Best-Estimate Low-Dose Extrapolation of Carcinogenicity Data

by H. A. Guess* and K. S. Crump*

At the 1974 NIEHS conference on low-dose extrapolation, Peto (*1*) reported that evidence concerning chronic exposure to direct carcinogens (i.e., those acting on the DNA, producing mutations) suggests that the excess rate of incidence over background should be given by the product of a function of the dose and a function of the exposure duration. The function of the dose should be a plynomial of the form

$$q_1 d^1 + q_2 d^2 + \ldots + q_K d^K$$

with $q_i \geq 0$ and $q_1 > 0$. Expressions of this form for the excess rate of incidence over background correspond to lifetime response probabilities of the form

$$P(d) = 1 - \exp\{-(q_0 + q_1 d^1 + \ldots + q_K d^K)\}$$
$$q_i \geq 0 \qquad (1)$$

We have developed a low-dose extrapolation technique for dichotomous data, using an assumed dose–response relation of the form (1) but without making any assumptions about whether the coefficient of the linear term is positive or about the degree of the polynomial. We use maximum likelihood to determine both the degree K of the polynomial and the coefficients $q_i \geq 0$, for $0 \leq i \leq K$. Allowance for nonzero background is made automatically by fitting the constant term, q_0. Our estimate of the increased response probability over background is given by

$$P(d) - P(0) = e^{-q_0}[1 - \exp\{-(q_1 d^1 +$$
$$\ldots + q_K d^K)\}] \qquad (2)$$

At very low doses this looks like

$$P(d) - P(0) = e^{-q_0} q_l d^l \qquad (3)$$

---

*Environmental Biometry Branch, National Institute of Environmental Health Sciences, Box 12233, Research Triangle Park, N. C. 27709.

for $d \ll 1$, where l is the smallest positive integer for which $q_l > 0$. At high doses, the shape of the curve is governed by the highest-order coefficient, q_K.

Unlike previous dose–response models, which contain only one or two free parameters to be determined, this model contains infinitely many, because the degree of the polynomial is left to be determined by the data. For any given set of data, only finitely many coefficients will be nonzero; however we do not make any *a priori* assumptions about which coefficients will nonzero. In view of Eq. (3), the resulting low-dose extrapolation will look like that of an l-hit model for some $l = 1, 2, \ldots$, depending on the data.

By considering at once all polynomials with nonnegative coefficients—unrestricted as to degree we are able to compare dose–response curves that are linear in the low-dose range with dose–response curves that are extremely flat in the low-dose range. For each set of data we calculate the (in most cases) uniquely determined set of nonnegative coefficients which best fits the data, in the sense of maximizing the likelihood function over the class of all polynomials with nonnegative coefficients. Maximum likelihood estimation by use of polynomials with nonnegative coefficients (absolutely monotonic) polynomials) is quite different from maximum likelihood estimation with polynomials with unrestricted coefficients. When the coefficients are unrestricted, it makes no sense to consider polynomials whose degree is greater than the number of doses at which tests are conducted. When the coefficients are constrained to be nonnegative, it is generally not possible to fit the data points exactly, no matter how high the degree. In general increasing the degree beyond a certain point can actually cause the likelihood function to start decreasing. Unlike the situation with unrestricted coefficients, increasing the degree of absolutely monotonic

polynomials does not lead to complex curve shapes. All that happens is that the high-dose part of the curve hooks up more sharply while the low-dose part of curve flattens out. The flatter the low-dose part of the curve, the higher will be the predicted dose corresponding to a given level of increased risk over background. Thus it is desirable not to exclude high degree polynomials because doing this limits artificially the extent to which data which are best fit by very flat curves can be fit by curves of the form (*1*). For this reason, we have developed the theory needed for maximizing the likelihood function globally over all absolutely monotonic polynomials, unrestricted as to degree. We have an efficient method for calculating such estimates and have obtained theorems describing their statistical properties, such as asymptotic unbiasedness, strong consistency, and asymptotic normality. In addition, we derive and calculate asymptotic (large-sample) confidence intervals for the risk estimates $P(d) - P(0)$ and compare the calculated confidence intervals with those obtained by computer simulation. These confidence intervals are not the ordinary binomial confidence limits at the test doses. Rather they are obtained by means of a limit theorem which we prove. Finally, we have developed a Monte Carlo goodness-of-fit test, which involves comparing the actual data with sets of artificial data simulated using the best estimate response probabilities.

All of these calculations have been assembled into one computer program which reads the dichotomous data and calculates estimates of the increased response probability over background, confidence intervals for these estimates, and goodness-of-fit-test results.

