistics give very satisfactory significance levels in this "worst case" situation of $D_0 = D_c = 1$.

5.3. *Utilization of Additional Measurements*

The Assumptions A9 - A11 are oversimplifications. Like assumptions A2 and A4 - A6, they are made for reasons of parameter identifiability rather than mathematical tractability. A9 is probably the most inaccurate, since cells which have been exposed to high doses of a toxic chemical usually require some time to resume normal growth, even after the chemical has been removed by washing. Also, a certain proportion of exposed cells are likely to be at a higher risk of death or mutation than the cells which were unexposed, even if they survive the exposure period as "normal" cells. The assumption of constant cloning efficiency in A10 may be more reasonable: CHO cells have been observed to return to 75-80% of normal cloning efficiency after a 5-day expression period (L. Thompson, personal communication). The assumption A10 is also not unreasonable. Although there is some cell death taking place in all normal cell populations, it is probably negligible compared to the effect of a toxic test chemical.

The methods of estimation and hypothesis testing for the mutation parameters which we have described employ only the final colony counts $M^{c,d}$ and $T^{c,d}$, $c = 0,1,\ldots,C$; $d = 1,\ldots,D_c$. A number of other measurements are made during the course of an assay, and their utilization in statistical inference may allow the assumptions A9 - A11 to be somewhat weakened.

First, the cell density measurements made at the end of the expression stages (i.e. at times t_i', $i = 3,\ldots,R+2$) can be used to estimate $d_i^c = \lambda_i^c - \mu_i^c$, for $i = 3,\ldots,R+1$ and $c = 0,1,\ldots,C$. These measurements are typically made by sampling a small fraction of the cell culture, and counting the number of cells in the sample mechanically, or by eye. Denote the fraction sampled in this manner at the end of the ith stage by ρ_i, and let $U_i^{c,d}$ be the number of cells counted at the end of the ith stage in the dth replicate at the cth dose. Then by the same argument which led to (3.3), in which $T^{c,d}$ was shown to be approximately Poisson distributed, we may also suppose that $U_i^{c,d}$ is approximately Poisson distributed, with mean

$$E\left(U_i^{c,d}\right) = \rho_i\, E\left((T_{i+1}')^{c,d}\right) = \rho_i E\left(T_i^{c,d}\right) e^{d_i^c r_i} = \rho_i \theta_{i-1}^{c,d}\, E\left((T_i')^{c,d}\right)\, e^{d_i^c r_i},$$

if the fraction ρ_i is small enough so that

$$(T'_{i+1})^{c,d} / \left[(T'_{i+1})^{c,d} - U_i^{c,d}\right] \approx 1.$$

Then we have the m.l.e.

$$\hat{d}_i^c = \frac{1}{r_i} \log \left[\frac{\sum_{d=1}^{D_c} U_i^{c,d}}{(\rho_i / \rho_{i-1}) \sum_{d=1}^{D_c} \theta_{i-1}^{c,d} U_{i-1}^{c,d}} \right] \tag{5.9}$$

of d_i^c, for $i = 3,\ldots,R+1$. Note that typically $\theta_i^{c,d}$ is chosen to provide some specified density for $T_{i+1}^{c,d}$, based on the density count $U_i^{c,d}$, thus, we may expect that $\theta_i^{c,d} U_i^{c,d} = T_i^c$, say, independent of d, unless cell killing is high and the required density cannot be attained without reducing the volume of culture.

If a cell density count is made at the end of the exposure period, d_1^c and d_2^c can be estimated similarly (in the estimate of d_1^c, replace $\sum_{d=1}^{D_c} \theta_{i-1}^{c,d} U_{i-1}$ by $D_c\ T_1$). If such a count is not made, as is the case in the protocol of Clive, *et al.* (1979), we can estimate d_2^c by $\hat{d}_3^c$, and d_1^c by formula (5.8), replacing the exponent by $-\sum_{i=2}^{R+1} \hat{d}_i^c\, r_i$, and removing the leading minus sign.

An estimate of the cloning efficiency may similarly be obtained at each dose by using

$$\hat{\pi}^c = \max\left[1, \frac{\sum_{d=1}^{D_c} T^{c,d}}{(1/\rho_{R+1}) \sum_{d=1}^{D_c} U_{R+1}^{c,d}\ \gamma^{c,d}}\right] \tag{5.10}$$

The measurements $U_i^{c,d}$, like $T^{c,d}$, provide no information about the mutation parameters.

We are now in a position to replace A9 - A11 by

A9* $\lambda_i^c = \lambda_1^0$ and $p_i^c = p_1^0$ for $i = 2,\ldots,R+1$

A10* $\lambda_1^c = \lambda$

for all c. Then $\hat{\mu}_1^c = \lambda - \hat{d}_i^c$ may be used in the variance expression (3.5) for the application of the statistical procedures described in Section 5.1. If $\hat{d}_i^c > \lambda$, we define $\hat{\mu}_1^c = 0$. We may also wish to use the ordering (5.3) to constrain the estimates. It may be noted that we still have not been able to dispense with the assumption that the chemical affects only the cell death rate and not the growth rate. This assumption is still required to overcome the non-identifiability which would arise if both λ_i^c and μ_i^c were allowed to be dose-dependent, and the estimation was to be based on (5.9).

The additional assumption

A11* $\mu_i^0 = 0,\ i = 1,\ldots,R+1$

allows λ to be estimated from the counts $U_i^{0,d}$ and $T^{0,d}$.

Use of these estimates in place of the true values is analogous to the use of estimates of μ_1^c as discussed in Section 5.2. We would thus expect the inference procedures developed in Section 5.1 to be valid with estimates of λ, π^c and μ_i^c in place of the true values, just as estimation of μ_1^c did not have a significant effect on the procedures (see Section 6 for further confirmation of this assertion regarding μ_1^c).

The main problem with using the counts $U_i^{c,d}$ arises from the fact that cell counting procedures cannot easily differentiate living cells from those which are recently dead but have not yet disintegrated. The size of the counting error will differ according to the amount of cell killing which has preceded the count, and estimation based on the counts may therefore not be particularly accurate.

A variation which is often incorporated in protocols which otherwise conform to the outline in Section 2 is a viable cell count after exposure. This involves plating a small fraction of the culture at time t_2', and counting colonies after incubation in a manner similar to the viable cell count made at t_{R+2}'. Such an extra measurement provides an alternative estimate of μ_1^c, if the

cloning efficiency immediately after exposure is known. If it is not, the cloning efficiency can be estimated at each dose level in the same way as the post-expression cloning efficiency is estimated by (5.10).

6. Simulation Study of Likelihood Procedures for Dose-Response Data

In this section we use computer simulation to examine the validity of some of the likelihood methods discussed in the previous section. We employ the normal approximation, and focus in particular on methods for interval estimation and hypothesis testing for β, the slope parameter in the relationship (5.4). This case is of particular interest, since β is the natural measure of mutagenic potency if (5.4) is valid, and the determination of whether the data indicate a statistically significant mutagenic effect can be made by testing H_0: $\beta = 0$. The simulations are based on the protocol of Clive, *et al.*, (1979) (see (5.1)).

For the experiments we simulated, we also used the following:

(i) $C = 5$, $D_c = 1$; $C = 5$, $D_c = 2$; $C = 10$, $D_c = 2$.

(ii) For $C = 5$, the actual doses were 0, 2, 8, 32, 128; For $C = 10$, the doses were 0, .5, 1, 2, 4, 8, 16, 32, 64, 128.

The numbers of doses and replicates in (i) were chosen to cover the range of values which are used in practice. Five dose levels with no replication is typical for routine screening of a chemical, while 10 dose levels with duplication would constitute the design for some U.S. Government testing programs. Spacing of dose levels equally on a logarithmic scale as in (ii) is also common practice.

For each set of parameter values, "experiments" were simulated using each of the three combinations of C and D_c. Observations on M, the number of mutant colonies counted at the end of expression, were generated from the density derived by Tan (1981), with appropriate parameter values substituted. Observations on T, the number of viable colonies counted, were generated from a Poisson distribution with parameters as specified by (3.3).

The methods evaluated in this section are all standard likelihood-based techniques for interval estimation and hypothesis testing, in which we used the normal approximation derived in Section 4 to provide the joint density of $\{M^{c,d},\ c = 0,1,\ldots,C;\ d = 1,2,\ldots,D_c\}$, and hence the likelihood. For each set of 5000 experiments in which $\beta = 0$, we compared nominal Type I error rates with those estimated from the simulation; for sets of 5000 experiments in which $\beta > 0$, we calculated confidence intervals of specified coverage, and compared nominal coverage with the coverage probabilities estimated from the simulations. The specific procedures used to test the hypothesis H_0: $\beta = 0$ against the one-sided alternative H_1: $\beta > 0$ were the following:

H1. Reject H_0 if $\hat{\beta} / \hat{\sigma}_\beta > Z_{(1-\nu)}$,

where $\hat{\beta}$ is the m.l.e. of β; $\hat{\sigma}_\beta$ is its estimated asymptotic standard error, obtained by substituting the m.l.e.'s of α and β into the appropriate term of the inverse of the information matrix, and ν is the nominal significance level of the test.

H2. Reject H_0 if $\sqrt{\Psi_0} > Z_{(1-\nu)}$,

where Ψ_{β_0} is the likelihood ratio statistic for testing H: $\beta = \beta_0$; namely

$$\Psi_{\beta_0} = 2\left\{ l(\hat{\alpha}, \hat{\beta}) - l(\hat{\alpha}_{\beta_0}, \beta_0) \right\},$$

where $l(\alpha, \beta)$ is the log-likelihood evaluated at α and β, $\hat{\alpha}$ is the m.l.e. of α, and $\hat{\alpha}_{\beta_0}$ is the m.l.e. of α when $\beta = \beta_0$.

The following $(1 - \nu)$ confidence intervals for β were calculated:

C1. $\{\beta: |\beta / \hat{\sigma}_\beta| < Z_{(1-\nu/2)}\}$.

C2. $\{\beta: \sqrt{\Psi_\beta} < Z_{(1-\nu/2)}\}$.

We would expect in both cases that the first procedure would be computationally simpler, while the second would be somewhat more accurate.

We first examined the accuracy of these procedures for the case where λ and μ_1^c, for $c = 1,\ldots,C$ are assumed to be known, and their exact values used to simulate the experiments

can therefore be substituted into the likelihood for the application of the above procedures. Table 3A reports results of the Type 1 error rate estimation for two sets of data in which $\beta = 0$. In the first of these, there is no toxicity; in the second, $\mu_1^c = .001 x_c$. Table 4A reports the results of the coverage probability estimation for a data set in which $\beta > 0$, with $\mu_1^c = .001 x_c$. For each set of parameter values, 5000 "experiments" were simulated.

It is clear from the tables that the methods for hypothesis testing and interval estimation based on Ψ_β (i.e. H2 and C2 respectively) are extremely accurate, while those based on an assumed normal distribution with known variance for $\hat{\beta}$ (i.e. H1 and C1) are only just satisfactory for the largest sample size. The increased complexity of H2 and C2 appears to be justified. Method H2 slightly overstated the significance, and C2 produced slightly narrow confidence intervals. However, the errors in both cases are very small, and generally within 2 standard deviations of the simulation error.

In further simulations, we retained the assumption of known λ, but estimated μ_1^c for $c = 1,\ldots,C$ as described in Section 5.2. Table 3B reports the results for one data set in which $\beta = 0$, and Table 4B contains results for a data set in which $\beta > 0$. For these data sets, 1000 experiments were simulated. Only methods H2 and C2 were examined, because of their greater accuracy demonstrated in the simulations with μ_1^c known.

Again, the nominal values are satisfactorily close to those estimated from the simulations. This result confirms the assertion made in the previous section that estimation of the toxicity parameters from the viable cell counts alone does not affect the distribution of the likelihood ratio statistic used in methods H2 and C2.

No simulations were carried out using the likelihood based on the density in the iterative form derived by Tan (1981). Although it would clearly be of interest to examine the small-sample accuracy of likelihood methods based on this distribution, the computational expense of repeatedly maximizing the likelihood in the iterative form described in Tan (1981) far exceeded the budget for this project. A small number of experiments were run, and it was found that the cost of estimation using the density exceeded that incurred using the normal approximation by

a factor of about 500.

7. Power and Design Considerations

In Sections 5 and 6 we introduced and evaluated procedures for analyzing data from *in vitro* mutagenesis assays which are carried out according to the protocol specified in Section 2. We now turn our attention to an investigation of the power of some of the proposed hypothesis tests, and develop guidelines for certain aspects of assay design.

The *design* of a mutagenesis assay refers to the values of all variables which are under the control of the experimenter. These variables may be conveniently classified into two groups.

(i) **Protocol variables:** These specify the experimental steps to be applied to each experimental cell culture, and include the length of the exposure stage, the number and length of expression stages, the initial number of cells exposed, the number of cells sampled for viable cell count, and the cellular density to be maintained at the start of each expression stage. Values of the protocol variables are usually specified by a formal assay protocol.

(ii) **Dosing variables:** These variables specify how many dose levels of chemical are tested, what specific doses are employed, and how many replicates are tested at each dose. Their values are decided upon by individual experimenters.

The values of the protocol variables specified in the assay protocol are generally determined by biological considerations, although there are economic and pragmatic factors involved as well. For example, the initial number of cells has to be large enough to produce an adequate number of mutant colonies by the end of selection even if the test chemical is not mutagenic, but small enough to be manageable in an average laboratory setting. The duration of the expression period is experimentally determined to be that required for the mutants produced during exposure to become phenotypically mutant, enabling their detection at selection. However, it is desirable to make the expression as short as possible while still being consistent with this objective, in order to minimize the overall duration of the assay.

The individual experimenter arrives at the values of the dosing variables in a somewhat more *ad hoc* manner. The total number of cultures employed is typically determined by the budget available for the assay; within this constraint, most experimenters test as many dose levels as possible, with either 1 or 2 replicates at each dose level. The spacing of doses is usually equal on a logarithmic scale, especially if the range of the chemical's toxic activity is unknown. If it is known, spacing equal on an arithmetic scale may be employed. The highest dose is chosen to be that which produces some specified maximum amount of toxicity during exposure ($S_\mu = .1$ and $S_\mu = .05$ are often used).

It is clear from the formula (3.5) that the variance of the observed number of mutant colonies is a function of the protocol variables. We would therefore expect that statistical optimality criteria such as the precision of parameter estimates and the power of hypothesis tests are in turn functions of these variables. If the test chemical affects cell growth, death or mutation rates, it is also apparent from (3.5) that the choice of dose levels can be a determinant of these optimality criteria. Thus, both the protocol and dosing elements of the design can influence the statistical precision of the experiment, and if we were to approach assay design from a purely statistical viewpoint, we would be faced with a very complex optimization problem. It is more realistic for the statistician to take the protocol variables of the design as given, and concentrate on assisting the biologist with strategies for choosing the dosing variables. It is concerning this latter aspect of assay design that the statistician is most likely to produce useful results; and it is here that the biologist is most likely to seek the statistician's assistance. Of course, this is not to say that the protocol aspects of design should be ignored by the statistician; within bounds of biological acceptability, there is almost certainly room for statistical improvement of existing protocols.

Reduction of the design problem to consideration of only the dosing variables still leaves us with substantial difficulties. First, the observations have variances which are in general functions of unknown parameters, and it is well known that in this situation, only local optimality is achievable (Fedorov, 1971). Another difficulty lies in the choice of optimality criteria. From a

theoretical point of view, a scalar function of the covariance matrix of parameter estimates, such as its trace or largest eigenvalue, is the most useful measure of experimental precision. On the other hand, from a practical point of view the power of certain hypothesis tests is of more interest. In particular, it is desirable to maximize the power of the test that the chemical has no mutagenic effect, and the power of a test to detect departures from a linear dose-response. In any reasonable situation the two criteria of power and estimation variance should produce nearly identical design rankings. The power, however, is a more meaningful quantity for the biologist, and we accordingly adopt it for the remainder of this section. Finally, we have to deal with the problem of the great diversity in specific protocols. In this section we once again use the mouse lymphoma protocol (see (5.1)) for illustrative purposes. However, the general principles should be broadly applicable, even if the specific conclusions are not.

