Review and Evaluation of Methods of Determining Risks From Chronic Low-Level Carcinogenic Insult *

by Kenny S. Crump and Marjory D. Masterman

BASIC PRINCIPLES AND CONSIDERATIONS

To aid in determining a proper regulatory action regarding a carcinogen that is present in man's environment, whether it be a feed additive, industrial pollutant, or otherwise, it is helpful to have some knowledge about the number of extra cancers that are likely to be caused by the presence of the carcinogen in the environment. It is also helpful to have some knowledge of the likely change in number of extra cancers that would accompany some projected increase or decrease in the level of human exposure occurring either as a result of regulatory action or inaction. This kind of information is usually impossible to obtain directly from human data. For this reason it is often necessary to use data from animal feeding experiments to estimate human risk. This procedure involves two difficult steps: 1) relating the animal risk at high doses to doses very near to zero and 2) relating the animal risk to risk in humans.

Typically, animal experiments use on the order of 100 animals at each experimental dose. If a particular experimental dose causes a lifetime increase in cancer risk of 1/10, this increase can be measured with a small degree of accuracy using 100 animals. But if the increased cancer risk is less than 1/100 this increase will often not even be detectable by an animal feeding experiment. For example, if the true risk is 1/100 it would require that over 400 animals be tested at that dose in order to be 99-percent sure of detecting any carcinogenic response at all (i. e., for there to be a probability of 0.99 that at least one animal gets cancer). If background or spontaneous carcinogenesis is present even larger numbers of animals will be required, On the other hand, the extra human risk that we may want to estimate resulting from environmental exposure is usually (and hopefully) smaller than 1/100 for any given chemical, perhaps on the order of 1/1,000,000. It is clear that it would not be practical to conduct an experiment with enough animals to measure directly an increase in risk this small.

For these reasons the procedure has been developed of conducting lifetime animal feeding experiments using, in addition to a control dose of zero, several doses at which the projected extra cancer risk may be 1/10 or larger, These high-dose data are then used to estimate the extra risk at a dose where the extra risk may be no larger than, say, 1/1,000,000. An equally important variant to this problem is the calculation of the so-called "safe" dose, that is, a dose for which there is some measure of statistical assurance that the extra risk at that dose is no more than, say, 1/1,000,000 These problems are often referred to collectively in the literature as the ''low-dose extrapolation problem. ''

IMPORTANCE OF THE MATHEMATICAL MODEL

Performing a low-dose extrapolation involves the choice of a mathematical function to model the dose-carcinogenic response relationship and the choice of statistical procedures to apply to the mathematical function. The choice for this mathematical function turns out to be extremely crucial to the outcome of low-dose risk estimation. If the assumed relationship between tumor occurrence and dose does not apply in the regions to which the extrapolation is being made, a serious over-

*Excerpt from OTA Working Paper entitled "Assessment of Carcinogenic Risks From PCBS in Food."' A complete copy of the paper can be obtained from the National Technical Information Service. (See app. J.)

estimate of the "safe" dose may result (Mantel and Bryan, **1961, p. 458).** Chand and Heel (1974) compared five standard dose-response models and observed that they could differ by many orders of magnitude at low dose levels for which extra risks are on the order of 1/100,000,000.

It might be supposed that it should be possible to discriminate among the various potential dose-response functions on the basis of experimental data but, unfortunately, two different dose-response functions can often fit experimental data equally well but still differ by several orders of magnitude at very low doses. Moreover, even if a particular dose-response function were to give a significantly better fit to data than several others this would still not furnish assurance that this function would necessarily correlate in any way with the true dose response at very low doses where it is not feasible to measure the true extra risk directly. As a consequence of the great disparity of dose-response functions at low doses it is imperative that the dose-response function be selected, neither arbitrarily nor solely on the basis of how well it can be made to fit experimental data, but, insofar as is possible, it should reflect known or at least plausible information regarding the biological mechanisms through which a chemical induces or promotes cancer.

WHAT SHAPE SHOULD BE EXPECTED FOR THE DOSE-RESPONSE CURVE AT LOW DOSES?

Tumors of so many different types arise in such a diversity of different tissues, their etiology is so little understood, and the agents that cause tumors affect a subject in such diverse ways, that it might seem that no general conclusions can be drawn. However, for a certain broad class of "directly acting" chemical carcinogens the range of uncertainty associated with the shape of the dose-response curve at low doses can be greatly narrowed. As used in this paper, the term "directly acting carcinogen" encompasses (Guess, Crump, and Pete, 1977) carcinogenic agents for which either the agent itself or a metabolize acts directly at the cellular level and produces a heritable change that eventually leads to the formation of a tumor. Carcinogens that are carcinogenic by reason of their mutagenicity should fall into the category of "directly acting carcinogens. " Accordingly, carcinogens that test positively using the Ames mutagenicity screening test for carcinogenicity are very likely to be directly acting (see McCann and Ames, 1976). In a recent study (McCann, Choi, Yamasaki, and Ames, 1975) in which about 300 carcinogens and noncarcinogens were tested using the Ames test, 90 percent (157 out of 175) of the carcinogens were mutagenic including almost all of the known human carcinogens. This indicates that the class of directly acting carcinogens may encompass most of the known carcinogens,

A partial solution to the low-dose extrapolation problem for the case of directly acting chemical carcinogens has been given in Peto (1977), Crump, Heel, Langley and Peto (1976), and Guess, et al. (1977). The key result is that, at least as long as background carcinogenesis is present, we should expect the dose-response curve not to be absolutely flat at zero dose. What this means is simply that when risk is plotted against dose response on ordinary linear scales, the tangent line to the dose-response curve at zero dose should have a positive slope. When a dose-response function has this property we will say it is linear at low dose. This simple property can have far-reaching consequences on low-dose extrapolation. For example, consider the two potential dose-response functions:

1. $0.1[(99/999)d + (900/999)d^2]$ and
2. $0.1d^2$ for the dose interval $0 \leq d \leq 3$.

