AGE-DEPENDENCE IN A STOCHASTIC MODEL OF CARCINOGENESIS

W. A. O'N. WAUGH
MCGILL UNIVERSITY

1. Introduction

This paper is concerned with a two-hit model for the initiation of neoplasia. The model is related to those discussed by H. G. Tucker [5], and by J. Neyman and E. L. Scott [4] in this Symposium, and also in recent papers by D. G. Kendall [1] and J. Neyman [3]. I should like to take the opportunity to thank Mr. Kendall for drawing my attention to these problems. The papers just mentioned all treat Markovian models, and this requires that the life lengths of cells be random variables with a negative exponential distribution. The main purpose of this paper is to discuss a model in which the probabilities for reproduction, mutation, and death of individual cells are supposed to have a general distribution, that is, to be functions of the ages of the cells, so that the process is non-Markovian.

2. Description of the model

We shall suppose that there are three types of cells, and for convenience we shall call them normal, gray, and black. Clones of black cells form malignant growths, while the gray cells are supposed to represent an intermediate stage between normality and malignancy. By mutation, a normal cell can be converted into a gray one, or a gray cell into a black one. We suppose that the population of normal cells is so large that its fluctuations can be ignored. The incidence of first-order mutations, from normal to gray, produces during a time interval (t_1,t_2) a number of gray cells which is a Poisson variable of expectation

$$\int_{t_1}^{t_2} f(t)dt, \tag{2.1}$$

where $f(t)$ is a function expressing the intensity of the carcinogenic action which causes the first-order mutations.

A gray cell will generate a clone of gray cells developing independently of one another according to the following age-dependent birth, death, and mutation process. Any individual when newly born has probability $1 - G(t)$ to live for a time longer than t, where $0 \leqq G(t) \leqq 1$ and $G(t) \uparrow 1$. Its life will be considered to end when it reproduces by binary fission, or ceases to be a gray cell as a result of mutation, or dies. Note that the probability for a cell's life to end is treated

as a function of its age, but that this probability is otherwise independent of the time at which the cell is newly born. This condition ensures that the process, though non-Markovian, is stationary. It corresponds to a system in which the gray cells grow in an unchanging environment.

If a gray cell's life ends at age t there are then three alternatives.

(a) It may be replaced by two newly born gray cells (reproduction by binary fission) with conditional probability $q_2(t)$.

(b) It may die, ceasing to be counted as a member of the population, with conditional probability $q_0(t)$. This may also include the possibility of mutation back to normality.

(c) It may be replaced by two newly born cells, one gray and one black, with conditional probability $q_1(t)$. Other possibilities may be considered. For example, a convenient way to treat the occurrence of mutations at times other than the epochs of cell division is to suppose that the gray cell is replaced by a single black cell.

A black cell will generate a clone of black cells developing by an age-dependent birth and death process. The difference between the two types of clone will be that the rates of birth, and of removal by death or mutation, for the gray cells, will be such that the process is subcritical; while for the black cells the birth and death rates will be such that the process is supercritical. That is to say, the probability of ultimate extinction for the gray clones is 1, while it is less than 1 for the black clones, and so after sufficient time some of those clones which we regard as malignant will grow indefinitely large. Conditions for this in terms of the birth and removal rates are known [6], and will be considered in detail later.

Now let us consider what might happen if the growth of gray and black clones took place as part of an experiment. At the end of a period of exposure to carcinogenic agents the experimenter might attempt to count the gray and black clones, but he would probably miss some, and it seems more likely that he would miss small clones than large ones. We suppose that $C_g(n)$ is the conditional probability that a gray clone is detected, given that its size is n; similarly $C_b(n)$ relates to black clones. In order to solve the integral equations for the probabilities for the observation of a gray clone and the black clones that may arise from it we are obliged to consider a particular form for $C_g(n)$ and $C_b(n)$ as well as for the generation time distribution $G(t)$, but the equations will be written down before these assumptions are introduced so that we shall be able to consider alternatives. We require one other conditional probability. Suppose a black cell is newly born at time $t = 0$, and let $q(t)$ be the probability that the resulting clone is detected when observation takes place at time t. Clearly $q(t)$ will be a function of the birth and death rates for the black cells, and of the conditional probabilities $C_b(n)$.

