

Review

From exposure to effect: a comparison of modeling approaches to chemical carcinogenesis

Ingeborg M.M. van Leeuwen*, Cor Zonneveld

*Department of Theoretical Biology, Faculty of Biology, Institute of Ecological Science, Vrije Universiteit,
De Boelelaan 1087, 1081 HV Amsterdam, The Netherlands*

Received 8 January 2001; received in revised form 24 April 2001; accepted 2 May 2001

Abstract

Standardized long-term carcinogenicity tests aim to reveal the relationship between exposure to a chemical and occurrence of a carcinogenic response. The analysis of such tests may be facilitated by the use of mathematical models. To what extent current models actually achieve this purpose is difficult to evaluate. Various aspects of chemically induced carcinogenesis are treated by different modeling approaches, which proceed very much in isolation of each other. With this paper we aim to provide for the non-mathematician a comprehensive and critical overview of models dealing with processes involved in chemical carcinogenesis. We cover the entire process of carcinogenesis, from exposure to effect. We succinctly summarize the biology underlying the models and emphasize the relationship between model assumptions and model formulations. The use of mathematics is restricted as far as possible with some additional information relegated to boxes. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Quantitative cancer risk assessment; Tumor incidence; Chemical carcinogen; Mathematical model; Dose–response relationship

1. Introduction

Despite improvements in prevention, diagnosis and treatment, cancer still strikes one in three people, and one in four will eventually die of the disease [1]. This high incidence is primarily due to two risk factors, namely cigarette smoking [2,3] and dietary habits [4]. As these concern lifestyle aspects which are in principle subject to personal choice, they can be avoided at the individual level. This does not hold for other risk factors, such as exposure to occupational or environmental agents. These factors constitute an

unintentional risk that can only be avoided through decisions made at the community level.

Since Pott conducted his historical study on chimney sweeps and identified soot as a carcinogenic agent [5], extensive testimony has accumulated with regard to the causal relationship between cancer incidence and exposure to chemical compounds. Evidence that chemicals may cause cancer has come, for instance, from experimental tests and epidemiological studies. Nowadays, the carcinogenic effect of a compound is crucial to restrictions on either its production or its emission into the environment. The success of these restrictions might explain the assertion by Ames et al. that food additives and industrial chemicals have had little impact on the overall cancer incidence to date [6].

Primary prevention of cancer can be accomplished by reducing the number of carcinogens to which

* Corresponding author. Tel.: +31-204446959;
fax: +31-204446959.
E-mail address: ingeborg@bio.vu.nl (I.M.M. van Leeuwen).

humans are exposed, and by reducing the levels of exposure to carcinogens [7]. Clearly, these approaches rely on accurate identification of carcinogens and on reliable quantification of their ability to elicit a carcinogenic response. Both hinge on the adequacy of the tests devised to estimate the carcinogenic potency of chemicals, and on the validity of the conclusions derived from them. These issues are far from settled as the often fierce debates about them illustrate. For two different reasons, society needs a strategy to accurately evaluate the implications of carcinogenicity tests for humans. From the perspective of human health, such a strategy is essential to rule out risk-underestimates. From an economic point of view, the avoidance of risk-overestimates averts costs associated with unnecessary reduction of exposure levels.

The various tests devised to estimate carcinogenic potency differ in many aspects, but they share a common feature: the analysis of their results is facilitated by the establishment of a functional relationship between dose and response [8]. The plethora of mathematical models used for this purpose may at times be disconcerting to experimental scientists. Some modelers have attempted to conciliate experimentalists with the models. Hanes and Wedel, for example, purport to “remove for the non-mathematician some of the mystery as to the derivation of the formulas” [9]. Even if they accomplished this aim, they still only consider a small subset of models that are currently in vogue. Other reviews are subject to similar restrictions in scope. For instance, van Ryzin only deals with dose–response models for risk assessment [10], while Kopp-Schneider focuses on multi-step models of tumor induction [11]. None of these works covers the entire area from exposure to effect. The most thorough overview of models involved in the estimation of human cancer risk is by Moolgavkar et al. [12]. Due to its in-depth treatment of the topic, this book is less suited as a gentle introduction to the field. In summary, we think that a general but succinct overview for the non-mathematician is still lacking. The aim of the present work is to provide such an overview.

The remainder of this paper is organized as follows. Section 2 provides a brief description of the biology of chemical carcinogenesis as a four-phase process. We realize that the informed reader will miss many important findings. Their omission derives from the fact that these new findings are not yet used by modelers;

we only present a skeletal outline of chemical carcinogenesis relevant for our overview. Section 3 discusses a representative selection of mathematical models for chemical carcinogenesis, organized according to the phases defined in Section 2. In our experience, existing explications of the models are either hard to find or hard to follow. Therefore, we aim to make explicit the relationship between model assumptions and model formulations. Finally, in Section 4, we make some concluding remarks that go beyond the individual models. Although in the paper we focus on the conceptual basis of the various models, we cannot always ignore mathematical formulations. Where these become cumbersome, they are confined to boxes. The reader not interested in the mathematical niceties may skip these boxes without losing continuity.

In this paper we compare some 30 representative modeling approaches. Understandably but regrettably, they do not share a common strategy with respect to their mathematical notation. Different models may therefore use the same symbol for unrelated concepts or variables. To minimize confusion related to the notation, we were forced at times to choose a representation that deviates from the customary. In the few instances where we used the same symbol for different concepts, the context should suffice to disambiguate its interpretation.

2. Chemical carcinogenesis

Theoretical studies on tumor biology are far outnumbered by experimental studies. Yet models still abound; for this review, we studied some 100 papers on a variety of models, as well as some monographs. To keep track of all these studies, a natural first step is to contrive some classification. ECETOC's Monograph 24 classifies models for chemical carcinogenesis loosely on the basis of their underlying statistical assumptions [13]. But as the authors remark themselves, “the division between the models is somewhat arbitrary as there is considerable overlap”. An alternative criterion classifies models as either descriptive or mechanistic. This criterion judges the amount of underlying biology involved. From a mathematical point of view, models can be classified as either deterministic (with a single outcome) or stochastic (with more than one possible outcome). Kopp-Schneider

classifies stochastic models for tumor induction on the basis of their intended use, level of biological detail and method used for their analysis [11]. As the method of analysis has only to do with the degree of mathematical complexity, we do not find this criterion very informative.

As an alternative we organize models according to a division of the process of chemical carcinogenesis, from exposure to effect, into four consecutive phases. We briefly define the phases here, while in the subsections below we provide further details. The first phase, referred to as *kinetics*, concerns the relationship between exposure to a (pro)carcinogen and internal dose of carcinogen. The second phase, referred to as *tumor induction*, comprises the toxico-dynamic mechanisms through which the carcinogen induces the transformation of normal cells into tumor cells. The third phase, referred to as *tumor growth*, relates to the clonal expansion of a tumor. During this phase, the tumor's malignancy may increase (tumor progression). The last phase, referred to as *effects*, involves the consequences of tumor development for the organism.

Not all the models discussed in this paper have been developed with the aim to contribute to a quantitative understanding of chemical carcinogenesis. For instance, most tumor growth models have been developed with the aim to improve cancer treatment, and have not yet been applied within the context of chemical carcinogenesis. Nevertheless, tumor growth models naturally fit the scheme outlined above because it is only after a tumor reaches a certain size that effects will become apparent. Many other models unambiguously address one of the steps defined above. Hence, the classification seems to offer a framework to keep track of the heterogeneous collection of models dealing with aspects of chemical carcinogenesis in its broadest sense.

2.1. Exposure and kinetics

As defined above, kinetics concerns the relationship between the exposure of an animal to a (pro)carcinogen and the internal dose of carcinogen at a target tissue. Here we use 'exposure' in a general sense that includes the exposure conditions. Kinetics constitutes an essential phase in chemical carcinogenesis because it is the tissue dose that is responsible for the carcinogenic response. For foreign chemicals, kinetics

comprises four biological processes, namely uptake, distribution, metabolic transformation, and elimination of the chemical compound.

Uptake consists of two subprocesses: intake and absorption. The result of intake is that the chemical enters either the lung cavity or the gastrointestinal tract. The intake rate depends on the exposure conditions and on the physiology of the animal. For instance, if administration of the chemical is via food, the concentration of the chemical in food and the food ingestion rate determine the intake rate. Intake is bypassed when the chemical is administered directly into the stomach (gavage). Furthermore, intake is absent when the chemical is applied directly on the skin.

Once the chemical is in a body cavity or on the skin, absorption may take place. The result of absorption is that the chemical actually enters the body. Chemicals administered via injections bypass both intake and absorption. There are two major absorption mechanisms, namely passive diffusion and carrier-mediated transport. The latter mechanism is saturable, whereas the former is not. The actual absorption rate, in contrast to the intake rate, depends on physico-chemical properties of the chemical, such as charge and molecular structure [14,15]. As it also depends on characteristics of the tissue involved (e.g. absorption surface area), the uptake route is important for the absorption rate. For instance, absorption from the lungs is usually rapid.

Distribution involves partitioning of the chemical among different body parts. It starts from the site where the chemical enters the body. Chemicals absorbed from the gastrointestinal tract first pass through the liver, whereas chemicals absorbed from the lungs directly enter the bloodstream. Consequently, uptake via the lungs usually results in a quick distribution among major organs [15]. Indeed, the site of entry is important for both the distribution rate and the final disposition of the chemical. Distribution of the chemical out of the bloodstream into the target tissues depends on physico-chemical properties of the chemical (e.g. lipophilicity) and on the relative characteristics of the tissues (e.g. relative size and lipid content). As the characteristics of the tissues vary with age, the same holds for partitioning among body parts [16].

Living organisms have a wide range of enzymatic defense mechanisms against toxic compounds. The enzymes involved usually convert lipophilic chemicals

into more hydrophilic and, therefore, easier excretable metabolites. Environmental and occupational (pro)-carcinogens are also subject to enzymatic transformation [17,18]. Metabolism of a carcinogen can give rise to transformation into non-carcinogenic metabolites (detoxication). The effect of metabolic transformation on the carcinogenic response of chemicals is not always favorable, however. Indeed, metabolism of a (pro)carcinogen can give rise to an active carcinogen (activation).

Several organs, including skin, kidney, and lung, have the ability to transform chemicals. However, the organ that has the largest capacity for metabolic transformation is the liver [15]. As an effect of detoxication in the liver, the carcinogenicity of a particular chemical is often less by oral ingestion than by other uptake routes. As an effect of activation in the liver, oral ingestion of (pro)carcinogens often results in the development of tumors in this organ. Thus, the uptake route influences metabolic transformation and, therefore, carcinogenic response. This influence may even cause a chemical to be carcinogenic only when uptake occurs via a particular route. A clear example is the situation in which the gastrointestinal microflora transforms a chemical into a potent carcinogen [19].

Detoxication of procarcinogens and carcinogens is clearly an important determinant for the internal concentration and, thus, for the carcinogenic response. Another important determinant is elimination of procarcinogens, carcinogens, and their metabolites from the body. The major elimination routes are urinary and biliary excretion. These routes can saturate, leading to accumulation [15].

2.2. Tumor induction

In this subsection we briefly deal with the biological processes underlying the transformation of normal cells into cancer cells. However, we neither attempt to summarize the latest advances in cancer research, nor pretend to deal with all the complexities of the genesis of the disease. Rather, we focus on results used as assumptions in the mathematical models discussed in Section 3.2.

The current view of tumorigenesis, as expressed by Hanahan and Weinberg, postulates that “tumor development proceeds via a process formally analogous to Darwinian evolution, in which a succession of genetic changes... leads to the progressive conversion of normal cells into cancer cells” [20]. This model, first proposed by Nowel [21], is illustrated in Fig. 1 for a tumor that originates from one single normal cell. This is consistent with the observation that most human and animal tumors are monoclonal in composition [21,22]. The genetic changes shared by all tumor cells accumulate along the single lineage preceding the final single founder cell, whereas heterogeneous changes occur during tumor growth [23] (see Fig. 1).

It is the progressive accumulation of multiple genetic changes that underlies the multi-step nature of tumorigenesis [20,24]. As many genes are targets of these changes, cancer is essentially a complex genetic disease. Cellular genes implicated in tumor development are usually referred to as cancer genes. It is possible to distinguish between two classes of cancer genes, those directly controlling cell proliferation (gatekeepers), and those maintaining the integrity of the genome (caretakers) [25]. The former include proto-oncogenes

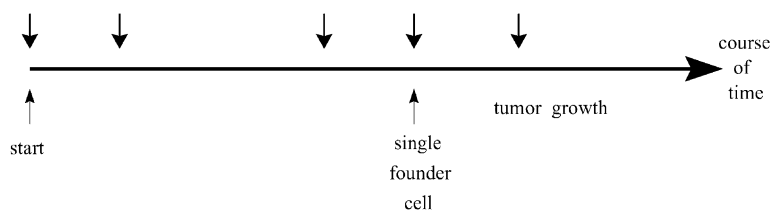


Fig. 1. Tumorigenesis. The start point is the moment at which one normal cell enters the process of tumor evolution. Any down pointing arrow represents the acquisition of one or more new physiological traits, each conferring an additional growth advantage [20]. The single founder cell is the last bottleneck along the evolutionary pathway [23]. The genesis of this cell indicates the beginning of final clonal expansion (tumor growth). In addition to the traits acquired in earlier stages, one or more alterations may occur during tumor growth leading to an increase in malignancy (tumor progression).

and tumor suppressor genes. In contrast to changes in gatekeepers, a change in the normal activity of a caretaker leads to cancer as a secondary effect. A change in activity of a cancer gene may concern either a change in the level of gene expression (e.g. [26,27]) or a disruption of the gene product's biological behavior (e.g. [28,29]). These changes may take place as the ultimate consequence of genetic events such as point mutations, rearrangements or major chromosomal aberrations. Recent studies further indicate that epigenetic events controlling the level of gene expression may play a more important role in tumorigenesis than was previously thought (see [30–32]).

The number of changes required to produce a tumor is specific to a particular tumor type. Indeed, the number of cancer genes involved in tumorigenesis varies from one tumor type to another [33,34]. Moreover, the number of alleles whose activity must change to lead to a phenotypic effect varies from one cancer gene to another. Aberrant genes that act in a recessive manner only have a phenotypic effect when present in the homozygous or hemizygous state, whereas aberrant genes that act in a dominant manner exert a phenotypic effect even when present in the heterozygous state. With regard to the temporal sequence of the changes, on the one hand it has been proposed that the total accumulation of genetic alterations, rather than their relative order, is most important for tumorigenesis (e.g. [22]). On the other hand, there is evidence that the nature and order of genetic changes can have impact on both tumor morphology and the likelihood of tumor progression (e.g. [35]).

2.2.1. Action of chemical carcinogens

Once present in a target tissue, chemical carcinogens can interfere with the process of tumorigenesis at one or more stages. Whichever mechanism is involved, tumor induction implies the interaction of the chemical with one or more cellular components. If the interaction results in DNA damage, the chemical is said to be 'genotoxic'. The potency of a genotoxic compound depends not only on its capacity to cause DNA damage, but also on the rate of cell replication and on the cell's capacity to repair the specific damage inflicted by the chemical compound [36,37]. Non-genotoxic carcinogens are able to act without causing DNA damage [31,38]. For instance, they can induce uncontrolled cell proliferation by altering inter-cellular communication

[39,40]. As a final remark, we notice that a carcinogen can have more than one mode of action [41,42]. For example, a chemical can act as a mutagen at low doses, while on top of this it may be cytotoxic at high doses [43,44].