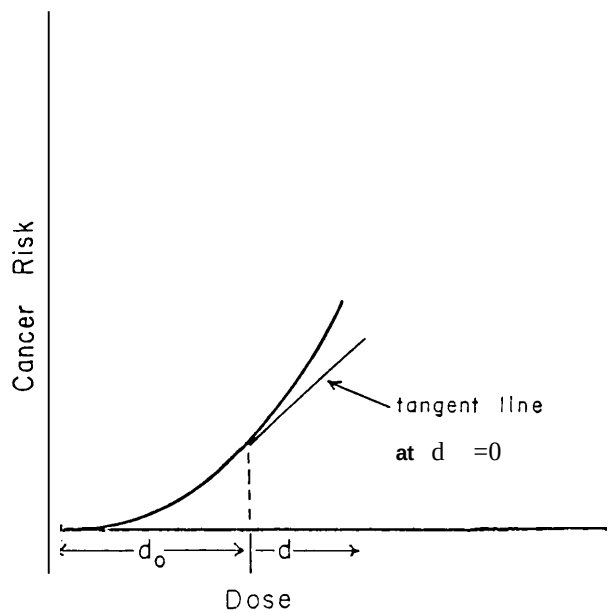
Both of the curves give a risk of 1/10 at a dose of d = 1 and are practically indistinguishable at higher doses. However, at a dose of d = 1/1,000: 1) predicts a risk of 1/100,000 and 2) predicts a risk of 1/10,000,000, a difference of two orders of magnitude. We note also that 1) has a tangent line with a positive slope at d = O whereas 2) does not.

One explanation of why the dose-response function should be linear at low dose when background is present may be found in Crump, et al. (1976) and Peto (1977) and will be briefly outlined here. When background carcinogenesis is present, the cellular mechanism through which the test agent produces cancer should already be operative in producing background tumors. When this is true the effect of the test agent is to add to an already ongoing process. The result of this additive effect is illustrated in figure D-1. The dose-response curve is for all tumors produced through the mechanism through which the test agent acts. Background carcinogenesis is allowed for in the figure by an effective background dose d_0. We see that, in this case, the added risk caused by a dose d of the test agent should be ex-

pected to increase approximately linearly near d = O (i.e., the tangent line at d = O will have a positive slope). Implicitly assumed by the way figure D-1 is drawn is the fact that an added dose of a carcinogen acting through this mechanism does not produce a smaller risk, If background carcinogenesis is allowed for as in figure D-1 by positing an effective background dose d_0 that is estimated from the data, then the wide range of risks obtained using different models effectively disappears (Pete, 1977). We note that the existence of a tangent line with a positive slope at zero dose does not, in itself, imply any lower bound for extra risk at low doses since the slope of the tangent line could possibly be very small.

The evidence given above for a positive slope to the dose-response function at zero dose applies particularly to the case in which background carcinogenesis is operative. This does not imply that we expect the dose-response curve not to be linear at low dose in the absence of background carcinogenesis. For example, the multistage models of cancer (Armitage and Doll, 1961) are a fairly broad class of models in which it is assumed that a number of events are required to occur at the cellular level to initiate cancer. Although all models in this class are linear at low dose provided background carcinogenesis is present, a sizable subclass of them are linear at low dose in the absence of background carcinogenesis (Crump, et al., 1976).

Figure D-1 .—Illustration of Why the Dose" Response Curve Should Be Linear at Low Dose in the Presence of Background Carcinogenesis. d_o Is the Effective Background Dose and d Is the Dose of the Carcinogen of Interest



Watson (1977) has recently proposed a more specialized model for cancer induction and promotion based on reversible epigenetic cellular changes. Watson concludes that his model supports the low-dose-linearity hypothesis and states "as suggested by a different argument of Crump, et al. . . . it is reasonable to assess the risk due to an additional carcinogen at low constant dosage by a linear relation. "

The evidence for low-dose linearity given above applies mainly to directly acting carcinogens. An indirectly acting carcinogen might be one that causes some gross physiological change such as suppression of ovulation that could predispose the subject to cancer. For such carcinogens the shape of the dose-response curve at low dose is highly speculative. There could possibly be a threshold dose below which the agent has no carcinogenic effect at all on an individual, However even if a threshold mechanism is operative, there is likely to be considerable variation in individual thresholds in a large population. Consequently the dose-response curve for the entire population could still exhibit a linear trend at risks as low as 1/1,000,000 or lower.

The effects of metabolic activation and detoxification on carcinogenic dose response have been recently considered by Cornfield (1977) through a kinetic model that encompasses free toxic substance, metabolize, deactivator, and the interactions of these substances. Only a steady-state situation is studied in that variation over time of the concentrations of these agents is not considered. The model predicts a threshold dose below which there is no carcinogenic risk under the assumption that the deactivator is 100-percent efficient in deactivating the carcinogen. However, in a naturally occurring process it is likely that deactivation would not be perfect and would be less than 100-percent effective in always combining with 100 percent of the carcinogen before an amount of the active metabolize reaches a cancer target site, Any of a number of modifications to the model to allow for nonperfect deactivation would rule out a threshold and would lead directly to a model for which carcinogenic response varies linearly with dose at low doses, Cornfield's own modification of perfect deactivation, that of al-

lowing the deactivating reaction to be reversible, leads, as Cornfield points out, to a model which is linear at low dose. This occurs regardless of how slowly the reverse reaction takes place, as long as the possibility is not eliminated entirely. Furthermore, even in the extremely unlikely case of perfect deactivation, an otherwise realistic model should still imply low-dose linearity since the theoretical time required for perfect deactivation would not be zero and would likely be infinite.

For most, perhaps all, carcinogens, the mechanisms through which cancer is produced are not sufficiently understood so that the shape of the carcinogenic-response curve can be theoretically predicted with certainty. As pointed out earlier, neither can experiments of sufficient size be conducted that would permit direct experimental investigation of the dose-response curve at low dose. We have noted that there are plausible arguments that the dose-response curve is linear at low dose for many carcinogens. On the other hand, this author knows of no serious proposal of a mechanism that would lead to a more conservative dose-response relationship such as the risk varying approximately as the square root of dose at low dose. In view of these uncertainties it would seem reasonable to base estimates of added risk of cancer on a mathematical model that encompasses low-dose linearity unless, of course, the mechanism through which the carcinogen operates is sufficiently understood so that low-dose linearity can be conclusively ruled out. Once the principle of low-dose linearity is accepted the problem of estimation of risks at low dose is nearly solved. This is because the disagreement between the upper statistical confidence bounds on risk at low doses based on a model that incorporates low-dose linearity and one that does not is typically several orders of magnitude whereas the corresponding disagreement between two reasonable models both of which incorporate low-dose linearity is usually much less than this.

