Environmental Health Perspectives
Vol. 50, pp. 285-291, 1983

Model for Human Carcinogenesis: Action of Environmental Agents

by Suresh H. Moolgavkar*

Current statistical prescriptions for low-dosage extrapolation of carcinogenic risk make no distinction between exposure to initiators and exposure to promoters despite the abundant data that these two classes of carcinogens have different modes of action. One reason for this is the lack of an appropriate model. In this paper, a model for carcinogenesis is presented which provides a framework for understanding the roles of "spontaneous" events, hereditary factors, and environmental agents in human carcinogenesis and for interpreting experimental carcinogenesis. This model incorporates two features: transition of target stem cells into cancer cells via an intermediate stage in two irreversible steps and growth and differentiation of normal target and intermediate cells. Cast in mathematical terms, the model can be fitted to age-specific incidence data on human cancers of both children and adults and can illuminate the relative importance of agents that affect transition rates, tissue growth and tissue differentiation. Within the context of the model, initiators act by affecting the transition rates, whereas promoters influence the kinetics of growth, especially of initiated cells. The model provides a good quantitative description of the epidemiology of carcinomas of the breast and of the lung. The data are consistent with the notion that hormones and cigarette smoke act as promoting agents in carcinoma of the breast and of the lung, respectively.

Much attention has been focused on the phenomena of "initiation" and "promotion" in carcinogenesis. The fact that agents with such disparate modes of action—initiators acting at the level of DNA as mutagens and promoters acting at the cell surface to cause cell proliferation—can facilitate carcinogenesis strongly suggests that both mutations and the dynamics of tissue growth and differentiation play important roles in carcinogenesis. The present paper is predicated on this assumption. I shall present a schematic model for carcinogenesis and I hope to convince you that, in our present state of knowledge, this model can accomodate much that is fact about carcinogenesis in animals and humans, and that it provides a framework for the understanding of environmental carcinogenesis in humans. In addition, the model makes interesting and testable predictions regarding the clonality of certain human neoplasms, and suggests new experiments in animals (*1*).

Evidence from diverse sources suggests that carcinogenesis is a multistage process. It is not my intention to review this evidence today, except to point out that some of the most intriguing evidence is derived from human cancers that occur in two forms: a spontaneous form, and one that is inherited in an autosomal dominant form on pedigree analysis. In such cancers inheritance of the gene increases enormously the risk of getting cancer. However, even though all the cells of the affected organ carry the gene, only a few tumors arise, suggesting that inheritance of the gene is not sufficient and that at least one other event is necessary for malignant transformation. An attractive hypothesis is that only one other event is necessary for malignant transformation: a two-stage model for malignant transformation is consistent with the development of homozygosity at a cancer gene locus.

I would like to present a two-stage model for carcinogenesis and deduce some of its consequences. There is strong evidence that the first critical event in carcinogenesis is a mutation: unfortunately, there is much less information on the nature of the subsequent events. Nevertheless, lurking in the background, behind the model, is the notion that both the critical events are mutations, probably at the same site on homologous chromosomes. Specifically, a genetic regulatory schema proposed by Comings (*2*) is attractive. Comings suggests that all cells contain oncogenes capable of coding for transforming factors that can release the cell from normal growth

*Institute for Cancer Research, Fox Chase Cancer Center, 7701 Burholme Avenue, Philadelphia, PA 19111.

constraints. These oncogenes, which are tissue-specific, are temporarily expressed during embryogenesis but are turned off by diploid pairs of regulatory genes. According to a two-event model then, most human malignant tumors arise when both regulatory genes in a diploid pair sustain mutations, thus releasing the oncogenes from control. Of course, the possibility that oncogenes may be turned back on in other ways must be admitted. For example, chromosomal rearrangement to bring an oncogene adjacent to a "promoter" site may do this. Insertion of a viral promoting sequence next to an oncogene would have the same effect. The word promoter here is not to be confused with promotion in chemical carcinogenesis. It is unfortunate that the same word is used in a technical sense in both molecular genetics and chemical carcinogenesis. Figure 1 is a schematic representation of the model.

