Multistage Models of Carcinogenesis

by P. Armitage*

The simple multistage model of carcinogenesis is outlined. It provides a satisfactory explanation of the power law for the age incidence of many forms of epithelial carcinoma, for the effects in human populations of changing exposures to supposed carcinogenic agents, and for many of the observed effects of applied carcinogens in animal experiments. In particular, the evidence on the effects of starting and stopping cigarette smoking suggests that both an early and a late stage may be affected. In the absence of direct evidence on the nature of the cellular changes there is some reluctance to accept a model with more than two stages, and several forms of two-stage models provide good general explanations of observed phenomena. Such a model has recently been applied to breast cancer; another approach to this disease, effectively involving transformations of the time scale, is discussed.

Introduction

Multistage and related models of carcinogenesis have been discussed for about 30 years, and the growth in the literature has been almost as rapid as the rise of cancer incidence with age. In a short paper I cannot attempt a comprehensive review, and I shall aim to outline the topic in a general way, making more specific comments about some of the points which happen to have interested me over this period. More comprehensive reviews of the mathematical theory have been given by Armitage and Doll (1), Whittemore (2), Whittemore and Keller (3), and Peto (4) has provided a stimulating general review.

Early Work

The flurry of work in the early 1950s, which led to the formulation of a number of related models, was probably motivated by evidence from various sources. First, there was the epidemiological evidence that the mortality or incidence rates for many forms of human cancer increased rapidly with age. This might be a general effect of aging, the body becoming more susceptible to insults of various sorts, or it might be because carcinogenesis is a complex process requiring time and perhaps involving several qualitatively different stages. There were two considerations favoring the second explanation: the fact that people exposed to a high but short-lived carcinogenic risk (for example from irradiation or industrial hazards) often acquire cancer after a long period of time; and animal experiments such as those of Berenblum and Shubik (5) showed that some chemicals are especially effective either early or late in the induction process (the present terms for these being

"initiators" and "promoters"), suggesting that qualitatively different processes were at work during the early and late phases. I shall discuss later some more recent work on the distinction between age *per se* and duration of exposure to carcinogens.

In the reviews mentioned earlier, fuller descriptions of some of the early models are given than can be presented here. They include the "multicell" theory of Fisher and Holloman (6) (requiring a mutationlike change to a specific number of neighboring cells in a tissue, and inconsistent with the unicellular nature of most tumors); and the "multistage" or "multihit" models of Stocks (7) and Nordling (8) (in which a specific number of changes in any order are required). The very similar model of Armitage and Doll (9) introduced the idea of a specific ordering of the changes, so as to accommodate the evidence from initiation–promotion experiments and also a number of features of the epidemiology of human cancer. There was also a series of papers by Iverson and Arley, starting with one (10) which postulated a randomly occurring initiating event followed by a randomly distributed induction period. This rather general formulation encompasses most of the other models, since in a multistage model the first stage can be taken as the initiating event while all subsequent events are subsumed into the induction period (1).

Derivation of the Basic Model

It will be useful to outline the theory of the Armitage-Doll model in slightly different terms to those of the original paper.

Suppose that, in a particular tissue, there are N cells (or cell lines, if they divide) that can potentially experience carcinogenic transformation. The final development of cancer is the k-th and last of a series of sudden and irreversible changes (or stages) which must take place in a specific order. The clinical detection of the

*Department of Biomathematics, University of Oxford, 5 South Parks Rd., Oxford OX1 3UB, England.

disease may be delayed by the period required for the tumor to grow to a detectable size: we shall assume that this is a relatively short lag and shall not consider it in any detail. (Many writers systematically replace the current time t by $t-w$, where w is the assumed lag.)

Suppose that, for any cell which has experienced $i-1$ changes [which we shall call an $(i-1)$-cell], the event rate for next change is λ_i, independent of time. That is the probability that the i-th change takes place in $(t, t+dt)$ is $\lambda_i dt + o(dt)$. This defines a time-homogeneous birth process. We should like to know $f(t)$, the event rate for the k-th change at time t (the process starting at time 0).