SPECIFIC METHODS FOR LOW-DOSE RISK ESTIMATION

Mantel= Bryan

The Mantel-Bryan procedure as originally proposed (Mantel and Bryan, **1961)** and "improved" (Mantel, Bohidar, Brown, Ciminera, and Tukey, 1975) is for the purpose of conservatively choosing a "safe" dose of a carcinogen, a "safe" dose being defined as one for which it can be expected that, with a given level of statistical assurance (e.g., 99 percent), the true dose producing a preassigned "safe" level (e. g., 1/1,000,000) of risk will lie above the "safe" dose. In the Mantel-Bryan procedure the mathematical model used for the dose-response model is the probit function:

$$P(d) = C + (1-C)\int_{-\infty}^{\alpha + b\log_{10} d} (2\pi)^{-1/2} \exp\{-x^2/2\}\, dx \qquad (1)$$

where d represents the dose of the carcinogen and P(d) represents the probability of a cancerous response in an animal subjected to a dose d. The parameters in the model are an intercept parameter a, a probit slope parameter b, and C, which represents the probability of a response in untreated animals. The parameter b is not estimated from the data but rather is arbitrarily set equal to 1. This choice is stated as being conservative (Mantel and Bryan, 1961), the argument for this being that typical dose data exhibit a probit slope in the experimentally observable region above 1-percent incidence that is greater than one, In Mantel and Schneiderman (1975) it was observed that a set of DES data (Gass, Coats, and Graham, 1964, C3H females) exhibited a probit slope of one-half, but the general use of a probit slope of b = 1 was still suggested.

With the probit slope parameter fixed at b = 1 the remaining parameters, a and C, are estimated from the data and then adjusted so as to produce a higher level of risk at a given dose that corresponds to an upper 99-percent statistical limit on the true risk at a given dose. The safe dose is then determined to be the one producing a given low risk (e. g., 1/1,000,000) based on the adjusted values of a and C.

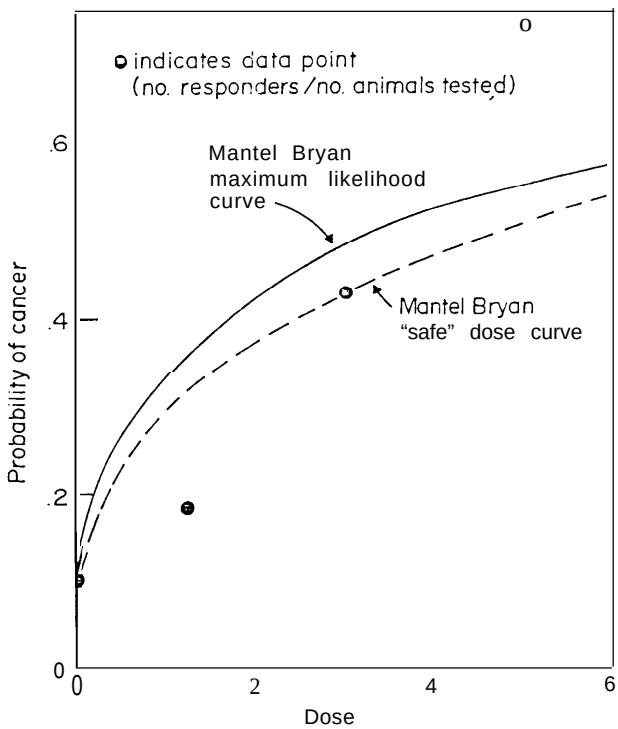
As pointed out in Mantel, et al. (1975), the Mantel-Bryan procedure rewards larger and better experiments in that the more evidence there is of safety, the higher the calculated safe dose will be. However this advantage should be shared by any extrapolation method that uses reasonable statistical procedures.

Some have considered the Mantel-Bryan procedure to be too conservative (Federal Register, vol. 42, 1977, p. 10419) in that it involves three conservative choices (99-percent statistical assurance, lifetime risk of 1/1,000,000, and probit slope set equal to 1) and that any one of these assumptions alone could provide adequate protection to the public. The first two of these choices are regulatory decisions that would have to be made with any extrapolation procedure. However, the arbi-

trary selection for the slope parameter seems to be peculiar to the Mantel-Bryan procedure. To investigate its effect upon the extrapolation procedure a typical fit is presented in figure D-2 (from Crump, 1977b) of the Mantel-Bryan probit curve (equation 1) to experimental carcinogenicity data when the probit slope parameter is fixed at b = 1. As can be readily seen the probit curve typically provides a very poor fit, curving downward even when the trend of the data is toward an increasingly upward curvature. This typically bad fit to data of the probit curve raises serious questions regarding the validity of statistical procedures associated with the Mantel-Bryan method (see Salsburg, 1977; and Crump, 1977b).

On the other hand, the Mantel-Bryan procedure utilizing the choice b = 1 was put forth as conservative procedure and it gives that appearance in figure D-1 since the probit curve appears to lie far above the trend of the data at the lowest doses. However as mentioned earlier and also pointed out by Mantel (Mantel and Bryan, 1961, p. 458) a procedure may, while appearing conservative at experimental dose levels, at the same time seriously overestimate the "safe" dose (i. e., be seriously anticonservative) if the assumed dose-response relationship does not apply at the low risk levels to which extrapolation is being made. Thus, before the degree of conservation can be evaluated for any procedure, the properties of the dose-response curve at very low doses must be evaluated,

Figure D-2.—Typical Fit of Mantel" Bryan Curve to Experimental Data (From Crump, 1977b)



