A STOCHASTIC MODEL FOR A TWO-STAGE THEORY OF CARCINOGENESIS

HOWARD G. TUCKER
UNIVERSITY OF CALIFORNIA, BERKELEY

1. Introduction to the theory

A two-stage theory of induced pulmonary carcinogenesis has been suggested by M. B. Shimkin and M. J. Polissar [6]. In brief, the preliminary evidence and the problem suggested by more detailed study are outlined as follows. Each mouse of a large collection of experimental mice is administered a dose of the carcinogenic agent, urethan. At certain times after administering the carcinogen a small group (5 or 10 or 15) of mice were selected from the large group; they were killed and their lungs were examined. In examinations undertaken shortly after administering the carcinogen, small distinct growths were noticed on the lungs which were not cancerous tumors and which Shimkin and Polissar call "hyperplastic foci." As time progressed these hyperplastic foci increased in both number and size. Eventually, the individual sizes of the hyperplastic foci became constant, while the total number of them per mouse lung began to decrease gradually. At about the same time, or shortly after the appearance of the first hyperplastic foci, tumors were observed on the lungs. For a while the tumors per mouse lung increased in both number and size. A time was reached, however, when the number of tumors on each mouse's lungs remained fairly constant, while continuing to increase in individual size. It was noticed at the early era of tumor formation that the smallest tumors observed were much larger in cross-sectional area than the smallest cross-sectional area of tumor that could still be observed with the microscopes. In their study, Shimkin and Polissar were thus led to consider that perhaps the hyperplastic foci and the tumors were not biologically independent of each other. It seemed to them that possibly the hyperplastic foci were precursors to the tumors and that as some hyperplastic foci attained a certain approximate age and size they changed into tumors. This is the two-stage theory of carcinogenesis as formulated by Shimkin and Polissar.

This paper is a preliminary attempt toward the verification or disproof of this theory. The method to be used is that of verifying (or, more accurately, not excluding) a particular theory by means of a mathematical model. This method is explained here for our particular problem. We first make the basic assumption that the number $X(t)$ of hyperplastic foci that can be counted on a mouse's lungs at time t after administering the carcinogen is a random variable. We make the

This investigation was supported (in part) by a research grant (No. RG-3666) from the National Institutes of Health, Public Health Service.

same basic assumption for $Y(t)$, the total number of tumors that can be counted on a mouse's lungs at time t, and for $Z(t)$, the total number of cells included in all observable tumors on a mouse's lungs at time t. We include in these assumptions that the joint distribution of $[X(t), Y(t), Z(t)]$ is well defined. So far we have assumed that the phenomenon behaves according to some rules of chance. It is within this assumption that we create a particular theory that incorporates a mechanism that the hyperplastic foci are indeed precursors to the tumors. Under the assumption that this particular theory is true, we shall derive the joint probability distribution of $[X(t), Y(t), Z(t)]$. Then we must derive a test of hypothesis that the experimental observations are distributed according to the joint distribution function just derived. Actual data are then needed, and the test must be performed. If on the basis of actual data the test tells us to reject the hypothesis, then of course the mathematical model is incorrect. In such a case, the theory is incorrect or, if it is correct, is governed by a mathematical model yet to be discovered. If on the basis of the test we do not reject the hypothesis, then we conclude that *we have no overwhelming evidence leading us to believe that the theory is incorrect.* It is to be remembered that whatever conclusions are drawn are in terms of the basic stochastic assumption. This means that once we have made this assumption we have removed ourselves from being able to make purely biological statements. We may now make only biological-mathematical assertions.

2. The stochastic model

We now state the basic assumptions behind our stochastic model for the two-stage theory of carcinogenesis. Justification of these assumptions will be given in the next section. It is therefore assumed that the hyperplastic foci are indeed the precursors of the tumors and that the formation of both is governed by the following rules.

We first assume (following Neyman [4]) that there is a nonnegative continuous function $f(t)$ which will be called the "feeding function." The function $f(t)$ is determined jointly by the animal and the carcinogen except for a multiplicative constant. The multiplicative constant is directly proportional to the dosage of carcinogen administered. For every $t \geqq 0$ and every $h > 0$ it is assumed that on one mouse's lungs

(i) the probability that the carcinogen causes the formation of one hyperplastic focus during the time interval $[t, t + h)$ is $f(t)h + o(h)$,

(ii) the probability that more than one hyperplastic focus is formed because of the carcinogen during the time interval $[t, t + h)$ is $o(h)$, and

(iii) if $0 < t_1 < t_2 < t_3$ are any three times, then the numbers of hyperplastic foci formed because of the carcinogen in the two intervals $[t_1, t_2)$ and $[t_2, t_3)$ are independent.

When a hyperplastic focus is first formed, it is assumed to be unobservable. An unobservable hyperplastic focus can remain unobservable or can undergo one

of two possible changes: (1) it can vanish (die, or disappear) or (2) it can become observable (under the microscope). We shall assume that there exists a positive constant μ_0 such that for every time $t \geqq 0$ and every $h > 0$, (i) the probability that any particular unobservable hyperplastic focus will die during the time interval $[t, t + h)$ is $\mu_0 h + o(h)$, and (ii) that all unobservable hyperplastic foci act independently of each other with regard to dying. It is also assumed that there is a positive constant ν_0 such that for every time $t \geqq 0$ and every $h > 0$, (i) the probability that a particular unobservable hyperplastic focus will become observable during the time interval $[t, t + h)$ is $\nu_0 h + o(h)$, and (ii) that all unobservable hyperplastic foci act independently of each other with respect to becoming observable or dying. We further assume that the values of μ_0 and ν_0 depend only on the carcinogen and the animal and *not* on the dosage.

Once a hyperplastic focus is observable, again we assume that one of two possible changes (if any) can occur to it: (1) it can die, or (2) it can be instantaneously transformed at some time thereafter into a tumor. We shall assume that there exists a constant $\mu > 0$ and a nonnegative continuous function $\nu(t)$ such that for every $t \geqq 0$ and every $h > 0$ the following holds true:

(i) the probability that any particular observable hyperplastic focus will die during the time interval $[t, t + h)$ is $\mu h + o(h)$,

(ii) the probability that any particular observable hyperplastic focus will change into a tumor during the time interval $[t, t + h)$ is $\nu(t)h + o(h)$, and

(iii) all observable hyperplastic foci act independently of each other with respect to dying or becoming tumors.