Initiation and Promotion

In chemical carcinogenesis, the words "initiation" and "promotion" have hitherto been used phenomenologically: a promoter is a substance which when applied after (but not before) an initiator gives rise to tumors. Thus, these notions, strictly speaking, have meaning only in the context of animal experiments. However, they have been picked up and

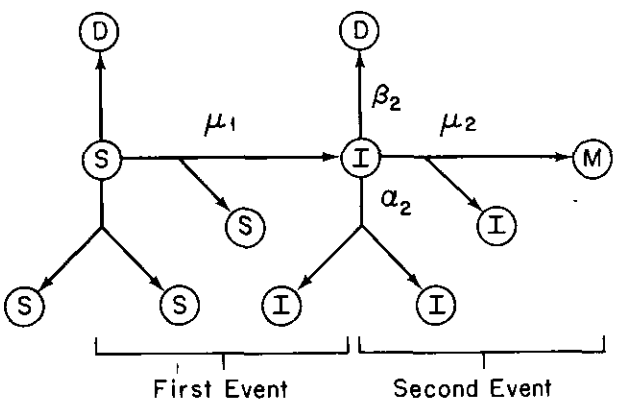

FIGURE 1. Two-stage model for carcinogenesis. S = normal stem cell, I = intermediate (one-hit) cell, D = differentiated (or dead) cell, M = malignant cell; μ_1 = rate at which first event occurs, μ_2 = rate at which second event occurs, α_2 = rate of division of intermediate cells, β_2 = rate of differentiation and death of intermediate cells. In a small time interval, a given stem cell (S) may divide with a certain probability to give rise to two daughter cells (S), or it may differentiate (or die) (D) and thus leave the pool of susceptible cells, or it may divide (with a small probability) into two cells one of which is normal (S) and the other of which has suffered the first event to become an intermediate cell (I). The intermediate cell may in turn give rise to two intermediate daughters (I); die or differentiate (D); or give rise (with a small probability) to one intermediate cell (I) and one malignant cell (M).

used rather loosely by scientists in other disciplines. Within the context of the model presented above, the words "initiator" and "promoter" can be given precise meaning. An initiator is any agent that increases μ_1, the probability of transition from a normal stem cell to a once-hit intermediate cell. A promoter is an agent that acts on intermediate cells to increase α_2, decrease β_2, or both. This results in a proliferation of intermediate cells to give rise to intermediate lesions. The papillomas that arise in the now-classical skin-painting experiments and the enzyme-altered foci that have been described in AAF-induced rat hepatocarcinogenesis can be considered to be examples of such intermediate lesions. Thus, promoters act on cells that have sustained a single critical event to cause their proliferation, i.e., they act as selective mitogens. This proliferation leads to an enormous increase in the number of intermediate cells, and thus increases the probability that one of these cells will sustain the second critical event and become malignant. Such a scenario has been proposed independently by Potter (*3*) for hepatocarcinogenesis in rats and suggests that an experiment in which application of an initiator is followed by several applications of promoter until intermediate lesions appear and then by another application of initiator should yield a larger number of malignant tumors than initiation-promotion alone (which should yield mainly intermediate lesions). The model for initiation-promotion presented above raises some fundamental questions. If both critical events in carcinogenesis are mutations, can there be agents that could be classified as "pure" initiators? Prolonged application of an initiator (a mutagen) should lead eventually to the development of true malignancies. If the initiator has some promoting activity, one should see both intermediate lesions (such as the papillomas) and true malignant lesions. In fact, it is known from the classical skin painting experiments, that a regimen of initiation-promotion leads to a large number of papillomas (intermediate lesions) and a few malignant lesions. On the other hand, prolonged application of initiators leads to the appearance of mainly malignant tumors (*4*). Thus a pure initiator would seem to be a mythical beast. To my knowledge, there is only one agent that has been labeled a pure initiator in the literature: urethane (*5*). However, the studies that conclude that urethane is a pure initiator are far from convincing: they were not carried out for long enough or with large enough populations of animals to rule out the eventual appearance of malignant tumors. All that one can conclude from these experiments is that repeated applications of urethane to mouse skin do not seem to produce papillomas.

