*Biostatistics* (2003), **4**, 2, *pp.* 231–248
*Printed in Great Britain*

# Age–time patterns of cancer to be anticipated from exposure to general mutagens

DONALD A. PIERCE*
*Department of Statistics, Radiation Effects Research Foundation, Hiroshima 732-0815, Japan*
pierce@rerf.or.jp

MICHAEL VÆTH
*Department of Biostatistics, Aarhus University, Aarhus DK-8000, Denmark*

## SUMMARY

We explore some stochastic considerations regarding accumulation of mutations in relation to carcinogenesis. In particular, we consider the effect of exposure to specific agents, especially ionizing radiation, that may increase mutation rates. The formulation and consequences are a further development of the Armitage–Doll model; both in terms of background cancer where assumptions are substantially weakened, and in terms of the effect of specific mutagenic exposures through generally increasing mutation rates. Under our model the effect of exposure is equivalent to a change in age scale, adding to age a parametric multiple of cumulative dose to the mutagen, which leads to useful formulae for the relative risk. In particular, the excess relative risk at age $a$ behaves approximately as a parametric multiple of the mean dose over ages prior to $a$. These results do not require assuming that some fixed number of mutations are required for malignancy. The implications are particularly useful in providing guidance for descriptive analyses since they have characteristics largely independent of parameter values. It is indicated that the model consequences conform remarkably well to observations from cohort studies of the A-bomb survivors, miners with prolonged exposure to radon, and cigarette smokers who stopped smoking at various ages.

*Keywords*: Armitage-doll model; Mechanistic cancer model; Multistage model; Mutations and cancer; Radiation and cancer; Smoking and cancer.

## 1. INTRODUCTION

We offer some mathematical considerations regarding cellular accumulation of mutations, which seem related to the age–time patterns of excess cancer due to specific acute and prolonged mutagenic exposures. A primary aim is to provide guidance for analyses of cohort data. These analyses are particularly challenging for prolonged exposures, and even for acute exposures it is difficult to distinguish between effects of age at exposure, time since exposure, and attained age. Carcinogenesis is extremely complicated, and we do not consider our formulation as a mathematical model for it. We focus on a fairly general stochastic process involving waiting times for accumulation of mutations in a cell, and the notion that some mutagens may quite generally increase the mutation rates. The generality of the process may implicitly

*To whom correspondence should be addressed

accommodate, to an extent, selective growth advantage of altered cells. Our primary interest is on age–time patterns in the relative increase in cancer rates, rather than on cancer rates themselves. The most notable aspect of the results is provision of insights into age–time patterns of exposure relative risk that are largely independent of parameter values in the detailed mathematical formulation. We do not recommend that the precise results here be used directly in substantive data analyses, but as indicated above only hope that they may provide some guidance in this. Our work is an instance of the suggestion of Aalen and Gjessing (2001) that hazard functions be considered in terms of underlying stochastic processes, and in fact we utilize what they refer to as phase-type models (Aalen, 1995).

The formulation and results are a further development of the Armitage and Doll (1954) multistage model that has for 50 years captured the imagination of statisticians and biologists interested in the age patterns of cancer rates. When considering observed cancer rates without regard to specific causes, the appeal is a possible reason why for most cancers, and for most of lifetime, there is the remarkable tendency for cancer rates to increase as a power of age. Many others have considered the extension of the multistage model to allow for specific carcinogenic exposures. Some of this work is considered further in Section 5, particularly that of Whittemore and Keller (1978), Whittemore (1977), Day and Brown (1980), and Ohtaki (1981), but some brief comment at this point may be useful.

In multistage models for carcinogenesis a stem cell, meaning its lineal descendency for all of lifetime, passes through several stages in becoming malignant. Note that it is a cell, not an organ or person, to which these stages correspond. The cancer rate for an organ or person is the product of the rate for a given stem cell and the relevant number of stem cells (probably roughly on the order of $10^7$ to $10^9$). Alternatively, this can be thought of in an extreme value sense, in terms of the first of many cells to become malignant. Multistage models may or may not allow explicitly for selective proliferation advantage of cells that are beyond some stage. The essential aspects of the Armitage–Doll multistage model are described in that section. The particular multistage model of interest here is illustrated by Figure 1 at the end of that section, where the stages correspond to the accumulation of mutations. A mutation is any cellular change transcending cell division. It is important to understand that much previous thinking has been in terms of stages of a more functional nature, e.g. initiation, promotion, progression. While we are by no means the first to consider the stages specifically as accumulation of mutations, there has been little emphasis on special issues arising from this. The most influential work in multistage models has—very differently from here—emphasized the notion, most sensible in the functional-stage thinking, that a given carcinogen can affect only one or two of the stages when identified in terms of the order of their occurrence, e.g. only initiation or only promotion. This formulation is less appealing when thinking of the stages as accumulation of mutations, and in particular we consider the notion that exposure to a fairly general mutagen might cause any of the required mutations not yet present in the cell.

Pierce and Mendelsohn (1999) considered a closely related formulation along these lines for the acute radiation exposures of A-bomb survivors followed up by the Radiation Effects Research Foundation (RERF) in Hiroshima and Nagasaki. Their results conform interestingly well to the age–time patterns seen in the RERF data for most radiation-related solid cancers. In concert with other more empirical developments (Kellerer and Barclay, 1992; Pierce *et al.*, 1996), this has led to substantial reinterpretation of descriptions of the excess cancer, with more emphasis on attained age and less on age at exposure or time since exposure. The initial aim of the present research was to directly extend the Pierce–Mendelsohn results to multiple and prolonged exposures, but as often happens, addressing a more general problem also led to improved understanding of the simpler one.

In the following section we present the mathematical formulation, results, and arguments for them. Section 3 provides application of the results to three settings: the A-bomb survivor data, lung cancer of miners exposed to radon, and the effects of cessation of cigarette smoking. The remainder of the paper is devoted to further discussion of the issues raised.

## 2. General formulation and result

### 2.1 *Formulation*

In the classical Armitage–Doll formulation a stem cell becomes malignant after passing through $k$ stages, where transitions occur according to a fixed sequence $\lambda_r; r = 1, 2, \ldots, k$ of age-constant transition rates. Their primary result was that the age-specific rate at which a cell becomes malignant is asymptotically, as transition rates uniformly approach zero, given by $\{(k-1)!\}^{-1}\{\prod_{r=1}^{k}\lambda_r\}a^{k-1}$, where $a$ denotes age. We consider the accuracy of this approximation in the Appendix. The cancer rate for a person, or an organ, is the product of this cellular rate and the relevant number of stem cells. Here and elsewhere we use 'cancer rate' to mean the rate at which a malignant cell occurs—there is of course some significant time lag between the existence of a malignant cell and the development of a diagnosable cancer, considered later.

This formulation in terms of a fixed sequence of transition rates is not very suitable when the transitions correspond to accumulation of mutations. The rate of the $r$th mutation in a cell will then be the sum of rates of those that could next occur, and different orderings of occurrence would lead to different $\lambda$-sequences unless all mutations had the same rates (a special case considered by Nordling (1953)). Generally, each rate $\lambda_r$ should thus depend in this manner on the mutational history of the cell up to that time. In dealing with this, we can with no loss extend the Armitage–Doll formulation to a very general age-homogeneous Markov process under which it remains true that cancer rates are asymptotically of form $\mu a^{k-1}$. The state of this process at any age is the mutational status of the cell, and the rate of the next transition is an arbitrary function of the current state, but is otherwise independent of age. This allows not only for different orders of occurrence, but more importantly for the likely possibility that certain mutations effectively increase the rate of subsequent ones by affecting repair mechanisms. The age homogeneity is a strong assumption and we show later that results along lines of this section can be obtained for specified age inhomogeneity.