Our strategy will be to develop approximate power functions, in some cases checking the approximation by exact calculation or simulation. We will then illustrate how the power functions may be used to both optimally select the dosing variables, and provide estimates of the power (or Type II error rates) inherent in currently used protocols when the statistical tests are applied.

We will first consider the test of no mutagenic effect for the case where a single treated culture is to be compared to a single control, and examine the effect of the amount of toxicity in the treated culture. We then treat the case where replicates and multiple dose levels are employed, and it is possible to test the goodness-of-fit of the linear model (5.4). The assumptions A1 - A11 and the approximations (3.2) and (3.3) will be retained.

7.1. *Single Treated and Control Culture*

In section 5, two methods were suggested for testing H_{0c}: $p_1^c = p_1^0$ against the one-sided alternative H_{1c}: $p_1^c > p_1^0$. These were

(i) The likelihood ratio test statistic, which we denoted Λ,

(ii) The statistic Γ (see equation (5.7)).

It was shown by exact calculation that when only p_1^0 and p_1^f are unknown, and $D_0 = D_c = 1$, the nominal significance level of the test is very close to the true level. Monte Carlo simulation revealed a similar finding for the case where a toxicity parameter in the treated culture is also estimated, from the viable cell determination. In order to investigate the power of these tests as functions of the level of the tests, the true values of unknown parameters and the design variables, we require closed form approximations for the power.

Once again, we initially restrict attention to the case where only the mutation parameters p_1^0 and p_1^f are unknown. First consider the test statistic Γ. The statistic is asymptotically (for large D_0 and D_c) normally distributed, with mean

$$E(\Gamma) = \frac{(p_1^f - p_1^0)\, r_1}{\left\{ \frac{1}{D_c^2} \sum_{d=1}^{D_c} V[c, d, \bar{p}] \,/\, (K^{c,d})^2 + \frac{1}{D_0^2} \sum_{d=1}^{D_0} V[0, d, \bar{p}] \,/\, (K^{0,d})^2 \right\}^{1/2}}$$

and variance

$$V(\Gamma) = \frac{\frac{1}{D_c^2} \sum_{d=1}^{D_c} V(M^{c,d}) \,/\, (K^{c,d})^2 + \frac{1}{D_0^2} \sum_{d=1}^{D_0} V(M^{0,d}) \,/\, (K^{0,d})^2}{\frac{1}{D_c^2} \sum_{d=1}^{D_c} V[c, d, \bar{p}] \,/\, (K^{c,d})^2 + \frac{1}{D_0^2} \sum_{d=1}^{D_0} V[0, d, \bar{p}] \,/\, (K^{0,d})^2}$$

Here $V[c, d, \bar{p}]$ is the variance of the number of mutants in the dth replicate at the cth dose as given by (3.5), but with $p_1^0 = p_1^f = \bar{p}$, and $\bar{p} = \frac{1}{D_0 + D_c}(D_0 p_1^0 + D_c p_1^f)$. Under H_{0c}, $E(\Gamma) = 0$. Therefore, the large-sample power of the test based on Γ for a critical value $Z_{(1-\nu)}$ is given by

$$1 - \Phi\left[\left\{ Z_{(1-\nu)} - E(\Gamma) \right\} / \sqrt{V(\Gamma)}\right],$$

where Φ is the standard normal distribution function.

The likelihood ratio statistic Λ is obtained as twice the difference between the log-likelihood maximized over p_1^f and p_1^0, and that maximized over p_1^f and p_1^0 subject to the constraint $p_1^f = p_1^0$. Standard theory suggests that Λ is asymptotically distributed as χ_1^2, with non-centrality parameter $\tau = (p_1^f - p_1^0)^2 / \sigma^2$, where σ^2 is the asymptotic variance of the m.l.e. of $p_1^f - p_1^0$. We obtain σ^2 by introducing the reparametrization $\delta = p_1^f - p_1^0$, and computing the information matrix with respect to the parameters p_1^0 and δ. The variance σ^2 is then the diagonal entry corresponding to δ in the inverse of the information matrix. The exact form of the information matrix under the normal approximation is given in Appendix 2. We then calculate the large-sample power for a critical value of $\chi^2_{1,(1-\nu)}$ as

$$1 - G_{1,\tau}(\chi^2_{1,(1-\nu)}),$$

where $G_{k,\tau}$ is the distribution function of a χ_k^2 random variable with non-centrality parameter τ.

There is some ambiguity in the literature regarding the value of δ at which the information matrix should be evaluated to obtain τ. As discussed, for example, in Cox and Hinkley (1974, p. 323), the non-central χ^2 approximation is valid for "local" alternatives; that is, δ_D such that $\delta_D = \delta_0 + O(D^{-1/2})$ where δ_0 is the true value of δ, $D = D_0 + D_c$, and the limit in D is taken as D_0 and D_c both become large. In this case, evaluation of the information matrix at δ_0 provides a satisfactory approximation up to terms of order $D^{-1/2}$ in Λ. On the other hand, Kendall and Stuart (1961, p. 231) seem to suggest that the information matrix should be evaluated at the alternative.

These asymptotic arguments are in fact somewhat irrelevant for the situation under consideration here, with $D_0 = D_c = 1$. In order to check the validity of the approximations in this case, exact power calculations were carried out in a manner identical to that described for the exact Type I error rate calculations reported in Section 5.1. We again employed the mouse lymphoma protocol (Clive, *et al.*, 1979) as a framework within which to make the calculations (see (5.1)). The distribution derived by Tan (1981) was used to evaluate the probability of the

outcomes (numbers of mutant colonies) in the treated and control cultures which result in rejection of H_{0c} using the statistics Γ and Λ. Table 5A displays the exact powers thus calculated, and those obtained via the appropriate approximation described in this section, for a range of values of p_0^f, p_1^f, μ_1^f and ν, the nominal significance level. We let $K_p = p_1^f / p_1^0$, the fold increase over the background mutation rate produced by the test chemical during exposure; and again $S_\mu = e^{-\mu_1^f r_1}$, the fraction of cells surviving exposure in the treated culture as compared with the control. For the approximation to the power of the test based on Λ, the information matrix was evaluated at the alternative.

Once again, it is apparent that even for this smallest possible sample size ($D_0 = D_c = 1$), the approximations for the power of the tests based on Γ and Λ are remarkably good. We can draw a number of other specific conclusions from this table. First, the tests based on Γ and Λ have very similar exact powers. The test based on Γ appears to be slightly more powerful, but this may simply be a reflection of the fact that for the same nominal type I error rate, the true error rate for the test is somewhat larger than that based on Λ. Second, we note that the approximation for the power of the test based on Γ is generally better than the approximation for Λ over the parameter values for which the power was calculated. The approximation for the likelihood ratio statistic starts to break down when both K_p and μ_1^f are large, while the approximation for Γ appears to be good over all the parameter values. The only area where the approximation for Γ performs badly is for alternatives close to the null hypothesis when $p_1^0 = 2.5\times10^{-7}$. In this case, the approximation shows power increasing with μ_1^f, while all the exact calculations demonstrate that increased toxicity produces lower power. However, the discrepancy is in a region of low power, and is anyway rather small. Third, it can be seen that the power to detect the same fold increase in p_1^f over p_1^0 is much higher for the higher level of p_1^0 (i.e. 7.5×10^{-7}). This is to be expected, since the absolute increase is obviously much greater.

Having obtained these approximations, we can use them to provide power curves for the hypothesis tests. In the remainder of this section, we will use the approximation based on Γ, since it is generally more accurate for the (roughly equal) power of both tests. We consider the

effect of toxicity during the exposure period. Figure 2 displays contours of equal power for varying values of K_p and $S_\mu = e^{-\mu_1 r_1}$, when $p_1^0 = 7.5 \times 10^{-6}$, the significance level is .05, and the protocol variables are as in (5.1). The figure also indicates corresponding values of K_{MF}, the ratio of treated to control expected mutant fraction. The mutant fraction $M^{c,d} / K^{c,d}$ is an unbiased estimate of

$$q^c = p_1^c \, r_1 + p_1^0 \sum_{k=2}^{R+1} r_k;$$

Thus

$$K_{MF} = q^c / q^0 = \left\{ p_1^c \, r_1 + p_1^0 \sum_{k=2}^{R+1} r_k \right\} / \, p_1^0 \sum_{k=1}^{R+1} r_k$$

$$= \frac{p_1^c}{p_1^0} \left\{ r_1 / \left(\sum_{k=1}^{R+1} r_k \right) \right\} + \sum_{k=2}^{R+1} r_k / \left(\sum_{k=1}^{R+1} r_k \right),$$

so K_{MF} is a linear transformation of $K_p = p_1^c / p_1^0$.

The ratio K_{MF} is of interest because biologists have suggested that an *observed* value of K_{MF} of greater than 2 be used as the minimum requirement for a chemical to be declared a mutagen. Clearly such a rule, which is independent of sample size, is of limited use in a statistical context. However the K_{MF} scale can be used to read off the power available to detect various *expected* values of K_{MF}.

It can be seen that the power to detect specific values of K_p (or K_{MF}) is not much affected by changes in μ_1 down to about $S_\mu = 0.25$. However, as the surviving fraction decreases from this level, there is a rapid increase in the value of K_p which is required for a mutagenic effect to be detected with the same power. It is of some note that a 2-fold increase in q (i.e. $K_{MF} = 2$) is detected with greater than 0.9 probability when survival as a fraction of control is only as low as about 0.7.

Figure 3 provides the same information in a more conventional format: Each curve is a plot of power against K_p, for a particular value of S_μ. It may be seen that a 10-fold increase in p_1^f will be detected with probability about .9 when the survival after exposure is 50% of control, but that at 5% survival, the same increase will be detected with probability less than 0.7.

It is common practice to test chemicals at doses up to that which causes no less than 10% survival as a percentage of control (e.g. Clive, *et al.*, 1979). Depending on the relationship between the chemical dose x, and μ_1 and p_1, there may be an 'optimal' dose, at which the power to detect a mutagenic effect is maximized. Even though increasing x will generally produce larger values of p_1, thereby increasing the power for fixed μ_1, the effect of increasing μ_1 for fixed p_1 has the effect of decreasing the power by increasing variability in the treated culture. If cell killing increases "faster" with x than does p_1, the power will start to decrease. For example, suppose a chemical at dose x_1 produced 50% killing and $K_p = 10$, but at dose x_2 produced 95% killing and $K_p = 12$. It can be seen from Figure 3 that the increase in dose would result in a drop in power from about 0.9 to 0.7. Obviously p_1 as a function of dose is not known in advance for an assay. However, preliminary toxicity tests are often carried out which provide some indication of the relationship between x and μ_1. An unreplicated assay which has desired power to detect a mutagenic effect in all pairwise comparisons of treated and control can be designed as follows: Choose the highest dose x_C to be that for which μ_1^C is such that a desired value of K_p (or K_{MF}) is detected with a specified probability (presuming the appropriate power contour in Figure 2 crosses the desired value of K_p).

Since μ_1^f is usually estimated, simulations were carried out for a subset of the parameter values in Table 5A, to determine the effect of estimation of μ_1^f on the power. It can be seen from Table 5B that the approximations still hold good in this situation.

7.2. *Replication and Multiple Doses*

It is well known from the theory of linear models that in least-squares linear regression of a homoscedastic observed variable against a single predictor, the estimate of slope will have minimum variance if half the observations are made at each of the two extremes of the scale of the predictor variable. It has also been shown (Chernoff, 1953) that for fairly general regression models, the asymptotic information is maximized when the number of points at which observations are made is equal to the number of unknown parameters; in the case of the linear relationship (5.4), this also implies that a design with two dose levels provides optimal estimates of α and β. However, it is important to know not only whether the chemical has a mutagenic effect, but whether the effect is linear. A response which is concave upward would suggest a multi-hit mutagenic mechanism, while a concave downward response would probably indicate to the biologist that a saturation effect is taking place (Myers, 1981). It is obviously not possible to compare the fit of the linear assumption with that of a more general alternative using data from an experiment in which only two dose points are tested.

While a visual examination of plotted data should always be made to check for linearity, it is also desirable to have available a means of formally testing for linearity. There are a number of ways this might be accomplished. First, we might fit a quadratic model and test for the significance of the second-degree term. An alternative is to fit a "saturated" model, in which p_1^c is estimated for each dose c, and compare the fit (or the maximized likelihood) under this model with that under the linear model using a likelihood ratio test. While the former test probably has somewhat better power against specifically quadratic alternatives, the saturated alternative has the advantage that it can be used to test the fit of any model which relates the dose of test chemical x to the mutation probability $p_1(x)$. We will adopt this goodness-of-fit test for the remainder of this section.

Having decided upon a means of testing for non-linearity, we can examine the effect of various dose allocations on the power of the tests of interest; namely, the likelihood ratio test for testing $H_0{:}\beta = 0$, and the goodness-of-fit test for linearity. We again use the non-central χ^2

distribution to provide approximate power functions for the test, and as before suppose that only the mutation parameters are unknown.

Under the linear relationship (5.4), there are two unknown parameters, namely α and β, to be estimated by maximum likelihood. The asymptotic approximation for the distribution of the likelihood ratio statistic for testing H_0: $\beta = 0$ is a χ_1^2 random variable with non-centrality parameter $\tau_1 = \beta^2 / \sigma_\beta^2$, where σ_β^2 is the lower right-hand member of the inverse of the information matrix with respect to the parameter vector $\boldsymbol{\omega}^T = (\alpha,\beta)$. The large-sample power for a critical value $\chi^2_{1,(1-\nu)}$ is therefore given by

$$1 - G_{1,\tau_1}(\chi^2_{1,(1-\nu)}).$$

The asymptotic power function for the goodness-of-fit test can be similarly derived. Under the saturated model, we estimate the $(C+1)$-dimensional parameter vector $\mathbf{p}_1^T = (p_1^0, p_1^1, \ldots, p_1^C)$. Define the $(C+1)\times(C+1)$-dimensional matrix

$$A = \begin{pmatrix} \mathbf{I}_2 & \mathbf{0} \\ \mathbf{B} & \mathbf{I}_{C-1} \end{pmatrix}$$

where $\mathbf{I}_l$ denotes the $l\times l$-dimensional identity matrix, $\mathbf{0}$ is a $2\times(C-1)$-dimensional matrix of zeroes and $\mathbf{B}$ is a $(C-1)\times 2$-dimensional matrix with elements $B_{l1} = x_{l+1}/x_1 - 1$ and $B_{l2} = -x_{l+1}/x_1$, for $l = 1,\ldots,C-1$. Then it can be seen that the vector $\boldsymbol{\xi} = \mathbf{A}\mathbf{p}_1$ represents a non-singular, linear transformation of the unknown parameter vector $\mathbf{p}_1$, such that $\xi_1 = p_1^0$, $\xi_2 = p_1^1$, and $\xi_l = p_1^{l-1} + p_1^0(x_{l-1}/x_1 - 1) - p_1^1 x_{l-1}/x_1$, for $l = 3,\ldots,C+1$. The null hypothesis that the linear model fits as well as the saturated model is equivalent to the hypothesis H_L: $\xi_l = 0$, $l = 3,\ldots,C+1$, since under H_L, the points (x_l, p_1^l) for $l = 2,\ldots,C$ lie on the unique line through $(0,p_1^0)$ and (x_1,p_1^1). Now because $\boldsymbol{\xi} = \mathbf{A}\mathbf{p}_1$, $Var(\hat{\boldsymbol{\xi}}) = \mathbf{A}\,Var(\hat{\mathbf{p}}_1)\mathbf{A}^T$, where $\hat{\boldsymbol{\xi}}$ is the m.l.e. of $\boldsymbol{\xi}$, and $\hat{\mathbf{p}}_1$ is the m.l.e. of $\mathbf{p}_1$. $Var(\hat{\mathbf{p}}_1)$ is in turn asymptotically given by the inverse of the $(C+1)\times(C+1)$-dimensional information matrix with respect to $\mathbf{p}_1$. Let $\mathbf{S}$ denote the lower right-hand $(C-1)\times(C-1)$-dimensional corner of $Var(\hat{\boldsymbol{\xi}})$. $\mathbf{S}$ is the asymptotic

covariance matrix of the (unconstrained) maximum likelihood estimate of the vector $\bar{\xi}^T = (\xi_3, \ldots, \xi_{C+1})^T$. Then standard large sample theory gives the asymptotic distribution of the likelihood ratio statistic for testing H_L against the saturated model as $\chi^2_{(C-1)}$, with non-centrality parameter $\tau_2 = \bar{\xi}^T S^{-1} \bar{\xi}$. The large sample power for a test using critical value $\chi^2_{(C-1),(1-\nu)}$ is given by

$$1 - G_{(C-1),\tau_2}(\chi^2_{(C-1),(1-\nu)}).$$

We now use the approximate power functions for the likelihood ratio tests of H_0: $\beta = 0$ and H_L, to gain some insight into the effect of dose spacing and allocation on the power of these tests. As before, the information matrices used in the calculation of non-centrality parameters are based on the normal approximation to the density of $M^{c,d}$. The matrices are evaluated at the alternative, since this seemed to provide slightly more accurate approximate power functions in the two-dose case described in Section 7.1.

