Research Article

New Cancer Stochastic Models Involving Both Hereditary and Nonhereditary Cancer Cases: A New Approach

Wai-Yuan Tan¹ and Hong Zhou²

¹ Department of Mathematical Sciences, The University of Memphis, Memphis, TN 38152, USA

² Department of Mathematics and Statistics, Arkansas State University, State University, AR 72467, USA

Correspondence should be addressed to Wai-Yuan Tan; waitan@memphis.edu

Received 24 August 2012; Accepted 10 October 2012

Academic Editors: T. LaFramboise, K. M. Page, I. Rogozin, and J. M. Starobin

Copyright © 2013 W.-Y. Tan and H. Zhou. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

To incorporate biologically observed epidemics into multistage models of carcinogenesis, in this paper we have developed new stochastic models for human cancers. We have further incorporated genetic segregation of cancer genes into these models to derive generalized mixture models for cancer incidence. Based on these models we have developed a generalized Bayesian approach to estimate the parameters and to predict cancer incidence via Gibbs sampling procedures. We have applied these models to fit and analyze the SEER data of human eye cancers from NCI/NIH. Our results indicate that the models not only provide a logical avenue to incorporate biological information but also fit the data much better than other models. These models would not only provide more insights into human cancers but also would provide useful guidance for its prevention and control and for prediction of future cancer cases.

1. Introduction

It is universally recognized that each cancer tumor develops through stochastic proliferation and differentiation from a single stem cell which has sustained a series of irreversible genetic and/or epigenetic changes (Little [1]; Tan [2, 3]; Tan et al. [4, 5]; Weinberg [6]; Zheng [7]). That is, carcinogenesis is a stochastic multistage model with intermediate cells subjecting to stochastic proliferation and differentiation. Furthermore, the number of stages and the number of pathways of the carcinogenesis process are significantly influenced by environmental factors underlying the individuals (Tan et al. [4, 5]; Weinberg [6]).

Another important observation in human carcinogenesis is that most human cancers cluster around family members. Further, many cancer incidence data (such as SEER data of NCI/NIH, USA) have documented that some cancers develop during pregnancy before birth to give new born babies with cancer at birth. This has been referred to as pediatric cancers. Well-known examples of pediatric cancers include retinoblastoma—a pediatric eye cancer, hepatoblastoma—a pediatric liver cancer, Wilm's tumor—a pediatric kidney

cancer, and medulloblastoma—a pediatric brain tumor. Epidemiological and clinical studies on oncology have also revealed that inherited cancers are very common in many adult human cancers including lung cancer, colon cancer [8], uveal melanomas (adult eye cancer, [9]), and adult liver cancer (HCC, [10]).

Given the above results from cancer biology and human cancer epidemiology, the objective of this paper is to illustrate how to develop stochastic models of carcinogenesis incorporating these biological and epidemiological observations. Based on these models and cancer incidence data, we will then proceed to develop efficient statistical procedures to estimate unknown parameters in the model, to validate the model, and to predict cancer incidence.

In Section 2, we illustrate how to incorporate segregation of cancer genes in multistage stochastic models of carcinogenesis to account for inherited cancer cases. In Section 3, we will develop stochastic equations for the state variables of the model described in Section 2. By using these stochastic equations we will derive probability distributions of the state variables (i.e., the number of intermediate cancer cells) and the probability distribution of time to detectable cancer

FIGURE 1: Histopathology lesions and genetic pathway of squamous cell carcinoma of NonSmall Cell Lung Cancer (NSCLC).

FIGURE 2: A Multistage Model of uveal melanoma (adult human eye cancer).

tumors. In Section 4, assuming that we have some cancer incidence data such as the SEER data from NCI/NIH, we proceed to develop statistical models for these data from these multistage models of carcinogenesis. In Section 5, by combining models in Sections 2–4, we proceed to develop a generalized Bayesian inference and Gibbs sampling procedures to estimate the unknown parameters, to validate the model, and to predict cancer incidence. As an example of application, in Section 6 we proceed to develop a multistage model of human eye cancer with inherited cancer cases as described in Figure 2. We will illustrate the model and methods by analyzing the SEER data of human eye cancer from NCI/NIH. Finally in Section 7, we will discuss the usefulness of the model and the methods developed in this paper and point out some future research directions.

