[CANCER RESEARCH 36, 2973-2979, September 1976]

Fundamental Carcinogenic Processes and Their Implications for Low Dose Risk Assessment

K. S. Crump,[1] D. G. Hoel,[2] C. H. Langley, and R. Peto

National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, North Carolina 27709 [K. S. C., D. G. H., C. H. L.], and University of Oxford, Oxford, England [R. P.]

Summary

Various possible models of carcinogenesis are analyzed with respect to low dose kinetics. The importance of background carcinogenesis upon the shape of the dose-response curve at low dose is emphasized. It is shown that, if carcinogenesis by an external agent acts additively with any already ongoing process, then under almost any model the response will be linear at low dose. Measures of the degree of linearity are obtained for multistage models of carcinogenesis, where it is shown that throughout the dose range where the extra risk is less than the spontaneous risk linear extrapolation must be quite accurate.

Introduction

The presence of carcinogenic agents in the environment is an accepted fact. Although most agents can be avoided once they are identified as carcinogenic, some may be avoided only at great expense or alternative risk, in which case "risk *versus* benefit" must be evaluated. One important aspect of the determination of risk is the estimation from animal experiments conducted at high doses on small to moderate numbers of animals of the risks to such animals of cancer associated with very low levels of exposure. This is likely to be a principal element in risk estimation for the myriad of chemicals that must be evaluated. Relevant human data are usually not available.

The estimation of attributable risk at a dose very much lower (say 1/1000th) than the smallest practical experimental dose involves the interpolation between 2 dose levels: the control and the experimental dose levels. Interpolation necessitates an assumption about the behavior of the risk with increasing dose. The assumption can be specified arbitrarily or it can be deduced from reasonable models of the carcinogenic process. An "estimate" of risk is as arbitrary as the interpolation scheme that produced it. We will attempt in what follows to relate the properties of various risk estimation procedures to several observations and assumptions about carcinogenesis.

Two properties of carcinogenesis are critical to low dose risk estimation.

1. Cancers are believed to be single cell in origin (6, 7). Of a large number of cells at risk in the individual organism, 1 undergoes certain changes that allow it to divide and grow into a tumor. Thus we can view the carcinogenic process as mechanistically single cell in origin even though, by the time a cancer is pathologically recognizable, very extensive changes may have developed.

2. It will be shown that it is important to know whether the causal processes associated with the particular carcinogen of interest are common to those involved in carcinogenesis due to other causes, either "spontaneous" or from other carcinogens. In other words, we need to know whether or not carcinogenesis due to a particular carcinogen is independent of other modes of carcinogenesis.

In the 1st section of this paper, the consequences of the manner of combining "spontaneous" and "induced" carcinogenesis will be explored. We will show that, if the addition of the test carcinogen merely increases the rates of processes that were occurring anyway, then dose-response relationships will be linear at low dose levels. In the 2nd section several models will be considered and their low-dose properties will be identified. We will find that every reasonable model of carcinogenesis is linear or sublinear at low dose. Finally, in the 3rd section we will look more closely at this linearity and determine the accuracy of linear approximations in "multistage" models of cancer.

It should be recognized that there may be agents that indirectly affect the carcinogenic process. An example might be some dietary alteration that led to a modification of gut flora that may change the carcinogenic process in a qualitative way. Although our analysis and conclusions might be appropriate for some of these indirect carcinogenic processes as well, we are chiefly discussing direct carcinogenic processes in which the compound or its metabolite acts at the cellular level to produce an irreversible and heritable (genetic or epigenetic) change.

Significance of the Relation of a Carcinogen to Occurrence of Cancer due to Other Causes

Throughout this paper we shall concentrate upon the case of a population chronically exposed to carcinogens at constant dose rates. We are interested in the individual response when the population is exposed to a particular carcinogen at an approximately constant dose rate d per unit time. This response can be described by the age-specific cancer incidence rate $I(t,d)$ which is the expected rate per unit time at which cancer will be discovered in individ-

---

[1] Present address: Louisiana Tech University, Ruston, La. 71270.

[2] To whom requests for reprints should be addressed, at National Institute of Environmental Health Sciences, P. O. Box 12233, Research Triangle Park, N. C., 27709.