Once a tumor is formed (or born) we shall assume that it cannot die or be transformed into something else. We shall assume that when a tumor is born it is composed of N cells, where N is a random variable. We shall assume that the distribution of N is one which is completely specified by its expectation; in the model considered here we shall consider it to be Poisson, but if this does not work one might try the binomial distribution, $B(1/2, m)$. We shall assume that there is a nonnegative continuous function $\rho(t)$, where $\rho(t)$ is asymptotically a positive constant such that the increase of cells of a tumor is a Poisson process with intensity $\rho(t)$. It should be noted here that the t in $\rho(t)$ denotes the time since administering the carcinogen and *not* the time since the tumor was born. We shall assume all tumors to be independent of each other, and the number, N, of cells of the newborn tumor to have the same distribution independent of the tumor and the time it was born. Also, for each tumor, N is assumed to be independent of the number of cells subsequently added.

At each time t we shall denote the number of observable hyperplastic foci, the number of tumors, and the total number of cells in all the tumors on a mouse's lungs by $X(t)$, $Y(t)$, $Z(t)$ respectively. We assume that, at each of a finite number of times, independent observations can be taken on the random vector $[X(t), Y(t), Z(t)]$, and that these sets of observations are independent of each other for different values of time.

It is to be noted at this point that although $[X(t), Y(t), Z(t)]$ is a vector

stochastic process, and its probabilistic treatment is for a stochastic process, the statistical inference cannot treat it as such. This is because one cannot observe the same sample function at more than one value of time. The particular kind of sampling involved and the inference problem it creates, will be discussed in greater detail in sections 5 and 6.

3. Justification of the model

On the basis of experimental results already reported (see Polissar and Shimkin [5], [6], [7]) there is considerable justification for the assumptions made in the preceding section. Also certain mathematical results deduced from the assumptions made are corroborated by experimental evidence, but this shall be reported in a later section. In this section an outline will be made of just the experimental evidence for the mathematical model.

The function $f(t)$, and its properties which were assumed, is not new to the study of carcinogenesis. It is called the "feeding function," and it has been used by Neyman [4] and Kendall [3] in their papers on this subject, in precisely the same way as it is used in this paper. The essential idea seems to be that the rate at which the hyperplastic foci are formed on a mouse's lungs is proportional to the amount of carcinogenic agent in the animal's body at time t. This amount is not usually constant. In the case of urethan (which was used by Shimkin and Polissar in their experiments) it is eliminated from the body rather quickly, but injections of S_r-90 are eliminated at a very slow rate. Since Shimkin and Polissar used urethan, it is expected that $f(t)$ is large for values of t close to zero and decreases rather rapidly to zero as t increases. A further justification for the use of $f(t)$ as the intensity function of a Poisson process lies in the fact that recent evidence (see [8]) indicated that cell nuclei of a vast number of cells are modified by the carcinogen and each such cell has therefore a very small chance of transforming itself into the nucleus of a hyperplastic focus. Thus the formation of hyperplastic foci is more realistically governed by the following process. At time $t = 0$ there are M elements of the population, where M is a very large integer. One assumes the existence of a nonnegative function $g(t)$ such that if at any time $t > 0$ and for every $h > 0$ the conditional probability that an element of the population will leave the population during $[t, t + h)$ given that n have left it prior to t is $(M - n)g(t)h + o(h)$. Further, the conditional probability that more than one element leaves the population during $[t, t + h)$ given that n have left it prior to t is $o(h)$. If M is very large and if $Mg(t)$ is of moderate size, then for small values of n, we have $ng(t)$ negligible. Thus the process of elements leaving the population is approximately a Poisson process with intensity $f(t) = Mg(t)$. In the model used in this study, we see that the Poisson process describing the size of the population of hyperplastic foci is just a reasonable approximation to a realistic model.

In assuming the existence of $\nu(t)$ and its properties, we need not have been so general and could actually have assumed that $\nu(t) = cf(t)$, where c is a constant.

For large values of t, $\nu(t)$ then should be close to zero, and this is verified by table 3, page 82 in Polissar and Shimkin [6], which shows that after 49 days the number of tumors remains relatively constant. We are justified in taking μ as a constant by noting on the same table referred to above that shortly after the 38th day after administering urethan the number of hyperplastic foci per mouse lung decreased linearly. Since by that time the value of $f(t)$ should be negligible, and so should $\nu(t)$, it is safe to assume that any change thereafter in the number of hyperplastic foci on a mouse's lungs is due to their dying. Thus μ very realistically is constant.

In both [6] and [7], Polissar and Shimkin noted that the rate of growth of the individual tumors does not depend on the amount of carcinogen administered. Actually, the number of tumors eventually formed was proportional to the carcinogen dosage, but the growth rate of a tumor, once it was born, was completely independent of the size of this dose. In addition, a tumor had a rapid rate of growth when it was first formed, but after a while its rate of growth became constant. Thus it seemed most reasonable to consider a tumor as starting with a random number N of cells and with cell additions being governed by a Poisson process with a time-dependent intensity $\rho(t)$ which is eventually constant $\rho_0 > 0$. (One should note that the t in $\rho(t)$ denotes time beginning with the administering of the carcinogen and not beginning from the time the hyperplastic focus changed into a tumor.)

The theory (model) formulated in section 2 and justified in section 3 is perhaps the simplest theory that can be used to explain the phenomena. Indeed, this model could be made more realistic by letting μ_0, ν_0, and μ depend on the time since the particular hyperplastic focus was born and $\rho(t)$ depend on the time since the particular tumor was born. Such concern for complete reality can prevent the mathematician from obtaining "answers" because of the complexity of the resulting mathematical model. In the model used in this study, some "answers" are obtainable. Whether the model is approximately correct and the answers meaningful will depend on gathering more data of the type indicated in [6] and applying the methods of inference of section 6 to this data.

4. Mutation processes

Before we proceed to derive the distribution of $[X(t), Y(t), Z(t)]$, we shall discuss the general process of which $[U(t), X(t), Y(t)]$ is a particular case, where $U(t)$ is the number of unobservable hyperplastic foci on a mouse's lungs. This process will be referred to here as a mutation process.