Figures 4A and 4B show two linear relationships between x and $p_1(x)$, over the dose range [0,64]. Superimposed on each line is a non-linear alternative, chosen from the family of curves $h(x) = a + bx + fe^{gx}$ in such a way that

$$\int_0^{64} (h(x) - [\alpha + \beta x])^2 \, dx$$

is minimized over α and β; thus the line $\alpha + \beta x$ is the "least-squares fit" to the non-linear curve $h(x)$. The line and curve in Figure 4B represent more extreme departures from, respectively H_0 and H_L, than do the line and curve in Figure 4A. For a number of allocations of 24 experimental cultures, Table 6A gives the approximate power to detect the non-zero slope β in Figure 4A, using the likelihood ratio test for H_0, and the approximate power to detect the curve $h(x)$ in Figure 4A as non-linear using the likelihood ratio test of H_L against the saturated model. The allocations chosen were equally spaced on either an arithmetic or a logarithmic scale, and were equally replicated at each dose. All allocations had the end-point doses

0 and 64. The toxicity was modelled by $\mu_1(x) = \nu x$, with ν chosen such that $S_\mu = .1$ at $x = 64$. The table indicates that when the doses are spaced equally on the arithmetic scale, the power is much greater for testing goodness-of-fit than when the doses are equally spaced on a log scale. The power for testing that $\beta = 0$ is also slightly higher. In both cases, the power for the test of zero slope increases as the number of doses becomes smaller. For the test of linearity, the power is maximized at three doses when spacing is equal on an arithmetic scale, while on a logarithmic scale, equal spacing results in maximum power at six doses. Table 6B shows similar results for the line and curve in Figure 4B.

These calculations argue strongly for usage of 3 dose points in mutagenesis testing. While the power to detect the slope as non-zero is somewhat decreased as compared with the 2-dose allocation, the power to detect non-linearity of the kind modelled by $h(x)$ appears to be maximized, as compared with the other equal-spacing allocations.

We then investigated the effect of the location of the middle dose on the power of the two tests. For the curves in Figure 4A, Figure 5A plots the power of the tests as a function of the fraction of the distance between dose 0 and dose 64 represented by the middle dose. There is, as we might expect, a far greater effect on the power of the goodness-of-fit test than on the power of the test for zero slope. The power of the latter test is maximized as the fraction approaches 1; this reflects the increasing variance with dose, and the consequent gain in power achieved by disproportionate replication in the high dose region. The power of the goodness-of-fit test is maximized at about .6 of the distance between the lowest and highest dose. Similar results were obtained for the curves in Figure 4B (see Figure 5B).

8. Data Analyses

In this section, some of the statistical methods introduced in Sections 5 - 7 will be illustrated by analyses of data obtained under the mouse lymphoma protocol of Clive, *et al.* (1979). We will also utilize replicated mouse lymphoma data to examine some of the assumptions which were made in Sections 3 and 5. The protocol variables are given in (5.1a) - (5.1d). All experiments

discussed in this section were carried out at SRI International. The author is grateful to Dennis Tajiri and David Low (formerly of SRI International), and Ann Mitchell and Pam Shrieve (presently at SRI International) for providing the data and assisting in their interpretation.

8.1. *Examples of Statistical Methods*

Table 7 gives the results of an assay carried out on chemical X. The chemical was tested in duplicate at 7 dose levels, including zero; the table records the mutant and viable cell colony counts, and the culture size based on Coulter counts at the end of each of the two expression stages. Under the mouse lymphoma protocol, no density measurement or viable cell count is made immediately after exposure. The recurrence of specific values for the culture size (e.g. 13.05 and 12.01) is the result of the use of conversion tables, which transform a range of Coulter counter density readings into the same estimate of overall density. Also reported in the table are estimates of π^c, the cloning efficiency at each dose, and of the growth parameter d_3^c, based on the formulae in Section 5.3. The procedure suggested in that section for estimating d_1^c and d_2^c when density counts immediately after exposure are unavailable does not provide meaningful results in this experiment: Use of $\hat{d}_2^c = \hat{d}_3^c$ results in inappropriately large negative values for $\hat{d}_1$, since even in the control cultures the growth between t_1 and t_3' appeared to be much slower than the growth between t_3 and t_4'. We therefore employed an alternative *ad hoc* procedure. Specifically, it was supposed that $d_1^0 = d_2^0 = \lambda_1^0$, and that $\mu_2^c = 0$ for all c.

The resulting estimates of d_1^c appear in Table 7. There is some inconsistency in allowing d_3^c to be dose-dependent, while forcing d_2^c to be constant over dose. However, there appears to be no alternative if we are to make some use of the density counts.

We also use the value $r_2 = 16$ hours rather than the 24 hours specified in the protocol, since the former is the value used in practice by the experimenters at SRI International.

The statistical methods were first applied with the estimates of π^c, d_1^c, d_2^c and d_3^c in place of their true values in the likelihood of the mutant colony counts. The parameters of the

relationship (5.4) were estimated by maximizing the likelihood based on the normal approximation to the density of $M^{c,d}$, resulting in $\hat{\alpha} = 2.36\times10^{-6}$ and $\hat{\beta} = 5.50\times10^{-6}$. The likelihood ratio test for zero slope, H_0, was highly significant ($p < .0001$), and the likelihood ratio test of the linear-fit against a saturated model resulted in rejection of the linear hypothesis ($p = .0001$). Figure 6A displays the m.l.e.s of p_1^c, $c = 0,1,...,C$, obtained under the saturated model, and the line $p_1(x) = \hat{\alpha} + \hat{\beta}x$. The lack of linear fit appears to be due to upward concavity, suggesting a multi-hit mutational mechanism. Figure 6B displays the observed counts of mutant colonies, and their corresponding expected values under the fitted linear model. The method C2 (see Section 6) was used to calculate a 95% confidence interval for β, resulting in the interval $(4.95\times10^{-6}, 6.03\times10^{-6})$.

We then analyzed the experiment without utilizing the information from the density counts, as outlined in Sections 5.1 and 5.2. The cloning efficiency was assumed to be constant ($\pi = .85$), the division parameter λ was estimated from the control viable counts, resulting in $\hat{\lambda} = .050$, and the toxicity parameters μ_1^c were estimated from the viable counts at the cth dose. The same procedures as above were applied, resulting in $\hat{\alpha} = 2.46\times10^{-6}$ and $\hat{\beta} = 4.84\times10^{-6}$ with 95% confidence interval $(4.22\times10^{-6}, 5.48\times10^{-6})$. The tests of zero slope and linearity were still significant, but less so ($p < .0001$ and $p = .018$, respectively). The m.l.e. of $\hat{\beta}$ is somewhat lower than that obtained when the information from the density counts was used, and its confidence interval is wider. Both of these results are due to the fact that when all of the difference in the growth of the cultures is ascribed to toxicity during exposure, as is the case in the second analysis here, the variance at each dose as calculated by (3.5) is higher than when the difference is distributed among the stages and the cloning efficiency. The effect on the variance is greatest at the higher 2 doses, and this results in the effective de-weighting of counts at these doses in the estimation of β. The estimate of β is lower than in the previous case, since it is the counts at the higher 2 doses that appear to be greater than would be predicted by a linear model fit to the first 5 doses.

8.2. *A Data Set with 60 Replicates*

In order to obtain some insight into the distribution of the number of mutant colonies counted, an experiment was carried out at SRI International in which 60 replicate cultures of mouse lymphoma cells were exposed to a fixed dose of the commonly used solvent, DMSO. It is generally believed that DMSO is not mutagenic, so that mutant counts should not be elevated above background. The compound is, however, somewhat toxic to the cells (Amacher, *et al.*, 1980). The usual procedure was followed of sampling a fraction of each culture to produce the required density at the end of each day of expression. Because the fraction sampled varies among the cultures, they are not true replicates, in the sense of being handled identically (to the best of the experimenter's ability). However, the sampling fractions do not vary greatly. They are in the range .1 - .16 at the end of the first stage of expression, and between .16 and .23 at the end of the second. We will therefore treat the cultures as true replicates in the consideration of distributional properties.

The cloning efficiency π and the parameters d_1, d_2 and d_3 were estimated from the cell density counts and viable cell counts, as described in Section 5.3; the resulting estimates were .85, .047, .072 and .072 respectively. It was assumed that there was no cell death during expression, and that the growth rate was constant. The mutation parameter p_1^0 was estimated from the data by the formula (5.5). These parameter estimates were then used to generate the fitted, or "theoretical" cumulative distribution function (c.d.f.) of the observations, by the method described by Tan (1981). Figure 7 plots this theoretical c.d.f. against the empirical c.d.f of the mutant colony counts. It is apparent that although the location and general shape of the theoretical c.d.f. closely match that of the empirical c.d.f., the variance of the fitted distribution is substantially less than that of the observations. In Figure 8, the c.d.f.'s are plotted on normal probability paper. While both distributions are apparently somewhat skewed, the plots are sufficiently linear to suggest that the normal density can provide a useful approximation. This is of course the conclusion we reached through the computations and simulations in Sections 5 - 7.

It is a matter of some concern that the fitted distribution understated the variance of the observations. If the cause is simply the variation in the sampling fractions, the inference procedures we have developed should not be affected, since the sampling fractions are incorporated in the density of each observation. On the other hand, if the cause is the uncontrollable within-dose variability discussed in Section 3.3, the procedures may be seriously affected. In particular, if the models used are understating the variance of the observations, hypothesis tests will have Type I error rates which are greater than their nominal level, and confidence intervals for unknown parameters will contain the true value of the parameter with a lower probability than the nominal coverage. It will be important for experimenters to generate more data sets of the kind examined briefly here, in order to resolve some of these questions.

9. Discussion, Conclusions and Areas for Further Research

Using a modelling approach to the analysis of mutagenesis assay data, we have presented a framework within which hypothesis testing and point and interval estimation of parameters can be carried out. Specifically, we have demonstrated that standard asymptotic methods based on a normal approximation provide acceptable results, provided the assumptions and approximations discussed in Sections 3 - 5 are valid.

It is possible that other statistical procedures may have produced equally good results within the assumed model. For example, estimation of the parameters α and β in the relationship (5.4) could be accomplished by iteratively-reweighted least squares, using the variance formula (3.5) to provide the inverses of the weights at each new step in the estimation. It is known that for generalized linear models such a scheme is equivalent to maximum likelihood estimation (Nelder and Wedderburn, 1972) and the success of the likelihood methods based on the normal approximation suggests that such an equivalence may also hold (at least approximately) for the model under consideration here.

Another way of dealing with heteroscedastic regression data is to derive or empirically estimate a transformation which produces homoscedasticity. This is the approach taken by Snee and Irr

(1981) for the analysis of CHO mutagenesis data. They estimate a Box-Cox-type power transformation from a large amount of historical data. However, it is unlikely that a single transformation can produce homoscedasticity of the observations $M^{c,d}$. This can be simply illustrated by the fact that testing a chemical which is non-mutagenic but toxic could result in cultures at two different doses, which had the same means but very different variances, because of differential toxicity. Therefore, to first order any transformation will produce random variables which again have the same mean but still have different variances.

We have also shown how approximate power functions can be calculated for hypothesis tests of interest, and used the functions to suggest guidelines for the dosing design of optimal experiments. The designs arrived at are only optimal in a restricted class. However, they appear to provide substantial improvement in power over currently used dosing allocations, while being still sufficiently "regular" to be acceptable to the biologist.

We now briefly discuss some topics which could be explored further than they have been taken in this paper, and mention areas for future research which have not been examined here at all.

(i) The robustness of the statistical procedures against violation of the assumptions A9 - A11 should be considered. As was noted in Section 5.3, the utilization of additional measurements can enable the weakening of the assumptions A9 - A11. However, the additional number of parameters which must then be estimated may have an adverse effect on the approximations to the distribution of test statistics, resulting in a trade-off in overall accuracy. If the methods proposed are sufficiently robust to the violation of assumptions A9 - A11, retention of these assumptions may be preferable to utilization of the additional measurements. Additional simulations would be required to resolve this question.

(ii) The assumption A8 can be removed if information on the growth of the experimental culture before exposure is available. For example, under the mouse lymphoma protocol as carried out by SRI International, the cells to be used for an assay are sampled from a much larger stock of cells which is kept in continuous exponential growth phase. The stock cultures are "cleansed" of any mutant cells which arise spontaneously, by weekly

treatment with methotrexate (A. Mitchell, personal communication). Under similar assumptions to those we have already made, the expected number of mutant cells in the stock culture at time t as measured from the cleansing is $E(M_t) = p_1^0 t N_t$. Then $M_1^{c,d}$, the number of mutant cells in the culture at the start of the exposure stage, can be assumed to be distributed as Poisson, with mean $p_1^0 t N_1$. This source of variability can be easily incorporated into the expressions (3.4) and (3.5) for the mean and variance of $M^{c,d}$, using standard formulae for iterated conditional means and variances. The statistical procedures suggested in Section 5 may still be applied, although the validity of the approximating distributions would again need to be checked. The presence of pre-existing mutants is not affected by dose. We could therefore expect that if A8 is employed when there are in fact pre-existing mutants, the effect would be overestimation, by a constant amount, of p_1^c, for $c = 0,1,\ldots,C$.

(iii) In any bioassay which repeatedly uses the same species of test organism, a large body of data will accumulate about the spontaneous level of response in the organism. We have been implicitly assuming that prior information is equally vague regarding spontaneous and induced mutation rates in the cell cultures, whereas in fact much more information about p_1^0 (or α in the relationship (5.4)) should be available. Recently Tarone (1982) proposed methods for incorporating historical control data in the analysis of binomial dose-response data, by in effect estimating a prior distribution for the background response rate. When this distribution has a large variance, it plays only a small role in the analysis; if the background has been historically rather stable, this information is weighted heavily in the analysis.

For mutagenesis assay data, a similar procedure could be utilized, by fitting a normal distribution to the historical estimates of p_1^0. We may wish to restrict the historical data to the particular laboratory whose experiment is being analyzed, as inter-laboratory variation in spontaneous rates may be substantial.

(iv) In the *Salmonella* mutagenesis assay (Ames test), dose-response curves typically turn downward at high doses due to cell killing. Bernstein *et al.* (1982) have suggested a means of eliminating data at doses which are toxic enough to cause downward departure from linearity. For the protocols considered here, an adjustment is made for cell killing by using the mutant fraction as the summary response measure, so there is not such a well-defined rationale for restricting attention to the initial portion of the dose-response curve. However, saturation effects may cause a plateauing of the dose-response, and some data even show a marked, though as yet unexplained, downturn at high doses (L. Thompson, personal communication). It may therefore be sometimes desirable to utilize a point-rejection rule such as that used by Bernstein *et al.* (1981) to define an initial region of linear dose-response.

Appendix 1

In this appendix, we demonstrate how standard techniques may be used to derive the moments of the number of mutant and normal cells under the assumptions A1-A7. The same techniques can be also used to obtain the moments for more general κ_{kl} under A1-A3 and A7. All quantities referred to are defined in section 3.

