A Branching-Process Model for Heterogeneous Cell Populations

ROGER S. DAY
Department of Biostatistics, Harvard School of Public Health, Boston, Massachusetts 02115

ABSTRACT

A multitype branching-process model is introduced for the growth of heterogeneous cell populations. This model includes events representing mitosis, death, mutation, and conversion from one cell type to another. Formulas for conditioning on interim events, and generalizations allowing parameters to be functions of time or cell counts, are presented. The probability generating function (p.g.f.) is solved approximately in a way that is both accurate and efficient enough to solve important problems in tumor biology. The uses of the p.g.f. in the setting of clinical oncology are described.

1. INTRODUCTION

A stochastic model for heterogeneous cell populations is presented, along with a solution. The novelty is not in the model, which is a particular subset of branching processes and a generalization of the "Goldie-Coldman model" for tumor growth and treatment [1]. The solution, however, is new. Though approximate, it is very useful for some important problems which other techniques cannot handle.

Guidelines for using the solution in applications are proposed as well. The practical use of the probability generating function (p.g.f.) has usually been mediated through the calculation of cumulants. Arguments are presented here for using other distributional summaries derived from the p.g.f. rather than the cumulants.

The range of application includes phenomena in which rare events (such as mutation) determine the outcome. It excludes phenomena in which the details of the cell cycle are important (such as fitting growth-fraction data or modeling the dissipation of cell synchronization). The techniques presented here complement the asymptotic techniques of branching-process theory, by providing finite-time solutions.

The work was motivated by the need for a flexible tool for understanding how tumor biology affects cancer-treatment choice and timing. (The presentation will be described in terms of this application, though the model has

MATHEMATICAL BIOSCIENCES 78:73–90 (1986) 73

©Elsevier Science Publishing Co., Inc., 1986
52 Vanderbilt Ave., New York, NY 10017 0025-5564/86/$03.50
applications in other contexts as well.) Goldie and Coldman [1, 2] stimulated
the interest of the oncology community in the mathematical modeling of
tumor heterogeneity, by demonstrating important consequences for treat-
ment choice in special circumstances. The robustness of their conclusions
under relaxed assumptions remained unanswered. In recent years the focus
has expanded to other aspects of the heterogeneity of tumors, such as
metastatic potential, growth rates, antigen expression, and degree of differ-
etiation [3–6]. New tools were required to do a “sensitivity analysis” not
just by varying parameters but by varying the population structure itself,
allowing more complex biology to enter the model. With the tools described
herein, current treatment strategies have been evaluated, and new ones
developed [7].

Mathematical thinking about the growth of cell colonies and mutation
became serious in the late 1940s. Luria and Delbrück [8] developed an
experimental technique, the fluctuation test, to measure bacterial rates of
mutation to drug resistance, and based the estimator on a simple mathemat-
ical model. This brought many statisticians to study a variety of estimators
and variations on the model [9–11]; these efforts still evolve [12–14]. Another
path was traversed as probabilists developed the theory of branching processes
[15–21], a class of models seemingly suitable for a huge class of biological,
and especially genetic, problems. These results are almost exclusively asym-
totic, and therefore of little use in modeling tumors. A third path has been
inspired by increased computational power. Cell-by-cell Monte Carlo com-
puter simulations have been attempted, but seem too costly in time and
memory to yield many practical results [22, 23].

The solution derived here is easily computed and useful for general
population structures. It allows rapid evaluation of the consequences of
assumptions and parameters, and makes model fitting and optimization
feasible. Results can be computed conditionally on interim events, allowing
the fitting of serial data. Also, parameters can be functions of time and cell
counts, allowing feedback among subpopulations, as in nonexponential
growth, immune responses, or the production of growth or inhibitory factors.

Phenomena which are easy to handle include heterogeneity in cell kinetics,
mutation, metastasis, cell death due to treatment, and simple heterogeneity in
drug sensitivity. In addition, quiescence, stem-cell differentiation, and gene
amplification can be modeled, although numerical problems can arise. One
can test the importance of the tumor population structure as well as individ-
ual parameter values, identifying the biological information needed to make
the best treatment choice. When sufficient information is available, one may
find near-optimal treatment strategies. (It is not necessary that predictions of
“cure rates” be accurate in the clinic; model-based comparisons of treatment
strategies are more likely to conform to the real world as the model becomes
more accurate in its details.) As better measures of cell subpopulations

become available (especially assays for drug resistance [24, 25]), the ability to fit serial data may become invaluable.

The model also allows estimation of parameters in \textit{in vitro} or animal experiments for which the population structure and most parameters are fairly well known.

This paper describes the basic model, derives the joint probability generating function (p.g.f.) for the number of cells of each type, and describes the use of the p.g.f. in typical applications. Finally, the assumptions, limitations, and potential uses of the model are discussed.