2.3. Tumor growth

Individual tumor cells are not immortal. Death of tumor cells occurs through the processes of apoptosis or necrosis [45,46]. The latter may take place as a consequence of insufficient supply of nutrients, or as a result of excessive accumulation of metabolic waste products. Survival of chemically induced tumor cells, however, is not only subject to natural death processes: they may also be killed by the immune system of the host organism [47,48]. The existence of cell loss implies that a tumor clone may regress prior to reaching a detectable size (see Fig. 2). Even if they are monoclonal in origin, tumor cells are often heterogeneous with respect to properties such as metabolism, cell division rate, and antigenicity. During tumor growth, tumor cells become heterogeneous as a result of the occurrence of additional genetic alterations (tumor progression). The phenotypic heterogeneity of tumor cells is also the result of significant differences among their local environments [49] which define, for instance, availability of nutrients. A limited supply of nutrients or oxygen commonly occurs in solid tumors, giving rise to a necrotic core [46]. The location of a cell within the tumor may determine its vulnerability to attacks of the immune system. Thus, phenotype

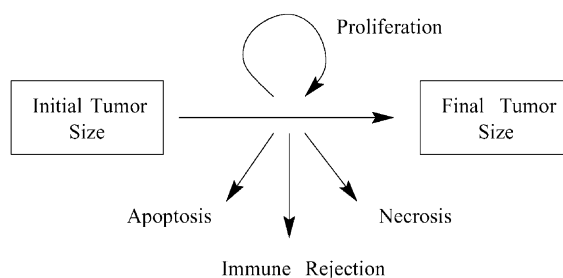


Fig. 2. Tumor growth. Tumor size increases due to cell proliferation. The cell-cycle time and growth fraction determine the rate at which cells are added to the tumor. Tumor size decreases due to cell death and cell killing. The final tumor size may be below the detection limit.

and location of individual cells determine the rates of cell gain and cell loss within the tumor.

For solid tumors it is possible to distinguish between two growth phases. During the initial phase, or avascular growth phase, the tumor cells obtain nutrients and oxygen by diffusion from the surrounding tissue. When the tumor is no longer able to obtain sufficient nutrients by diffusion alone, tumor cells may start to produce several factors to stimulate angiogenesis [50,51]. This defines the beginning of the so-called vascular tumor growth phase. During this second phase, tumor cells obtain oxygen and nutrients from the newly formed tumor blood vessels. After vascularization the tumor may become larger.

Once a tumor reaches a detectable size, its growth may be quantified by measuring tumor size as a function of time. However, only few body sites allow more than one measurement of tumor size [52]. The sparsity of data on human tumor growth is even more accentuated because treatment is rarely withheld [53]. As a consequence, available experimental data mainly relate to growth of tumors *in vitro* (tumor spheroids), growth of tumors inoculated in animal models, or growth of natural tumors *in vivo* during a short period of their growth ontogeny.

Experimental results have revealed that growth of solid tumors *in vivo* is often characterized by a late phase of declining growth rate [52]. The same growth pattern has been observed in studies on tumor spheroids *in vitro* [49]. The growth deceleration has been attributed to several factors such as increase in cell loss, increase in cell-cycle time or decline in the growth fraction [54,55]. The growth fraction, defined as the ratio of proliferating cells to total cells [56], is a concept frequently used to compare tumors in terms of their growth capacity. Another concept used for this purpose is the tumor doubling time, that is the time a tumor needs to double its size. Tumor doubling time is a useful concept if a tumor grows exponentially, because it is then a constant. As soon as exponential growth is no longer realistic, the tumor doubling time becomes less informative.

2.4. Effects

In many carcinogenicity tests the time to tumor onset is not observable. The presence of a tumor can only be detected after the death (or sacrifice) of the

animal. Thus, though the goal of carcinogenicity studies is to evaluate tumor incidence, one has to confront the topic of tumor lethality. How a tumor causes a decrease in survival is not always clear. The size of the tumor is relevant, but not the cause in itself. Rather, when a tumor reaches a certain size it may either impair the normal function of the host organ [57], or exhaust the organism due to its unrestrained use of resources. Moreover, the probability of invasion and metastasis, which are the most life-threatening aspects of tumor progression [58,59], increases with the size of the primary tumor [60]. The detrimental effect of a tumor thus depends on its size, but it may also depend on its location in the body. For instance, it is unlikely that a brain tumor causes death through attrition.

3. Overview of existing models

To begin with, we want to distinguish between a conceptual model, a mathematical model and a mathematical description. We view a conceptual model as a set of assumptions regarding a certain phenomenon. If the conceptual model is translated into equations, it becomes a mathematical model. The mathematical model thus comprises both the mathematical description and the underlying assumptions (i.e. the conceptual model). Note that this accounts for the possibility that two different mathematical models share the same mathematical description. A good example of a mathematical description that appears in different contexts is the Weibull equation [61]:

$$f(y) = 1 - e^{-vy^r} \quad (1)$$

where y is some variable. Sometimes $f(y)$ describes the fraction of tumor bearing animals as a function of dose (e.g. Eq. (4)), sometimes it describes the fraction of tumor bearing animals as a function of time (e.g. Eq. (5)), i.e. the interpretation of the variable y differs. The associated conceptual models are clearly different as they concern different phenomena. More confusing is the situation in which two models for the same phenomenon result in an analogous mathematical description. Despite their outward similarity, we consider such models to be different, because they differ in their underlying assumptions. Naturally, such models are indistinguishable when fitted to experimental data.

3.1. Kinetic models

Although an organism is exposed to a certain environmental concentration of a (pro)carcinogen, the relevant concentration for tumor induction is at the target site. This implies that we have to relate the external concentration to the internal. Kinetic models deal with this problem.

Any kinetic model consists of a set of mass balance equations, each equation describing the change in the amount of chemical in a body ‘compartment’. A compartment does not necessarily correspond to an organ. For example, the most simple kinetic model treats the entire body as one compartment. The models assume that the chemical is well mixed within each compartment, so it makes sense to define the concentration of the chemical in each compartment. The concentration of the chemical in compartment i is $C_i = Q_i/V_i$, where Q_i and V_i denote the mass of the chemical and the volume of that compartment, respectively. A general mass balance equation for the change in Q_i is

$$Q'_i = \text{flux}_{\text{in}} + \text{flux}_{\text{ma}} - \text{flux}_{\text{out}} - \text{flux}_{\text{md}} \quad (2)$$

where flux_{ma} stands for production flux (metabolic activation) and flux_{md} for the metabolic detoxication flux. The actual expressions for the fluxes depend on the specific choice for the kinetic model. In steady state the total positive flux equals the total negative flux, and the mass of the chemical in the compartment is constant.

To give some flavor of kinetic models, we here briefly illustrate the linear one-compartment model that treats the whole body as one compartment. If no metabolic transformation takes place, the amount of chemical in the body is determined by the uptake and elimination processes only. The model assumes that both uptake and elimination follow simple linear kinetics, or equivalently, it assumes that flux_{in} is proportional to external concentration, while flux_{out} is proportional to internal concentration. Since the physics of transport suggests that flux_{in} and flux_{out} are also proportional to the areas of the surfaces involved in absorption and excretion [62], Eq. (2) yields

$$Q' = \text{flux}_{\text{in}} - \text{flux}_{\text{out}} = \delta_u A_u d - \delta_\eta A_\eta C$$

where d represents the external concentration, C the internal concentration, and A_u and A_η are the effective surface areas for absorption and excretion, respectively. The interpretation of the proportionality constants δ_u and δ_η depends on the uptake and elimination routes and on the transport mechanisms.

Body growth can substantially affect the kinetics of a chemical and, thus, its internal concentration. Indeed, if the organism’s size (V) is not constant, the effective surface areas A_u and A_η are also functions of time. Moreover, due to the increase in size dilution of the chemical occurs. In mathematical terms this means that the relation $C'(t) = Q'(t)/V(t)$ does not hold. For a discussion on a one-compartment model that accounts for body growth, see [62,63]. We here focus on the simple situation where (i) the organism does not grow, (ii) the external concentration is constant, and (iii) the initial internal concentration is zero. Deviation from these conditions complicates the mathematical expressions somewhat. These complications are beyond the aim of our presentation.

If the organism’s body size remains constant, the relation $C'(t) = Q'(t)/V$ holds and the effective surface areas A_u and A_η are constant. The mass balance equation above can then be rewritten as $C' = ud - \eta C$, with $u = \delta_u A_u/V$ and $\eta = \delta_\eta A_\eta/V$ the (constant) uptake and elimination coefficients, respectively. The solution of this linear differential equation is $C(t) = (u/\eta)(1 - e^{-\eta t})d$, which satisfies $C(0) = 0$. The equation gives a saturating curve when internal concentration is plotted against exposure time (see Fig. 3).

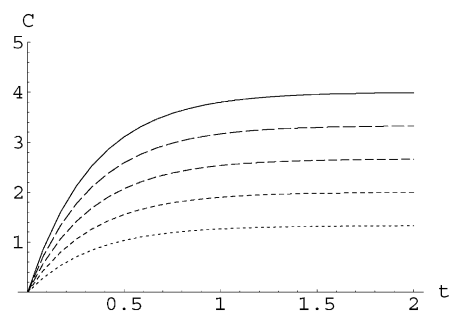


Fig. 3. One-compartment model. Internal concentration as a function of exposure time ($u = 1$ and $\eta = 3$). From top downwards d equals 12, 10, 8, 6, and 4, respectively. For each curve, the asymptotic maximum internal concentration is given by $C_{\text{max}} = d/3$.

After some time the term $e^{-\eta t}$ dies out, and the internal concentration becomes proportional to the external concentration with proportionality coefficient (u/η). In ecotoxicology this ratio is usually called bioconcentration factor [14,63,64]. The steady-state proportionality between external and internal concentration is a generic characteristic of linear compartment models. This property breaks down, for example, if metabolic transformation follows the more realistic nonlinear Michaelis-Menten kinetics.

If any of the assumptions that underlie the one-compartment model does not hold, a multi-compartment model can be used. Two main approaches have been pursued in developing multi-compartment kinetic models, namely data-based compartmental modeling and physiologically-based compartmental modeling (PBPK, where PK stands for pharmacokinetics) [65,66]. The former includes empirical models, whose compartments often lack a biological interpretation. The latter includes biologically-based models, whose compartments correspond more closely to anatomical structures. Indeed, a compartment comprises a single organ, or a group of organs that share relevant physiological features.

Most PBPK models define a central (blood) compartment that is responsible for the distribution of the chemical (e.g. [67]). The amount of chemical entering (or leaving) a compartment via the circulatory system depends on the concentration in the blood and in the compartment, and on the solubility of the chemical in the blood and in the compartment. If distribution among organs is fast in comparison with uptake and elimination, blood flows can be omitted from the model [68]. The alternative is that blood flows function as model parameters.

Even with a small number of compartments, PBPK models require a substantial number of parameters [65]. These include physiological parameters such as blood flows, pulmonary ventilation, and organ volumes, as well as biochemical and physico-chemical parameters such as partition coefficients, tissue clearances, and the rates of metabolism [65,69]. As they have a biological interpretation, most of them can be directly measured by experimental techniques. The remaining parameters have to be estimated. In empirical models all the parameters, as they lack a biological interpretation, have to be estimated from experimental data.

3.2. Tumor induction models

The models we presented so far were all deterministic. From here on they are either deterministic or stochastic. A deterministic model yields a single outcome, whereas a stochastic model yields multiple outcomes and assigns a probability to each of the different outcomes. Before going into the description of the models we briefly introduce a few basic concepts that crop-up in most of the stochastic models. Among these concepts are cumulative distribution function, survivor function and hazard rate. To introduce these concepts, we consider a relevant example.

Let T denote a variable representing the ‘time to first tumor’. The random variable T , which may adopt any positive value, is exhaustively characterized by its cumulative distribution function $F_T(t) = \text{prob}\{T \leq t\}$. This expression reads ‘the probability that the time to first tumor is less than or equal to t ’. Ignoring mathematical exactness, this amounts to a prediction of the fraction of tumor bearing animals at time t . Closely related to F_T is the survivor function $G_T(t) = 1 - F_T(t) = \text{prob}\{T > t\}$ that provides ‘the probability that an individual is tumor free at time t ’. Finally, let h_T denote the hazard rate. Intuitively, the hazard rate concerns the probability per unit time that a tumor develops in a individual of age t , given that the individual is still tumor free. Mathematically, the hazard rate relates to the survivor function as follows:

$$G_T(t) = e^{-\int_0^t h_T(s) ds} \quad (3)$$

Thus, if the hazard rate is known, the survivor function and the cumulative distribution function $F_T(t) = 1 - G_T(t)$ are also known, and vice versa. Hence, the hazard rate constitutes an alternative way to exhaustively characterize a random variable. Most of the stochastic models described below provide expressions for the hazard rate.

The cumulative distribution function, the survivor function, and the hazard rate are denoted in the example above as F_T , G_T , and h_T , respectively, where the subscript indicates the random variable. The same concepts can be defined in a more general sense for any random variable X (for further details see, for example, Cox and Oakes [70]). For example, in the next section we will use F_U , where U is a random variable with the same dimension as the external dose. Finally,

we notice that the biological and mathematical interpretations of survival only coincide if the random variable represents ‘time to death of an individual’.

3.2.1. Tolerance distribution models

The models we present in this subsection are often motivated by the concept of tolerance distribution [71]. To introduce this concept, let us consider a group of mice that have been exposed to a chemical for a particular period of time. Any tolerance distribution model treats the group of mice as heterogeneous with respect to their susceptibility to the chemical: each individual has a different threshold-dose below which no response occurs. No hypothesis about possible mechanisms underlies such a threshold. The models treat the ‘threshold-dose of an individual’ (or tolerance, for short) as a random variable, say U .

For the given exposure time, let $P(d)$ denote the probability that an individual responds to a dose d (i.e. $P(d)$ amounts to a prediction of the fraction of tumor bearing animals). If the group of mice have been exposed to a dose d , only the animals with threshold dose below d will respond. Thus, the probability that an individual responds is $\text{prob}\{U \leq d\} = F_U(d)$, i.e. in this context $P(d) = F_U(d)$. The actual expression for F_U depends on the distribution of U , the so-called tolerance distribution.

Any continuous statistical distribution can be used as tolerance distribution, the only constraint being that it covers only positive values ($d \geq 0$). Experimental results often show that a few animals have a very high tolerance. To account for this, skewed distributions are preferred. The choice for one or another distribution is further motivated by the desired simplicity of the expression for F_U . The log-normal, log-logistic and Weibull distributions offer the desired shape with relatively simple expressions. Therefore, these are the statistical distributions most frequently used in dose–response analysis. The log-normal, log-logistic and Weibull distributions give rise to the log-probit, log-logistic and dose-Weibull models, respectively.