We first consider a single stage. Let (N_t^1, M_t^1) and (N_t^2, M_t^2) respectively denote the number of normal and mutant cells at time t arising from a single normal and a single mutant cell at time $t = 0$, with corresponding p.g.f.s $\phi_t^1(u, v)$ and $\phi_t^2(u, v)$. Then, if $G_t(u, v)$ is the p.g.f. of (N_t, M_t),

$$G_t(u, v) = [\phi_t^1(u, v)]^{N_0} [\phi_t^2(u, v)]^{M_0}.$$

Following Karlin and Taylor (1975, p. 424), ϕ_t^1 and ϕ_t^2 are the solutions of the backward equations

$$\frac{\partial \phi_t^i}{\partial t} = w^i(\phi_t^1, \phi_t^2), \tag{A1.1}$$

where w^i for $i = 1,2$ are the infinitesimal generating functions

$$w^1(u, v) = \mu + (\lambda - p)u^2 + puv - (\lambda + \mu)u \tag{A1.2}$$

$$w^2(u, v) = \mu + \lambda v^2 - (\lambda + \mu)v \tag{A1.3}$$

The equation (A1.1) with $i = 2$ only involves ϕ_t^2. The solution is the p.g.f. of a simple birth-and-death process starting from a single individual at time $t = 0$, with birth parameter λ and death parameter μ. The equation (A1.1) with $i = 1$ does not appear to have a closed-form solution, unless $\lambda = 0$ (e.g. Bartlett, 1978, p.). However, successively differentiating (A1.1) with $i = 1$ with respect to u and v and setting $u = v = 1$ at each stage results in linear differential equations for the factorial moments of (N_t^1, M_t^1) which may be solved directly, if somewhat tediously. The moments of (N_t, M_t) are simple functions of the moments of (N_t^1, M_t^1) and (N_t^2, M_t^2). The following are the first and second joint moments of (N_t, M_t) conditional on (N_0, M_0), under the assumptions A1-A6.

$$E(N_t) = N_0 e^{(\lambda - \mu - p)t} \tag{A1.4}$$

$$E(M_t) = N_0 \left\{ e^{(\lambda - \mu)t} - e^{(\lambda - \mu - p)t} \right\} + M_0 e^{(\lambda - \mu)t} \tag{A1.5}$$

$$Var(N_t) = \frac{\lambda + \mu - p}{\lambda - \mu - p} \left\{ e^{2(\lambda - \mu - p)t} - e^{(\lambda - \mu - p)t} \right\} \tag{A1.6}$$

$$\begin{aligned} Var(M_t) = N_0 \Bigg\{ & \frac{\lambda^2 - \mu^2 + 2p^2 - 3\lambda p - \mu p}{(\lambda - \mu)(\lambda - \mu - p)} e^{(\lambda - \mu - p)t} \\ & + \frac{\lambda + \mu}{\lambda - \mu} \left[e^{2(\lambda - \mu)t} - e^{(\lambda - \mu)t} \right] \\ & - \frac{2(\lambda + \mu - p)}{\lambda - \mu} e^{(2(\lambda - \mu) - p)t} \\ & + \frac{\lambda + \mu - p}{\lambda - \mu - p} e^{2(\lambda - \mu - p)t} \Bigg\} \\ & + M_0 \frac{\lambda + \mu}{\lambda - \mu} \left[e^{2(\lambda - \mu)t} - e^{(\lambda - \mu)t} \right] \end{aligned} \tag{A1.7}$$

$$Cov(M_t, N_t) = N_0 (\lambda + \mu - p) \left\{ \frac{p}{(\lambda - \mu)(\lambda - \mu - p)} e^{(\lambda - \mu - p)t} + \frac{1}{(\lambda - \mu)} e^{(2(\lambda - \mu) - p)t} - \frac{1}{\lambda - \mu - p} e^{2(\lambda - \mu - p)t} \right\} \tag{A1.8}$$

The higher moments can be similarly obtained. In particular, the third moment of M_t which was required for the skewness coefficient in Table 1 was derived in this manner, although the formula is not given here.

Returning to the full assay protocol, we now derive the first and second joint moments of (N_i, M_i), $i = 2,3,\ldots,R+1$ conditional on (N_1, M_1). Consider for example $E(M_i)$. By A7, M_i is distributed binomially, with parameters M_i' and θ_{i-1}, conditional in M_i'. Therefore,

$$E(M_i) = E[E(M_i \mid M_i')] = E[\theta_{i-1} M_i'] = \theta_{i-1} E(M_i') = \theta_{i-1} E[E(M_i' \mid M_{i-1}, N_{i-1})]$$

So

$$\begin{aligned} E(M_i) &= \theta_{i-1} E\left[N_{i-1}\left(e^{d_{i-1} r_{i-1}} - e^{(d_{i-1} - p_{i-1}) r_{i-1}}\right) + M_{i-1} e^{d_{i-1} r_{i-1}}\right] \\ &= \theta_{i-1}\left[e^{d_{i-1} r_{i-1}} E(M_{i-1}) + \left(e^{d_{i-1} r_{i-1}} - e^{(d_{i-1} - p_{i-1}) r_{i-1}}\right) E(N_{i-1})\right] \end{aligned}$$

where $d_i = \lambda_i - \mu_i$, $r_i = t_{i+1}' - t_i$ and making use of (A1.5).

Similarly, we can obtain difference equations for the other first and second moments of (N_i, M_i). With $\bar{\theta}_i = 1 - \theta_i$,

$$\begin{aligned} E(N_i) &= \theta_{i-1} e^{(d_{i-1} - p_{i-1}) r_{i-1}} E(N_{i-1}) \\ Var(N_i) &= \bar{\theta}_{i-1} E(N_i) + \theta_{i-1}^2 A_{i-1} E(N_{i-1}) + \theta_{i-1}^2 e^{2(d_{i-1} - p_{i-1}) r_{i-1}} Var(N_{i-1}) \\ Var(M_i) &= \bar{\theta}_{i-1} E(M_i) + \theta_{i-1}^2 \left(e^{d_{i-1} r_{i-1}} - e^{(d_{i-1} - p_{i-1}) r_{i-1}}\right)^2 Var(N_{i-1}) \\ &\quad + 2\,\theta_{i-1}^2 \left(e^{2 d_{i-1} r_{i-1}} - e^{2(d_{i-1} - p_{i-1}) r_{i-1}}\right) Cov(N_{i-1}, M_{i-1}) \\ &\quad + \theta_{i-1}^2 B_{i-1} E(M_{i-1}) + \theta_{i-1}^2 C_{i-1} E(N_{i-1}) \\ &\quad + \theta_{i-1}^2 e^{2 d_{i-1} r_{i-1}} Var(M_{i-1}) \\ Cov(N_i, M_i) &= \left(e^{(2 d_{i-1} - p_{i-1}) r_{i-1}} - e^{2(d_{i-1} - p_{i-1}) r_{i-1}}\right) Var(N_{i-1}) \\ &\quad + \theta_{i-1}^2 D_{i-1} E(N_{i-1}) + \theta_{i-1}^2 e^{(2 d_{i-1} - p_{i-1}) r_{i-1}} Cov(M_{i-1}, N_{i-1}) \end{aligned}$$

where A_{i-1} and D_{i-1} are obtained from respectively, formulae (A1.6) and (A1.8) with λ_{i-1}, μ_{i-1}, p_{i-1} and r_{i-1} in place of λ, μ, p and t, and with $N_0 = M_0 = 1$; and B_{i-1} and C_{i-1} are the coefficients of M_0 and N_0 respectively in (A1.7), with λ, μ, p and t similarly replaced.

These equations are all of the form $a_i = \gamma_i a_{i-1} + \tau_i$, which has the solution

$$a_i = \sum_{j=2}^{i} \left(\prod_{m=j+1}^{i} \gamma_m\right) \tau_j + \left(\prod_{j=2}^{i} \gamma_j\right) a_1$$

where the first product is defined as unity for $j = i$. The above difference equations may be solved in this way, yielding successively $E(N_i)$, $E(M_i)$, $Var(N_i)$, $Cov(N_i, M_i)$ and $Var(M_i)$.

The moments of (N_t, M_t) for $t_i < t \leqslant t_{i-1}'$ can be obtained by letting $r_i' = t - t_i$ and $\theta_i' = 1$; then the moments of (N_t, M_t) are the same as those of (N_{i+1}, M_{i+1}) when r_i and θ_i are replaced by r_i' and θ_i' respectively.

In Section 3, we discussed the assumptions which lead to the distribution of M (the number of mutant colonies counted at the end of selection) as derived by Tan (1981); Section 4 introduced a normal approximation for this distribution. In Section 5 it was suggested that "standard likelihood methods", based on the normal approximation, be employed to carry out the estimation and hypothesis testing for unknown parameters, using the counts $(M^{c,d}, T^{c,d})$, $c = 0,1,\ldots,C$ and $d = 1,\ldots,D_c$. While the accuracy in small samples of some of these methods was validated by exact calculation and Monte Carlo simulation in Sections 5 and 6, it is still important to check that the methods satisfy the asymptotic properties ascribed to them.

In particular, we would like to confirm the asymptotic normality of maximum likelihood estimates (m.l.e.'s) of the mutation parameters and the asymptotic χ^2 distribution of the likelihood ratio statistic for testing various hypotheses about the parameters, when the m.l.e.'s are obtained by maximizing the likelihood based on the normal approximation to the iteratively defined distribution for M derived by Tan (1981). To this end, we will prove the following result:

Theorem 1: Let $y_{c,d}$ have the density function $f_c(y_{c,d}; \boldsymbol{\omega})$ with respect to Lebesgue measure on R, where

$$f_c(y_{c,d}; \boldsymbol{\omega}) = [2\pi V_c(\boldsymbol{\omega})]^{-1/2} \exp\left\{ - \frac{[y_{c,d} - Y_c(\boldsymbol{\omega})]^2}{2 V_c(\boldsymbol{\omega})} \right\}$$

for $c = 0,1,\ldots,C$ and $d = 1,\ldots,D_c$, where Y_c and V_c are known functions of the unknown k-dimensional parameter $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_k)^T$, and $k \leqslant C$. We also suppose

(a) $\boldsymbol{\omega}$ is contained in Ω, an open interval in k-dimensional space.

(b) Y_c and V_c are twice differentiable functions of $\boldsymbol{\omega}$ on Ω, with continuous second derivatives; $V_c(\boldsymbol{\omega}) \neq 0$ for all $\boldsymbol{\omega}$ in Ω.

(c) The information matrix J, whose r,s^{th} element is

$$j_{r,s} = - E\left(\frac{\partial^2 l}{\partial \omega_r \, \partial \omega_s} \right),$$

is positive definite for all $\boldsymbol{\omega}$ in Ω, where the log-likelihood

$$l(\boldsymbol{\omega}) = \sum_{c=0}^{C} \sum_{d=1}^{D_c} \log f_c(y_{c,d}; \boldsymbol{\omega}).$$

(d) $\lim_{D \to \infty} \frac{D_c}{D} = \eta_c > 0$, where $D = \sum_{c=0}^{C} D_c$.

Then, under (a) - (d), the maximum likelihood equations $\partial l / \partial \boldsymbol{\omega} = \mathbf{0}$ have a unique root $\hat{\boldsymbol{\omega}}$ which is strongly consistent for $\boldsymbol{\omega}_0$, the true value of the parameter, as $D \to \infty$. Furthermore, $D^{1/2}(\hat{\boldsymbol{\omega}} - \boldsymbol{\omega}_0)$ is normally distributed with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma} = \mathbf{J}^{-1} \mid \boldsymbol{\omega}_0$.

The proof of Theorem 1 is via standard techniques (see for example Rai and Van Ryzin, 1981). We first show that the likelihood equation has a consistent root. Suppose $\boldsymbol{\omega}^*$ is contained in $B = \{\boldsymbol{\omega} : \| \boldsymbol{\omega} - \boldsymbol{\omega}_0 \| = \delta\}$, where the metric is the standard Euclidean. By the result 1*e*6.6 in Rao (1973, p. 59),

$$E\left[\log \left(\frac{f_c\,(y_{c,d}; \boldsymbol{\omega}^*)}{f_c\,(y_{c,d}; \boldsymbol{\omega}_0)} \right) \right] < 0,$$

i.e.

$$E\,[\log fc(y_{c,d}; \boldsymbol{\omega}^*)] < E\,[\log fc(y, d; \boldsymbol{\omega}_0)] = 0. \qquad \text{(A2.1)}$$

Therefore, by the Strong Law of Large Numbers,

$$\frac{1}{D_c}\sum_{d=1}^{D_c}\log f_c(y_{c,d};\,\boldsymbol{\omega}^*) < \frac{1}{D_c}\sum_{d=1}^{D_c}\log f_c(y_{c,d};\,\boldsymbol{\omega}_0)$$

as $D_c \to \infty$, with probability 1.

Moreover, asymptotically,

$$\begin{aligned}\frac{1}{D}\,l(\boldsymbol{\omega}^*) &= \sum_{c=0}^{C}\eta_c\left\{\frac{1}{D_c}\sum_{d=1}^{D_c}\log f_c(y_{c,d};\,\boldsymbol{\omega}^*)\right\}\\ &< \sum_{c=0}^{C}\eta_c\left\{\frac{1}{D_c}\sum_{d=1}^{D_c}\log f_c(y_{c,d};\,\boldsymbol{\omega}_0)\right\}\\ &= \frac{1}{D}\,l(\boldsymbol{\omega}_0)\end{aligned}$$

with probability 1, and by the compactness of B, it is clear that in fact

$$\sup_{\boldsymbol{\omega}^*\in B}\frac{1}{D}\,l(\boldsymbol{\omega}^*) < \frac{1}{D}\,l(\boldsymbol{\omega}_0)$$

as $D \to \infty$. Thus, since l is differentiable, it attains a local maximum at $\hat{\boldsymbol{\omega}}$ within B, at which $\frac{\partial l}{\partial \boldsymbol{\omega}} = \mathbf{0}$. Since δ can be made arbitrarily small, $\hat{\boldsymbol{\omega}} \to \boldsymbol{\omega}_0$ with probability 1. The matrix $\mathbf{H}$, whose r,s^{th} element is $h_{r,s} = \frac{\partial^2 l}{\partial\omega_r\,\partial\omega_s}$, is negative definite when evaluated at $\hat{\boldsymbol{\omega}}$, with probability 1.

The consistent root is asymptotically unique, since if there exists $\tilde{\boldsymbol{\omega}} \neq \hat{\boldsymbol{\omega}}$ such that $\frac{\partial l}{\partial \boldsymbol{\omega}} = \mathbf{0}$ is also satisfied by $\tilde{\boldsymbol{\omega}}$, $\mathbf{H}|_{\boldsymbol{\omega}^*}$ must have a zero determinant, for some $\boldsymbol{\omega}^* = \epsilon\tilde{\boldsymbol{\omega}} + (1-\epsilon)\hat{\boldsymbol{\omega}}$, where $0 \leqslant \epsilon \leqslant 1$ (this is a consequence of a multivariate form of Rolle's theorem, quoted in the Appendix of Rai and Van Ryzin (1981)). But then $\boldsymbol{\omega}^*$ is also consistent for $\boldsymbol{\omega}_0$, and thus $\mathbf{H}|_{\boldsymbol{\omega}^*}$ converges to a negative definite matrix. This contradiction proves the uniqueness.

