For WSPA Contract No. EG 004-07: "PREPARE JOURNAL-QUALITY ARTICLES ON NEW DIRECTIONS IN CANCER MODELING FOR BENZENE REPORT"

CANCER RISK ASSESSMENT FOR CHEMICAL CARCINOGENS: FROM STATISTICAL TO BIOLOGICALLY-BASED MODELS

Louis Anthony Cox, Jr.

November 20, 1989

Abstract

Chemical carcinogens act by a variety of biological mechanisms to increase the rates of cancer induction and cancer expression. Despite this diversity, several unifying "biologically based" risk-modeling approaches have been developed recently that can help to make more accurate and realistic the quantitative assessment of cancer risks from exposures to chemicals. This paper summarizes the traditional linearized multistage (LMS) statistical model used in a majority of regulatory cancer risk assessments and then presents three more recent approaches -- the Moolgavkar *et al* two-stage (MVK) model, physiologically-based pharmacokinetic (PB-PK) modeling, and molecular epidemiology approaches -- that are starting to change the ways in which cancer risks are assessed. These approaches share an emphasis on incorporating more biology into risk assessment to obtain better risk estimates. This paper reviews and evaluates these approaches to biologically-based cancer risk assessment and identifies several new directions for research.

Introduction

Chemical carcinogens are one of only three broad classes of agents -- chemicals, radiation, and viruses -- that are now recognized to be possible causes of cancer. They can act by many different mechanisms to increase the rates of cancer induction, promotion, or progression. Known mechanisms range from purely "genotoxic" effects (e.g., covalent binding of an alkylating agent directly to DNA, which can "initiate" cancer by producing point mutations, chromosomal breaks or translocations, or other heritable damage in somatic cell lines) to purely "epigenetic" effects, e.g., binding of chemical molecules to cell surface membrane receptors, which may thus be either artificially stimulated or artificially inhibited in sending signals to the proliferation control mechanisms in the cell's nucleus (Castagna and Martelly, 1989). Some chemicals, such as 1,3-butadiene (an olefin widely used in synthetic rubber manufacturing) are metabolized by enzyme-catalyzed reactions into reactive metabolites (monoepoxides and diepoxides) that can then activate cancer-inducing retroviruses in susceptible mouse strains, as well as binding chemically directly to DNA. Others, such as benzene metabolites (phenol and hydroquinone) stimulate compensating proliferation of bone marrow stem cells in response to cytotoxic (cell-poisoning) damage, thus increasing the numbers of cells at risk of leukemic transformation, apart from any direct genotoxic effects. Although the biochemical details of such mechanisms are generally not yet well understood, substantial progress has been made in understanding how to quantify the effects of exposure to chemical carcinogens on human cancer risks. Key to this progress

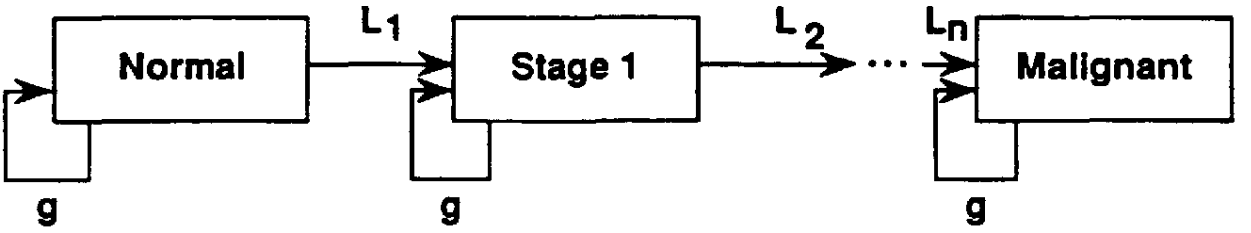
has been the increasing use of biological measurements and knowledge in cancer risk modeling. This paper describes the growing use of biological insights, molecular-biological data, and biomathematical modeling techniques in quantitative cancer risk assessment for chemical carcinogens. It also elucidates some issues that have been incompletely explored in modeling cancer dose-response functions.

It is now possible to contrast an emerging generation of "biologically motivated" cancer risk assessment models with the older generation of statistical "curve-fitting" models still widely used by regulatory agencies (Thorslund et al, 1987.) Whereas the former use plausible hypotheses about how cancer induction might work to suggest algebraic forms for dose-response models, the latter assume algebraic forms without detailed biological justification. Model parameters must then be estimated from empirical data by maximum-likelihood estimation (MLE) or similar statistical methods, rather than being measured directly in the laboratory. The distinction between "biologically motivated" and "statistical" dose-response modeling is often exaggerated, however. Since the 1950's, the most successful statistical cancer risk assessment models have been based on assumptions about the biology of carcinogenesis. In particular, the linearized multistage (LMS) model -- a family of dose-response functions developed during the 1970's and still widely applied by the EPA's Carcinogen Assessment Group (CAG) and other Federal and State regulatory groups -- is based on a simple, speculative model of cancer causation.

The [illegible]riz[illegible] ultistage model began with simple probability models of carcinogenesis developed by Armitage and Doll (1961) during the 1950's and 1960's. It was extended for EPA's CAG into a practical tool for routine risk assessment of genotoxic chemical carcinogens by K.S. Crump and colleagues during the 1970's (Crump and Watson, 1979; Crump, 1984.) The LMS paradigm today dominates regulatory risk assessments for chemical carcinogens. However, an alternative modeling framework for cancer risk assessment has been growing up throughout the 1980's that seeks to account for additional biological phenomena -- e.g., cell proliferation, differentiation, and death -- that play key roles for some carcinogens, especially for tumor "promoters" and non-genotoxic carcinogens. Beginning with pioneering investigations by Knudson on the inheritance of retinoblastoma (Weinberg, 1988) and biomathematical modeling by Moolgavkar and Venzon in the late 1970's, this "Moolgavkar-Venzon-Knudson" (MVK) model has become a powerful alternative to LMS for understanding qualitative cancer dynamics and the quantification of chemically-induced cancer risks (Moolgavkar et al, 1988, 1989.)

This paper first reviews and critically evaluates the biological motivation for the Armitage-Doll model and the LMS framework. Then, three more recent biologically-based risk modeling approaches are described -- the Moolgavkar-Venzon-Knudson (MVK), physiologically-based pharmacokinetics (PB-PK), and "molecular epidemiology" frameworks -- that offer additional understanding to quantitative cancer modeling. Strengths and limitations of these new approaches are summarized. Finally, several new directions for research are

FIGURE 1: Armitage-Doll Model (1961)



Assumptions

- Stages traversed in a fixed order
- Constant transition rate for all cells in a stage. Independent transitions.
- Constant growth rates, same for all nonmalignant cells.
- Rare transitions (small L_i)

Conclusion

- Hazard function for tumor arrival is

$$h(t) = KL_1L_2\ldots L_n(t-w)^{n-1}$$

tumor growth time

$$\approx Kt^{n-1}$$

- Fits empirical data well for n = 5.

TABLE 1: BIOLOGICAL ASSUMPTIONS OF E[illegible] LINEARIZED MULTISTAGE CANCER RISK MODEL

1. Dose rate is constant over time, i.e., the biologically effective dose acting on cell populations per unit time is treated as constant. If dose rate varies, then it is assumed that only the equivalent constant "average" dose rate (averaged over a subject's lifetime) is relevant in determining lifetime probability of response: the time pattern of exposure is ignored. (Taken at face value, this assumption implies that the cancer process must anticipate future exposure so that it can respond to the correct lifetime average level.)

2. There is only one path (set of cellular transformations) that can lead a cell line to malignancy. Cell lines undergo stochastic (random) transitions among a fixed set of stages along the path from normalcy to malignancy, with each stage representing the acquisition of another heritable transformation (broadly speaking, a mutation.) A necessary and sufficient condition for cancer induction is that all of the stages be traversed. (This precludes the possibility of several paths, involving different transformations, leading to malignancy.)