The log-probit model (frequently abbreviated to probit) assumes the logarithm of the tolerance has a normal distribution [72]. The tolerance ($U = e^W$, with W the logarithm of the tolerance) then has a so-called log-normal distribution. The resulting cumulative distribution function for U (see Fig. 4) is often expressed in terms of two parameters, θ_1 and

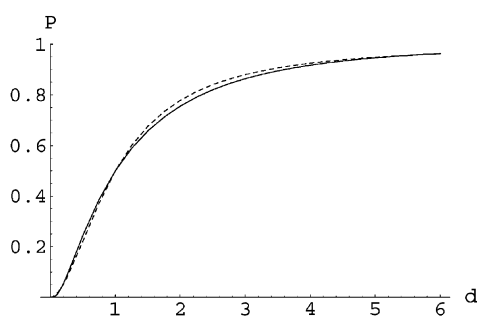


Fig. 4. Fraction of tumor bearing animals as a function of dose, $P(d)$: (solid line) prediction according to the log-probit model (W normally distributed); (dashed line) prediction according to the logit model (W logistically distributed). For both distributions the underlying stochastic W has zero expectation and unit variance (W represents the logarithm of the tolerance).

θ_2 , which relate to the mean and the variance of W as shown in Box 1. The values of parameters θ_1 and θ_2 , which have to be estimated by fitting experimental data, implicitly depend on the duration of the exposure. This follows from the fact that although time does not figure in the model, longer exposure times increase the chances of an animal developing a tumor.

The log-logistic model (usually called logit, on the analogy of probit) assumes that the tolerance U has a log-logistic distribution, or equivalently, that the logarithm of the tolerance W has a logistic distribution. The resulting cumulative distribution function for U is often expressed in terms of two parameters, ϱ_1 and ϱ_2 , which relate to the mean and the variance of W as shown in Box 1. Like in the log-probit model, the values of the parameters implicitly depend on the duration of the exposure. It can be seen from Fig. 4 that the logit and log-probit models provide very similar predictions for the fraction of tumor bearing animals. The choice among them is therefore largely arbitrary. Motivations for the use of one or the other are rarely given. We suspect that the choice for a particular tolerance distribution is mainly due to habit.

Finally, the ‘dose-Weibull model’ assumes that the tolerance has a Weibull distribution. The resulting model is

$$P(d) = 1 - e^{-\lambda d^\beta} \quad (4)$$

where again the values of the parameters implicitly depend on the duration of the exposure. Although Eq. (4) is the usual representation of the model, there

Box 1: Log-probit and log-logistic models

Let W denote the logarithm of the tolerance, $E[W]$ the mean, and $\text{Var}[W]$ the variance. Let us assume that W has a normal distribution and let us denote the mean and variance as μ and σ^2 , respectively. For the log-probit model, F_U can then be written as below, with $\theta_1 = \mu/\sigma$ and $\theta_2 = 1/\sigma$. Thus, $F_U(d) = \Phi(-\theta_1 + \theta_2 \ln d)$, where Φ represents the cumulative distribution function of the standard normal distribution. Alternatively, let us assume that W has a logistic distribution and let us denote the mean and variance as μ and $\pi^2 z^2/3$, respectively. For the log-logistic model, F_U can then be written as below, with $\varrho_1 = \mu/z$ and $\varrho_2 = 1/z$. Thus, $F_U(d) = \Psi(-\varrho_1 + \varrho_2 \ln d)$, where Ψ represents the logistic function.

Model	W	$E[W]$	$\text{Var}[W]$	Cumulative distribution function for U
Log-probit	Normal	μ	σ^2	$F_U(d) = \int_{-\infty}^{-(\mu/\sigma) + (1/\sigma) \ln d} (2\pi)^{-1/2} \exp\{-\frac{x^2}{2}\} dx$ $= \int_{-\infty}^{-\theta_1 + \theta_2 \ln d} (2\pi)^{-1/2} \exp\{-\frac{x^2}{2}\} dx$ $= \Phi(-\theta_1 + \theta_2 \ln d)$
Log-logistic	Logistic	μ	$\frac{\pi^2 z^2}{3}$	$F_U(d) = \left(1 + \exp\left\{\frac{\mu}{z} - \frac{1}{z} \ln d\right\}\right)^{-1}$ $= (1 + \exp\{\varrho_1 - \varrho_2 \ln d\})^{-1}$ $= \Psi(-\varrho_1 + \varrho_2 \ln d)$

is an alternative in which the exponent λd^β is replaced by $(d/d_*)^\beta$. The motivation for this alternative representation is that in Eq. (4) the dimension of λ depends on the value of β . As the value of β derives from experimental data, the dimension of λ varies depending on the data considered, which renders the parameter λ uninterpretable. In contrast, the parameter d_* has always the same dimension as d , and has the interpretation of a reference dose. The reference dose must depend on the exposure time, as zero exposure time cannot result in a tumor. For a given exposure time, the corresponding d_* is the level of exposure at which the fraction of tumor bearing animals is $P(d_*) = 1 - e^{-1} \approx 0.632$.

3.2.2. Empirical ‘time to tumor’ models

Survival analysis is the branch of statistical modeling that deals with the analysis of failure time data. The failure time of an individual is the time until a particular event occurs. Any event that occurs at most once to each individual defines a failure time. Because the occurrence of a first tumor is such an event, survival analysis techniques can be applied to time to tumor data.

Any continuous statistical distribution can be used as failure time distribution, the only constraint being

that it covers only positive values ($t \geq 0$). As for the tolerance models the log-normal, log-logistic and Weibull are the statistical distributions most frequently used in time–response analysis. This should not come as a surprise, as again the only motivation for their choice is in the shape and simplicity of the distributions. The Weibull distribution, for instance, is now given by

$$F_T(t) = 1 - e^{-at^b} \tag{5}$$

where the variable T represents the time to tumor. Time t (and not dose d) is now the independent variable. Therefore, we refer to this expression as ‘time-Weibull model’ in order to avoid confusion with the dose-Weibull model above. The values of parameters a and b , which have to be estimated by fitting experimental data, implicitly depend on the level of exposure.

3.2.3. One-hit and multi-hit models

Let us again consider a group of mice that have been exposed to a chemical for a particular period of time. Contrary to the tolerance distribution models discussed above, the hit models assume that the group

of animals is homogeneous with regard to their susceptibility to a process generating ‘hits’. One might think of a ‘hit’ as any of the changes discussed in Section 2.2. Let us assume that an individual develops a tumor when a hit occurs, and that the occurrence of a hit is a random event. In this special case the random variables ‘time to first hit’ and ‘time to first tumor’ are thus interchangeable. As long as a mouse is still tumor free, it may develop a first tumor with a certain probability during the next (small) time unit. In the simplest scenario, this probability per time unit (hit rate) remains constant; the variable ‘time to first hit’ then has an exponential distribution, and thus

$$h_T(t) = \mu, \quad F_T(t) = 1 - e^{-\mu t} \quad (6)$$

where T represents ‘time to first hit (or tumor)’, and μ the hit rate. The hazard rate h_T and the cumulative distribution function F_T relate to each other as explained in the introduction to Section 3.2. Although Eq. (6) is often referred to as the one-hit model in survival analysis, we refer to it as the one-hit failure-time model (OHFT model) to avoid confusion with the one-hit dose–response model presented below. The OHFT model is characterized by a constant hazard rate, μ . This implies that *susceptibility of developing* a tumor does not increase with time (age). Note, however, that the (cumulative) *chances of developing* a tumor do increase with time (age)!

A natural extension of the model above is to assume that more than one hit is required before a tumor develops, say k hits. In this special case the random variables ‘time to the k th hit’ and ‘time to first tumor’ are interchangeable. With the occurrence of a first hit, the process generating hits does not change, so that the hit rate μ still is the same. This assumption implies that the variable ‘waiting time between the first and the second hit’ also has an exponential distribution with hazard rate μ , and more generally, any waiting time between two successive hits has an exponential

distribution with hazard rate μ (see Fig. 5). In this context the parameter μ is the (mean waiting time)⁻¹.

The variable ‘time to the k th hit’ now has a so-called Erlang distribution. For further details see Box 2. Although this extension of the OHFT model is often referred to as multi-hit model in survival analysis, we refer to it as multi-hit failure-time model (MHFT model) to avoid confusion with the multi-hit dose–response model presented below.

3.2.3.1. The ‘hit’ models with dose-dependent parameters. The hit models described above do not yet account for the level of exposure, or dose. Obviously, the dose is an important determinant of the carcinogenic effect of the chemical, so it cannot be ignored. To account for the dose we have to specify its relationship with the hazard rate. Hanes and Wedel use the most simple approach to do this: they assume that the internal concentration is constant and proportional to the constant external dose (see Section 3.1), and that the hit rate is proportional to the chemical’s internal concentration [9]. These assumptions lead to a constant hit rate proportional to the external dose. The hazard rate, which equals the hit rate in the one-hit failure-time model (Eq. (6)), then becomes αd . Substitution of the expression for the hazard rate in Eq. (6) yields

$$F_T(t, d) = 1 - e^{-\alpha dt} \quad (7)$$

The probability that an animal exposed to a dose d develops a tumor before time t thus is a function of both exposure time and dose. Consequently, for a single fixed exposure time t^* it becomes a function of external dose alone. The fixed exposure time now plays the role of a model parameter with a known value. In sum, $F_T(t^*, d)$ provides a prediction of the fraction of tumor bearing animals after an exposure period t^* , given an exposure to a dose d :

$$P(d) = 1 - e^{-\lambda d} \quad (8)$$

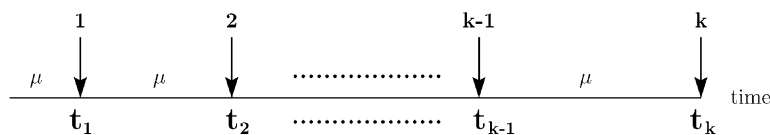


Fig. 5. The multi-hit failure-time model: it is assumed that the random variable ‘waiting time between two successive hits’ has an exponential distribution with parameter μ (k represents the number of hits required for tumor development; t_i represents the time until the i th hit). Note that the hit models are on the individual level, so μ is a probability of hit per time unit per individual.

Box 2: Multi-hit models

If any waiting time between two successive hits has an exponential distribution (with parameter μ), the variable ‘number of hits in a fixed time interval’ has a Poisson distribution (with parameter μt), and vice versa. Let T denote the variable ‘time to the k th hit’ (or ‘time to first tumor’) and let Z denote the variable ‘number of hits in a time interval of length t ’. The event in which Z is less than k is equivalent to the event in which T is greater than t . That is, $\text{prob}\{Z < k\} = \text{prob}\{T > t\} = G_T(t)$. Further, $F_T(t) = 1 - G_T(t) = 1 - \text{prob}\{Z < k\} = 1 - \sum_{i=0}^{k-1} (e^{-\mu t} (\mu t)^i / i!)$, because Z has a Poisson distribution. This expression can be written in the form shown below, where Γ represents the Gamma function:

$$F_T(t) = \int_0^t \frac{\mu^k s^{k-1} e^{-\mu s}}{\Gamma(k)} ds$$

For a fixed exposure time t^* , the MHFT model gives rise to the multi-hit dose–response model:

$$P(d) = \int_0^d \frac{\lambda^k x^{k-1} e^{-\lambda x}}{\Gamma(k)} dx$$

with $\mu = \alpha d$ and $\lambda = \alpha t^*$.

where $\lambda = \alpha t^*$ and $P(d) = F_T(t^*, d)$. Because the hit rate μ has the interpretation of the inverse of mean waiting time, and $\mu t^* = \lambda d$, the product λd stands for the mean number of hits in a time interval of length t^* . Eq. (8) is referred to as one-hit model in dose–response analysis. Normally the one-hit model is only used because of its mathematical simplicity, and an interpretation of the model is rarely given.

Likewise, substitution of $\mu = \alpha d$ into the expression for the MHFT model yields the so-called multi-hit dose–response model (see Box 2). If the number of hits required for tumor development k is equal to one, the multi-hit model reduces to the one-hit model (Eq. (8)). Thus, the multi-hit dose–response model is an extension of the one-hit dose–response model.

Above it was assumed that the hit rate is proportional to the dose ($\mu = \alpha d$). Substitution of this relation in Eq. (6) gave rise to the one-hit dose–response model. Other assumptions are also possible. For instance, one might argue that the hit rate is proportional to a power of dose ($\mu = \alpha d^\beta$). Substitution of this alternative expression for the hazard rate in Eq. (6) gives rise to the same mathematical expression for $P(d)$ as the dose-Weibull model (Eq. (4)) with $\lambda = \alpha t^*$.

3.2.4. Multi-stage models

Many epidemiologic studies have revealed that age-specific cancer-incidence rates increase with age.

Plots of the age-specific incidence rate against age yield straight lines when logarithmic axes are used. This suggests that age-specific incidence rates increase proportionally with a power of age. To explain this result, Nordling proposed that several mutations in the same cell are required to induce a tumor: “If three mutations were required, a cancer frequency proportional to the second power of age might be expected, with four mutations to the third power of age, and so on” [24]. In 1954, Armitage and Doll examined Nordling’s work and presented a mathematical formulation of his hypothesis [73]. The resulting model is now widely known as the Armitage–Doll multi-stage model (AD model), one of the first mathematical models for carcinogenesis.

The AD model assumes that several successive ‘changes’ in one cell are required to transform it into a tumor cell (see Fig. 6). Nordling maintains that the



Fig. 6. Armitage–Doll model. A normal cell (N) goes through several intermediate stages before becoming a tumor cell (M). The transition from any state to the next is determined by the occurrence of a specific change. An intermediate cell type i (Y_i) is a cell that has incurred exactly i changes. k denotes the number of changes required to transform a normal cell into a tumor cell.

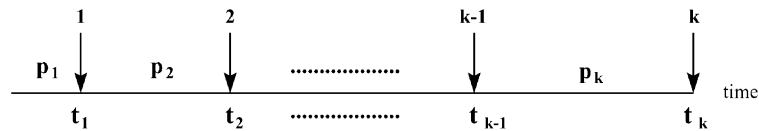


Fig. 7. Comparison of the AD model with the MHFT model. Here k is the number of changes required for malignant transformation, p_i the transition rate i , and t_i is the time to occurrence of the i th change. (a) The AD model is a model on the cellular level, whereas the MHFT model is a model on the individual level. Indeed, p_i is rate per time per cell, whereas μ is a rate per time per individual. (b) In the AD model the waiting time for a cell to go from state i to state $i + 1$ has an exponential distribution with parameter p_{i+1} , whereas in the MHFT model any waiting time for an individual to go from any state to the next has a exponential distribution with the *same* parameter μ . (c) In the AD model the changes must take place in a unique order, whereas in the MHFT model no restriction is placed on the order of the hits.

changes are mutations, but this specification is overly restrictive for the mathematical development of the AD model [73]. The only constraint on the nature of the changes is that they must be irreversible and take place independently of each other. Let us assume that k changes are required for transformation of a normal cell into a malignant one. This implies that a normal cell (N) goes through $k - 1$ intermediate stages before becoming a tumor cell (M). For any $i < k$, an intermediate cell type i (Y_i) is a cell that has incurred exactly i changes. With regard to the time course, the AD model postulates that the waiting time between any two successive changes is exponentially distributed with transition rate p_i (see Fig. 7). Finally, the AD model posits that the changes must proceed in a unique order. None of the other multi-step models impose restrictions on the order of the steps and, therefore, the last assumption characterizes the AD model.

