BIOMETRICS 40, 27–40
March 1984

Calculation of Carcinogenic Potency from Long-Term Animal Carcinogenesis Experiments

Charles Sawyer

Department of Biochemistry, University of California, Berkeley, California 94720, U.S.A.

Richard Peto

Cancer Studies Unit, Nuffield Department of Clinical Medicine, Radcliffe Infirmary, Oxford, England

and

Leslie Bernstein and Malcolm C. Pike

Department of Preventive Medicine, School of Medicine, University of Southern California, Los Angeles, California 90033, U.S.A.

SUMMARY

An index of carcinogenic potency for chemicals tested in chronic animal experiments is described. By analogy with the well-known 'lethal dose 50' (LD50) of quantal bioassay, a 'tumorigenic dose 50' (TD50) may be defined (in the absence both of tumors in the control group and of intercurrent deaths) as that (daily) dose of chemical which gives 50% of the test animals tumors by some fixed age. Tumors in the control (zero-dose) group are handled exactly as for the LD50, and intercurrent deaths are handled by life-table methods. Nonparametric procedures are developed for estimating the TD50 and for constructing confidence intervals. These are based on likelihoods which assume that the tumor hazard is linear in dose.

1. Introduction

For reasons discussed in detail elsewhere (Gold *et al.*, 1984; Peto *et al.*, 1984), a particularly appropriate index of the apparent potency of chemicals in long-term animal carcinogenesis experiments is the 'TD50'. This is defined as the daily dose rate required to halve the probability of remaining tumorless at the end of a standard lifespan. A choice of standard lifespans for each species accompanied the original definition, and is necessary because a 20% increase in lifespan can easily double the cumulative incidence of tumors. Note that this definition automatically makes due allowance for the effects of intercurrent mortality, unrelated to treatment, on tumor yields. It contains, however, no explicit statement of precisely how the TD50 is to be estimated from real data. The present paper concerns estimation of the TD50 from such experiments by an extension of the methods of Cox (1972) in one important special case. This is the case of an animal experiment that is terminated at the end of the standard experimental lifespan, and in which the hazard function can be assumed to be approximately linearly related to the dose rate, and all tumors of the relevant type(s) found in animals dying before then can be analyzed as if they had caused the death of the host. Strictly speaking, one should ignore all tumors found incidentally at the postmortem of an animal that died prematurely of unrelated causes, and count all tumors, without exception, found at the terminal sacrifice (Peto *et al.*, 1980). There have been previous attempts to quantify potency (Twort and Twort, 1930, 1933;

Key words: Life-table methods; Carcinogenic potency; Linear hazard rate; Animal carcinogenesis experiments; TD50.

Iball, 1939; Bryan and Shimkin, 1943; Irwin and Goodman, 1946; Druckrey, 1967; Meselson and Russell, 1977). Our proposed index of carcinogenic potency (TD50) differs in that it takes into account both the incidence of spontaneous tumors and intercurrent mortality.

In this paper we first briefly develop the rationale for TD50. We then present a general linear dose–response model for time-to-tumor data, discuss the estimation of the TD50 for this model following the general approach of Cox (1972), and suggest a method for calculating confidence limits for the TD50.

2. TD50

The notion of the lethal dose 50 (LD50) is well-defined in standard quantal bioassay as that dose that kills 50% of the test animals (Loomis, 1978). This definition assumes that no animals die in the zero-dose (control) group. However, if animals need to be observed for an extended period of time, deaths may occur in the control group and the definition of the LD50 must be changed accordingly. This situation was discussed in detail by Finney (1949, 1964) who, in effect, suggested defining the LD50 as that dose that will halve the proportion of animals remaining alive at the end of the test period. Let P_0 $(= 1 - Q_0)$ be the proportion of control animals dying and let P_d $(= 1 - Q_d)$ be the proportion of animals dying at a dose of size d; then the LD50 is that dose, D, such that

$$Q_D = \tfrac{1}{2}Q_0. \tag{1}$$

This definition assumes that a proportion P_0 of animals respond whatever dose is given, and that the remaining proportion, Q_0, respond according to a tolerance distribution which has a structure that is *independent* of the size of Q_0. The LD50 is that dose that will kill 50% of this remaining proportion Q_0.

By analogy to the LD50, the TD50 may be defined (in the absence both of tumors in the control group and of intercurrent deaths) as that (daily) dose that gives 50% of the test animals tumors by some fixed age, T (the 'standard lifespan' for the given species). Tumors in the control group are handled exactly as in (1), and intercurrent deaths are handled by life-table methods.

3. Mathematical Model

We assume that an animal exposed daily to a specific chemical at dose d has a tumor-hazard rate (age-specific incidence rate) at age t, which may be written

$$\lambda(t; d) = (a + bd)\mu_0(t), \tag{2}$$

where $a > 0$ and $b \geq 0$ are unknown parameters and $\mu_0(t)$ is an unknown function of age such that $a\mu_0(t)$ is the tumor-incidence rate at age t at zero dose. Equation (2) states that an animal that is exposed to dose d, and is still alive and tumor-free immediately before age t, has a probability, $\lambda(t; d)\Delta$, of being diagnosed with a tumor in the next small age interval, Δ. This model has been shown to adequately describe many carcinogenesis data sets in man and in animals (Brown, 1976; Crump *et al.*, 1976; Hoel, 1980).