2. A DESCRIPTION OF THE MODEL

The basis for the model is a multitype branching process in continuous time (cf. Bellman and Harris [15]). A cell colony is imagined to contain a small number of subpopulations, each one being homogeneous (of a single "cell type"). Each individual cell exists until one of several events occur:

- (1) renewal: replacement by two cells of the same type as the original cell;
- (2) mutation: replacement by two cells, one of the same type and the other of a different type;
- (3) conversion: replacement by one cell of a different type;
- (4) death: disappearance of the cell.

The time to an event is assumed to be exponentially distributed. (This limits the choice of growth curves only slightly; See section 6.) The rates at which these events occur are allowed to depend on the cell type. The tumor cells present at any one instant are assumed to have statistically independent futures.

In addition to these kinetic events, we allow instantaneous death of cells at previously fixed time points (episodes), representing for example killing of tumor cells by treatment. In such an episode, each cell is assumed to have an independent chance to die (disappear), with a probability which depends on the cell type. The killing probabilities may vary with episode.

3. COMPUTATION OF THE PROBABILITY GENERATING FUNCTION

This section derives the probability generating function (p.g.f.) for the joint distribution of the number of cells of each cell type at a given time. The use of the p.g.f. is discussed in the next section.

Cell types will be denoted $X$ and $Y$, and the numbers of cells will be denoted $N_X$, $N_Y$, etc. A vector containing the number of cells of each cell type will be denoted $\mathbf{N} = (N_X, N_Y, \ldots)$. The joint probability generating function for counts of all cell types at time $t$, evaluated at the dummy
argument \( s = (\ldots, s_Y, \ldots) \), is defined by

\[
\Lambda_t(s) = \mathbb{E}\left( \prod_Y s_Y^{N_Y(t)} \right),
\]

where \( \mathbb{E} \) signifies expectation over the distribution of \( \mathbf{N}(t) \), and \( Y \) indexes the cell types. We assume that \( \Lambda_0 \) is known (e.g. \( \mathbf{N} \) is known at time 0).

We make extensive use of the following multivariate composition rule for the p.g.f. of a random sum of i.i.d. random variables. Let \( D_Y(X; j) \) signify the number of \( Y \)-type cells present at time \( u \) which are descendants of the \( j \)-th \( X \)-cell present at time \( t \). Then the number of cells of type \( Y \) at the end of an interval \( (u, v) \) is

\[
N_Y(v) = \sum_X \sum_{j=1}^{N_X(u)} D_Y(X; j).
\]

Let \( \mathbf{D}(X; j) = (\ldots, D_Y(X; j), \ldots) \) represent all the progeny, and let \( \Phi_X \) be its joint p.g.f., which depends on \( X \) but not on \( j \). (It also depends on the interval, but this is not central at the moment.) Assume the vectors \( \{\mathbf{D}(X; j)\} \) for all \( X \), all \( j = 1, \ldots, N_X(u) \) are independent. Then

\[
\Lambda_u(s) = \sum_n \text{prob}(\mathbf{N}(u) = n) \mathbb{E}\left( \prod_Y \prod_X \prod_{j=1}^{n_X} s_Y^{D_Y(X; j)} \right)
\]

\[ = \sum_n \text{prob}(\mathbf{N}(u) = n) \prod_X \left[ \Phi_X(s) \right]^{n_X} \]

\[ = \Lambda_u(\ldots, \Phi_X(s), \ldots).
\]

Here the \( X \)-coordinate of the argument is shown, and the expectation is over the distributions of the \( \mathbf{D} \)'s.

Let \( \tau_1 < \tau_2 < \cdots < \tau_m \) be a list of times at which treatment is given (epochs). The instant before or after treatment is symbolized by a superscripted \(-\) or \(+\). With the assumption that killing of cells of type \( X \) at \( \tau_i \) is binomial with proportion \( K_{iX} \), \( D_Y(X; j) \) is binomial for \( X = Y \) and zero otherwise. With \( u = \tau_i^- \) and \( v = \tau_i^+ \) the composition rule above becomes

\[
\Lambda_{\tau_i^+}(s) = \Lambda_{\tau_i^-}(\ldots, K_{iX} + (1 - K_{iX}) s_X, \ldots).
\]

Equation (1) also handles growth, mutation, and conversion in intervals between treatments (from \( u = \tau_i^- \) to \( v = \tau_i^- \)). Our assumption that cells at any given time have independent futures implies that any two \( \mathbf{D} \)'s representing progeny of two different cells will be independent. Therefore the composition rule (1) is valid between epochs as well.
It remains only to find \( \Phi_X \) for \( D(X; j) \) between epochs, based on the growth kinetics. (This depends on the duration of the interval between epochs; we now add this duration as a superscript for emphasis: e.g. \( \Phi_X^T \).

The solution for \( \Lambda, \) at \( s \) will then be computable by successive functional composition, in the following way. Define \( \Omega_i(s) = (\ldots, \Phi_{X_i}^\Delta i(s), \ldots) \) with \( \Delta i \) the interval length \( \tau_i^- - \tau_{i-1}^+ \), and let \( K_i(s) = (\ldots, K_{iX} + (1 - K_{iX})s_X, \ldots) \).