One can translate the above assumptions into an expression for the probability that a certain cell becomes a tumor cell before time t (see, for example, [74]). One needs three additional assumptions to extrapolate this result from single cells to entire organisms. First, one has to assume that cells transform independently of each other. Box 3 shows how this assumption is used. Second, one has to know which cells are susceptible to the changes. According to the so-called stem cell theory, only proliferative cells qualify for this. The effective number of normal cells thus equals the number of ‘stem cells’. The third and final assumption maintains that the number of stem cells is constant. These assumptions, together with the other assumptions of the AD model, lead to an expression for the probability that the time to first tumor cell is less than or equal to t . In general, this

probability differs from the probability that the time to first tumor is less than or equal to t . However, on the assumption that a tumor cell constitutes a detectable tumor (for a further explanation on this assumption, see Section 3.3), the same expression describes both probabilities. This expression, usually referred to as the AD exact formula, is a rather awkward page-filling equation (see, for example, [74]). Box 3 provides a derivation of the exact formula for a two-stage model.

The original AD model is an approximation of the AD exact formula. It holds if any transition rate is small in comparison with the organism’s life span, and malignant transformation is a rare phenomenon. Box 3 includes some explanatory information on these assumptions and their implications. The approximate expression for the AD model is given by

$$h_T(t) \approx \mu t^{k-1}, \quad F_T(t) \approx 1 - e^{-(\mu/k)t^k} \quad (9)$$

where the parameter μ is proportional to the product of the transition rates p_i and proportional to the number of stem cells. According to this expression, an age-specific incidence proportional to a $(k - 1)$ th power of age indicates that malignant transformation requires k steps, and vice versa. The mathematical expression for the survivor function (Eq. (9)) is a special form of the time-Weibull model (Eq. (5)), with $b = k$ an integer. Moreover, if the number of required changes to transform a normal cell into a tumor cell is one, the AD model (Eq. (9)) reduces to the OHFT model (Eq. (6)).

3.2.4.1. The AD model with dose dependent parameters. So far we have not mentioned the level of exposure. To use the AD model in risk assessment, one

Box 3: Multi-stage models

Let N_0 denote the number of susceptible normal cells (stem cells) and J the random variable ‘time until a certain cell gives rise to a tumor cell’. The probability that an organism is tumor free at time t equals the probability that not any cell transforms into a tumor cell before time t . Under the assumption that cells transform independently of each other, this implies that $G_T(t)$ equals the product of N_0 times $G_J(t)$, or equivalently, $G_T(t) = G_J(t)^{N_0} = \{1 - F_J(t)\}^{N_0}$. In terms of the hazard rates this means $h_T(t) = N_0 h_J(t)$.

Exact formula: If two changes are required to transform a certain cell, the time to transformation equals the sum of the waiting time until the first change (K_1) and the waiting time between the first and the second change (K_2). The variables K_1 and K_2 follow an exponential distribution with parameters p_1 and p_2 , respectively. F'_J can be expressed in terms of F'_{K_1} and F'_{K_2} , as follows: $F'_J(t) = \int_0^t F'_{K_1}(t-s)F'_{K_2}(s) ds = p_1 p_2 / \{e^{-p_1 t} - e^{-p_2 t}\} = (p_2 - p_1)$. Integration gives an exact expression for F_J and, thus, also for $G_T = \{1 - F_J\}^{N_0}$.

Approximate formula: From Eq. (3), we have $G'_J(t) = -h_J(t)G_J(t)$. Because of the relation $F_J = 1 - G_J$, this is equivalent to $F'_J(t) = h_J(t)\{1 - F_J(t)\}$. In this context, the assumption that transformation is a rare phenomenon means $(1 - F_J) \approx 1$, or equivalently, $h_J(t) \approx F'_J(t)$. The hazard for T then yields: $h_T(t) = N_0 h_J(t) \approx N_0 F'_J(t) = p_1 p_2 N_0 / \{e^{-p_1 t} - e^{-p_2 t}\} = (p_2 - p_1)$. Based on expansion in Taylor series about $t = 0$ and the assumption that p_1 and p_2 are small, this expression reduces to $h_T(t) \approx p_1 p_2 N_0 t$. Thus, for the two stage model $\mu = p_1 p_2 N_0$ (Eq. (9)).

needs to assume something about the relation between the hazard rate and the dose. For instance, one might argue that each transition rate is proportional to external dose, $p_i = \alpha_i d$. More frequently each transition rate is assumed to be a linear function of dose [75], $p_i = a_i + b_i d$, where the a_i have the interpretation of background transition rates (see Section 3.2.6). F_T can then be viewed as a function of exposure time and dose. Moreover, the probability of tumor at a fixed exposure time t^* can be seen as a function of dose only:

$$P(d) = 1 - \exp \left\{ - \sum_{i=0}^k q_i d^i \right\} \quad (10)$$

where any compound parameter q_i is a product of $(t^*)^k$, the number of normal cells, and a function of the coefficients a_j and b_j . Please note that this approach disregards the step from an external dose to an internal dose. This is only justified when these two quantities are constant and proportional to each other. The only kinetic models that satisfy this constraint are linear compartment models (see Section 3.1). Eq. (10) is known as the linearized multi-stage (LMS) dose–response model [76]. If the dose is low, the following approximation holds: $P(d) \approx 1 - e^{-q_0 - q_1 d}$. An analogous expression can be obtained from the

OHFT model (Eq. (6)) by assuming that the hit rate is a linear function of dose.

3.2.4.2. Some modifications of the AD model. In the original AD model, a single cell undergoes successive changes before becoming a tumor cell. That is, the model does not account for proliferation and death of intermediate cells. In 1957, Armitage and Doll proposed a two-stage model that incorporates cell kinetics [77]. This model assumes that once an intermediate cell is generated, it starts to proliferate at a constant rate. In 1993, Chen extended the two-stage model to account for age-dependent parameters [78]. Two years later, Little generalized the two-stage model to account for an arbitrary number of stages and time-varying parameters [79].

3.2.5. Multi-event models

In 1971 Knudson conducted a statistical study on hospital patients and concluded that two mutations must occur before retinoblastoma can develop [80]. He also proposed that the first mutation is germinal in the inherited form of the disease, whereas both mutations are somatic in the non-inherited form. It is now widely accepted that this childhood cancer is caused by the biallelic inactivation of the *RB* tumor

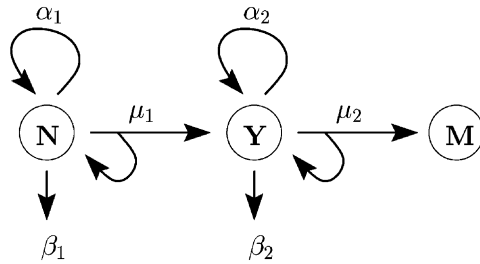


Fig. 8. Two-event model. Normal cells progress to intermediate and then to tumor cells (0, 1 and 2 mutations, respectively). The mutational events are irreversible. N, normal susceptible cell (stem cell); Y, intermediate cell; M, malignant cell; α_1 , rate (per cell per year) of cell division of normal cells; β_1 , rate (per cell per year) of death or differentiation of normal cells; μ_1 , rate (per cell per year) of division into one normal and one intermediate cell. α_2 , β_2 , and μ_2 are defined similarly for intermediate cells.

suppressor gene [81,82]. Thus, Knudson's two mutations correspond to mutational events at homologous loci of the *RB* gene. This result, generalized to the hypothesis that most tumors arise by mutation of recessive tumor suppressor genes, constitutes the basis of a two-event carcinogenesis model proposed by Moolgavkar et al. [83,84]. On the basis of the initials of the authors, this model is called the MVK model.

Like the Armitage–Doll model (AD model), the MVK model starts from the cellular level. It is for this reason that the models share some basic assumptions. For instance, both assume that only mutations in stem cells lead to cancer, and that cells transform independently of each other. However, in contrast to the AD model, the MVK model accounts for both cell proliferation and cell death. Indeed, cell kinetics plays a major role in the MVK model. Clonal expansion of intermediate cells significantly affects the probability of tumor induction, because it increases the number of target cells for the second mutational event [85,86]. Moreover, in the context of the MVK model, a 'mutational event' is equivalent to a cell division producing one mutant daughter cell. This interpretation of a mutational event was first suggested by Kendall [87]. It is based on the observation that fixation of a mutation requires at least one cycle of cell division [36,88]. Hence, in the MVK model an intermediate cell arises when a normal cell divides into one normal and one intermediate cell (such a division does not change the number of normal cells). In a similar way the genesis

of a tumor cell occurs during the division of an intermediate cell. It should be noted that a mutational event in the MVK model concerns the occurrence of an *effective* mutation for the tumor type of interest. That is, the model's mutation rates do not correspond to mutation rates measured by experimental techniques.

All tumor induction models described in the previous sections view tumor induction as a stochastic process. In the tolerance models, an individual has a probability to respond to a dose. In the multi-hit and multi-stage models, hits and changes may occur with a certain probability. The MVK model also views tumor induction as a stochastic process. It incorporates stochasticity in a different manner, though. It assumes that the mutational events as well as cell division and cell death are random events. Hence, in any small time interval, normal cells may divide into two normal cells, die or differentiate, or divide into one normal cell and one intermediate cell. Likewise, intermediate cells may divide into two intermediate cells, die or differentiate, or divide into one intermediate cell and one tumor cell. Each of these events may occur with a certain probability. Further, the model assumes that the probability of more than one event occurring in the small time interval is negligibly small. Finally, the MVK model assumes that a tumor cell constitutes an observable tumor. The AD model also uses the terms tumor and tumor cell interchangeably. For a further explanation on this assumption, see Section 3.3.

A model based on the assumptions above was considered difficult to apply. For this reason, an approximation has been used based on the assumption that the probability of malignant transformation is small (the resulting expression for the hazard rate is shown in Box 4). The same assumption is also used in the AD model to obtain an approximate expression. Such an approximation can be useful if the result does not deviate significantly from the full model. This appears not to hold for this approximation, though. Moolgavkar and Dewanji pointed out that the approximation is unlikely to be adequate when dealing with animal experiments in which the probability of tumor is high [89]. Furthermore, several studies have revealed that the approximate MVK model can deviate significantly from the full model for certain parameter values [74,90–92]. To avoid misleading results, the use of the full model is recommended by these authors. Several studies exemplify how the full model can be applied

Box 4: The MVK model

The following parameters figure in the MVK model (see Fig. 8): N_0 , initial number of normal susceptible cells (stem cells); α_1 , rate of cell division of normal cells; β_1 , rate of death or differentiation of normal cells; μ_1 , rate of division into one normal and one intermediate cell. α_2 , β_2 , and μ_2 are defined similarly. For a detailed mathematical development of the model see, for example, [83,93]. For a description of the mathematical techniques see, for example, [11,120].

Approximate formula: It relies on the assumption that the probability of malignant transformation is small. In the particular case that the rates of mutation, cell division, and cell death remain constant in time, the following expression for the (approximate) hazard rate can be derived:

$$h_T(t) \approx \begin{cases} \frac{s_1}{s_2 - s_3} \{e^{s_2 t} - e^{s_3 t}\} & \text{for } s_2 - s_3 \neq 0 \\ s_1 t & \text{for } s_2 - s_3 = 0 \end{cases}$$

where $s_1 = \mu_1 \mu_2 N_0$, $s_2 = \alpha_1 - \beta_1$, and $s_3 = \alpha_2 - \beta_2$.

Exact formula: An exact analytical expression for the hazard rate can be derived on that assumptions that (i) the number of normal cells is constant, and (ii) the parameter values remain constant. The hazard rate is then given by [94,95]

$$h_T(t) = \frac{1}{2x_1} \frac{(x_3 - x_2^2)[e^{\sqrt{x_3}t} - 1]}{(\sqrt{x_3} - x_2) + (\sqrt{x_3} + x_2)e^{\sqrt{x_3}t}}$$

where $x_1 = \alpha_2 / \mu_1 N_0$, $x_2 = \beta_2 + \mu_2 - \alpha_2$, and $x_3 = [(\alpha_2 + \beta_2 + \mu_2)^2 - 4\alpha_2\beta_2]$ are identifiable parameters (i.e. these compound parameters can be uniquely estimated from experimental data) [115]. Note that the exact formula has the same number of parameters and can be implemented as easily as the approximate formula. Again, the hazard rate and the survivor function relate to each other as shown in Eq. (3).

to epidemiological and experimental data (for review, see [74,93]).

One way to get closer to a workable expression for the full model is to make the additional assumption that the number of normal cells (N) is constant. This is approximately true if the number of normal cells is large. If, in addition, the rates of mutation, cell division and cell death remain constant in time, the model yields exact analytical expressions for the hazard rate and the survivor function [94,95] (see Box 4). Less restrictive is the assumption that the parameters are piece-wise constant. This means that the parameters are constant for a certain time interval; they then may change, after which they are again constant for some time. A closed form expression can be found for such a model [92,93], but this expression is “not easy for non-mathematicians” [91]. Because of this difficulty, Clewell et al. and Hoogeveen et al. developed an improved approximate model with arbitrarily

time-varying parameters [91,96]. For time-constant parameters, the derived expression is exact [91].

Above we used the MVK model to obtain an expression for the hazard rate and the survivor function, which predict the fraction of tumor bearing animals. Interestingly, we can also use the model to obtain expressions regarding the size and number of intermediate clones (foci) [97–99]. This is relevant for those experiments in which information on the number of premalignant clones and their sizes is available. The proper way to analyze data on foci is currently topic of research [100–103].

3.2.5.1. The MVK model with dose-dependent parameters. If one wants to use the MVK model in risk assessment, one needs to specify how the parameters of the model depend on the level of exposure. Two choices need to be made for this. One has to decide which parameters are affected by a particular

chemical, and one must specify how they are affected. Thorslund et al. presented an overview of possible answers to the first question [104]. However, in practice only two of the possibilities are considered: genotoxic carcinogens can act by altering the mutation rates, whereas non-genotoxic carcinogens can act by changing cell kinetics.

Before we explain how carcinogens can affect the parameters, we introduce some compound parameters. The effects of non-genotoxic carcinogens are most easily characterized in terms of these parameters. The compound parameters are simple functions of the basic parameters of the MVK model, which are shown in Fig. 8. The first compound parameter is the mutation probability, m_1 . If α_1 denotes the rate (per cell per year) of cell division of normal cells and μ_1 the rate (per cell per year) of aberrant division into one normal and one intermediate cell, then the mutation probability at cell division is $m_1 = \mu_1/(\mu_1 + \alpha_1)$. The second compound parameter characterizes the net proliferation of a normal cell. If β_1 denotes the rate (per cell per year) of death or differentiation of normal cells, then the net proliferation rate of a normal cell is $(\alpha_1 - \beta_1)$. The parameters α_2 , β_2 , μ_2 , and m_2 describe the behavior of intermediate cells in a similar way.