Since $\mu_0(t)$ is an arbitrary function of age, (2) may be written in a more convenient form by incorporating a into $\mu_0(t)$ so that

$$\lambda(t; d) = (1 + \beta d)\lambda_0(t). \tag{3}$$

Equation (3) implies that

$Q_0(t)$ = pr(animal at zero dose has not been diagnosed with a tumor by age t in the absence of all other causes of death)

$$= \exp\left\{-\int_0^t \lambda_0(x)\,dx\right\}. \tag{4}$$

Similarly, at dose d,

$Q_d(t)$ = pr(animal at dose d has not been diagnosed with a tumor by age t in the absence of all other causes of death)

$$= \{Q_0(t)\}^{1+\beta d}. \tag{5}$$

For fixed T, the TD50 is thus defined from (4) and (5) as that dose, D, such that

$$\{Q_0(T)\}^{1+\beta d} = \tfrac{1}{2}Q_0(T),$$

that is,

$$D = \log \tfrac{1}{2}/[\beta \log\{Q_0(T)\}]. \tag{6}$$

If the 'cause' of tumors in the absence of the chemical could somehow be removed, then it is reasonable to assume that the tumor-incidence rate at dose d would be

$$\xi(t;\, d) = \beta d\lambda_0(t). \tag{7}$$

Under this assumption the TD50 dose, D, would cause 50% of the animals to have tumors by age T. This is the 50% cumulative single-risk lifetime incidence of Meselson and Russell (1977), adjusted for intercurrent mortality. Note that the dose, TDρ, that will cause a ρ% cumulative single-risk lifetime incidence is obtained by multiplying the TD50 by the factor $\log(1 - \rho/100)\log\frac{1}{2}$, that is, TD$\rho = \log(1 - \rho/100)/[\beta \log\{Q_0(T)\}]$.

4. Maximum Likelihood Estimation of TD50

We consider an experiment at r dose levels, $d_1 \equiv 0, d_2, \ldots, d_r$, where

N_j = number of animals exposed to dose d_j

and

n_j = number of animals exposed to dose d_j that are diagnosed with tumors occurring at ages $X_{j1} \leqslant X_{j2} \cdots \leqslant X_{jn_j}$.

Write

$$N = \sum_{j=1}^{r} N_j$$

= total number of animals in the experiment

and

$$n = \sum_{j=1}^{r} n_j$$

n = total number of animals diagnosed with tumors in the experiment.

The n ages at which animals are diagnosed with tumors are often not recorded precisely, and more than one such tumor occurrence may be recorded at a given age. We write the

distinct ages at which data may be recorded as $\tau_0 \equiv 0 < \tau_1 < \cdots < \tau_K$. For example, if the experiment continues for 104 weeks and data are recorded daily, then $K = 728$.

Equation (3) implies that, for an animal exposed to the dose d_j and alive at τ_{i-1}, the conditional probability that it will be diagnosed with a tumor in the interval (τ_{i-1}, τ_i) is

$$\begin{aligned} P_i(d_j) &= 1 - \exp\{-(1 + \beta d_j) \int_{\tau_{i-1}}^{\tau_i} \lambda_0(t)\, dt\} \\ &= 1 - q_i(d_j) \\ &= 1 - \{q_i(0)\}^{1+\beta d_j} \\ &= 1 - q_i^{1+\beta d_j}, \end{aligned} \tag{8}$$

where $q_i(d_j)$ is the (conditional) probability of remaining tumor-free in the interval (τ_{i-1}, τ_i) and we write q_i for $q_i(0)$. We assume that any animal dying from a nontumor cause in (τ_{i-1}, τ_i) would have remained tumor-free until age τ_i in the absence of the nontumor death. In this situation $\lambda_0(t)$ can only be estimated in the form of the q_i terms, and β and the q_i terms must be estimated simultaneously.

The maximum likelihood estimate (MLE) of q_i at any age at which no tumors are diagnosed is 1, independent of β. We therefore need only consider the q_i terms corresponding to the k distinct ages of tumor diagnosis, $t_{(1)} < t_{(2)} < \cdots < t_{(k)}$. Denote the associated k conditional probabilities of remaining tumor-free as $q_{(1)}, q_{(2)}, \ldots, q_{(k)}$.

The total likelihood is

$$l(q_{(1)}, \ldots, q_{(k)}, \beta) \propto \prod_{i=1}^{k} \prod_{j=1}^{r} (1 - q_{(i)}^{(1+\beta d_j)})^{n_{ij}} (q_{(i)}^{(1+\beta d_j)})^{(N_{ij} - n_{ij})}, \tag{9}$$

where, for Dose Group j, n_{ij} and N_{ij} are, respectively, the number of animals diagnosed with a tumor at age $t_{(i)}$ and the number of animals at risk at age $t_{(i)}$. Note that if $t_{(i)} = \tau_h$, then the number of animals at risk at age $t_{(i)}$ is taken to be the number of animals alive immediately after τ_{h-1}. The associated log likelihood is thus

$$\begin{aligned} L &= L(q_{(1)}, \ldots, q_{(k)}, \beta) \\ &= \sum_{i=1}^{k} \sum_{j=1}^{r} \{n_{ij} \log(1 - q_{(i)}^{(1+\beta d_j)}) + (N_{ij} - n_{ij})(1 + \beta d_j) \log q_{(i)}\}. \end{aligned} \tag{10}$$

Writing $\hat{\beta}$ and $\hat{q}_{(i)}$ for the MLEs we find that

$$\hat{Q}_0(T) = \prod_i \hat{q}_{(i)}, \tag{11}$$

where the product of i extends from 1 to i_T, with $t_{(i_T)} \leq T$ and $t_{(i_T+1)} > T$. We also find that

$$\hat{D} = \log \tfrac{1}{2} / [\hat{\beta} \log\{\hat{Q}_0(T)\}]. \tag{12}$$

5. Confidence Intervals for TD50

From (6) and general likelihood theory, we may obtain approximate $100(1 - \alpha)\%$ confidence intervals for the TD50 by finding the maximum and minimum values of