A mutation process is a finite dimensional stochastic process

$$\mathbf{X}(t) = [X_1(t), \cdots, X_r(t)], \qquad r \geqq 2, \tag{4.1}$$

where $X_i(t)$ is a nonnegative integer-valued random variable for every t and denotes at time t the number of particles of the ith kind. We now define this process. Particles of type 1 are formed according to a Poisson process with time-

dependent intensity $\lambda(t)$. Each of the particles of type 1 is stochastically independent of all others; for every $t \geqq 0$ and every $h > 0$, it will die during $[t, t + h)$ with probability $\mu_1(t)h + o(h)$ and will change into a particle of type 2 with probability $\nu_{12}(t)h + o(h)$ where $\mu_1(t)$ and $\nu_{12}(t)$ are nonnegative continuous functions of t. For every i, with $2 \leqq i \leqq r - 1$, each particle of type i is stochastically independent of every other particle of type i and all particles of type $1, 2, \cdots, i - 1$. For every $t \geqq 0$ and every $h > 0$ a particular particle of type i will die during $[t, t + h)$ with probability $\mu_i(t)h + o(h)$ and will change into a particle of type $i + 1$ during $[t, t + h)$ with probability $\nu_{i,i+1}(t)h + o(h)$, where $\mu_i(t)$ and $\nu_{i,i+1}(t)$ are nonnegative continuous functions of t. Finally, every particle of type r is independent of every other particle of type r and all particles of type $1, 2, \cdots, r - 1$, and, for every $t \geqq 0$ and every $h > 0$, will die during the time interval $[t, t + h)$ with probability $\mu_r(t)h + o(h)$, where $\mu_r(t)$ is a continuous function of t. It is assumed that $P\{\mathbf{X}(0) = \mathbf{0}\} = 1$, where $\mathbf{0}$ denotes an r-dimensional vector composed entirely of zeros.

The problem treated here is to find the joint distribution of $\mathbf{X}(t) = [X_1(t), \cdots, X_r(t)]$ and to investigate its properties. For the sake of simplicity and economy of words in deriving this joint distribution in the general case, we do it first for the case $r = 2$. If we let

$$p(m, n, t) = P\{X_1(t) = m, X_2(t) = n\}, \tag{4.2}$$

then, by the same method as that used in the last chapter of Feller [7], one obtains the infinite system of differential equations

$$\begin{aligned} \frac{d}{dt} p(m + 1, n + 1, t) &= -(m + 1)[\nu_{12}(t) + \mu_1(t)]p(m + 1, n + 1, t) \\ &\quad - (n + 1)\mu_2(t)p(m + 1, n + 1, t) - \lambda(t)p(m + 1, n + 1, t) \\ &\quad + \lambda(t)p(m, n + 1, t) + \mu_1(t)(m + 2)p(m + 2, n + 1, t) \\ &\quad + \nu_{12}(t)(m + 2)p(m + 2, n, t) + \mu_2(t)(n + 2)p(m + 1, n + 2, t) \end{aligned} \tag{4.3}$$

for $m, n = -1, 0, 1, 2, \cdots$. Let us denote the joint probability generating function of $[X_1(t), X_2(t)]$ by

$$\psi(x, y, t) = E[x^{X_1(t)}y^{X_2(t)}]. \tag{4.4}$$

If we multiply both sides of (4.3) by $x^{m+1}y^{n+1}$ and sum both sides from $m = -1$, $n = -1$ to ∞, we obtain the partial differential equation

$$\begin{aligned} \frac{\partial}{\partial t}\psi(x, y, t) + \{\mu_1(t)(x - 1) + \nu_{12}(t)(x - y)\}\frac{\partial}{\partial x}\psi(x, y, t) \\ + \mu_2(t)(y - 1)\frac{\partial}{\partial y}\psi(x, y, t) = \lambda(t)(x - 1)\psi(x, y, t). \end{aligned} \tag{4.5}$$

In order to solve (4.5) we begin by considering the system of differential equations

$$\frac{dy}{dt} = \mu_2(t)(y - 1)$$

$$\frac{dx}{dt} = \mu_1(t)(x - 1) + \nu_{12}(t)(x - y) \tag{4.6}$$

$$\frac{d\psi}{dt} = \lambda(t)(x - 1)\psi.$$

We shall also make use of the notation

$$F(t) = \int_0^t \lambda(\tau)\, d\tau, \qquad G(t) = \int_0^t \tau\lambda(\tau)\, d\tau,$$

$$N(t) = \int_0^t \nu_{12}(\tau)\, d\tau, \qquad R(t) = \int_0^t \mu_2(\tau)\, d\tau, \tag{4.7}$$

$$M(t) = \int_0^t \mu_1(\tau)\, d\tau.$$

One easily solves the system of equations (4.6) in the order in which they are presented to obtain

$$C_1 = (y - 1)e^{-R(t)},$$

$$C_2 = (x - 1)e^{-[M(t)+N(t)]} + (y - 1)e^{-R(t)} \int_0^t \nu_{12}(\tau)e^{R(\tau)-M(\tau)-N(\tau)}\, d\tau, \tag{4.8}$$

$$C_3 = \psi(x, y, t)e^{-V(x,y,t)},$$

where C_1, C_2, C_3 are the so-called arbitrary constants, and where

$$\begin{aligned} V(x, y, t) = {} & (x - 1)e^{-[M(t)+N(t)]} \int_0^t e^{M(\tau)+N(\tau)}\lambda(\tau)\, d\tau \\ & + (y - 1)e^{-R(t)} \left[\int_0^t e^{M(\tau)+N(\tau)}\lambda(\tau)\, d\tau\right] \left[\int_0^t \nu_{12}(\tau)e^{R(\tau)-M(\tau)-N(\tau)}\, d\tau\right] \\ & - (y - 1)e^{-R(t)} \int_0^t e^{M(\alpha)+N(\alpha)}\lambda(\alpha)\left[\int_0^\alpha \nu_{12}(\tau)e^{R(\tau)-M(\tau)-N(\tau)}\, d\tau\right] d\alpha. \end{aligned} \tag{4.9}$$

The general solution of (4.5) is obtained as

$$C_3 = \Omega(C_1, C_2), \tag{4.10}$$

where $\Omega(C_1, C_2)$ is an arbitrary function of C_1 and C_2. The precise form of Ω is obtained in the following way. One first notices that, by (4.7),

$$F(0) = G(0) = N(0) = R(0) = M(0) = 0; \tag{4.11}$$

using this fact in (4.9) we obtain

$$V(x, y, 0) = 0 \qquad \text{for all } (x, y). \tag{4.12}$$

In (4.8), if we let $t = 0$ we obtain

$$C_3 = 1, \qquad C_2 = x - 1, \qquad C_1 = y - 1. \tag{4.13}$$

If we substitute (4.13) into (4.10) we obtain $1 \equiv \Omega(y - 1, x - 1)$, from which we deduce that $C_3 = \Omega(C_1, C_2) = 1$. At last we obtain

$$\psi(x, y, t) = e^{V(x,y,t)}, \tag{4.14}$$

where $V(x, y, t)$ is defined in (4.9).