3. Progeny of a single gray cell; integral equation

Suppose a single gray cell is newly born at time $t = 0$. Let g_t be the number of gray *cells*, B_t be the number of black *clones*, and B'_t be the number of *detected* black clones when observation is made at time t. Let

$$\varphi(z, w; t) = E\{z^{g_t}, w^{B'_t}\}. \tag{3.1}$$

We shall obtain an integral equation for this double generating function by considering possible events during the period $(0, t)$. Suppose the initial gray cell's life ends during the time interval $(u, u + du)$, which is an event of probability $dG(u)$. If $0 < u < t$, then with probability $q_2(u)$ two newly born gray cells may replace it and give rise to two independent subpopulations, each developing under the same system of probabilities for the period $t - u$. The generating function for the total numbers of gray cells and detected black clones that result is $\varphi^2(z, w; t - u)$. In a similar way the other alternatives, including that of survival of the initial individual beyond the period $(0, t)$, contribute the other terms to the equation, which is

$$\varphi(z, w; t) = \int_0^t \{q_2(u)\varphi^2(z, w; t - u) + q_0(u) + q_1(u)[1 - (1 - w)q(t - u)]\varphi(z, w; t - u)\}dG(u) + z[1 - G(t)]. \tag{3.2}$$

In order to obtain detailed information about $\varphi(z, w; t)$ from equation (3.2), it is necessary to make some assumptions about the nature of the functions $q_j(t)$ and $G(t)$, and about the system of growth and detection of the black clones, which determines $q(t)$. In the present work we shall concentrate on the limiting behavior of the system after a long period of development, and attempt to solve for age-dependent processes the same problems as were investigated by Kendall [1] for Markovian processes. To do this it is sufficient to make quite general assumptions.

Let us first consider the black clones. Suppose the generation time distribution for a black cell is $H(t)$ and the probabilities for fission or death at the end of its life are $r_2(t)$ and $r_0(t)$ respectively. Let

$$M = \int_0^\infty r_0(t)dH(t) \quad \text{and} \quad L = \int_0^\infty r_2(t)dH(t). \tag{3.3}$$

Then it is known that the probability of ultimate extinction for a black clone starting from one newly born black cell is M/L (see [6]). Henceforward we shall suppose that the detection probabilities have the form suggested by Neyman and used by Kendall [1]; that is, $C_g(n) = 1 - \gamma^n$ and $C_b(n) = 1 - \beta^n$ where $0 < \gamma < 1$ and $0 < \beta < 1$. As in Kendall's work, this makes it possible to use generating functions. Let b_t be the number of black cells in a clone that started at time $t = 0$ and let $\psi(z; t) = E\{z^{b_t}\}$. Then clearly the probability of detection for such a clone, at time t, is given by $q(t) = 1 - \psi(\beta; t)$. An argument exactly similar to the one that leads to (3.2) gives

$$\psi(z;t) = \int_0^t \{r_2(u)\psi^2(z;t-u) + r_0(u)\}dH(u) + z[1 - H(t)] \tag{3.4}$$

and we can put $z = \beta$ in this.

Letting $t \to \infty$ in (3.4), and using (3.3), we get

$$\psi(\beta;\infty) = L\psi^2(\beta;\infty) + M, \tag{3.5}$$

which is the same as the equation for the probability M/L of ultimate extinction. By analytical methods similar to those in section 4 of [6] it can be shown that $\psi(\beta;t)$ must converge to M/L, but we will merely note the obvious interpretation, that after a long period of time a black clone either becomes extinct or becomes so large that its detection is virtually certain. Thus $q(\infty) = 1 - M/L$.