$\beta \log\{Q_0(T)\} = \gamma$ such that

$$2\{L(\hat{\gamma}) - L(\gamma)\} = z^2_{\alpha/2}, \tag{13}$$

where $z_{\alpha/2}$ is the upper $100(1 - \frac{1}{2}\alpha)\%$ point of the standard normal distribution, and

$$L(\gamma) = \max\{L(q_{(1)}, \ldots, q_{(k)}, \beta)\} \tag{14}$$

subject to $\beta \sum \log q_{(i)} = \gamma$ and $\hat{\gamma} = \hat{\beta} \sum \log \hat{q}_{(i)}$, with $\sum$ denoting summation from 1 to (i_T).

To accomplish this we first determine approximate $100(1 - \alpha)\%$ confidence limits (β_L, β_U) for β. Writing $L(\hat{q}_{(1)}, \ldots, \hat{q}_{(k)} \mid \beta)$ for the likelihood evaluated at the MLEs of the $q_{(i)}$ terms for the given β, the lower limit, β_L $(<\hat{\beta})$, is that value of β for which

$$L(\hat{q}_{(1)}, \ldots, \hat{q}_{(k)} \mid \beta) = L(\hat{q}_{(1)}, \ldots, \hat{q}_{(k)}, \hat{\beta}) - \tfrac{1}{2}z^2_{\alpha/2} \tag{15}$$

holds. The upper limit, β_U $(>\hat{\beta})$, is defined similarly. Values of β outside the range (β_L, β_U) are not acceptable and are excluded from consideration in defining the confidence limits for the TD50.

For any fixed β in the interval (β_L, β_U) we must obtain the maximum and minimum values of γ, $\gamma_L(\beta)$ and $\gamma_U(\beta)$, subject to the constraint

$$L(q_{(1)}, \ldots, q_{(k)} \mid \beta) = L(\hat{q}_{(1)}, \ldots, \hat{q}_{(k)}, \hat{\beta}) - \tfrac{1}{2}z^2_{\alpha/2}. \tag{16}$$

When the method of Lagrange multipliers is used, this implies solving (16) and the i_T equations

$$1 - \delta\left[\sum_{j=1}^{r} (1 + \beta d_j)\{N_{ij} - n_{ij}/(1 - q_{(i)}^{(1+\beta d_j)})\}\right] = 0 \tag{17}$$

simultaneously for δ and the q_i terms $(i = 1, \ldots, i_T)$, where $\delta > 0$ for $\gamma_L(\beta)$ and $\delta < 0$ for $\gamma_U(\beta)$, leaving $q_{(i_T+1)}, \ldots, q_{(k)}$ at their ML values for fixed β.

The desired confidence interval (γ_L, γ_U) for γ is then

$$\gamma_L = \min\{\gamma_L(\beta)\},$$

where minimization is over all β in (β_L, β_U), and

$$\gamma_U = \max\{\gamma_U(\beta)\},$$

where maximization is again over all β in (β_L, β_U).

The desired confidence limits, D_L and D_U, for the TD50 are

$$D_L = \log \tfrac{1}{2}/\gamma_L \tag{18}$$

and

$$D_U = \log \tfrac{1}{2}/\gamma_U. \tag{19}$$

Computational methods for estimating these confidence intervals are described in §7. The adequacy of these confidence intervals for the TD50 is discussed in §9.

6. Rough Probabilities

If the ages of tumor diagnosis are recorded continuously so that $n_{(i)} = \sum_j n_{ij} \equiv 1$, then in the manner of Cox (1972), β can be estimated independently of the terms $q_{(i)}$. Conditional on the set $\{t_{(i)}\}$ of instants at which tumors occur, the probability that the tumor is in the animal as observed is

$$\prod_{j=1}^{r} (1 + \beta d_j)^{n_{ij}} \Big/ \sum_{j=1}^{r} N_{ij}(1 + \beta d_j).$$

Each tumor contributes a factor of this nature, and hence the required 'conditional' likelihood is

$$l(\beta) = \prod_{i=1}^{k} \left\{ \prod_{j=1}^{r} (1 + \beta d_j)^{n_{ij}} \Big/ \sum_{j=1}^{r} N_{ij}(1 + \beta d_j) \right\}. \tag{20}$$

Once the MLE of β is obtained, we consider the estimation of the distribution associated with $\lambda_0(t)$, calculating the MLEs of the terms $q_{(i)}$ conditional on β.

Time-to-tumor data are usually recorded discretely so that the terms $n_{(i)}$ may be greater than 1. If the data are not recorded continuously, then the 'exact' methods of obtaining the TD50 and its confidence interval (described in §§4 and 5) can be computationally prohibitive when several animals have the same ages of tumor diagnosis. However, the required 'conditional probability' for β at a given time point is an 'average' of the probabilities associated with each possible ordering of tumor occurrence. Here, each term has a numerator

$$\prod_{j=1}^{r} (1 + \beta d_j)^{n_{ij}},$$

and a denominator

$$\left\{ \sum_{j=1}^{r} N_{ij}(1 + \beta d_j) \right\} \left[\prod_{m=1}^{n_{(i)}-1} \left\{ \sum_{j=1}^{r} (N_{ij} - u_{ijm})(1 + \beta d_j) \right\} \right],$$