In the context of the MVK model, a genotoxic carcinogen increases the mutation rates (μ_1 and μ_2). Obviously, an increase in either of the mutation rates (or both of them) accelerates tumorigenesis. Theoretical interest in this possibility is limited, probably due to the trivial nature of the mechanism. In the last decade, modeling the effect of non-genotoxic carcinogens has received much more attention, due to the increasing interest in the role of cell proliferation in tumorigenesis (e.g. [43,105,106]). The architecture of the MVK model suits the study of this problem, as it explicitly accounts for cell kinetics. A non-genotoxic carcinogen increases the parameters involved in cell kinetics, without changing the mutation probabilities m_1 and m_2 . Note that such chemicals indirectly increase the mutation rates per cell per year. That is, if α_2 increases, μ_2 must increase for the probability $m_2 = \mu_2/(\mu_2 + \alpha_2)$ to remain constant.

A non-genotoxic carcinogen may increase both cell division and death rates in such a way that the net proliferation rate of an intermediate cell does not change. Indeed, this occurs if the chemical increases α_2 and β_2 such that $(\alpha_2 - \beta_2)$ remains constant. This

leads to a rather small effect on tumor incidence [107]. In contrast, even small changes in the net proliferation rate of intermediate cells ($\alpha_2 - \beta_2$) lead to a rather profound effect on tumor incidence [107]. In this situation the non-genotoxic chemical affects tumor incidence by at least two mechanisms, namely increasing the mutation rates (μ_1 and μ_2) while simultaneously increasing the net change in the number of intermediate cells [86]. Moolgavkar suggested that the action of hormones exemplifies this phenomenon [85].

3.2.5.2. Some modifications of the MVK model.

Since its first publication, the MVK model has received considerable attention (e.g. [108–113]). Many investigators either sought to extend the model to incorporate further biological details, or to facilitate its practical use in cancer risk assessment. Several theoretical studies on the full MVK model deal with improved implementation (e.g. [91,114]) and parameter identifiability (e.g. [92,115–117]). Model extensions accounting for tumor growth are treated in Sections 3.3 and 3.5. Other examples of model extensions are discussed below.

A first example of model extension is the three-event model proposed by Moolgavkar [118]. The motivation for this extension was the classic paper on colorectal cancer by Fearon and Vogelstein [22]. In the model the first two events correspond to mutations at homologous loci of the *DCC* gene, whereas the third is a mutation at one allele of the *p53* gene [118]. To account for the fact that different cancers may involve different numbers of mutations, Little has provided an expression for a multi-event model with an arbitrary number of steps [79]. The practical interest of this expression is somewhat doubtful. With any additional step included, the number of parameters piles up, making practical application of the model impossible. For practical application the two-event model is used even if it is known that more than two steps are involved. This use is motivated by the casual assumption that only two steps are rate limiting.

Multiple pathway models were first developed by Tan and Chen [119]. The motivation for such models was the observation that the same type of tumor might arise from different pathways. Multi-variate models account for the possibility that a single agent may induce two or more different types of tumors [120].

Mixed models allow for different individuals in the population either to start the process of carcinogenesis at different steps of the same pathway, or to involve different pathways [121]. For an exhaustive study on multiple-pathway, mixed and multi-variate models, see Tan [120].

Attempts have also been made to describe in more detail the interaction between the carcinogen and the cell. For instance, a few models explicitly account for DNA repair. Among them are the damage-fixation model formulated by Portier and Kopp-Schneider [122], and the model developed by Bois and Compton-Quintana [123]. Both models describe DNA repair as a random process. Alternatively, Conolly incorporated DNA repair in a deterministic way [124] by describing the formation of DNA adducts. The adduct formation rate is assumed to be proportional to the amount of genotoxic carcinogen and to the amount of nucleotides, whereas the adduct repair rate is assumed to be proportional to the amount of adducts. The MVK mutation rates (μ_1 and μ_2) are then assumed to depend on the amount of adducts.

3.2.6. Background tumor incidence

The dose–response models we discussed above aim to relate tumor incidence to the dose the animals are exposed to. Experiments concerning this relationship always include a control group of non-exposed animals ($d = 0$). Most of the dose–response models above predict absence of tumor incidence in this group, that is, no dose implies no response ($P(0) = 0$). This contradicts the observational evidence that tumors often develop in control animals. Background incidence can be easily incorporated into the models, however. It requires a choice between two assumptions [75,125]. One assumption is frequently referred to as ‘additive background assumption’, whereas the other is frequently referred to as ‘independent background assumption’.

An additive background means that the same mechanism is responsible for both spontaneous and induced tumors. This assumption holds when the carcinogen acts by accelerating naturally occurring processes. To account for an additive background response, one often introduces a dummy dose d_0 . So, one postulates an unknown background dose to be responsible for background tumor incidence. The fraction of animals bearing either a spontaneous or induced tumor at dose

d is then $P^*(d) = P(d_0 + d)$, where P represents some dose–response model.

On the basis of a time-dependent model, whose parameters have an interpretation, a more realistic approach is possible. This approach requires two additional choices. First, one has to decide which parameters are affected by the chemical. We already addressed this topic for the hit, multi-stage, and multi-event models. Second, one has to specify how the parameters are affected. For instance, to use the Armitage–Doll model as dose–response model, the transition rates are assumed to depend linearly on dose, $p_i = a_i + b_i d$, where a_i has the interpretation of a background transition rate. Thus, a linear dose-dependence accounts for background incidence. Moreover, any dose-dependence of the form $p_i = a_i + g_i(d)$, where g_i is an arbitrary function satisfying $g_i(0) = 0$, accounts for an additive background incidence.

An independent background means that different mechanisms are responsible for spontaneous and induced tumors, and that both mechanisms take place independently of each other. In this context, one often uses what is known as Abbott’s correction [126] to predict the fraction of animals bearing either a spontaneous or an induced tumor:

$$P^*(d) = P_0^* + \{1 - P_0^*\}P(d) \quad (11)$$

where P represents some dose–response model describing the occurrence of induced tumors; P_0^* is the tumor probability at dose 0. Box 5 provides a derivation of this expression.

3.3. Tumor growth models

Tumor induction models at the cellular level (such as the AD model and the MVK model) characterize the random variable ‘time to first tumor cell’. As we showed above, they are used to analyze time to tumor data on the assumption that a single cell constitutes a detectable tumor. Such use is warranted if the tumor type fulfills two conditions. The first is that a tumor arises from a single cell; the observation that most tumors are monoclonal supports this. The second is that the time span a tumor cell requires to become a detectable tumor is negligibly small in comparison with the duration of tumor induction. The time to observing a tumor then roughly equals the time to

Box 5: Independent background assumption

Let R , I and T denote the random variables ‘time to spontaneous tumor’, ‘time to induced tumor’, and ‘time to tumor’ (as a consequence of the independence assumption, R does not depend on the level of exposure). At any point in time, the fraction of tumor-free animals is the fraction of animals that have neither a spontaneous nor an induced tumor. Under the independence assumption, this implies that $G_T^*(t, d)$ equals the product of $G_R(t)$ and $G_I(t, d)$. Expressed in terms of the cumulative tumor probabilities this translates into

$$F_T^*(t, d) = F_R(t) + \{1 - F_R(t)\}F_I(t, d)$$

For a fixed exposure time, this equation reduces to Eq. (11) (Abbott’s correction), with $P(d) = F_I(t, d)$. As the original dose–response model only accounts for induced tumors, it provides an expression for $F_I(t, d)$. In terms of the hazard rates the relation above implies $h_T = h_R + h_I$. Swanyer et al. relate h_I to h_R through a linear proportional hazard assumption, $h_I = \alpha dh_R$. In this particular situation, the survivor functions for R and T relate to each other as $G_T^* = \{G_R\}^{1+\alpha d}$.

developing a tumor, and the terms ‘tumor cell’ and ‘tumor’ become interchangeable.

For monoclonal fast growing tumors the growth period can thus be ignored. However, neglect of tumor growth is less realistic for slowly growing tumors as well as for rapidly induced tumors. If tumor growth cannot be neglected, the simplest way to account for it is to assume that the time it takes a tumor cell to reach a detectable size is constant, say t_g . The fraction of tumor bearing animals at time t is then the fraction of animals with a tumor cell at time $t - t_g$. A prediction for the latter fraction is provided by the original model. Iversen and Arley considered t_g , the time delay between the genesis of a tumor cell and the emergence of a detectable tumor, but did not assume its value to be constant [127]. In contrast, they assumed it to be a normally distributed random variable.

Disregard tumor growth rules out the possibility that a tumor regresses before reaching a detectable size [128]. In other words, it implies that once a tumor cell arises, it will certainly give rise to a detectable tumor. This also holds for the models above that incorporate growth as a time delay. Moreover, it also holds for any model that describes tumor growth in a deterministic way that does not account for a decrease in size. In the next sections we compare some deterministic growth models. As an alternative, some models for chemical carcinogenesis account for tumor growth in a stochastic way (e.g. [99,121,129,130]). In such models tumor cells are subject to stochastic birth-and-death processes, i.e. in any small time interval, tumor cells

may divide or die with a certain probability. The actual probabilities depend on the tumor growth model.

Sherman and Portier modeled stochastic tumor growth on the basis of clones [130]. So, the process of tumor growth starts when an intermediate cell gives rise to a malignant clone of size one. If the cell dies, the clone becomes extinct. If the cell divides, a clone of size two results. If then either cell dies the clone reverts to a single-cell clone, and so on. In summary, if an observable clone consists of M_{\min} cells, to become observable a single-cell clone has to go through M_{\min} stages of increasing size [130,131]. The model assumes that once a clone reaches a detectable size, it can no longer regress in size.

3.3.1. Classic growth models

Several classic growth models from various disciplines have been used to describe tumor growth (see [132]). In this section we deal with four such models: exponential growth, Von Bertalanffy growth, Gompertz growth, and logistic growth. Mathematically these models share a common pattern: $V' = r(V)V$, where V denotes tumor volume, and $r(V)$ denotes the relative growth rate or the increase in volume per unit volume per unit time. The relative growth rate is thus some function of the size of the tumor; its actual expression differs among the different growth models. With the additional assumption that all cells within the tumor have the same volume, the equation above can be rewritten in terms of the total number of tumor cells. This is specially relevant for

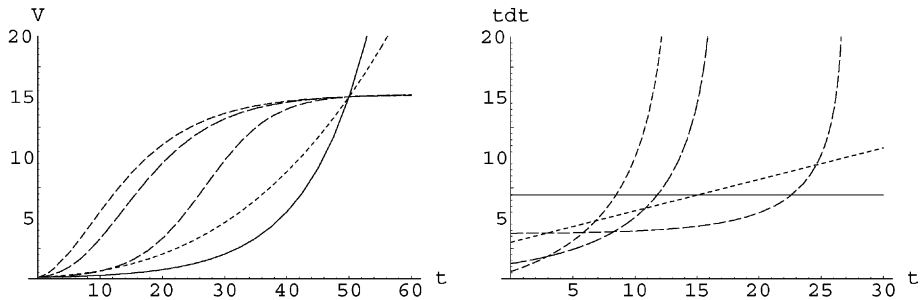


Fig. 9. (Left) Tumor volume as a function of time: (solid line) exponential growth; (dotted line) cube root growth; (broken lines from top downwards) Von Bertalanffy, Gompertz, and logistic growth. For all the models we chose parameters values such that $V(0) = 0.1$ and $V(50) = 15$ (and $V_{\max} = 15.20$ when relevant). (Right) Tumor doubling time (tdt) as a function of time. The saturating curves reach half the maximum volume at $t \approx 13.14$, $t \approx 16.68$ and $t \approx 26.87$, respectively. Exponential growth results in a constant tumor doubling time (solid line). Cube root growth results in a tumor doubling time that increases linearly with time (dotted line). The other growth models result in a tumor doubling time that increases more and more rapidly with time (broken lines).

disseminated or dispersed tumors such as leukemias and lymphomas.

The simplest growth model, the exponential growth model, results from the assumption that a constant fraction of tumor cells divide and die per time unit. If so, the relative growth rate is constant in time, and the tumor doubling time is also constant. The tumor doubling time is $(\ln 2/r)$, with r the relative growth rate. Another relevant characteristic of the exponential growth model is that there is no maximum tumor volume (see Fig. 9).

The Von Bertalanffy growth model [133] defines growth as the net result of gains and losses in volume due to anabolic and catabolic processes, respectively. The gain in volume is proportional to tumor surface area, whereas volume loss is proportional to tumor volume. An additional assumption states that a tumor maintains the same shape during growth (isomorphic growth), so that its surface area is proportional to its $(\text{volume})^{2/3}$ [134]. Contrary to the exponential growth curve, the Von Bertalanffy curve is S-shaped with an asymptotic maximum tumor volume (see Fig. 9).

The model most widely used to describe tumor growth is the Gompertz growth model [135,136] (see Fig. 9). As early as 1934, Casey used Gompertz curves to analyze experimental results on tumor transplantation [137]. Likewise, in 1964, Laird fitted the Gompertz growth model to tumor growth data with success [138]. It is intriguing that this model originally conceived as a “law of human mortality” [135] gives such an accurate description of tumor growth.

Logistic growth arises from the assumption that the relative growth rate declines linearly with tumor volume. The resulting growth model is a S-shaped curve with an asymptotic maximum tumor volume (see Fig. 9). The logistic growth equation was originally used by Verhulst to describe the growth of biological populations [139]. No biological mechanisms underlie its formulation.

As long as tumor size remains small with respect to its maximum, the relative growth rate remains approximately constant for both the logistic and the Gompertz model. As a consequence, it may not be possible to discriminate between the exponential, logistic, and Gompertz models during the early growth period on the basis of experimental data.

There are at least two ways to compare growth models. A fit of the different models to the same data is useful to reveal similarities. Vaidya and Alexandro have carried out such an analysis for solid tumors [140], and Afenya and Calderón have done the same for disseminated tumors [141]. As an alternative, in Fig. 9, we sought to reveal the differences. To do this we forced the growth curves to include two values, $V(0) = 0.1$ and $V(50) = 15$. In addition, for the saturating curves, we chose a fixed asymptotic maximum volume, $V_{\max} = 15.20$. For the three S-shaped growth curves the doubling time becomes larger as the growth process continues. Laird observed that for the Gompertz model the doubling time increases slowly at the beginning of the growth process, but more and more rapidly as the tumor becomes larger

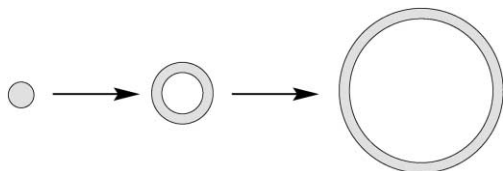


Fig. 10. Living layer model. (1) Early tumor without a dead kernel (radius smaller than or equal to the maximum thickness of the living layer). The whole tumor grows exponentially. (2) The tumor's radius has exceeded the maximum thickness of the living layer and, consequently, the tumor has developed a dead core (inner white sphere). (3) Advanced tumor: the tumor has increased in size, whereas the thickness of the living layer has remained constant.

[138]. This also applies to the logistic and Von Bertalanffy models. Another common characteristic of the S-shaped curves is that the doubling time becomes infinitely large when the tumor approaches half its maximum size. Fig. 9 depicts these results.

3.3.2. Living layer model

The living layer model, proposed by Mayneord, is based on the experimental observation that solid tumors often have a dead kernel surrounded by a shell of viable tumor cells (see Fig. 10). The model is the first one that relates tumor growth to the tumor's structural heterogeneity. It involves three assumptions; first, the tumor is spherical in shape; second, there is a fixed maximum thickness (L_{\max}) for the living outer layer; and third, the cell population within the living layer grows exponentially [142].