Then the joint p.g.f. for the entire process is

\[
\Lambda_i(s) = \Lambda_0(\Omega_0 \circ K_1 \circ \Omega_1 \circ \cdots \circ K_m \circ \Omega_m(s)). \tag{3}
\]

The rest of this section will derive \( \Phi_X^T \), the joint p.g.f. for all progeny of a single \( X \)-cell at the end of an interval of length \( T \) between epochs.

**DERIVATION OF \( \Phi \)**

Now consider a single time interval, of length \( T \), and a single \( X \)-cell at the start. We will obtain \( \Phi = \Phi_X^T(s) \), the joint p.g.f. for all progeny of this cell at time \( T \). For now, consider only the events occurring to \( X \)-cells. It will be useful in the exposition to introduce a couple of new terms: the events which leave no \( X \)-cells, death and conversion, we will call **culling events**, and the events which result in a non-\( X \)-cell, mutation and conversion, we will call **altering events**. Let \( \Theta = \Theta_{X,C,M,T} \) be the joint p.g.f. for the number of cells of type \( X \) at the end of the interval \( (N_X) \), the number of culling events \( (N_C) \), and the number of mutation events \( (N_M) \) in the interval. That is,

\[
\Theta_{X,C,M}(s_X,s_C,s_M) = E(s_X^{N_X}s_C^{N_C}s_M^{N_M}).
\]

At this point, an approximation is needed. For mathematical ease, only a single altering event is “allowed” in the interval in any single chain of descent. That is, in a chain of descent including two altering events in the same interval, the descendants past the second events will not be counted; they are “lost.” This is best thought of not as an “assumption,” which is biologically meaningless here, but instead as a conceptual trick to get a good approximate solution. If intervals are very long, or if mutation or conversion rates are very high, this approximation may lead to serious errors. However, one can improve the accuracy of the model by forcing the time intervals to be short. In practice, this is easily done by doing the functional iteration for \( \Lambda, \) Equation (3), at several time points within each treatment interval, as well as at the treatment times, i.e. by augmenting the list of epochs (the \( \tau \)'s) so that no interval is too long. The limitations introduced by this approximation are discussed in Section 6.

The number of \( Y \)-cells at the end of the interval is given by a sum of birth-death processes for \( Y \)-cells, each extending from the time of an event creating a \( Y \)-progenitor to the end of the interval. Let \( t \) denote the (random)
time of such a creating event, and let \( g(t) \) denote the probability density function for the time of the event; then the p.g.f. for the number of \( Y \)-cells produced as a result of a creating event occurring to an \( X \)-type cell in the interval is given by

\[
\Psi_Y(s_Y; X, T) = \int_0^T \Theta_{Y,C,M; (T-t)}(s_Y, 1, 1) g(t) \, dt.
\]

(4)

Here we are ignoring contributions to \( Y \) from any source except a creating event followed by a birth-death process; thus the appearance of \( \Theta_{Y,C,M} \).

Now let \( \alpha(X,Y) \) equal the rate of mutation from type \( X \) to type \( Y \), and \( \beta(X,Y) \) be the rate of conversion from type \( X \) to type \( Y \), and \( \delta(X) \) be the death rate for \( X \)-cells. Then the probability that a given mutation event yields a \( Y \)-cell is

\[
\frac{\alpha(X,Y)}{\sum_Y \alpha(X,Y)},
\]

(5a)

and the probability that a given culling event yields a \( Y \)-cell is

\[
\frac{\beta(X,Y)}{\delta(X) + \sum_Y \beta(X,Y)}.
\]

(5b)

The number of \( Y \)'s, then, can be regarded as a sum, over all \( X \)-altering events, of a quantity which is the number of progeny if the progenitor created at the event is a \( Y \), and zero otherwise. Therefore, if we could assert that the number of progeny stemming from the various creating events were independent of each other, then several applications of the composition rule would yield

\[
\Phi(s) = \Theta_{X,C,M}(s_X, \frac{\delta(X) + \sum_Y \beta(X,Y) \Psi_Y(s_Y)}{\delta(X) + \sum_Y \beta(X,Y)}, \frac{\sum_Y \alpha(X,Y) \Psi_Y(s_Y)}{\sum_Y \alpha(X,Y)}).
\]

(6)

(The dependence of \( \Phi \) and \( \Psi \) on \( X \) and \( T \) is suppressed.) In fact, \( g(t) \) is proportional to the number of \( X \)-cells at time \( t \), and therefore the event time is not exactly independent of \( N_X(T) \). Furthermore, the times of two creation events are correlated, because they both are related to the history of \( X \)-cells. We avoid these complications by assuming that the probability density \( g \) of the creation-event time is proportional to the mean number of \( X \)-cells, rather
than to the actually occurring number, and that two event times are independent. This allows us to retain Equation (6). It introduces a source of error that appears to be miniscule in practice.

This approach represents an improvement over the filtered-Poisson-process model which is commonly used [1, 12], in which the event times are independent, not just of each other, but also of the number of X-cells. This would imply that the number of mutants is independent of the number of X-cells at the end of the interval. The method presented here corrects this, but slightly overcompensates for the correlation between mutants and X-cells. We favor the overestimate. In reality, the hospitality of the environment varies between tumors, and with time and location within a tumor, inducing correlations in increments of source cells (X) and altered cells (Y). Therefore, in overestimating the correlation, we err in the direction of biological plausibility.