2. The Stochastic Multistage Model of Carcinogenesis with Inherited Cancer Cases

The k -stage multistage model of carcinogenesis views carcinogenesis as the end point of k ($k \geq 2$) discrete, heritable, and irreversible events (mutations, genetic changes or epigenetic changes) with intermediate cells subjected to stochastic proliferation and differentiation (Little [1]; Tan [2, 3]; Tan et al. [4, 5]; Weinberg [6]). Let $N = J_0$ denote normal stem

cells, T the cancer tumors, and J_i the i th stage initiated cells arising from the $(i - 1)$ th stage initiated cells ($i = 1, \dots, k$) by some genetic and/or epigenetic changes. Then the model assumes $N \rightarrow J_1 \rightarrow J_2 \rightarrow \dots \rightarrow J_k \rightarrow \text{Tumor}$ with the J_i cells subject to stochastic proliferation (birth) and differentiation (death). Further, it assumes that each stem cell proceeds independently of other cells and that cancer tumors develop from primary J_k cells by clonal expansion (stochastic birth and death), where primary J_k cells are J_k cells which arise directly from J_{k-1} cells; see Yang and Chen [11].

For example, Figure 1 is a multistage pathway for the squamous NSCLC (NonSmall Cell Lung cancer) as proposed by Osada and Takahashi [12] and Wistuba et al. [13]. Similarly, Figure 2 is the multistage model for uveal melanoma proposed by Landreville et al. [14] and Mensink et al. [15] while Figure 3 is the APC- β -Catenin-Tcf pathway for human colon cancer (Tan et al. [8], Tan and Yan [16]).

Remark 1. To develop stochastic multistage models of carcinogenesis, in the literature (Little [1], Tan [2], Zheng [7]) it is conveniently assumed that the J_k cells grow instantaneously into cancer tumors as soon as they are generated. In this case, the number of tumors is equal to the number of J_k cells and one may identify J_k cells as tumors. It follows that the number of tumors is a Markov process and that the J_k cells are

FIGURE 3: The APC- β -catenin-Tcf-myc pathway for human colon cancer.

transient cells. In these cases, one needs only to deal with $T(t)$ and J_j cells with $j = 1, \dots, k-1$. However, as shown by Yang and Chen [11], the number of tumors is much smaller than the total number of J_k cells. Also, in many animal models and in cancer risk assessment of radiation, Klebanov et al. [17], Yakovlev and Tsodikov [18], and Fakir et al. [19] have shown that $T(t)$ are in general not Markov.

To extend the above model to include hereditary cancers, observe that mutants of cancer genes exist in the population and that both germline cells (egg and sperm) and somatic cells may carry mutant alleles of cancer genes [2, 20]. Further, without exception, every human being develops from the embryo in his/her mother's womb (embryo stage, denote time by 0), where stem cells of different organs divide and differentiate to develop different organs respectively (see Weinberg [6], Chapter 10). If both the egg and the sperm generating the embryo carry mutant alleles of relevant cancer genes, then the individual is an J_2 -stage person at the embryo stage; if only one of the germ line cells (egg or sperm) generating the embryo carries mutant alleles of cancer genes, then at the embryo stage the individual is an J_1 -stage person. Similarly, the individual is a normal person ($N = J_0$ person) at the embryo stage if both the egg and the sperm generating the embryo do not carry mutant alleles of cancer genes. Refer to the person in the population as an J_i ($i = 0, 1, 2$) person if he/she is an J_i -stage person at the embryo stage. Then with respect to the cancer development in question, people in the population can be classified into 3 types of people: normal people ($N = J_0$ people), J_1 people, and J_2 people. Based on this classification, for normal people in the population the stochastic model of carcinogenesis is a k -stage multievent model given by $J_0 \rightarrow J_1 \rightarrow \dots \rightarrow J_k \rightarrow \text{Tumor}$; for J_1 people in the population the stochastic model of carcinogenesis is a $(k-1)$ -stage multievent model given by $J_1 \rightarrow J_2 \rightarrow \dots \rightarrow J_k \rightarrow \text{Tumor}$ and for J_2 people in the population, the stochastic model of carcinogenesis is a $(k-2)$ -stage multievent model given by $J_2 \rightarrow \dots \rightarrow J_k \rightarrow \text{Tumor}$.