A notable aspect of our development is that useful results on the relative cancer rates of exposed and unexposed persons arise without the usual assumption, as above, that a certain number $k$ of mutations are required for malignancy. All that is required is that, in some more general manner, the occurrence of malignancy depends only on the mutational state of the cell. Results involving a specific $k$ follow from these more general ones.

The assumption explored here for effects of mutagenic exposure is that an increment of exposure momentarily increases transition rates by a common dose-dependent factor not depending on the mutational status of the cell. That is, for prolonged exposure at age-specific dose rates $d(a)$, the rate $\lambda_r$ of the transition being awaited at age $a$ is increased by the exposure to $\lambda_r[1 + \beta d(a)]$. Note that for a given age the rate $\lambda_r$ is a random quantity given by the sum of rates of mutations that could happen next. The assumption made is slightly weaker than one under which the exposure increases all mutation rates by a common factor. Modification of the results for nonlinearity of the factor in dose rate is straightforward. However, any dose-dependent increase by a factor independent of $r$ is a strong assumption, and it is not our intention to claim that it should hold in any generality. For radiation exposure there is some case for it, but not an ironclad one since radiation may be more effective in causing some mutations than others. Our intention is only to explore the consequences of this assumption, and compare them to what is seen in some data.

Figure 1 illustrates the essence of our model. We recapitulate and amplify on what is involved.

(1) Malignancy of a cell results from accumulation of mutations, perhaps any one of several collections of these. If the chance of a malignant cell developing to a cancer does not depend on age (or exposure) then the *RR* (relative risk) results are unchanged, but for cancer rates themselves this chance would need incorporation.

(2) For the results we consider most important, it is not necessary that the required number of mutations for malignancy is some given number $k$. Some further results require this and we will make clear which ones do.
(3) Without the mutagenic exposure in question, the rate of the next mutation in a cell depends in a given but arbitrary manner on its current mutational status, but not otherwise on age (results under relaxation of the age homogeneity are given in Section 4). The rate of the next mutation in a cell is the sum of rates of those particular mutations which could next occur, the formulation allowing that there may be order restrictions and that rates of particular mutations may depend on which have occurred.
(4) The effect of the mutagenic exposure is that an increment of dose at age $a$, at rate $d(a)$, momentarily increases by a factor $1 + \beta d(a)$ the randomly determined transition rate $\lambda_r$ currently faced by each cell (the results generalize readily when the factor is not linear in dose rate). The essential points are that this increase is multiplicative rather than additive, and that this factor does not depend on $r$.
(5) It presents no problem if some mutations act recessively (must occur in both alleles) and others act dominantly (need only occur in one allele). If it is assumed that some fixed $k$ must occur, then both recessive mutations count towards this. For the recessive case, the mutagenic exposure under consideration would, with dominant probability, cause at most one of the mutations at a given locus.
(6) The model does not explicitly allow for selective proliferation of cells having only some of the required mutations, nor indeed for any proliferation or death prior to malignancy. To some extent, the adaptiveness described in item 3 implicitly allows for this.

### 2.2 *Results*

Under the above assumptions, with malignancy of a cell depending only on some unspecified aspect of its mutational status, we have the following result. Write $r_e(a)$ and $r_u(a)$ for the cancer rates for exposed and unexposed persons (or organs), $D(a)$ for the cumulative dose by age $a$, and consider the change of age scale to $a^* = a + \beta D(a)$. We will see that on the age scale $a^*$ exposed and unexposed persons have (with totally negligible error) the same cancer rates, so on transforming back to the age scale we have that

$$r_e(a) = r_u\{a + \beta D(a)\}\{1 + \beta d(a)\}. \tag{1}$$

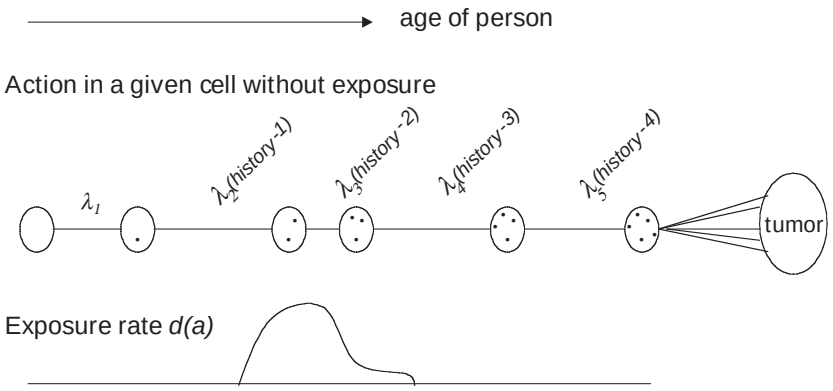
The only approximation involved is the virtual equivalence of the hazard rate and the density function in the very extreme left tail of the distribution of ages at which a given cell becomes malignant (this is not the distribution of ages at which the *first* cell becomes malignant). We are mainly concerned with the relative risk, thus given by

$$RR(a) = \frac{r_e(a)}{r_u(a)} = \frac{r_u\{a + \beta D(a)\}}{r_u(a)}\{1 + \beta d(a)\}. \tag{2}$$

Subsequently to exposure the final factor of course disappears, and for acute exposure the formulae then hold with $D(a)$ being simply the dose.

Relations (1), (2) are interesting in their own right, but more so when something is known about the function $r_u(a)$. In cohort studies this can be estimated without regard to consideration of mechanistic models, and it can be useful to estimate $RR(a)$ in this manner. However, it is often true that an approximation of form $r_u(a) = \mu a^p$ is useful over most of lifetime, leading to the rather empirical approximation

$$RR(a) = \left\{1 + \beta \frac{D(a)}{a}\right\}^p \{1 + \beta d(a)\}. \tag{3}$$




During exposure period, $\lambda_r$ are currently in effect int eh various cells are increased to $\lambda_r[1 + \beta d(a)]$.

Fig. 1. Illustration of model for carcinogenic process both without and with mutagenic exposure. A cell becomes malignant due to accumulation of mutations, not necessarily requiring a certain fixed number of them, and then proliferates to become a tumor. Without exposure, the inter-event times are exponentially distributed with rates $\lambda_r$ that can depend arbitrarily on what mutations have occurred in the cell, indicated by the (history-r) notations. Mutagenic exposure increases the transition rates as indicated, in terms of an exposure rate function that depends on age.

Subsequent to exposure to total dose $D$, acute or prolonged, this becomes

$$RR(a) = \left\{1 + \beta \frac{D}{a}\right\}^p = 1 + p\beta \frac{D}{a} + \cdots \tag{4}$$

and the linear approximation indicated is adequate for doses and ages such that *RR* is no greater than about 3, and for reasonable values of $p$ in the range 4–7. Note that for acute exposures, the *RR* given by (4) does not depend on age at exposure or time since exposure, but only on the attained age. In equation (4) the 'excess' *RR* decreases approximately as $1/a$, and of course the quantity $p\beta$ would be estimated without regard to the value of $p$. We note, however, that the decrease of the dose-linear term in (4) as $1/a$ is sensitive to departures from the approximation of log–log linearity of $r_0(a)$. A moderate decrease in the log–log slope of this at old ages, commonly seen, can result in the dose-linear approximation to $RR(a)$ from (2) to decrease much faster than $1/a$ (roughly like $1/a^2$, for example). We note in the final section, though, that such a decrease in log–log slope may be seen for populations, due to heterogeneity and selection, when it does not exist for individuals and thus it may be inappropriate to focus on this in mechanistic models.