where u_{ijm} $(m = 1, \ldots, n_{(i)} - 1)$ is a possible sequence of tumors. The denominators are thus slightly less than the rough approximation to them given by

$$\left\{ \sum_{j=1}^{r} N_{ij}(1 + \beta d_j) \right\}^{n_{(i)}}.$$

In his discussion of the paper by Cox (1972), Peto suggested using this simple 'rough-probabilities' approach. Experience with the Cox model has shown that this approximation gives very accurate answers. This rough-probabilities approach suggests that $\hat{\beta}$ may be approximated by obtaining the MLE of β from the appropriately modified version of (20), namely,

$$l_{\mathrm{RP}}(\beta) = \prod_{i=1}^{k} \left[\prod_{j=1}^{r} (1 + \beta d_j)^{n_{ij}} \Big/ \left\{ \sum_{j=1}^{r} N_{ij}(1 + \beta d_j) \right\}^{n_{(i)}} \right]. \tag{21}$$

Conditional on this MLE of β, approximate values of each $\hat{q}_{(i)}$ may be found from (9) by ML methods. The solution to the k ML equations derived from (9), i.e.

$$\sum_{j=1}^{r} [N_{ij}(1 + \beta d_j)/q_{(i)} - n_{ij}(1 + \beta d_j)/\{q_{(i)}(1 - q_{(i)}^{1+\beta d_j})\}] = 0, \tag{22}$$

requires iteration. An approximate method, requiring no iteration, may be developed in the manner suggested by Breslow (1974), in which the hazard function $\lambda(t; d)$ is approximated by a step function with discontinuities at each observed age of tumor diagnosis (see Kalbfleisch and Prentice, 1980, §4.2.4). This is equivalent to replacing (9), for fixed β, with

$$l_\eta(\eta_{(1)}, \eta_{(2)}, \ldots, \eta_{(k)} \mid \beta) = \prod_{i=1}^{k} \left[\left\{ \eta_{(i)} \sum_{j=1}^{r} (N_{ij} - \tfrac{1}{2}n_{ij})(1 + \beta d_j) \right\}^{n_{(i)}} \right.$$

$$\left. \exp\left\{ - \eta_{(i)} \sum_{j=1}^{r} (N_{ij} - \tfrac{1}{2}n_{ij})(1 + \beta d_j) \right\} \Big/ (n_{(i)}!) \right], \tag{23}$$

where $\eta_{(i)} = -\log q_{(i)}$. [Equation (23) is a Poisson approximation.] The solution for $\hat{\eta}_{(i)}$ is immediate:

$$\hat{\eta}_{(i)} = n_{(i)} \Big/ \left\{ \sum_{j=1}^{r} (N_{ij} - \tfrac{1}{2}n_{ij})(1 + \hat{\beta}d_j) \right\}. \tag{24}$$

Equations (21) and (23) may also be used in the obvious way as the basis for developing confidence intervals as described in §5. The adequacy of these intervals is discussed in §9.

7. Computational Methods

We have found that using a Newton–Raphson approach to obtain the MLEs of β and the $q_{(i)}$ terms for either (9) or (21) is not satisfactory, and we have resorted to direct search techniques. We consider β in the form $\theta = \arctan \beta$, so as to restrict our search to a finite range $(0, \frac{1}{2}\pi)$. For both the exact and the rough-probabilities methods, $\hat{\theta}$ (or, equivalently, $\hat{\beta}$) is located by a golden section (GS) search (Overholt, 1967), and the $q_{(i)}$ terms for the exact method are obtained at each θ in the search by the same GS method.

To compute the confidence intervals described in §5, β_L and β_U are obtained by a GS search in a straightforward manner. For a given β in (β_L, β_U), we solve (16) and (17) simultaneously to obtain $\gamma_L(\beta)$ and $\gamma_U(\beta)$, by realizing that, for any given δ, each of the i_T equations (17) may be solved independently for a $q_{(i)}$. Thus, for a given β, the problem then reduces to a one-dimensional search for δ. We then use a GS search on β to obtain γ_L and γ_U.

The GS method assumes that the function being maximized is unimodal. However, the procedure continuously monitors a set of four points to detect any gross violation of this assumption. We have not proved that the respective likelihoods behave as assumed, but the checking of many individual cases, by using fine subdivisions of the intervals in which we search, has always given the same answer as the GS search.

8. Goodness of Fit

We test the fit of the observed data to the hypothesized model (3) by comparing the observed number of tumor deaths on Treatment j, n_j, to its expected number, e_j, based on the model, where

$$e_j = \sum_{i=1}^{k} N_{ij}(1 - \hat{q}_{(i)}^{(1+\hat{\beta}d_j)}). \tag{25}$$

If the model is an adequate description of the experimental results, then

$$X^2 = \sum_{j=1}^{r} (n_j - e_j)^2/e_j \tag{26}$$

should have an approximate chi square distribution on $r - 2$ degrees of freedom (df). The accuracy of this assumption is discussed in §9.

Under the null hypothesis, that is, $\beta = 0$, write the expected number of tumors in Group j as E_j (Mantel, 1966; Peto and Pike, 1973). A simple test of the null hypothesis is then

$$X^2 = \sum_{j=1}^{r} (n_j - E_j)^2/E_j \tag{27}$$

which has an approximate chi square distribution on $r - 1$ df. This test can be strengthened without recourse to the sophistication of (3) by using the test for trend, due to Mantel (1963) with respect to dose.