If the two critical mutations in carcinogenesis oc-

cur at the same site on homologous chromosomes, an intriguing possibility arises. There could be agents that are not promoters in the sense defined above but still increase the probability of the second event without affecting the first. Homozygosity at the critical site could be brought about not only by mutations but by homologous chromosome exchange during mitosis. Any agent that favors mitotic recombination could act in this way. Thus, within the context of the model, in addition to initiators and promoters, it is possible to postulate the existence of a third class of agents that facilitate carcinogenesis. Perhaps these hypothetical agents should be called completers. As I shall point out later, there is at least one genetic condition in humans that seems to predispose to cancer in this way. Of course, as has been noted above, agents may act in more than one way to facilitate tumorigenesis. Thus, many initiators have some promoting activity, and it has been suggested that the phorbol esters may facilitate mitotic recombination and thus have "completer" activity (*6*).

Genetic Predisposition to Cancer

Thus, within the framework of the model, environmental agents and genetic predisposition to cancer could, broadly speaking, facilitate carcinogenesis in one of two ways: by increasing the transition rates μ_1 or μ_2 or both or by increasing the proliferation of normal or intermediate cells. I shall come back to the implications of this for environmental carcinogenesis shortly. I want to point out first that there are genetic conditions in humans that predispose to cancer in each of these ways. Perhaps the most important examples of genetic predisposition in humans are provided by those cancers that, on segregation analysis, are inherited in an autosomal dominant fashion. Examples are hereditary retinoblastoma and carcinoma of the colon in familial polyposis. According to the model, affected individuals have inherited the first event and are born with the cells of the affected tissue in the intermediate stage (*7*). Then, there are recessive conditions such as xeroderma pigmentosum, in which there is defective repair of ultraviolet-induced DNA damage. Such a condition would simply increase the probability of mutations thus increasing both μ_1 and μ_2. In one recessive condition, Bloom's syndrome, there is a great increase in sister chromatid and homologous chromosome exchanges. As discussed earlier, this would lead to an increase in μ_2 without an increase in μ_1. A possible example of a genetic condition that affects the kinetics of normal tissue leading to an increased susceptibility to cancer is provided by Fanconi's anemia, an autosomal recessive condition. The kinetics of bone marrow cells are severely disturbed in individuals with this disease, and afflicted individuals are at increased risk of leukemia. There is no evidence that DNA repair is deficient in individuals with Fanconi's anemia (*8*).

Age-Specific Incidence in Human Populations

I would like to turn my attention now to some aspects of the epidemiology of human cancers. First, the two-stage model presented above can be cast into mathematical form to yield an expression for the age-specific incidence rate of cancer of a specific site. The age-specific incidence rate per 100,000 individuals in the population is given by $I(t) \times 10^5$, where

$$I(t) \approx \mu_1 \mu_2 \int_0^t X(s) \{(\alpha_2 - \beta_2)(t - s)\} ds \qquad (1)$$