To prove normality of $\hat{\boldsymbol{\omega}}$ under (a) - (d), we make the Taylor expansion

$$\mathbf{0} = \frac{\partial l}{\partial \boldsymbol{\omega}}\Big|_{\hat{\boldsymbol{\omega}}} = \frac{\partial l}{\partial \boldsymbol{\omega}}\Big|_{\boldsymbol{\omega}_0} + \mathbf{H}|_{\boldsymbol{\omega}^*}\,(\hat{\boldsymbol{\omega}} - \boldsymbol{\omega}_0).$$

where $\boldsymbol{\omega}^* = \epsilon\boldsymbol{\omega}_0 + (1-\epsilon)\hat{\boldsymbol{\omega}}$ for some $0 < \epsilon < 1$, which can be rewritten as

$$D^{1/2}(\hat{\boldsymbol{\omega}} - \boldsymbol{\omega}_0) = \left\{-\frac{1}{D}\,\mathbf{H}|_{\boldsymbol{\omega}^*}\right\}^{-1} D^{-1/2}\,\frac{\partial l}{\partial \boldsymbol{\omega}}\Big|_{\boldsymbol{\omega}_0}. \tag{A2.2}$$

Now

$$\frac{\partial \log f_c(y_{c,d};\,\boldsymbol{\omega})}{\partial \omega_r} = -\tfrac{1}{2}\left\{\frac{\partial V_c}{\partial \omega_r}\left[\frac{1}{V_c} - \frac{(y_{c,d}-Y_c)^2}{V_c^2}\right] - \frac{\partial Y_c}{\partial \omega_r}\left[\frac{2(y_{c,d}-Y_c)}{V_c}\right]\right\},$$

so

$$E\left\{\frac{\partial \log f_c(y_{c,d};\,\boldsymbol{\omega})}{\partial \omega_r}\right\}\Bigg|_{\boldsymbol{\omega}_0} = 0$$

Also,

$$E\left\{\frac{\partial \log f_c(y_{c,d};\,\boldsymbol{\omega})}{\partial \omega_r}\cdot\frac{\partial \log f_c(y_{c,d};\,\boldsymbol{\omega})}{\partial \omega_s}\right\}\Bigg|_{\boldsymbol{\omega}_0}$$

$$= \tfrac{1}{2}\left[\frac{\partial V_c}{\partial \omega_r}\frac{\partial V_c}{\partial \omega_s}\frac{1}{V_c^2} + \frac{\partial Y_c}{\partial \omega_r}\frac{\partial Y_c}{\partial \omega_s}\frac{2}{V_c}\right]$$

$$= -E\left\{\frac{\partial^2 \log f_c(y_{c,d};\, \boldsymbol{\omega})}{\partial \omega_r\, \partial \omega_s}\right\}\Bigg|_{\boldsymbol{\omega}_0};$$

Therefore, the Multivariate Central Limit Theorem applies to the vector $\mathbf{h}_c$, whose r^{th} component is

$$h_{cr} = D_c^{-1/2} \sum_{d=1}^{D_c} \frac{\partial \log f_c(y_{c,d};\, \boldsymbol{\omega})}{\partial \omega_r},$$

and therefore, as $D \to \infty$,

$$D^{-1/2} \frac{\partial l}{\partial \boldsymbol{\omega}}\Bigg|_{\boldsymbol{\omega}_0} = \sum_{c=0}^{C} \eta_c^{1/2}\, \mathbf{h}_{cr} \to N(0, \boldsymbol{\Sigma}) \tag{A2.3}$$

in distribution, since a finite linear combinations of sequences which converge to multivariate normality does likewise. We obtain the r,s^{th} element of $\boldsymbol{\Sigma}^{-1} = \mathbf{J}$ as

$$j_{r,s} = \tfrac{1}{2} \sum_{c=0}^{D_c} D_c\left\{\frac{\partial V_c}{\partial \omega_r}\frac{\partial V_c}{\partial \omega_s}\frac{1}{V_c^2} + \frac{\partial Y_c}{\partial \omega_r}\frac{\partial Y_c}{\partial \omega_s}\frac{2}{V_c}\right\}. \tag{A2.4}$$

We now rewrite

$$-D^{-1}\, \mathbf{H}|_{\boldsymbol{\omega}^*} = \left\{D^{-1}\, \mathbf{H}|_{\boldsymbol{\omega}_0} - D^{-1}\, \mathbf{H}|_{\boldsymbol{\omega}^*}\right\} - D^{-1}\, \mathbf{H}|_{\boldsymbol{\omega}_0} \tag{A2.5}$$

The strong consistency of $\hat{\boldsymbol{\omega}}$ and the continuity of $\mathbf{H}$ as a function of $\boldsymbol{\omega}$ (by (b)) imply that the term in brackets on the right of (A2.5) approaches zero almost surely as $D \to \infty$; the Strong Law of Large Numbers immediately gives

$$D^{-1}\, \mathbf{H}|_{\boldsymbol{\omega}_0} \to \boldsymbol{\Sigma} \tag{A2.6}$$

with probability 1. Combining (A2.2), (A2.3), (A2.5) and (A2.6) shows that

$$D^{1/2}(\hat{\boldsymbol{\omega}} - \boldsymbol{\omega}_0) \to N(0, \boldsymbol{\Sigma})$$

in distribution, as required; this concludes the proof.

We now return to the particular models under consideration in the text. Suppose once again that only the mutation parameters p_i^c, $c = 0,1,\ldots,C$ (or a smaller number of parameters of which the p_i^c are functions) are unknown; that for a fixed i and c, the sampling fractions $\theta_i^{c,d}$ do not depend on d; and that the normal approximation for the distribution of $M^{c,d}$ is in fact exact. Then the m.l.e.'s obtained by maximizing the likelihood based on the normal approximation to the distribution of $M^{c,d}$ will be asymptotically normal, provided the conditions of Theorem 1 are satisfied. Verification of the conditions (a), (b) and (c) must be done separately for each parametrization of p_i^c.

We consider estimation of the parameter vector $\boldsymbol{\omega}$, where $p_i^c = p_i^c(\boldsymbol{\omega})$ for $c = 0,1,\ldots,C$. In the case of the relationship (5.4), $k = 2$ and $\boldsymbol{\omega}^T = (\alpha, \beta)$. For the saturated model, $k = C+1$ and $\boldsymbol{\omega}^T = (p_1^0, p_1^1, \ldots, p_1^C)$. To satisfy condition (c) of Theorem 1, we need to show that the information matrix $\mathbf{J}$ is positive definite.

The information matrix may be written as

$$\mathbf{J} = \mathbf{F}_1\, \mathbf{F}_1^T + \mathbf{F}_2\, \mathbf{F}_2^T, \tag{A2.7}$$

where $\mathbf{F}_1 = \mathbf{G}_1\mathbf{K}_1$ and $\mathbf{F}_2 = \mathbf{G}_2\mathbf{K}_2$; here $\mathbf{G}_1$ and $\mathbf{G}_2$ are $k\times(C+1)$-dimensional matrices, whose r,c^{th} elements are respectively

$$(G_1)_{r,c} = \frac{\partial V_{c-1}}{\partial \omega_r}, \quad (G_2)_{r,c} = \frac{\partial Y_{c-1}}{\partial \omega_r},$$

and $\mathbf{K}_1$ and $\mathbf{K}_2$ are diagonal $(C+1)\times(C+1)$ matrices with c^{th} diagonal elements $D_{c-1}^{1/2} / (V_{c-1}\sqrt{2})$ and $(D_{c-1} / V_{c-1})^{1/2}$, respectively. Now since $\mathbf{K}_1$ and $\mathbf{K}_2$ are non-singular when V_c is defined and non-zero for $c = 0,1,\ldots,C$, $\mathbf{F}_1\mathbf{F}_1^T$ and $\mathbf{F}_2\mathbf{F}_2^T$ are positive definite if, respectively, $\mathbf{G}_1$ and $\mathbf{G}_2$ are of full rank k. This follows by an argument used in Lemma 3 of the Appendix of Krewski and Van Ryzin (1981) [A result in Rao (1973, p. 30) gives rank(**AB**) = rank(**A**) for **B** square and non-singular; Searle (1971, p. 37) shows that the product of a matrix of full row rank with its transpose is positive definite.]. Thus if $\mathbf{G}_1$ and $\mathbf{G}_2$ are of full rank k, **J** being the sum of two positive definite matrices will likewise be positive definite. We have thus proven

Lemma 1: If $\mathbf{G}_1$ and $\mathbf{G}_2$ in the decomposition (A2.7) for **J** are both of rank k, **J** is positive definite.

For the saturated model, $Y_c = E(M^{c,d})$ and $V_c = Var(M^{c,d})$ are linear combinations of p_1^0 and p_1^c, and do not involve $p_1^{c'}$ for $c' \neq 0$ and $c' \neq c$ (see equations (3.4) and (3.5)). Therefore, $(G_1)_{rc} = (G_2)_{rc} = 0$ for $r > 1$ and $r \neq c$, but $(G_1)_{1c} > 0$ and $(G_2)_{1c} > 0$ for $c = 1,\ldots,C+1$; and $(G_1)_{cc} > 0$ and $(G_2)_{cc} > 0$ for $c = 1,\ldots,C+1$, provided $p_1^c > 0$, $c = 0,1,\ldots,C+1$. Thus $\mathbf{G}_1$ and $\mathbf{G}_2$ are of the form

$$\begin{pmatrix} w & \mathbf{X}^T \\ 0 & \mathbf{Z} \end{pmatrix}$$

where **Z** is diagonal with non-zero diagonal elements, $w > 0$, and the elements of the vector **X** are positive; and are clearly of full rank $k = C+1$. Thus the theorem applies, for $\boldsymbol{\omega} = (p_1^0, p_1^1, \ldots, p_1^C)$ contained in $\boldsymbol{\Omega}$, the positive quadrant of $(C+1)$-dimensional space.

For any model of p_1^c as a function of x_c, the cth dose of test chemical, we can use results on Tchebycheff sets of functions, also quoted in the Appendix of Krewski and Van Ryzin (1981). Specifically, we will need

Lemma 2: A set of functions $\{\phi_r(x),\ r = 1,\ldots,k\}$ is a Tchebycheff set on $[0, X]$ (i.e. $\sum_{r=1}^{k} a_r \phi_r$ has at most $k-1$ zeroes on $[0, X]$ for all $\mathbf{a} = (a_1,\ldots,a_k)^T \neq \mathbf{0}$) if and only if the matrix whose r,c^{th} element is $\phi_r(x_c)$ is of full rank for every set of k distinct points $\{x_c\}$ in $[0, X]$.

By Lemmas 1 and 2, a sufficient condition for **J** to be positive definite for a particular choice of functional relationship between p_1^c and x_c is that under the relationship,

$$V_1 = \left\{\frac{\partial Y_c}{\partial \omega_1}, \ldots, \frac{\partial Y_c}{\partial \omega_k}\right\} \text{ and } V_2 = \left\{\frac{\partial V_c}{\partial \omega_1}, \ldots, \frac{\partial V_c}{\partial \omega_k}\right\}$$

are Tchebycheff sets of functions of x on $[0, X]$. Suppose for example that

$$p_1(x) = \alpha + \sum_{r=1}^{n} \beta_r x^r, \tag{A2.8}$$

a polynomial in dose. Now from (3.4) and (3.5), $Y_c = E(M^{c,d}) = \xi_1 p_1^0 + \psi_1(x_c)\, p_1^c$ and $V_c = Var(M^{c,d}) = \xi_2 p_1^0 + \psi_2(x_c)\, p_1^c$, where ξ_1 and ξ_2 are positive constants, and ψ_1 and ψ_2 are positive and (possibly) functions of dose x_c. Thus, $k = n+1$, and

$$\frac{\partial Y_c}{\partial \alpha} = \xi_1 + \psi_1, \quad \frac{\partial Y_c}{\partial \beta_r} = \psi_1 x_c^r, \quad r=1,\ldots,n$$

$$\frac{\partial V_c}{\partial \alpha} = \xi_2 + \psi_2, \quad \frac{\partial V_c}{\partial \beta_r} = \psi_2 x_c^r, \quad r=1,\ldots,n.$$

As noted by Krewski and Van Ryzin, $\{1,x,...,x^n\}$ is a Tchebycheff set on any interval $[0, X]$, and a Tchebycheff set multiplied by a positive function (i.e. ψ_1 and ψ_2) is still a Tchebycheff set. Since adding a constant (i.e. ξ_1 and ξ_2) to one of the functions of a Tchebycheff set also maintains the property, the sets V_1 and V_2 are Tchebycheff sets, and the information matrix J is positive definite under the polynomial relationship (A2.8), provided $n+1 \leqslant C+1$. If we restrict $\boldsymbol{\omega} = (\alpha_1, \beta_1, \ldots, \beta_k)$ to the positive quadrant of $(n+1)$-dimensional space, the conditions of Theorem 1 are again satisfied. Thus under the stated conditions, the m.l.e. of $\boldsymbol{\omega}$ is asymptotically normal as $D \rightarrow \infty$, and the result holds for the linear relationship (5.4) as a special case.

We now briefly indicate other conditions under which it would be useful to prove asymptotic normality of the m.l.e.'s, and where possible, suggest how this may be done.

(i) If other parameters such as the toxicity parameters $\{\mu_1^c, c = 1,...,C\}$ are in fact unknown, the asymptotic normality of the mutation parameter estimates will hold , as long as a $\sqrt{n}$-consistent estimate of the parameters is available. For the toxicity parameters, such an estimate is provided by formula (5.6), for example.

(ii) We have assumed in effect that the observations comprise a fixed number $(C+1)$ of independently and identically distributed (i.i.d.) samples of size D_c, $c = 0,1,...,C$. This has been achieved by assuming that the number of doses is fixed as $D \rightarrow \infty$, and that within dose levels, the sampling fractions $\theta_i^{c,d}$ at the end of the ith stage do not change with d. If the first of these assumptions is relaxed, we require conditions on the first and second derivatives of the likelihood under which the Strong Law of Large Numbers and the Central Limit Theorem for not necessarily i.i.d. random variables can be applied. We can, however, relax the second assumption and remain in a situation of having a fixed number of i.i.d. samples, by supposing, for example, that $\theta_i^{c,d}$ is distributed as $N(\theta_i^c, \sigma_{ic}^2)$, where θ_i^2 and σ_{ic}^2 are either known or may be estimated consistently.

(iii) If the true value of one of the parameters lies on a boundary of the parameter space such as would be the case if $\beta = 0$ in (5.4), the asymptotic distribution of the m.l.e. for that parameter is a truncated normal distribution, with half its mass lying at the boundary point (see Moran, 1971). Although a careful proof of this assertion in the present context has not been attempted here, the model is sufficiently regular to suggest that it could be accomplished straightforwardly.

(iv) Perhaps of most theoretical interest would be to show that the m.l.e.'s obtained using the likelihood based on the normal approximation also have the asymptotic normal distribution when the true distribution of the observations $M^{c,d}$ is given by the density of Tan (1981). Huber (1967) has shown that m.l.e.'s from a family of densities which does not include the true density of the observations are asymptotically normal, if the true density is sufficiently regular. It will be a subject of further research to prove such a result in this case.

Having verified asymptotic normality of parameter estimates, the asymptotic χ^2 distribution of the likelihood ratio can be obtained for parameter values within $O(D^{-1/2})$ of the true value, by the argument used in Cox and Hinkley (1974, p. 322).

Appendix 3

The following result justifies the use of the isotonic regression test as discussed in Section 5.1.

Theorem 2: Let $y_{c,d}$, $c = 0,1,...,c$, $d = 1,...,D_c$ be independent, with finite mean Y and finite, non-zero variance σ_c^2. Also let $D_c = \eta_c D$, where $0 < \eta_c < 1$, and suppose s_c^2 is a consistent estimate of σ_c^2.

For a vector $\boldsymbol{\nu}^T = (\nu_1, \ldots, \nu_n)$, let $(\dot{\nu}_1, \ldots, \dot{\nu}_n)$ denote the isotonic regression estimate of $\boldsymbol{\nu}$ for weights $\boldsymbol{\omega}^T = (\eta_0/s_0^2, \ldots, \eta_C/s_C^2)$, with respect to a partial ordering of the set $\{0,1,...,C\}$.

Then

$$\lim_{D\rightarrow\infty} Pr\left\{\sum_{c=0}^{C} D_c\,(\dot{\bar{y}}_c - \bar{y})^2 / s_c^2 \geqslant u\right\}$$

$$= \sum_{k=1}^{C} P(k, C; \boldsymbol{\omega})\ Pr(\chi^2_{k-1} \geqslant u),$$

where $\bar{y}_c = \frac{1}{D_c} \sum_{d=1}^{D_c} y_{c,d}$, $\bar{y} = \frac{1}{D} \sum_{c=0}^{C} \sum_{d=1}^{D_c} y_{c,d}$, χ^2_l is a chi-square random variable on l d.f., and $P(t, r; n)$ is given in Barlow, *et al.* (1972, p. 126).