These null-hypothesis expected values, E_j, may also be used to suggest deviations from

an increasing tumor risk with increasing dose. Write

$$R_j = n_j/E_j, \tag{28}$$

and note that R_1 refers to the zero-dose group. R_j is termed the 'relative tumor rate' for Group j. The ratio, R_j/R_1, is an excellent estimate of the ratio of the tumor rate in Group j to the tumor rate in the controls (Pike, in his discussion of the paper by Cox, 1972; Breslow, 1975; Bernstein, Anderson and Pike, 1981). A plot of R_j/R_1 against dose will show obvious deviations from a linear trend, as will residual plots of $(n_j - e_j)/(e_j)^{1/2}$.

9. Simulation Study

A simulation study was performed to investigate the amount of bias in the estimates of the TD50 calculated by the exact method (§§4 and 5) and by the rough-probabilities method (§6) as well as the coverage probabilities of the respective confidence intervals and the behavior of the $r - 2$ df goodness-of-fit test of the model (26).

For each specific case described below, data were generated for 1000 bioassays in which groups of 50 animals were tested at three dose levels,

$$\{d_1, d_2, d_3\} = \{0, 1, 2\}.$$

Since most theoretical models of carcinogenesis predict a Weibull distribution (Pike, 1966; Peto, 1977), this distribution was selected as the appropriate model for generating trial times. For each simulation, tumor times were generated according to the Weibull survival distribution,

$$F(x) = \exp\{-c(1 + \beta d)(x - w)^s\} \tag{29}$$

with $s = 4$ and $w = 50$. We set T at 104 (weeks). The TD50 based on (29) is given by

$$\text{TD50} = \log 2/\{c\beta(T - w)^s\} \tag{30}$$

so that

$$c\beta = \log 2/\{\text{TD50}(T - w)^s\}. \tag{31}$$

In addition to fixing the TD50, we also required the probability of tumor in the control group, by age T, to be p_T in order to solve for c and β. From (29),

$$p_T = 1 - \exp\{-c(T - w)^s\}. \tag{32}$$

Thus, from (32), we obtain

$$c = -\log(1 - p_T)/(T - w)^s \tag{33}$$

and, using (31) and (33),

$$\beta = \log \tfrac{1}{2}/\{\text{TD50} \log(1 - p_T)\}. \tag{34}$$

For the simulation study, we investigated a variety of cases in which $p_T = 10^{-6}$ or $p_T = .25$ and TD50 = 1 or TD50 = 5.

We investigated the effects of intercurrent mortality due to two independent causes of death, one independent of dose and one related to chemical toxicity. For the cause of death unaffected by chemical dose, we assumed a competing cause of death with survivor function given by (29), with $s = 4$, $w = 50$, $\beta = 0$, and c calculated so that the net probability of death due to this cause was p_U. For the cause of death affected by chemical dose, we assumed a competing cause of death with survivor function given by (29), with $s = 4$, $w = 50$, c given by (33) with p_T set equal to p_A, and β calculated to give a fixed LD50 (in the same manner as described above for the TD50). Details of the specific cases studied are shown in Table 1.

Table 1
Specific situations studied by simulation

Case	TD50	p_T	p_U	LD50	p_A
1	1	10^{-6}	10^{-6}	∞	10^{-6}
2	1	.25	10^{-6}	∞	10^{-6}
3	5	10^{-6}	10^{-6}	∞	10^{-6}
4	5	.25	10^{-6}	∞	10^{-6}
5	1	.25	10^{-6}	1	.25
6	1	.25	10^{-6}	5	.25
7	5	.25	10^{-6}	1	.25
8	5	.25	10^{-6}	5	.25
9	1	.25	.25	5	.25

In order to generate random samples of trial times, independent uniform random variates generated by the pseudorandom number algorithm of Pike and Hill (1966), as refined by Hill (in Atkinson and Pearce, 1976, p. 451) were transformed according to (29) as required. For all situations studied, the same stream of random variables was generated. Three trial times (time to tumor, time to dose-independent intercurrent mortality, time to dose-dependent intercurrent mortality) were generated for each animal. The minimum trial time was then rounded up to the nearest week, with trial times exceeding T considered as censored observations at T. For each of the specific cases studied (see Table 1), estimates of the TD50 and 95% confidence intervals were calculated, and the results of the goodness-of-fit test for the model at the $\alpha = .05$ level of significance were obtained.

Normal probability plots for $\hat{D}$, log $\hat{D}$ and $1/\hat{D}$, based on the simulation results for the exact method for Case 8, are shown in Figs 1 to 3, respectively. The distribution of $1/\hat{D}$ was found to be approximately normal, as indicated in Fig. 3, suggesting that this transformation is the natural scale to use in presenting simulation results for the mean estimates of the TD50. The results of these simulations, presented in Table 2, indicate that both the exact method and the rough-probabilities method give nearly identical minimally

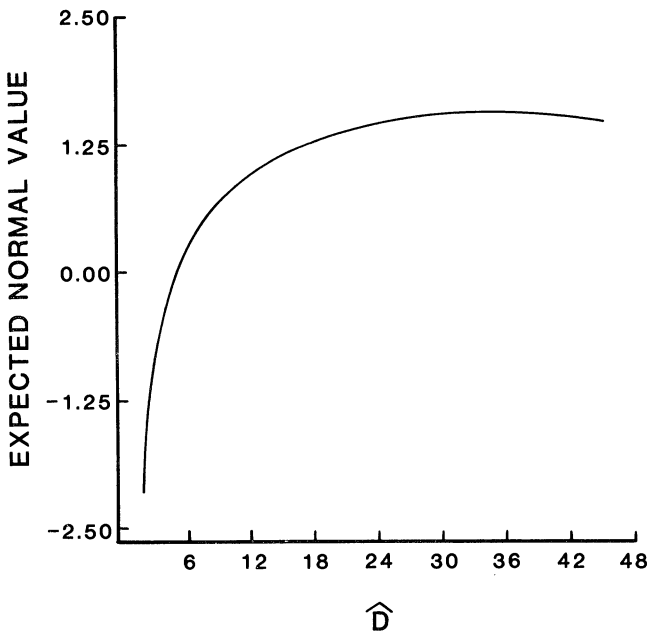

Figure 1. Normal probability plot for $\hat{D}$ based on simulation results for Case 8 analyzed by the 'exact' method.