General and particular solutions for this problem are well-known (11–14) but are algebraically cumbersome. Fortunately, an approximation is adequate for almost all purposes. Consider the position for values of t small enough to make the probabilities of any of the changes in $(0,t)$, in any one cell, very small. We can either take the limit of the general expression as $t \to 0$ (13,14), or use a straightforward argument (9) to show that

$$f(t) \sim \frac{\lambda_1\lambda_2 \ldots \lambda_k}{(k-1)!} t^{k-1} \qquad (1)$$

The cumulative probability, $F(t)$, that the k-th change has taken place by time t is

$$F(t) = \int_0^t f(u)du \sim \frac{\lambda_1\lambda_2 \ldots \lambda_k}{k!} t^k \qquad (2)$$

Clearly, Eq. (2) cannot hold indefinitely as t increases. However, in most cases lifetime values of t will still be sufficiently small for the limiting assumption (which concerns single cells) to be adequate.

For the particular tissue with N cells, the probability that cancer (i.e., the k-th change) has not appeared by time t is

$$1 - G(t) = \{1 - F(t)\}^N$$
$$\sim \left\{1 - \frac{\lambda_1\lambda_2 \ldots \lambda_k}{k!} t^k\right\}^N$$
$$\sim \exp\left\{-\frac{\lambda_1\lambda_2 \ldots \lambda_k N}{k!} t^k\right\}$$
$$= \exp\left\{-\frac{\alpha_t{}^k}{k}\right\} \qquad (3)$$

Thus, the distribution function $G(t)$ for the time to appearance of the first cancer is a Weibull distribution, with a density function

$$g(t) = G'(t) = \alpha t^{k-1} \exp\left\{-(\alpha/k)t^k\right\}$$

and hazard function

$$h(t) = g(t)/\{1 - G(t)\} = \alpha t^{k-1} \qquad (4)$$

We have here the familiar power law. The limiting approximation on which it depends seems reasonably secure, since it assumes small rates per cell, but Moolgavkar (15) has pointed out that, for some values of the parameters which are plausible for human cancer, it may appreciably overestimate the hazard to be expected at high ages.

Human Cancer

The age-specific mortality rates for cancers at a particular site, or more directly the age-specific incidence rates obtained from cancer registries, can be regarded as roughly analogous to the hazard functions described mathematically by Eq. (4), since the denominators of the rates are the numbers of people alive at the ages in question. From Eq. (4),

$$\log h(t) = \log \alpha + (k-1) \log t \qquad (5)$$

and this linear log-log relation has been widely observed for a wide range of sites and human populations (16,17). It seems to be the usual finding in most epithelial carcinomas, but a variety of quite different relationships is seen for many nonepithelial tumours and for epithelial tumours at sex-specific sites (4). The slope in Eq. (5) is commonly in the range four to six, suggesting there may be around five to seven discrete stages.

However, there are several reasons for caution. In the first place, several other diseases show rapidly increasing age-incidence curves, and one would not seek to explain them all by models of this sort. Secondly, a power law, with a slope of $k-1$, or something very close to it, could be obtained with fewer than k stages. Suppose some of the stages had rates increasing as powers of the time elapsing since the previous stage. Then the slope $k-1$ would be the sum of the (power + 1) for all stages before the last plus the power for the last; for instance, $k-1 = 4$ would arise from five constant rates, or two linearly increasing rates followed by a constant rate, or a quadratic rate followed by a linear rate. A reductio ad absurdum is to postulate one state with $\lambda_1 \propto t^{k-1}$; the model then becomes purely tautological. Two-stage models are discussed below.

Third, a similar effect (of a high slope with a small number of stages) will be obtained if one or more of the event rates increases with age (rather than with time since last event).

Fourth, the Weibull hazard, Eq. (4), can be obtained more generally, on the argument (18) that the time to first tumor in a tissue is the minimum of N random variables (the time to tumor in the N cells), and that Eq. (4) is a standard limit of the distribution of minima in large samples. However, for this limiting form to be valid there are restrictions on the shape of the extreme left-hand tails of the distributions of the cell-specific times, namely, that they are power functions like Eq. (1), and this might be taken to provide at least weak

support for the multistage theory.

Confidence in a multistage model must clearly depend on wider considerations than the power law. In particular, we need to consider the effects of external carcinogenic agents, data from animal experiments, and biological plausibility. These and other topics are taken up in later sections of the paper.