3. Stages must be traversed in a fixed order. In some variants, the order of traversal can be arbitrary (Whittemore and Keller, 1978). But the general case of partially ordered stages, in which some stages must precede some specific others in order to produce malignancy, has not been considered.

4. Cells divide and undergo transformations independently of each other, i.e., regardless of what other cells are doing or of the relative sizes of different cell subpopulations. (This assumption rules out compensating proliferation, in which stem cells spend more time actively dividing, and hence at risk of e.g., leukemic transformation.)

5. All nonmalignant cells proliferate at the same rate, independent of dose or stage. (This precludes the possibility of cytotoxic effects, e.g., in which a dose of benzene delays or stimulates cell divisions.)

6. Transformation rates are linearly related to dose. The transformation rate for cells in stage j (i.e., the transition rate into stage j+1) increases linearly with the biologically effective dose acting on the cells in stage j. (This is essentially the "one hit" hypothesis.)

7. The number of "normal" (non-transformed) cells in a target organ remains approximately constant over time. This equilibrium level, say N, may be homeostatically regulated to maintain its value. (This is probably not a useful assumption for leukemias, since the number of blood cells at risk responds to changing physiological conditions of the subject at risk.)

References: Armitage and Doll (1961), Whittemore and Keller (1978), Crump (1984), Brown and Chu (1989).

identified and discussed.

1. The Armitage-Doll Cancer Model

The main aspects of the Armitage-Doll (1961) model are outlined in Figure 1. It relies heavily on several simplifying assumptions that are listed, with critical comments, in Table 1. Using versions of these assumptions that apply to the case of no exposure, i.e., zero external dose, Armitage and Doll mathematically showed that the background hazard function for cancer induction should increase in proportion to a power of age. When this conclusion was tested empirically, it was found that appropriate choices of the power and the constant of proportionality would allow empirical age-specific cancer incidence curves for many types of cancers -- especially carcinomas -- to be accurately reproduced. Thus, the basic Armitage-Doll risk model $h(t) = at^{k-1}$, where a and k were parameters to be estimated statistically from known human incidence data, was established as a useful tool for predicting age-specific cancer hazard functions.

2. Allowing Dose-Dependent Transformation Rates: The LMS Model

During the 1970's, EPA propelled the further development of the Armitage-Doll model to allow for dose-dependent cellular transformation rates. Using the assumptions in Table 1, researchers from the Carcinogen Assessment Group (CAG) obtained the following natural

generalization of the Armitage-Doll model:

$$h(t) = c(a_1 + b_1x)(a_2 + b_2x)\ldots(a_k + b_kx)t^{k-1}.$$

Here, x is dose rate, k is the number of stages that a cell line must traverse to malignancy; $c = N/k!$ where N is the (constant) number of normal cells at risk; and $(a_j + b_jx)$ is the dose-dependent transformation rate of cells at stage j-1 into stage j. This is the basis for most of the EPA's subsequent policies on generic methodologies for cancer risk assessment.

To enable cancer risk assessment, EPA supported development of a series of computer programs (the GLOBAL programs) for estimating the parameters of the LMS model (Crump and Wasson, 1989.) These programs combine maximum likelihood estimation (MLE) and some statistically ad hoc extrapolation procedures to establish estimated dose-response functions and confidence limits. To simplify calculations, the terms in the above expression for the hazard function are usually multiplied out and then integrated over the lifetime of the exposed subject to obtain the following algebraic expression for lifetime cumulative hazard:

$$H(x) = q_0 + q_1x + q_2x^2 + \ldots + q_kx^k,$$

where x is the lifetime average dose rate (e.g., in mg of carcinogen per kg of body weight per day) and the q_j are the parameters of the model (constrained to be nonnegative) estimated from the data --

usually, animal bioassays. A quantal dose-response model for subject's lifetime probability of cancer, say p(x), follows directly from the probability formula $p(x) = 1 - \exp[-H(x)]$. For very small values of x (e.g., for $x < 0.1$), $1 - \exp[-H(x)]$ is approximated by H(x), quadratic and higher terms approach zero, and the model simplifies to $p(x) = q_0 + q_1 x$. This is the algebraic form most often used by the EPA in practical applications. In this case, q_1 can be interpreted as the carcinogenic potency of the chemical carcinogen, defined as the rate of increase in lifetime probability of cancer per unit increase in x (i.e., q_1 is the slope of the dose-response curve at very low doses.) Often, the GLOBAL program's upper confidence limit for the potency parameter, denoted q*, is used as a conservative estimate of q_1. Scaling q* to humans, e.g., through an interspecies allometric formula -- which might, for example, scale the dose level x according to the 2/3 or 3/4 power of the ratio of human to animal body weights to adjust for differences in metabolic rates and other factors (Travis and White, 1988; Chappell, 1989) -- gives a unit risk estimate for humans, defined as the increase in lifetime probability of cancer from continual exposure for 70 years to 1 ug/m^3 of the chemical inhaled by a 70 kg subject.

The combined statistical and theoretical approach that has grown out of the assumptions in Table 1 is the linearized multistage (LMS) risk model. EPA uses the LMS model as the regulatory method of choice for quantitative cancer risk assessment (Andersen, 1983.)

3. Critique of the Armitage-Doll and LMS Models

Progress in understanding carcinogenesis has revealed serious flaws in the Armitage-Doll model and in the LMS model. Some of the main objections are as follows:

1. The biological assumptions on which the Armitage-Doll and LMS risk models are based are incorrect. Specifically, addressing in order the assumptions in Table 1,

(a) Intermittent doses are important in practice and can not be accurately replaced with "equivalent" lifetime average doses (Crump and Howe, 1984; Murdoch and Krewski, 1988.) In general, there will be no constant exposure level that is "equivalent" in its health effects to a specific time-varying dose pattern, e.g., because of the important transient responses of metabolite levels produced by time-varying exposure concentrations. For example for some chemicals, including butadiene and benzene, it has been found in laboratory animals that reducing exposure concentration and increasing exposure duration in such a way that concentration x duration (e.g., in ppm-weeks) remains constant can dramatically decrease carcinogenic response (Cronkite, 1987.)

(b) More than one set of mutations or other critical molecular events may lead to the same carcinogenic endpoint. A clear understanding of the mechanisms of induction of retinoblastoma has now emerged (Weinberg, 1988) to show that a fixed order of mutations is not required for this or similar neoplasms.

(c) Lesions in some cells may make it easier for neighboring cells to acquire lesions (e.g., because of impaired cellular contact control or secretion of unusual growth factors) so that cell lines do not necessarily undergo transformations independently of each other. Moreover, deaths or deficiency of proliferation in a mature cell subpopulation can stimulate compensating proliferation in a less differentiated subpopulation, as has been observed experimentally for, e.g., benzene bone-marrow toxicity (Cronkite, 1987.) Thus, cells do not necessarily divide independently of each other: a change in the proliferation rate in one cell population may be informative about the probable rates in other populations.

(d) Cell proliferation rates can be sensitive to dose through various cytotoxic mechanisms, e.g., selective poisoning of normal cells compared to transformed (resistant) cells or delay of mitosis by cycle-specific cytotoxins, including metabolites of benzene (Eastmond et al, 1987.) Thus, cell proliferation rates are not necessarily independent of dose; moreover, they are probably not independent of stage in the sequence of transformations leading to full malignancy. For example, in the case of leukemia, blood cells in different lineages and stages of transformation typically divide at different rates and respond to different molecular regulatory signals.

(e) The one-hit hypothesis is incorrect for many cancers, e.g., for bladder cancers caused by chemically-induced bladder stones (a non-hit mechanism.) It may be incorrect even for some cancers that are caused by genotoxic damage. If each strand of DNA in a chromosome must be

hit within a certain time window (namely, before DNA repair mechanisms heal the damage or the cell divides) to sever the chromosome and cause heritable damage in the cell line, then a two-hit (quadratic) dose-response relation results (Thorslund et al, 1988.) Thus, a linear relation between internal (e.g. cell-specific) dose and transformation rate can not be assumed in general. Although EPA has long used the one-hit model based on an assumption that it is "conservative" (unlikely to underestimate true risk) (Andersen et al, 1983), even this assumption has recently been questioned (Bailar et al, 1988.)