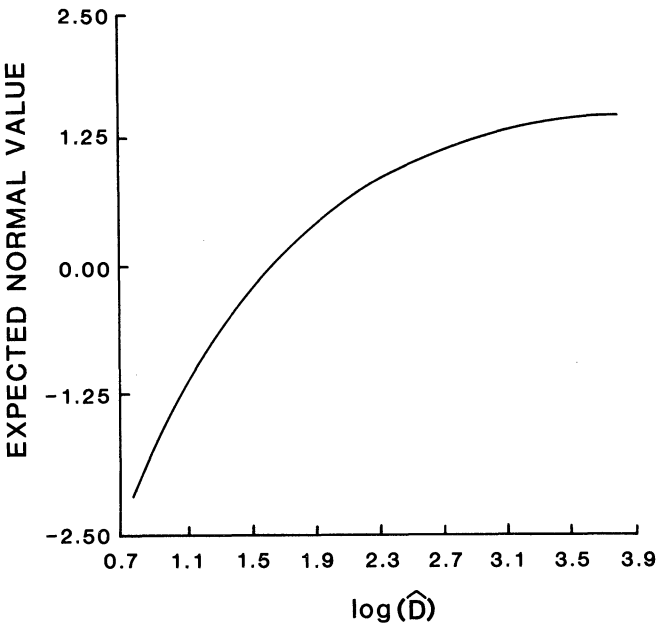

Figure 2. Normal probability plot for log $\hat{D}$ based on simulation results for Case 8 analyzed by the 'exact' method.

biased mean estimates of the TD50. For both methods, coverage probabilities of the confidence intervals give approximately the right coverage probabilities. The last column of Table 2 shows that results for the goodness-of-fit test for the model also are all very close to the nominal level.

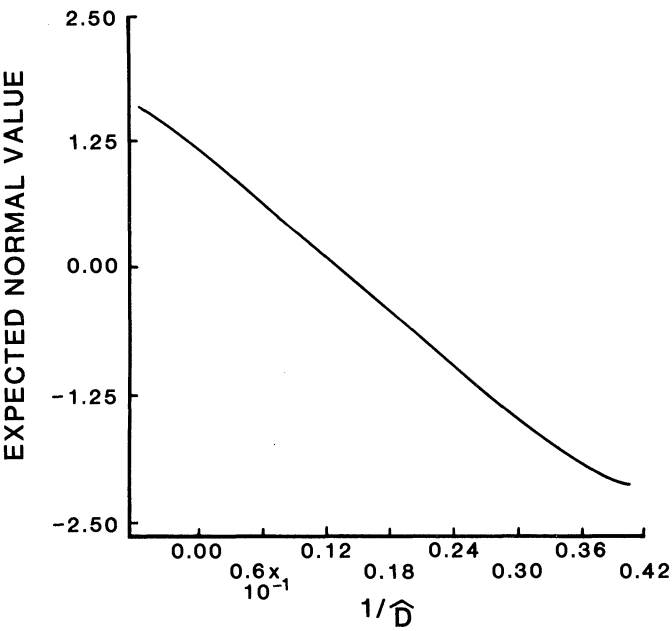

Figure 3. Normal probability plot for $1/\hat{D}$ based on simulation results for Case 8 analyzed by the 'exact' method.

Table 2

Harmonic mean values of simulation estimates of TD50, actual coverage probabilities of 95% confidence intervals, and actual probabilities of rejecting goodness-of-fit test for the model at $\alpha = .05$ level of significance

Case	TD50	Method of analysis	$\hat{D}$	pr(TD50 < D_L)	pr(TD50 > D_U)	pr($X^2 > \chi^2_{.95,1}$)
1	1	Exact	0.990	.020	.004	.030
		RP*	0.990	.020	.030	.034
2	1	Exact	0.988	.019	.028	.037
		RP	0.997	.021	.023	.045
3	5	Exact	4.939	.003	.024	.045
		RP	4.937	.024	.024	.045
4	5	Exact	4.896	.018	.025	.048
		RP	4.951	.011	.021	.052
5	1	Exact	0.998	.026	.024	.040
		RP	1.016	.027	.020	.047
6	1	Exact	0.984	.019	.024	.045
		RP	0.999	.021	.022	.055
7	5	Exact	4.973	.009	.025	.034
		RP	5.003	.002	.024	.034
8	5	Exact	4.956	.016	.034	.048
		RP	5.014	.016	.029	.050
9	1	Exact	0.985	.017	.027	.047
		RP	1.001	.022	.020	.053

* RP: Rough probabilities.

The confidence intervals for the TD50 discussed in §5 require extensive iteration. When computing costs must be considered, the approximate normality of the distribution of $D^{-1} = 1/\text{TD50}$ suggests that one might use an asymptotic approach to the calculation of these intervals. The approximate variance of D^{-1} can be found by first applying the 'delta method' (Rao, 1965) to estimate the variance, $\hat{V}$, of $\log(\hat{D}^{-1})$ and then noting that

$$\hat{V}(\hat{D}^{-1}) = \hat{V}/\hat{D}^2.$$

We have investigated these asymptotic confidence intervals by simulation for the cases shown in Table 1 and find that their total coverage probabilities are close to the nominal level. However, the lower and upper coverage probabilities are not equal with pr(TD50 < D_L) being roughly twice pr(TD50 > D_U).