We briefly investigate the expression for $V(x, y, t)$ in (4.9). We notice that $V(x, y, t)$ is a linear combination of $x - 1$ and $y - 1$. The coefficient of $x - 1$ is

$$\theta(t) = e^{-M(t)-N(t)} \int_0^t e^{M(\tau)+N(\tau)} \lambda(\tau)\, d\tau. \tag{4.15}$$

By integration by parts the coefficient of $y - 1$ may be more conveniently expressed by

$$\varphi(t) = e^{-R(t)} \int_0^t \nu_{12}(\tau) e^{R(\tau)-M(\tau)-N(\tau)} \left[\int_0^\tau \lambda(\omega) e^{M(\omega)+N(\omega)}\, d\omega\right] d\tau. \tag{4.16}$$

Thus we may write the joint probability generating function of $X_1(t)$ and $X_2(t)$ as

$$\psi(x, y, t) = \exp\{\theta(t)(x - 1) + \varphi(t)(y - 1)\}, \tag{4.17}$$

where $\theta(t)$ and $\varphi(t)$ are given by (4.15) and (4.16).

It is seen immediately at each time t that $X_1(t)$ and $X_2(t)$ are independent random variables, each distributed according to a Poisson distribution with parameters $\theta(t)$ and $\varphi(t)$ respectively. It should be emphasized at this point that $X_1(t)$ and $X_2(t)$ are not independent stochastic processes.

We now find the joint distribution of $X_1(t), \cdots, X_r(t)$ for arbitrary $r \geqq 2$. For this purpose, let $\mathbf{n} = (n_1, \cdots, n_r)$ denote an r-tuple of integers, let $\mathbf{1} = (1, 1, \cdots, 1)$, and let $\mathbf{N}_i = (0, \cdots, 0, 1, 0, \cdots, 0)$, where 1 is in the ith place and the jth component is 0 if $j \neq i$. Then if we let

$$p(\mathbf{n}, t) = P\left\{\bigcap_{i=1}^{r} [X_i = n_i]\right\}, \tag{4.18}$$

we obtain in the same way we derived (4.1) the system

$$\begin{aligned}\frac{d}{dt} p(n + 1, t) &= -\left[\lambda(t) + (n_r + 1)\mu_r(t) + \sum_{i=1}^{r-1} (n_i + 1)(\mu_i(t) + \nu_{i,i+1}(t))\right] p(\mathbf{n} + \mathbf{1}, t) \\ &\quad + \lambda(t) p(\mathbf{n} + \mathbf{1} - \mathbf{N}_1, t) + \sum_{i=1}^{r} (n_i + 2)\mu_i(t) p(\mathbf{n} + \mathbf{1} + \mathbf{N}_i, t) \\ &\quad + \sum_{i=1}^{n-1} (n_i + 2)\nu_{i,i+1}(t) p(\mathbf{n} + \mathbf{1} + \mathbf{N}_i - \mathbf{N}_{i+1}, t).\end{aligned} \tag{4.19}$$

It should be understood that if any m_i in $\mathbf{m}$ is negative, then $p(\mathbf{m}, t) = 0$. Now let us multiply both sides of (4.19) by

$$\prod_{i=1}^{r} x_i^{n_i+1} \tag{4.20}$$

and sum up both sides for all n_i from $n_i = -1$ to ∞. Let us denote the joint probability generating function of $\mathbf{X}(t) = [X_1(t), \cdots, X_r(t)]$ by

$$\psi(\mathbf{x}, t) = E\left[\prod_{i=1}^{r} x_i^{X_i(t)}\right]. \tag{4.21}$$

After some rearranging we obtain the linear partial differential equation

$$\frac{\partial}{\partial t}\psi(\mathbf{x}, t) + \sum_{i=1}^{r-1}[\nu_{i,i+1}(t)(x_i - x_{i+1}) + \mu_i(t)(x_i - 1)]\frac{\partial}{\partial x_i}\psi(\mathbf{x}, t) + \mu_r(t)(x_r - 1)\frac{\partial}{\partial x_r}\psi(\mathbf{x}, t) = \lambda(t)(x_1 - 1)\psi(\mathbf{x}, t). \tag{4.22}$$

We solve (4.22) by solving the system

$$\begin{aligned}\frac{dx_r}{dt} &= \mu_r(t)(x_r - 1)\\ \frac{dx_i}{dt} &= \mu_i(t)(x_i - 1) + \nu_{i,i+1}(t)[(x_i - 1) - (x_{i+1} - 1)]\end{aligned} \tag{4.23}$$

for $i = r - 1, r - 2, \cdots, 2, 1$, and

$$\frac{d\psi}{dt} = \lambda(t)(x_1 - 1)\psi. \tag{4.24}$$

One first observes that the variables x_i always occur in (4.23) and (4.24) in the form $(x_i - 1)$. On solving the first equation of the system, we obtain that $x_r - 1$ equals an arbitrary constant times a nonzero function of t. If, then, the system (4.23) is solved successively when $i = r - 1$, $i = r - 2, \cdots, i = 1$, it is easily seen that each $(x_i - 1)$ is a linear combination of $C_i, C_{i+1}, \cdots, C_r$, the coefficients being certain functions of t. Also, when $t = 0$, then $C_i = x_i - 1$, $i = 1, 2, \cdots, r$. The last equation in (4.23) is solved to yield

$$\log \psi(x, t) - \log C_0 = \begin{cases}\text{some linear combination}\\ \text{of } C_1, C_2, \cdots, C_r.\end{cases} \tag{4.25}$$