**DERIVATION OF Θ_{X.C,M}**

The computation of Θ_{X.C,M} is accomplished using a technique described by Bailey [26]. In Bailey’s notation, for a vector of integer increments, ε = (ε_X, ε_C, ε_M) in the vector N = (N_X, N_C, N_M), one defines

\[
f_ε = \frac{1}{dt} \text{prob}[N(t + dt) - N(t) = ε].
\]

Let

\[
ξ = \text{birth rate for X-cells},
\]

\[
γ = \text{total culling rate } = δ(X) + \sum_Y β(X, Y),
\]

\[
μ = \text{total mutation rate } = \sum_Y α(X, Y).
\]

The only nonzero f’s are

\[
\frac{1}{dt} \text{prob[an X-cell divides in } (t, t + dt)] = f_{+1,0,0} = ξ N_X(t),
\]

\[
\frac{1}{dt} \text{prob[an X-cell is culled in } (t, t + dt)] = f_{-1,+1,0} = γ N_X(t),
\]

and

\[
\frac{1}{dt} \text{prob[an X-cell mutates in } (t, t + dt)] = f_{0,0,+1} = μ N_X(t).
\]
The corresponding equation for \( \Theta = \Theta_{X,C,M}(s_X, s_C, s_M) \) can be shown to be

\[
\frac{\partial \Theta}{\partial t} = \left( (s_X - 1) \xi + \left( \frac{s_C}{s_X} - 1 \right) \gamma + (s_M - 1) \mu \right) s_X \frac{\partial \Theta}{\partial s_X} = \xi Q \frac{\partial \Theta}{\partial s_X}
\]  

(7)

where \( Q = s_X^2 - \nu s_X + (\gamma / \xi)s_C, \nu = 1 + B \xi^{-1}, B = \gamma + (1 - s_M)\mu. \) \( \Theta \) is constant along contours defined by the auxiliary equation \( ds_X/dt = \xi Q \). Integrating, one finds that a constant of the integration is

\[
\eta = \eta(t) = e^{D \xi t} \frac{r_1 - s_X}{r_2 - s_X},
\]

where \( r_1 \geq r_2 \) are the solutions to \( Q(s_X) = 0, \) and \( D = [\nu^2 - 4(\gamma / \xi)s_C]^{1/2} = r_1 - r_2. \) Fitting the boundary condition \( \Theta = s_X \) at \( t = 0, \) we get

\[
\Theta = \frac{r_1 - r_2 \eta}{1 - \eta}.
\]

(8)

This formula fails for certain special cases. The full solution is given in Appendix A.

**DERIVATION OF \( \Psi_Y \)**

It remains to derive \( \Psi_Y = \Psi_Y(s_Y; X, T), \) the p.g.f. for the number of cells of type \( Y \) which are the progeny of a single creating event (conversion or mutation) creating a \( Y \)-progenitor.

Recall that we are assuming the time of the event, \( t, \) to be distributed proportionally to the mean number of \( X \)-cells, rather than the actual number in any single realization. Ignoring mutations or conversions to \( X \)-cells, the mean of \( N_X \) grows exponentially; letting the net growth rate \( \xi(X) - \gamma(X) \) be denoted by \( \lambda(X), \) we have

\[
g(t) = \frac{\lambda(X) e^{\lambda(X) t}}{e^{\lambda(X) T} - 1},
\]

(9a)

but in the special case \( \lambda(X) = 0 \) we have

\[
g(t) = T^{-1}.
\]

(9b)

The expressions for \( g \) and for \( \Theta \) can be used in Equation (4). Let \( \rho_Y = \gamma(Y) / \xi(Y). \) In the general case covered by (8) and (9a), we get

\[
\Psi_Y = 1 - (1 - \rho_Y)(z(T) - 1)^{-1} \xi I,
\]

(10)
\[ \zeta = \eta(T), \quad z(t) = e^{\lambda(X)t}, \quad \text{and} \quad I = \int_1^{z(T)} \frac{1}{\zeta - z^{\lambda(Y)/\lambda(X)}} \, dz. \]

The integral \( I \) can be evaluated exactly for \( \lambda(X) = \lambda(Y) \), and numerically otherwise.

There are many special cases in which (10) fails; solutions are provided in Appendix B.

4. EXTENSIONS

VARIATIONS ON CONVERSION

An examination of the derivation of \( \Theta \) shows that the crucial element is that the polynomial \( Q \) is quadratic. To guarantee this, it is only necessary that the number of \( X \)-cells change by \(+1\), \(0\), or \(-1\), as in renewal, mutation, or culling, respectively. The culling event can be generalized in a useful way to include events in which an \( X \)-cell is replaced by a \( Y \)- and a \( Z \)-cell. The only change is in \( s_C \), the second argument to \( \Theta \); in the numerator, the factor \( \Psi_Y(s_Y) \) is replaced by the product \( \Psi_Y(s_Y)\Psi_Z(s_Z) \), since the progeny of the \( Y \)- and the \( Z \)-cell are independent of each other.