(f) The number of target cells at risk in some neoplasms, including leukemia, is not constant. For example, blood cell populations respond with great sensitivity to cytotoxic effects of exposures to chemicals such as benzene and can fluctuate by substantial fractions of their normal values (Mehlmān, 1989.)

(g) The mathematical analysis of the linearized multistage model, as implemented in the GLOBAL programs, depends on an assumption that carcinogenic responses are rare. In fact, in the animal experiments to which the model is usually applied, carcinogenic responses are usually very common, invalidating the use of the standard LMS model (Moolgavkar and Dewanji, 1988.)

Thus, many of the assumptions in the LMS model can no longer be considered speculative (i.e., possibly true and possibly false) with respect to current scientific knowledge: they are now known to be incorrect as general postulates for specific aspects of carcinogene-

for this. On the other hand, the Armitage-Doll/LMS model fails to provide adequate representation for some other biological aspects of carcinogenesis that are now recognized to play a critical role in many cases of cancer causation. These omissions include

o Cell death and differentiation. As stem cells differentiate into increasingly mature, decreasingly proliferative, lineages they lose their susceptibility to carcinogenic transformation. An accurate accounting of the number of stem cells at risk of malignant transformation over time must allow for losses due to stem cell death and differentiation.

o Feedback control loops (e.g., based on molecular signal strength from contact regulation) from cell population sizes to the proliferation and differentiation rates of individual cells.

o Effects of the subject's age at the beginning of dosing on his or her susceptibility to carcinogenic transformations (Murdoch and Krewski, 1988.)

o The critical role of mitosis (cell division) in locking-in carcinogenic damage.

Qualitatively, the predictions of the Armitage-Doll/LMS model conflict both with some biological evidence and with other models. For example,

[illegible] stages. To fit empirical data, the Armitage-Doll model requires that carcinogenesis for most carcinomas involves betweent k = 4 and k = 7 stages. But there is no biological evidence that more than two stages are involved in the majority of carcinomas. A k between 4 and 7 "does not have experimental justification and seems somewhat implausible" (Murdoch et al, 1987.) Exactly two stages are required for retinoblastoma and similar carcinomas (Weinberg, 1988.)

o The LMS model is algebraically inconsistent with the general relative risk hypothesis unless it is assumed that the dose variable acts on exactly one stage and is algebraically inconsistent with the general absolute risk hypothesis unless the background hazard function is constant.

Finally, there are several pragmatic and statistical objections to the LMS framework. These include the following two:

o The predictions of the LMS model are often very sensitive to slight variations in its assumptions, e.g., about the number of stages involved in carcinogenesis, the specific stages whose transformation rates are affected by the exposure variable, and so on, as discussed by Sielken (1987).

o The predictions of the LMS model are sometimes insensitive to relevant empirical data. The polynomial form of the hazard function, implied by the assumption of a static, linear increase in transformation rates at each stage, is a mathematical straight jacket

th [illegible] even the best-fitting LMS model from fitting the observed data adequately (Sielken, 1987.) For example, if different terms hold at high doses than at low doses (perhaps because normal metabolic pathways become saturated at high doses) then the LMS framework's implicit assumption that the coefficients $q_0,\ldots,q_k$ have the same values throughout the range of doses will be incorrect and may produce "maximum-likelihood" risk estimates (for this class of models) that do not fit the observed data. The LMS methodology seeks to avoid this problem by successively deleting high-dose observations until an adequate fit is obtained. But doing so requires ignoring some relevant data. A better approach would be to use a less restrictive class of models, e.g., one based on simulation, that does not require a fixed algebraic form for the dose-response function.

In summary, the LMS model has problems in its foundations, as well as statistical and pragmatic drawbacks, that make its predictions suspect in many applications. It has a biological motivation, but one based on assumptions and speculations, once scientifically plausible, that are now known to be wrong for many carcinogens. Perhaps more importantly, the LMS framework fails to capture several important cell-biological phenomena underlying carcinogenesis.

4. A More Recent Approach: The MVK Framework

Numerous attempts have been made to expand the practical applicability of the original Armitage-Doll framework, for example by

allowing for

-- Different proliferation rates at different premalignant stages. (Some of the original work by Armitage and Doll considered this possibility. However, it is now recognized that differential growth rates are probably not critical in most cancer induction.)

-- Dose-dependent transition rates (Crump et al, 1976). This effort eventually led to the linearized multistage risk modeling framework.

-- Exposure to multiple carcinogens acting at different stages (Whittemore and Keller, 1978);

-- Time-dependent dosing patterns (Crump and Howe, 1984; Murdoch and Krewski, 1988.)

-- Exposures to mixtures of carcinogens that interact in affecting stage-specific transition rates (Thorslund and Charnley, 1986; Brown and Chu, 1989).

All of these extensions have been made within the framework of assumptions outlined in Table 1. All suffer from (i) Ad hoc algebraic assumptions about the effect of administered dose on stage-specific transition rates (i.e., the rates at which cell lines at one stage are transformed into cell lines at the next stage); and (ii) Inability to represent the effects of cell birth, death and differentiation in the target tissues (Murdoch et al, 1987.) Thus, they do not overcome

these basic limitations of the Armitage-Doll/LMS framework.

An alternative two-stage cancer risk modeling framework has been developed over the past twenty years that extends and refines the Armitage-Doll model by explicitly accounting for cell birth and death (e.g., chemically induced cell proliferation) as well as cell differentiation and cell transformation via genotoxic (DNA-damaging) chemical reactions ("hits"). Originally proposed for retinoblastoma by Knudson on empirical grounds in 1971 and later developed mathematically by Moolgavkar and Venzon as a general approach for modeling carcinomas, this Moolgavkar-Venzon-Knudson (MVK) model has attracted a great deal of attention recently as a more biologically realistic alternative to Armitage-Doll/LMS. However, debates over its theoretical soundness and practical usefulness are ongoing. The remainder of this section summarizes and evaluates the MVK model.

4.1 Assumptions of the MVK Model

The structure of the MVK model is indicated diagrammatically in Figure 2. Embedded in this structure are the following key assumptions.

A. First Assumption: Two Stages are Involved in Carcinogenesis

Two transformations are required to convert a normal cell line to a malignant one. As in the Armitage-Doll model, each successive transformation corresponds to another "stage" in the progress of the