Here μ_1 and μ_2 are the transition rates per cell per year, and $X(s)$ represents the number of normal susceptible cells at time (age) s (*1,9*). This expression for age-specific incidence is not as simple as that derived from the Armitage and Doll (*10*) model. However, it should be noted that the growth of normal and intermediate cells enters explicitly into the equation. It should also be noted that (1) the transition rates μ_1 and μ_2 are multiplicative factors and are important in determining the overall incidence rates of the cancer in question; however, they do not influence the shape of the incidence curve; and (2) the shape of the incidence curve is strongly influenced by the growth curve of the normal tissue and the cellular kinetics of intermediate cells. Moreover, it is the difference between "birth rate" and "death rate," i.e., $\alpha_2 - \beta_2$, that affects incidence, not the individual parameters. The expression (1) for age-specific incidence rate applies only when μ_1, μ_2, α_2 and β_2 remain constant. If these parameters change in response to changes in environment, then appropriate changes must be made in the expression for $I(t)$. However, it is fair to say that initiators (mutagens) which affect μ_1 and μ_2, have a much smaller effect on the shape of the incidence curve than promoters, which affect $\alpha_2 - \beta_2$. I would like to note also that, unlike the Armitage-Doll model which generates the age-specific incidence curves only of many adult carcinomas, the two-stage model described here generates the age-specific incidence curves of childhood tumors and exceptional tumors such as breast cancer in females as well (*1,9,11*). Moreover, the model shows that the shape of the age-specific incidence curve of cancer of a specific tissue is dependent upon the kinetics of growth and differentiation of that tissue, a biologically appealing result.

It is known that many human cancers have age-specific incidence curves with characteristic shapes

that are similar in various populations, even though the rates of these cancers may vary enormously (*12,13*). Thus, within the framework of the model, the international variation in rates is due to differences in μ_1 or μ_2, or both. An obvious exception is lung cancer in which the promotional effect of cigarette smoke is probably important. This is discussed later.

Action of Environmental Agents

Let me turn my attention now to the action of environmental agents. I shall restrict my discussion to chronic exposure. The implications of the model for short-term high intensity exposure to agents such as radiation has been discussed in a recent publication (*1*). In environmental carcinogenesis, we need to consider not only the consequences of chronic exposure to agents, but also the consequences of discontinuing exposure to these agents. For simplicity, I shall only discuss the case in which exposure begins at birth. Exposure beginning later in life yields essentially the same conclusions. Suppose that an agent acts directly on the transition rates and that exposure to such an agent starts at time *0*. Then the expression for the age-specific incidence rates shows that the risk in persons exposed to a given dose relative to nonexposed individuals remains constant with time (regardless of whether the first or the second transition rate is affected) provided that exposure to that dose increases the transition rates to new constant levels. On the other hand, if an agent increases the proliferation of intermediate cells by a constant amount (i.e., increases $\alpha_2 - \beta_2$ without affecting transition rates), then the risk in persons exposed relative to those not exposed increases with time. Thus, within the context of the model, if an agent (promoter), affects the kinetics of intermediate cells, then duration of exposure to this agent is an effect modifier in standard epidemiologic parlance (Fig. 2). This is because the ratio of the number of intermediate cells in exposed individuals to that in nonexposed individuals increases with time. This ratio remains constant with time when the agent affects transition rates alone. The model also predicts that promoting agents are remarkably efficient in affecting incidence: small changes in $\alpha_2 - \beta_2$ lead to large changes in incidence. Small changes in the growth kinetics of normal cells have little effect on incidence. However, in some instances the changes may be large enough to affect cancer incidence. Thus, nonspecific stimuli for cell proliferation such as chronic irritation may act in this fashion.

Suppose now that chronic exposure to an environmental agent is stopped. If an agent affects only the first transition rate, then stoppage of exposure leads to an incidence curve that lies in between the incidence curves in exposed and nonexposed individuals. In other words, even after exposure is stopped, the incidence rate never goes back to pre-exposure levels. The reason for this is intuitively clear: during exposure, there is a build-up of cells in the intermediate compartment. On the other hand, if the agent affects the second transition rate alone, discontinuance of the exposure leads to a quick reversion of the incidence rate to pre-exposure levels.