To account for inherited cancer cases, let p_1 be the proportion of J_1 people in the population and p_2 the proportion of J_2 people in the population. In general large human populations under steady-state conditions, one may practically assume that the p_i is a constant independent of time (Crow and Kimura [21]). Then $p_0 = 1 - p_1 - p_2$ ($0 < p_1 + p_2 < 1$) is the proportion of normal people (i.e., $N = J_0$ people) in the population. Let n be the population size and n_i ($i = 0, 1, 2$) the number of J_i people in the population so that $\sum_{i=0}^2 n_i = n$. Assume that n is very large and that marriage between people in the population is random with respect to cancer genes; then as shown in Crow and Kimura [21] (see also Tan [22], Chapter 2), the conditional probability distribution of (n_1, n_2) given n is 2-dimensional multinomial with parameters (n, p_1, p_2) . That is,

$$(n_1, n_2) \mid n \sim \text{Multinomial}(n; p_1, p_2). \quad (1)$$

To derive probability distribution of time to cancer under the above model, observe that during pregnancy the proliferation rates of all stem cells are quite high. Thus, with positive probability J_2 people in the population may acquire additional genetic and/or epigenetic changes during pregnancy to become J_3 -stage people at birth. Similarly, J_1 people may acquire genetic and/or epigenetic changes during pregnancy to become J_2 people at birth; albeit the probability is very small, normal people at the embryo stage may acquire some genetic and/or epigenetic changes during pregnancy to become J_1 people at birth. Because the probability of genetic and epigenetic changes is small, one may practically assume that an J_i ($i = 0, 1, 2$) person at the embryo stage would only give rise to J_i stem cells and possibly J_{i+1} stem cells at birth. This is equivalent to assuming that J_i people at the embryos stage would not generate J_{i+j} ($j > 1$) stem cells at or before birth. This model is represented schematically in Figure 4. Notice that if $k = 2$, one may practically assume that with probability one an J_2 person at the embryo stage would develop cancer at or before birth (t_0). If $k = 3$, then

FIGURE 4: Embryo genotypes and their frequencies at embryo stage and at birth.

with probability α ($\alpha > 0$), an J_2 person at the embryo stage would develop cancer at or before birth.

3. The Stochastic Process of Carcinogenesis with Hereditary Cancer Cases and Mathematical Analysis

Because tumors are developed from primary J_k cells, for the above stochastic model, the identifiable response variables are $T(t)$ and $\{J_u(t; i), i = 0, 1, 2, u = i, i + 1, \dots, k - 1\}$, where $T(t)$ is the number of cancer tumors at time t and $J_u(t; i)$ is the number of J_u ($u = i, i + 1, \dots, k - 1$) cells at time t in people who are J_i people at the embryo stage (see [3, 5, 8, 23], Remarks 1 and 2). For people who have genotype J_i ($i = 0, 1, 2$) at the embryo stage, the stochastic model of carcinogenesis is then given by the stochastic process $\{X_i(t), T(t), t > 0\}$, where $X_i(t) = \{J_u(t; i), u = i, i + 1, \dots, k - 1\}'$. For these processes, in the next subsections, we will derive stochastic equations for the state variables ($J_u(t; i), i = 0, 1, 2, u = i, \dots, k - 1$); we will also derive the probability distributions of these state variables and the probabilities of developing cancer tumors. These are the basic approaches for modeling carcinogenesis used by the first author and his associates; see Tan [3], Tan et al. [4, 5, 8, 23], Tan and Zhou [9], Tan and Yan [16], and Tan and Chen [24, 25] and Remark 3.