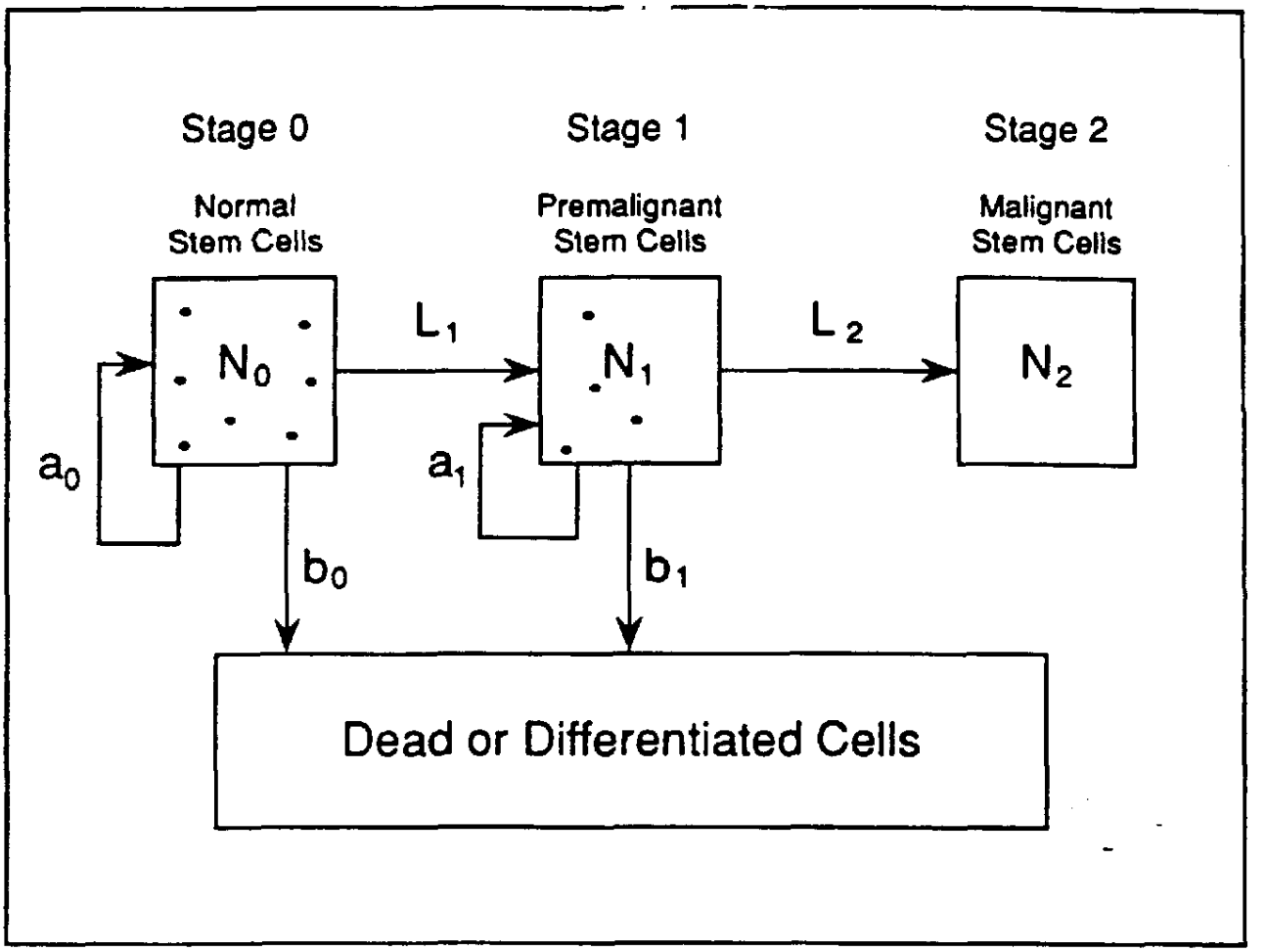


FIGURE 2: MKV Model

cell line toward malignancy. Each cell [illegible] division) into the same stage as its parent. If it acquires an additional, heritable transformation during its life that is not repaired before it divides, then it can be thought of as having made a "transition" to the next stage. In Figure 2, normal cells belong to stage 0. Stage 1 cells have acquired exactly one of the two transformations needed for malignancy. These are called "intermediate" cells in the MVK model (see e.g., Moolgavkar et al. 1988), but are referred to as "premalignant" cells in this paper. Stage 2 cells are the fully malignant ones.

In the MVK framework, the first transformation is generally considered to be a "mutation" in the broad sense of altering a cell line's genotype. The second transformation may or may not involve genotoxic mechanisms.

B. Second Assumption: Cell lines undergo transformations independently of each other. These transformation rates from stage 0 to stage 1 and from stage 1 to stage 2 are denoted in Figure 2 as L1 and L2, respectively. More specifically, each normal cell at risk dies or differentiates, divides into two normal cells, or divides into one normal and one premalignant cell, with average hazard rates (per cell per unit time) of b0, a0, and L1, respectively. (Differentiation is lumped with cell death because either removes the cell line from further risk of cancerous transformation.) Similarly, each premalignant cell at risk either dies (or differentiates), produces two premalignant offspring, or produces one premalignant and one fully

malignant daughter cell, with rates [illegible], and [illegible]. (The probability that a cell will produce two offspring in the next stage is considered to be negligible compared to the probability that it will produce one.) The probability that a dividing cell produces two daughter cells in its own stage is orders of magnitude greater than the probability that it produces a daughter in the next stage.

C. Third Assumption: As soon as a fully malignant stem cell is produced, it begins to proliferate, leading to a detectable tumor some time, w, later. Moolgavkar et al (1988) treat w as a constant, called the latency period for the carcinoma, but note that it could also be modeled as a random variable.

D. Fourth Assumption: The carcinogenic process takes place in the context of normal tissue cell growth and differentiation. This process can be represented by making N0(t), the number of cells in stage 0 (normal stem cells), a deterministic function of time t; or a stochastic model for {N0(t)} can be specified. Moolgavkar et al (1988) consider both possibilities.

E. Fifth Assumption: Transition parameters can be dose-dependent. As pointed out by Thorslund et al, 1987, the six parameters of the MVK model, namely a0, a1, b0, b1, L1, and L2 in our notation, can all be dose-dependent, thus producing a framework suitable for quantitative risk assessment. While the transition parameters in LMS are all interpreted as transformation rates along the path from normalcy to malignancy, analogous to the L1 and L2 parameters in the MVK model,

the MVK model also has transition parameters that can show the effects of dose on stem cell proliferation rates, a0 and a1, and on stem cell death rates, b0 and b1. This substantially increases the expressive power of the MVK model: it can express (and calculate the effects) of phenomena that can not be represented in Armitage-Doll approaches to cancer modeling.

4.2 Qualitative Insights and Contributions from the MVK Model

The structure of the MVK model suggests several qualitative conclusions even in the absence of detailed quantitative modeling. These are summarized in Table 2. The conceptual framework provided by the model provides a powerful tool for thinking about carcinogenesis in general and about the directions and time patterns of effects of different carcinogens. The MVK model also contributes to cancer modeling in the following ways:

o Explanation of carcinogenesis is organized around a few key parameters. The six (possibly time-varying) parameters in Figure 2 provide the basic input data for the model from which cancer hazard rates over time are predicted. Carcinogens affect the probable time to tumor only through these six parameters, although each parameter may be supported by biologically causal models. These parameters provide a parsimonious structure for integrating diverse research efforts and results and for clearly focusing and characterizing the expected benefits of different lines of empirical research.

TABLE 2: QUALITATIVE CONCLUSIONS FROM THE MVK MODEL

1. The hazard rate cancer initiation at time t is $h(t) = N1(t)L2(t)$, i.e., it is the number of premalignant cells at risk of malignant transformation times the transformation rate. Technically, $h(t) = 1 - \exp[-N1(t)L2(t)]$, but for $L2(t)$ small enough so that $N1(t)L2(t)$ is small, this is numerically almost exactly $N1(t)L2(t)$.

2. A chemical carcinogen is a chemical that increases any of the quantities L1, L2, or N1. Thus, a chemical that increases the proliferation rate a1 or decreases the differentiation rate b1, thereby increasing the equilibrium value of N1, will be a carcinogen, even though it may leave L1 and L2 unchanged. A classical "initiator" can be thought of as a chemical that increases L1 or L2, e.g., by genotoxically induced mutations; while a classical "promoter" might increase the proliferation rate a1.

3. A carcinogen that increases L1 produces a permanent increase in the hazard function h(t) among exposed subjects. The effect persists even after exposure ceases because an increase in L1 permanently increases the pool N1 of premalignant cells awaiting malignant transformation. By contrast, a carcinogen that increases L2 leads to a temporary increase in the cancer hazard function during exposure that stops when exposure ceases. (See Figure 2.)

4. Older people, who have accumulated a larger population of premalignant cells N1(t) over the years, will be disproportionately sensitive to carcinogens that increase L2.

5. The effects of low doses of a chemical carcinogen on the cancer induction hazard function depend very much on the specific mechanisms by which the carcinogen acts -- e.g., by increasing L2 vs. increasing a1 or L1.