This theorem is a generalization of the theorem in Appendix 2 of Collings, *et al.* (1981), in which σ_c^2 does not depend on c, and is hence not used in the weights for the isotonic regression. By analogy with the proof of those authors, we write

$$\sum_{c=0}^{C} D_c\ (\dot{\bar{y}}_c - \bar{y})^2 / s_c^2 = \sum_{c=0}^{C} \eta_c \left\{ D_c^{1/2} \left(\frac{\bar{y}_c \dot{-} \bar{y}}{\sigma_c} \right) \right\}^2 \sigma_c^2 / s_c^2$$

The consistency of s_c^2 implies both that $\sigma_c^2/s_c^2 \rightarrow 1$ and that the weights $\eta_c/s_c^2 \rightarrow \eta_c/\sigma_c^2$ in probability. The rest of the argument used by Collings, *et al.* (1981) now applies to give the required result.

For the special case $C = 1$, the test reduces to a one-sided two sample test, and for the particular application in Section 5.1, the test statistic Θ is equivalent to the statistic Γ for testing H_{0c} against H_{1c}.

In order to apply these tests to assay data, we must suppose that $\theta_i^{c,d}$ does not depend on d (or, in effect, that $Var(M^{c,d})$ does not depend on d). Alternatively, $\theta_i^{c,d}$ may be considered to be random, as in point (ii) at the end of Appendix 2.

Appendix 4

This appendix discusses the program used to calculate maximum likelihood estimates of the mutation parameters p_i^c, or the parameters α and β in the relationship (5.4). The following notes explain particular aspects of the program.

(i) Newton-Raphson iteration is used to search for the maximum likelihood estimates. Thus the kth step in the iteration consists of evaluating

$$\boldsymbol{\omega}_{k+1} = \boldsymbol{\omega}_k - \mathbf{H}_k^{-1} \frac{\partial l}{\partial \boldsymbol{\omega}_k},$$

where $\boldsymbol{\omega}_k$ is the vector of unknown parameters, at the kth step in the iteration; $\partial l / \partial \boldsymbol{\omega}_k$ is the vector of derivatives of the likelihood with respect to the parameter vector evaluated at $\boldsymbol{\omega}_k$; and $\mathbf{H}_k$ is the matrix of second derivatives evaluated at $\boldsymbol{\omega}_k$.

(ii) Starting values $\boldsymbol{\omega}_1$ are needed to use the Newton-Raphson procedure. For the saturated model, these are obtained from formulae (5.5) and (5.6). For the linear model, we make use of the fact that

$$E(M^{c,d}) = K^{c,d} \left\{ (\alpha + \beta x_c) r_1 + \alpha \sum_{i=2}^{R+1} r_i \right\}$$

so that

$$E\left[\frac{M^{c,d}}{K^{c,d} \sum_{i=1}^{R+1} r_i} \right] = \alpha + \beta\ r_1\ x_c / \sum_{i=1}^{R+1} r_i \qquad \text{(A4.1)}$$

Then defining $y^{c,d}$ to be the expression inside the braces on the left hand side of (A4.1), and z^c to be the coefficient of β on the right, we can easily obtain the usual least squares estimates

$$\hat{\beta} = \frac{\sum_{c=0}^{C} \sum_{d=1}^{D_c} (z^c - \bar{z})\,(y^{c,d} - \bar{y})}{\sum_{c=0}^{C} D_c\,(z^c - \bar{z})^2},$$

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{z},$$

where $\bar{y} = \frac{1}{D} \sum_{c=0}^{C} \sum_{d=1}^{D_c} y^{c,d}$, and $\bar{z} = \frac{1}{D} \sum_{c=0}^{C} D_c\, z^c$. We use these estimates as starting values.

(iii) Boundary conditions on the parameters are dealt with by resetting a parameter to the boundary if it exceeds it at a step in the iteration. If the parameter returns to the boundary a certain number of times (we use 10) it is fixed at the boundary.

(iv) Because the normal distribution is defined by its mean and variance, and because $E(M^{c,d})$ and $Var(M^{c,d})$ are linear functions of the mutation parameters with coefficients which are known (or at least, estimated from other data), fitting alternative models is relatively straightforward. Once the coefficients have been evaluated, they are stored and can thereafter be used to construct the mean and variance and hence the likelihood and its derivatives, as functions of particular values of the mutation parameters.

Table 1

*Characteristics of the distribution of (N_t, M_t) under A1 - A6, for various parameter values**

Parameter Values				$E(M_t)$ †	$V(M_t)$	$Sk(M_t)$	$E(N_t)$	$CV(N_t)$	Cor
t	N_0	μ/λ	p/λ						
4	6×10^6	0	10^{-5}	17.4802	21.8520	.3213	7.6×10^6	.0002	.0004
				17.4803	21.8520	.3213			
		0	10^{-3}	1747.8243	2185.0049	.0321	7.6×10^6	.0002	.0036
				1747.0264	2185.2029	.0321			
		.1	10^{-5}	17.0807	21.3186	.3244	7.4×10^6	.0002	.0004
				17.0808	21.3186	.3244			
		1	10^{-5}	13.8664	17.0719	.3539	6.0×10^6	.0003	.0005
				13.8665	17.0719	.3539			
		10	10^{-3}	173.1446	195.4466	.0897	7.4×10^5	.0012	.0050
				173.1646	195.4608	.0897			
24	3×10^6	0	10^{-6}	16.6615	55.4908	.8267	1.2×10^7	.0050	.0007
				16.6615	55.4908	.8267			
		0	10^{-5}	166.6137	554.9085	.2614	1.2×10^7	.0050	.0022
				166.6148	554.9081	.2614			
48	3×10^6	0	10^{-5}	1334.1077	13120.0860	.2177	4.8×10^7	.0006	.0046
				1334.1262	13119.8752	.2177			

* The column headings are, respectively, the mean, variance and coefficient of skewness of M_t, the mean and coefficient of variation of N_t, and the correlation between N_t and M_t.

† In the first three columns, the upper of each pair is the approximate quantity obtained under assumptions A1-A6, while the lower is the same quantity obtained under A1-A6 and (3.1), neglecting terms which are $o(1/N_0)$.

$\lambda = .0578$ and $M_0 = 0$ in all cases.

Table 2

*Type I error rates for the test of H_{0c} against the alternative H_{1c} using the test statistics Γ and Λ at nominal size ν **

(A) Exact calculation, with only $p_1^0 = p_1^f = p$ unknown

p	S_μ	$\nu = .05$		$\nu = .01$		Fraction of Distribution Evaluated
		Γ	Λ	Γ	Λ	
2.5×10^{-7}	1.0	.053	.051	.009	.009	.9997
	.5	.052	.049	.010	.010	.9997
	.1	.050	.042	.011	.010	.9998
7.5×10^{-7}	1.0	.054	.051	.011	.010	.9996
	.5	.053	.050	.011	.010	.9996
	.1	.052	.044	.011	.010	.9995

(B) Estimated error rates, with only $p_1^0 = p_1^f = p$ and μ_1^f unknown; $p = 7.5\times10^{-7}$, $S_\mu = 0.1$

$\nu = .05$		$\nu = .01$	
Γ	Λ	Γ	Λ
.051	.044	.012	.011

*Protocol variables and parameter values not specified in the Table are as in (5.1). See text for definition of symbols.

Table 3

*Estimated Type I error rates for testing H_0: $\beta = 0$ against the alternative H_1: $\beta > 0$ using the normal approximation, with nominal size ν **

(A) Only α and β unknown

Parameters	C	D_c	$\nu = .05$ H1	$\nu = .05$ H2	$\nu = .025$ H1	$\nu = .025$ H2	$\nu = .01$ H1	$\nu = .01$ H2
$\alpha = 7.5\times10^{-7}$	10	2	.045	.053	.021	.028	.008	.013
$\mu_1^c = 0$	5	2	.039	.049	.018	.025	.007	.012
	5	1	.032	.046	.015	.025	.006	.012
$\alpha = 7.5\times10^{-7}$	10	2	.036	.054	.018	.031	.006	.015
$\mu_1^c = .005x_c$	5	2	.024	.047	.011	.029	.003	.014
	5	1	.021	.045	.009	.025	.003	.013

(B) Method H2, with only α, β and μ_1^c, $c = 1,\ldots,C$ unknown; $\alpha = 7.5\times10^{-7}$, $\mu_1^c = .005x_c$.

C	D_c	$\nu = .05$	$\nu = .025$	$\nu = .01$
10	2	.055	.032	.013
5	2	.048	.029	.015
5	1	.046	.025	.014

*Protocol variables and parameter values not specified in the Table are as in (5.1). See text for definition of symbols.

Table 4

*Estimated coverage probabilities of the confidence interval for β, using the normal approximation, with nominal coverage $1-\nu$ **

(A) Only α and β unknown; $\alpha = 7.5\times10^{-7}$, $\beta = 1.0\times10^{-7}$, $\mu_1^c = .005x_c$

C	D_c	$1-\nu = .95$ C1	$1-\nu = .95$ C2	$1-\nu = .975$ C1	$1-\nu = .975$ C2	$1-\nu = .99$ C1	$1-\nu = .99$ C2
10	2	.933	.937	.965	.970	.985	.987
5	2	.934	.943	.960	.972	.980	.988
5	1	.926	.946	.954	.971	.974	.988

(B) Method C2, with only α, β and μ_1^c, $c = 1,\ldots,C$ unknown; $\alpha = 7.5\times10^{-7}$, $\beta = 1.0\times10^{-7}$, $\mu_1^c = .005x_c$

C	D_c	$1-\nu = .95$	$1-\nu = .975$	$1-\nu = .99$
10	2	.945	.975	.990
5	2	.944	.974	.989
5	1	.950	.970	.986

*Protocol variables and parameters not specified in the Table are as in (5.1). See text for definitions of symbols.

Table 5

*True and approximate power of the tests of H_{0c} against the alternative H_{1c}, using the test statistics Γ and Λ with nominal size ν **

(A) Exact calculation for "True", with only p_1^0 and p_1^c unknown.

p_1^0	K_p	S_μ	$\nu = .05$				$\nu = .01$				Fraction of distribution Evaluated
			Γ		Λ		Γ		Λ		
			t†	a	t	a	t	a	t	a	
2.5×10^{-7}	2	1.0	.082	.079	.080	.078	.019	.018	.018	.018	.9998
		.5	.081	.079	.077	.078	.019	.019	.018	.018	.9998
		.1	.078	.081	.067	.076	.019	.020	.019	.017	.9997
	5	1.0	.219	.218	.215	.211	.074	.075	.070	.069	.9998
		.5	.215	.218	.207	.205	.073	.076	.069	.066	.9998
		.1	.194	.211	.173	.175	.069	.079	.066	.053	.9997
	10	1.0	.546	.542	.539	.523	.294	.296	.285	.266	.9997
		.5	.531	.532	.519	.496	.284	.291	.275	.245	.9997
		.1	.445	.465	.416	.371	.232	.261	.225	.155	.9997
7.5×10^{-7}	2	1.0	.110	.106	.106	.104	.029	.027	.027	.026	.9996
		.5	.109	.106	.103	.103	.029	.027	.027	.026	.9995
		.1	.103	.104	.090	.096	.028	.028	.026	.023	.9994
	5	1.0	.432	.431	.422	.417	.202	.201	.196	.186	.9995
		.5	.421	.422	.409	.401	.198	.198	.189	.175	.9995
		.1	.360	.373	.333	.316	.165	.176	.156	.123	.9993
	10	1.0	.911	.909	.907	.898	.757	.754	.751	.721	.9994
		.5	.898	.895	.893	.875	.737	.733	.728	.679	.9994
		.1	.803	.795	.784	.709	.604	.607	.591	.448	.9992

(B) Estimates from simulations for "True", with only p_1^0, p_1^c and μ_1^c unknown; $p_1^0 = 7.5\times10^{-7}$, $S_\mu = 0.1$

K_p	$\nu = .05$				$\nu = .01$			
	Γ		Λ		Γ		Λ	
	t	a	t	a	t	a	t	a
2	.098	.104	.086	.096	.029	.028	.027	.023
5	.352	.373	.324	.316	.156	.176	.144	.123
10	.809	.795	.788	.709	.613	.607	.600	.448

*Protocol variables and parameter values not specified in the Table are as in (5.1). See text for definition of symbols.

† t = true; a = approximate.

Table 6

*Approximate power of likelihood-ratio tests of the hypotheses H_0 and H_L at nominal size ν, for equal spacing and allocation of 24 experimental cultures on the dose range [0, 64]**

(A) $\alpha = 7.5\times10^{-7}$, $\beta = 2.5\times10^{-8}$, $f = 2\times10^{-8}$, $g = .08$

Number of Dose Levels	Arithmetic Scale				Logarithmic Scale			
	H_0		H_L		H_0		H_L	
	$\nu = .05$	$\nu = .01$	$\nu = .05$	$\nu = .01$	$\nu = .05$	$\nu = .01$	$\nu = .05$	$\nu = .01$
12	.47	.22	.12	.03	.40	.17	.11	.03
8	.49	.24	.17	.06	.43	.20	.15	.05
6	.52	.26	.23	.08	.47	.22	.17	.06
4	.57	.31	.39	.18	.54	.28	.14	.05
3	.63	.36	.53	.29	.63	.36	.09	.03
2	.76	.51	-	-	.76	.51	-	-

(B) $\alpha = 7.5\times10^{-7}$, $\beta = 5\times10^{-8}$, $f = 10^{-8}$, $g = 1$

Number of Dose Levels	Arithmetic Scale				Logarithmic Scale			
	H_0		H_L		H_0		H_L	
	$\nu = .05$	$\nu = .01$	$\nu = .05$	$\nu = .01$	$\nu = .05$	$\nu = .01$	$\nu = .05$	$\nu = .01$
12	.90	.72	.32	.14	.82	.59	.26	.10
8	.92	.76	.48	.25	.85	.64	.41	.20
6	.93	.79	.64	.40	.88	.69	.49	.26
4	.96	.86	.87	.69	.93	.79	.39	.19
3	.98	.91	.95	.85	.97	.89	.20	.07
2	.99	.97	-	-	.99	.97	-	-

*Protocol variables and parameter values not specified in the Table are as in (5.1). See text for definition of symbols.

Table 7

*Data and estimates from an assay of chemical X using the mouse lymphoma protocol**

Dose	Culture Size ($\times 10^{-6}$) t_3'	t_4'	$T^{c,d}$	$M^{c,d}$	$\hat{\pi}^c$	$\hat{d}_1^c$	$\hat{d}_3^c$
0	11.85	13.05	557	284	.916	.035	.060
	12.51	12.01	542	337			
3.9	10.47	12.01	497	395	.836	.0	.060
	10.62	13.05	506	396			
4.9	10.00	11.55	392	445	.666	-.027	.056
	9.00	11.55	407	419			
6.1	9.37	7.32	414	436	.643	-.049	.039
	8.04	8.11	358	432			
7.7	6.00	7.50	366	424	.613	-.164	.040
	4.96	8.34	370	475			
8.5	4.89	7.32	296	454	.517	-.210	.035
	4.24	6.67	324	539			
9.5	3.27	4.84	306	502	.484	-.288	.017
	3.42	4.23	275	504			

*For each dose, data from duplicate cultures are given in successive lines. See text for definition of symbols.

References

Abbondandolo, A. (1977). Prospects for evaluating genetic damage in mammalian cells in culture. *Mutation Research* 42, 279-298.

Amacher, D.E., Paillet, S.C., Turner, G.N., Ray, V.A. and Salsburg, D.S. (1980). Point mutations at the thymidine kinase locus in L5178Y mouse lymphoma cells. Test validation and interpretation. *Mutation Research* 72, 447-474.

Ames, B.N., McCann, J. and Yamasaki, E. (1975). Methods for detecting carcinogens and mutagens with the Salmonella/mammalian-microsome mutagenicity test. *Mutation Research* 31, 347-364.

Armitage, P. (1952). The statistical theory of bacterial populations subject to mutation. *Journal of the Royal Statistical Society,* Series B 14, 1-40.

Barlow, R.E., Bartholomew, D.J., Bremner, J.M. and Brunk, H.D. (1972). *Statistical Inference under Order Restrictions.* New York: Wiley.

Bartlett, M. (1978). *Introduction to Stochastic Processes,* 3rd Edition. Cambridge University Press.

Bernstein, L., Kaldor, J., McCann, J. and Pike, M.C. (1982). An empirical approach to the statistical analysis of mutagenesis data from the *Salmonella* test. *Mutation Research,* in press.

Chase, G.R. and Hoel, D.G. (1975). Serial dilutions: Error effects and optimal designs. *Biometrika* 62, 329-334.