Interestingly, the living layer model predicts three growth phases. During the first phase tumor radius is smaller than or equal to L_{\max} . As a consequence of the third assumption, the tumor grows exponentially. The second phase starts when tumor radius becomes larger than L_{\max} , giving rise to the development of a dead core. Finally, when the ratio of L_{\max} to tumor radius became very small, tumor growth tends to follow the cube root law (i.e. the cube root of tumor volume increases linearly with time; see Fig. 9).

3.3.3. Complex tumor growth models

For clinical studies growth of tumors is of prime relevance. With the motivation to improve understanding and treatment of cancer, several complex models for tumor growth have been developed. Most of them are diffusion-limited growth models. These models

describe the growth of avascular tumors (or tumorspheroids), assuming that nutrients must be taken up across the surface of the tumor. Uptake of nutrients at the surface together with their use by the tumor tissue creates a nutrient gradient within the tumor. The diffusion-reaction theory allows one to predict this gradient [143,144]. Along the gradient, both proliferation and death of cells vary, according to availability of the nutrient. Most models assume that cells grow exponentially if the concentration of nutrients is above some critical level. When the tumor becomes larger, this approach predicts a structurally heterogeneous tumor with a outer living layer that remains constant in thickness and a necrotic core [145]. Basic diffusion models have been extended to account for more realistic biological details such as presence of growth inhibitors [146], non-uniform nutrient consumption [147], and tumor-immune system interactions [132].

The complex growth models have the potential to increase our understanding. For practical applications such as the description of experimental data on chemical carcinogenesis they are less suited, due to the complexity of the mathematics they use. For this reason we only briefly mention these models.

3.4. Effects

In carcinogenicity studies on skin tumors the time to tumor onset (i.e. time to detectable tumor) is directly observable. This is the exception rather than the rule, however. In most carcinogenicity tests the time to tumor onset cannot be observed and, therefore, the presence of a tumor can only be detected after the death (or sacrifice) of the animal. Consequently, there is a need to make inferences about the actual time to tumor occurrence T , using the time to death caused by the tumor, \dagger . As tumor bearing animals have a certain probability to die from the tumor, such inferences concern tumor lethality.

Sawyer et al. assume that the tumor of interest is instantly lethal [148]. This implies that time of death with tumor equals time to tumor onset. This method is appropriate for highly aggressive tumors, because they cause death shortly after their onset. However, many carcinogenicity tests involve tumors that do not significantly affect survival. The use of the time to death from tumor as a surrogate for time to tumor occurrence is then unrealistic.

Data including cause-of-death information usually distinguish between five outcomes for the cause of death, only one of which is death from tumor of interest. The possible competing causes are natural death with incidental tumor, natural death without tumor, sacrifice with tumor, and sacrifice without tumor. Modelers usually cluster these outcomes into (1) death from tumor; (2) death from competing risk, with tumor present; and (3) death from competing risk, with no tumor present.

Some authors have sought to use cause-of-death information (e.g. [149,150]). However, because cause-of-death determinations are frequently unreliable, other authors have sought to carry out the required inferences without using cause-of-death information (e.g. [151–154]). Both approaches mainly have resulted in purely statistical (non-parametric) methods that do not rely on any conceptual model. A few authors assume a mathematical model for the time to tumor onset, whereas they use non-parametric estimates for both the time to death from tumor and time to death from a competing risk.

Dewanji et al. proposed a model-based approach that accounts for the three types of death [155]. The use of cause-of-death data is possible but not necessary. They relate time to tumor onset and time to death from tumor through a lethality parameter ρ lying between 0 (incidental tumor) and 1 (rapidly fatal tumor). They assume that the hazard rate for \dagger is proportional to the hazard rate for T , with proportionality coefficient ρ . Consequently, the corresponding survivor functions relate to each other as $G_{\dagger}(t) = \{G_T(t)\}^\rho$. If $\rho = 0$, there are no deaths from tumor ($G_{\dagger} = 1$), whereas if $\rho = 1$, the tumor is instantly lethal ($G_{\dagger} = G_T$).

3.5. Combined models

Models for chemical carcinogenesis aim to characterize the relation between tumor incidence and level of exposure to a certain (pro)carcinogen. The first model for chemical carcinogenesis, which was proposed by Iversen and Arley in the early 1950s, includes kinetics, tumor induction and tumor growth [127]. However, most of the models that are currently used focus on tumor induction only. That is, they use fairly elaborate assumptions on the induction process, whereas they use shallow ‘default assumptions’

for the remaining phases. Only a few models embrace more than one phase in similar detail. In this subsection we deal with these models.

3.5.1. Kinetics + induction

All dose–response models in Section 3.2 rely on the default assumption for kinetics, which states that the internal concentration is constant and proportional to the external dose. As discussed in Section 3.1, this only holds for linear kinetics and constant external dose. A more realistic alternative is estimate the internal concentration with the aid of a kinetic model. Such an approach combines kinetics and induction in one model.

To illustrate how a kinetic and an induction model can be combined, we consider the one-compartment model (Section 3.1) and the one-hit failure-time model (Section 3.2.3). If we assume that the hit rate is proportional to internal concentration ($\mu = \omega C$), it is no longer constant in time. The hazard rate, which equals the hit rate in the one-hit model, now becomes

$$h_T(t, d) = \frac{u\omega}{\eta}(1 - e^{-\eta t})d$$

Because the hazard rate is not constant, the variable time to tumor T is no longer exponentially distributed. If after some time the internal concentration reaches steady-state conditions, the hazard rate becomes constant. The model then reduces to the one-hit dose–response model (Eq. (7)), with $\alpha = u\omega/\eta$.

Van Ryzin and Rai developed a combined model that embraces the phases of kinetics and induction [156]. They described the internal concentration in the target organ by assuming Michaelis-Menten kinetics for both the incoming and outgoing processes. To solve the resulting mass balance equation, they assumed the external dose and internal concentration to be constant (steady-state conditions). Consequently, the internal concentration depends hyperbolically on the external dose. The model relates induction to internal concentration through a Weibull equation, $P(d) = 1 - e^{-\gamma - \omega C(d)^\beta}$, where γ is a parameter that accounts for background incidence.

Still other combinations between kinetic models and induction models have been explored. For instance, to analyze the effect of metabolic transformation on tumor induction, Tan and Singh combined Michaelis-Menten kinetics with the MVK model

[157]. Conolly et al. also used the MVK model, but combined it with a PBPK model [124]. In this model the internal amount of active metabolite affect the MVK parameters, according to the two types of carcinogens discussed in Section 3.2.5. Bogen used approximate multi-event models to study tumor induction associated with exposure to chlorinated methanes [158]. To predict the effective liver concentration, he used a PBPK model. Reitz et al. combined a PBPK model with the LMS dose–response model, in an attempt to predict liver angiosarcoma incidence due to vinyl chloride exposure [159]. Some of these attempts are difficult to evaluate, because they do not present either the model structure or the model equations. Indeed, most of the articles cited above focus on model results, rather than on model descriptions.

3.5.2. Induction + growth

Multi-stage models (Section 3.2.4) and multi-event models (Section 3.2.5) characterize the random variable ‘time to tumor cell’. These models are directly used to analyze carcinogenicity tests on the assumption that the terms ‘time to tumor cell’ and ‘time to tumor’ are interchangeable. This constitutes the default assumption for tumor growth. As argued in Section 3.3, it only holds for monoclonal fast growing tumors (see Section 3.3).

A few attempts have been made to combine tumor induction and tumor growth in one model. For instance, Iversen and Arley define time to tumor onset as the sum of an ‘excitation-time’ and a ‘growth-time’ [127]. Excitation involves the interaction of the chemical with a cell. The model describes excitation-time (i.e. time to tumor cell) on the basis of the one-hit theory (Section 3.2.3). The growth-time plays the role of a stochastic delay between developing and observing a tumor. Alternatively, Yang combined a multi-event model with a tumor growth model [128]. More recently, Sherman et al. extended the MVK model to incorporate tumor growth [130,131]. The model describes tumor growth on the basis of stochastic division and death of tumor cells. Like Yang’s model, it treats the size of a detectable tumor as a constant (see Section 3.3). Luebeck and Moolgavkar consider a threshold tumor size for which the probability of extinction is negligibly small [129]. Beyond this threshold, the tumor can be assumed to grow deterministically. The difference between the detection

threshold and the viability threshold is that the latter changes with the parameter values.

Models that combine induction and growth where both are treated as a stochastic process tend rapidly to become complicated. This might be one of the reasons that so few of these models have been formulated. Although models that combine kinetics and induction are slightly more common, in general, the combined models are far outnumbered by models that focus on induction alone.

4. Conclusions

In this paper we gave an overview of models describing any part of the chemical carcinogenesis process. We structured the overview according to a division of the entire process, from exposure to effect, into four consecutive phases. An alternative criterion to classify the models is whether a model is descriptive or mechanistic. This hardly is an all-or-none criterion; rather it defines a continuum with descriptive and mechanistic as endpoints. Tolerance distribution models, empirical time to tumor models, and classic growth models obviously cluster at the descriptive end of the continuum, whereas PBPK models and multi-event models move some distance towards the mechanistic endpoint.

All descriptive models or, as we argued in the introduction to Section 3.2, mathematical descriptions, have in common that they are rather simple mathematical expressions with a small number of parameters. They are useful to neatly summarize results of experiments in a few numbers, the parameter estimates. The summarizing parameter estimates lack a biological interpretation, however. This is different for mechanistically oriented models that rely on a set of assumptions on biological aspects of the process. Their parameter estimates provide quantitative information on the rates of processes and the factors that affect them [160].

Tests to determine the carcinogenic potency of a chemical aim to reveal the relationship between exposure to the chemical and occurrence of a carcinogenic response. This response may involve appearance of unusual tumors, increase in incidence of normal tumors, earlier occurrence of normal tumors, or increase in multiplicity of normal tumors. As we stated in the

Introduction, models may facilitate the analysis of the sought relationship. Here we will evaluate to what extent current models actually contribute to this aim. But let us first briefly consider the data that have to be described.

In a standardized long-term carcinogenicity test several groups of animals are administered different levels of the chemical. Besides these dose-groups, the test includes an unexposed control group. The experiment has a fixed duration; those animals still alive at the end of the experiment are sacrificed to determine whether they bear tumors. The observed carcinogenic response, or data, vary from one study to another. If the tumor of interest is directly observable, tumor-onset times are recorded. In contrast, if the tumor is not directly observable, time to death (or sacrifice) and tumor pathology at the time of death (or sacrifice) are recorded.

Between exposure and effect is a chain of processes, summarized in Fig. 11. In short, everything starts with the presence of a (pro)carcinogen. Kinetic processes lead to a certain internal concentration, which may induce the appearance of a tumor cell. This cell may become a detectable tumor in due time. The presence of this tumor may eventually result in the death of the animal.

In Section 3 we gave an overview of models describing any part of chemical carcinogenesis. As shown in Fig. 11, models for the different phases derive from different disciplines. Kinetic models are mainly in the domain of toxicology; cancer risk assessment deals with tumor induction models; models for tumor growth are in the realm of clinical oncology; and effect models are the topic of epidemiology. Moreover, much of the work on different phases apparently proceeds in isolation of each other, as can be inferred from the paucity of cross-references.

Models for chemical carcinogenesis aim to characterize the relation between dose and carcinogenic response. It seems desirable that such models deal with the entire process, from exposure to effect. Ideally, the models for the four phases should form a chain, in which the output of one model serves as the input for the next. Indeed, the very first model for chemical carcinogenesis by Iversen and Arley started to do just this [127] (the model accounts for kinetics, induction, and growth). Current modeling is apparently not motivated by such a desire, though. Models for chemical

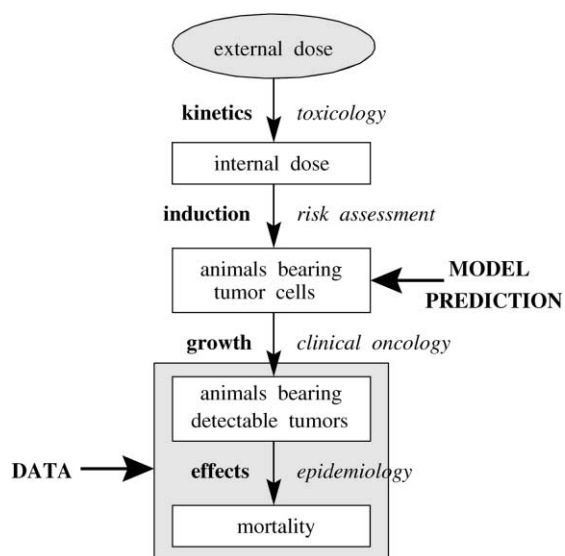


Fig. 11. Chemically induced carcinogenesis as a four-phase process: kinetics concerns the relationship between exposure and effective internal dose. Tumor induction comprises the chemically induced transformation of normal cells into tumor cells. Tumor growth relates to the clonal expansion of a tumor. Effects involves the consequences of tumor development for the organism. As indicated in italics, models devised for the different phases derive from different disciplines. The large arrows indicate that tumor induction models make predictions concerning tumor cells, whereas one can only detect tumors of a certain minimum size. Even worse, one may not be able to detect tumors till the animal dies.

carcinogenesis, with exception of the few combined models, focus on the tumor induction phase. This situation has some undesirable consequences.

One consequence is that most models for chemical carcinogenesis suffer from an imbalance between depth and width. For instance, a model for chemical carcinogenesis that covers tumor induction aspects in depth may shallowly treat, or even virtually ignore, kinetic aspects. Much new work tends to aggravate this imbalance, as it expands existing models by adding more biological details on tumor induction only. Biological details are probably added to increase the realism of the model. To judge whether a model has indeed become more realistic after addition of some new elements, it has to be confronted with data. Here we stumble upon another undesirable consequence. The predictions of an induction model concern tumor cells, whereas the data concern detectable tumors. In other words, the model predictions do not directly

bear on the data! (This is also indicated in Fig. 11.) This makes the interpretation of parameter estimates rather uncertain. It also hampers a straightforward interpretation of the realism of the models involved.

Let us now return to the original question: do current models contribute to the aim to relate exposure to carcinogenic response? Although sweeping generalizations over heterogeneous collections are always difficult to make, we think the answer is at best 'to some extent'. The models used are not overly realistic for the purpose of data description, because they ignore essential processes. An increase in realism should be sought in accounting for these processes in the first place. This may bridge the gap between model predictions and data. Other prospect for improvement is offered by the use of alternative information, especially on foci, as end-point in carcinogenicity testing. If the data concern foci rather than tumors or mortality, the gap between data and prediction could be eliminated. The model then only needs to cover kinetics and development of foci; there is no need to pay attention to tumor growth and the impact of the tumor on the animal. With such improvements the models may become much more useful for the purpose of estimation of cancer risk from carcinogenicity tests.

Acknowledgements

We should like to take this opportunity to thank Bas Kooijman, head of the Department of Theoretical Biology, for his support. Our thanks are also due to Jacques Bedaux, Bernd Brandt, Jan Kooter, Henk van Kranen, Wout Slob and two anonymous referees for their insightful comments on an earlier draft of the paper. The work presented in this paper was part of the research project VBI-4692, financially supported by the Netherlands Technology Foundation (STW) and DSM.