When it is further assumed that $k$ mutations are required for malignancy, we will show that asymptotically $r_{\mathrm{u}}(a) = \mu a^{k-1}$, so the result becomes

$$RR(a) = \left\{1 + \beta \frac{D(a)}{a}\right\}^{k-1} \{1 + \beta d(a)\} \tag{5}$$

and so forth as in (4). The asymptotic approximation for $r_{\mathrm{u}}(a)$, given the assumptions made if not in reality, is usually adequate for mutation rates compatible with cancer rates. But due to cancellation of errors in the numerator and denominator cancer rates involved, the asymptotic approximation (5) is even better than for the rates themselves.

### 2.3 *Arguments*

We now give the argument for (1). Write $\mathbf{s} = \{s_1, \ldots, s_r, \ldots\}$ for the sequence of mutational states of a cell while awaiting mutations $1, \ldots, r, \ldots$. Until stated otherwise, the argument is conditional on $\mathbf{s}$. Write $\boldsymbol{\lambda}_\mathbf{s} = \{\lambda_{1,s_1}, \ldots, \lambda_{r,s_r}, \ldots\}$ for the sequence of transition rates for an unexposed cell, which are by the age-homogeneous Markov process assumption determined by the state vector $\mathbf{s}$ and otherwise independent of $a$. Now suppose that malignancy of a cell results in some way from the mutations considered, but not necessarily upon the occurrence of the $k$th one. Then the age-specific cancer rate for the unexposed cell is determined by $\boldsymbol{\lambda}_\mathbf{s}$, a function of age written as $r_{\mathrm{u},\mathbf{s}}(a; \boldsymbol{\lambda}_\mathbf{s})$. For an exposed cell the transition rates become $\boldsymbol{\lambda}_{\mathrm{e},\mathbf{s}}(a) = \{\lambda_{1,s_1}[1 + \beta d(a)], \ldots, \lambda_{r,s_r}[1 + \beta d(a)], \ldots\}$. On the age scale $a^* = a + \beta D(a)$ the transformation rates become $\boldsymbol{\lambda}^*_{\mathrm{e},\mathbf{s}}(a^*) = \boldsymbol{\lambda}_{\mathrm{e},\mathbf{s}}(a)\mathrm{d}a/\mathrm{d}a^* = \boldsymbol{\lambda}_\mathbf{s}$. Note that this involves not only the differential element canceling the factors $[1 + \beta d(a)]$ but also the age-homogeneity of the transition rates for the unexposed cell. Thus the cancer rate function for an exposed cell on the transformed age scale is $r^*_{\mathrm{e},s}(a^*) = r_{\mathrm{u},s}(a^*; \boldsymbol{\lambda}_\mathbf{s})$, and the cancer rate function for an exposed cell on the original age scale is thus $r_{\mathrm{e},\mathbf{s}}(a) = r_{\mathrm{u},\mathbf{s}}(a^*; \boldsymbol{\lambda}_\mathbf{s})\mathrm{d}a^*/\mathrm{d}a = r_{\mathrm{u},\mathbf{s}}\{a + \beta D(a); \boldsymbol{\lambda}_\mathbf{s}\}\{1 + \beta d(a)\}$. Given this, the only approximation for the result leading to (1) is that unconditionally $r_\mathrm{u}(a) \sum r_{\mathrm{u},\mathbf{s}}(a; \boldsymbol{\lambda}_\mathbf{s})\mathrm{pr}(\mathbf{s})$. Such an averaging is exact for densities rather than rates, but these are hazard rates in the extreme left tail of the distribution of cancer ages for individual cells, where the hazard rate and density are virtually the same.

This establishes (1), and for (5) we now turn to the asymptotic result that under the assumptions for cancer without exposure, now assuming that a cell is malignant upon the $k$th mutation, $r_\mathrm{u}(a) = \mu a^{k-1}$. It is precisely the Armitage–Doll result that, conditionally, $r_{\mathrm{u},\mathbf{s}}(a) = \mu_\mathbf{s} a^{k-1}$. Again, since densities and rates are virtually identical then have that $r_\mathrm{u}(a) = \sum r_{\mathrm{u},\mathbf{s}}(a; \boldsymbol{\lambda}_\mathbf{s})\mathrm{pr}(\mathbf{s}) = \left(\sum \mu_\mathbf{s}\mathrm{pr}(\mathbf{s})\right) a^{k-1} = \mu a^{k-1}$. The combination of this result and (2) leads to (5).

Of course there is some delay, or latent period, between existence of a malignant cell and a cancer, and most work in this area gives at least passing attention to this. More importantly, it is likely that a substantial fraction of malignant cells fail to develop into a cancer, something that would have a large effect on cancer rates as considered above but would largely cancel out in the relative risk. The variable latent period can have a particularly notable effect on the above results, due to the factor $[1 + \beta d(a)]$ in the relative risk formulae. In particular, an essentially acute exposure leads to a 'spike' in the relative risk there, and smoothing this out with a variable lag time is important. Although, strictly speaking, the smoothing should be done on the cancer rates in the numerator and denominator of relative risks, we find that smoothing of the *RR* formulae themselves is usually quite adequate.

## 3. EXAMPLES

### 3.1 *A-bomb survivor data*

The data used here pertain to cancer incidence during 1958–94, from the follow-up by RERF. Documentation of similar data on cancer mortality and incidence, and the basic statistical approach, is given in Pierce *et al.* (1996) and Thompson *et al.* (1994). We follow Pierce and Mendelsohn (1999) in considering the class of major, non-sex-specific cancers, excluding also thyroid cancer: namely cancers of the stomach, lung, liver, colon, rectum, gallbladder, pancreas, bladder and esophagus. The omissions are made because cancers that are hormonally related often have distinctive age patterns. For this class of cancers there are about 8000 cases for the current follow-up, of which about 400 are estimated as radiation related. About 45 000 of the cohort of 80 000 utilized here have positive dose estimates. Restricting to those, the quartiles of the dose distribution are about 0.01, 0.05, and 0.18 Sv. Of the 45 000 about 10% and 5%, respectively, had doses greater than 0.50 and 1.0 Sv. (Note: A seivert (Sv) is a rather large radiation dose that for whole-body exposure would require medical attention for acute effects. Annual limits for nuclear workers

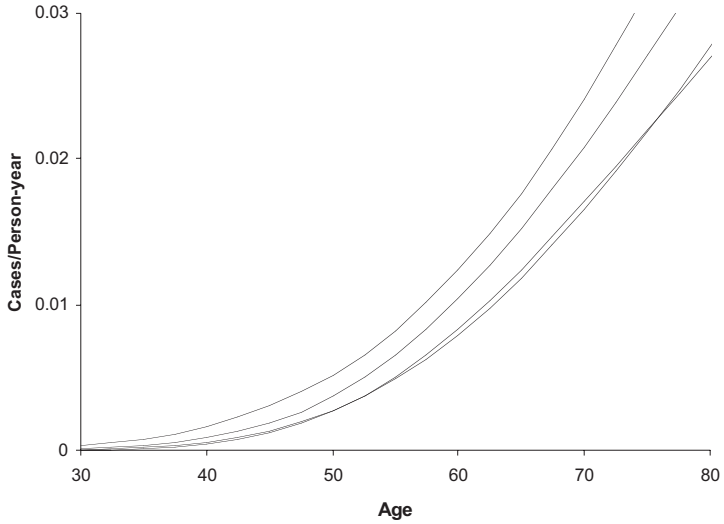



Fig. 2. Male age-specific cancer rates for four dose categories with cut-points 0.005, 0.5, and 1 Sv, to provide perspective for Figure 3. Higher lines correspond to higher dose categories. This is for the pooled nine cancer sites described in the text.

are 0.02 Sv with further limitations so that workers would not ordinarily accumulate 0.10 Sv). The excess cancer risk is very linear in dose. The currently standard analytical methods used at RERF, differing from those used in Pierce *et al.* (1996) only to allow for age trends in the relative risk, provide estimates of the excess relative risk ($ERR = RR - 1$) at age 60 of 0.41/Sv for males and 0.72/Sv for females, for any prior age at exposure. This sex difference is largely an artifact of the relative risk representation. That is, males have substantially larger background cancer rates than females during most of life, and as indicated in Pierce *et al.* (1996) the radiogenic absolute increase in cancer rates is about the same for the sexes.