As already noted, the cancers of sex-specific organs tend not to follow the power law. This is understandable since many of these organs are subject to changes in their hormone dependence at various periods throughout life or, like the uterine cervix, are affected by changes in sexual habits. Some tentative explanations of age-incidence can often be given in qualitative terms (9). Some recent quantitative modeling for breast cancer, in terms of a two-stage model is discussed in a later section.

Animal Experiments and the Effects of Applied Carcinogens

In experiments in which animals receive continuous application of a carcinogenic agent, the time to first tumor commonly follows a distribution close to the Weinbull (19,20). Such experiments not only provide a measure of support for the general theory, but also enable one to study the dose–response relationship. In the simple multistage model, suppose that m of the k stages are affected by the carcinogen, so that, for these values of i, $\lambda_i = d\lambda_{oi}$, where d is dose intensity. Then, from Eq. (4) and the definition of α in Eq. (3), the hazard function should be proportional to d^m. It is common to find $m < k$, suggesting that some but not all of the stages are affected by a particular carcinogen.

Carcinogenic agents may, of course, not be applied at constant rates, and the question arises how the hazard function $h(t)$ is affected if a particular rate constant, say λ_i, is an arbitrary function of time $\lambda_i(t)$, which in the simplest case might be proportional to the dose intensity of $d(t)$ of an applied carcinogen. The answer (9) is that $h(t)$ is proportional to a weighted mean of $\lambda_i(t)$ in $(0,t)$, the weight at time τ $(0 < \tau < t)$ being proportional to $\tau^{i-1}(t - \tau)^{k-i-1}$. This means that, for small values of i (early stages affected), what matters is the value of $\lambda_i(\tau)$ at low τ, whereas for high values of i (say k or $k - 1$) the more recent values of $\lambda_i(\tau)$ carry most weight. These effects are explored more fully by Whittemore and Keller (3) and by Day and Brown (21).

In this context one could broadly explain an initiator–promoter experiment by saying that the initiator affects primarily the first step and the promoter primarily a later step (perhaps the second of two). However, Stenbäck et al. (22–24) have shown that the interpretation of these experiments may be complicated by aging and other effects. Earlier, Peto et al. (25) had carried out some experiments with regular benzpyrene applications to mice, which showed that under these circumstances the incidence of tumors depended on the time since start of exposure and not on age. This result is consistent

with the view that the first stage is affected and that its enhanced event rate in the presence of benzpyrene is much greater than the natural background rate. Thus, whether or not it also affects some late stage(s), benzpyrene appears at least to "initiate" the first stage. In contrast with the age-independent effect of an initiator, however, Stenbäck et al., in experiments similar to (but much longer than) those of Berenblum and Shubik (5), found that the "promoting" effect of a TPA declined with age, suggesting a systemic aging effect in the response to TPA. Finally, to illustrate that the opposite effect is possible, Gray et al. (26) in experiments on radon inhalation by rats, found the incidence at a fixed time after start of exposure to increase with age. This is what might be expected if radon affected the second or a later stage, since with increasing age at start of exposure there would be more cells that had already undergone one or more of the early stages spontaneously. We return to the question of age effects in the next section.

Human Data and Exposure to Carcinogens

The considerations outlined in the first two paragraphs of the section titled "Animal experiments and the effects of applied carcinogens" would be expected to apply to human exposures as well as to animal experiments. One of the most illuminating examples is provided by the effects of starting and stopping smoking at different ages (4, 27–29).

Data from prospective studies, such as the British and American data analyzed by Doll (16), show that nonsmokers have a log-log relationship for lung cancer with a slope $k - 1$ of about four. For cigarette smokers, the same slope is obtained if time is measured not from birth but from the start of smoking. This is reasonable if smoking enhances one or more of the k_i to such high levels that the naturally occurring changes are very much less frequent than those induced by smoking.

Consider now the effect of stopping smoking. Smokers who stop retain their high rates, but at a constant level, perhaps until the nonsmokers' rates rise to that level. This is precisely what would be expected if smoking affected the $(k - 1)$th of k stages, for there would be a pool of ex-smokers with $(k - 1)$-cells, waiting for the final change which would occur at a constant rate. In due course, the pool will be augmented by naturally occurring $(k - 1)$-cells and the rate will start to rise.