6. By suitable choice of the six MVK parameters a0,a1,b0,b1,L1, and L2, the empirical age-specific hazard functions for most real carcinomas can be successfully reproduced. Thus, a two-stage model suffices to explain empirical data. Moreover, the time pattern of association between smoking and breast cancer, or between other carcinogens and cancers, can also be explained very successfully using this model.

o Many of the key parameters (or surrogates for them) can potentially be measured directly in the laboratory (or in cellular systems) rather than being estimated statistically from whole-animal bioassay response data. Thus, the MVK model can potentially use the empirical data from "molecular epidemiology," as discussed in Section 5.

o The MVK model suggests an integration of causal biological mechanisms with more traditional statistical approaches to understanding and quantifying cancer risks. The structure of the model suggests the qualitative behavior that a particular carcinogen's age-specific hazard function should exhibit, given some information about its mechanism(s) of action. Predicted qualitative behavior can be useful in selecting among alternative classes of statistical models for application to whole-organ reponse data from traditional animal bioassays. The MVK framework also invites direct measurement or estimation of its parameters through the application of statistical techniques to cell-level laboratory measurements. Thus, it provides a potential unifying framework for the synthesis of internal/causal/biological and external/descriptive/statistical approaches to quantitative risk assessment.

o Useful predictions are made without a simple analytic formula. The simple logical structure of the MVK model shown in Figure 2 does not in general lead to simple, closed-form expressions for the cancer hazard function h(t) representing the expected arrival rate of the first fully malignant cell into the Stage 2 compartment. However, useful quantitative results can be produced using numerical

techniques. An important lesson is that it is not necessary to restrict biological assumptions to those that lead to analytically tractable formulas (e.g., the LMS polynomial) for quantitative cancer risk assessments.

4.3 Limitations and Enhancements for the MVK Framework

The MVK model has unquestionably made a major contribution to understanding chemical carcinogenesis. However, viewed as a tool for quantitative risk assessment, it has limitations. These center around the areas of (i) Speculative assumptions; (ii) Data gaps; (iii) Incomplete description of biological processes; and (iv) Questionable analytic methods. The first two of these have made many regulators hesitate to accept the MVK approach as a basis for quantitative risk assessment. But it seems likely that all four types of limitations can and will be overcome in many practical applications over the next few years. In the rest of this subsection, specific limitations and possible improvements and extensions of the MVK model are identified.

A. First Limitation: Speculative Assumptions

Adherents of the LMS model have sometimes criticized the MVK model on the grounds that it is speculative, unproven, and inadequately tested to support regulatory risk assessments. Since the LMS model itself is speculative (or in many aspects wrong; see Section 3), however, accepting the MVK framework is not a matter of choosing a

speculative framework over a non-speculative alternative. Yet, it is true that the central MVK hypothesis of only two stages, although correct for retinoblastoma and plausible for some other cancers, may not be correct as a general model for cancers (or even for all carcinomas.) The provision for cell proliferation and differentiation (and death) seems unexceptionable: most stem cells undergo these processes. Modeling cancer growth, following the creation of a fully malignant cell, as a constant or random latency period is simplistic, as acknowledged by Moolgavkar et al (1988), but, as they also point out, a more accurate stochastic growth model can easily be substituted for a simple latency period model.

Extending the MVK model to allow for more than two stages makes it general enough to subsume LMS as a special case (by setting the death/differentiation parameters and the proliferation parameters equal to constants or zero.) Thus, an extended, multi-stage MVK framework is strictly better able than the LMS framework to describe biologically realistic cancer processes.

B. Second Limitation: Incomplete Description of Biological Processes

The essential contribution of the MVK framework compared to the Armitage-Doll/LMS framework is its inclusion of some additional aspects of cancer biology -- namely, cell line proliferation, death, and differentiation -- that are important in carcinogenesis and that enter directly into the biologically correct determination of cancer hazard rates. It is possible to carry further this idea of including

biologically important features in a risk model: not all the essential determinants of age-specific cancer hazard functions are represented in the MVK model. Important omissions include the following:

1. Internal dose. Cell transformation rate parameters in the MVK model must by definition depend on internal, biologically effective dose (e.g., quantities of metabolites reacting with DNA or other target macromolecules per unit time) rather than on administered dose. The MVK model therefore requires a "front end" physiologically-based pharmacokinetic (PB-PK) model to convert administered dose of a chemical to internal doses of relevant metabolites (since in general they are not proportional to each other.) This is equally true for LMS and other dose-response models.

2. Cytotoxic effects. Stem cell proliferation and differentiation rates are homeostatically regulated, in part by molecular control signals fed back by more mature cell populations (Metcalf, 1988.) The dynamic response of stem cell population sizes to the cytotoxic (cell-killing and cell-damaging) effects of chemicals can not be modeled without incorporating additional compartments to represent more mature cell populations and different cell lineages. The MVK model must therefore be coupled to a cell kinetics and cytotoxicity model if it is correctly to account for the fluctuating sizes of the cell population in its first compartment. This is particularly an issue for leukemias and other neoplasms of the hematopoietic (blood-forming) system, whose cells can be especially sensitive to cytotoxic effects. The MVK model should also have state-dependent

parameter values, in which the numbers of cells in different compartments affect cell birth and death rate parameters as well as differentiation vs. proliferation probabilities.

3. Transient malignancies. The simple MVK model assumes that the first fully malignant cell eventually creates a tumor (or leukemia) by clonal expansion. In fact, however, it is likely that a malignant cell may lie dormant for some time and then die (or be killed, e.g., by cytotoxic action of continued exposure to the chemical carcinogen, or differentiate into a harmless lineage) before it can proliferate. The simple MVK model must be modified to take into account such spontaneous extinction of malignant cell lines. As noted by Moolgavkar et al (1988), it also must be linked to a back-end model of tumor growth so that its quantitative predictions can be easily compared to animal bioassay and epidemiological data.

4. Heterogeneous transition parameters. Not all cells in the same compartment or stage proliferate at the same expected rate (Webb, 1986.) The proliferative capacity of a cell depends in part on how many generations of successive cell divisions separate it from its original ancestor (the start of the cell line or clone to which it belongs.) The probablity that a cell will divide also varies with the amount of time since its last division, i.e., with its own age or "maturity" within its current compartment. Cell kinetics models have recognized and incorporated this heterogeneity for many years (Eisen, 1979.)

In summary, to continue the move toward more biologically realistic and expressive models that has been started in the MVK modeling literature, it is necessary to extend the original MVK model to incorporate cell kinetics, PB-PK, and tumor growth modeling techniques (Conolly et al, 1989.)

C. Third Limitation: Incomplete Mathematical Analysis

The next limitation of MVK modeling, at least as it has been developed through 1989, is not intrinsic to the model. The analytic work of Moolgavkar et al (Moolgavkar, Dewanji, and Venzon, 1988, 1989) has so far used deterministic differential equations to propagate expected values of random variables. However, the qualitative dynamic behavior predicted by this type of analysis does not correctly reflect the probable behavior of the underlying stochastic system. Despite its power and elegance and the valuable insights that it has produced, such analysis can create artificially high risk estimates because, for example, it does not correctly account for extinction probabilities in small cell populations.

Recall that in Figure 2, the hazard rate at time t for the first entry of a cell into stage 2 (malignant), given that there are $N_1(t)$ premalignant cells in stage 1 at time t, is $h(t) = L_2(t)N_1(t)$. This is the hazard rate when $N_1(t)$ is known. But since $\{N_1(t)\}$ follows a stochastic birth-and-death process, its value $N_1(t)$ for $t > 0$ will not be known at time 0. Assuming that $N_1(t)$ can not be observed, $\{h(t)\}$ will be a randomly evolving hazard function, driven

by the randomly changing process $\{N_1(t)\}$. This is what complicates the analysis. (Braces { } enclose time series.)

To get around this difficulty, Moolgavkar et al (1988, 1989) proceeded roughly as follows. (Their development used a different technical device -- moment generating functions -- that is more rigorous but that leads to the same conclusions.) From applied probability theory, it is known that the expected value of the probability of an event is the conditional probability of the event. A hazard rate is essentially a conditional probability per unit time. Hence, the expected value of h(t), denoted by E[h(t)], should (by this line of reasoning) be identical with h(t) itself, i.e., $h(t) = E[h(t)]$. Taking expectations of random variables on both sides of $h(t) = L_2(t)N_1(t)$ gives the fundamental equation used by Moolgavkar and coworkers, $h(t) = E[h(t)] = L_2(t)E[N_1(t)]$. The remaining challenge is to find $E[N_1(t)]$.