We first explore the adequacy of the result in equation (1) that the effect of exposure is equivalent to an age increment linear in dose. Reference to equations (2)–(5) indicates that for reasons discussed above, this can only be adequate when the dose coefficient depends on sex. Estimates of these coefficients can be obtained as values that, when actually applied for the age increments, result in null estimates of *ERR*/Sv in standard analyses referred to above. This fitting inevitably results in much of the dose effect disappearing on the transformed age scale, and the question is whether careful analysis can show any residual effects at all of radiation exposure. This should include both looking for such effects due to dose level, and to age at exposure.

Figure 2 shows age-specific cancer rates for men in four dose categories indicating clearly the radiation effect; the pattern for women is similar but at lower levels. Figure 3 shows the corresponding plot for both sexes on the transformed age scale, where we see that the dose effect is completely removed. That plot ignores age at exposure, usually considered to be an important factor. Figure 4 shows by age-at-exposure categories' cancer rates on the transformed age scale, without regard to dose, compared to those for unexposed persons. The pattern seen is birth cohort variation in background cancer rates, and for each age-at-exposure group the radiation effect vanishes on the transformed age scale. This analysis indicates that there is virtually no systematic departure from the predictions of equation (1). For all these plots cancer rates are fitted as quadratic functions on the log–log scale. Results of this aspect of the analysis are similar when all solid cancers are used, without the omissions made here.

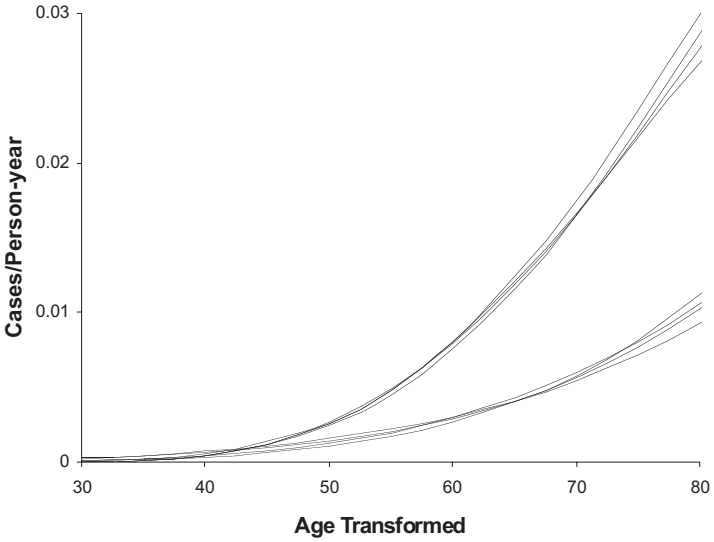



Fig. 3. Male and female age-specific cancer rates on transformed age scales, for the four dose categories of Figure 2. The four upper curves are for males, with age scale $age + 4.1\,dose$, and lower curves are for females with age scale $age + 6.4\,dose$. The dose-dependent age transformation removes any indication of a remaining radiation effect

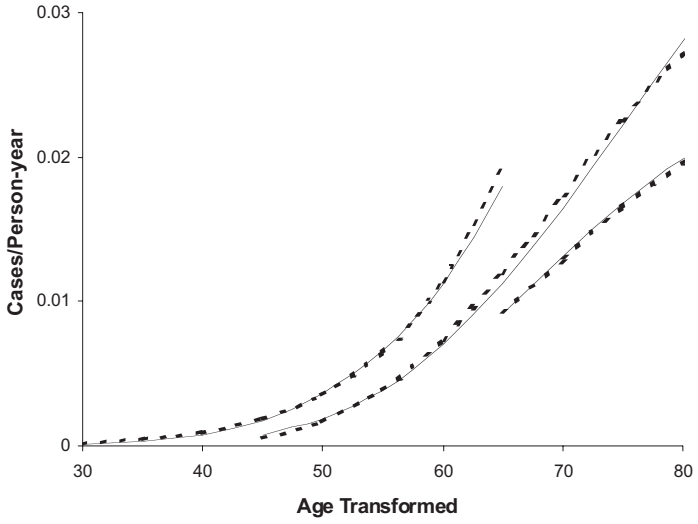



Fig. 4. Comparison of cancer rates on the transformed age scale, all doses, with those for unexposed on the original age scale, for age-at-exposure categories $< 20, 20 - 40, > 40$. There are thus two curves for each age-at-exposure category, with age ranges over follow-up indicating these categories. There is a substantial birth cohort effect for the unexposed, but no other effect for the exposed on the transformed age scale.

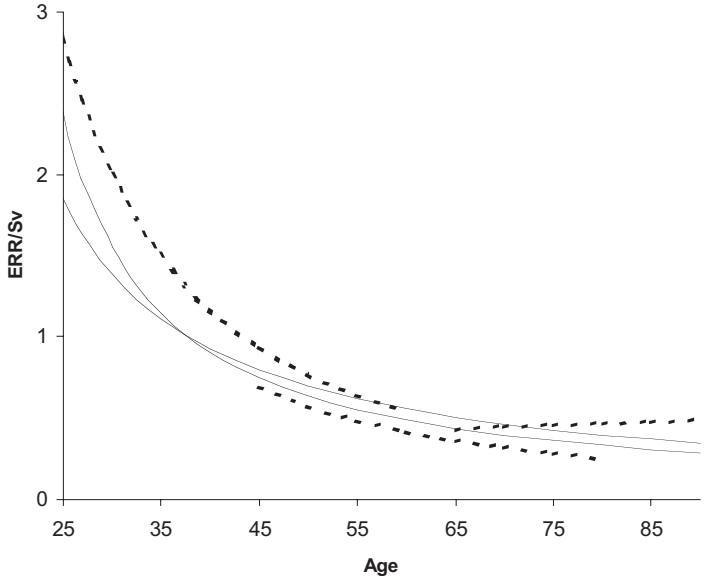



Fig. 5. Description of *ERR* and model fit for the 9-site collection of cancer types, averaging over sex. The three dashed lines are empirically estimated age-specific *ERR*s over the follow-up for ages at exposure 10, 30, and 50. The steeper solid curve is the fit of equation (4) with $p = 4.5$, and the other is the fit to the first term in expansion of that, decreasing as 1/age. There is no statistically significant departure from the model predictions, but possible reasons for the apparent lack of fit at young ages are discussed in the text.

We now turn to the adequacy of *RR* age patterns suggested by equation (4). Some background will clarify our approach to this. At least until recently, it has been generally considered that the *ERR* for A-bomb survivors decreases substantially with increasing age at exposure. The *ERR* for those exposed as adults has been considered essentially constant in age, while that for those exposed as children is initially higher but decreases with age—although only recently has this decrease become statistically significant. Thus it is useful to provide empirical descriptions of the *ERR* over the follow-up within age-at-exposure groups, for comparison to the fit of relations with no age-at-exposure effect derived in this paper. The empirical description used here is along lines commonly used at RERF, using rather minimal smoothing through an *ERR* model of form $\beta_{\text{sex}} d \exp[\theta e + \alpha_j + \delta_j \log(a)]$, where $j = 1, 2, 3$ denotes age-at-exposure categories $< 20, 20 - 40, > 40$ years and $e$ denotes age at exposure. For all analysis here, sex-specific parameters are used with results presented as the average over sex.