Our main conclusions are as follows: We can easily specify hypothetical sets of dichotomous dose-response data for which the best estimate (maximum likelihood) low dose extrapolation of the form of Eq. (1) is much closer to a probit extrapola-

---

*Probit curves vanish to infinite order at the origin in the sense that $d^{-k}\Phi(a + b \log d) \to 0$ as $d \to 0$ for all $k = 1,2, \ldots$, where Φ is the standard normal distribution function. Curves of the form (1), being analytic, can only vanish to finite order. Thus, given any curve of the form (1) and any probit curve, the probit curve will eventually undershoot the curve of form (1) for d sufficiently close to 0. However, as Table 1 shows, it is not difficult to fit a probit curve with a curve of the form (1) down to doses corresponding to risk levels of about 10^{-7}.

†With the dimethylnitrosamine data it is possible to make the linear coefficient q_1 vanish by excluding the response at the highest dose tested. However, the goodness-of-fit, as measured by our Monte Carlo goodness-of-fit test, was not rejectable at the 10% level with the highest dose included. For this reason, it did not seem valid to exclude this response.

tion than to a linear extrapolation. Table 1 illustrates such data. We cite these results to demonstrate that it is mathematically possible for risk estimates based on Eq. (1) to be approximately the same as risk estimates based on a probit model and to differ by several orders of magnitude from risk estimates based on a linear model.*

Table 1A. Hypothetical data for which the multistage model approximates a probit curve.[a]

Dose	Animals tested	Animals with positive responses
0	1000	0
5	1000	0
10	1000	0
18	1000	1
30	1000	11
45	1000	42
60	1000	95
75	1000	161

[a]The data are expected values computed from the probit curve $P(d) = \Phi(-7.18 + 3.30 \log d)$. This curve was also used to get the probit low-dose extrapolation. The multistage and linear extrapolations were obtained from the above hypothetical data by maximum likelihood.

Table 1B. Low-dose extrapolations of data of Table 1A.

	Dose d corresponding to given risk		
Risk $P(d)$	Probit Model	Multistage model[a]	One-hit model
10^{-8}	2.99	0.28	7.5×10^{-6}
10^{-7}	3.99	0.61	7.5×10^{-5}
10^{-6}	5.44	1.32	7.5×10^{-4}
10^{-5}	7.65	2.85	7.5×10^{-3}
10^{-4}	11.19	6.13	7.5×10^{-2}
10^{-3}	17.35	13.22	0.75
10^{-2}	29.56	28.52	7.55
10^{-1}	61.24	62.42	79.14

[a]The multistage model is the model based on Eq. (1).

In every set of actual data we have analyzed, the linear term q_1 is positive, and the best-estimate low-dose extrapolation is much closer to a linear extrapolation than a conservative Mantel-Bryan extrapolation. We have analyzed data for vinyl chloride, dieldrin, DDT, dimethylnitrosamine, and ionizing radiation. Figures 1 and 2 illustrate the results for dieldrin.† In the range of increased response between 10^{-5} and 10^{-8} the best-estimate dose-response curve of the form of Eq. (1) for these agents is typically one or two orders of magnitude above a conservative Mantel-Bryan extrapolation.

Our results have implications which should be considered by anyone who intends to design a large-scale experiment to measure the shape of dose–response curves for chemical carcinogens in the very low-dose range. When both very flat curves and gradually sloping (linear or nearly linear) curves are considered together, it is extremely dif-
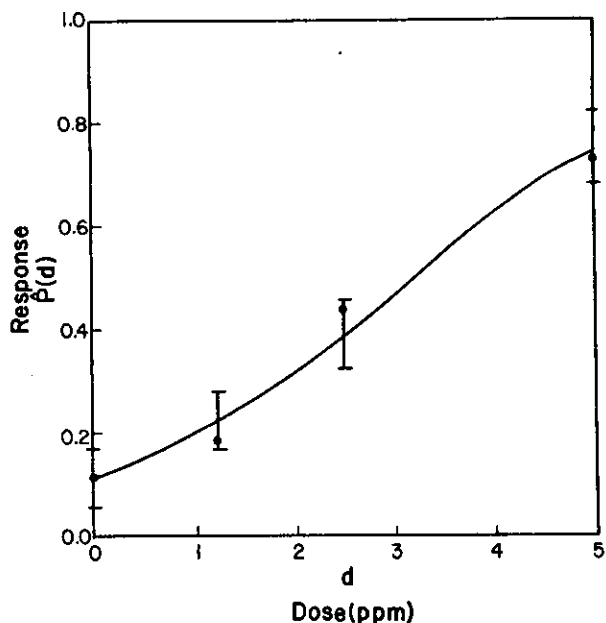
FIGURE 1. Plot in which dieldrin data have been fitted by use of Eq. (1). The dark dots indicate data points. The lines with bars at each end represent 95% two-sided confidence intervals for the risk at the test doses.
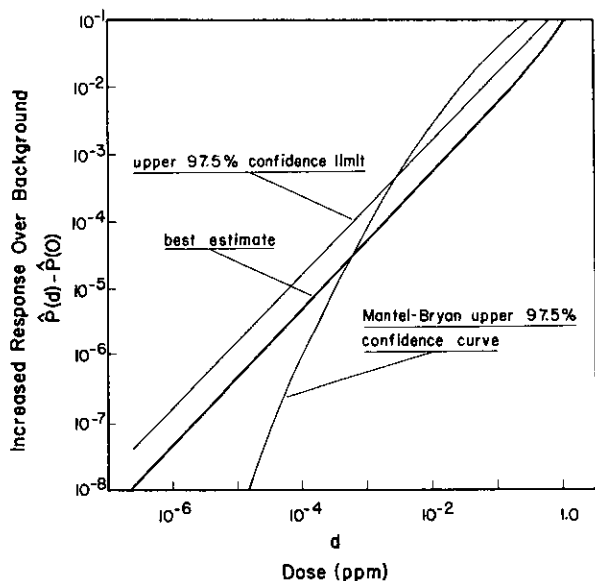