10. An Example

To illustrate the use of the exact method and the rough-probabilities method for calculating the TD50 and its confidence intervals, we present the skin-tumor data (Table 3) and the analysis (Table 4) of the National Cancer Institute bioassay on the tumorigenic effects of 5-nitro-*o*-anisidine administered to male rats in their diet. Three groups of male rats were studied: a control group of 99 animals, a low-dose group (118 mg/kg body weight/day) of 49 animals, and a high dose group (235 mg/kg body weight/day) of 50 animals. For this experiment, we assumed the standard lifespan for the animals to be $T = 104$ weeks.

The estimates obtained for the TD50 and its confidence intervals are little different. The good-of-fit tests for the model and plot of the relative tumor-rate ratio against dose indicate that the assumption of linearity is valid for this experiment.

11. Conclusions

The two methods of calculating the TD50 described here are based on a model that assumes that the tumor-incidence rate is a linear function of dose. The TD50 differs from earlier

Table 3

Time-to-skin-tumor data for the National Cancer Institute bioassay on the tumorigenic effects of 5-nitro-o-anisidine in male rats

Time point	Time to tumor (weeks)	Dose rate (mg/kg/day)					
		0.0		118.0		235.0	
		n_{i1}	N_{i1}	n_{i2}	N_{i2}	n_{i3}	N_{i3}
1	47	0	99	0	47	1	49
2	48	0	99	0	47	1	48
3	60	0	98	1	47	0	46
4	66	0	98	1	43	2	44
5	68	0	98	1	42	1	42
6	69	0	98	1	41	0	40
7	70	0	98	0	39	2	40
8	71	0	98	1	39	6	38
9	72	0	98	0	37	10	32
10	73	0	97	1	37	1	21
11	74	0	97	1	36	1	20
12	75	0	96	0	35	1	19
13	76	0	96	1	35	1	18
14	79	0	91	1	33	0	17
15	81	0	86	1	30	0	16
16	83	0	86	1	28	0	15
17	84	0	86	0	26	1	15
18	87	0	84	0	25	2	14
19	92	0	80	0	25	3	12
20	93	0	79	1	25	0	9
21	94	0	79	1	24	0	9
22	96	0	77	0	23	1	9
23	97	0	76	0	22	3	8
24	98	0	76	9	22	0	5
25	99	1	75	0	12	0	5
26	100	0	74	2	11	2	5
27	101	0	72	1	9	1	3
28	102	0	71	4	7	2	2
29	106	1	68	1	3	0	0

indices of potency in that both intercurrent mortality and the incidence of spontaneous tumors are taken into account. Simulation results for Weibull data indicate that both the exact method and the rough-probabilities method provide minimally biased estimates of 1/TD50 and close to the nominal confidence interval coverage. These methods can be applied to the results of the many chronic animal studies that have adequate time-to-tumor data available. 'Incidental tumors' (Peto *et al.*, 1980), i.e. tumors that are discovered at autopsy but that have not contributed to the death of the animal, should not strictly be counted as tumors when using these methods. Our method is applicable to either 'fatal tumors', i.e. tumors that are the direct or indirect cause of death of the animal, or 'mortality-independent tumors', i.e. tumors detected at a standard point in the development of the tumor in living animals. Peto *et al.* (1984) discussed the analysis of actual experiments where the distinction between fatal and incidental tumors is not made.

Confidence limits for the TD50 can be used to examine the sensitivity of 'negative' (i.e. not statistically significant) bioassays, so that the bioassay can be described as excluding TD50s below a certain limit, rather than simply as 'negative'. Such calculations may be used to reconcile seemingly conflicting positive and negative results on the same compound.

Table 4

Results of analyses of skin-tumor data for the National Cancer Institute bioassay on the tumorigenic effects of 5-nitro-o-anisidine in male rats, assuming $T = 104$ weeks

	Method	
	Exact	Rough probabilities
$\hat{D}$	28.0	29.7
Test for $\beta = 0$*	187.8	175.4
Two-sided *P*-value	0.00	0.00
(D_L, D_U)†	(14.9, 48.6)	(16.2, 50.9)
Fit of model		
X^2 on 1 df for non-linearity	2.37	2.42
Two-sided *P*-value	0.12	0.12
Mantel test for trend		
X^2 on 1 df for trend	182.39	
One-sided *P*-value	0.00	

* Based on the likelihood ratio test using (10): the tabulated values are of twice the change in log-likelihood on 1 df.

† 99% Confidence intervals for *D*.

Acknowledgements

The authors wish to thank Dr Bruce Ames for realizing the need for and the possibility of developing a standard method of analysis of animal carcinogenesis experiments, and Lois Gold and colleagues for arranging for the collection of the data from the published literature on these experiments. This work was supported by Grant P01 CA 17054 from the National Cancer Institute; Grant SIG-2 from the American Cancer Society; Department of Energy Contract DE-AT03-80EV70156; and Grant ES01896 from the National Institute of Environmental Health Sciences.