Figure 5 shows these results, where semi-parametric relative risk regression (Cox regression with age-dependent covariates, using the *ERR* function above) is used to fit all the models. Thus, as we prefer, the asymptotic result $\mu a^{k-1}$ for background rates plays no primary role in this analysis. Two fits to the data related to equation (4) are shown: one using the full formula with $p = 4.5$ and the other only the first term of the expansion, where specification of $p$ is not required. Recall that age at exposure has no effect in (4). The extent of agreement between the two curves from (4) indicates that the *RR* is insensitive to the value of $p$, so this cannot be usefully estimated from analysis of only the *RR*. The value $p = 4.5$ was estimated by Pierce and Mendelsohn (1999) in analysis of a previous version of these data that considered background cancer rates for ages 35–80 as well. We have adapted equation (4) to allow for a fixed latent period of 7.5 years between malignancy and cancer. This is roughly the maximum likelihood estimator, but note that the follow-up begins 12 years after exposure so these data provide little information on that issue and the choice used here has very minor effect.

There is no statistically significant lack of fit for results based on equation (4), in comparison to the statistical model used for the empirical description ($\chi^2 = 4$ on 6 df). That is, although the empirical description has somewhat higher *ERR* than the model fit for the youngest age-at-exposure group, this is quite imprecisely estimated since background rates are so small for that group. The more rapid decline with age in the empirical description than from the model prediction may be partly due to the point raised following equation (4) regarding departures from log–log linearity of background rates. As discussed in the final section, lack of fit for that reason might not be an undesirable aspect of a model. However, we note that for analysis of all solid cancers together without the selection made here, there is a larger and significant lack of fit to the model results, in the sense of a residual age-at-exposure effect. The issues regarding an age-at-exposure effect in the *RR* are complex, since this tends to reflect birth cohort variations in background rates and is not a straightforward modifier of the actual radiation effect. It is shown in the following section that departures from the age homogeneity of our model would introduce age-at-exposure effects in the *ERR*. Thus for various reasons the traditional age-at-exposure considerations referred to above remain important, but results here do suggest emphasis on generally age-declining *RR* and possible biological reasons for this.

Kai *et al.* (1997) considered a very different two-stage mathematical model for A-bomb survivor data, where cells undergo only two mutational transformations, with proliferative advantage for transformed cells following the first. Their model, without specification of parameter values, is less predictive of the general character of age–time patterns than results here, but it can be fitted to the data reasonably well.

### 3.2 *Underground miners exposed to radon*

Two committees of the National Research Council (1988, 1999) and also a National Cancer Institute working group (Lubin *et al.*, 1994, 1995) have done extensive analysis of pooled data on lung cancer from several large cohorts of underground miners exposed to radon. Because exposures were prolonged, the description of age–time patterns of excess risk has been challenging, but substantial progress has been made. A primary characteristic of the descriptive modeling has been to relate the age-specific *ERR* to the dose within each of several 'windows' of time prior to the age at risk, meaning that time since exposure has been a primary focus. The models considered also allow for a general decrease of the *ERR* with attained age, and the selected one an effect of exposure duration aside from accumulated dose. That used here is referred to as the exposure-age-duration model, in Table 3-3 of National Research Council (1999), which differs only in minor changes of parameter values from that presented on page 821 of Lubin *et al.* (1995).

The approach illustrated in Figures 6–8 is to graph the age–time pattern of *ERR* given by this descriptive model, for several typical exposure scenarios, and compare to this the results of equation (3), adapted to allow for a random latent period. The dose rate is taken in terms of the conventional measure for the descriptive model, something called *WLM*/year. We have not attempted to fit our model to the actual data (not publicly available), which seems unnecessary for our purposes. There are three choices involved in implementing the model: the values of $p$ and $\beta$, and a distribution for the latent period. Results are insensitive to the value of $p$, taken here as 5, and $\beta$ was chosen based on more exposure scenarios than shown here. In contrast to the previous example, the choice of a distribution for the latent period is here important, since we are interested in cancer rates during the exposure period. The distribution used for the latent period, chosen roughly to obtain a generally best fit to the descriptions, is as follows:

| years | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| density | 1 | 2 | 3 | 4 | 5 | 4 | 4 | 3 | 3 | 2 | 2 | 1 |

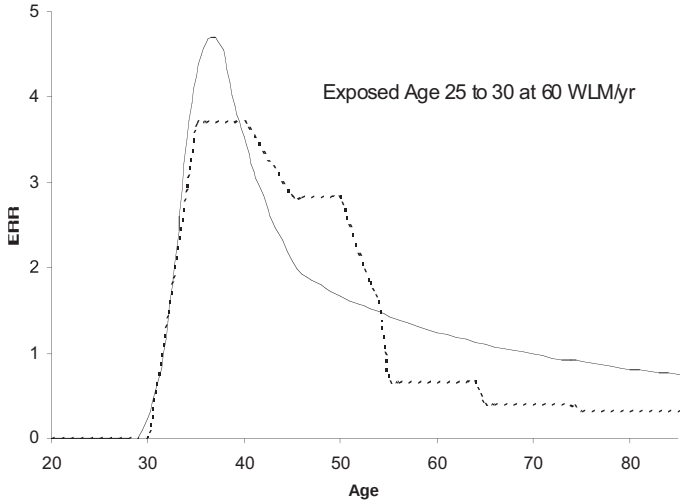



Fig. 6. Empirical description and model fit for lung cancer among miners with prolonged radiation exposure, under exposure scenario indicated. Dashed curve is from the application of the empirical model of the NRC committee, solid curve is fitted using results of this paper.

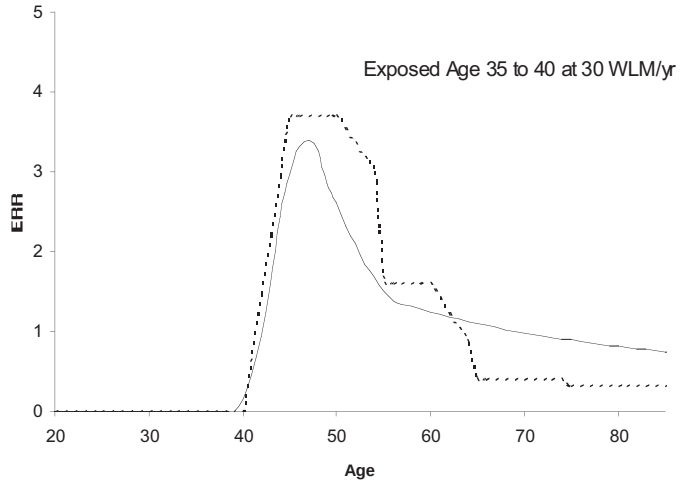



Fig. 7. Results as in Figure 6 for the indicated alternative exposure scenario.

The primary effect of this distribution pertains to the final factor in equation (3), which we note later is the effect of the exposure causing the final required mutation. Considering the plots, the rapid increase in the *ERR* a few years after start of exposure, and the correspondingly rapid decrease ending 15 years after exposure ends, are largely the result of lagging and smoothing the effect of the exposure causing the final mutation. The noticeable change in the decline at 15 years after the end of exposure is where the effect of that final factor has disappeared, and the *ERR* then decreases roughly as 1/age.