Received September 25, 1975; accepted June 1, 1976.

uals of age t who were previously cancer free. In considering this response, it is important to keep in mind that individuals at risk will ordinarily be exposed to a large number of carcinogens, and we are interested in the effect upon the response of a single one of these which we shall for convenience call the primary carcinogen. As we shall see, the independence or equivalence of the mechanism(s) of action of the primary carcinogen and the other carcinogens can be important in determining the response due to a low dose rate of the primary carcinogen.

We can reasonably divide the totality of carcinogens into 2 groups: Group 1, containing all of those carcinogens that cause a response of cancer in a way that is completely independent of the mechanisms by which the primary carcinogen causes a response; and Group 2, consisting of those carcinogens (including spontaneous biochemical accidents) that somehow act in conjunction with the primary carcinogen in causing cancer. Let I_1 be the incidence rate of new cancers at a fixed time t due to a carcinogen in Group 1 or via an inherent spontaneous phenomenon that is mechanistically related to the effects of the carcinogens of Group 1. Let I_2 be the incidence rate of new cancers at time t due either to a carcinogen in Group 2 or to an inherent spontaneous phenomenon that is mechanistically related to the effects of the carcinogens of Group 2. Then, because of the assumed independence, we can write

$$I(d) = I_1 + I_2 \tag{A}$$

where, as shall be done throughout the paper when convenient, the argument t has been omitted.

Now suppose that Group 2 consists of m carcinogens at dose levels $d^{(1)}, \ldots, d^{(m)}$ in addition to the primary carcinogen at a dose level d. The simplest assumption with regard to the interactive effect of these carcinogens would be to suppose that the effect is additive, *i.e.*, the rate I_2 at which cancer occurs due to a carcinogen in Group 2 is a function of an effective dose rate

$$D = D_0 + \beta d \tag{B}$$

where D_0 is some function of $d^{(1)}, \ldots, d^{(m)}$. Now we can write

$$I_2 = H(D)$$

and we will assume that H is a nondecreasing analytical function. (This merely implies that any increase in the dose rate does not decrease the age-specific cancer incidence rate.) Now we can write

$$\begin{aligned} I(d) &= I_1 + H(D_0 + \beta d) \\ &= I_1 + H(D_0) + \beta H'(D_0)d + o(d) \end{aligned} \tag{C}$$

as the dose rate d approaches zero where $o(d)$ denotes a function with the property $o(d)/d$ approaches zero as d approaches zero. Thus we see that $I(d)$ will be a linear function of the dose rate d at low dose rates provided $H'(D_0) > 0$.

Other authors (11) have allowed for cancers in the models due to causes other than the primary carcinogen of interest by using the formula of Abbott (1) for correcting for response due to extraneous causes. In terms of the above formulation, this is equivalent to supposing that the 2nd group of carcinogens contains only the primary one, or, in other words, the carcinogen of interest acts in some manner completely independent of the mechanism by which all other cancers are formed. If this does, in fact, turn out to be the true biological situation, then the response function $I(d)$ is still represented by Equation C except now $D_0 = 0$. Thus we see that even in this case $I(d)$ can still be linear at low dose provided the slope of the function H is positive at zero. This turns out to be true for some models of carcinogenesis, *e.g.*, 1-hit models (2), but not for others, *e.g.*, some multihit models (15) and the Mantel-Bryan model (11). However, when other carcinogens act in conjunction with the primary one ($D_0 > 0$), the linearity of the response merely depends upon H having a positive slope at the point D_0. This seems intuitively likely and, in fact, is the case in all models of carcinogenesis with which we are familiar.

There are, of course, a number of questions about the biological validity of the assumptions.

1. Do all carcinogens act independently or do certain subgroups act in conjunction with each other? It should not be difficult to answer this question experimentally with regard to the effects of specific carcinogens on specific cancers.

2. Given that at least some carcinogens act in conjunction with the primary carcinogen, is it reasonable to assume that their individual effects are additive in the sense of Equation B? This question is probably much more difficult than the 1st one. However, the idea of complete additivity of effects is not essential to our arguments, and a variety of other assumptions would lead to effectively the same conclusion.