As described earlier, there are strong arguments that indicate the dose-response curve should be "linear at low dose" particularly for directly acting carcinogens in the presence of background carcinogenesis. This has led Peto (1974) to recommend extrapolation procedures using only dose-response functions from a class containing only dose-response functions that are linear at low dose, At the very least, however, it would seem prudent not to go to the opposite extreme and use a dose-response function that rules out linearity at low dose by assumption, However, the Mantel-Bryan procedure, through its use of the probit curve (equation 1) rules out linearity at low dose in favor of a "flatness property" (see Hartley and Sielken, 1977; Mantel, 1977; and Crump, 1977b) at low dose which is anticonservative to the extreme. This property implies that mathematical derivatives of all orders of the probit curve approach zero (through positive values) as the dose approaches zero. This unusual property is more often discussed within the context of mathematical oddities rather than in connection with a scientific investigation. It implies that if the true dose-response curve comes from an extremely broad class of functions known as analytic functions and which pervade scientific applications of mathematics, then the probit curve will eventually underestimate the true risk at low doses. Furthermore, at low enough doses, the degree to which the risk will be underestimated will be arbitrarily large (i.e., the ratio of the true risk to the probit estimate will grow arbitrarily large).

It was emphasized earlier that it is important when extrapolating to low doses for the assumed dose-response function to incorporate known or at least plausible facts about the mechanisms of carcinogenesis. In neither the original paper (Mantel and Bryan, 1961) nor in the paper outlining the improved version is biological justification given for the selection of a curve having the above described "flatness property." It should be mentioned at this point that the incorporation of background carcinogenesis into the Mantel-Bryan probit model (equation 1) using the parameter C implies that the mechanism through which the test carcinogen produces cancer is independent of the mechanisms through which all of the back-

ground cancers are produced (Crump, et al., 1976). In keeping with the discussion in the last section it would seem more proper to incorporate background into the probit model by positing an effective background dose d_o which adds to the dose d of the test carcinogen. If background is incorporated in this way the probit curve no longer has the "flatness property" and becomes linear at low dose (Guess, et al., 1977). In fact. with background incorporated in this way, the probit curve assumes a shape similar to the one-hit model, sometimes referred to as the most conservative of all procedures (e.g., Mantel, 1977).

Even though the "flatness property" implies the probit curve should at suitably low doses be anticonservative to the extreme, the "flatness" property holds only for doses approaching zero and the feature of arbitrarily fixing the probit slope at 1 mitigates the anticonservativeness implied by the "flatness" property at any given low dose (although the property itself will hold for all choices of the parameters a, b, C), The cancer risks that are typically extrapolated to are in the risk ranges 1/10,000 to 1/100,000,000. We will examine the outcome of Mantel-Bryan extrapolations to these risk levels in a later section when we compare them with extrapolations based on the multistage model.

Linear Extrapolation

The technique for linear extrapolation was recommended by Heel, Gaylor, Kirschstein, Safiotti, and Schneiderman (1975) for use on an interim basis until better procedures could be developed. The procedure is straightforward and is based on an assumed linear relationship between dose and response at low dose. The procedure utilizes only the data for the group of control animals and a single other dose group, usually either the highest dose that elicits no response or else the lowest dose that elicits some response. In the case there are no cancers in the control animals the "safe" dose, based on a maximum risk of 1/1,000,000 and 99-percent statistical assurance, is calculated as follows: An upper 99-percent confidence bound is calculated for the cancer risk in the dose group of animals. From this risk and dose one extrapolates back toward zero dose and zero risk using a straight line relationship. The dose corresponding to a risk of 1/1,000,000 on this straight line is taken to be the "safe" dose. If there are cancers in the control animals this procedure is modified to allow for the statistical treatment of the response in the control group while retaining the straight line relationship. When data at other experimental doses are available this method of linear extrapolation has the obvious shortcoming of not fully utilizing the available data.

A linear dose-response curve is linear at low dose but the converse is not necessarily true. A curve can be linear at low dose and still have a high degree of nonlinearity at higher doses.

Linear extrapolation is viewed by some as a very conservative procedure. For example, comments were made during the decision on which extrapolation procedure to incorporate into the SOM document to the effect that linear extrapolation is the most conservative of all procedures. Crump, et al. (1976) examined the extent of the conservatism of a linear dose-response function when compared with a multistage dose-response model (Armitage and Doll, 1961). The multistage model assumes that a cell must go through a number of different stages before cancer is initiated in that cell and the model can encompass a high degree of nonlinearity. It was determined that the maximum possible degree of conservatism of a linear model relative to a multistage model depended rather heavily on the incidence at the experimental dose relative to the background incidence. (This is consistent with the general relationship between background carcinogenesis and linearity at low dose as discussed earlier.) For example, when the incidence at the experimental dose is four times the incidence at zero dose the extra incidence at low dose derived from the linear dose response differs from the incidence derived from the multistage model by, at most, less than a factor of 2.5 regardless of the number of stages in the multistage process. Thus, when background carcinogenesis is present, the linear dose-response curve is not overly conservative relative to the multistage dose-response curve. In fact the linear dose-response curve is anticonservative when compared to the one-stage or one-hit models.

Linear extrapolation has long been proposed for use in radiation carcinogenesis (see Brown, 1976, for a review of the relevant reports). The BIER (1972) report on radiation risks from the National Academy of Science recommended linear extrapolation as a "best estimate" approach as opposed to a conservative approach. Brown reviewed arguments both for and against linearity and concluded that "linear extrapolation of human data from high dose of low LET radiation cannot be said to overestimate the risk at low

doses. In fact, there is some doubt as to whether the risk is not underestimated. "

Certainly much remains to be learned about both radiation and chemical carcinogenesis. However, if both radiation and chemicals cause cancer through similar mechanisms then it should be expected that there would also be similarities between the respective carcinogenesis dose-response functions. Direct damage to DNA by the carcinogenic agent has been implicated as one cancer-initiating mechanism for both radiation and chemicals (Brown, 1976; and McCann and Ames, 1976). Thus, the findings related to the potential linearity of the dose-response function for radiation has implications for chemical carcinogens as well, particularly for "directly acting" carcinogens.