Remark 2. At any time (say t) the total number of J_k cells is equal to the total number of J_k cells generated from J_{k-1} cells at time t plus the total number of J_k cells generated by cell division from other J_k cells at time t ; the former J_k cells are referred to as primary J_k cells while the latter are not primary J_k cells. Since each tumor is developed from a single primary J_k cell through stochastic birth and death process, each primary J_k cell will generate at most one tumor. It follows that at any time the total number of J_k cells is considerably greater than the number of cancer tumors (see also Yang and Chen [11]). Thus, for generating cancer tumors the only identifiable state variables are the number of J_j cells with ($j = 0, 1, \dots, k - 1$) and the number of detectable cancer tumor.

Remark 3. To model stochastic multistage models of carcinogenesis, the standard traditional approach is to assume that the last stage cells (i.e., the J_k cells in the model $N \rightarrow J_1 \cdots \rightarrow J_k \rightarrow \text{Tumor}$) grow instantaneously into a cancer tumor as soon as they are generated and then apply the standard Markov theory to $T(t)$ and to the state variables $X(t) = \{J_i(t), i = 0, 1, \dots, k - 1\}$. This approach has been described in detail in Tan [2], Little [1], and Zheng [7]; see also Luebeck and Moolgavkar [26] and Durrett et al. [27]. However, in some cases the assumption of instantaneous growth into cancer tumors of J_k cells may not be realistic

(Klebanov et al. [17], Yakovlev and Tsodikov [18], and Fakir et al. [19]); in these cases, $T(t)$ is not Markov so that the Markov theory method is not applicable to $T(t)$. To develop analytical results and to resolve many difficult issues, Tan and his associates [4, 5, 24] have proposed an alternative approach through stochastic equations and have followed Yang and Chen [11] to assume that cancer tumors develop by clonal expansion from primary last stage cells. Through probability generating function method, Tan and Chen [24] have shown that if the Markov theory is applicable to $T(t)$, then the stochastic equation method is equivalent to the classical Markov theory method but is more powerful. Also, through stochastic equation method we have shown in the Appendix that the classical approach provides a close approximation to discrete time model under the assumption that the primary last stage cells develop into a detectable tumor in one time unit. This provides a reasonable explanation why the traditional approach (see [2, 22]) can still work well even though the Markov assumption for $T(t)$ may not hold. In this paper we will thus basically use the stochastic equation method and assume that cancer tumors develop from primary last stage cells through clonal expansion.

3.1. Stochastic Equations for the State Variables. Assume now that an individual is an J_i ($i = 0, 1, 2$) person at the embryo stage. Then in this individual, cancer is developed by a $(k-i)$ -stage multievent model given by $J_i \rightarrow \dots \rightarrow J_k \rightarrow \text{Tumor}$ and the identifiable response variables are given by $\{X_i(t)', T(t)\}$. To derive stochastic equations for the staging variables in $X_i(t)$ in this individual, observe that for each $i = 0, 1, 2$, $X_i(t)'$ is in general a Markov Process although $T(t)$ may not be Markov; see Remark 1, Tan [3] and Tan et al. [4, 5], Tan and Zhou [9], and Tan and Yan [16]. It follows that $X_i(t + \Delta t)'$ derive from $X_i(t)'$ through stochastic birth-death processes of J_u ($u = i, i+1, \dots, k-1$) cells and through stochastic transition $\{J_u \rightarrow J_{u+1}, u = i, i+1, \dots, k-1\}$ during $(t, t + \Delta t]$. Let $\{B_u(t; i), D_u(t; i), M_u(t; i)\}$ be the number of birth, the number of death of J_u cells, and the number of transition from $J_u \rightarrow J_{u+1}$ cells during $(t, t + \Delta t]$, respectively in people who are J_i people at the embryo stage. Let $M_0(t)$ denote the number of transitions from $N \rightarrow J_1$ during $(t, t + \Delta t]$. Because the transition of $J_u \rightarrow J_{u+1}$ would not affect the number of J_u cells but only increase the number of J_{u+1} cells (see Remark 4), by the conservation law we have the following stochastic equations for $\{J_u(t; i), u = i, \dots, k-1, i = 0, 1, 2\}$ (see Tan [3], Tan et al. [4, 5, 8], Tan and Zhou [9], and Tan and Yan [16]):