References

- [1] L.A.G. Ries, C.L. Kosery, B.F. Hankey, B.A. Miller, L. Clegg, B.K. Edwards (Eds.), SEER Cancer Statistics Review, 1973–1996, National Cancer Institute, Bethesda, 2000.
- [2] J.H. Lubin, W.J. Blot, F. Berrino, R. Flamant, C.R. Cillis, M. Kunze, D. Schmahl, G. Visco, Modifying risk of developing cancer by changing habits in cigarette smoking, *Br. Med. J.* 288 (1984) 1953–1956.
- [3] R. Peto, A.D. Lopez, J. Boreham, M. Thun, C. Heath, R. Doll, Mortality from smoking worldwide, *Br. Med. Bull.* 52 (1996) 12–21.
- [4] V.L. Fulgoni, A.G. Ramirez, Cancer: the role of diet, nutrition, and fitness, *Cancer* 15 (1998) 1775–1783.
- [5] P. Pott, Chirurgical Observations Relative to the Cataracts, the Polypus of the Nose, the Cancer of the Scrotum, the Different Kinds of Ruptures and the Mortifications of the Toes and Feet, Hawes, Clarke and Collins, London, 1775.
- [6] B.N. Ames, L.S. Gold, Environmental pollution, pesticides, and the prevention of cancer: misconceptions, *FASEB J.* 11 (1997) 1041–1052.
- [7] L. Tomatis, J. Huff, I. Hertz-Picciotto, D.P. Sander, J. Bucher, P. Boffeta, O. Axelson, A. Blair, J. Taylor, L. Stayner, J.C. Barret, Avoided and avoidable risks of cancer, *Carcinogenesis* 18 (1997) 97–105.
- [8] D. Krewski, E. Cardis, L. Zeise, V.J. Feron, Empirical approaches to risk estimation and prediction, in: S.H. Moolgavkar, D. Krewski, L. Zeise, E. Cardis, H. Møller (Eds.), *Quantitative Estimation and Prediction of Human Cancer Risks*, IARC Scientific Publications, Lyon, 1999, pp. 131–178.
- [9] B. Hanes, T. Wedel, A selected review of risk models: one hit, multihit, multistage, probit, Weibull, and pharmacokinetic, *J. Am. College Toxicol.* 4 (1985) 271–278.
- [10] J. van Ryzin, Quantitative risk assessment, *J. Occup. Med.* 22 (1980) 321–326.
- [11] A. Kopp-Schneider, Carcinogenesis models for risk assessment, *Statist. Methods Med. Res.* 6 (1997) 317–340.
- [12] S.H. Moolgavkar, D. Krewski, L. Zeise, E. Cardis, H. Møller (Eds.), *Quantitative Estimation and Prediction of Human Cancer Risks*, IARC Scientific Publications, Lyon, 1999.
- [13] ECETOC, Risk assessment for carcinogens, Monograph 24, European Centre for Ecotoxicology and Toxicology of Chemicals, Brussels, 1996.
- [14] G.M. Rand (Ed.), *Fundamentals of Aquatic Toxicology*, 2nd Edition, Taylor and Francis, Washington, 1995.
- [15] J. Timbrell, *Principles of Biochemical Toxicology*, 3rd Edition, Taylor and Francis, London, 2000.
- [16] R.E. Vestal, Aging and pharmacology, *Cancer* 80 (1997) 1302–1310.
- [17] F.P. Guengerich, Metabolism of chemical carcinogens, *Carcinogenesis* 21 (2000) 345–351.
- [18] E.C. Miller, J.A. Miller, Searches for ultimate chemical carcinogens and their reactions with cellular macromolecules, *Cancer* 47 (1981) 2327–2345.
- [19] B.R. Goldin, Intestinal microflora: metabolism of drugs and carcinogens, *Ann. Med.* 22 (1990) 43–48.
- [20] D. Hanahan, R.A. Weinberg, The hallmarks of cancer, *Cell* 100 (2000) 57–70.
- [21] P.C. Nowel, The clonal evolution of tumor cell populations, *Science* 194 (1976) 23–28.
- [22] E.R. Fearon, B. Vogelstein, A genetic model for colorectal tumorigenesis, *Cell* 61 (1990) 759–767.

- [23] J.L. Tsao, Y. Yatabe, R. Salovaara, H.J. Järvinen, J.P. Mecklin, L.A. Aaltonen, S. Tavarè, D. Shibata, Genetic reconstruction of individual colorectal tumor histories, *Proc. Natl. Acad. Sci. U.S.A.* 97 (2000) 1236–1241.
- [24] C.O. Nordling, A new theory on the cancer-inducing mechanism, *Br. J. Cancer* 7 (1953) 68–72.
- [25] K. Kinzler, B. Vogelstein, Gatekeepers and caretakers, *Nature* 386 (1997) 761–762.
- [26] J.G. Herman, A. Umar, K. Polyak, J.R. Graff, N. Ahuja, J.P.J. Issa, S. Markowitz, J.K.V. Willson, S.R. Hamilton, K.W. Kinzler, M.F. Kane, R.D. Kolodner, B. Vogelstein, T.A. Kunkel, S.B. Baylin, Incidence and functional consequences of *hMLH1* promoter hypermethylation in colorectal carcinoma, *Proc. Natl. Acad. Sci. U.S.A.* 95 (1998) 6870–6875.
- [27] Y.E. Whang, X. Wu, H. Suzuki, R.E. Reiter, C. Tran, R.L. Vessela, J.W. Said, W.B. Isaacs, C.L. Sawyers, Inactivation of the tumor suppressor *PTEN/MMAC1* in advanced human prostate cancer through loss of expression, *Proc. Natl. Acad. Sci. U.S.A.* 95 (1998) 5246–5250.
- [28] A.J. Levine, J. Momand, C.A. Finlay, The *p53* tumour suppressor gene, *Nature* 351 (1991) 453–456.
- [29] B. Lewin, Oncogenic conversion by regulatory changes in transcription factors, *Cell* 64 (1991) 303–312.
- [30] P.A. Jones, P.W. Laird, Cancer epigenetics comes of age, *Natl. Genet.* 21 (1999) 163–167.
- [31] D.G. MacPhee, Epigenetics and epimutagens: some new perspectives on cancer, germ line effects and endocrine disrupters, *Mutat. Res.* 400 (1998) 369–379.
- [32] J.M. Zingg, P.A. Jones, Genetic and epigenetic aspects of DNA methylation on genome expression, evolution, mutation and carcinogenesis, *Carcinogenesis* 18 (1997) 869–882.
- [33] J.M. Bishop, Molecular themes in oncogenesis, *Cell* 64 (1991) 235–248.
- [34] T. Hunter, Cooperation between oncogenes, *Cell* 64 (1991) 249–270.
- [35] J. Jen, S.M. Powel, N. Papadopoulos, K.J. Smith, S.R. Hamilton, B. Vogelstein, K.W. Kinzler, Molecular determinants of dysplasia in colorectal lesions, *Cancer Res.* 54 (1994) 5523–5526.
- [36] C. Borek, L. Sachs, The number of cell generations required to fix the transformed state in X-ray-induced transformation, *Proc. Natl. Acad. Sci. U.S.A.* 59 (1967) 83–85.
- [37] W.K. Lutz, Dose–response relationships in chemical carcinogenesis: superposition of differential mechanisms of action, resulting in linear–nonlinear curves, practical thresholds, J-shapes, *Mutat. Res.* 405 (1998) 117–124.
- [38] A.J. Ingram, P. Grasso, Evidence for and possible mechanisms of non-genotoxic carcinogenesis in mouse skin, *Mutat. Res.* 248 (1991) 333–340.
- [39] H. Yamasaki, Role of disrupted gap junctional intercellular communication in detection and characterization of carcinogens, *Mutat. Res.* 365 (1996) 91–105.
- [40] H.S. Rosenkranz, N. Pollack, A.R. Cunningham, Exploring the relationship between the inhibition of gap junctional intercellular communication and other biological phenomena, *Carcinogenesis* 21 (2000) 1007–1011.
- [41] S.F. Velazquez, R. Schoeny, G.E. Rice, V.J. Cogliano, Cancer risk assessment: historical perspectives, current issues, and future directions, *Drug Chem. Toxicol.* 19 (1996) 161–185.
- [42] M.D. Waters, H.F. Stack, M.A. Jackson, Genetic toxicology data in the evaluation of potential human environmental carcinogens, *Mutat. Res.* 437 (1999) 21–49.
- [43] B.N. Ames, L.S. Gold, Too many rodent carcinogens: mitogenesis increases mutagenesis, *Science* 249 (1990) 970–971.
- [44] L.S. Gold, T.H. Slone, B.N. Ames, What do animal cancer tests tell us about human cancer risk: overview of analysis of the carcinogenic potency database, *Drug Metabol. Rev.* 30 (1998) 359–404.
- [45] S.W. Lowe, A.W. Lin, Apoptosis in cancer, *Carcinogenesis* 21 (2000) 485–495.
- [46] I.F. Tannock, Biology of tumor growth, *Hosp. Practice* 18 (1983) 81–93.
- [47] A.F. Ochsenbein, P. Klenerman, U. Karrer, B. Ludwig, M. Pericin, H. Hengartner, R.M. Zinkernagel, Immune surveillance against a solid tumor fails because of immunological ignorance, *Proc. Natl. Acad. Sci. U.S.A.* 96 (1999) 2233–2238.
- [48] M.M. Remedi, E. Hliba, M. Demarchi, M. Depiante-Depaoli, Relationship between immune state and tumor growth rate in rats bearing progressive and non-progressive mammary tumors, *Cancer Immunol. Immunother.* 46 (1998) 350–354.
- [49] R.M. Sutherland, Cell and environment interactions in tumor microregions: the multicell spheroid model, *Science* 240 (1988) 177–182.
- [50] J. Folkman, M. Klagsbrun, Angiogenic factors, *Science* 235 (1987) 442–447.
- [51] R.S. Kerbel, Tumor angiogenesis: past, present and the near future, *Carcinogenesis* 21 (2000) 505–515.
- [52] G.G. Steel, *Growth Kinetics of Tumors*, Clarendon Press, Oxford, 1977.
- [53] G.G. Steel, L.F. Lamerton, The growth rate of human tumours, *Br. J. Cancer* 20 (1966) 74–86.
- [54] P.K. Lala, H.M. Patt, Cytokinetic analysis of tumor growth, *Proc. Natl. Acad. Sci. U.S.A.* 56 (1966) 1735–1742.
- [55] I.F. Tannock, R.P. Hill (Eds.), *The Basic Science of Oncology*, 3rd Edition, McGraw-Hill, New York, 1998.
- [56] M.L. Mendelsohn, The growth fraction: a new concept applied to tumors, *Science* 123 (1960) 1496.
- [57] R.J.B. King, *Cancer Biology*, 2nd Edition, Pearson Education, Harlow, 2000.
- [58] L.A. Liotta, P.S. Steeg, W.G. Stetler Stevenson, Cancer metastasis and angiogenesis: an imbalance of positive and negative regulation, *Cell* 64 (1991) 327–336.
- [59] J. Yokota, Tumor progression and metastasis, *Carcinogenesis* 21 (2000) 497–503.
- [60] M. Kimmel, B.J. Flehinger, Nonparametric estimation of the size–metastasis relationship in solid cancers, *Biometrics* 47 (1991) 987–1004.
- [61] W. Weibull, A statistical distribution of wide applicability, *J. Appl. Mech.* 18 (1951) 293–297.
- [62] S.A.L.M. Koopman, *Dynamic Energy and Mass Budgets in Biological Systems: Theory and Applications*, Cambridge University Press, Cambridge, 2000.