Consider, secondly, the incidence rate at a fixed time after start of smoking, as a function of age at starting. The data are sparse but seem to indicate either little effect of age at starting or at most a rather modest positive effect. This would be consistent with an early stage being affected; (if the first stage were affected so that λ_1, were increased dramatically by smoking, the process would effectively start at that point, but if the second stage were affected the number of 1-cells available for further transformation would increase approx-

imately linearly with age at starting). Moreover, general considerations about the delay in the effect on a *population of a marked increase in smoking* suggest that an early stage is affected.

Thus, different arguments support effects on both early and late stages. Both effects could, of course, be present. Some skin-painting experiments with benzpyrene on mice (*19*) have suggested an incidence proportional to (dose)^2, in turn suggesting that two stages are affected by benzpyrene or that there is one stage with a quadratic effect. A preliminary analysis by Whittemore and Altshuler (*30*) of the study on British doctors (*31*) suggested that the incidence rate was proportional to the number of cigarettes, which provisionally implied that one state was affected proportionally to dose, i.e., that $m = 1$. However, an analysis (*32*) of a "reliable" subset of the doctors' data suggests a response more than proportional to dose; the estimate of m may be reduced by errors of measurement of smoking habits; and the effect of smoking on a particular λ_i may be less than proportional to the daily consumption of cigarettes. The evidence thus points, somewhat loosely, toward the involvement of two stages.

A useful discussion of the effects of removal of carcinogenic exposure in a range of human cancers, as well as in animal experiments, is given by Day and Brown (*21*).

The concept that a carcinogen may affect only some of the rate constants helps us to understand some of the observed interactions between different carcinogens. In some instances, as in the interaction between smoking and asbestos exposure (*33*), the effects of the two agents appear to be multiplicative. This would be expected if they acted, with proportionate effects on the rate constants, for two different stages, say the *i*-th and *j*-th, since the hazard function, Eq. (4), involves the product $\lambda_i\lambda_j$. On the other hand, if both agents affect the same rate constant λ_i, their effects could well be additive.

Two-Stage Models: Breast Cancer

In the absence of direct biological evidence about a *succession of stages, models with several* (five to seven) stages have often been regarded as implausible. A two-stage model with exponential proliferation of the 1-cells has been discussed (*34*). The exponential growth in the rate constant for the second stage has much the same effect as a low-order polynomial and it is not surprising that the two-stage model with proliferation mimics *fairly closely the multistage model with constant rates.* Other two-stage models are detailed elsewhere (*35–37*).

Moolgavkar and Venzon have studied a generalization of the model (*39*) permitting growth also of the O-cells and have been able to fit data for a wide range of human cancers. The model has been adapted for breast cancer by Moolgavkar, Day, and Stevens (*40*) who postulate growth in the rate constant for the first initiation during puberty (with menarche following a logistic curve), subsequent proliferation of 1-cells with an enhanced rate during pregnancy, a reduced rate after menopause, and

a protective effect of first birth by a subsequent reduction in 1-cell proliferation. They provide extremely impressive fits to data.

Pike and his colleagues (*41,42*) have obtained equally impressive fits with a model conceptually different from, but very similar in its consequences, to that of Moolgavkar et al. Pike et al. adopt a power law with an index $k - 1$ of 4.5, but assume that "time" (as used in the formula) is effectively expanded or contracted during a woman's life. Exposure starts at menarche (for which again a logistic curve is assumed), "time" moves more rapidly during reproductive life, with a temporary spurt *during pregnancy, a fall after the first birth and a further fall after menopause.* These authors suggest that the constancy of breast cancer rates in postmenopausal Japanese women (in contrast to the rise in other population groups) may be an effect of their low weights and low estrogen levels. The changes in the rate of passage of "time"are equivalent, in the simple multistage model, *to the multiplication of all the rate constants* by some factor varying throughout a woman's life, and may be motivated by the view that the rate constants depend on the rate of metabolism of stem cells, which may vary in the way indicated. It would, of course, be a rather strong assumption that all the rate constants should remain in the same ratios to each other although *varying greatly with time.*

Low-Dose Extrapolation

Considerable interest has been expressed in recent years in the assessment of low-dose carcinogenicity on the basis of extrapolation toward zero dose from the *results of animal experiments in which high doses of* test substances are used (*43*). Setting aside the important questions of the extrapolation from laboratory animal to man, there are serious problems about downward extrapolation within one animal species. The results depend heavily on the assumed nature of the dose–response curve at very low doses (*44–47*).