$$\begin{aligned} J_i(t + \Delta t; i) &= J_i(t; i) + B_i(t; i) - D_i(t; i), \quad i = 0, 1, 2, \\ J_u(t + \Delta t; i) &= J_u(t; i) + B_u(t; i) - D_u(t; i) \\ &\quad + M_{u-1}(t; i), \quad i < u \leq k-1. \end{aligned} \quad (2)$$

Because $\{B_v(t; i), D_v(t; i), M_v(t; i), i = 0, 1, 2, v = i, \dots, k-1\}$ are random variables, the above equations are basically stochastic equations. To derive probability distributions of these variables, let $b_u(t)$ and $d_u(t)$ denote the birth rate and the death rate at time t of the J_u ($u = 0, 1, \dots, k-1$) cells, respectively. Let $\beta_u(t)$ ($u = 0, 1, \dots, k-1$) be the transition

rate at time t from $J_u \rightarrow J_{u+1}$. Then, as shown in Tan [3], for $(i = 0, 1, 2, u = i, \dots, k-1)$ we have, to the order of $o(\Delta t)$,

$$\begin{aligned} \{B_u(t; i), D_u(t; i), M_u(t; i)\} &| J_u(t; i) \\ &\sim \text{Multinomial} \{J_u(t; i); b_u(t) \Delta t, d_u(t) \Delta t, \beta_u(t) \Delta t\}. \end{aligned} \quad (3)$$

It follows that to the order of $o(\Delta t)$,

$$\begin{aligned} \{B_u(t; i), D_u(t; i)\} &| J_u(t; i) \\ &\sim \text{Multinomial} \{J_u(t; i); b_u(t) \Delta t, d_u(t) \Delta t\}, \\ M_u(t; i) &| J_u(t; i) \\ &\sim \text{Binomial} \{J_u(t; i); \beta_u(t) \Delta t\} \\ &\sim \text{Poisson} \{J_u(t; i) \beta_u(t) \Delta t\} + o(\beta_u(t) \Delta t), \\ &\text{independently of } \{B_u(t; i), D_u(t; i)\}, \\ &u = 0, 1, \dots, k-1. \end{aligned} \quad (4)$$

From these distribution results, by subtracting from the random transition variables its conditional means, respectively, we obtain the following stochastic equations for the state variables $X_i(t)$ ($i = 0, 1, \dots, \text{Min}(2, k-1)$):

$$\begin{aligned} dJ_i(t; i) &= J_i(t + \Delta t; i) - J_i(t; i) = B_i(t; i) - D_i(t; i) \\ &= \{J_i(t; i) \gamma_i(t)\} \Delta t + e_i(t; i) \Delta t, \quad i = 0, 1, 2, \\ dJ_u(t; i) &= J_u(t + \Delta t; i) - J_u(t; i) = M_{u-1}(t; i) \\ &\quad + B_u(t; i) - D_u(t; i) \\ &= \{J_{u-1}(t; i) \beta_{u-1}(t) \\ &\quad + J_u(t; i) \gamma_u(t)\} \Delta t + e_u(t; i) \Delta t, \\ &i = 0, 1, \dots, \text{Min}(2, k-1); \quad i < u \leq k-1, \end{aligned} \quad (5)$$