The correspondence of model predictions and empirical results is remarkable, even though there is some systematic overprediction of *ERR* at older ages. Again, this could be partly due to the point raised

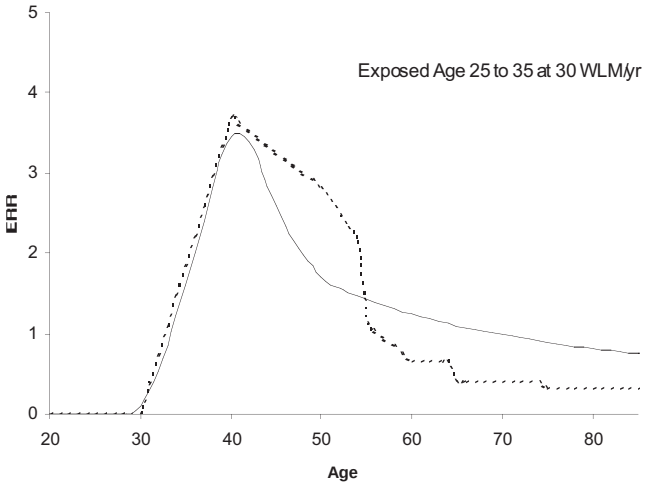



Fig. 8. Results as in Figure 6 for the indicated alternative exposure scenario.

following equation (4) regarding departure from log–log linearity of background rates, and as discussed in the final section, lack of fit for that reason might not be an undesirable aspect of a model. The empirical description, evolving from earlier developments in National Research Council (1988), was quite difficult to obtain without guidance regarding what to expect, and is subject to considerable statistical error. It is interesting to consider what description would have resulted had there been less emphasis on time since exposure, as suggested by results here.

Among others who have analysed related data, Luebeck *et al.* (1999) and Hazelton *et al.* (2001) fitted a very different two-stage model, similar to that of Kai *et al.* (1997) referred to above, for two of the cohorts involved (Colorado and Chinese). The most noteworthy feature of these analyses was finding an apparent radiation effect on the rate of cellular proliferation that was important in comparison to its effect on mutation rates. Although such an effect cannot be evaluated within our formulation, the results seen here suggest that such considerations may not be necessary to explain the data. Incidentally, they and others have estimated a mean latent period for lung cancer similar to that of the distribution chosen here.

### 3.3 *Effects of cessation of cigarette smoking*

The American Cancer Society follows up a cohort of about a million persons in regard to lung cancer mortality and cigarette smoking. Halpern *et al.* (1993) present in their Table 4 age-specific lung cancer death rates for non-smokers, continuing smokers, and those who stopped smoking at various ages, ratios of which are plotted as the points in Figure 9. We note that the follow-up used is only from 1982–88, so the age variation is largely that for different birth cohorts. The age-specific rates for nonsmokers are quite linear on a log–log plot, with slope $p = 5$. The curves in Figure 9 are computed according to equation (3), incorporating a fixed latent period $t_0 = 8$ years, with $p = 5$, and $\beta d = 0.9$ during the smoking ages assumed to begin at 18 years. There is substantial time lag between cessation of smoking and evaluation of cancer risk, and the role of the latent period is less important than in the previous example. Taking $t_0 = 5$ improves the fit slightly, but the larger value was taken to conform to the previous examples. The value of $\beta d$ was chosen for an overall best fit to the points in the plot. Results are insensitive to the value of $p$. There is no doubt statistically significant lack of fit, but to evaluate this would require more detailed analysis. However, it is remarkable that such a simple model with essentially only one parameter could predict the risk patterns this well.

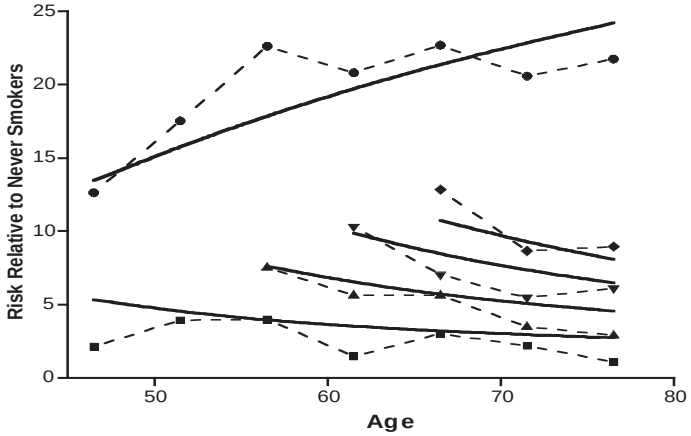



Fig. 9. Empirical description and model fit for lung cancer risks of continuing smokers, and those stopping at ages (rounded) of 35, 45, 52.5, and 57.5. Points linked by dashed lines are from ACS cohort data. The solid lines are fitted using the model of this paper, involving essentially the fitting of just one parameter.

## 4. Further comments on theoretical results

In our main results the *RR* is a polynomial in dose, and the terms have interesting interpretation not evident from the derivation given. Under the assumption that malignancy occurs upon the $k$th mutation, the arguments of Section 2 show that after termination of exposure to total dose $D$ the excess absolute malignancy rate is given by

$$EAR(a) = \mu\left[\{a+\beta D\}^{k-1} - a^{k-1}\right] = \mu(k-1)\beta D a^{k-2} + \cdots . \tag{6}$$

It can be shown that the terms in this expansion correspond to the effect of the exposure having caused $1, 2, \ldots, k-1$ mutations in the same cell. Although the terms decrease rapidly for moderate doses, it is interesting that those beyond the first may not be totally negligible. The underlying idea is that without exposure malignancy occurs after $k$ exponential waiting times resulting in cancer rate increasing as $a^{k-1}$, and for cells where exposure causes mutations eliminating $j$ of the waiting times, the malignancy rate increases as $a^{k-1-j}$. Pierce and Mendelsohn (1999) derived the first term in (6) and we initially derived subsequent terms, noticing that they corresponded to the expansion above before arriving at the much simpler argument of Section 2. When exposure has not ended, (6) is multiplied as in (1) by $[1+\beta d(a)]$, arising in the earlier argument as a differential element $\mathrm{d}a^*/\mathrm{d}a$, but in considerations of this paragraph corresponding to the exposure causing the $k$th and final mutation.

We noted that the essential results here on age–time patterns of risk carry through when the factor $[1+\beta d(a)]$ is replaced by $[1+\beta g\{d(a)\}]$ for any function $g(\cdot)$. The only change is that $D(a)$ in our results is replaced by the integral of $g\{d(a)\}$. But there is an equally tractable extension in another direction. There is substantial current interest in the idea that an increment of exposure might cause some long-lasting instability in a cell, resulting in increased mutation rates for some extended period. All results here can be applied directly to such situations by defining $d(a)$ not at the dose rate at age $a$, but a measure of the instability effect extending beyond the increment of exposure.

The results can also be generalized to allow for a certain type of non-age-homogeneity in background mutation rates. Let the age-homogeneous transformation rates $\lambda_r$ be replaced by $\lambda_r s(a)$, where $s(a)$ is an arbitrary function representing age-specific variation in mutation rates. Then essentially the same

argument as before shows that on the transformed age scale $a^* = \int^a s(e)[1 + \beta d(e)]\mathrm{d}e$, exposed and unexposed cells have the same cancer rates. Thus (3) is replaced by

$$RR(a) = \left\{ 1 + \beta \frac{\int^a s(e)d(e)\mathrm{d}e}{S(a)} \right\}^p \{1 + \beta d(a)\}$$

where $S(a) = \int^a s(e)\mathrm{d}e$.