Returning to the problem of the offspring of a gray cell, suppose

$$\sigma_j = \int_0^\infty q_j(u)dG(u), \qquad j = 0, 1, 2. \tag{3.6}$$

Once again applying the analytical methods of [6], we can show that $\Phi(z;w) = \lim_{t\to\infty} \varphi(z, w; t)$ exists and satisfies the equation that we get by letting $t \to \infty$ in (3.2) and by replacing $q(t-u)$ by $q(\infty)$ under the integral sign. This gives

$$\sigma_2\Phi^2(z;w) + \left[w\sigma_1\left(1 - \frac{M}{L}\right) - \left(1 - \sigma_1\frac{M}{L}\right)\right]\Phi(z;w) + \sigma_0 = 0. \tag{3.7}$$

In this quadratic the coefficients are functions of w alone and hence we have

$$\Phi(z, w) = c(w), \tag{3.8}$$

say, which is independent of z. This result corresponds to one given by Kendall [1] for his Markovian model of carcinogenesis, and he has given an interpretation of it. After a long time, the gray clone will have died out, but before doing so it may give rise to some black cells, from each of which a black clone may develop. Some of these may become extinct but others may survive and grow very large. The latter will be those that the experimenter counts, and the number of them will have the probability distribution generated by $c(w)$. We have

$$c(w) = \frac{1}{2\sigma_2}\left[\left(1 - \sigma_1\frac{M}{L}\right) - w\sigma_1\left(1 - \frac{M}{L}\right)\right] - \frac{1}{2\sigma_2}\left\{\left[w\sigma_1\left(1 - \frac{M}{L}\right) - \left(1 - \sigma_1\frac{M}{L}\right)\right]^2 - 4\sigma_2\sigma_0\right\}^{1/2}, \tag{3.9}$$

which clearly can be expanded in powers of w.

The coefficients are complicated but it can be seen that the limiting probability of the gray cell giving rise to no detected black clone is

$$\frac{1}{2\sigma_2}\left(1 - \sigma_1\frac{M}{L}\right)\left[1 - \left\{1 - 4\sigma_2\sigma_0\left(1 - \sigma_1\frac{M}{L}\right)^{-1}\right\}^{1/2}\right] \tag{3.10}$$

and the probability of just one is

$$\frac{\sigma_1}{2\sigma_2}\left(1 - \frac{M}{L}\right)\left[\left\{1 - 4\sigma_2\sigma_0\left(1 - \sigma_1\frac{M}{L}\right)^{-2}\right\}^{-1/2} - 1\right], \tag{3.11}$$

while it is clear that the probability of n, where $n \geqq 2$, will involve a factor σ_1^n. Now σ_1 is the probability that the life of a single newly born gray cell will end in mutation and it will usually be very small, so that it is almost certain that a single mutation of a normal cell to a gray one will give rise, if at all, to just one detected black clone.

It is of interest to consider an alternative model here. Suppose the effect of mutation, instead of being as described under (c) of section 2, is to change a gray cell instantaneously into a black cell at any period of its life. It is conceivable that this accident might speed up or delay its reproduction or death. In general we shall have to consider the conditional probability $q(t)$ as being replaced by $q(t|v)$, for a black clone to be detected at time t, given that it started with the mutation of a gray cell at time zero, the gray cell then being of age v. The equation (3.2) will be replaced by

$$\varphi(z, w; t) = \int_0^t \{q_2(u)\varphi^2(z, w; t-u) + q_0(u) + q_1(u)[1 - (1-w)q(t-u|v)]\}dG(u) + z[1 - G(t)]. \tag{3.12}$$

Unfortunately we cannot let $t \to \infty$ in this equation as we have done in (3.2) because $q(t-u|v)$ does not depend only on the difference $t-u$. One can imagine various situations in which it tends to a limit independent of v, for example in the special case when the *black* cells reproduce in a Markovian manner. In that case the limiting function $\Phi(z, w)$ must satisfy

$$\Phi = \sigma_2\Phi^2 + \sigma_0 + \sigma_1\left[1 - (1-w)\left(1 - \frac{M}{L}\right)\right], \tag{3.13}$$

which is equivalent to Kendall's equation (2.8A) in [1].