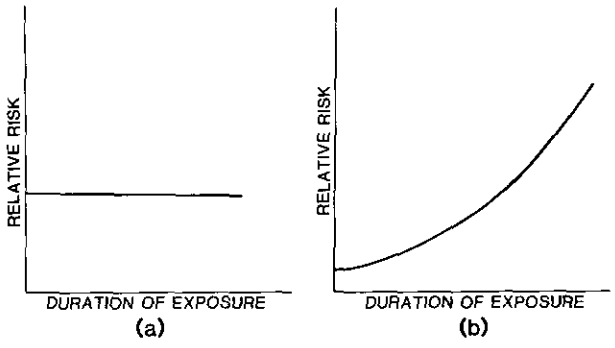

FIGURE 2. Relative risk (a) as a function of time when chronic exposure to an agent (from time 0) increases μ_1 or μ_2 (or both) to a new constant level and (b) as a function of time when chronic exposure to an agent increases $\alpha_2 - \beta_2$.

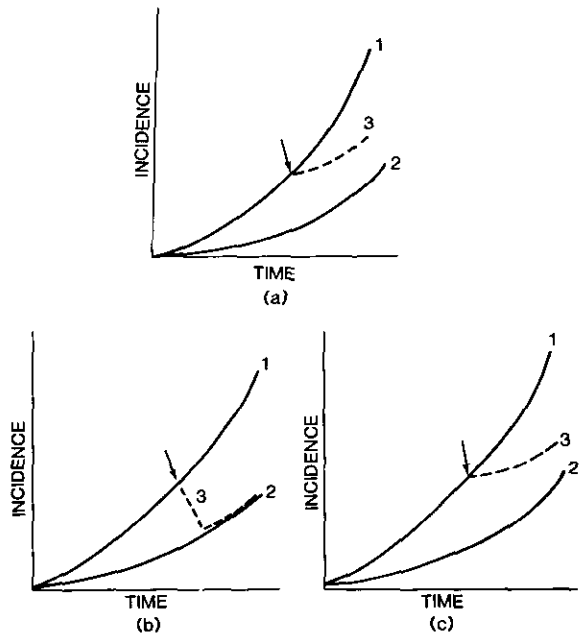

FIGURE 3. Effect of stopping chronic exposure to environmental agents: (*a*) exposure to an agent that increases μ_1 (initiator); (*b*) exposure to an agent that increases μ_2; (*c*) exposure to an agent that increases $\alpha_2 - \beta_2$ (promoter). In all three figures, (1) is the incidence curve in the exposed population, (2) is the incidence curve in the unexposed population and (3) is the incidence curve after exposure is stopped at the time indicated by the arrow.

Finally, if exposure to a promoter is discontinued, the incidence rates will again lie between the incidence curves among exposed and nonexposed individuals (Fig. 3).

Epidemiology of Lung and Breast Cancer

Doll (*14*) has constructed an age-specific incidence curve for lung cancer in nonsmokers. Although this curve is based on a small number of age groups, it seems likely that the age-specific incidence rates in nonsmokers increases with about the fourth power of age. As can be seen from Figure 4, the model provides an excellent description of the data. With regard to the influence of cigarette smoke, various studies (*15,16*) suggest that for a given daily level of smoking, the relative risk increases with duration of smoking. That is, duration of smoking is an effect modifier. In addition, data are now becoming available on smokers who quit. These data indicate that the risk in exsmokers has not reverted to the risk in the nonsmoking population in 15 years. However, the sharp increase in risk seen in continuing smokers is averted. Thus, the minimal hypothesis that is consistent with all the facts is that cigarette smoke increases the proliferation of the intermediate cells, i.e., acts as a promoter. Indeed there is evidence that cigarette smoke causes hyperplasia of the bronchial epithelium (*17*) and is known to contain tumor promoters (*18*). Thus, even though cigarette smoke contains mutagens and would, therefore, be expected to increase the transition rates to some extent, the most important effect of smoke would seem to be promotion.