3. Is the assumption that $H'(D_0)$ is positive valid? For example, if there were some type of threshold effect operating so that $H(D) = 0$ for D less than some threshold value D_{th}, then if D_0 were less than D_{th} the argument would break down and $I(d)$ would, in fact, not be linear at the lowest dose rates. On the other hand, if cancer is single cell in origin, then the threshold D_{th} is a property of a single cell rather than of the whole organism. Viewed in this light, it is entirely plausible that, even if a threshold effect does exist for each cell, nevertheless in the entire organism the probability of response may be linear at low dose rates. These cellular thresholds will not all be constant but will be distributed over some range of doses. If this range includes D_0, the organismic response will, generally speaking, be linear at $D_0 + 0$. (The same conclusion would follow from postulating that each person in a large population has a particular threshold but that individual thresholds have a random distribution.)

4. Even if one is willing to accept the fact that the response curve is linear for low dose rates, this in itself may be of little value unless there is some knowledge about "how" linear and "how" low the dose rates must be. To answer such quantitative questions as these, one must make more specific assumptions than are incorporated into the very general discussion presented here. We shall return to this question in the light of some particular models for carcinogenesis.

Particular Models for Carcinogenesis

Experimental evidence (6, 7) indicates that cancers originate from a single cell. The models we shall look at will be based upon this premise. First, consider the time to response of a single cell, where the response might possibly be detection of or death due to a cancer originating with this cell. This time to response can be written as the sum of the period of genetic and/or epigenetic alteration of the cell to a malignant phenotype, plus the growth period from the time at which the cell is completely altered to the time of the observed response. The time to cell alteration is presumably dependent upon the dose rate *d*. This may also be true for the growth time, but we shall assume that the latter effect is negligible. If these 2 times are independent of each other, we can in general write

$$I_c(t,d) = \int_0^t I_a\,(t-u,d)f(u)du \qquad \text{(D)}$$

where $I_a\,(t,d)$ is the incidence rate of the alteration of a single cell, $f(t)$ is the density of cancer growth time, and $I_c(t,d)$ is the observed incidence rate for cancer response. The approximation in Equation D is valid because $I_a(t,d)$ applies to a single cell and will be very small. Now the observed incidence rate is for an entire tissue and, as pointed out by Armitage and Doll (4), insufficient attention has been given in some earlier models to the distinction between cell response and tissue response. If a tissue is composed of *n* cells, then the time to response of a tissue is the minimum of the associated *n* cell response times. If we suppose that these *n* cells all respond in the same manner but independently, then we arrive at the formula

$$I(t,d) = nI_c(t,d) \qquad \text{(E)}$$

1. Multihit and Multistage Models. Suppose $k \geq 1$ different events (hits) must occur in a cell before it is sufficiently altered and suppose the *i*th event occurs at a constant rate λ_i, $i = 1,2,\ldots,k$. Suppose further that cells that have suffered some, but not all, of the relevant events have no selective advantage or disadvantage relative to normal cells. Nordling (15) used the multihit model for the total response time, time to alteration, plus growth time, but here we generalize his approach by using the multihit process to model only the time to cellular alteration. The effect of dose rate is introduced in the manner of Neyman and Scott (14) by taking $\lambda_i = \alpha_i + \beta_i d$ (subject to $\alpha_i \geq$ o and $\beta_i \geq$ o). The incidence rate of the alteration of a single cell is [see Armitage and Doll (*4*)]

$$I_a(t,d) \doteq kt^{k-1}\left\{\prod_{i=1}^{k}(\alpha_i + \beta_i d)\right\} = kt^{k-1}Q_k(d) \qquad \text{(F)}$$

where $Q_k(d)$ is a *k*th degree polynomial in *d* with constant coefficients. For the observed incidence rate of cancer we obtain, using Equations D, E, and F,

$$I(t,d) \doteq Q_k(d)S_k(t) \qquad \text{(G)}$$

where

$$S_k(t) = nk\int_0^t f(t-u)u^{k-1}du \qquad \text{(H)}$$

At low dose rates this response is linear in the dose rate. To see this, we note that it is possible to write

$$I(t,d) = S_k(t)Q_k(d) = S_k(t)\,\{A + Bd + \mathrm{o}(d)\}$$

where

$$A = \prod_{i=1}^{k}\alpha_i \text{ and } B = \sum_{i=1}^{k}\beta_i \prod_{j \neq i}\alpha_j$$

The incidence rate will be linear in dose rate at low dose rates whenever the constant *B* is positive. In order that there be both background carcinogenesis and also some effect of the dose rate *d*, it is necessary that all of the α_i's and at least 1 of the β_i's be positive. However, under these conditions the constant *B* is seen to be positive, and thus the incidence rate is linear at low dose.