To calculate $E[N_1(t)]$, one can construct a system of ordinary differential equations representing the balance of flows among compartments in a diagram such as Figure 2. The flows relating expected values of $N_1(t)$, $N_2(t)$, and $N_3(t)$ in Figure 2 correspond to the system of differential equations

$$dE[N_0(t)]/dt = [a0(t) - b0(t) - L1(t)]E[N_0(t)]$$

$$dE[N_1(t)]/dt = [a1(t) - b1(t) - L2(t)]E[N_1(t)] + L1(t)E[N_0(t)]$$

$dE[N_2(t)]/dt = L2(t)E[N_1(t)]$; or in vector-matrix notation,

$$dE[N(t)]/dt = A(t)E[N(t)] \text{ where } A(t) = \begin{bmatrix} a0(t)-b0(t)-L1(t) & 0 & 0 \\ L1(t) & a1(t)-b1(t)-L2(t) & 0 \\ 0 & L2(t) & 0 \end{bmatrix}$$

(Moolgavkar et al set the L1(t) term in the first equation and the L2(t) term in the second equation equal to zero, presumably because L1 has a negligible influence on the evolution of cell counts in stage 0, while the number of L2 transitions is 0 if conditioned on $N_2 = 0$.) The solution of this system of equations (or vector differential equation) is a matrix exponential. Moolgavkar et al use this solution as the basis for numerical calculations that can successfully reproduce many observed cancer risk curves.

Critique of the Deterministic Mathematical Analysis

The analysis just presented, although plausible, is mathematically incomplete. The problem is in the basic identity E[h(t)] = h(t), on which everything else depends. This identity does not hold for stochastic hazard functions (Arjas, 1981, Schechner, 1982, Brown and Ross, 1982.) The problem is that at time 0, h(t) is a random variable, whereas E[h(t)] is constant.

In slightly more detail, let I(t) denote the information set (the sigma algebra) with respect to which E[h(t)] is calculated, so that E[h(t)] is an abbreviation of E[h(t);I(t)]. Then E[h(t);I(t)] is not the same as E[h(t);I(0)] [and, technically, is not even measurable with respect to I(0), and so is a random variable with respect to I(0).] In the

terminology of modern stochastic process theory, the expected value propagated through the Moolgavkar et al differential equations are no adapted to their underlying filtration. A correct formulation woul require stochastic differential equations to represent the flows i Figure 2 and special (martingale calculus) techniques to solve the (Schechner, 1982; Brown and Ross, 1982.)

The following simple example demonstrates that the type o mathematical analysis described above can give incorrect prediction for the qualitative behavior of MVK cancer risk hazard functions.

Example: Deterministic Differential Equations Do Not Predic Extinction Probabilities Correctly

Problem Setting: In the MVK flow diagram in Figure 2, set L1 = 0, L2 0.01, $N_1(0) = 4$, $N_2(0) = 0$, a1 = 0.5, and b1 = 0.5. Thus, th transition rates are constant; the inflow from stage 0 to stage 1 i eliminated for convenience of exposition; the stochastic outflow frc stage 1 to stage 2 occurs at an expected rate of one transformation pe hundred cells in stage 1 per unit time (this is completely arbitrary) and each cell in stage 1 is equally likely to produce a daughter cel or to die or differentiate in each unit of time. At time 0, there ar four premalignant cells in stage 1 and 0 malignant cells in stage 2.

Problem: What is the probability that a malignant cell will eventuall be created?

MVK Solution: If the formula $h(t) = E[N_1(t)]L2(t)$ were correct, malignant cell would eventually be created with probability 1. Fc $(N_1(t))$ evolves according to a random walk (with zero as an absorbir boundary) in which the number of premalignant cells is equally likel to increase by 1 as to decrease by 1 in unit of time prior to th creation of a malignant cell. Hence, in the Moolgavkar et al analysis the expected number of premalignant cells remains constant over time at $E[N_1(t)] = 4$ in this example. [In terms of the differentia equations used to propagate expected values, notice that when a1(t) b1(t) and L1(t) = 0, the equation $dE[N_1(t)]/dt = [a1(t) - b1(t)$ $E[N_1(t)] + L1(t)E[N_0(t)]$ used by Moolgavkar et al (Equation 11, p 387 of their 1988 paper) reduces to $dE[N_1(t)]/dt = 0$, implying tha the expected number of cells in stage 1 remains constant.] But ths the hazard rate h(t) must also remain constant over time. A constar hazard rate implies that eventually a malignant cell will be create with probability 1. In this example, the constant hazard rate would l $h(t) = E[N_1(t)]L2(t) = 4(0.01) = 0.04$. The time until a malignar cell is created therefore has an exponential distribution with a expected value of 1/(0.04) = 25 periods.

Correct Solution: Since the number of premalignant cells, $(N_1(t))$, essentially follows a symmetric random walk stochastic process with an absorbing boundary at $N_1 = 0$, it follows that with probability 1, there will eventually be zero cells left in stage 1, i.e., there will come a "stopping time," say T. for which $N_1(T) = 0$ for the first time. (See e.g., Ross, 1983, p. 245, Example 7.6(b) for a proof.) Less formally, the initial premalignant clone will become extinct with probability 1, given the choices of parameter values we have made. Thus, it has only a finite number of periods in which to generate a fully malignant cell. If the premalignant clone becomes extinct before a malignant cell has been generated, then no malignant cell will ever occur. Thus, the correct probability that a malignant cell will ever be generated is less than 1. The reason is that the hazard function $(h(t);I(t))$ is not constant, but evolves randomly according to a stochastic process that eventually crosses the line $h(t) = 0$ and stops. The expected number of premalignant cells $E[N_1(t);I(0)]$ does not stay constant, as the Moolgavkar et al differential equations prescribe, but decreases with t as the probability that extinction has occurred, $Pr[T \leq t;I(0)]$, increases.

In summary, Moolgavkar et al (1988, 1989), although making a major contribution to cancer risk modeling by incorporating biological phenomena into a parsimonious stochastic compartmental flow model, do not analyze the resulting model fully because $(h(t))$ is a stochastic hazard function.

An Alternative Analytic Strategy: Stochastic Simulation

The exact mathematical analysis of randomly evolving hazard functions with state-dependent transition parameters can be very difficult. Fortunately, for the purposes of practical numerical calculations, an alternative approach is available. Stochastic simulation of the randomly evolving system can be used to obtain accurate numerical results even when the analytic formulation is too hard to solve (Yakovlev and Zoren, 1988.)

The basic ideas of stochastic simulation for extended MVK models are described in Cox (1989.) By use of appropriate simulation techniques such as uniformization (Ross, 1983) it is possible to construct a discrete-time stochastic simulation of the MVK model (or more general multicompartment birth-death processes with state-dependent transition parameters) whose probabilistic behavior obeys exactly the same probability laws as the continuous-time system being simulated. Analysis of the MVK model via stochastic simulation is preferable to the type of mathematical analysis begun by Moolgavkar et al insofar as it can produce accurate results without making the simplifying approximations usually required to get tractable mathematical results. On the other hand, stochastic simulation often requires many runs to build up enough sample paths to obtain accurate numerical risk estimates.

5. New Uses of Biological Data in Quantitative Risk Assessments

Both the LMS framework and the MVK framework require model parameters to be estimated from empirical data before predictions can be made. Historically, quantitative risk assessments have been based exclusively on two types of data: animal bioassays and human epidemiological studies. Other potential sources of relevant data include

o Structure-activity relations (SAR's), indicating that a chemical has a structure associated with carcinogenic activity (United States Dept. of Health and Human Services, 1986.) [A recent extension of this idea is to examine metabolites for structural similarities as possible sites

for enzyme activity leading to carinogenesis (M. Bird, 1989, personal communication.)]

o In vitro tests in microorganisms such as the Ames test for mutagenicity in Salmonella typhimurium (Margolin et al, 1989); and

o Biological markers and "molecular epidemiology" assays, such as the formation of covalent adducts in purified DNA exposed to a chemical (Hattis, 1986, Perera, 1986.)