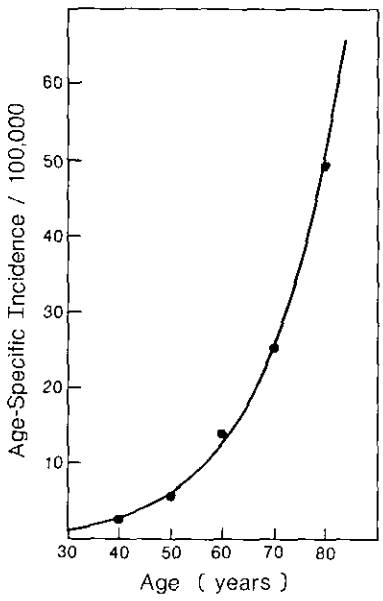

FIGURE 4. Fit of the model to age-specific incidence data on carcinoma of the lung in nonsmokers. From Moolgavkar and Knudson, (*1*).

I would now like to discuss another example from human epidemiology. Breast cancer in females is a most interesting tumor in that it has a peculiar age-specific incidence curve and age at first full-term pregnancy is a well recognized risk factor. If the age-specific incidence of breast cancer is observed during a single time period, two distinct patterns emerge. In Western populations, in which risks are high and more or less stable, the age-specific incidence rates rise till about menopause, level off and then continue to rise though less steeply than before menopause. In Eastern populations, exemplified by the prefecture of Osaka in Japan, in which risks are low and increasing, the age-specific incidence curve rises till menopause, levels off and then shows an actual fall. However, this fall in rates in cross-sectional data is due to the cohortwise increase in risks in Osaka. After adjustment for cohort effects, the age-specific incidence curve continues to rise after menopause as it does in the West (*13*). Thus, the shapes of the age-specific incidence curves of female breast cancer are identical in Eastern and Western populations: it is the magnitude of the rates that sets these populations apart.

When the physiological responses of the breast tissue to menarche and menopause are incorporated into the two-stage model it generates an age-specific incidence curve that describes well the shape of the age-specific incidence curve of female breast cancer (Fig. 5). The cellular transition rates μ_1 and μ_2 determine the magnitude of the incidence rates in a population, and adjustment of the transition rates then generates curves that are in close quantitative agreement with those observed in six test populations: Connecticut, Denmark, Finland, Slovenia, Iceland, and Osaka (*11*). According to the model, hormones influence the epidemiology of breast cancer in females by their actions on the kinetics of growth of nonneoplastic (normal and intermediate cells) breast tissue. The breast grows in response to hormonal stimuli at puberty, and it involutes when these stimuli are removed at menopause. It is these two major kinetic changes imposed on the breast by hormonal influences that, according to the two-stage model, determine the basic shape of the age-specific incidence curve of carcinoma of the breast.

A full-term pregnancy is assumed to cause a certain fraction of normal and intermediate cells to undergo differentiation and thus remove this fraction from the population of susceptible cells. Experimental evidence that such a mechanism operates in rat mammary glands has recently been obtained by