FIGURE 2. Low dose estimates of extra risk based upon the same dieldrin data as Figure 1 by use of both expression (2) and also the Mantel-Bryan approach.

ficult on mathematical grounds (even with nearly perfect data) to reject the hypothesis that the dose-response curve is nearly linear in the dose range corresponding to increased risks over background of about 10^{-6} or less. We demonstrate this with two examples, one in which no background is present

and the other in which there is an expected background effect of about 10%.

Figure 3 shows low dose extrapolations for the hypothetical data simulated as described in Table 2. The expected response frequencies at the test doses lie on the curve

$$P(d) = 1 - \exp\{-1.54 \times 10^{-14}d^8\} \qquad (4)$$

One hundred sets of data were simulated binomially for the eight doses shown in Table 2 with 1000 animals per dose. A maximum likelihood best estimate dose–response curve of the form (1) was calculated for each data set. At each of eight levels of increased risk over background $P(d) - P(0) = 10^{-1}, 10^{-2}, \ldots, 10^{-8}$ the doses corresponding to these risk levels were ranked in decreasing order. The curve labeled "true dose response curve from
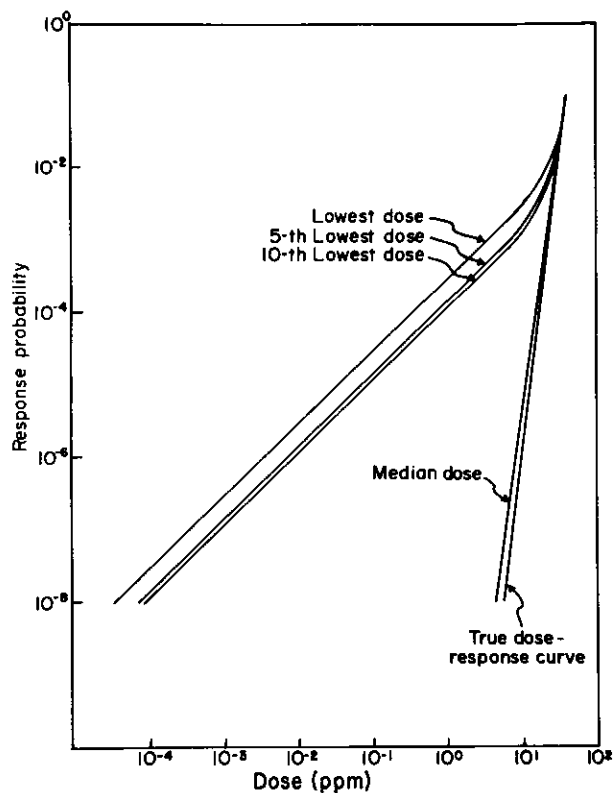

FIGURE 3. Best-estimate low-dose extrapolations for 100 sets of simulated data. Test parameters are shown in Table 2.

which the test data were drawn" is the graph of Eq. (4). The other four curves in Figure 3 were obtained from the best-estimate low-dose extrapolations of the 100 sets of simulated data. For example, the curve labeled "10th lowest dose (out of 100) at each risk level" is the curve constructed by connecting the 10th lowest of the 100 best-estimate extrapolated doses corresponding to a risk level of 10^{-1},

Table 2. Parameters for simulation study (low dose extrapolations are shown in Fig. 3).[a]