One plausible and helpful assumption is, however, suggested by many multistage models. Suppose, as before, that *m* stages are affected by the carcinogen. Since *we are dealing with very low doses it will be inappro-* priate to assume the λ_i to be proportional to the dose-intensity *d*, because there may well be background effects, but a linear relation seems reasonable. At fixed *t*, therefore, from Eq. (3), the cumulative incidence at dose *d* will be

$$P(d) = 1 - \exp\left\{-\prod_{i=1}^{m} (\alpha_i + \beta_i d)\right\} \qquad (6)$$

where the parameters α_i and β_i absorb the constants and terms involving *t* in Eq. (3). A slight generalization of Eq. (6) is to write

$$P(d) = 1 - \exp\left\{-\sum_{i=0}^{m} \theta_i d^i\right\} \qquad (7)$$

where all the θ_i are nonnegative, and this model has been studied in detail (48–51). In Eq. (7), if $\theta_1 > 0$, the response curve is essentially linear at low doses. This restriction will give more "conservative" assessments (i.e., a given excess risk will be reached at lower doses) than most or all other models proposed. Now, a maximum likelihood estimate (50) of θ_1 may be 0, in which case a steeper curve will be fitted at low doses; however, some nonzero value of θ_1 will always be consistent with the data, and so linear extrapolation can scarcely be excluded as a reasonable procedure.

Other models have been proposed. Hartley and Sielken (52,53) generalize Eq. (7) to include time. Cornfield and his associates (54–56) have studied a multihit model (i.e., one involving hits occurring in an arbitrary order), which, with certain assumptions about background effects, has similar consequences to Eq. (6). However, Van Ryzin (57) points out that the low-dose linearity of Eq. (6) depends essentially on the assumption that the λ_i are asymptotically linear in d. This linearity would follow if the background incidence were due to a carcinogenic agent, the dose of which combined additively with the applied dose. This need not be true.

Biological Evidence and Conclusions

In the construction of mathematical models for biological phenomena it is not uncommon to find that theories of quite disparate types provide good fits to the same data. Discrimination between models must then depend partly on general biological plausibility and partly on the ability of the models to explain new data.

This is essentially the position with our present topic. As a statistician I can offer no authoritative guide to biological mechanisms. There seems little doubt, though, that the multistage theory, in some form or another, has provided a useful framework for hypothesis formation and for the design of observational and experimental studies. A number of experimental biologists maintain that carcinogenesis is a multistage process (the term 'multistep' is often used) (58,59), with perhaps an initial mutationlike stage of initiation being followed by one or more steps of a different nature (such as the activation of an oncogene). Evans and DiPaolo (60,61) have identified a number of specific stages in the progression of guinea pig fetal cells to neoplasia, such as morphological transformation, anchorage-independent growth, colony forming in agar, etc.

Until and unless we obtain direct evidence about the presence and nature of intermediate stages, any statistical theory is likely to remain largely unfalsifiable, particularly if it is allowed to be modified with the flexibility to which we have become accustomed.

The main contenders for generally applicable theories seem to be (a) the multistage theory, (b) some form of two-stage theory, and (c) the time-transformation theory of Pike and his colleagues. Until we have clear evidence for more than two states, it seems best to regard the multistage theory, like the dogmas of certain religions, as permitting either a literal or a figurative interpretation. That is, one can either assume that there really are $k > 2$ separate stages or one can regard some of the intermediate stages as being fictional shorthand for a single proliferative stage. There does seem a need to preserve at least two stages, so that we can distinguish between "early" and "late" effects of carcinogens.

The explicit two-stage models, with appropriate assumptions about proliferation, seem to explain many of the known facts. However, the observations on the effects of starting and stopping smoking, described above, suggest that at least three stages are involved for lung cancer (two affected by smoking and a final stage). Moreover, the multiplicativity of the effects of asbestos and radiation with that of smoking suggests at least a third stage. In any proliferative model, the precise nature of the proliferative parts of the process is likely to remain indeterminate until and unless direct biological observations become available.