Extrapolation Methods Based on the Multistage Model

Two methods of low-dose extrapolation which are alternatives to the Mantel-Bryan or linear procedures have recently been proposed independently by Guess, Crump, and Deal (Guess and Crump, 1976, 1978; and Crump, Guess, and Deal, 1977) and Hartley and Sielkin (1977). Both of these methods utilize a multistage dose response function of the form:

$$P(d) = 1 - \exp\{-(q_0 + q_1 d + q_2 d^2 + \ldots + q_k d^k)\} \quad (2)$$

where $q_0, q_1, \ldots, q_k$ are all non-negative parameters to be estimated from the data. This dose-response function is general enough to yield a considerably wide range of responses at low dose. On the one hand, if $q_1 > 0$ and $q_i = 0$ for $i \geq 2$ the dose-response function (equation 2) becomes the one-stage model which yields risks at low doses comparable to what would be obtained with linear extrapolation, On the other hand, the model can produce risks even as low as the probit curve (equation 1) down to any fixed positive low dose. Thus this model is capable of fitting both highly linear and highly nonlinear dose-response relations. Since the model has the property of "linear at low dose" if $q_1 > 0$ and does not have this property if $q_1 = 0$, use of this model does away with having to make the arbitrary but crucial decision of having to either assume linear at low dose as in linear extrapolation. Thus "safe" doses computed using this model should provide a more realistic measure of the true uncertainty of low-dose extrapolation than would "safe" doses based on either an assumed linear curve shape or an assumed highly nonlinear curve shape.

The dose-response relation (equation 2) contains all of the Armitage and Doll **(1961)** multistage dose-response models as special cases but also contains curves that are much flatter **at low** dose than any of the multistage curves.

The two extrapolation procedures based on equation 2 have some features that are different. When computing "most likely" estimates the procedure of Guess, Crump, and Deal uses an infinite dimensional maximization procedure so that it is not necessary to specify a value of k, the degree of the polynomial in equation 2. However the two methods differ chiefly in the way the statistical confidence intervals are computed. There have not yet been sufficient comparative calculations made to determine how safe doses may differ using the two approaches, Mantel (1977) has made a critical review of the statistical procedure used by Hartley and Sielkin for calculating the "safe' dose.

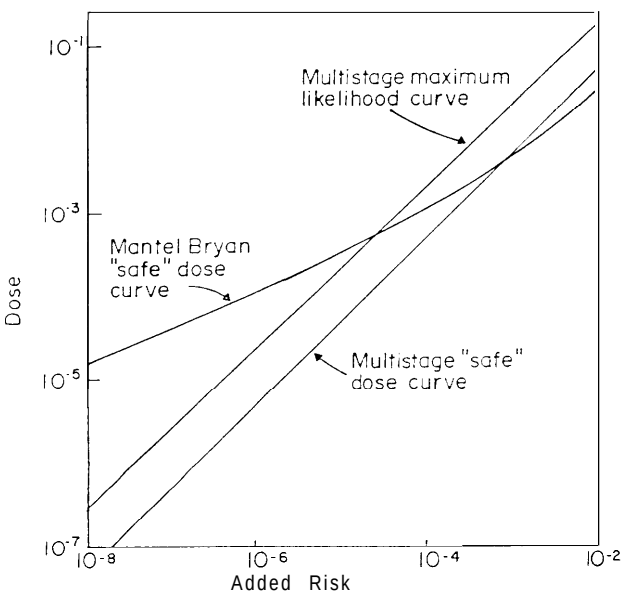
Both the Hartley and Sielkin and the Guess, Crump, and Deal (as extended by Daffer et al., 1979) procedures can utilize times at which cancer is detected in the experimental animals rather than just the dichotomous information of whether or not an animal contracted cancer before it died of some other cause or before the termination of the experiment, The utilization of time data in low-dose extrapolation is important for at least two reasons: 1) the age at which cancers occur should be important in assessing the magnitude of the harmful effect of a carcinogen on man (e.g., cancers that occur early in life should be viewed as more serious than those that occur in extreme old age); and 2) in many animal carcinogenicity experiments the response data at the highest doses lies below the trend of the lower dose data, This sometimes appears to be due to the fact that at the highest doses some of the animals are being poisoned by the chemical before they have a chance to develop cancer, When this occurs the high-dose data is often just deleted from the analysis. However if the times at which the animals die are properly used the high-dose data might not appear anomalous. More research needs to be done on the proper utilization of animal time of death data to assess the harmful effects of chemicals to man.

Comparisons Between Mantel-Bryan and Multistage Extrapolations

To compare low-dose extrapolations using the Mantel-Bryan probit model (equation 1) with

those using the multistage model (equation 2) we present figure D-3 based on the same data as figure D-2. In this figure, the Mantel-Bryan "safe" dose is plotted on a log-log scale as well as both the multistage "most likely" curve and the "safe" dose based on (equation 2) and computed as outlined in Crump, et al. (1977). A 99-percent statistical assurance was used for both "safe" dose curves. We note that the Mantel-Bryan "safe" dose lies above the multistage safe dose curve at values of added risk below 5 x 10^{-4}. The Mantel-Bryan "safe" dose curve lies above the multistage "safe" dose curve by a factor greater than 20 for an added risk of 10^{-6} and by a factor greater than 300 for an added risk of 10^{-8}. Because of the "flatness" property of the Mantel-Bryan probit function (equation 1) described earlier, the Mantel-Bryan "safe"' dose curve will lie above the multistage safe dose curve by arbitrarily large factors at extreme low doses. Guess, et al. (1977) have compared the Mantel-Bryan "safe" dose curves to the multistage "safe" dose curves and found this to be a typical situation. Thus it is clear that if the true dose-response curve could be similar to the multistage dose-response function (equation 1) then the Mantel-Bryan procedure could not be justifiably called conservative (see also Crump, 1977 for further discussion of this point.) On the other hand, we have seen that there are quite plausible arguments for the true dose-response curve to have the same shape at low dose (linear) as the estimated multistage curve.