Dose	Animals tested	Response probability used
0	1000	0.000
30	1000	0.010
35	1000	0.034
40	1000	0.096
45	1000	0.228
50	1000	0.451

[a]The response probabilities used lie on the curve: $P(d) = 1 - \exp\{-1.54 \times 10^{-14}d^8\}$. One hundred sets of data were simulated, and a best-estimate dose–response function of the form (1) was calculated for each set. The doses corresponding to increased risks of $10^{-1}, 10^{-2}, \ldots, 10^{-8}$ were calculated for each set and each of these 8 sets of 100 doses was tabulated in decreasing order.

with the 10th lowest at a risk level of 10^{-2}, and so on up to the 10th lowest at a risk level of 10^{-8}. The median curve was constructed by connecting the median of the 100 best-estimate extrapolated doses at each of the risk levels $10^{-1}, \ldots, 10^{-8}$.

The median curve is quite close to the true dose–response curve (4), as it should be. However, the striking result of this study is that the envelope curve for the 10th lowest dose at each risk level is almost perfectly linear at risk levels below 10^{-3}. This curve represents roughly an upper 90% confidence interval on the true dose response curve. The study shows that if a carcinogen with the highly nonlinear (almost thresholdlike) curve (4) as its true dose–response curve were to be tested in a perfectly conducted large-scale (8000 animal) test with no background effects present, using the doses shown in Table 2, the test data would probably not reject the hypothesis of linearity at the 10% significance level. In fact, the upper 90% confidence curve from the test would probably not be very different from a simple linear extrapolation of the expected response frequency at the lowest positive dose tested.

When background is present, it is even more difficult to reject the hypothesis of near linearity in the risk range of 10^{-6}. Table 3 illustrates this dramatically. By changing the outcomes of only 11 out of 8000 animals the best-estimate dose corresponding to an increased risk of 10^{-6} over background changes by more than three orders of magnitude.

Intuitively, what is happening here is that slight changes in the experimental outcome involving a few tenths of a percent of the animals at each dose have an appreciable probability of occurring by pure chance. Such changes can transform data which are best fit by an extremely flat dose–response curve into data which are best fit by a dose–response curve with just the slightest hint of a positive linear term. When one extrapolates back to risks of 10^{-6} or less, this small difference in slope is magnified by several orders of magnitude.

Table 3A. Hypothetical data illustrating extreme sensitivity of best-estimate extrapolation to minute changes in data when background is present and true curve is flat.

Dose	Animals tested	Responses Experiment 1	Responses Experiment 2
0	1000	103	100
2	1000	99	99
15	1000	100	105
30	1000	109	112
35	1000	131	131
40	1000	187	187
45	1000	305	305
50	1000	506	506

[a]Best-estimate dose response curves: experiment 1: $P(d) = 1 - \exp\{-(0.105 + 1.1 \times 10^{-15}d^7 + 1.5 \times 10^{14}d^8)\}$; experiment 2: $P(d) = 1 - \exp\{-0.106 + 1.6 \times 10^{-4}d^1 + 1.2 \times 10^{-14}d^8 + 7.2 \times 10^{-17}d^9)\}$.

Table 3B. Low dose extrapolations of data in Table 3A.

Increased risk $P(d) - P(0)$	Dose Experiment 1	Dose Experiment 2
10^{-8}	5.40	7.20×10^{-5}
10^{-7}	7.20	7.20×10^{-4}
10^{-6}	9.60	7.20×10^{-3}
10^{-5}	12.80	7.20×10^{-2}
10^{-4}	17.09	0.72
10^{-3}	22.78	7.20

Summary

An important tentative conclusion from our work is that when dose–response curves that are extremely flat (probit-like) in the low-dose range are compared with dose response curves that are nearly linear in the low-dose range, the nearly linear curves fit the data better for each of the four chemical compounds for which we have analyzed data. A second conclusion is that if anyone has data to support a very flat dose–response curve in the low-dose range, this technique would permit one to infer the flat shape from the data rather than having to assume a flat curve shape, as one implicitly does when using a probit technique. (However, as the discussion above indicates, it is likely that any valid upper 5 or 10% confidence limits on the risk estimates from the data would not differ much from linear extrapolations.) Finally, we believe that simulation studies such as are discussed above could be useful in designing animal carcinogenicity tests and in helping to decide whether a given test design has much chance of accomplishing its objectives.

In our opinion, these results have implications both for test design and for risk assessment. They suggest that if the hypothesis of a nearly linear dose–response curve in the low-dose range cannot be ruled out by assumption, then it seems questionable that it can be rejected by the data, even in cases where it may be false.