4. The general problem

In section 3 we have investigated the growth of the gray and black offspring of a single gray cell supposed to have been born at time zero. Returning to section 2 we recall that gray cells are supposed to come into existence in a random manner so that the number appearing between times t_1 and t_2 is a Poisson variable of expectation $\int_{t_1}^{t_2} f(u)du$. We must now investigate the result of the complete process in which such gray cells appear and give rise to clones. We shall suppose that the gray cells that appear are newly born, which, as in section 2, corresponds to a situation in which a mutant appears at the instant of fission. Let X_t be the number of gray clones and Y_t be the number of black clones *that are detected at time t.* Let

$$\psi(z, w; t) = E\{z^{X_t}, w^{Y_t}\} \tag{4.1}$$

be the double generating function for the joint distribution of these variables. Then it can be evaluated, and moments of X_t and Y_t obtained from it, in

exactly the same way as in Kendall's paper [1]. For convenience we shall reproduce the results since they are to be used in what follows.

Suppose a gray cell is formed during $(u, u + du)$, an event which has probability $f(u)du$, and that at time t it yields g gray cells and B' detected black clones. Suppose it yields G' *detected* gray clones, where $G' = 0$ or 1. Then

$$E\{z^{G'}, w^{B'}|g, B'\} = (1 - \gamma^g)zw^{B'} + \gamma^g w^{B'}, \tag{4.2}$$

whence

$$E\{z^{G'}, w^{B'}\} = \varphi(\gamma, w; t - u) + z[\varphi(1, w; t - u) - \varphi(\gamma, w; t - u)]. \tag{4.3}$$

Thus the contribution to the double generating function for the distribution of the numbers of gray and black clones formed during $(0, t)$ and detected at the epoch t, due to the possibility of a gray cell being formed during $(u, u + du)$, is

$$1 - f(u)du + E\{z^{G'}, w^{B'}\}f(u)du + o(du). \tag{4.4}$$

Forming a product-integral from this we obtain

$$\psi(z, w; t) = E\{z^{X_t}, w^{Y_t}\} = \exp\{-R(w, t) - (1 - z)S(w, t)\}, \tag{4.5}$$

whence

$$R(w, t) = \int_0^t [1 - \varphi(1, w; t - u)]f(u)du \tag{4.6}$$

and

$$S(w, t) = \int_0^t [\varphi(1, w; t - u) - \varphi(\gamma, w; t - u)]du. \tag{4.7}$$

From these we can obtain the moments of X_t and Y_t as follows

$$E\{X_t\} = S(1, t) = \int_0^t [1 - \varphi(\gamma, 1; t - u)]f(u)du \tag{4.8}$$

$$E\{Y_t\} = -\frac{\partial}{\partial w}R(1, t) = \int_0^t [\varphi_w(1, w; t - u)]_{w=1}f(u)du \tag{4.9}$$

$$E\{X_t^2\} = [S(1, t)]^2 + S(1, t) \tag{4.10}$$

$$\text{Var } (X_t) = E\{X_t\} \tag{4.11}$$

$$E\{Y_t^2\} = \left[\frac{\partial}{\partial w}R(1, t)\right]^2 - \frac{\partial^2}{\partial w^2}R(1, t) - \frac{\partial}{\partial w}R(1, t) \tag{4.12}$$

$$\begin{aligned}\text{Var } (Y_t) &= -\left(\frac{\partial^2}{\partial w^2} + \frac{\partial}{\partial w}\right)R(1, t) \\ &= \int_0^t [\varphi_{ww}(1, w; t - u) + \varphi_w(1, w; t - u)]_{w=1}f(u)du\end{aligned} \tag{4.13}$$

$$\text{Cov}\,(X_t, Y_t) = \frac{\partial}{\partial w} S(1, t) \tag{4.14}$$

$$= \int_0^t [\varphi_w(1, w; t-u) - \varphi_w(\gamma, w; t-u)]_{w=1} f(u)\, du.$$