If $k = 1$ (1-hit model) then

$$I(d) = (\alpha_1 + \beta_1 d)S_1(t) \qquad \text{(I)}$$

and the incidence rate is exactly linear in dose for all dose rates.

Alternatively, one could consider the time to cellular alteration to be the result of a multistage process. This process has been applied to carcinogenesis by Armitage and Doll (5). As in the multihit process, *k* events must occur in a cell to initiate cancer, these events occurring with fixed rate constants. The only difference between the multistage and the multihit processes is that the *k*-initiating events in a multistage process must occur in some particular time sequence. It can be shown (3) that Expression G still holds provided the right side of this equation is divided by *k*. Consequently, the comments on the linearity of the response for the multihit model hold true for the multistage model also. Our comments would, of course, also hold for intermediate models in which some stages must occur in a fixed sequence while others may occur in various orders.

2. Generalized Multievent Model. More generally, a large number of events related to the initiation of cancer could occur in a cell. However, rather than its being necessary for all of the events to occur to initiate cancer, there could be a (possibly quite large) number of subcollections $A_1,\ldots,A_n$ of these events so that cancer is initiated as soon as all of the events in any 1 of these subcollections occur. Thus, in this model there are many paths through which cancer can be initiated, a single path corresponding to a particular subcollection A_i of events. A path containing *k* events can be called a *k*-hit path. The subcollections need not be disjoint so that 1 particular event could be included in a number of different paths. The general consequences of this model are readily understood without going through all of the details. It makes no difference for our purposes whether or not some of the events must occur in a specified order since adjustments necessary to go from one case to the other are quite like the adjustment necessary for going from a multihit to a multistage model. It is again supposed that the ith event occurs at a constant rate $\alpha_i + \beta_i d$ where the α_i's may be functions of dose rates of other carcinogens but are not functions of d. In the terminology of the previous section, all of the paths in which the rates are not functions of d (that is in which the β_i's are all zero) represent the "1st-group" mechanisms, which are independent of the mechanisms by which the primary carcinogen causes cancer. The incidence rate associated with the union of these paths represents I_1 of Equation A, and the incidence rate associated with the union of those paths in which at least 1 of the rates is a function of d represents I_2. It can be shown that the response will be linear in d for small dose rate d unless all paths depending upon d contain at least 2 events that occur only in the presence of the specific primary carcinogen.

How Linear Is "Linear"?

In this section we attempt to describe quantitatively the range of dose rates for which the linear approximations are valid. To do this one must obviously be somewhat model specific, and we will assume the hit models. Two different approaches to the question of "how linear?" will be considered.

For a k-hit model with an arbitrary distribution for cancer induction time, we found (Equation G) that $I(t,d)$ can be expressed as the product of a function of age t only, and the polynomial $\overset{k}{\underset{i=1}{\pi}} (\alpha_i + \beta_i d)$ in the dose rate d. It is of interest to compare this exact expression with its linear approximation

$$I_l(d) = I(0) + dI'(0) \tag{J}$$

where $I(d)$ is given exactly by Equation G. To do this we shall consider the ratio r of $I(d)$ to the linear approximation $I_l(d)$ when $I(d)$ is a certain prescribed proportional excess p of $I(0)$, the age-specific incidence rate at a zero dose rate. Symbolically, we have

$$\begin{aligned} r &= \frac{I(d_p)}{I_l(d_p)} = \frac{I(0)\,(1+p)}{I(0) + I'(0)d_p} \\ &= \frac{1+p}{1 + \dfrac{I'(0)d_p}{I(0)}} \end{aligned} \tag{K}$$

where d_p satisfies $I(d_p) = (1+p)I(0)$.