Figure D-3.—Comparisons of "Safe" Doses Computed From the Mantel" Bryan Procedure and From a Procedure Based On the Multistage Model (From Crump, 1977b)



We note that both the multistage "safe" dose curve and "most likely" curve have a slope 1 in figure D-3 as plotted on the log-log scales which is equivalent to the curves being linear at low dose. The fact that the "most likely" curve has slope 1 is due to the fact that with this particular data set the linear coefficient q_1 in equation 2 will always be linear at low dose regardless of whether or not the linear coefficient q_1 is estimated to be positive. This property should be shared by any valid statistical procedure based on a dose-response function that does not rule out linearity at low dose by assumption as Mantel-Bryan does. Just as it is not possible to prove statistically the existence of a threshold, it is likewise not possible to rule out the possibility that the true dose-response curve is linear at low dose on the basis of statistical analysis. (See Guess, et al. (1977) for a thorough discussion of this important point.) The Mantel-Bryan obtains "safe'" dose estimates which are considerably higher than those obtained using the multistage model because it assumes away linearity at low dose, an assumption that we have seen is probably unwarranted for that majority of carcinogens which are classified as ''directly acting" carcinogens.

Since extrapolation based on a model such as the multistage model (equation 2) must always be linear at low dose, the question arises as to how different the result will be from simple linear extrapolation. For some data the difference will be minimal. For example, for the data upon which figure D-2 is based, "safe doses" computed using the multistage model are almost identical with *'safe'" doses based on linear extrapolation. For some data sets, however, the difference could be considerable. For example, with the Gass, et al. (1964) DES using the C3H female mice, the "safe" dose based upon linear extrapolation is lower than the "safe" dose based on the multistage model by a factor of about five and there are other data sets where this difference is greater than an order of magnitude.

The Gamma Multihit Carcinogenesis Dose-Response Model

Rai and Van Ryzin (1978) have proposed basing risk estimation on the gamma multihit model:

$$P(d) = C + (1-C)\int_0^{\lambda d} \frac{u^{k-1}e^{-u}}{(k-1)!}\,du \qquad k \geq 0 \qquad (3)$$

$$= C + (1-C)\left\{1 - e^{-\lambda d}\left[1 + \lambda d + \frac{(\lambda d)^2}{2!} + \ldots + \frac{(\lambda d)^{k-1}}{(k-1)!}\right]\right\} \qquad k = 1, 2, \ldots$$

where P(d) is the lifetime probability of cancer in a tissue when subjected to a constant dose rate, d, of the carcinogen. This model is obtained by assuming that cancer due to the carcinogen occurs randomly according to a Poisson distribution. The manner in which background carcinogenesis is incorporated into the model is equivalent to assuming that the event "cancer occurs due to the action of the carcinogen" is independent of the event "cancer occurs spontaneously, " This assumption would not apply, for example, to processes in which the effect of the carcinogen is to speed up the rate at which the "spontaneous" events occur which lead to the background cancers. At low dose rates, the response is approximately given by:

$$P(d) \cong \lambda^k d^k / k!.$$

Consequently, this dose response model is linear at low dose rates when and only when k = 1. Rai and Van Ryzin calculate confidence intervals for added risk at a given dose and for the dose producing a fixed added risk using asymptotic maximum likelihood theory. Although in the theoretical development k must be an integer, in the applications k is allowed to assume any positive value.

Lower statistical confidence limits on the dose producing a given low amount of extra risk ("virtually safe dose" or "VSD") computed using this procedure can be compared with those computed from the multistage model (equation 2) by considering two general classes of data.

If the data exhibit a general downward curvature as illustrated in figure D-4a, lower confidence limits on a VSD computed from the gamma multihit model (equation 3) should generally be less than or equal to corresponding lower confidence limits computed from the multistage model (equation 2). In certain instances the gamma multihit lower limits could be much less than the corresponding multistage lower limits. This could occur when the data is consistent with k <1 in the gamma multihit model (presumably corresponding to a fraction of a hit).

On the other hand, if the data exhibit a general upward curvature as illustrated in figure D-4b the reverse situation will hold; gamma multihit lower confidence limits on VSDs will generally be greater than or equal to corresponding multistage limits, Gamma multihit limits will not in general share the low-dose linearity of multistage limits by orders of magnitude at low doses. The reason these large differences can occur is as follows. The multistage family of models contains members which are simultaneously linear at low dose and exhibit upward curvature at moderate doses. For example, the particular multistage model:

$$1 - e^{-(q_1 d + q_2 d^2)} \qquad q_1, q_2 > 0$$

is linear at low dose since q_1 >0 and still can exhibit upward curvature at moderate doses since q_2 >0. This means there are dose-response curves in the multistage class which are both linear at low dose and can adequately describe data of the type exemplified by figure D-4b. On the other hand, this will generally not be true of gamma multihit models. All dose-response curves in the

Figure D-4a.— Example of Data Exhibiting Downward Curvature

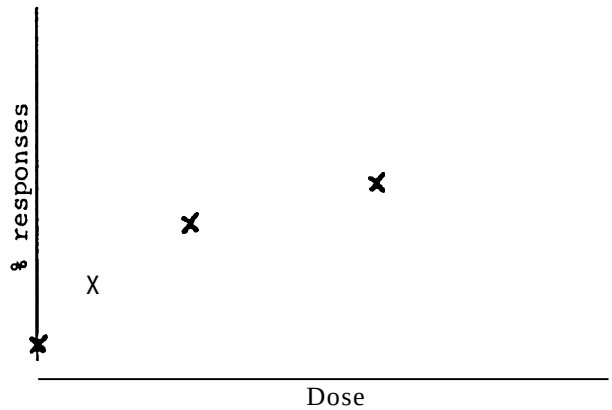


Figure D-4b.—Example of Data Exhibiting Upward Curvature

gamma multihit class that are linear at low dose must exhibit downward curvature and consequently would generally not be consistent with the data in figure D-4b. When this is true gamma multihit lower confidence limits on VSDs will be sublinear (e.g., quadratic) at low dose and consequently much larger than the multistage confidence limits.