In order to use these formulas we need to know

$$\varphi(\gamma, 1; t), \tag{4.15}$$

$$\varphi_w(1, w; t) \quad \text{and} \quad \varphi_{ww}(1, w; t) \qquad \text{when} \quad w = 1, \tag{4.16}$$

$$\varphi_w(\gamma, w; t) \qquad \text{when} \quad w = 1. \tag{4.17}$$

The obstacle to further progress is (4.15). Explicit evaluation of this function, equivalent to $q(t)$ in sections 2 and 3, is possible in the case of Markovian growth. In even the simplest generalization, where the generation time distribution is supposed to be of χ^2_{2k} form, $\varphi(\gamma, 1; t)$ can be shown to be the solution of a higher order nonlinear equation which has been treated elsewhere [2], [6], and the most promising line of approach seems to be to linearize it and calculate moments. This is precisely what is wanted for (4.16) and (4.17), so we can go further with (4.9) and (4.12). An alternative, which has been used before in studies of mutation, is to suppose that one type of clones (in this case the gray ones) grow in a Markovian manner while the others grow in an age-dependent manner. If this is done, however, the results are almost identical with those already obtained by Kendall [1]. Solutions for finite time would involve explicit solutions of the linearized equations for (4.16) and (4.17) but in view of the difficulty over (4.15) we shall simply consider (4.9) and (4.13) for large values of time t. We have, differentiating (3.2)

$$\begin{aligned}\varphi_w(z, w; t-u) = \int_0^t \{&2q_2(u)\varphi(z, w; t-u)\varphi_w(z, w; t-u) \\ &+ q_1(u)q(t-u)\varphi(z, w; t-u) \\ &+ q_1(u)[1 - (1-w)q(t-u)]\varphi_w(z, w; t-u)\}\, dG(u).\end{aligned} \tag{4.18}$$

Putting $z = w = 1$ and letting $t \to \infty$, we have

$$\varphi_w(1, 1; \infty) = 2\sigma_2\varphi_w(1, 1; \infty) + \sigma_0\left[1 - \frac{M}{L}\right] + \sigma_1\varphi_w(1, 1; \infty). \tag{4.19}$$

In (4.9) we see that $E\{Y_t\}$ depends on the feeding function $f(u)$, as well as on the solution of (4.19). In the case where $f(u)$ is equal to a constant f, we obtain

$$\lim_{t\to\infty} E\{Y_t/t\} = f\frac{\sigma_1}{\sigma_0 - \sigma_2}\left[1 - \frac{M}{L}\right]. \tag{4.20}$$

We can differentiate (4.18) again to obtain an equation for $\varphi_{ww}(z, w; t-u)$ and once again consider the limit as $t \to \infty$, when $z = w = 1$. Together with (4.13) and (4.19) this gives

$$\lim_{t\to\infty}\frac{1}{t}\operatorname{Var}(Y_t) = f\frac{\sigma_1}{\sigma_0-\sigma_2}\left[1-\frac{M}{L}\right]+2f\frac{\sigma_1^2}{(\sigma_0-\sigma_2)^3}\left[1-\frac{M}{L}\right]^2. \tag{4.21}$$

Some qualitative deductions made by Kendall [1] for the Markovian case can be reproduced for the present model; for instance, obviously X_t is a Poisson variable and Cov (X_t, Y_t) is positive. Also his remarks on the subject of a "threshold" require only trivial changes in the formulas, which are due to the present model being worked out for the case where mutation takes effect at the instant of fission.