where $\gamma_u(t) = b_u(t) - d_u(t)$ for $u = 0, 1, \dots, k-1$ and where $e_i(t; i) \Delta t = [B_i(t; i) - J_i(t; i) b_i(t) \Delta t] - [D_i(t; i) - J_i(t; i) d_i(t) \Delta t]$ for $i = 0, 1, 2$, $e_u(t; i) \Delta t = [B_u(t; i) - J_u(t; i) b_u(t) \Delta t] - [D_u(t; i) - J_u(t; i) d_u(t) \Delta t] + [M_{u-1}(t; i) - J_{u-1}(t; i) \beta_{u-1}(t) \Delta t]$ for $i = 0, 1, \dots, \text{Min}(2, k-1), i < u \leq k-1$.

From the above equations, by dividing both sides by Δt and letting $\Delta t \rightarrow 0$ we obtain

$$\begin{aligned} \frac{J_i(t; i)}{dt} &= J_i(t; i) \gamma_i(t) + e_i(t; i), \quad i = 0, 1, \dots, \text{Min}(2, k-1); \\ \frac{J_u(t; i)}{dt} &= J_u(t; i) \gamma_u(t) + J_{u-1}(t; i) \beta_{u-1}(t) + e_u(t; i), \\ &\text{for } i = 0, 1, \text{Min}(2, k-1); \quad i < u \leq k-1. \end{aligned} \quad (6)$$

In the above equations, using the distribution results in (4) it can easily be shown that the random noises $\{e_u(t; i), u = i, \dots, k-1, i = 0, 1, 2\}$ have expected value zero and are

uncorrelated with the staging variables and $T(t)$. The initial conditions at birth (t_0) for the above stochastic differential equations are $\{J_u(t_0; i) > 0, u = i, i + 1; J_u(t_0; i) = 0, u > i + 1\}$.

Given the initial conditions ($J_u(t_0; i) > 0, u = i, i + 1$) and ($J_u(t_0; i) = 0, u > i + 1$) at birth (t_0), the solution of the equations in (6) is given, respectively, by

$$\begin{aligned} J_i(t; i) &= J_i(t_0; i) e^{\int_{t_0}^t \gamma_i(x) dx} + \eta_i(t; i), \\ i &= 0, 1, \dots, \text{Min}(2, k - 1), \\ J_u(t; i) &= J_u(t_0; i) e^{\int_{t_0}^t \gamma_u(x) dx} \\ &+ \int_{t_0}^t J_{u-1}(x; i) \beta_{u-1}(x) e^{\int_x^t \gamma_u(y) dy} dx \\ &+ \eta_u(t; i) = \dots = J_u(t_0; i) e^{\int_{t_0}^t \gamma_u(x) dx} \\ &+ \sum_{v=1}^{u-i} J_{u-v}(t_0; i) \phi_u^{(v)}(t; i) + \sum_{v=1}^{u+1-i} \eta_{u-v}^{(v)}(t; i), \\ u &= i + 1, \dots, k - 1, \end{aligned} \quad (7)$$

where $i = 0$ if $k = 2$,

$i = 0, 1$ if $k = 3$,

$i = 0, 1, 2$, if $k > 3$,

where

$$\begin{aligned} \phi_u^{(1)}(t; i) &= \int_{t_0}^t e^{\int_x^t \gamma_u(y) dy + \int_{t_0}^x \gamma_{u-1}(y) dy} \beta_{u-1}(x) dx \\ u &= i, \dots, k - 1, \\ \phi_u^{(v)}(t; i) &= \int_{t_0}^t e^{\int_x^t \gamma_u(y) dy} \beta_{u-1}(x) \phi_{u-1}^{(v-1)}(x; i) dx, \\ v &= 2, \dots, u - i, \\ \eta_u(t; i) &= \eta_u^{(1)}(t; i) \\ &= \int_{t_0}^t e^{\int_x^t \gamma_u(y) dy} e_u(x; i) dx, \\ u &= i, \dots, k - 1, \\ \eta_{u-v}^{(v)}(t; i) &= \int_{t_0}^t e^{\int_x^t \gamma_u(y) dy} \beta_{u-1}(x) \eta_{u-1}^{(v-1)}(x; i) dx, \\ v &= 2, \dots, u - i. \end{aligned} \quad (8)$$