Résumé

On décrit un index de la puissance cancerigène de produits chimiques testés dans des expériences sur animaux suivis dans le temps. Par analogie avec la dose lethale 50 (LD50) bien connue dans les essais biologiques, on peut définir une 'dose tumeur 50' (TD50) (en absence de tumeur dans le groupe contrôlé et de mort dans la période d'étude) comme la dose (journalière) de produit chimique qui donne 50% de tumeurs pour les animaux testés à un âge fixé. Les tumeurs dans le groupe contrôlé (dose zéro) sont traitées exactement comme pour LD50 et on utilise les méthodes des tables de survie pour tenir compte des morts. On développe des procédures non paramétriques pour estimer le TD50 et construire des intervalles de confiance. Ces procédures supposent que les probabilités de l'apparition de la tumeur sont linéaires de la dose.

References

Atkinson, A. C. and Pearce, M. C. (1976). The computer generation of beta, gamma, and normal random variables (with discussion). *Journal of the Royal Statistical Society, Series A* **139,** 431–461.

Bernstein, L., Anderson, J. A. and Pike, M. C. (1981). Estimation of the proportional hazard in two treatment group clinical trials. *Biometrics* **37,** 513–519.

Breslow, N. E. (1974). Covariance analysis of censored survival data. *Biometrics* **30,** 89–99.

Breslow, N. E. (1975). Analysis of survival data under the proportional hazards model. *International Statistical Review* **43,** 45–58.

Brown, C. C. (1976). Mathematical aspects of dose–response studies in carcinogenesis: The concept of thresholds. *Oncology* **33,** 62–65.

Bryan, W. R. and Shimkin, M. B. (1943). Quantitative analysis of dose–response data obtained with

three carcinogenic hydrocarbons in strain C3H mice. *Journal of the National Cancer Institute* **3,** 503–531.

Cox, D. R. (1972). Regression models and life tables (with discussion). *Journal of the Royal Statistical Society, Series B* **34,** 187–220.

Crump, K. S., Hoel, D. G., Langley, C. H. and Peto, R. (1976). Fundamental carcinogenic processes and their implications for low-dose risk assessment. *Cancer Research* **26,** 2973–2979.

Druckrey, H. (1967). Quantitative aspects in chemical carcinogenesis. In *Potential Carcinogenic Hazards from Drugs*, R. Truhaut (ed.). UICC Monograph No. 7, 60-78. New York: Springer Verlag.

Finney, D. J. (1949). The adjustment for a natural response rate in probit analyses. *Annals of Applied Biology* **36,** 187–195.

Finney, D. J. (1964). *Statistical Methods in Biological Assay.* London: Griffin.

Gold, L. S., Sawyer, C. B., McGaw, R., Buckman, G. M., deVecidna, M., Levinson, R., Hooper, N. K., Havender, W. R., Bernstein, L., Peto, R., Pike, M. C. and Ames, B. N. (1984). A carcinogenic potency database of the standardized results of animal bioassays. *Environmental Health Perspectives.* To appear.

Hoel, D. G. (1980). Incorporation of background in dose–response models. *Federation Proceedings* **39,** 73–75.

Iball, J. (1939). The relative potency of carcinogenic compounds. *American Journal of Cancer* **35,** 188–190.

Irwin, J. O. and Goodman, N. (1946). The statistical treatment of measurements of the carcinogenic properties of tars (Part I) and mineral oils (Part II). *Journal of Hygiene* **44,** 362–420.

Kalbfleisch, J. D. and Prentice, R. L. (1980). *The Statistical Analysis of Failure Time Data.* New York: Wiley.

Loomis, T. A. (1978). *Essentials of Toxicology.* Philadelphia: Lea and Febiger.

Mantel, N. (1963). Chi-square tests with one degree of freedom. Extensions of the Mantel–Haenszel procedure. *Journal of the American Statistical Association* **58,** 690–700.

Mantel, N. (1966). Evaluation of survival data and two new rank order statistics arising in its consideration. *Cancer Chemotherapy Reports* **50,** 163–170.

Meselson, M. and Russell, K. (1977). Comparisons of carcinogenic and mutagenic potency. In *Origins of Human Cancer*, H. H. Hiatt, J. D. Watson and J. A. Winsten (eds), 1473–1481. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory.

Overholt, K. J. (1967). Algorithm 17—GOLDSEC. *Computer Journal* **9,** 415.

Peto, R. (1977). Epidemiology, multistage models and short-term mutagenicity tests. In *Origins of Human Cancer*, H. H. Hiatt, J. D. Watson and J. A. Winsten (eds.), 1403–1428. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory.

Peto, R. and Pike, M. C. (1973). Conservatism of the approximation $\sum (0 - E)^2/E$ in the logrank test for survival data or tumor incidence data. *Biometrics* **29,** 579–584.

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

Peto, R., Pike, M. C., Bernstein, L., Gold, L. S. and Ames, B. N. (1984). The TD50: A proposed general convention for the numerical description of the carcinogenic potency of chemicals in chronic-exposure animal experiments. *Environmental Health Perspectives.* To appear.

Pike, M. C. (1966). A method of analysis of a certain class of experiments in carcinogenesis. *Biometrics* **22,** 142–161.

Pike, M. C. and Hill, J. D. (1966). Remark on algorithm 266(G5) pseudo-random numbers. *Communications of the Association for Computing Machinery* **9,** 687.

Rao, C. R. (1965). *Linear Statistical Inference and Its Applications.* New York: Wiley.

Twort, C. C. and Twort, J. M. (1930). The relative potency of carcinogenic tars and oils. *Journal of Hygiene* **29,** 373–379.

Twort, C. C. and Twort, J. M. (1933). Suggested methods for the standardization of the carcinogenic activity of different agents for the skin of mice. *American Journal of Cancer* **17,** 293–320.

Received December 1981; *revised March* 1983