It is easily seen that $r \geq 1$ is independent of t. Moreover, it can be shown using the method of Lagrange multipliers that r assumes its largest value when $\alpha_1/\beta_1 = \alpha_2/\beta_2 = \ldots = \alpha_k/\beta_k$. When this condition holds, we find that

$$r = \frac{1+p}{1 + k\{(1+p)^{1/k} - 1\}} \tag{L}$$

This is an appealing result in that r depends only on the number of stages k and the proportion of background p, parameters that are easily interpretable. This expression is increasing in k and approaches

$$\frac{1+p}{1 + \log(1+p)}$$

as $k \to \infty$. Thus, we have the bounds

$$1 \leq r \leq (1+p)/\{1 + \log(1+p)\} \tag{M}$$

This upper bound for r holds for any k, and for any multipath model with different values of k for different paths. The upper bound for the ratio is increasing in p. For a "doubling dose" $p = 1$ we have $r \leq 1.18$ regardless of the value of k, and if $k = 2$ we have $r \leq 1.09$. For a 10% increase in incidence over background, we set $p = 0.1$ and find that $r \leq 1.004$. These results are quite interesting and useful for 2 reasons: (*a*) they indicate the closeness to linearity for a very general class of models; (*b*) the results depend only on the proportion over background. If a finer bound is desired, the number of stages of the carcinogenic process or an upper bound to the number of stages is needed.

Our next approach to the question "how linear?" could be of interest in the following situation. Suppose estimates for $I(d)$ and $I(0)$ are available from experiments where d is a known experimental dose and information is desired about the incidence rate curve at dose rates much lower than d. There are 2 possible problems to consider. First of all, one might ask what dose rate d_a would yield a prescribed incidence rate I_a which may represent "a given acceptable" increase in the incidence rate over the background incidence rate $I(0)$. It should be expected that, if d is so small as to be on the linear portion of the dose-response curve, then d_a can be approximated by fitting a straight line through the points $[0, I(0)]$ and $[d, I(d)]$ and using the dose rate corresponding to I_a on the line. The dose computed using this linearization process is

$$d_L = \frac{I_a - I(0)}{I(d) - I(0)}\, d \tag{N}$$

Since $I_a - I(0)$ will usually be very small, $d_L \doteq d_a I'(0)d/\{I(d) - I(0)\}$, and thus the ratio of the true dose d_a and the approximation d_L is

$$\frac{d_a}{d_L} \doteq \frac{I(d) - I(0)}{I'(0)d} \equiv R \tag{O}$$

On the other hand, one may be interested in estimating the incidence rate I_E which corresponds to a very small environmental dose d_E. If I_E is approximated by I_L, the incidence rate corresponding to d_E on the line joining $[0, I(0)]$ and $[d, I(d)]$, then the ratio of the approximation $I_L - I(0)$ of the increase over background incidence to the true value $I_E - I(0)$, is the same R, since

$$\frac{I_L - I(0)}{I_E - I(0)} \doteq \frac{I(d) - I(0)}{dI'(0)} = R \tag{P}$$

Let us now consider the k-hit model and express R in terms of p, the excess response over background, determined by $1 + p = I(d)/I(o)$. It can be shown that

$$1 \le R \le \frac{p}{k\{(1+p)^{1/k} - 1\}} \qquad \text{(Q)}$$

the upper bound being attained when all β_i/α_i are equal. This upper bound increases steadily from 1 for a 1-hit model ($k = 1$) to $\frac{p}{\log(1+p)}$ as $k \to \infty$. We have tabulated the upper bound for R for different values of p and k in Table 1. We see that the linear approximations are reasonable over a wide range of values of k and p. The only circumstances in which linear approximation might be inappropriate are seen to be those where the background rate is vanishingly small and thus p is very large. An example of this might be the induction of angiosarcomas by vinyl chloride; however, if recent evidence (13) that common tumors are also caused by vinyl chloride is confirmed, p will not be extreme, and linear approximation will be adequate even in this case. The excess p of experimental incidence rate over background incidence $I(o)$ thus plays a key role in the accuracy of the linear approximations. In particular, the linear approximations improve as p decreases towards zero.

DISCUSSION

We have shown under some reasonable assumptions about carcinogenic mechanisms and processes that dose responses will be approximately linear at low doses. Let us examine the evidence in favor of these assumptions and review the generality of models considered.

Single-Cell Origin. If individual cancers arise from an original, single, "transformed" cell, then the statistical nature of the carcinogenic dose response will be governed by the extreme tail of the "transformation" response distribution. The effect of this is to make virtually any process of discrete events approximately linear at low dose.