The time-transformation model is relatively new and its full consequences have not, as far as I know, been explored. In one sense, it avoids some of the assumptions of the other models, in that the power law can be invoked as an empirical observation without any reference to stages. On the other hand, the particular way in which the response function is modified by changing circumstances (which we have seen would be equivalent to changing the rate for each of a number of stages by the same multiple) seems more specific than is required by other models, and it is unclear whether the model provides a suitable explanation for initiator–promoter data or other data in which an early or a late effect in indicated.

In many areas of biomathematics the ingenuity of the mathematician often seems to run ahead of the ability of the biological scientist to provide the data needed to validate the mathematical models. In the study of carcinogenesis it is encouraging to see, to an increasing extent, the close cooperation between mathematicians and statisticians on the one hand, and biologists on the other, and I believe that in this sort of collaboration lies the key to the solution of some of the problems I have discussed in this paper.

REFERENCES

1. Armitage, P., and Doll, R. Stochastic models for carcinogenesis. In: Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Vol. IV (J. Neyman, Ed.), University of California Press, Berkeley and Los Angeles, 1961, pp. 19–38.
2. Whittemore, A. Quantitative theories of carcinogenesis. Adv. Cancer Res. 27: 55–58 (1978).
3. Whittemore, A., and Keller, J. B. Quantitative theories of carcinogenesis. SIAM Rev. 20: 1–30 (1978).
4. Peto, R. Epidemiology, multistage models and short-term mutigenicity tests. In: Origins of Human Cancer, Book C: Human Risk Assessment (H. H. Hiatt, J. D. Watson and J. A. Winston, Eds.), Cold Spring Harbor Laboratory, Cold Spring Harbor,

1977, pp. 1403–1428.

5. Berenblum, I., and Shubik, P. An experimental study of the initiating stage of carcinogenesis, and a re-examination of the somatic cell mutation theory of cancer. Brit. J. Cancer 3: 109–118 (1949).

6. Fisher, J. C., and Hollomon, J. H. A hypothesis for the origin of cancer foci. Cancer 4: 916–918 (1951).

7. Stocks, P. A study of the age curve for cancer of the stomach in connection with a theory of the cancer producing mechanism. Brit. J. Cancer 7: 407–417 (1953).

8. Nordling, C. O. A new theory on the cancer-inducing mechanism. Brit. J. Cancer 7: 68–72 (1953).

9. Armitage, P., and Doll, R. The age distribution of cancer and a multi-stage theory of carcinogenesis. Brit. J. Cancer 8: 1–12 (1954).

10. Iversen, S., and Arley, N. On the mechanism of experimental carcinogenesis. Acta Pathol. Microbiol. Scand. 27: 773–803 (1950).

11. McKendrick, A. G. Studies on the theory of continuous probabilities with special reference to its bearing on natural phenomena of a progressive nature. Proc. London Math. Soc. 13: 401–416 (1914).

12. Lundberg, O. On Random Processes and their Application to Sickness and Accident Statistics. Almqvist & Wicksell, Uppsala, 1940.

13. Armitage, P. A note on the time-homogeneous birth process. J. Roy. Statist. Soc. B15: 90–91 (1953).

14. Moolgavkar, S. H. The multistage theory of carcinogenesis and the age distribution of cancer in man. J. Natl. Cancer Inst. 61: 49–52 (1978).

15. Moolgavkar, S. H. The multistage theory of carcinogenesis (Letter to the Editor). Int. J. Cancer 19: 730 (1977).

16. Doll, R. The age distribution of cancer: implications for models of carcinogenesis (with discussion). J. Roy. Statist. Soc. A134: 133–166 (1971).

17. Cook, P., Doll, R., and Fellingham, S. A. A mathematical model for the age distribution of cancer in man. Int. J. Cancer 4: 93–112 (1969).

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

19. Lee, P. N., and O'Neill, J. A. The effect both of time and dose applied on tumour incidence rates in benzpyrene skin painting experiments. Brit. J. Cancer 25: 759–770 (1971).

20. Peto, R., and Lee, P. N. Weibull distributions for continuous-carcinogenesis experiments. Biometrics 29: 457–470 (1973).