One use of such an event is as a model for point mutation. Suppose a diploid cell homozygous for some character of interest undergoes an intermitotic mutation, which in our model would be classified as a "conversion" event. This cell is now heterozygous \((Aa \text{ in the classical notation})\), and upon mitosis becomes a pair of homozygous cells, \((AA \text{ and } aa)\).

Another application arises in which \( Y \) and \( Z \) are the same cell type. In modeling differentiation, one might allow exactly one cell division per differentiation level. This may be appropriate, for example, for some aspects of hemopoiesis. Then each division produces two cells of the next cell type in the differentiation pathway.

CONDITIONING ON SUBPOPULATION SIZES

There are situations where one may need to condition on population sizes. There may be a small number of cells at time \(0\), and a substantial probability that the tumor never reaches macroscopic size due to the death component in the kinetics. Then the results should be conditional on the establishment of a diagnosable tumor. Furthermore, there may be data on subpopulation sizes, and results may change if this information is taken into account. The conditioning may be required in the middle of the time sequence as well as at the end.

We present a method which requires only the p.g.f. To simplify the exposition, suppose there is only one cell type, and the number of cells \( N \) is
required to be less than $10^k$ at time $t$. To condition on this event is too
difficult computationally, so we will condition instead on an event $A$ which is
nearly the same, but much easier to work with. This event is defined by the
outcome of a thought experiment at time $t$. The cells are imagined to be
subjected (temporarily) to a “dose” of $k$ logs, with binomial kill, as in our
model for treatment. Let $A$ be the event that all the cells are thereby
“killed.” We will condition on $A$, which is an approximation of the event
that $N < 10^k$, in the sense that, when one occurs but not the other, it is either
a rare occurrence ($N$ far from $10^k$), or else of no practical importance ($N$
near $10^k$). When $k$ equals zero, $A$ is exactly the event $N = 0$.

Let $\phi_0$ be the p.g.f. for the process prior to $t$. Then the probability of $A$ is
given by $\phi_0(s_A)$, where $s_A = 1 - 10^{-k}$, and the conditional p.g.f. at $t$ is

$$\phi_0(s | A) = \sum_n \text{prob}(A | N = n) \frac{\text{prob}(N = n)}{\text{prob}(A)} s^n = \frac{\phi_0(s_A s)}{\phi_0(s_A)}. \quad (11a)$$

One can condition instead on the event $B$, that not all cells are “killed” by a
test probability of $s_B$, as an approximation to $N > 1/(1 - s_B)$. The resulting
p.g.f. is

$$\phi_0(s | B) = \sum_n \text{prob}(B | N = n) \frac{\text{prob}(N = n)}{\text{prob}(B)} s^n = \frac{\phi_0(s) - \phi_0(s_B s)}{1 - \phi_0(s_B)}. \quad (11b)$$

Conditioning on a range, $N \in [m_1, m_2]$, is accomplished through two such
thought experiments in tandem, with $s_A = 1 - m_2^{-1}$, $s_B = (1 - m_1^{-1})/(1 -
m_2^{-1})$. Substitute in the right side of Equation (11a) for $\phi_0$ in Equation (11b)
to get

$$\phi_0(s | A \text{ and } B) = \frac{\phi_0(s_A s) - \phi_0(s_A s_B s)}{\phi_0(s_A) - \phi_0(s_A s_B)}. \quad (11c)$$

In the examples given above, the time of the conditioning is generally
earlier in the history of the tumor than the time of evaluation. In that case, if
$\phi_1$ is the p.g.f. for the process after $t$, Equations (11a), (11b), and (11c) are
altered by replacing $s$ with $\phi_1(s)$.

These results generalize directly to the case of multiple cell types in the
obvious way.

The approximation described here may be inadequate when $k$ is small but
not zero.
FEEDBACK AMONG SUBPOPULATIONS

The model we have proposed requires that the kinetic parameters be constants, i.e., although they may vary with time (as step functions), they may not depend on the sizes of the subpopulations, which are random. This limitation makes certain phenomena impossible to model. However, the approach outlined above for dealing with conditioning also works for this purpose.

To simplify the discussion, we suppose only one cell type, because the crucial aspect is the dependence of the kinetics parameters on the state of the process; the generalization will be obvious. As an example, suppose we wish to model the situation in which the tumor’s growth rate switches to a lower value when it reaches $10^4$ cells. Since we are working with p.g.f.’s, this is replaced by an easier problem. Suppose that, at a given time, the value which the growth rate assumes depends on the result of an “experiment” to test the population size, in which each cell is temporarily culled with probability $0.9999 = 1 - 10^{-4}$. If the entire population is thereby culled, the tumor is declared “smaller than $10^4$,” and the growth assumes the higher value. Otherwise it assumes the lower value.