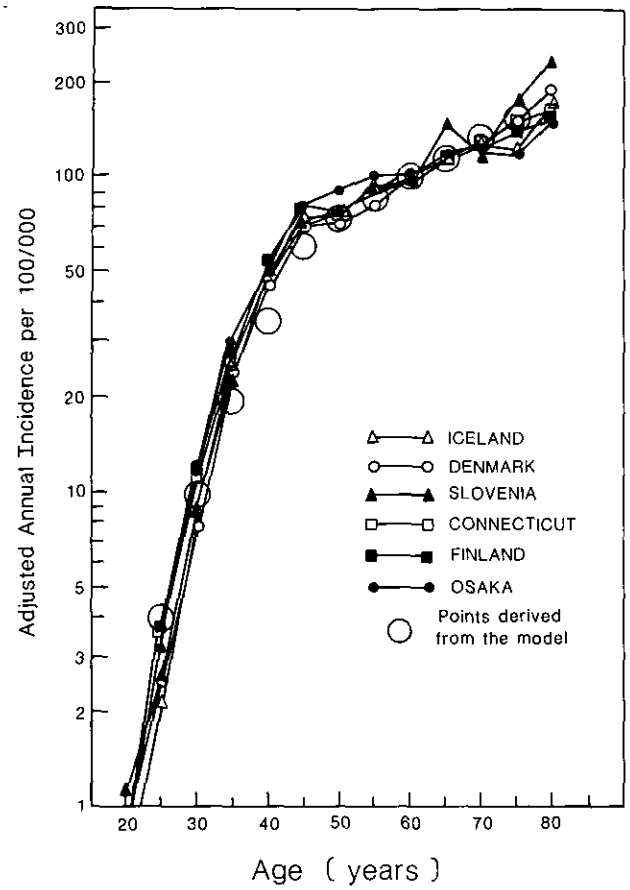

FIGURE 5. Fit of the model to age-specific breast cancer incidence data from various populations. All curves have been normalized so that the sum of the rates over all age groups is the same. From Moolgavkar et al. (*11*).

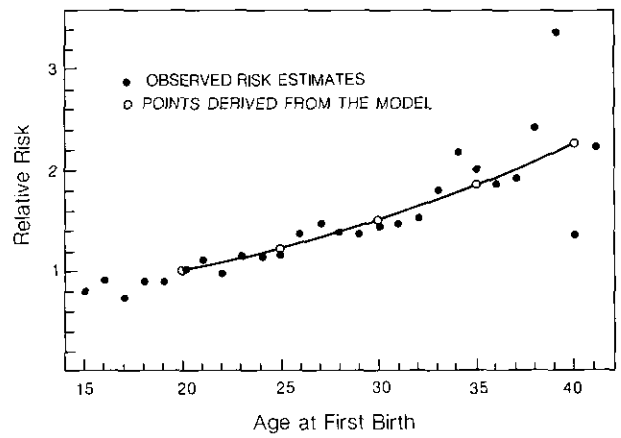

FIGURE 6. Risk of breast cancer for different ages at first full-term pregnancy relative to first full-term pregnancy at age 20 compared with predictions from the model. From Moolgavkar et al. (*11*).

Russo and Russo (*19*). With this assumption, the model predicts a protective effect of an early first-birth that is in good agreement with data from a multinational study (Fig. 6) (*11*). Full-term pregnancies after the first could also stimulate differentiation and further decrease risk, and indeed, a recent prospective study in Iceland concluded that full-term pregnancies after the first decreased risk. Thus, in our view, the features of the epidemiology of breast cancer that indicate that risk is mediated hormonally arise from the hormonal effect on tissue kinetics. Since both the age-specific incidence curve and the first-birth effects, which are strongly dependent on tissue kinetics, are similar in populations with different levels of risk, one might infer that the cell kinetics of breast tissue and hormone levels are similar in the different populations. The studies that have reported no noteworthy differences in hormone levels between low-risk and high-risk populations are consistent with this viewpoint.

Finally, the model provides a logical explanation of genetic predisposition to breast cancer and of breast cancer in irradiated women. These issues are dealt with in detail in Moolgavkar et al. (*11*.)

To summarize, I have presented a biologically based two-stage model for carcinogenesis. According to this model, carcinogenesis is the result of two critical events, probably mutations. Initiation, i.e., acquisition of the first event leads to an improperly controlled proliferation of cells that have sustained this event. It is on these intermediate, or first stage, cells that promoters act as selective mitogens. Thus, the two stages in the model refer to the occurrence of the two critical events. Promotion is not a stage; rather promoters act on cells in the first stage to cause their expansion into clones. Human epidemiologic data are consistent with the view that estrogens act as promoters in female breast cancer, and that cigarette smoke acts as a promoter in lung cancer. The model is consistent with facts both from human epidemiology and from animal experiments. In view of the different modes of action of initiators and promoters, it is clear that single statistical prescriptions for extrapolation of risk to low doses cannot possibly apply in all situations. Sensible extrapolations can be made only when the mode of action of the specific agent under study is known.