Two primary observations indicate the single-cell origin of cancers. In women who are heterozygous for electrophoretic variants of X-linked glucose-6-phosphate dehydrogenase, cancers are uniformly of one phenotype or the other (6), whereas a comparable amount of normal tissue is composed of a mixture of cells of the 2 phenotypic classes. Further evidence for the single-cell origin of cancers comes from experimental efforts in which "transformed" cells are transplanted into whole animals. Although there is much controversy associated with various aspects of this line of research, it seems that the ability of a single cell to give rise to a cancer is well demonstrated (7). Thus, 2 lines of evidence indicate that cancer can be most reasonably assumed to arise from events associated with or occurring inside single cells.

Table 1

Values of the upper bound of the ratio R from Equation Q

k \ p	0.1	0.5	1	4	10	100
1	1	1	1	1	1	1
2	1.02	1.11	1.21	1.62	2.16	5.52
5	1.04	1.18	1.35	2.11	3.25	13.18
∞	1.05	1.23	1.44	2.49	4.17	21.67

We note again here that this analysis is appropriate only for those agents that affect cancer incidence through the alteration of single cells in an irreversible and hereditable manner (*e.g.*, chronic exposure to low-level ionizing radiaiton). Those agents that increase cancer by anatomical and/or physiological alteration of whole tissues and organs (*e.g.*, dietary modification of gut flora) may or may not be described by these models. Since we do not know the relative proportion of these 2 types of carcinogens and often do not know into which category a particular agent falls, we must stress the importance of understanding basic carcinogenic mechanisms.

Relationship between "Spontaneous" and Induced Carcinogenesis. As we have shown, the independence or dependence of "spontaneous" and "induced" carcinogenesis is critical to the shape of the low dose-response curve. Two types of evidence indicate that these 2 processes share many common mechanistic steps if they are not identical.

Cancers thought to be induced are generally indistinguishable from "spontaneous" cancers. This obviously does not demonstrate that the cancers arise by a common mechanism, but it is consistent with a common pathway to "induced" and "spontaneous" carcinogenesis.

The view of carcinogenesis as a fundamentally mutational phenomenon, as recently reviewed by Knudson (8, 9), supports the assumption that induced and spontaneous steps are mechanistically identical. That is, experimental induction of cancer is the speeding up or the increasing of the probability of the various steps.

The most important observation relevant to the relationship between "induced" and "spontaneous" is that humans demonstrate a high background incidence of cancer. Whether these are due totally to "induction" by environmental agents or also to some truly spontaneous process is immaterial when considering the effects of a small amount of increased human exposure to a particular carcinogen.

Approximately 1 of 5 Americans develops a cancer, and for any particular environmental carcinogen we are interested in a very small associated increase in risk. This 20% background must surely provide some significant "spontaneous" processes that are shared with carcinogenesis by the carcinogen in question; from a public health standpoint the assumption that "induced" and "spontaneous" are not independent is conservative, as well as being biologically plausible. Small extra doses of a carcinogen will therefore elicit linear increases in risk for virtually any response model.

One practical implication of the fact that different carcinogens share many mechanistic steps is that enhancement of certain carcinogenic processes may have a more readily detectable effect on cancer incidence in animals with high background levels of all other carcinogenic processes. Therefore carcinogenicity tests of various substances should possibly include tests on high-spontaneous-incidence strains or experiments to see whether the test substance enhances the carcinogenic effect of a standard carcinogen.

Induction Time and Dose. In our discussion of stochastic models, we assumed that induction time is variable but independent of dose. This assumption is unfortunately weak in that high doses could well affect induction time. This, then, is an area of research that is in need of further effort. If we consider the low doses at which individual environmental carcinogens are experienced, however, it seems reasonable to us to assume relatively little effect on induction times.

Now let us examine the generality of the result: linear dose response at low dose. Given the uncertainties and complexities of carcinogenesis, it is conceivable that several distinct mechanistic phenomena will eventually be discovered to contribute to the appearance of cancer. Thus, we must have an open mind about our modelling and attempt to present the least model-dependent result that we can. This we have done.

Virtually all models of carcinogenesis that depict the exposure as affecting an already ongoing process will lead to linearity at low dose. We have discussed the validity of this assumption above. This result then implies that, no matter what the biological mechanism we might imagine, if the carcinogen increases some part of the already ongoing process, then we should expect the response to be approximately linear at low dose.

As pointed out above, this assumption of dependence or common mechanism is not trivial. It can make orders of magnitude differences in the estimated risk associated with low dose exposure.