If we start with each C_r in the right side of (4.25) and in turn replace each C_i by what it is equal to in terms of the $(x_i - 1)$ and functions of t, we obtain

$$\log \psi(\mathbf{x}, t) - \log C_0 = V(\mathbf{x}, t) \tag{4.26}$$

where $V(\mathbf{x}, t)$ is a linear combination of $x_1 - 1, \cdots, x_r - 1$, or

$$\psi(\mathbf{x}, t) = C_0 \exp [V(\mathbf{x}, t)]. \tag{4.27}$$

In obtaining (4.25) it is seen that if $t = 0$, then $C_0 = 1$. Thus, as in the case of $r = 2$, we know that $V(\mathbf{x}, 0) = 0$. The general solution of (4.23) is

$$C_0 = \Omega(C_1, \cdots, C_r), \tag{4.28}$$

where Ω is an arbitrary function. If we let $t = 0$, this equation becomes

$$1 \equiv \Omega(x_1 - 1, x_2 - 1, \cdots, x_r - 1), \tag{4.29}$$

and thus we deduce that

$$\psi(\mathbf{x}, t) = \exp [V(\mathbf{x}, t)], \tag{4.30}$$

where $V(\mathbf{x}, t)$ may be represented as

$$V(\mathbf{x}, t) = \sum_{i=1}^{r}\theta_i(t)(x_i - 1), \tag{4.31}$$

and where the $\theta_i(t)$ have continuous derivatives with respect to t. Thus we have shown that for each t, the $X_1(t), \cdots, X_r(t)$ are independent random variables, each distributed according to a Poisson distribution and where the Poisson parameter for $X_i(t)$ is $\theta_i(t)$. As was remarked before in the case for $r = 2$, these r stochastic processes are not independent but are dependent. It is just that the r random variables at each time t are independent.

The problem remains to find the functional form of each $\theta_i(t)$. These could be obtained by the tiresome procedure of solving system (4.23). However, we may obtain each $\theta_i(t)$ more easily by reasoning with expectations. Indeed, one may write

$$\theta_1(t + h) = \theta_1(t) + E[\Delta X_1(t)], \tag{4.32}$$

where $\Delta X_1(t) = X(t + h) - X(t)$. For any nonnegative integer n, if it is known that $X_1(t) = n$, then $\Delta X_1(t)$ takes values 1, 0, and -1 with probabilities $\lambda(t)h + o(h)$, $1 - \lambda(t)h - n[\mu_1(t) + \nu_{12}(t)] + o(h)$, and $n[\mu_1(t) + \nu_{12}(t)]h + o(h)$ respectively, and takes any other value with probability $o(h)$. Thus

$$E[\Delta X_1(t)|X_1(t) = n] = \{\lambda(t) - n[\mu_1(t) + \nu_{12}(t)]\}h + o(h), \tag{4.33}$$

and

$$\begin{aligned} E[\Delta X_1(t)] &= \sum_{n=0}^{\infty} E[\Delta X_1(t)|X_1(t) = n]P\{X_1(t) = n\} \\ &= \{\lambda(t) - \theta_1(t)[\mu_1(t) + \nu_{12}(t)]\}h + o(h). \end{aligned} \tag{4.34}$$

This gives us the differential equation

$$\frac{d}{dt}\theta_1(t) + [\mu_1(t) + \nu_{12}(t)]\theta_1(t) = \lambda(t). \tag{4.35}$$

For $i = 2, \cdots, r - 1$ we again consider

$$\theta_i(t + h) = \theta_i(t) + E[\Delta X_i(t)], \tag{4.36}$$

where $\Delta X_i(t) = X_i(t + h) - X_i(t)$. In this case let us note that if $X_{i-1}(t) = m$ and $X_i(t) = n$, then $\Delta X_i(t)$ takes values $+1$, 0, and -1 with conditional probabilities $m\nu_{i-1,i}(t)h + o(h)$, $1 - m\nu_{i-1,i}(t) - n[\mu_i(t) + \nu_{i,i+1}(t)]h + o(h)$ and $n[\mu_i(t) + \nu_{i,i+1}(t)]h$ respectively, and all other values with probability $o(h)$. Since $X_{i-1}(t)$ and $X_i(t)$ are independent we may write

$$\begin{aligned} &E[\Delta X_i(t)] \\ &= \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} E[\Delta X_i(t)|X_{i-1}(t) = m, X_i(t) = n]P\{X_{i-1}(t) = m\}P\{X_i(t) = n\}. \end{aligned} \tag{4.37}$$

Thus we obtain, for $i = 2, 3, \cdots, r - 1$,

$$\frac{d}{dt}\theta_i(t) + [\mu_i(t) + \nu_{i,i+1}(t)]\theta_i(t) = \nu_{i-1,i}(t)\theta_{i-1}(t). \tag{4.38}$$

In a similar manner,

$$\frac{d}{dt}\theta_r(t) + \mu_r(t)\theta_r(t) = \nu_{r-1,r}(t)\theta_{r-1}(t). \tag{4.39}$$

Under conditions of continuity imposed on the parameters of the process, the system of linear differential equations (4.35), (4.38) and (4.39) are easily solved. Since $\theta_1(0) = \cdots = \theta_r(0) = 0$, then

(4.40)

$$\theta_1(t) = \exp\left\{-\int_0^t [\mu_1(\tau) + \nu_{12}(\tau)]\, d\tau\right\} \int_0^t \lambda(\tau) \exp\left\{\int_0^\tau [\mu_1(\omega) + \nu_{12}(\omega)]\, d\omega\right\} d\tau.$$

For $i = 2, \cdots, r-1$, we have

$$\theta_i(t) = \exp\left\{-\int_0^t [\mu_i(\tau) + \nu_{i,i+1}(\tau)]\, d\tau\right\} \int_0^t \nu_{i-1,i}(\tau)\, \theta_{i-1}(\tau) \exp\left\{\int_0^t [\mu_i(\omega) + \nu_{i,i+1}(\omega)]\, d\omega\right\} d\tau \tag{4.41}$$

and

$$\theta_r(t) = \exp\left[-\int_0^t \mu_r(\tau)\, d\tau\right] \int_0^t \nu_{r-1,r}(\tau)\theta_{r-1}(\tau) \exp\left[\int_0^\tau \mu_r(\omega)\, d\omega\right] d\tau. \tag{4.42}$$

It might be of interest to find the limiting expectations of $\mathbf{X}(t)$ as $t \to \infty$ when the $\mu_i(t)$, the $\nu_{i,i+1}(t)$ and $\lambda(t)$ are constant (or asymptotically constant). Accordingly, let $\mu_i(t) = \mu_i$, $\nu_{i,i+1}(t) = \nu_{i,i+1}$, $\lambda(t) = \lambda$, that is, these $2r$ functions become nonnegative constants. Then

$$\theta_1(t) = \begin{cases} \lambda t, & \mu_1 + \nu_{12} = 0, \\ \dfrac{\lambda}{\mu_1 + \nu_{12}} \{1 - \exp[-(\mu_1 + \nu_{12})t]\}, & \mu_1 + \nu_{12} > 0. \end{cases} \tag{4.43}$$