- [63] S.A.L.M. Kooijman, J.J.M. Bedaux, *The Analysis of Aquatic Toxicity Data*, VU University Press, Amsterdam, 1996.
- [64] C.J. van Leeuwen, J.L.M. Hermens (Eds.), *Risk Assessment of Chemicals: An Introduction*, Kluwer Academic Publishers, Dordrecht, 1995.
- [65] M.E. Andersen, Development of physiologically-based pharmacokinetic and physiologically-based pharmacodynamic models for applications in toxicology and risk assessment, *Toxicol. Lett.* 79 (1995) 35–44.
- [66] R.B. Conolly, M.E. Andersen, Biologically-based pharmacodynamic models: tools for toxicological research and risk assessment, *Annu. Rev. Pharmacol. Toxicol.* 31 (1991) 503–523.
- [67] J.J. Mills, M.E. Andersen, Dioxin hepatic carcinogenesis: biologically motivated modeling and risk assessment, *Toxicol. Lett.* 68 (1993) 177–189.
- [68] G.W. van der Molen, A generic toxicokinetic model for persistent lipophilic compounds in humans: an application to TCDD, *Fundam. Appl. Toxicol.* 31 (1996) 83–94.
- [69] M.E. Andersen, K. Krishnan, Physiologically-based pharmacokinetics and cancer risk assessment, *Environ. Health Perspect.* 102 (1994) 103–108.
- [70] D.R. Cox, D. Oakes, *Analysis of Survival Data*, Chapman & Hall, London, 1984.
- [71] J.A. Hoekstra, Estimation of the LD50: a review, *Environmetrics* 2 (1991) 139–152.
- [72] J. Cornfield, Carcinogenic risk assessment, *Science* 198 (1977) 693–699.
- [73] P. Armitage, R. Doll, The age distribution of cancer and a multistage theory of carcinogenesis, *Br. J. Cancer* 8 (1954) 1–12.
- [74] S. Moolgavkar, D. Krewski, M. Schwarz, Mechanisms of carcinogenesis and biologically-based models, in: S.H. Moolgavkar, D. Krewski, L. Zeise, E. Cardis, H. Møller (Eds.), *Quantitative Estimation and Prediction of Human Cancer Risks*, IARC Scientific Publications, Lyon, 1999, pp. 179–237.
- [75] K.S. Crump, D.G. Hoel, C.H. Langley, R. Peto, Fundamental carcinogenic processes and their implications for low dose risk assessment, *Cancer Res.* 36 (1976) 2973–2979.
- [76] D.P. Lovell, G. Thomas, Quantitative risk assessment and the limitations of the linearized multistage model, *Hum. Exp. Toxicol.* 15 (1996) 87–104.
- [77] P. Armitage, R. Doll, A two-stage theory of carcinogenesis in relation to the age distribution of human cancer, *Br. J. Cancer* 9 (1957) 161–169.
- [78] C.W. Chen, Armitage–Doll two-stage model: implications and extension, *Risk Anal.* 13 (1993) 273–279.
- [79] M.P. Little, Are two mutations sufficient to cause cancer? Some generalizations of the two-mutation model of carcinogenesis of Moolgavkar, Venzon, and Knudson, and of the multistage model of Armitage and Doll, *Biometrics* 51 (1995) 1278–1291.
- [80] A.G. Knudson, Mutation and cancer: statistical study of retinoblastoma, *Proc. Natl. Acad. Sci. U.S.A.* 68 (1971) 820–823.
- [81] S.V. Hodgson, E.R. Maher, *A Practical Guide to Human Cancer Genetics*, 2nd Edition, Cambridge University Press, Cambridge, 1999.
- [82] H.J.S. Huang, J.K. Yee, J.Y. Shew, P.L. Chen, R. Bookstein, T. Friedmann, W.H. Lee, E.Y.H.P. Lee, Suppression of the neoplastic phenotype by replacement of the *RB* gene in human cancer cells, *Science* 242 (1988) 1563–1566.
- [83] S.H. Moolgavkar, A. Dewanji, D.J. Venzon, A stochastic two-stage models for cancer risk assessment. I. The hazard function and the probability of tumor, *Risk Anal.* 8 (1988) 383–392.
- [84] S.H. Moolgavkar, D.J. Venzon, Two-event models for carcinogenesis: incidence curves for childhood and adult tumors, *Math. Biosci.* 47 (1979) 55–77.
- [85] S.H. Moolgavkar, Hormones and multistage carcinogenesis, *Cancer Surveys* 5 (1986) 635–648.
- [86] S.H. Moolgavkar, Cell proliferation and carcinogenesis models: general principles with illustrations from the rodent liver system, *Environ. Health Perspect.* 101 (1993) 91–94.
- [87] D.G. Kendall, Birth-and-death processes, and the theory of carcinogenesis, *Biometrika* 47 (1960) 13–21.
- [88] B.E. Henderson, R.K. Ross, M.C. Pike, Toward the primary prevention of cancer, *Science* 254 (1991) 1131–1138.
- [89] S.H. Moolgavkar, A. Dewanji, Biologically-based models for cancer risk assessment: a cautionary note, *Risk Anal.* 8 (1988) 5–6.
- [90] A. Kopp-Schneider, C.J. Portier, A note on approximating the cumulative distribution function of the time to tumor onset in multistage models, *Biometrics* 45 (1989) 1259–1263.
- [91] R.T. Hoogenveen, H.J. Clewell, M.E. Andersen, W. Slob, An alternative exact solution of the two-stage clonal growth model of cancer, *Risk Anal.* 19 (1999) 9–14.
- [92] W.F. Heidenreich, E.G. Luebeck, S.H. Moolgavkar, Some properties of the hazard function of the two-mutation clonal expansion model, *Risk Anal.* 17 (1997) 391–399.
- [93] S.H. Moolgavkar, E.G. Luebeck, Two-event model for carcinogenesis: biological, mathematical and statistical considerations, *Risk Anal.* 10 (1990) 323–341.
- [94] Q. Zheng, On the exact hazard and survival functions of the MVK stochastic carcinogenesis model, *Risk Anal.* 14 (1994) 1081–1084.
- [95] A. Kopp-Schneider, C.J. Portier, C.D. Sherman, The exact formula for the tumor incidence in the two-stage model, *Risk Anal.* 14 (1994) 1079–1080.
- [96] H.J. Clewell III, D.W. Quinn, M.E. Andersen, R.B. Conolly, An improved approximation to the exact solution of the two-stage clonal growth model of cancer, *Risk Anal.* 15 (1995) 467–473.
- [97] A. Dewanji, D.J. Venzon, S.H. Moolgavkar, A stochastic two-stage model for cancer risk assessment. II. The number and size of premalignant clones, *Risk Anal.* 9 (1989) 179–187.
- [98] E.G. Luebeck, S.H. Moolgavkar, Stochastic analysis of intermediate lesions in carcinogenesis experiments, *Risk Anal.* 11 (1991) 149–157.
- [99] A. Dewanji, S.H. Moolgavkar, E.G. Luebeck, Two-mutation model for carcinogenesis: joint analysis of premalignant and malignant lesions, *Math. Biosci.* 104 (1991) 97–109.

- [100] E.G. Luebeck, S.H. Moolgavkar, A. Buchmann, M. Schwarz, Growth kinetics of enzyme-altered liver foci in rats treated with phenobarbital or α -hexachlorocyclohexane, *Toxicol. Appl. Pharmacol.* 130 (1995) 304–315.
- [101] M.C. de Gunst, E.G. Luebeck, A method for parametric estimation of the number and size distribution of cell clusters from observations in a section plane, *Biometrics* 54 (1998) 100–112.
- [102] A. Dewanji, M.J. Goddard, D. Krewski, S.H. Moolgavkar, Two stage models for carcinogenesis: number and size distributions of premalignant clones in longitudinal studies, *Math. Biosci.* 155 (1999) 1–12.
- [103] B. Grasl-Kraupp, E.G. Luebeck, A. Wagner, A. Löw-Baselli, M. de Gunst, T. Waldhör, S.H. Moolgavkar, R. Schulte-Hermann, Quantitative analysis of tumor initiation in rat liver: role of cell replication and cell death (apoptosis), *Carcinogenesis* 21 (2000) 1411–1421.
- [104] T.W. Thorslund, C.C. Brown, G. Charnley, Biologically motivated cancer risk models, *Risk Anal.* 7 (1987) 109–119.
- [105] S.M. Cohen, L.B. Ellwein, Cell proliferation in carcinogenesis, *Science* 249 (1990) 1007–1011.
- [106] R. Schulte-Hermann, W. Bursch, B. Marian, B. Grasl-Kraupp, Active cell death (apoptosis) and cellular proliferation as indicators of exposure to carcinogens, in: D.B. McGregor, J.M. Rice, S. Venitt (Eds.), *The Use of Short- and Medium-term Tests for Carcinogens and Data on Genetic Effects in Carcinogenic Hazard Evaluation*, IARC Scientific Publications, Lyon, 1999, pp. 273–285.
- [107] S.H. Moolgavkar, Biologically motivated two-stage model for cancer risk assessment, *Toxicol. Lett.* 43 (1988) 139–150.
- [108] K.M. Boucher, A.Y. Yakovlev, Estimating the probability of initiated cell death before tumor induction, *Proc. Natl. Acad. Sci. U.S.A.* 94 (1997) 12776–12779.
- [109] G. Charnley, J.D. Wilson, Evaluation of the form of the cell growth rate function of the two-stage model for carcinogenesis, in: B.E. Butterworth, T.J. Slaga, W. Farland, M. McClain (Eds.), *Chemically Induced Cell Proliferation: Implications for Risk Assessment*, Wiley-Liss, New York, 1991, pp. 291–301.
- [110] J.J. Chen, R.L. Kodell, D.W. Gaylor, Using the biological two-stage model to assess risk from short-term exposures, *Risk Anal.* 8 (1988) 223–230.
- [111] L.A. Cox, Extending the stochastic two-stage model of carcinogenesis to include self-regulation of the non-malignant cell population, *Risk Anal.* 12 (1992) 129–138.
- [112] A. Kopp-Schneider, C.J. Portier, A stem cell model for carcinogenesis, *Math. Biosci.* 120 (1994) 211–232.
- [113] Q. Zheng, On the MVK stochastic carcinogenesis model with Erlang distributed cell life lengths, *Risk Anal.* 15 (1995) 495–502.
- [114] J. Denes, D. Krewski, An exact representation for the generating function for the Moolgavkar-Venzon-Knudson two-stage model of carcinogenesis with stochastic stem cell growth, *Math. Biosci.* 131 (1996) 185–204.
- [115] L.G. Hanin, A.Y. Yakovlev, A nonidentifiability aspect of the two-stage model of carcinogenesis, *Risk Anal.* 16 (1996) 711–715.
- [116] W.F. Heidenreich, On the parameters of the clonal expansion model, *Radiat. Environ. Biophys.* 35 (1996) 127–129.
- [117] C.D. Sherman, C.J. Portier, The two-stage model of carcinogenesis: overcoming the nonidentifiability dilemma, *Risk Anal.* 17 (1997) 367–374.
- [118] S.H. Moolgavkar, A population perspective on multistage carcinogenesis, in: C.C. Harris, S. Hirohashi, N. Ito (Eds.), *Proceedings of the 22nd International Symposium of the Princess Takamatsu Cancer Research Foundation on Multistage Carcinogenesis*, Japan Scientific Societies Press, Tokyo, 1992, pp. 381–391.
- [119] W.Y. Tan, C.W. Chen, Multiple-pathway model of carcinogenesis involving one- and two-stage models, in: O. Arino, D.E. Axelrod, M. Kimmel (Eds.), *Mathematical Population Dynamics*, Marcel Dekker, New York, 1990, pp. 469–482.
- [120] W.Y. Tan, *Stochastic Models of Carcinogenesis*, Marcel Dekker, New York, 1991.
- [121] W.Y. Tan, C.W. Chen, Stochastic modeling of carcinogenesis: some new insights, *Math. Comput. Modeling* 28 (1998) 49–71.
- [122] C.J. Portier, A. Kopp-Schneider, A multistage model of carcinogenesis incorporating DNA damage and repair, *Risk Anal.* 11 (1991) 535–543.
- [123] F.Y. Bois, P.J.E. Compton-Quintana, Sensitivity analysis of a new model of carcinogenesis, *J. Theor. Biol.* 159 (1992) 361–375.
- [124] R.B. Conolly, R.H. Reitz, H.J. Clewell, M.E. Andersen, Pharmacokinetics, biochemical mechanism and mutation accumulation: a comprehensive model of chemical carcinogenesis, *Toxicol. Lett.* 43 (1988) 189–200.
- [125] D.G. Hoel, Incorporation of background in dose–response models, *Federation Proc.* 39 (1980) 73–75.
- [126] W.S. Abbott, A method of computing the effectiveness of an insecticide, *J. Econ. Entomol.* 18 (1925) 265–267.
- [127] S. Iversen, N. Arley, On the mechanism of experimental carcinogenesis, *Acta Pathol. Microbiol. Scand.* 27 (1950) 773–803.
- [128] G.L. Yang, A stochastic two-stage carcinogenesis model: a new approach to computing the probability of observing tumor in animal bioassays, *Math. Biosci.* 104 (1991) 247–258.
- [129] E.G. Luebeck, S.H. Moolgavkar, Simulating the process of malignant transformation, *Math. Biosci.* 123 (1994) 127–146.
- [130] C.D. Sherman, C.J. Portier, Calculation of the cumulative distribution function of the time to a small observable tumor, *Bull. Math. Biol.* 62 (2000) 229–240.
- [131] C.D. Sherman, C.J. Portier, A. Kopp-Schneider, Multistage models of carcinogenesis: an approximation for the size and number distribution of late-stage clones, *Risk Anal.* 14 (1994) 1039–1048.
- [132] J.A. Adam, N. Bellomo (Eds.), *A survey of models for tumor-immune system dynamics*, Modeling and Simulation in Science, Engineering and Technology, Birkhäuser, Boston, 1997.
- [133] L. von Bertalanffy, Quantitative laws in metabolism and growth, *Q. Rev. Biol.* 32 (1957) 217–231.

- [134] D. Thompson, *On Growth and Form*, Cambridge University Press, Cambridge, 1961.
- [135] B. Gompertz, On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies, *Philos. Trans. R. Soc.* 115 (1825) 513–585.
- [136] C.P. Winsor, The Gompertz curve as a growth curve, *Proc. Natl. Acad. Sci. U.S.A.* 18 (1932) 1–8.
- [137] A.E. Casey, The experimental alteration of malignancy with an homologous mammalian tumor material, *Am. J. Cancer* 21 (1934) 760–775.
- [138] A.K. Laird, Dynamics of tumor growth, *Br. J. Cancer* 18 (1964) 490–502.
- [139] P.F. Verhulst, Notice sur la loi que la population suit dans son accroissement, *Curr. Math. Phys.* 10 (1838) 113–121.
- [140] V.G. Vaidya, F.J. Alexandro, Evaluation of some mathematical models for tumor growth, *Int. J. Bio-Med. Comput.* 13 (1982) 19–35.
- [141] E.K. Afenya, C.P. Calderón, Diverse ideas on the growth kinetics of disseminated cancer cells, *Bull. Math. Biol.* 62 (2000) 527–542.
- [142] W.V. Mayneord, On a law of growth of Jensen's rat sarcoma, *Am. J. Cancer* 16 (1932) 841–846.
- [143] K. Groebe, W. Mueller-Klieser, Distributions of oxygen, nutrient, and metabolic waste concentrations in multicellular spheroids and their dependence on spheroid parameters, *Eur. Biophys. J.* 19 (1991) 169–181.
- [144] J.P. Ward, J.R. King, Mathematical modelling of avascular-tumour growth, *IMA J. Math. Appl. Med. Biol.* 14 (1997) 39–69.
- [145] R.H. Thomlison, L.H. Gray, The histological structure of some human lung cancers and the possible implications for radiotherapy, *Br. J. Cancer* 9 (1955) 539–549.
- [146] H.P. Greenspan, Models for the growth of a solid tumor by diffusion, *Stud. Appl. Math.* 51 (1972) 317–340.
- [147] D.L.S. McElwain, P.J. Ponzio, A model for the growth of a solid tumor with non-uniform oxygen consumption, *Math. Biosci.* 35 (1977) 267–279.
- [148] C. Sawyer, R. Peto, L. Bernstein, M.C. Pike, Calculation of carcinogenic potency from long-term animal carcinogenesis experiments, *Biometrics* 40 (1984) 27–40.
- [149] J.H. Farmer, R.L. Kodell, D.W. Gaylor, Estimation and extrapolation of tumor probabilities from a mouse bioassay with survival/sacrifice components, *Risk Anal.* 2 (1982) 27–34.
- [150] D.M. Finkelstein, Modeling the effect of dose on the lifetime tumor rate from an animal carcinogenicity experiment, *Biometrics* 47 (1991) 669–680.
- [151] H. Ahn, R.L. Kodell, H. Moon, Attribution of tumor lethality in the absence of cause-of-death information, *Appl. Statist.* 49 (2000) 157–169.
- [152] A.J. Bailer, C.J. Portier, An index of tumorigenic potency, *Biometrics* 49 (1993) 357–365.
- [153] G.E. Dinse, Constant risk differences in the analysis of animal tumorigenicity data, *Biometrics* 47 (1991) 681–700.
- [154] K.L. Meier, A.J. Bailer, C.J. Portier, A measure of tumorigenic potency incorporating dose–response shape, *Biometrics* 49 (1993) 917–926.
- [155] A. Dewanji, D. Krewski, M.J. Goddard, A Weibull model for the estimation of tumorigenic potency, *Biometrics* 49 (1993) 367–377.
- [156] J. van Ryzin, K. Rai, A dose–response model incorporating nonlinear kinetics, *Biometrics* 43 (1987) 95–105.
- [157] W.Y. Tan, K.P. Singh, Assessing the effects of metabolism of environmental agents on cancer tumor development by a two-stage model of carcinogenesis, *Environ. Health Perspect.* 74 (1987) 203–210.
- [158] K.T. Bogen, Risk extrapolation for chlorinated methanes as promoters vs. initiators of multistage carcinogenesis, *Fundam. Appl. Toxicol.* 15 (1990) 536–557.
- [159] R.H. Reitz, M.L. Gargas, M.E. Andersen, W.M. Provan, T.L. Green, Predicting cancer risk from vinyl chloride exposure with a physiologically-based pharmacokinetic model, *Toxicol. Appl. Pharmacol.* 137 (1996) 253–267.
- [160] C.J. Portier, H. El Masri, Statistical research needs in mechanistic modelling for carcinogenic risk assessment, *Statist. Methods Med. Res.* 6 (1997) 305–315.