If we conceive of the cell alteration process as a series of discrete single-cellular events that can occur in sequence or randomly in any given cell and that a dose-independent induction period follows, then we should expect dose response over background to be linear. We have required neither that all steps be affected by the carcinogen (only some) nor that these steps be all mechanistically similar in quantity or quality. This general class incorporates most of the reasonable models that have been proposed. The keys to this result are the assumptions of the single-cell origin and the lack of any appreciable dose dependence in the induction period.

A further extension of this group of models allows the incorporation of threshold models into the class of "linear at low dose." We have indicated that, if we conceive of single cells as the biological unit at risk and that the initiation response is a threshold phenomenon, then by assuming that the threshold is randomly distributed in dose we find that the low dose response of the whole tissue over background will be approximately linear. If, rather implausibly, we do suppose that some sort of cellular thresholds exists, then clearly all cells do not have the same threshold since all cells do not all become cancers simultaneously. Here again, we have assumed that the carcinogen acts in conjunction with the "spontaneous" or background effects.

The Mantel-Bryan procedure (11) may be interpreted as a random threshold model (albeit without our assumption that substances equivalent to the suspect carcinogen are already present in the environment), although this interpretation was not made by Mantel and Bryan. This requires that the whole organism or tissue be interpreted as the biological unit with a threshold. However, if the single cell is the unit at risk, it must be tentatively accepted that even the threshold concept of carcinogenesis (if it were appropriate) might not be excluded from the class of "linear at low dose."

We have also attempted to answer the question of how linear is "linear at low dose" for particular models. For the multihit and multistage models, linearity is dependent on background incidence. If the background is within a typically observable range, then the linear model provides a reasonable estimate of the true state of nature, while approaching this estimate from the conservative side.

All these considerations clearly demonstrate the importance of explicit and realistic modelling in the development of low-dose extrapolation schemes. Many may feel that we have not considered certain biological observations or hypotheses in the models presented above. We have tried to embrace as much relevant information about carcinogenesis as possible and to obtain results that were the least model dependent.

The weight of these results for human risk assessment is difficult to judge. It is likely that the error in the acceptable dose associated with simple linear extrapolation will be much less than that associated with the species-to-species extrapolation to man from the laboratory animal data. The BEIR report (16) recommended linear extrapolation on pragmatic grounds. The theoretical conclusions of the present paper are that linear extrapolation to low dose levels is generally valid as a realistic yet slightly conservative procedure.

Practical Implications. Our results may be crudely summarized by the observation that, in environments already containing appreciable amounts of carcinogenic processes, the effects of any slight addition to these processes will be proportional to the amount added. Both control laboratory animals and wild humans already suffer a considerable incidence of cancer; thus the extra incidence caused by a small amount of a new carcinogen will be proportional to the dose rate of that carcinogen. This thought is not particularly remarkable, but its implications are that much previous investigation of the form of the dose-response relationship at infinitesimal doses is irrelevant to the interpretation of animal studies for the formulation of social policy.

Unfortunately, the implications of linear extrapolation are bleak. Mantel *et al.* (10–12) have proposed that safe doses be defined on the basis of "probit" extrapolation from upper confidence limits defined by the experimental results, arguing that such a procedure would reward good experimental investigations (by allowing industry bigger permitted doses) while enabling regulators to guarantee to the public that permitted doses were so small that they would cause cancer in less than 1 person in 10^8. Probit extrapolation may be scientifically valid for a few very special "indirect" carcinogenic processes, but our arguments suggest that in general it is not correct. The social implications of our results are best understood by considering, as an extreme case, the use of linear extrapolation to define a "safe" level after doing a large experiment in which no carcinogenic effects were observed. The most definite such negative experiment that is practical might compare animals fed with the order of 10% of the test substance in their diet with a control group and might conclude that the extra risk of cancer was less than something like 1%. Linear extrapolation, even from such ideal results as these, implies that a dose level below 100 ppb is needed for the risk to be less than 10^{-8}, and in real experiments dose levels below 10 or even 1 ppb are likely to be indicated by linear extrapolation in order to guarantee a risk below 10^{-8}. Our arguments that linear extrapolation is generally appropriate at least suffice to demonstrate that linear extrapolation may be appropriate; thus dietary concentrations as low as a few ppb or less will always be needed to guarantee a risk below 10^{-8}.[3]