In the case that $\mu_i + \nu_{i,i+1} = 0$, then $X_{i+1}(t), \cdots, X_r(t)$ are all identically zero in t. We exclude this case from now on. Thus

$$\lim_{t\to\infty} \theta_1(t) = \frac{\lambda}{\mu_1 + \nu_{12}}. \tag{4.44}$$

One could perform the successive laborious integrations on the expressions in (4.41) and (4.42) and obtain $\lim_{t\to\infty} \theta_i(t)$ for $i = 2, \cdots, r$. We find these limits by a different and easier method. This easier method depends on the

LEMMA: *If $g(t)$ is Lebesgue integrable over every bounded interval $[0, T]$, and if $\lim_{t\to\infty} g(t) = K$, then*

$$\lim_{t\to\infty} e^{-t} \int_0^t g(\tau) e^\tau\, d\tau = K. \tag{4.45}$$

PROOF. It is clearly sufficient to prove this only in the case $K = 0$. Let $\epsilon > 0$ be arbitrary, and let $T_0 > 0$ be large enough so that $|g(\tau)| < \epsilon/2$ for all $\tau > T_0$. If $g(t) \equiv 0$ over $[0, T_0]$, let $T_1 = T_0$; if not, let $T_1 > T_0$ be large enough so that for all $t > T_1$,

$$0 < e^{-t} < \frac{\epsilon}{2}\left[\int_0^{T_0} |g(\tau)| e^\tau\, d\tau\right]^{-1}. \tag{4.46}$$

Hence for all $t > T_1$,

$$|e^{-t}\int_0^t g(\tau)e^{\tau}\,d\tau| \leqq |e^{-t}\int_0^{T_0} g(\tau)e^{\tau}\,d\tau| + |e^{-t}\int_{T_0}^t g(\tau)e^{\tau}\,d\tau| \tag{4.47}$$

$$\leqq \frac{\epsilon}{2} + \frac{\epsilon}{2}e^{-t}(e^t - e^{T_0}) < \epsilon,$$

which proves the assertion.

Using the above lemma and (4.44) in (4.41) and (4.42), we obtain for $i = 2, 3, \cdots, r-1$,

$$\lim_{t\to\infty} \theta_i(t) = \frac{\lambda\nu_{12}\nu_{23}\cdots\nu_{i-1,i}}{\prod_{j=1}^{i}(\mu_j + \nu_{j,j+1})}, \tag{4.48}$$

and

$$\lim_{t\to\infty} \theta_r(t) = \frac{\lambda\prod_{i=2}^{r}\nu_{i-1,i}}{\mu_r\prod_{j=1}^{r-1}(\mu_j + \nu_{j,j+1})}. \tag{4.49}$$

Thus we have shown that the expected numbers of particles of each kind is asymptotically constant when the parameters are constant.

5. Derivation of the joint distribution of $X(t)$, $Y(t)$, $Z(t)$

It should be recalled that although the subject of this study is a stochastic process with continuous time, it cannot be observed as such. In other words, the experimenter cannot observe an entire sample function over a time interval. Even worse, he cannot make more than one observation on a particular sample function. In the stochastic process considered here each mouse constitutes a sample function. Thus, in the experimentation described in section 1, the experimenter selects m values of t, say $0 < t_1 < \cdots < t_m$ and decides to select n_i mice at random at time t_i. If $\mathbf{W}(t) = [X(t), Y(t), Z(t)]$, then he observes $\mathbf{W}_{i,1}, \cdots, \mathbf{W}_{i,n_i}$, n_i independent observations on $\mathbf{W}(t_i)$. The totality of random vectors $\{\mathbf{W}_{i,j}, 1 \leqq i \leqq m, 1 \leqq j \leqq n_i\}$ are independent. These are the only observations possible, and the problems of estimating the parameters and verifying the model can only be in terms of these observations. From these considerations, it is seen that our sole distributional problem is that of deriving the joint probability distribution of $[X(t), Y(t), Z(t)]$ at any time t. The purpose of this section is to derive this joint distribution by finding the formula of the joint probability generating function.

An immediate application of the results developed for mutation processes in section 4 to the model for the two-stage theory of carcinogenesis constructed in section 2 yields the joint probability generating function of $U(t)$, $X(t)$, $Y(t)$ to be

$$E[u^{U(t)}x^{X(t)}y^{Y(t)}] = \exp[\theta_0(t)(u-1) + \theta(t)(x-1) + \varphi(t)(y-1)], \tag{5.1}$$

where $U(t)$ denotes the number of unobservable hyperplastic foci at time t, and

$\theta_0(t) = EU(t)$. Using (4.35), (4.38) and (4.39) the expectations $\theta_0(t)$, $\theta(t)$ and $\varphi(t)$ are related by the following differential equations:

$$\begin{aligned} \frac{d\theta_0(t)}{dt} + (\mu_0 + \nu_0)\theta_0(t) &= f(t), \\ \frac{d\theta(t)}{dt} + [\mu + \nu(t)]\theta(t) &= \nu_0\theta_0(t), \\ \frac{d\varphi(t)}{dt} &= \nu(t)\theta(t). \end{aligned} \tag{5.2}$$

The solutions for $\theta_0(t)$, $\theta(t)$ and $\varphi(t)$ are obtained easily using (4.40), (4.41), and (4.42). We shall have no need for these solutions and shall only make use of (5.2).