Chernoff, H. (1953). Locally optimal designs for estimating parameters. *Annals of Mathematical Statistics* 24, 586-602.

Clive, D., Johnson, K.O., Spector, J.F.S., Batson, A.G. and Brown, M.M.M. (1979). Validation and characterization of the $L5178Y/TK^{+/-}$ mouse lymphoma mutagen/carcinogen assay system. *Mutation Research* 59, 61-108.

Collings, B.J., Margolin, B.H. and Oehlert, G.W. (1981). Analyses for binomial data, with application to the fluctuation test for mutagenicity. *Biometrics* 37, 775-794.

Cox, D.R. and Hinkley, D. (1974). *Theoretical Statistics.* London: Chapman and Hall.

Crump, K. and Hoel, D. (1974). Mathematical models for estimating mutation rates in cell populations. *Biometrika* 61, 237-252.

Fedorov, V.V. (1972). *Theory of Optimal Experiments.* Translated and edited by W.J. Studdess and E.M. Klimko. New York: Academic Press.

Fox, M. (1975). Factors affecting the quantitation of dose-response curves for mutation induction in V79 Chinese hamster cells after exposure to chemical and physical mutagens. *Mutation Research* 24, 449-466.

Grafe, A. and Vollmar, J. (1979). The collaborative Ames Test Study, 1977/1978, analyses and interpretation of results. *Proceedings of the Second Workshop on Bacterial In Vitro Mutagenicity Test Systems,* Mannheim, Germany, April 2-3, 1979.

Green, M.H.L., Muriel, W.J. and Bridges, B.A. (1976). Use of a simplified fluctuation test to detect low levels of mutagens. *Mutation Research* 38, 33-42.

Hsie, A.W., Brimer, P.A., Mitchell, T.J. and Goslee, D.G. (1975). The dose-response relationship for ethyl methane sulfonate-induced mutations at the hypoxanthine-guanine phosphoribosyl transferase locus in Chinese hamster ovary cells. *Somatic Cell Genetics* 1, 247-261.

Huber, P.J. (1967). The behavior of maximum likelihood estimates under nonstandard conditions. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability,* Volume 1, L. Lecam and J. Neyman (eds.), 221-233. Berkeley: University of California Press.

Huberman, B. and Sachs, L. (1974). Cell-mediated mutagenesis of mammalian cells with chemical carcinogens. *International Journal of Cancer* 13, 326-333.

Jagers, P. (1975). *Branching Processes with Biological Applications.* New York: Wiley.

Jagers, P. and Norrby, K. (1974). Estimation of the mean and variance of cycle times in cinemicrographically recorded cell populations during balanced exponential growth. *Cell Tissue Kinetics* 7, 201-211.

Karlin, S. and Taylor, H.M. (1975). *A First Course in Stochastic Processes,* 2nd Edition. New York: Academic Press.

Kendall, M.G. and Stuart, A. (1960). *The Advanced Theory of Statistics,* Volume II. New York: Hafner Publishing Company.

Krewski, D. and Van Ryzin, J. (1981). Dose-response models for quantal response toxicity data. To appear in *Current Topics in Probability and Statistics,* M. Csorgo, D. Dawson, J.N.K. Rao and E. Saleh (eds.), Amsterdam: North Holland.

Kubitschek, H. (1967). Cell generation times; ancestral and internal controls. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability,* Volume 4, L. Lecam and J. Neyman (eds.), 549-572, Berkeley: University of California Press.

Luria, S.E. and Delbruck, M. (1943). Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* 28, 491-511.

Margolin, B.H., Kaplan, N. and Zeiger, E. (1981). Statistical analysis of the Ames *Salmonella/* microsome mutagenicity test. *Proceedings of the National Academy of Sciences, U.S.A.* 78, 3779-3783.

Moran, P.A.P. (1971). Maximum likelihood estimation in non-standard conditions. *Proceedings of the Cambridge Philosophical Society* 70, 441-450.

Myers, L.E., Sexton, N.H., Southerland, L.I. and Woolff, T.J. (1981). Regression analysis of Ames Test data. *Environmental Mutagenesis* 3, 575-586.

Nelder, J.A. and Wedderburn, R.W.M. (1972). Generalized linear models. *Journal of the Royal Statistical Society,* Series A 135, 370-384.

O'Neill, J.P., Brimer, P.A., Machanoff, R., Hirsch, G.P. and Hsie, A.W. (1977). A quantitative assay of mutation induction at the hypoxanthine-quanine phosphoribosyl transferase locus in Chinese hamster ovary cells (CHO/HGPRT System): Development and definition of the system. *Mutation Research* 45, 91-101.

Rao, C.R. (1973). *Linear Statistical Inference and Its Application,* 2nd edition. New York: Wiley.

Searle, S.R. (1971). *Linear Models.* New York: Wiley.

Snee, R.D. and Irr, J.D. (1981). Design of a statistical method for the analysis of mutagenesis at the hypoxanthine-quanine phosphoribosyl transferase locus of cultured Chinese hamster ovary cells. *Mutation Research* 85, 77-93.

Tan W.Y. (1980). On distribution theories for the number of mutants in cell populations. Technical report 80-8, Department of Mathematical Science, Memphis State University.

Tan, W.Y. (1981). The probability distribution of the number of mutants in the CHO/HGPRT assay. Presented at the National Toxicology Program Workshop on the Statistical Analysis of Mutagenesis Assays. Chapel Hill, North Carolina, April 21-22, 1981.

Tarone, R.E. (1982). The use of historical control information in testing for a trend in proportions. *Biometrics* 38, 215-220.

Figure Legends

1. Schematic of mutagenesis assay protocol.
2. Contours of approximate power for (K_p, S_μ) using the test statistic Γ at nominal level .05. The protocol variables are in (5.1), $D_0 = D_c = 1$, and S_μ, the survival in the treated culture, is expressed as a percentage of control.
3. Approximate power of the test based on the statistic Γ at nominal level .05, as a function of K_p, for various values of S_μ. The protocol variables are as in (5.1), $D_0 = D_c = 1$, and S_μ is expressed as a percentage of control.
4. Linear relationship between dose x and $p_1(x)$ given by $\alpha + \beta x$; a non-linear relationship $a + bx + fe^{gx}$, whose line of best least squares fit is $\alpha + \beta x$.
   (A) $\alpha = 7.5\times10^{-7}$, $\beta = 2.5\times10^{-8}$, $f = 2\times10^{-8}$, $g = .08$.
   (B) $\alpha = 7.5\times10^{-7}$, $\beta = 5.0\times10^{-8}$, $f = 10^{-8}$, $g = 1$.
5. Power of the likelihood ratio tests for zero slope (H_0) and linear fit (H_L), as functions of the location of the middle dose in a 3-dose, equally replicated design. Protocol variables as in (5.1). H_0 tested at nominal level .01, H_L tested at nominal level .05.
   (A) Linear and non-linear alternative as in Figure 4A.
   (B) Linear and non-linear alternative as in Figure 4B.
6. Results of the analysis of the data in Table 7.
   (A) Maximum likelihood estimates of p_c^1, $c = 0,1,\ldots,C$ under the saturated model, and estimates of α and β under the linear model.
   (B) Observed (open circles and squares) and expected (solid circles and squares) mutant colony counts under the linear model.
7. Empirical (jagged line) and fitted (smooth line) cumulative distribution functions for the replicated mutant colony counts, plotted on a linear scale.
8. Empirical and fitted cumulative distribution functions for the replicated mutant colony counts, plotted on a normal probability scale.

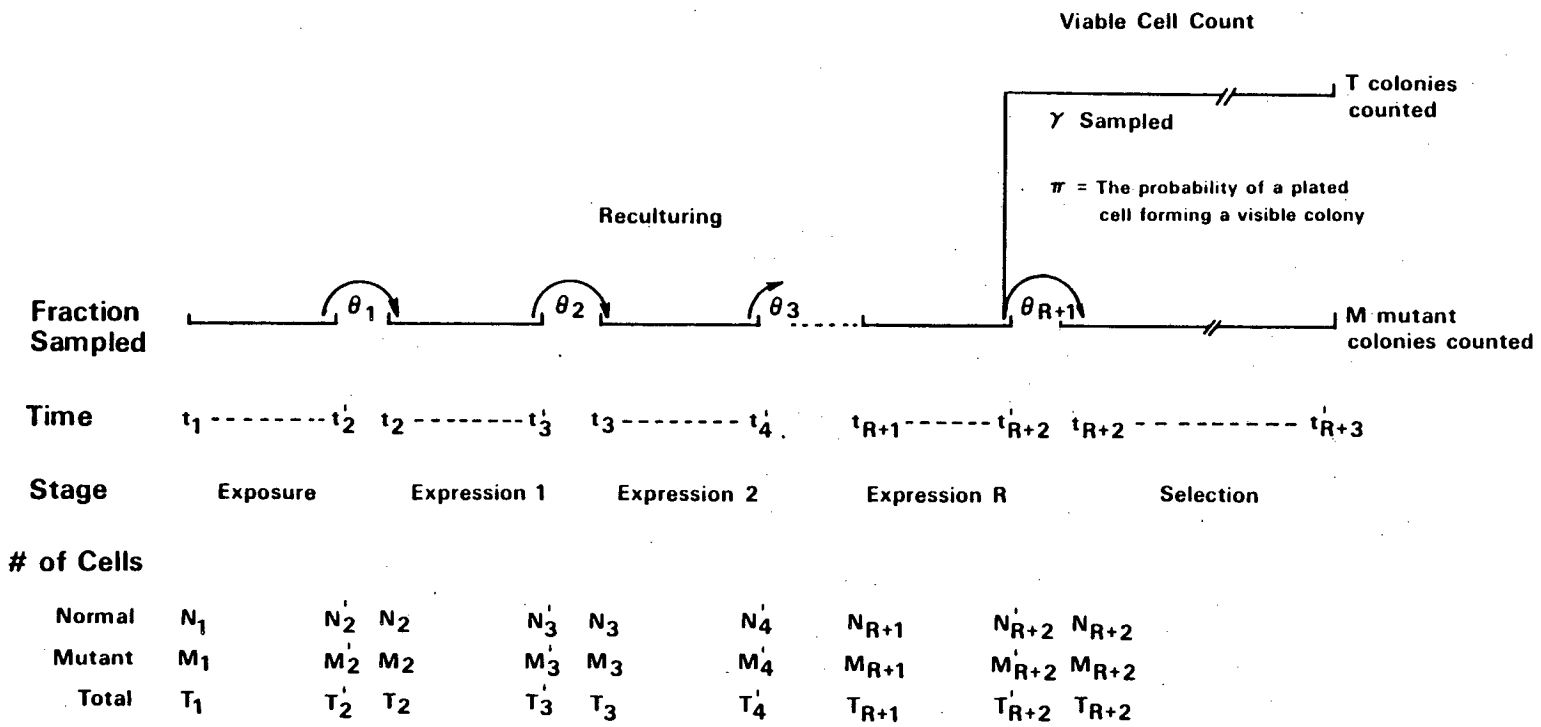

Viable Cell Count
T colonies counted
γ Sampled
π = The probability of a plated cell forming a visible colony
Reculturing
Fraction Sampled
θ1
θ2
θ3
θR+1
M mutant colonies counted
Time
t1 - - - - - - - - t'2 t2 - - - - - - - - t'3 t3 - - - - - - - - t'4 tR+1 - - - - - - t'R+2 tR+2 - - - - - - - - - t'R+3
Stage
Exposure
Expression 1
Expression 2
Expression R
Selection
# of Cells
Normal N1 N'2 N2 N'3 N3 N'4 NR+1 N'R+2 NR+2
Mutant M1 M'2 M2 M'3 M3 M'4 MR+1 M'R+2 MR+2
Total T1 T'2 T2 T'3 T3 T'4 TR+1 T'R+2 TR+2

Fig. 1

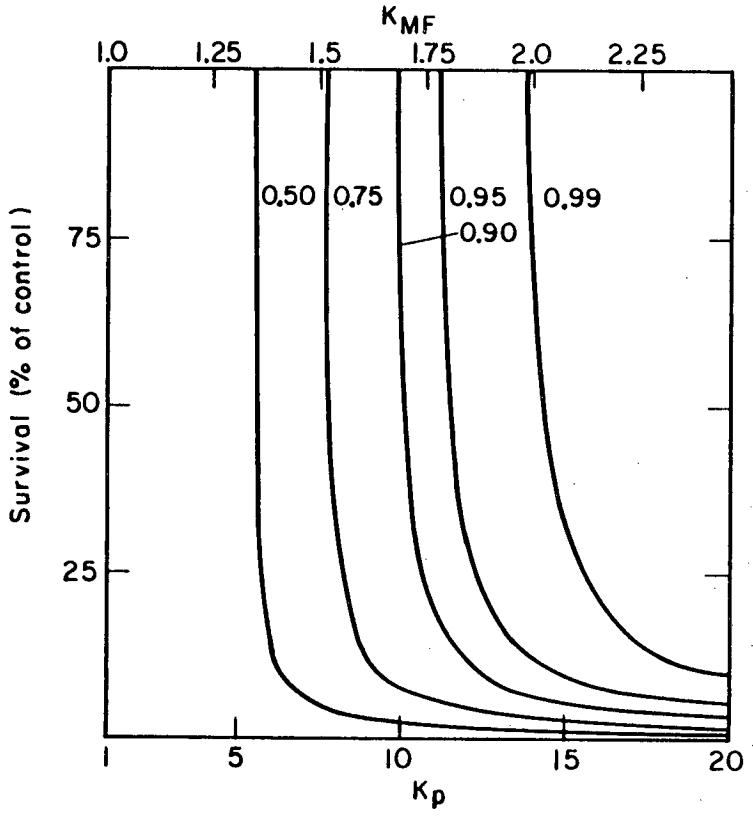

K_{MF}
1.0
1.25
1.5
1.75
2.0
2.25
0.50
0.75
0.95
0.90
0.99
Survival (% of control)
75
50
25
1
5
10
15
20
K_p

Fig. 2

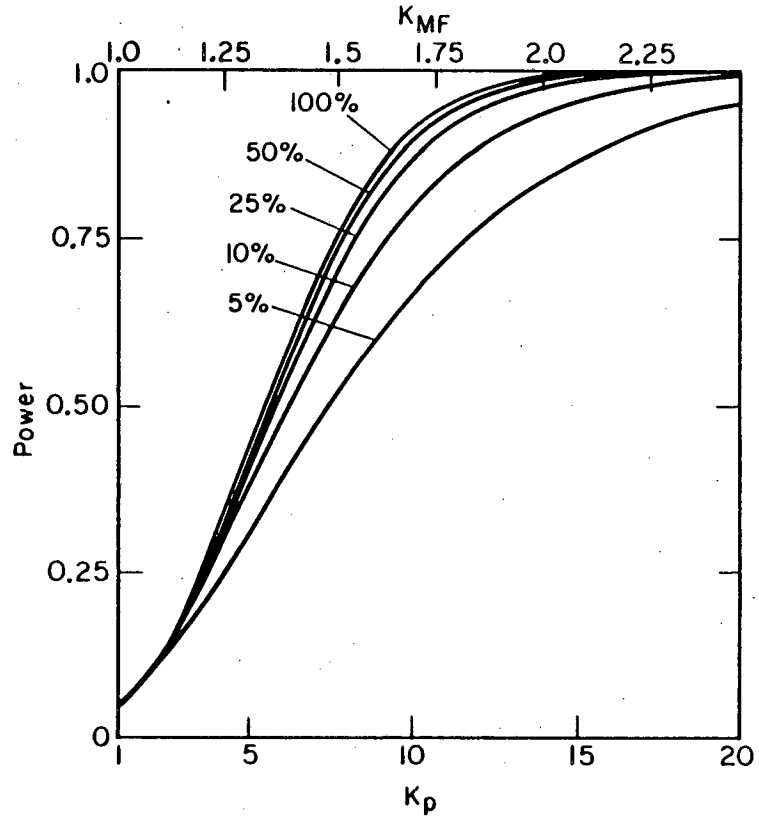

K_{MF}
1.0
1.25
1.5
1.75
2.0
2.25
100%
50%
25%
10%
5%
Power
1.0
0.75
0.50
0.25
0
1
5
10
15
20
K_p