For an acute dose $D$ at age $e$ the relative risk subsequent to exposure is then given by

$$RR(a) = \left\{ 1 + \beta \frac{s(e)D}{S(a)} \right\}^p,$$

the crucial point being that the *RR* now depends on age at exposure $e$. Although we have not explored the matter fully, this may be useful for cancer types excluded here for the A-bomb survivors, where there are indeed age-at-exposure effects. For prolonged exposures the non-age-homogeneous extension is related to considerations of Pike *et al.* (1983). They showed that for female breast cancer the Armitage–Doll result applies better on a transformed age scale corresponding to the number of menstrual cycles.

## 5. DISCUSSION

Many others have considered the extension of the Armitage–Doll model to include effects of specific carcinogenic exposures. Whittemore and Keller (1978), followed by Day and Brown (1980), considered fixed stage transition rates $\lambda_r, r = 1, \ldots, k$, taken as $\lambda_r = v_r + \delta_r d(a)$. Their general results were very complicated, and the key to the simplicity of ours is the further assumption that $\delta_r = v_r \beta$: that is, representing the effects as multiplicative rather than additive. Ohtaki (1981, 1985) did capitalize on this assumption, obtaining results in the direction of ours. However, he did not allow transition rates to depend on the mutational history of the cell, nor did he consider the simplifications arising from focus on the relative risk.

The most influential results of Whittemore (1977), Whittemore and Keller (1978) and Day and Brown (1980), considered by many others including Thomas (1982), were for the case of $\lambda_r = v_r + \delta_r d(a)$, where $\delta_r d(a) \ll v_r$ except for one value of $r$. Thus, characteristic age patterns of excess cancer rates were derived under assumptions that the carcinogen in question affected only an early stage, a late stage, and so forth. In this regard, there was largely misguided emphasis, initiated by Armitage and Doll (1954) and perpetuated by most workers, on the notion that if $\delta_r d(a)$ were substantial in relation to $v_r$ for $m$ values of $r$, then the cancer dose response would involve an $m$th-degree polynomial in dose. That this seems untenable for values of $m$ larger than about 2 was taken to suggest that in the multi-stage model a given carcinogen is unlikely to affect more than one or two transitions. The flaw in this reasoning is that what is formally an $m$th-degree polynomial in dose can in fact be dominantly quite linear, as noted in regard to our results.

The major aspect of current views of carcinogenesis not explicitly included in the idealized formulation here pertains to proliferative advantage of cells having acquired some, but not all, of a collection of mutations required for full malignancy. Generalization of our approach to allow explicitly for this seems intractable. However, we believe that to some approximation our results allow implicitly for modest intermediate cell proliferation. Proliferation of altered cells is largely equivalent to increased mutation rates (for clones) and we do allow mutation rates to depend arbitrarily on the current state of each cell. We note that our analysis applies exactly if, as an approximation, proliferation prior to full malignancy is considered only to occur immediately upon the penultimate of the required mutations (not

requiring that this be some fixed number). Note further than when carcinogenesis involves accumulation of several mutations, proliferative clones of interest are unlikely to become large, since cell ceases to be a clone member when it acquires a further mutation. The alternative models referred to in the examples of Kai *et al.* (1997) and Luebeck *et al.* (1999), follow work of Moolgavkar and various co-workers to allow explicitly, in a highly idealized manner, for proliferation following a single mutagenic transformation.

It is well known that the form $\mu a^p$ for population cancer rates is often inadequate when the age range includes the very young and very old, and in particular there is very often seen a substantial reduction in the log–log slope at old ages. But this is to be expected for population cancer rates even if a relation $\mu a^p$ holds for individuals, with heterogeneity in the values of $\mu$. That is, as age increases the surviving population includes an increasing fraction of those with smaller $\mu$ values. Mechanistic cancer modeling is concerned with cancer rates for individuals, rather than for heterogeneous populations, and modeling driven by aims to accommodate the reduction of log–log slope at old ages may be counterproductive.

We conclude with some discussion of our attitude towards the value of results such as here, and about mathematical modeling more generally. Carcinogenesis is very complicated, and a formulation such as used here ignores not just details but generally important aspects. However, science is not primarily the analysis of complexity. As pointed out in Pierce and Mendelsohn (1999), our attitude towards models is that they should usually be highly idealized, not attempting to reflect many of the actual complexities of the process, or even of broadly considered data. What is important about a model is that it be useful, rather than that it reflect a complex reality all that faithfully; indeed its usefulness is often inversely related to its complexity. It may provide new insights into data, providing a useful framework for further thought, and suggest interesting consequences that can be verified more directly—thus perhaps transcending the validity of the model.

In particular, we think that one would almost never want to fit to observed data, for substantive descriptive purposes as opposed to evaluating a model, the relations derived in this paper. But there is need for guidance in descriptive analyses, which are not easily done even for acute exposures and become truly daunting for prolonged exposures. We think most would agree with us that the best aim is some hybrid approach involving empirical description along with what seems useful from mathematical modeling. Further, the most important issues often involve interpretation of descriptions, where one is faced with the need to sort out causality in the presence of too many highly correlated age–time–dose covariables. Results here have had substantial effect in this sense on work at RERF, contributing to much greater emphasis on relative risks generally declining with age.

## ACKNOWLEDGEMENTS

This publication is based on research performed at the Radiation Effects Research Foundation (RERF), Hiroshima and Nagasaki, and the University of Aarhus, Denmark. RERF is a private, nonprofit foundation funded equally by the Japanese Ministry of Health, Labor, and Welfare and the US Department of Energy through the National Academy of Sciences. We are grateful to Megu Ohtaki for helping us to appreciate the generalization mentioned at the end of Section 4. The authors also thank Dale L. Preston for his encouragement and support.

## APPENDIX

Although our formulation for background carcinogenesis is much more general than the Armitage–Doll process, we do use the Armitage–Doll approximation in the argument and it is useful to consider its validity. In regard to our more general model, we note that Aalen (1995) has formally investigated a class of time-homogeneous Markov processes, called phase-type models, that includes ours as a special case. He does not consider asymptotics as transition rates approach zero, but does provide results useful for this.


Fig. A1. Exact result and Armitage–Doll approximation (upper line) for the rate of malignancy under mutation rates of the Moolgavkar example given in the appendix text. The log–log slopes are $k1 = 6$ and 5.8, respectively.

For fixed transition rates $\lambda_1, \ldots, \lambda_k$ it is a classical result that the exact density of the time until the $k$th event, i.e. the sum of $k$ exponential variates, is

$$f(a; \lambda_1, \ldots, \lambda_k) = \lambda_1 \ldots \lambda_k (-1)^k \sum_{j=1}^{k} \frac{\exp(-\lambda_j a)}{\prod_{i \neq j} (\lambda_i - \lambda_j)}$$

with limiting values used when some of the rates are equal. This is the result from which Armitage (1953) proved the standard asymptotic result. It is straightforward to calculate the hazard function, which here we refer to as the cancer rate, ignoring both the latent period and more seriously the chance that a malignancy does not result in cancer. For computation of the above, recursive relations based on the theory of divided differences are useful, particularly when some of the mutation rates are equal. For exposed cells the exact formula can be used through the change of age scale argument. Although such formulae are not difficult to use in calculating cancer rates for given ages, they are not very useful in understanding the age patterns of excess risk. The asymptotic result provides much more insight, the exact result being mainly useful for evaluating the adequacy of the asymptotics.

In a strict sense, the Armitage–Doll approximation $\{(k-1)!\}^{-1} \left\{ \prod_{r=1}^{k} \lambda_r \right\} a^{k-1}$ can easily be inaccurate for relevant values of transition rates. This has been stressed by Moolgavkar (1978), who considered $k = 7$ with transition rates $\lambda$ of 1, 2, 3, 34, 70, 80, 90, all times $10^{-4}$, leading roughly to human cancer rates. In this case Armitage–Doll approximation overestimates cancer rates by 17%, 27%, and 37% at ages 40, 60, and 80. The heterogeneity of transition rates is an important factor: about 1/3 of the error here is due to the size of the rates, and 2/3 is due to their heterogeneity. That is, the Armitage–Doll approximation (which depends only on the geometric mean of the transition rates) has about 1/3 the error indicated above if the $\lambda$ are taken as all equal to the geometric mean of those used here.