The objection may be raised that this process is lacking in realism. In addition, kinetics parameters almost certainly vary smoothly with subpopulation sizes. Note, however, that this artificial process actually results in smoothing out the dependence on population size, so it probably mimics reality better than the deterministic switch model first introduced. Furthermore, the computations, being based on the p.g.f., are efficient. It also conforms to our philosophy of providing a concrete, well-defined stochastic process as the goal of the computing effort.

Now let $t$ be the time at which the switch may occur, let $A$ be the event that all cells are “killed” in our test experiment at time $t$, let $B$ be the complementary event, and let $s_A$ be the probability that any particular cell is “killed” in the test experiment. Let $\phi_0$ be the p.g.f. for the process prior to $t$; let $\phi_A$ be the p.g.f. for the process after $t$ if $A$ occurs at $t$ and $\phi_B$ if $B$ occurs at $t$. Then the p.g.f. of the entire process is

$$\phi(s) = \text{prob}(A) \phi_0(\phi_A(s)|A) + \text{prob}(B) \phi_0(\phi_B(s)|B)$$

$$= \phi_0(s_A\phi_A(s)) + \phi_0(\phi_B(s)) - \phi_0(s_A\phi_B(s)).$$

(12)

To allow more than two choices for the parameters, let our test experiment now be multinomial, with probabilities $s_1, \ldots, s_f$ of landing in each of the sets $a_1, \ldots, a_f$. Say the event $A_i$ occurs if there are no cells in $a_{i+1}, \ldots, a_f$ and $a_i$ is not empty. This event is an approximation for the event $N \in [1/\Sigma_{j \geq i} s_j, 1/\Sigma_{j > i} s_j]$. Letting $\phi_i$ be the p.g.f. after $t$ conditional on $A_i$, the
p.g.f. for the entire process is easily shown to be

\[ \phi(s) = \sum_{i=0}^{I} \text{prob}(A_i) \phi_0(\phi_i(s)|A_i) \]

\[ = \phi_0(s_1 \phi_1(s)) + \sum_{i=2}^{I} \left\{ \phi_0(\phi_i(s) \sum_{j \leq i} s_j) - \phi_0(\phi_i(s) \sum_{j < i} s_j) \right\} \]  

(13)

The computing time rises exponentially with the number of switching time points; therefore the method is impractical unless this number is kept quite small. Within this limitation, the effects of feedback may be assessed. The effects of feedback will likely be underestimated.

5. CALCULATING MODEL PREDICTIONS WITH THE PROBABILITY GENERATING FUNCTION

There are two standard manipulations of p.g.f.'s. First, the calculation of moments is a routine if sometimes tedious procedure traditionally performed in branching-process models. But the distribution of the number of cells is typically highly skewed. Also, in the case of cell populations with repeated "treatment" epochs, there may be a significant probability of zero cells. These facts conspire to make moments potentially deceptive representations of the distribution. Much less common is to transform the p.g.f. to get the distribution itself. This is generally difficult and computationally intensive.

We will argue for more direct ways of using the p.g.f. in practice. The uses of the p.g.f. will be discussed in terms of four questions which an oncologist would want a tumor model to answer:

1. How well does a particular treatment regimen perform?
2. What would be the result of adding or augmenting a treatment at a specific time?
3. How big are particular subpopulations at various times?
4. When is the tumor at its most vulnerable; in other words, when would a change in treatment schedule be most likely to improve prognosis?

The first question requires a criterion for judging performance. A natural one is the probability that all cells have been eradicated by the end of the treatment regimen—a probability of "cure." This is easily computed to be

\[ \text{prob}(N_X(T) = 0 \text{ for each cell type } X) = \Lambda_T(0, \ldots, 0). \]  

(14)

In some applications, some subset of all cells is more important than the total count. For example, there may be a cell-counting technique available which does not distinguish between living and dead cells. In that case, the model would include a cell type explicitly representing the dead cells which are
counted, and "death" in the model would represent the lysing of dead cells. Then in (14) we would want to exempt dead or doomed cells. This is accomplished for a cell type $Y$ by setting $s_Y = 1$.

Answering the second question is equally straightforward. As suggested by Equation (2), the consequences of adding a dose can be assessed by evaluating $\Lambda$ at the value $s = K$, with $K_X$ equal to the probability of killing cells of type $X$ associated with the extra treatment being contemplated. A physician may ask, "If a two-log dose were added right now, what would be the probability that every cell would be killed?" or "How much of a dose (in logs of kill) would be required to get a 50% chance of killing all cells?" Each answer to such a question provides a point on a "dose-response" curve, describing the tumor operationally at any time. The graph of the p.g.f. is simply converted into a dose-response curve; one regards the argument axis as a dose axis, perhaps transforming the kill probability scale ($s$) to a "logs of kill" scale ($-\log s$), and replaces the p.g.f. axis by $1 - p.g.f.$, the probability of surviving the test dose.