These sources have not been used in (or generally considered appropriate for) quantitative cancer risk assessment, but have been used in qualitative hazard identification (United States Dept. of Health and Human Services, 1986.)

During the 1980's, increasing enthusiasm has been expressed for the idea of using such biological data as part of quantitative risk assessment. Instead of basing parameter estimates and risk estimates entirely on statistical extrapolations from whole-animal data, several proposals have been made to use laboratory measurements of biological processes inside the organism (and interpolations based on them) as the primary relevant data for quantitative risk assessment (Hattis, 1986, Perera, 1986.)

Measurement-based approaches to quantitative risk assessment now fall into two broad categories, as follows.

o Physiologically-based pharmacokinetic (PB-PK) approaches draw on metabolic rate parameters, compartment sizes, biochemical reaction rates, and chemical equilibrium partition coefficients measured in the laboratory to predict the internal doses of metabolites at target sites that will be generated by an external (administered) dose (Reitz, 1989; Conolly, 1989.)

o "Molecular epidemiology" approaches (including the use of biological markers) attempt to directly measure internal doses, or surrogates for internal doses, by laboratory techniques (Perera, 1986; Warren and Beck, 1988; Dragsted, 1989.)

Both approaches offer potential advantages over whole-organism data and extrapolations for quantifying risks. However, so far, only the PB-PK approach has started to prove its value in practical applications by providing a new, useful framework for making interspecies dose conversions and risk extrapolations. Both approaches will now be described.

5.1 PB-PK Modeling for Interspecies Dose Conversion

In brief, a compartmental flow model is used to describe how an inhaled, ingested, or injected dose of a chemical carcinogen is distributed to the relevant compartments (e.g., fatty tissue, liver, bone marrow, etc.) within the body by the blood stream, and how it is transformed into successive metabolites. A system of ordinary differential equations describing flows across compartment boundaries is then set up and solved numerically.

The most relevant point about PB-PK modeling is that the biochemical reaction constants and other data that it requires can in many cases be measured experimentally and applied successfully to predict internal doses in various species (Reitz et al, 1989, Travis, 1989.) It thus provides a constructive approach to incorporating some relevant biological measurements into quantitative risk estimates. PB-PK modeling is perhaps most useful for predicting internal dose patterns of metabolites in humans from observed internal dose patterns in animals and from knowledge of the relevant metabolic pathways involved. However, it does not explain which metabolites are involved in cancer induction or model the cancer induction, promotion, or progression processes themselves. So PB-PK modeling is a useful component in cancer modeling. It must be combined with other modeling techniques to form a complete description and predictive framework for cancer risk modeling (Conolly et al, 1989.) At present, the PB-PK approach has proved itself well enough to start winning some limited recognition and acceptance among members of national and international research and regulatory communities (Travis, 1989.)

5.2 Molecular Epidemiology

"Molecular epidemiology" uses laboratory techniques to study the propagation of effects at the biochemical and molecular levels. The strengths, weaknesses, and uses of molecular epidemiology techniques coincide with those of biological markers. Some representative molecular epidemiology assays for carcinogenicity are listed in Table 3 (Warren and Beck, 1988.)

TABLE 3: SOME REPRESENTATIVE MOLECULAR EPIDEMIOLOGICAL ASSAYS FOR MEASURING CARCINOGENIC ACTIVITY (Warren and Beck, 1988)

1. Measurement of covalent binding of a chemical to purified DNA *in vitro*.

2. Alkylation of other cellular macromolecules, including RNA and especially amino acids in cellular proteins.*

3. Alkylation of other, more accessible proteins, such as hemoglobin. These proteins may (with questionable validity) be considered proxies for the relevant cellular macromolecules.

4. Measurement of chemical-DNA adducts (e.g., by monoclonal antibodies or competetive radioimmunoassay.)

5. Measurement of sister chromatid exchanges (SCE's) in lymphocytes

6. Measurement of mutation for 6-thioguanine resistance in lymphocytes

7. Measurement of chromosomal aberrations in lymphocytes

8. Sperm morphology tests

9. Induction of dark basal cells in skin (as assay for tumor promoters)

* Note: As explained by Warren and Beck (1988), the rationale for this assay is based on the classic Miller and Miller theory of genotoxic carcinogenicity in which carcinogens, which tend to be strongly electrophilic, act on nucleophilic cellular macromolecules including DNA to create point mutations, chromosomal translocations and aberrations, and other forms of genetic damage.

All of the techniques in Table 3 have been used in hazard identification for chemical carcinogens. For example, several of them were applied to workers occupationally exposed to benzene, and the resulting evidence was considered by EPA and OSHA in revising benzene standards.

The essentially new idea proposed by advocates of molecular epidemiology for quantitative risk assessments is that instead of being used only to indicate carcinogenic potential, these same techniques can also be used to help quantify the strengths of carcinogenic responses at the cellular and biochemical levels. Hattis (1986) and Perera (1986) discuss opportunities to improve quantitative risk analysis by use of such methods. Important potential benefits include

1. The ability to measure and quantify individual susceptibilities to carcinogens. This offers a chance to study quantitatively the variability in human susceptibility to carcinogens within human populations. (Variabilities spanning two to five orders of magnitude are not uncommon in the study of drug reactions.)

2. Improved interspecies dose-conversion based on appropriate adjustment of biological parameter values, rather than on more or less ad hoc extrapolations from body weight or surface area. This advantage is already starting to be realized to some extant through PB-PK modeling.

3. Improved dose reconstruction in cases where biological effects (e.g., certain chromosomal aberrations in lymphocyte cell lines) persist for years after direct exposure to a chemical has ceased.

4. Direct measurement of biologically effective doses (or surrogates for them) in animal bioassay experiments.

5. Early warnings of carcinogenic responses.

Achieving any of these benefits would be significant. However, techniques of molecular epidemiology are still being developed and validated: they are still generally too new for practical use in quantitative risk assessment. Moreover, there are obstacles to trying to use these techniques quantitatively. For example, the fact that a chemical forms covalent adducts with purified DNA in a test tube is no guarantee that it will do so in vivo, where the cellular DNA may be far less accessible.

Most practical molecular epidemiological techniques today must settle for proxy measurements that can be observed in place of the quantities of real interest -- for example, binding of a chemical to hemoglobin as a proxy for its binding to target cellular macromolecules (which may be unknown.) Sound inference then depends on the strength of the correlation between the proxy measurement on variables and the true values of the variables. In many cases, this correlation provides only weak predictive power. For example, Warren and Beck (1988) note that the level of DNA alkylation by methyl bromide predicted on the basis of its alkylation of hemoglobin in vitro is far greater than the level actually observed.

In summary, it appears that as long as knowledge of the causal

paths leading from exposures to a chemical to cancer induction has major gaps, molecular epidemiology techniques that use correlation as a proxy for causation will be of limited use for quantitative risk assessment. On the other hand, the techniques can contribute qualitatively to some of the benefit categories just described. Moreover, as causal paths are filled in -- for example, by using PB-PK modeling to predict the relation between a chemical's measured concentration in a bio- chemical pathway and the concentrations of its metabolites -- these techniques can be expected to become much more powerful aids to quantitative risk assessment.

6. Conclusion: What Lies Ahead for Chemical Carcinogen Risk Assessment?

The previous sections suggest that the field of quantitative cancer risk assessment is undergoing a crucial historical transition. What will be left behind is a statistical analysis framework, crystalized in the EPA's LMS methodology, that is characterized by

- o Static, closed-form algebraic models of the dose-response relation;
- o Whole-animal, species-specific response data;
- o Statistical estimation of model parameters (typically by maximum-likelihood estimation applied to animal bioassay or epidemiological data);
- o A motivating biological rationale based on incorrect and incomplete simplifying assumptions for a single (genotoxic) mechanism of cancer induction;
- o Ignored biological processes, including metabolism, cell kinetics, and cell line proliferation, differentiation, and death, that are essential to describing the process of carcinogenesis;

o Statistical extrapolation of observed responses beyond the range of observed conditions eliciting them.