21. Day, N. E., and Brown, C. C. Multistage models and primary prevention of cancer. J. Natl. Cancer Inst. 64: 977–989 (1980).

22. Stenbäck, F., Peto, R., and Shubik, P. Initiation and promotion at different ages and doses in 2200 mice. I. Methods, and the apparent persistence of initiated cells. Brit. J. Cancer 44: 1–14 (1981).

23. Stenbäck, F., Peto, R., and Shubik, P. Initiation and promotion at different ages and doses in 2200 mice. II. Decrease in promotion by TPA with aging. Brit. J. Cancer 44: 15–21 (1981).

24. Stenbäck, F., Peto, R., and Shubik, P. Initiation and promotion at different ages and doses in 2200 mice. III. Linear extrapolation from high doses may underestimate low-dose tumour risks. Brit. J. Cancer 44: 24–33 (1981).

25. Peto, R., Roe, F. J. C., Lee, P. N., Levy, L., and Clack, J. Cancer and ageing in mice and men. Brit. J. Cancer 32, 411–426 (1975).

26. Gray, R. G., Lafuma, J., Parish, S. E., and Peto, R. Lung tumours and inhalation in over 2000 rats: approximate linearity across a wide range of doses, and potentiation by age and by tobacco smoke. In: Proceedings of the Twenty-Second Hanford Life Sciences Symposium: Life-Span Radiation Effects Studies in Animals: What can they tell us? (R. C. Thompson and J. A. Mahaffey, Eds.), Technical Information Services, U.S. Department of Energy (1986).

27. Armitage, P. Discussion on Paper by R. Doll (*16*). J. Roy. Statist. Soc, A134: 155–156 (1971).

28. Doll, R. An epidemiological perspective on the biology of cancer. Cancer Res. 38: 3573–3583 (1978).

29. Doll, R., and Peto, R. The Causes of Cancer. Oxford University Press, Oxford, 1981, Appendix E, pp. 1292–1305.

30. Whittemore, A., and Altshuler, B. Lung cancer incidence in cigarette smokers: further analysis of Doll and Hill's data for British physicians. Biometrics 32: 805–816 (1976).

31. Doll, R., and Peto, R. Mortality in relation to smoking: 20 years' observations on male British doctors. Brit. Med. J. 2: 1525–1536 (1976).

32. Doll, R., and Peto, R. Cigarette smoking and bronchial carcinoma: dose and time relationships among regular smokers and life-long non-smokers. J. Epidemiol. Commun. Health 32: 303–313 (1978).

33. Selikoff, I. J., and Hammond, E. C. Multiple risk factors in environmental cancer. In: Persons at High Risk of Cancer (J. F. Fraumeni, Jr., Ed.), Academic Press, New York, 1975, pp. 467–483.

34. Armitage, P., and Doll, R. A two-stage theory of carcinogenesis in relation to the age distribution of human cancer. Brit. J. Cancer 11: 161–169 (1957).

35. Kendall, D. G. Birth-and-death processes, and the theory of carcinogenesis. Biometrika 47: 13–21 (1960).

36. Neyman, J., and Scott, E. L. Statistical aspects of the problem of carcinogenesis. In: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Vol IV (L. L. LeCam and J. Neyman, Eds.), University California Press, Berkeley and Los Angeles, 1967, pp. 745–776.

37. Knudson, A. G. Mutation and cancer: statistical study of retinoblastoma. Proc. Natl. Acad. Sci. (U.S.) 68: 820–823 (1971).

38. Knudson, A. G., Hethcote, H. W., and Brown, B. M. Mutation and childhood cancer: a probabilistic model for the incidence of retinoblastoma. Proc. Natl. Acad. Sci. (U.S.) 72: 5116–5120 (1975).

39. Moolgavkar, S. H., and Venzon, D. J. Two-event models for carcinogenesis. Incidence curves for childhood and adult tumors. Math. Biosci. 47: 55–57 (1979).

40. Moolgavkar, S. H., Day, N. E., and Stevens, R. G. Two-stage model for carcinogenesis: epidemiology of breast cancer in females. J. Natl. Cancer Inst. 65: 559–569 (1980).