Fig. 3

Fig. 4

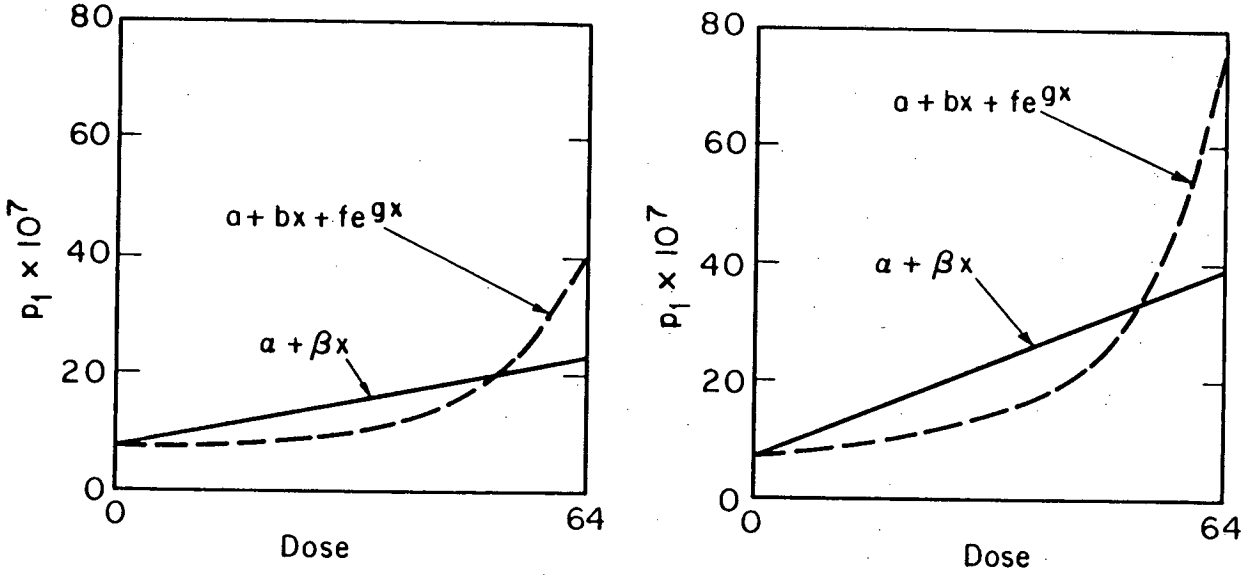
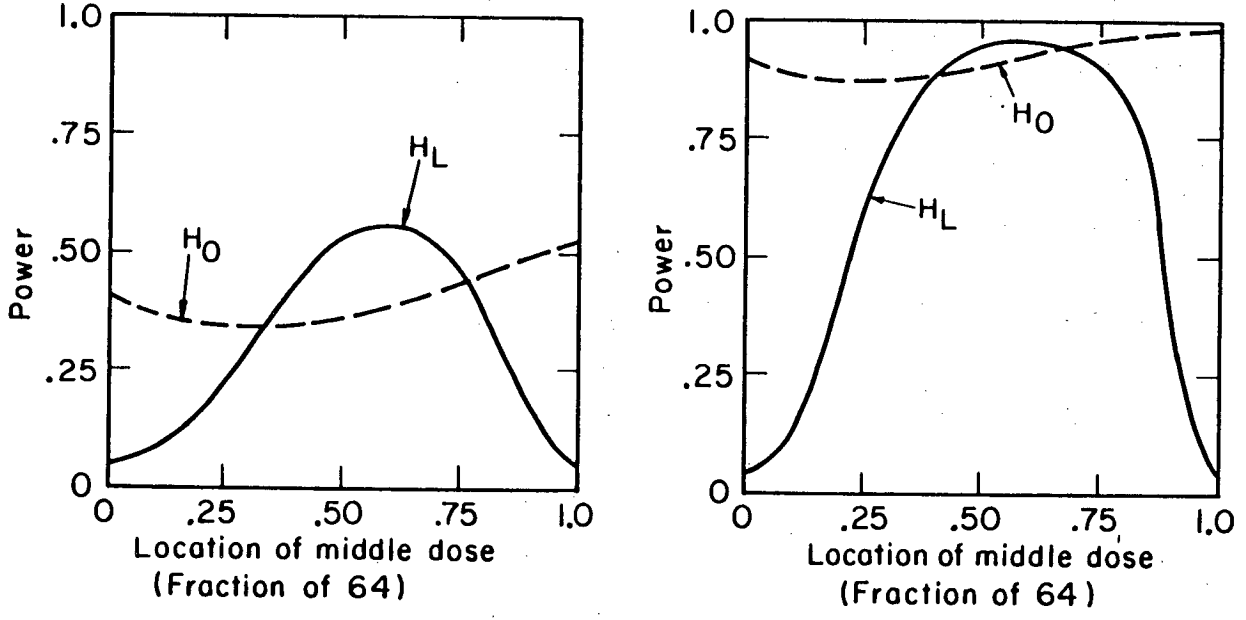
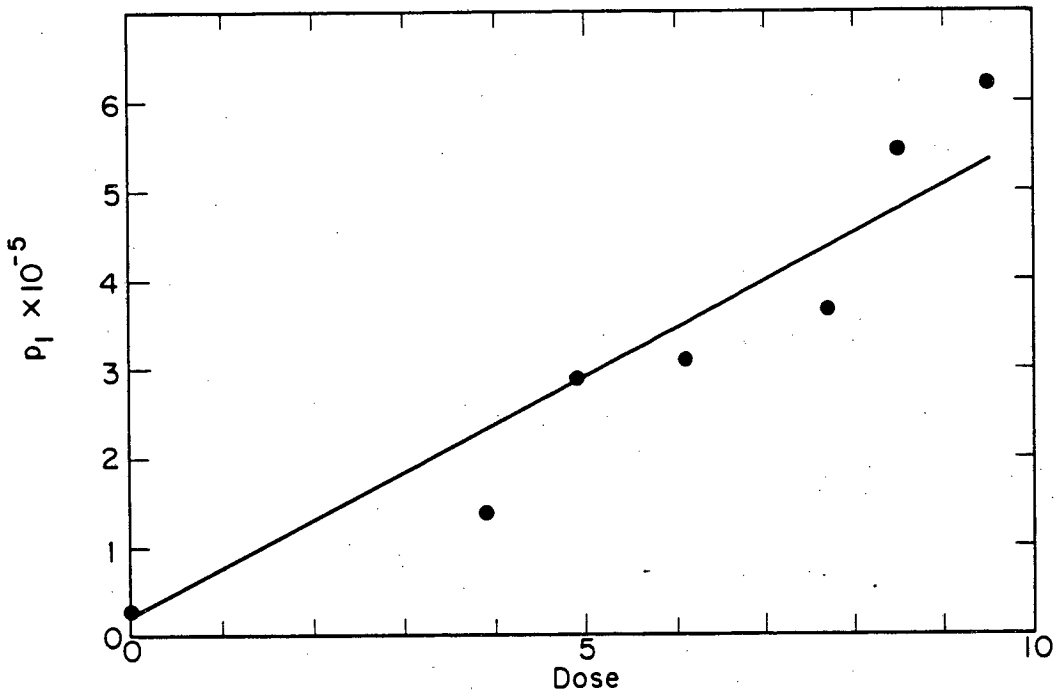

80
60
40
20
0
$p_1 \times 10^7$
$a + bx + fe^{gx}$
$a + \beta x$
0
64
Dose
80
60
40
20
0
$p_1 \times 10^7$
$a + bx + fe^{gx}$
$a + \beta x$
0
64
Dose

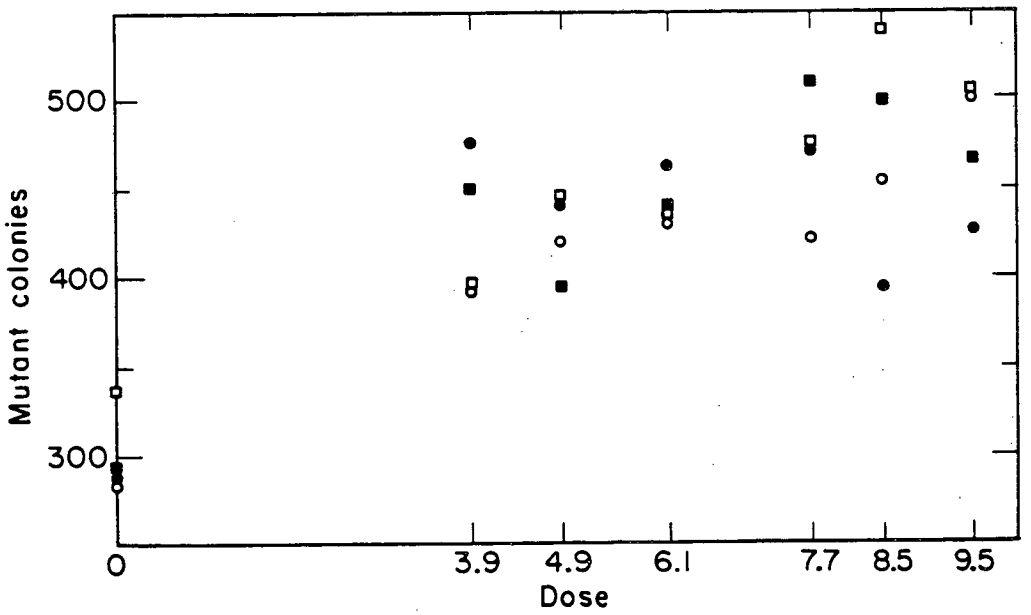
Fig. 5


1.0
.75
.50
.25
0
Power
H_L
H_0
0
.25
.50
.75
1.0
Location of middle dose
(Fraction of 64)
1.0
.75
.50
.25
0
Power
H_0
H_L
0
.25
.50
.75
1.0
Location of middle dose
(Fraction of 64)

(A) (B)

(A)





XBL827-3976

(B)

Fig. 6

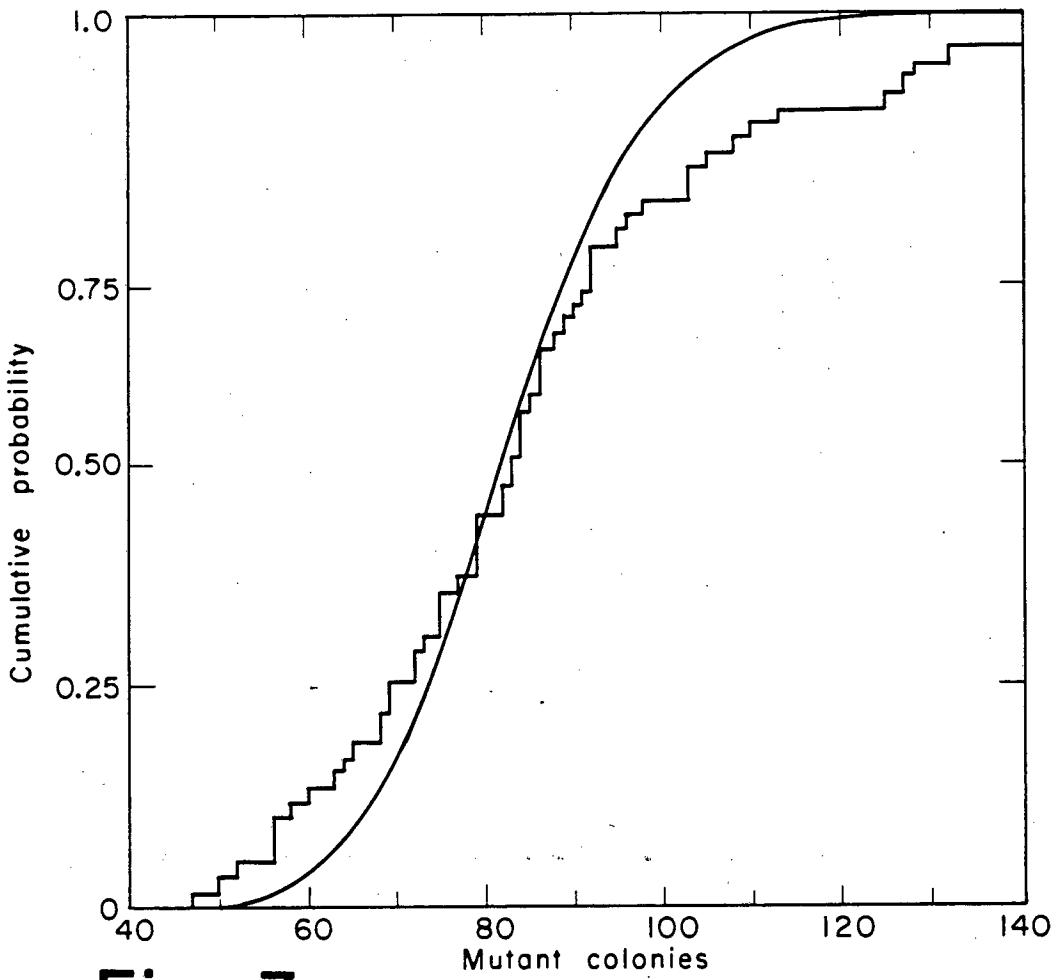


Fig. 7

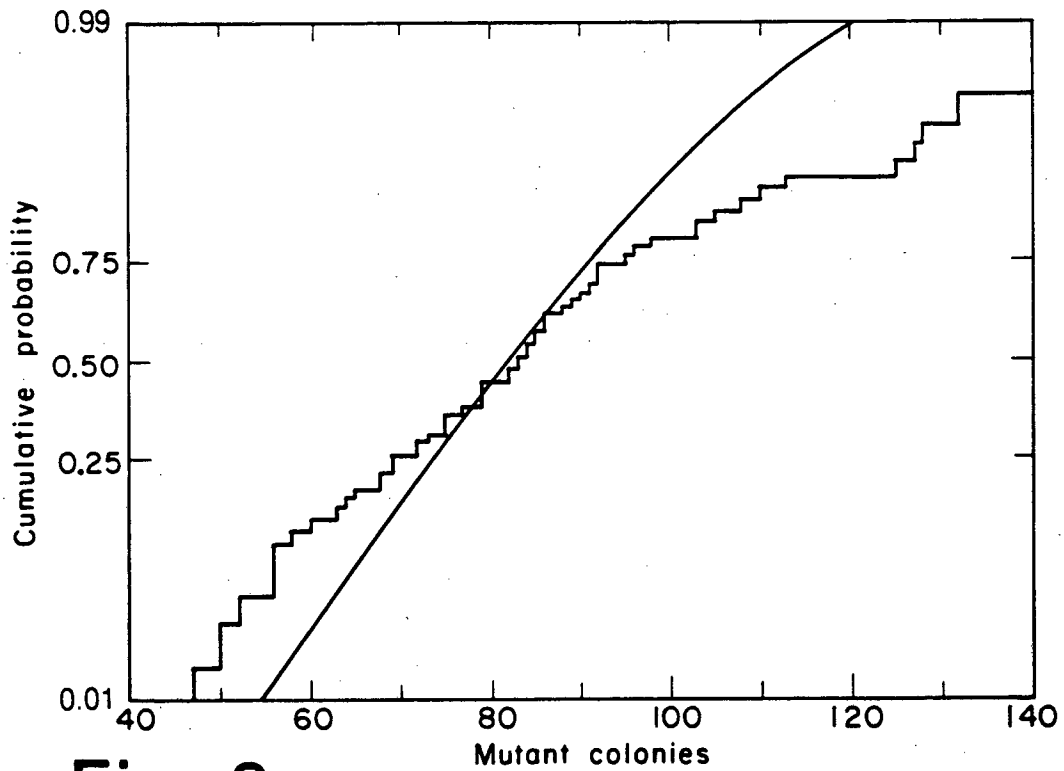


Fig. 8

XBL827-3977

This report was done with support from the Department of Energy. Any conclusions or opinions expressed in this report represent solely those of the author(s) and not necessarily those of The Regents of the University of California, the Lawrence Berkeley Laboratory or the Department of Energy.

Reference to a company or product name does not imply approval or recommendation of the product by the University of California or the U.S. Department of Energy to the exclusion of others that may be suitable.

TECHNICAL INFORMATION DEPARTMENT
LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720