What will replace this status quo, if current trends continue to a successful conclusion, will be a new set of biologically based risk models characterized by

o Analytically intractable, biologically more realistic computer models of the dose-response relationship and the mechanisms behind it (to the extent that these are known);

o Within-organism biological process data supplemented by data on between-organism variability in important process parameters;

o Empirical measurement of model parameters and functions (typically by laboratory techniques of molecular biology);

o A motivating biological rationale that includes PB-PK, cell kinetics, cytotoxic, stochastic induction, and tumor progression relationships and their interactions;

o Statistical interpolation of measured parameters within the range of conditions for which measurements have been made.

The new theme is that by incorporating more cancer biology and empirical biological measurements into risk modeling, better risk estimates can be achieved (Travis, 1989; Wilson, 1988.) Technical development of laboratory and biomathematical risk modeling techniques that will supply the tools needed to convert this vision into reality is already well underway. The next substantial advances in biologically-based cancer risk modeling can be expected in the area of computer models that start to better integrate the PB-PK, cell kinetics modeling, and MVK-type cancer induction models, and molecular biological data bases that already exist.

REFERENCES

Andersen, E.L., and members of CAG, "Quantitative approaches in use to assess cancer risks," Risk Analysis, 3, 4, 1983, 277-295.

Arjas, A., "The failure and hazard processes in multivariate reliability systems," Mathematics of Operations Research, 6, 2, 1981, 263-276.

Armitage, P. and R. Doll, "Stochastic models for carcinogenesis," in Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Vol. 4, University of California Press, Berkeley, CA, 1961.

Bailar, J.C. et al, "One-hit models of carcinogenesis: Conservative or not?," Risk Analysis, 8, 4, 1988, 485-497.

Brown, C.C., and K.C. Chu, "Additive and multiplicative models and multistage carcinogenesis theory," Risk Analysis, 9, 1, 1989, 99-105.

Brown, M., and S.M. Ross, "The observed hazard and multicomponent systems," Naval Logistics Research Quarterly, 29, 4, 1982, 679-683.

Chappell, W.R., "Interspecific scaling of toxicity data: A question of interpretation," Letter to the Editor, Risk Analysis, 9, 1, 1989, 13-14.

Conolly, R.B., et al, Computer simulation of chemical carcinogenesis, in C.C. Travis (Ed.), Biologically Based Methods for Cancer Risk Assessment, Plenum Press, New York, 1989.

Cox, L.A., New Directions in Cancer Modeling for Benzene, Draft Final Report to the American Petroleum Institute and the Western States Petroleum Association, August, 1989.

Cronkite, E., "Chemical leukemogenesis: Benzene as a model," Seminars in Hematology, 24, 1, 1987, 2-11.

Crump, K.S., "Mechanisms leading to dose-response models," in P.F. Ricci (ed), Principles of Health Risk Assessment, Prentice-Hall, Englewood Cliffs, N.J., 1984.

Crump, K.S., and R.B. Howe, "The multistage model with a time-dependent dose pattern: Applications to carcinogenic risk assessment," Risk Analysis, 4, 3, 1984.

Crump, K.S., and W.W. Watson, GLOBAL79: A FORTRAN Program to Extrapolate Dichotomous Animal Carcinogenicity Data to Low Doses, National Institute of Environmental Health Sciences, Contract No. 1-ES-2123, 1979.

M. Castagna and I. Martelly, "Overview of promotion as a mechanism in carcinogenesis," in C. Travis (ed), Biologically Based Methods for Cancer Risk Assessment, Plenum Press, New York, 1989.

Dragsted, L.O., "Determination of carcinogen exposure by immunological techniques," in C.C. Travis (Ed.), Biologically Based Methods for Cancer Risk Assessment, Plenum Press, New York, 1989.

Eastmond, D.A., M.T. Smith, and R.D. Irons, "An interaction of benzene metabolites reproduces the myelotoxicity observed with benzene exposure," Toxicology and Applied Pharmacology, 91, 1, 1987, 85-89.

Eisen, M., Mathematical Models in Cell Biology and Cancer Chemotherapy, Springer-Verlag, New York, 1979.

Hattis, D.B., "The promise of molecular epidemiology for quantitative risk assessment," Risk Analysis, 6, 2, 1986, 181-194.

Margolin, B.H., "The Ames Salmonella/microsome mutagenicity assay: issues of inference and validation," J. American Statistical Association, 84, 407, 1989, 651-661.

Mehlman, M.A. (Ed.), Benzene: Occupational and Environmental Hazards -- Scientific Update, Princeton Scientific Publishing Co., Princeton, N.J., 1989.

Moolgavkar, S., and A. Dewanji, "Biologically based models for cancer risk assessment: A cautionary note," Letter to the Editor, Risk Analysis, 8, 1, 1988, 5-6.

Moolgavkar, S.H., A. Dewanji, and D.J. Venzon, "A stochastic two-stage model for cancer risk assessment. I. The hazard function and the probability of tumor," Risk Analysis, 8, 3, 1988, 383-392.

-- "A stochastic two-stage model for cancer risk assessment. II. The number and size of premalignant clones," Risk Analysis, 9, 2, 1989, 179-187.

Murdoch, D.J., and D. Krewski, "Carcinogenic risk assessment with time-dependent exposure patterns," Risk Analysis, 8, 4, 1988, 521-530.

Murdoch, D.J., Krewski, D.R., and K.S. Crump, "Quantitative theories of carcinogenesis," Chapter 2 in J.R. Thompson and B.W. Brown (Eds.), Cancer Modeling, Marcel Dekker, 1987.

Perera, F., "New approaches in risk assessment for carcinogens," Risk Analysis, 6, 2, 1986, 195-202.

Ross, S.M., Stochastic Processes, Wiley, New York, 1983.

Schechner, Z., "A note on randomly evolving hazard rate functions," Naval Logistics Research Quarterly, 29, 4, 1982, 693-695.

Sielken, R.L., "The capabilities, sensitivity, pitfalls, and future of quantitative risk assessment," in R.S. McColl (Ed.), Environmental Health Risks: Assessment and Management, University of Waterloo Press, Waterloo, Ontario, Canada, 1987.

Thomas, F.B., "An Emerging View of Benzene Hematotoxicity and Leukemogenesis: Comments of the American Petroleum Institute," API, Washington, D.C., January, 1989.

Thorslund, T.W., C.C. Brown, and G. Charnley, "Biologically motivated cancer risk models," Risk Analysis, 7, 1, 1987, 109-120.

Thorslund et al, Quantitative Re-evaluation of the Human Leukemia Risk Associated with Inhalation Exposure to Benzene, Final Report, Clement Associates, Inc., Fairfax, Va., October, 1988.

Travis, C.C. (Ed.), Biologically Based Methods for Cancer Risk Assessment, Plenum Press, New York, 1989.

Travis, C.C., and R.K. White, "Interspecific scaling of toxicity data," Risk Analysis, 8, 1988, 119-125. (See also comment by Chappell, 1989.)

United States Department of Health and Human Services, Determining Risks to Health: Federal Policy and Practice, Auburn House Publishing Company, Dover, MA, 1986.

Warren, A.J., and B.D. Beck, "Screening and monitoring for exposure and susceptibility to carcinogens," Chapter 17 in J.D. Brain et al (Eds.), Variations in Susceptibility to Inhaled Pollutants: Identification, Mechanisms, and Policy Implications, Johns Hopkins University Press, Baltimore, Md., 1988.

Webb, G.F., "A model of proliferating cell populations with inherited cycle length," J. Math. Biol., 23, 1986.

Weinberg, R.A., "Finding the anti-oncogene," Scientific American, 259, 3, 1988, 44-53.

Whittemore, A.S., and J.B. Keller, "Quantitative theories of carcinogenesis," SIAM Review, 20, 1978, 1-30.

Wilson, J.D., "Time for a change," Guest Editorial, Risk Analysis, 6, 2, 1988, 44-53.

Yakovlev, A.Y., and A.V. Zorin, Computer Simulation in Cell Radiobiology, Springer-Verlag, New York, 1988.