However, for our purposes this type of error is essentially irrelevant. We have no interest in the relation between specific transition rates and cancer rates, but only a relation $\mu a^{k-1}$ where $\mu$ as defined in Section 2.3's is a much more complicated function than in the Armitage–Doll result above, and either cancels in the relative risk or is estimated from cancer rate data. Further, our concern is with the log–log linearity of the relation, rather than whether the exponent is slightly different from some $k - 1$. Figure A1 compares

exact cancer rates and the Armitage–Doll approximation for the above example. Most of the error seen would be removed simply by adjusting $\mu$. More generally, there is no more departure than this from log–log linearity, even for a much wider age range, for values of $k$ less than about 8 and transition rates compatible with human cancer rates.

A more serious concern with multi-mutation models is some tension between presumed spontaneous mutation rates and observed cancer rates. Arguments given by biologists that mutation rates are much too small to explain cancer rates under models such as here are often not made correctly, and it is important to consider explicitly the Armitage–Doll approximation, and cancer rates rather than the lifetime risk of cancer. There does remain some tension, and we believe that a substantial part of the resolution of this must be that some mutations enormously increase the effective rate of subsequent ones by seriously impairing cellular repair mechanisms. Further, intermediate expansion of affected cells may mean that the log–log slope of cancer rates is somewhat larger than $k - 1$, allowing the entertainment of smaller values of $k$ than is usual and further reducing this tension.

## REFERENCES

AALEN, O. (1995). Phase type distributions in survival analysis. *Scandinavian Journal of Statistics* **22**, 447–463.

AALEN, O. AND GJESSING, H. K. (2001). Understanding the shape of the hazard rate: A process point of view. *Statistical Science* **16**, 1–13.

ARMITAGE, P. (1953). A note on the time-homogeneous birth process. *Journal of the Royal Statistical Society, Series B* **15**, 90–91.

ARMITAGE, P. AND DOLL, R. (1954). The age distribution of cancer and a multi-stage theory of carcinogenesis. *British Journal of Cancer* **VIII**, 1–12.

DAY, N. E AND BROWN, C. C. (1980). Multistage models and primary prevention of cancer. *Journal of the National Cancer Institute* **64**, 977–989.

HALPERN, M. T., GILLESPIE, B. W. AND WARNER, K. E. (1993). Patterns of absolute risk of lung cancer mortality in former smokers. *Journal of the National Cancer Institute* **85**, 457–464.

HAZELTON, W. D., LUEBECK, E. G., HEIDENREICH, W. F. AND MOOLGAVKAR, S. H. (2001). Analysis of a historical cohort of Chinese tin miners with arsenic, radon, cigarette smoke, and pipe smoke exposures using the biologically based two-stage clonal expansion model. *Radiation Research* **156**, 78–94.

KAI, M., LUEBECK, E. G. AND MOOLGAVKAR, S. H. (1997). Analysis of the incidence of solid cancer among atomic bomb survivors using a two-stage model of carcinogenesis. *Radiation Research* **148**, 348–358.

KELLERER, A. M. AND BARCLAY, D. (1992). Age dependencies in the modeling of radiation carcinogenesis. *Radiation Protection Dosimetry* **41**, 273–281.

LUEBECK, E. G., HEIDENREICH, W. F., HAZELTON, W. D., PARETZKE, H. G. AND MOOLGAVKAR, S. H. Biologically based analysis of the data for the Colorado uranium miners cohort: age, dose, and dose-rate effects. *Radiation Research* **152**, 339–351.

LUBIN, J. H., BOICE, J. D., EDLING, C., HORNUNG, R. W., HOWE, G., KUNZ, E., KUSIAK, R. A., MORRISON, H. I., RADFORD, E. P., SAMET, J. M., TIRMARCH, M., WOODWARD, A., XIANG, Y. S. AND PIERCE, D. A. (1994). *Radon and Lung Cancer Risk: A Joint Analysis of 11 Underground Miner Studies*. Washington, DC: National Institutes of Health Publication No. 94-3644.

LUBIN, J. H., BOICE, J. D., EDLING, C., HORNUNG, R. W., HOWE, G., KUNZ, E., KUSIAK, R. A., MORRISON, H. I., RADFORD, E. P., SAMET, J. M., TIRMARCH, M., WOODWARD, A., XIANG, Y. S. AND PIERCE, D. A. (1995). Lung cancer in radon-exposed miners and estimation of risk from indoor exposure. *Journal of the National Cancer Institute* **87**, 817–827.

MOOLGAVKAR, S. H. (1978). The multistage theory of carcinogenesis and the age distribution of cancer in man. *Journal of the National Cancer Institute* **61**, 49–52.

NATIONAL RESEARCH COUNCIL (1988). *Committee on the Biological Effects of Ionizing Radiation: Health Effects of Radon and Other Internally-Deposited Alpha-Emitters: BEIR-IV*. Washington, DC: National Academy Press.

NATIONAL RESEARCH COUNCIL (1999). *Health Effects of Exposure to Radon: BEIR-VI*. Washington, DC: National Academy Press.

NORDLING, C. O. (1953). A new theory on the cancer induction mechanism. *British Journal of Cancer* **VII**, 68–72.

OHTAKI, M. (1981). An approximation to the left tail of a distribution for waiting-time in an irreversible point process. *Journal of the Japan Statistical Society* **11**, 111–118.

OHTAKI, M. (1985). The age distribution of human adult cancer and an initiation-manifestation model for carcinogenesis. *Japanese Journal of Clinical Oncology* **15**, 325–343.

PIERCE, D. A., SHIMIZU, Y., PRESTON, D. L., VÆTH, M. AND MABUCHI, K. (1996). Studies of the mortality of atomic-bomb survivors, Report 12, Part I. Cancer: 1950–1990. *Radiation Research* **146**, 1–27.

PIERCE, D. A. AND MENDELSOHN, M. L. (1999). A model for radiation-related cancer suggested by the atomic bomb survivor data. *Radiation Research* **152**, 642–654.

PIKE, M. C., KRAILO, M. D., HENDERSON, B. E., CASAGRANDE, J. T. AND HOEL, D. G. (1983). 'Hormonal' risk factors, 'breast tissue age' and the age-incidence of breast cancer. *Nature* **303**, 767–770.

THOMAS, D. C. (1982). Temporal effects and interactions in cancer: Implications of carcinogenic models. Page 107–121 in *Environmental Epidemiology: Risk Assessment*. Philadelphia: SIAM Institute for Mathematics and Society.

THOMPSON, D. E., MABUCHI, K., RON, E., SODA, M., TOKUNAGA, M., OCHIKUBO, S., SUGIMOTO, S., IKEDA, T., TERASAKI, M., IZUMI, S. AND PRESTON, D. L. (1994). Cancer incidence in atomic bomb survivors. Part II: Solid tumors, 1958–1987. *Radiation Research* **137**, S17–S67.

WHITTEMORE, A. (1977). The age distribution of human cancer for carcinogenic exposures of varying intensity. *American Journal of Epidemiology* **106**, 418–432.

WHITTEMORE, A. AND KELLER, J. B. (1978). Quantitative theories of carcinogenesis. *SIAM Review* **20**, 1–30.

[*Received January 25, 2002; revised July 1, 2002; accepted for publication July 1, 2002*]