Confidence limits based on the gamma multihit model will be approximately correct whenever this model is the correct model, A similar statement could be made for the probit model, multistage model, or any other model to which reasonable statistical methods are applied. However, there may be considerable uncertainty as to what the true model may be in a particular situation. The multistage model not only reflects some reasonable assumptions regarding the carcinogenic process which dovetail nicely with epidemiological data for many cancers (Pete, 1977 b), but it also reflects some of the uncertainty with regard to the true model by virtue of encompassing a relatively large class of dose-response functions, For example, as noted earlier, the multistage class contains dose-response functions which are linear at low dose and also exhibit upward curvature at moderate doses. On the other hand, the gamma multihit class is more restrictive at this point in that it does not permit such behavior. Is this extra restrictiveness of the gamma multihit model justified? To help answer this question, consider the following modification to this model. Suppose that the hits (phenomenological events which are required to occur in a tissue in order that a cancer appear) can possibly occur spontaneously in the absence of the carcinogen. Part of the effect of the carcinogen could then be to speed up the rates at which the spontaneous hits are occurring. For example, if one of the "hits" is in incorrect base substitution in DNA during mitosis, the carcinogen could speed up the rate at which these "hits" are occurring in individual cells by providing more aberrant bases for use in such an incorrect substitution. With this modification to the model, the gamma multihit lower confidence limits on VSDs will no longer be sublinear at low dose and should be at least as small as corresponding limits calculated from the multistage model (equation 2). Thus, in order to obtain sublinear lower confidence limits on VSDs with the gamma multihit model a modification such as the one described above must be ruled out, not on the basis of data, but by assumption.

By way of summary, confidence limits based on the multistage model will always be linear at low dose. Confidence limits based upon the gamma multihit model may be either "super linear" (corresponding to $k < 1$) or "sublinear" (corresponding to $k > 1$). Superlinearity is achieved by making the model too broad in that a fraction of a hit is allowed which has no biological basis. On the other hand, sublinearity is achieved by making the model possibly too restrictive in that models which are reasonable from a biological viewpoint are ruled out by assumption.

In addition to questions as to the appropriateness of the gamma model, there is also a very significant difficulty with the statistical procedure applied to the model in the Food Safety Council (1978) report. This difficulty is illustrated in table D-1. Here are listed lower confidence limits for VSDs for the gamma multihit model which were taken from table D-1 in Food Safety Council (1978). For each of the 14 data sets a goodness-of-fit test was conducted for compatibility of the data to the subclass of gamma models for which $k = 1$. As indicated by table D-1, the subclass of models with $k = 1$ provided an adequate fit in 10 out of the 14 data sets. For these 10 data sets are listed the VSD predicted by the best fitting gamma model with $k = 1$. The significant point is that in 8 of these 10 cases, the lower confidence limit on the VSD is greater than the VSD from the model with $k = 1$, greater in some cases by enormous factors, e.g., a factor of 3,600 for aflatoxin B_1, a factor of 1,000 for diethyl-nitrosamine, and a factor of 18,000 for sodium saccharin. Put another way, in the case of the sodium saccharin data, there is a member in the gamma multihit class which fits the data quite adequately and for which the VSD is less by a factor of 18,000 than the lower confidence limit for the VSD reported in Food Safety Council (1978). This obviously implies there is a serious problem with the computation of the lower confidence limits for the VSD. These lower confidence limits are much too large, even based on assuming the gamma multihit model is the correct model, for these 8 data sets. This problem also exists for the 4 data sets which could not be fit adequately by a model with $k = 1$. This is well-illustrated by the graph in Food Safety Council (1978) for the NTA data. In this graph, the lower 97.5-percent confidence limit for the VSD corresponding to a risk of 10^{-6} falls almost directly on top of a data point for which the measured risk is 1/91. Put another way, even though the measured risk at this dose is 1/91, the statistical procedure used indicates, with 97.5-percent assurance, that the risk is no greater than **1/1,000,000.**

These difficulties can be overcome by using a different statistical procedure. However, when this is done, the lower confidence limits will be

Table D-1.—Comparison of Lower 97.5-Percent Confidence Limits on Virtually Safe Doses From Gamma Multihit Model Given in to Virtually Safe Doses From Best Fitting Gamma Multihit Model With k = 1

No.	Substance	No. of experimental dose levels	Dose units	Lower 97.5% confidence limit on virtually safe dose at risk level 10^{-6}	Asymptotic sign. probability of X^2 goodness of fit test with k = 1	Virtually safe doses from best fitting gamma model with k = 1
1	NTA	6	% in diet	.70	p < .005 (4 d.f.)	significant lack of fit
2	Alfatoxin B_1	6	ppb	.12	p = .25 (4 d.f.)	.000033
3	Eythlenethiourea	6	mg/kg	1.80	p < .005 (4 d.f.)	significant lack of fit
4	2,3,7,8 tetrachlorodibenzo-p-dioxin	5	mg/kg	.0013	p = .58 (3 d.f.)	$.52 \times 10^{-5}$
5	Dimethylnitrosamine	5	ppm	.035	p = .12 (3 d.f.)	.000032
6	Vinyl chloride	6	ppm	1.2×10^{-10}	p = .05 (4 d.f.)	.020
7	Hexachlorobenzene	5	ppm	.000029	p > .99 (3 d.f.)	.00020
8	Botulinum toxin-type A	11	mg	.0120	p <.005(9 d.f.)	significant lack of fit
9	Bischloromethyl ether	6	No. of 6 hr exposures	.011	p = .49 (4 d.f.)	.00016
10	Sodium saccharin	5	% in diet	.76	p = .41 (3 d.f.)	.000042
11	Ethylenethiourea	6	ppm	25.3	p < .005 (4 d.f.)	significant lack of fit
12	Dieldrin	4	ppm	.0022	p = .22 (2 d.f.)	$.57 \times 10^{-5}$
13	DDT	5	ppm	.013	p = .30 (3 d.f.)	.00028
14	Span oil	5	% in diet	.00037	p = .84 (3 d.f.)	.000037

many times smaller than those reported in Food Safety Council (1978). For data sets such as the 10 referred to in table D-1 which are compatible with k= 1, lower confidence limits for VSDs must essentially be calculated from a model with k <1 and consequently be smaller than those calculated from the one-hit model. This will lead, in many cases, to super-small confidence limits, as illustrated by the lower confidence limit for vinyl chloride, For those data sets, such as botulinum toxin-type A, which are incompatible with k = 1 because of their strong upward curvature, a lower confidence limit for the VSD would still be larger than one computed from the one-hit model. However, the theoretical objections raised earlier to the mathematical form of the gamma model would still apply.