Let $Z_j(t)$ denote the number of cells at time t in the jth tumor to be formed on a mouse's lungs. If the jth tumor was formed (or born) after time t, then $Y_j(t) = 0$. Because of our assumptions that these tumors are stochastically independent, it follows that $\{Z_1(t), Z_2(t), \cdots\}$ are independent. Further

$$Z(t) = Z_1(t) + Z_2(t) + \cdots + Z_{Y(t)}(t) \tag{5.3}$$

is the total number of cells in all the tumors at time t. Let τ_j denote the time at which the jth tumor was born, and let the event H be defined as follows:

$$H = [X(t) = m, Y(t) = n, \tau_1 = t_1, \cdots, \tau_n = t_n], \tag{5.4}$$

where $0 < t_1 < \cdots < t_m < t$. Also, let $g(z)$ denote the probability generating function of the random variable N discussed in section 2. If $f_H(z)$ denotes the conditional probability generating function of $Z(t)$ given H, then easily

$$f_H(z) = g^n(z) \exp\,(z - 1) \sum_{i=1}^{n} \int_{t_i}^{t} \rho(\alpha)\, d\alpha, \tag{5.5}$$

where $\rho(\alpha)$ is as described in section 2. Also, the conditional probability density of $\tau_1, \cdots, \tau_n$ given that $X(t) = m$ and $Y(t) = n$ is given by

$$\begin{aligned} &f(t_1, \cdots, t_n | X(t) = m, Y(t) = n) \\ &\quad = \begin{cases} \dfrac{n! \prod_{i=1}^{n} \varphi'(t_i)}{\varphi^n(t)}, & \text{if } 0 < t_1 < \cdots < t_n < t, \\ 0, & \text{otherwise.} \end{cases} \end{aligned} \tag{5.6}$$

Using this, one finally arrives at the expression for $\psi(x, y, z)$, the joint probability generating function of $X(t)$, $Y(t)$, $Z(t)$,

$$\begin{aligned} &\psi(x, y, z) \\ &\quad = \exp\left\{\theta(t)(x - 1) + yg(z) \int_0^t \exp\left[(z - 1) \int_w^t \rho(v)\, dv\right] \varphi'(w)\, dw - \varphi(t)\right\}. \end{aligned} \tag{5.7}$$

It will be useful in the sequel to have formulas for the expectation and variance of $Z(t)$, their corresponding derivatives, and the covariance of $Y(t)$ and $Z(t)$. From (5.7) we therefore deduce the following formulas

$$E[Z(t)] = \int_0^t \left[\int_w^t \rho(v)\,dv\right] \varphi'(w)\,dw + E(N)\varphi(t), \tag{5.8}$$

or, using integration by parts,

$$E[Z(t)] = \int_0^t \varphi(w)\rho(w)\,dw + E(N)\varphi(t). \tag{5.9}$$

Thus

$$\frac{d}{dt} E[Z(t)] = \varphi(t)\rho(t) + E(N)\varphi'(t). \tag{5.10}$$

The variance of $Z(t)$ is given by

$$\begin{aligned}\sigma^2[Z(t)] = {} & \int_0^t \left(\int_w^t \rho(v)\,dv\right)^2 \lambda(w)\,dw \\ & + \{E[Z(t)] - E(N)\varphi(t)\}[2E(N) + 1] \\ & + \sigma^2(N)\varphi(t) + [E(N)]^2\varphi(t).\end{aligned} \tag{5.11}$$

The covariance of $Y(t)$ and $Z(t)$ is obtained by the formula

$$\text{Cov}\,[Y(t), Z(t)] = \frac{\partial^2}{\partial y\,\partial z}\psi(1, y, z)\Big|_{y=z=1} - E[Y(t)]E[Z(t)]. \tag{5.12}$$

Applying this formula we obtain

$$\text{Cov}\,[Y(t), Z(t)] = E(N)\varphi(t) + \int_0^t \varphi(w)\rho(w)\,dw. \tag{5.13}$$

An important relationship to note is that

$$\frac{\text{Cov}\,[Y(t), Z(t)]}{E[Z(t)]} = 1. \tag{5.14}$$

6. Tests of the model and estimates of the parameters

In order to suggest tests for the model developed in the previous sections, the type of experimental observations should be specified. The mice involved in the experiment should be divided into two populations which we shall call population I and population II. The mice in population I should each be given the same heavy dose of urethan (or whatever carcinogen is being used), and the mice of population II should each be given the same light dose of urethan. The "principal times" for sampling each of the two populations may be denoted by $0 < t_1 < \cdots < t_m$, where the t are more or less equally spaced, with the t being (perhaps) a little more frequent during the first 60 days. However, each t_i really refers to three successive days of sampling. (The units of t are in days.) To be more specific, the $3m$ days of sampling are $t_1 - 1, t_1, t_1 + 1, \cdots, t_m - 1, t_m, t_m + 1$. On each of these days, n mice should be selected from each population. They should all be sacrificed at the same time and at the same time of day for each day's sampling experiment. Each mouse's lungs should be examined in order to determine the following three numbers: the number of hyperplastic foci, the

number of tumors, and the total number of cells in (or the total volume of) the tumors. Then at each t_i the parameters $\theta(t_i)$, $\varphi(t_i)$, $EZ(t_i)$ and their derivatives can be estimated in the obvious way for each of the two populations.

The first test to make is that $X(t_i)$ and $Y(t_i)$ are independent and each with a Poisson distribution. Polissar and Shimkin [5] report that for low doses of carcinogen $Y(t_i)$ is approximately Poisson distributed, while the fit does fall off for large doses. They suggest, however, that this might be due to the heterogeneity of the mice and conclude that if a pure strain were used, then at larger doses the distribution of tumors would be Poisson. The tests for independence and for a sample being from a Poisson distribution are standard and need not be elaborated on here.

Once this is done and the conclusion of it is successful, one should compute estimates for Cov $[Y(t_i), Z(t_i)]/E[Z(t_i)]$, for $i = 1, \cdots, m$ in the obvious way and verify that these estimates "hover" around unity. [See (5.13) for this requirement.] We shall call this the second test of the model despite the fact that it is not exactly a test of hypothesis. If the first and second tests prove to be successful, then one may proceed to test the entire model.

In order to test the model itself, we must first estimate EN and the function $\rho(t)$. We do this by using equation (5.10). In this equation, estimates can be constructed in the obvious manner for $dEZ(t)/dt$ at $t = t_i$, $\varphi(t_i)$ and $\varphi'(t_i)$ for each of the two populations. Since $\rho(t)$ and EN do not depend on the dosage (see section 2), we have at each t_i two equations in the two unknowns $\rho(t_i)$ and EN. Thus $\rho(t_i)$ and EN can be estimated at each t_i. The function $\rho(t)$ can then be estimated through appropriate curve fitting, and EN can be estimated in an over-all manner.