REFERENCES

1. Moolgavkar, S. H., and Knudson, A. G. Mutation and cancer: A model for human carcinogenesis. J. Natl. Cancer Inst. 66: 1037-1052 (1981).
2. Comings, D. E. A general theory of carcinogenesis. Proc. Natl. Acad. Sci. (U.S.) 70: 3324-3328 (1973).
3. Potter, V. R. Initiation and promotion in cancer formation: the importance of studies on intercellular communication. Yale J. Biol. Med. 53: 367-384 (1980).

4. Stenback, F., Garcia, H., and Shubik, P. Present status of the concept of promoting action and co-carcinogenesis in skin. In: Physiopathology of Cancer, Vol. I (P. Shubik, Ed.), S. Karger, Basel; 1974, pp. 155-225.
5. Salaman, M. H., and Roe, J. F. C. Incomplete carcinogens: ethyl carbamate (urethane) as an initiator of skin tumor formation in the mouse. Brit. J. Cancer 7: 472-481 (1953).
6. Kinsella, A. R., and Radman, M. Tumor promoter induces sister chromatid exchanges: Relevance to mechanisms of carcinogenesis. Proc. Natl. Acad. Sci. (U.S.) 75: 6149-6153 (1978).
7. Knudson, A. G. Mutation and cancer: statistical study of retinoblastoma. Proc. Natl. Acad. Sci. (U.S.) 68: 820-823 (1971).
8. Knudson, A. G. Genetics and etiology of human cancer. In: Advances in Human Genetics, Vol. 8 (H. Harris and K. Hirschhorn, Eds.), Plenum Press, New York, 1977, pp. 1-66.
9. Moolgavkar, S. H., and Venzon, D. J. Two-event models for carcinogenesis: incidence curves for childhood and adult tumors. Math. Biosci. 47: 55-77 (1979).
10. Armitage, P., and Doll, R. The age distribution of cancer and a multistage theory of carcinogenesis. Brit. J. Cancer, 8: 1-12 (1954).
11. Moolgavkar, S. H., Day, N. E., Stevens, R. G. Two-stage model for carcinogenesis: epidemiology of breast cancer in females. J. Natl. Cancer Inst. 65: 559-569 (1980).
12. Cook, P. J., Doll, R., and Fellingham, S. A. A mathematical model for the age distribution of cancer in man. Int. J. Cancer 4: 93-112 (1969).
13. Moolgavkar, S. H., Stevens, R. G., and Lee, J. A. H. Effect of age on incidence of breast cancer in females. J. Natl. Cancer Inst. 62: 493-501 (1979).
14. Doll, R. The age distribution of cancer: implications for models of carcinogenesis. J. Roy. Soc. A134: 133-166 (1971).
15. Hammond, E. C. Smoking in relation to the death rates of one million men and women. Natl. Cancer Inst. Monographs, Vol. 19, 1966, pp. 127-204.
16. Stevens, R. G., and Moolgavkar, S. H. Estimation of relative risk from vital data: smoking and cancers of the lung and bladder. J. Natl. Cancer Inst. 63: 1351-1357 (1979).
17. Auerbach, O., Stout, A. P., Hammond, E. C., and Garfinkel, L. Changes in bronchial epithelium in relation to cigarette smoking and in relation to lung cancer. New Engl. J. Med. 265: 253-267 (1961).
18. Van Duuren, B. L., and Goldschmidt, B. M. Cocarcinogenic and tumor-promoting agents in tobacco carcinogenesis. J. Natl. Cancer Inst. 56: 1237-1243 (1976).
19. Russo, J., and Russo, I. H. Susceptibility of the mammary gland to carcinogenesis II. Pregnancy interruption as a risk factor in tumor incidence. Am. J. Pathol. 100: 497-512 (1980).