In fact, such a curve provides information on the distribution of the number of cells, suggesting an approach to question 3. Suppose cell were not binomial, but instead tumors were eradicated if and only if they had fewer "logs of cell" than the number of logs of kill. Then the dose-response curve, with "logs of cell" replaced by "logs of kill," would be precisely the cumulative distribution of the number of cells. Let $\phi(s) = \Lambda(s)$, with $s_X$ equaling $s$ for cell types included in the subpopulation of interest, and 1 for those excluded. This gives us the approximation

$$\text{prob}\left( \sum_{X \in S} N_X < n \right) \approx \phi(1 - n^{-1}).$$

The right-hand side is equal to a convolution of the true distribution with the smoothing kernel $K(N, n) = N n^{-2} (1 - n^{-1})^N n^{-1}$. For practical purposes this will be quite a reasonable approximation, unless the number of cells is small.

Finally, we consider question 4. A common way of thinking among oncologists is to model tumor growth in a simple nonstochastic way, and examine a plot of cell number over time for the "nadir," the time at which the tumor is smallest and presumably most vulnerable. To provide a comparable tool, we develop a summary, $\hat{n}$, for the distribution of tumor size, which is neither the mean, the mode, nor the median, but which has a clear clinical meaning and is easy to compute. The fundamental idea is to use a quantile of the approximate cumulative distribution discussed above. We let

$$\hat{n} = \frac{1}{1 - \phi^{-1}(P)}.$$
for some suitable $P$. A natural choice is $P = \frac{1}{2}$, in analogy with the median, but as we shall see, $P = 1/e = 0.358$ has a special merit.

For a tumor of nonrandom size $n$, we would generally not have $\hat{n} = n$, surely an undesirable characteristic for a summary statistic. However, forcing $\hat{n} = n$ leads to $n = 1/(1 - P^{1/n})$, so that $P = (1 - 1/n)^n \approx e^{-1}$. This suggests the choice $P = 1/e = 0.368$.

Unfortunately, $\hat{n}$ is undefined when the probability of eradication is greater than $P$ before the probe. We can extend $\hat{n}$ to cover this case in a somewhat artificial way. Note that a Poisson variable with mean equal to 1 has probability $1/e$ of being zero. Let $Q = \text{prob(\text{no cells before the probe})}$. If $Q > P$, we can choose our summary statistic to be the mean of the Poisson variable $X$ for which $\text{prob}(X = 0) = Q$. That is,

$$\hat{n} = -\log Q.$$ 

While this is a smooth extrapolation of $\hat{n}$, it has the disadvantage that, when $Q > P$, it contains no information about the distribution except $Q$ itself; therefore subsequent growth of those tumors which have not been eradicated is not reflected in $\hat{n}$.

The estimate $\hat{n}$ will be deceptive when the distribution is dramatically bimodal, which occurs when $Q$ is substantial. It is important in these cases to compute $\hat{n}$ conditioning on the number of cells being nonzero.

6. DISCUSSION

LIMITATIONS DUE TO COMPUTATIONAL PROBLEMS

The approximation due to the “assumption” of one altering event per interval can cause computational problems. For applications where altering events are rare (for example, where “mutations” represent real mutations), the error is easily vanquished by interpolating points of iteration, i.e. by forcing the intervals to be short enough, so that a line of descent with two conversion or mutation events within a single interval will be rare. When there are very high conversion or mutation rates, or a very long time horizon with a long sequence of successive cell-types, the numerical accuracy will be degraded by the large number of iterations required in Equation (3). (Models for stem-cell differentiation are examples.) However, these are precisely the conditions under which the asymptotics of branching-process theory perform well. In this sense, the two approaches complement each other. It would be desirable to have an approach which works with hybrid models which include both rare and common events.

The replacement of the true $g(t)$ by the mean number of source cells, in the computation of $\Psi_Y(s; g; X, T)$, is almost always inconsequential. The replacement will be dangerous if the primary change in the number of source
cells during the interval is due to new source cells which result from conversion or mutation events. These cells are not counted in computing \( g(t) \). For example, if cell type \( Y \) has a zero growth rate, yet increases rapidly because of a high rate of conversion from cell type \( Z \), then \( g(t) \) will be constant throughout the interval, whereas it should be weighted more heavily towards the end of the interval. Thus if \( Y \)-cells make a conversion to \( Z \)-cells, which are rapidly growing, then the number of \( Z \)-cells at the end of the interval will be overestimated by the model. Again, this happens only in models with high rates of altering events.

An analysis of error estimates, with a discussion of the implications for the range of applicability, is available from the author [27].

**LIMITATIONS DUE TO ASSUMPTIONS**

Several assumptions were required to make the computation of the p.g.f. tractable.

One is that the time to each event is exponentially distributed. Replacement by more realistic distributions would affect the behavior of the model only in certain situations: e.g., if we asked about the state of affairs after only a few divisions, and if the cells were synchronized at the outset. However, cell synchronization may be an important factor; phase-specific drugs may partially synchronize cells, resulting in apparent interactions with other drugs.

We have assumed throughout that kinetic parameters are constant over time, but this is not a true limitation in most cases. Clearly the derivation requires only that they are constant within an interval over which \( \Phi(s; X, T) \) is computed. If iteration intervals are made short enough, several additional phenomena can be investigated, including nonexponential (e.g. Gompertzian) growth, and changes in mutation rate due to introduction of mutagenic agents.