It may not be generally realized that it is uncommon to find data sets which are incompatible with k = 1 by reason of strong upward curvature. The four data sets in table D-1 which fall into this category all involved a large number of animals and a number of different experimental dose levels. These four data sets contained from 330 animals distributed over 11 dose levels to 720 animals distributed over 6 dose levels. Most carcinogenicity bioassays are smaller than these, For example, the standard bioassay for a given sex and strain in the National Cancer Institute program seems to be 200 animals distributed over 4 dose levels. For experiments of this size it is more difficult to rule out k= 1. Consequently it would seem that, for most of the data sets that are likely to result from a large screening program, appropriate statistical procedures applied to the gamma multihit model would yield lower confidence limits for VSDs which are even smaller than those calculated from the one-hit model.

APPENDIX D REFERENCES

1. Armitage, P. and R. Doll. Stochastic models for carcinogenesis. In Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, vol. 4 (Berkeley and Los Angeles, Calif.: University of California Press, 1961), pp. 19-38.
2. BIER. The effects on populations of exposures to low levels of ionizing radiation, National Academy of Sciences. National Research Council (Washington, D. C., 1972).
3. Brown, J. M. Linearity vs. nonlinearity of dose response for radiation carcinogenesis. Health Physics 31: 231-245, 1976.
4. Chand, N., and D. G. Heel. A comparison of models for determining safe levels of environmental agents. In Reliability and Biometry, F. Prochan and R. J. Serfling (eds.) (Philadelphia, Pa.: SIAM, 1974), pp. 382-401.
5. Cornfield, Jerome. Carcinogenic risk assessment. Science 198:693-699, 1977.
6. Crump, K. S., D. G. Heel, C. H. Langley, and R. Pete. Fundamental carcinogenic processes and their implication for low dose risk assessment. Cancer Research 36:2973-2979, 1976.
7. Crump, K. S., H. A. Guess, and K. L. Deal. Confidence intervals and tests of hypotheses concerning dose response relations inferred from animal carcinogenicity data. Biometrics 33:437-451. 1977.
8. Crump, K. S. Response to open query: Theoretical problems in the modified Mantel-Bryan procedure. Biometrics 33:752-755, 1977.
9. Daffer, P. Z., K. S. Crump, and M. D. Masterman. Asymptotic theory for analyzing dose response survival data with application to the low-dose extrapolation problem (submitted), 1979.
10. Food Safety Council. Proposed system for food safety assessment. Food and Cosmetic Toxicology (December), 1978.
11. Gass, G. H., D. Coats, and N. Graham. Carcinogenic dose response curve to oral diethylstilbestrol. J. Natl. Cancer Inst. 33:971-977, 1964.
12. Guess, H. A. and K. S. Crump. Low-dose rate extrapolation of data from animal carcinogenicity experiments—analysis of a new statistical technique. Mathematical Biosciences 30:15-36, 1976.
13. Guess, H. A. and K. S. Crump. Maximum likelihood estimation of dose-response functions subject to absolutely monotonic constraints. Annals of Statistics 6:101-111, 1978.
14. Guess, H. A., K. S. Crump, and R. Pete. Uncertainty estimates for low-dose rate extrapolations of animal carcinogenicity data. Cancer Research 37:3475-3483, 1977.
15. Hartley, H. O., and R. L. Sielken. Estimation of "safe doses" in carcinogenic experiments. Biometrics 33:1-30, 1977.
16. Heel, D. G., D. W. Gaylor, R. L. Kirschstein, U. Saffiotti, and M. A. Schneiderman. Estimation of risks of irreversible, delayed toxicity. J. of Toxicology and Environmental Health 1:1 33-151, 1975.
17. Mantel, N. and W. R. Bryan. Safety testing of carcinogenic agents. J. Natl. Cancer Inst. 27:455-470, 1961.
18. Mantel, N., and M. Schneiderman. Estimating "safe" levels, a hazardous undertaking. Cancer Research 35:1379-1386, 1975.
19. Mantel, N., N. R. Bahidar, C. C. Brown, J. L. Ciminera, and J. W. Tukey. An improved Mantel-Bryan procedure for the "safety" testing of carcinogens. Cancer Research 35:865-872, 1975.
20. Mantel, N. Aspects of the Hartley-Sielken approach for setting "safe doses'" of carcinogens. In origins of Human Cancer Book C. Human Risk Assessment, H. H. Hiatt, J. D. Watson, and J. A. Winsten (eds.) (Cold Springs Harbor, N. Y.: Cold Springs Harbor Laboratory, 1977), pp. 1397-1401.
21. McCann, J., E. Choi, E. Yamasaki, and B. N. Ames. Detection of carcinogens as mutagens in the Salmonella/microsome test. Assay of 300 chemicals. Proc. Nat. Acad. Sci. USA 72:5135-5139, 1975.
22. McCann, J., and B. N. Ames. Detection of carcinogens as mutagens in the Salmonella/microsome test. Assay of 300 chemicals: Discussion. Proc. Nat. Acad. Sci. USA 73:950-954, 1976.
23. Pete, R. Paper delivered at the NIEHS Wrightsville Beach conference on Low-Dose Risk Estimation, 1974.
24. Pete, R. The carcinogenic effects of chronic exposure to very low levels of toxic substances. Presented at an NIEHS conference on cancer risk estimation, Pinehurst, N. C., Mar. 10-12, 1976. Environmental Health Perspectives, 1977.
25. Pete, R. Epidemiology, multistage models, and short-term mutagenicity y tests. In Origins of Human Cancer, Book C, Human Risk Assessment, H. H. Hiatt, J. D. Watson, and J. A. Winsten (eds.) (Cold Springs Harbor, N. Y.: Cold Springs Harbor Laboratory, 1977 b), pp. 1397-1401.
26. Rai, K. and J. Van Ryzin. Risk assessment of toxic substances based on a generalized multihit dose response model. Proceedings of SIMS Conference, Alta, Utah, June 26-30, 1978.
27. Salsburg, D. S. Open query: Theoretical problems in the modified Mantel-Bryan procedure. Biometrics 33:4 19-421, 1977.
28. Watson, G. S. Age incidence curves for cancer. Proc. Nat. Acad. Sci. USA 74:1341-1342, 1977.