These levels are of little practical use for the regulatory control of deliberate food additives, although they might occasionally be of practical value for the legislative control of certain special contaminants. Our investigations have thus led us, reluctantly, away from Mantel's hopes and back to the familiar ground where the discovery of an apparently relevant carcinogenic effect of a putative food additive in animals is likely to cause its rapid withdrawal rather than merely a slight reduction of the permitted dose level, unless there are compelling reasons otherwise. This does encourage any carcinogenicity testing sponsored by the industrial users of an additive to be as bad as possible. Moreover, it also means that essentially absolute human safety (*e.g.*, a risk of 10^{-8}) in general cannot be guaranteed by extrapolation from the results of animal carcinogenicity tests. This makes rational public action more difficult than it would be if we lived in a universe where probit extrapolation was true, but it is not a step towards rationality to adopt probit extrapolation because of this. In our opinion, linear dose-response relationships are likely to be approximately correct for many environmental carcinogens, and this should be publicly agreed for such substances, as it was for radiation 20 years ago.

REFERENCES

1. Abbott, W. S. A Method of Computing the Effectiveness of an Insecticide. J. Econ. Ent., *18:* 265-267, 1952.
2. Arley, N., and Iversen, N. On the Mechanism of Experimental Carcinogenesis. Acta Pathol. Microbiol. Scand., *31:* 164-171, 1952.
3. Armitage, P. A Note on the Time-homogenous Birth Process. J. Roy. Statist. Soc. Ser. B, *15:* 90-91, 1953.
4. Armitage, P., and Doll, R. Stochastic Models for Carcinogenesis. *In:* Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Vol. 4, pp. 19-38. Berkeley and Los Angeles: University of California Press, 1961.
5. Armitage, P., and Doll, R. The Age Distribution of Cancer and a Multistage Theory of Carcinogenesis. Brit. J. Cancer, *18:* 1-12, 1954.
6. Fiaklow, P. J. The Origin and Development of Human Tumors Studied with Cell Markers. New Engl. J. Med., *219:* 26-35, 1974.
7. Gartler, S. M. Utilization of Mosaic Systems in the Study of the Origin and Progression of Tumors. *In:* J. German (ed.), Chromosomes and Cancer, pp. 313-334, New York, Wiley/Interscience, 1974.
8. Knudson, A. G. Mutation and Human Cancer. Advan. Cancer Res., *17:* 317-352, 1973.
9. Knudson, A. G. Heredity and Cancer. Am. J. Pathol., *77:* 77-84, 1974.
10. Mantel, N., Bohidar, N., Brown, C., Ciminera, J., and Tukey, J. An Improved "Mantel-Bryan" Procedure for "Safety" Testing of Carcinogens. Cancer Res., *35:* 865-872, 1975.
11. Mantel, N., and Bryan, W. R. Safety Testing of Carcinogenic Agents. J. Natl. Cancer Inst., *27:* 455-470, 1961.
12. Mantel, N., and Schneiderman, M. Estimating "Safe" Levels, a Hazardous Undertaking. Cancer Res., *35:* 1379-1386, 1975.
13. Monson, R. R., Peters, J. M., and Johnson, M. N. Proportional Mortality among Vinyl Chloride Workers. Environmental Health Perspectives, *11:* 75-77, 1975.
14. Neyman, J., and Scott, E. G. Statistical Aspect of the Problem of Carcinogenesis. *In:* Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Vol. 4, pp. 745-776. Berkeley and Los Angeles: University of California Press, 1965.
15. Nordling, C. O. A New Theory on the Cancer-inducing Mechanism. Brit. J. Cancer, *7:* 68-72, 1953.
16. The Effect on Populations of Exposure to Low Levels of Ionizing Radiation (BEIR Report). *In:* Report of the Advisory Committee of the Biological Effects of Ionizing Radiations, National Academy of Sciences, National Research Council. Publication No. 0-489-797, p. 217. Washington, D. C., U. S. Government Printing Office, 1972.

---

[3] There is no compelling reason why the guarantee should be of an extra risk of 10^{-8} or less; if, for example, we merely wanted to guarantee an extra risk of less than 10^{-5}, a few ppm might be permitted, and, since human cancer death rates in middle age exceed 10^{-3} *per annum* anyway, a limit substantially greater than 10^{-8} might be considered for substances that are particularly valuable.