The assumption that cells have independent futures also limits applications, by precluding biological interactions between tumor subpopulations. For example, antagonisms may include competition for nutrients, secretion of autotoxins, or antigenic cross-reactivity, while synergisms may include mutually beneficial angiogenesis, secretion of autocrine growth factors, or secretion of immunosuppressive factors. In principle, the feedback methods presented here can handle these situations, but may be too crude or too computationally intensive to be practical.

Certain aspects of tumor growth which appear in other models are absent from the one presented here. In each case, an approximate model can be constructed within our framework which should be satisfactory. One example is the phenomenon of “phenotypic delay” [9, 12], in which the phenotype (e.g. increased resistance to a treatment) appears not immediately at the moment of mutation, but significantly later. To deal with this situation, we
can define a cell type corresponding to the sensitive intermediate, and postulate a conversion event to model the appearance of the new phenotype. Another example is the modeling of the physical organization of the tumor, which has been approached using Monte Carlo simulation [22]. Cell-type definitions can include their physical location in the body, if the compartments are not required to be too finely drawn. Distinguishing between, say, interior tumor cells, cells on the tumor periphery, metastasizing cells, and established metastatic cells may be very important. The success of these approaches in extending the range of our model remains to be seen.

The strength of the approach presented here is the freedom to explore a wide variety of possibilities in the context of a single class of models, in a time-efficient way.

7. CONCLUSION

A program implementing the formulas presented here, tailored to applications in cancer treatment policy, has been tested and used extensively to evaluate current concepts in cancer treatment policy [7].

APPENDIX A. FULL SOLUTION FOR EQUATION (7)

If \( t = 0 \), if \( \xi = B = 0 \), or if \( s_X = r_2 \), then

\[ \Theta = s_X. \] (A1)

If \( \xi = 0 \), then

\[ \Theta = \gamma \frac{s_C}{B} (1 - e^{-Bt}) + s_X e^{-Bt}. \] (A2)

If \( \eta = 1 \) [which occurs when \( D = 0 \); that is, when \( s_C = 1 \), \( \mu(1 - s_M) = 0 \), and \( \gamma = \xi \)], then

\[ \Theta = 1 - \frac{1 - s_X}{\xi \xi (1 - s_X) + 1}. \] (A3)

Otherwise, \( \Theta \) is given by Equation (8).

APPENDIX B. FULL SOLUTION FOR EQUATION (4)

There are four cases for \( \Theta \) (Appendix A); each leads to a solution to (4), which needs modification when \( \lambda(X) \) takes on special values.

If \( t = 0 \), \( \lambda(Y) = 0 \), or \( s_Y = \rho_\gamma \), then (A1) leads to

\[ \Psi = s_Y. \] (B1)

This will also hold if \( \lambda(X) = \lambda(Y) = \xi(Y) = 0 \).
If $\xi(Y) = 0$, then (A2) leads to

$$\Psi = 1 - (1 - s_Y)(z(T) - 1)^{-1} \frac{\lambda(X)}{\lambda(X) - \lambda(Y)} \left( e^{\lambda(X)T} - e^{\lambda(Y)T} \right) \tag{B2}$$

in the general case,

$$\Psi = 1 - (1 - s_Y) \gamma(Y)^{-1} T^{-1} \left( 1 - e^{-\gamma(Y)T} \right) \tag{B3}$$

when $\lambda(X) = 0$, and

$$\Psi = 1 - (1 - s_Y) z(T)(z(T) - 1)^{-1} \lambda(X) T \tag{B4}$$

when $\lambda(X) = \lambda(Y)$.

If $\eta = 1$ [so that $\lambda(Y) = 0$], then (A3) leads to

$$\Psi = 1 - (z(T) - 1)^{-1} (1 - s_Y) \kappa e^{-\kappa} \int_{\kappa}^{\kappa + \lambda(X)T} u^{-1} e^u \, du, \tag{B5}$$

[with $\kappa = \lambda(X) \xi(Y)^{-1} (1 - s_Y)^{-1}$] in the general case, and

$$\Psi = 1 - \xi(Y)^{-1} T^{-1} \ln \left[ 1 + (1 - s_Y) \xi(Y) T \right] \tag{B6}$$

when $\lambda(X) = 0$.

Finally, if the general solution for $\Theta$ holds [Equation (8)], $\Psi$ is given by Equation (10) in the general case, and by

$$\Psi = 1 - (\rho_Y - 1) T^{-1} \lambda(Y)^{-1} \ln \left| \frac{1 - \eta(0)}{1 - \eta(T)} \right| \tag{B7}$$

when $\lambda(X) = 0$.

In the case $\lambda(X) = \lambda(Y)$, Equation (10) can be written more simply:

$$\Psi = 1 - (\rho_Y - 1) \eta(T) \left[ z(T) - 1 \right]^{-1} \ln \left| \frac{z(T) - \eta(T)}{1 - \eta(T)} \right|. \tag{B8}$$

REFERENCES