Let (x_0, y_0, z_0) be a triplet of numbers, each unequal to unity. Suggested values are for those greater than unity. Now it is possible to consider $\psi_i(x_0, y_0, z_0|t_j, \hat{\Omega})$. This is $\psi(x_0, y_0, z_0)$ of (5.7) with $\theta(t)$, $\varphi(t)$, $\varphi'(t)$, $\rho(t)$ and EN in $g(z_0) = \exp(EN)(z_0 - 1)$ {or $g(z_0) = [(1 + z_0)/2]^{2EN}$} replaced by their estimates for population i, for $i = 1, 2$. At time t_i, let $[X_{ij}^{(k)}, Y_{ij}^{(k)}, Z_{ij}^{(k)}]$, with $1 \leqq j \leqq n$, denote n independent observations on $[X(t_i), Y(t_i), Z(t_i)]$ on the kth population. Let

$$J_n^{(k)}(i, \alpha) = n^{-1} \sum_{j=1}^{n} x_0^{\alpha X_{ij}^{(k)}} y_0^{\alpha X_{ij}^{(k)}} z_0^{\alpha Y_{ij}^{(k)}} \tag{6.1}$$

and let, for each i,

$$W_n(x_0, y_0, z_0|k, t_i) = \frac{J_n^{(k)}(i, 1) - \psi_k(x_0, y_0, z_0|t_i, \Omega)}{(n^{-1}\{J_n^{(k)}(i, 2) - [J_n^{(k)}(i, 1)]^2\})^{1/2}}. \tag{6.2}$$

For large values of n, this statistic should be approximately normally distributed with mean zero and variance unity, *provided that the model is correct.* (For development of the theory behind the use of this statistic and for an application, see [1].) The inference made is as follows. Compute W_n for each population at each of the t_i and for convenient values of (x, y, z). If the array of these values appears to have come from a normal population with mean zero and variance possibly even a little greater than unity, then we conclude that there is no over-

whelming evidence for rejecting the model. If, on the other hand, the values of W_n are nowhere in or near the interval $(-3, +3)$, then we conclude that this model does not adequately fit the mechanism for which it was intended, and consequently it must be altered. The test just described will be called the third test of the model.

If the results of the first, second and third tests seem promising and do not give us violent cause for rejecting the model, we may proceed to obtain crude estimates for $\nu(t)$, μ and $f(t)$. From the third equation in (5.2) we can estimate $\nu(t_i)$ in an obvious way for each i and each population, and from this, $\nu(t)$ can be approximated by drawing a smooth curve through the points.

In order to estimate μ, we first must settle a question concerning $\theta_0(t)$, the expected number of unobservable hyperplastic foci at time t. It is safe to assume that for large values of t, say $t \geqq 100$ days, $\theta_0(t)$ vanishes. If one looks at column 3 on page 80 of Polissar and Shimkin [6], and if one keeps in mind that 10 square cellular units is still visible under the microscope, then this assumption becomes a reasonable one to make. Thus, for a $t_i \geqq 100$ days, and assuming $\theta_0(t_i) = 0$ for both populations, one can use the second differential equation of (5.2) to obtain an estimate for μ in the obvious way. Then one can estimate $\nu_0\vartheta_0(t_i)$ for all t_i by using the second differential equation of (5.2).

We now estimate $\nu_0 f(t)$. Let $s_j = (t_{j-1} + t_j)/2$. Then $\nu_0\theta_0'(s_j)$ and $\nu_0\theta_0(s_j)$ can be estimated by

$$\nu_0\theta_0'(s_j) \fallingdotseq \frac{\nu_0\theta_0(t_j) - \nu_0\theta_0(t_{j-1})}{t_j - t_{j-1}}, \tag{6.3}$$

$$\nu_0\theta_0(s_j) \fallingdotseq \frac{1}{2}[\nu_0\theta_0(t_{j-1}) + \nu_0\theta_0(t_j)]. \tag{6.4}$$

The first equation in (5.2) can be rewritten

$$\nu_0 \frac{d\theta_0(t)}{dt} + (\mu_0 + \nu_0)\nu_0\theta_0(t) = \nu_0 f(t). \tag{6.5}$$

At each s_j we have estimates now for $\nu_0\theta_0'(s_j)$ and $\nu_0\theta_0(s_j)$ for each population. Let us suppose that the dosage of urethan in population II is K times the dosage of urethan administered to each mouse in population I. Then for every s_j we may write the two equations

$$\begin{aligned} -(\mu_0 + \nu_0)\nu_0\theta_0(s_j) + \nu_0 f(s_j) &= \nu_0\theta_0'(s_j) \\ -(\mu_0 + \nu_0)\nu_0\theta_0(s_j) + \nu_0 K f(s_j) &= \nu_0\theta_0'(s_j) \end{aligned} \tag{6.6}$$

for populations I and II respectively. A word of explanation is necessary now. The number $f(s_j)$ refers to the value of the feeder function at time s_j for population I only. The pair of numbers $\nu_0\theta_0(s_j)$ and $\nu_0\theta_0'(s_j)$ are not the same for the two equations of (6.6). In the first equation these have the values computed for population I, and in the second equation they have the values computed for population II. Thus (6.6) gives us two equations in the two unknowns, $\mu_0 + \nu_0$ and

$\nu_0 f(s_j)$. In this way estimates are obtainable for both $\mu_0 + \nu_0$ and $\nu_0 f(s_j)$. Thus we can estimate the curve of $f(t)$ except for a multiplicative constant.

REFERENCES

[1] E. K. DALTON, W. D. JAMES, and H. G. TUCKER, "Stochastic models for the electron multiplier tube," to be published.
[2] W. FELLER, *An Introduction to Probability Theory and Its Applications*, Vol. I, New York, Wiley, 1957 (2nd ed.).
[3] D. G. KENDALL, "Birth and death processes, and the theory of carcinogenesis," *Biometrika*, Vol. 47 (1960), pp. 13-21.
[4] J. NEYMAN, "A stochastic model for carcinogenesis," mimeographed lecture presented at The National Institutes of Health, 1958.
[5] M. J. POLISSAR and M. B. SHIMKIN, "A quantitative interpretation of the distribution of induced pulmonary tumors in mice," *J. Nat. Cancer Inst.*, Vol. 15 (1954), pp. 377–403.
[6] ———, "Some quantitative observations on the induction and growth of primary pulmonary tumors in strain A mice receiving urethan," *J. Nat. Cancer Inst.*, Vol. 16 (1955), pp. 75–97.
[7] ———, "Growth of pulmonary tumors in mice of strains A and C3H," *J. Nat. Cancer Inst.*, Vol. 21 (1958), pp. 595–610.
[8] W. SULLIVAN, "New cell clues found in cancer," *New York Times*, January 16. 1960, p. 10.