41. Pike, M. C., Henderson, B. E., and Casagrande, J. T. The epidemiology of breast cancer as it relates to menarche, pregnancy and menopause. In: Hormones and Breast Cancer, Banbury Report No. 8 (M. C. Pike, P. K. Siiteri and C. W. Welsch, Eds.), Cold Spring Harbor Laboratory, New York, 1981, pp. 3–18.

42. Pike, M. C., Krailo, M. D., Henderson, B. E., Casagrande, J. T., and Hoel, D. G. "Hormonal" risk factors, "breast tissue age" and the age-incidence of breast cancer. Nature 303: 767–770 (1983).

43. Armitage, P. The assessment of low-dose carcinogenicity. Biometrics 38 (Suppl): 119–129 (1982).

44. Mantel, N., and Bryan, W. "Safety testing" of carcinogenic agents. J. Natl. Cancer Inst. 27: 455–470 (1961).

45. Chand, N., and Hoel, D. A comparison of models for determining safe levels of environmental agents. In: Reliability and Biometry, (F. Proschan and R. J. Serfling, Eds.), SIAM, Philadelphia, 1974, pp. 681–700.

46. Schneiderman, M. A. Safe dose? Problems of the statistician in the world of trans-science. J. Wash. Acad. Sci. 64: 68–78 (1974).

47. Hoel, D. G., Gaylor, D., Kirschstein, R., Saffiotti, U., and Schneiderman, M. Estimation of risks of irreversible delayed toxicity. J. Toxicol. Environ. Health 1: 133–151 (1975).

48. Crump, K. S., Hoel, D., Langley, C., and Peto, R. Fundamental carcinogenic processes and their implications for low dose risk assessment. Cancer Res. 36: 2973–2979 (1976).

49. Guess, H. A., Crump, K. S., and Peto, R. Uncertainty estimates for low-dose extrapolations of animal carcinogenicity data. Cancer Res. 37: 3475–3483 (1977).

50. Crump, K. S., Guess, H., and Deal, K. Confidence intervals and test of hypotheses concerning dose response relations inferred from animal carcinogenicity data. Biometrics 33: 437–451 (1977).

51. Brown, C. Statistical aspects of extrapolation of dichotomous dose response data. J. Natl. Cancer Inst. 60: 101–108 (1978).

52. Hartley, H. O., and Sielken, R. L., Jr. Estimation of 'safe doses' in carcinogenic experiments. Biometrics 33: 1–30 (1977).

53. Hartley, H. O., and Sielken, R. L., Jr. Response to Professor Mantel's "Discussion of presentation by Hartley on the estimation

of safe doses and carcinogenic experiments." J. Environ. Pathol. Toxicol. 1: 267–269 (1977).

54. Cornfield, J., Carlborg, F. W., and Van Ryzin, J. Setting tolerances on the basis of mathematical treatment of dose-response data extrapolated to low doses. In: Proceedings of the First International Congress on Toxicology (G. L. Plaa and W. A. M. Duncan, Eds.), Academic Press, New York, 1978, pp. 143–164.

55. Rai, K., and Van Ryzin, J. Risk assessment of toxic environmental substances using a generalized multi-hit dose response model. In: Energy and Health (N. Breslow and A. Whittemore, Eds.), SIAM, Philadelphia, 1979, pp. 99–117.

56. Cornfield, J., Rai, K., and Van Ryzin, J. Procedures for assessing risk at low levels of exposure. Arch. Toxicol. (Suppl). 3: 259–303 (1980).

57. Van Ryzin, J. Discussion on paper by P. Armitage (*43*). Biometrics 38 (Suppl): 130–139 (1982).

58. Cooper, G. M. Cellular transforming genes. Science 217: 801–806 (1982).

59. Kennedy, A. R., Cairns, J., and Little, J. B. Timing of the steps in transformation of C3H 10T½ cells by X-irradiation. Nature 307: 85–86 (1984).

60. Evans, C. H., and DiPaolo, J. A. Neoplastic transformation of guinea pig fetal cells in culture induced by chemical carcinogens. Cancer Res. 35: 1035–1044 (1975).

61. Evans, C. H., and DiPaolo, J. A. Independent expression of chemical carcinogen-induced phenotypic properties frequently associated with the neoplastic state in a cultured guinea pig cell line. J. Natl. Cancer Inst. 69: 1175–1182 (1982).