5. Model involving more than two hits

We can extend some of the preceding results to a model in which there are two (or in general n) types of gray cells, and the course of mutation can be schematically represented by normal $\rightarrow$ gray$_1$ $\rightarrow$ gray$_2$ $\rightarrow$ black. We shall outline the generalization of section 3 to this model. Let $J(t)$ be the generation time distribution for a gray$_1$ cell, that is to say, the probability that a gray$_1$ cell newly born at time $t = 0$ lives longer than time t is $1 - J(t)$. Let $s_j(t)$, where $j = 0, 1, 2$, be the conditional probabilities for what follows the end of life of a gray$_1$ cell. In particular, $s_1(t)$ is to be the probability that the gray$_1$ cell is replaced by a newly born pair consisting of a gray$_1$ and a gray$_2$ cell. Let g_{1t} be the number of gray$_1$ cells, g_{2t} be the number of gray$_2$ cells, and B_t' be the number of detected black clones when observation is taken at time t, there being a single newly born gray$_1$ cell at time 0.

Let

$$\chi(x, y, z; t) = E\{x^{g_{1t}}, y^{g_{2t}}, z^{B_t'}\}. \tag{5.1}$$

Referring back to the argument leading to (3.2) it can be seen that the *double* generating function for g_{2t} and B_t', *when it is supposed that one newly born gray$_2$ cell is in existence at time* $t = 0$, is the same function $\varphi(y, z; t)$ as appears there, provided that we replace g_t by g_{2t}. Note that the dummy variables are now y and z. A similar argument then shows that $\chi(x, y, z; t)$ and $\varphi(y, z; t)$ are connected by the equation

$$\begin{aligned}\chi(x, y, z; t) = \int_0^t \{&s_2(u)\chi^2(x, y, z; t-u) + s_0(u)\\ &+ s_1(u)\chi(x, y, z; t-u)\varphi(y, z; t-u)\}dJ(t) + x[1 - J(t)].\end{aligned} \tag{5.2}$$

The equation (5.2) can be treated by limiting operations such as have been used before. Let

$$\chi(x, y, z; \infty) = \lim_{t\to\infty}\chi(x, y, z; t), \tag{5.3}$$

and

$$\int_0^\infty s_j(t)dJ(t) = \Sigma_j, \qquad j = 0, 1, 2. \tag{5.4}$$

Then

$$\chi(x, y, z; \infty) = \Sigma_2 \chi^2(x, y, z; \infty) + \Sigma_0 + \Sigma_1 \chi(x, y, z; \infty)\varphi(y, z; \infty), \tag{5.5}$$

whence it follows that $\chi(x, y, z; \infty)$ is a function of z alone, given by

$$\chi(x, y, z; \infty) = \frac{1}{2\Sigma_2}\left(1 - \Sigma_1 c(z) - \{[1 - \Sigma_1 c(z)]^2 - 4\Sigma_2\Sigma_0\}^{1/2}\right), \tag{5.6}$$

where $c(z) = \varphi(y, z; \infty)$ is the limiting generating function obtained in section 3. This argument can clearly be repeated to cover more than two types of gray cells.

REFERENCES

[1] D. G. KENDALL, "Birth-and-death processes, and the theory of carcinogenesis," *Biometrika*, Vol. 47 (1960), pp. 13–22.

[2] ———, "On the role of variable generation time in the development of a stochastic birth process," *Biometrika*, Vol. 35 (1948), pp. 316–330.

[3] J. NEYMAN, "A stochastic model of carcinogenesis," approximate text of seminar talks delivered at the National Institutes of Health, November 12 and 16, 1958.

[4] J. NEYMAN and E. L. SCOTT, "Two-stage mutation theory of carcinogenesis," unpublished.

[5] H. G. TUCKER, "A stochastic model for a two-stage theory of carcinogenesis," *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, University of California Press, 1961, Vol. 4, pp. 387–403.

[6] W. A. O'N. WAUGH, "An age-dependent birth and death process," *Biometrika*, Vol. 42 (1955), pp. 291–306.