If the model is time homogeneous so that $\{\beta_u(t) = \beta_u, b_u(t) = b_u, d_u(t) = d_u, \gamma_u(t) = b_u - d_u = \gamma_u, u = 0, 1, \dots, k - 1\}$ and if $\gamma_i \neq \gamma_u$ if $i \neq u$, the above solutions under the initial

conditions ($J_{i+u}(t_0; i) > 0, u = 0, 1; J_{i+u}(t_0; i) = 0, u > 1$) then reduce, respectively, to

$$\begin{aligned} J_i(t; i) &= J_i(t_0; i) e^{\gamma_i(t-t_0)} + \eta_i^{(1)}(t; i), \\ i &= 0, 1, \dots, \text{Min}(2, k - 1), \\ J_u(t; i) &= J_u(t_0; i) e^{\gamma_u(t-t_0)} \\ &+ \beta_{u-1} \int_{t_0}^t J_{u-1}(x; i) e^{\gamma_u(t-x)} dx + \eta_u^{(1)}(t; i) \\ &= \dots = J_u(t_0; i) e^{\gamma_u(t-t_0)} \\ &+ \sum_{r=i}^{u-1} J_r(t_0; i) \left(\prod_{v=r}^{u-1} \beta_v \right) \sum_{l=r}^u A_{ru}(l) e^{\gamma_l(t-t_0)} \\ &+ \sum_{r=i}^u \eta_{u-r}^{(u+1-r)}(t; i) \\ &= \sum_{r=i}^{i+1} J_r(t_0; i) \left(\prod_{v=r}^{u-1} \beta_v \right) \sum_{l=r}^u A_{ru}(l) e^{\gamma_l(t-t_0)} \\ &+ \sum_{r=i}^u \eta_{u-r}^{(u+1-r)}(t; i), \quad i < u \leq k - 1, \end{aligned} \quad (9)$$

where for $i \leq v \leq u$,

$$\begin{aligned} A_{iu}(v) &= 1, \quad \text{if } i = u \\ &= \prod_{l=i, l \neq v}^u (\gamma_l - \gamma_v)^{-1}, \quad \text{if } i < u. \end{aligned} \quad (10)$$

Obviously, $E[\eta_u^{(r)}(t; i)] = 0$ for all $(i = 0, 1, 2, u = i, \dots, k - 1, r = 1, \dots, u - i + 1)$. It follows that for $(i = 0, 1, \dots, \text{Min}(k - 1, 2))$, the expected values $E[J_{k-1}(t; i)]$ of $J_{k-1}(t; i)$ for homogeneous models with $\gamma_i \neq \gamma_u$ if $i \neq u$ are given by

$$\begin{aligned} E[J_{k-1}(t; i)] &= E[J_{i+1}(t_0; i)] \\ &\times \left(\prod_{u=i+1}^{k-2} \beta_u \right) \sum_{v=i+1}^{k-1} A_{(i+1)(k-1)}(v) e^{\gamma_v(t-t_0)} \\ &+ E[J_i(t_0; i)] \left(\prod_{u=i}^{k-2} \beta_u \right) \sum_{v=i}^{k-1} A_{i(k-1)}(v) e^{\gamma_v(t-t_0)}, \\ i &= 0, 1, \text{Min}(k - 1, 2), \quad k \geq 2, \end{aligned} \quad (11)$$

where as a convention, $(\sum_{j=i+1}^i c_i = 0, \prod_{j=i+1}^i d_j = 1)$ for all (c_i, d_j) .

Remark 4. Because genetic changes and epigenetic changes occur during cell division, to the order of $o(\Delta t)$, the probability is $\beta_u(t) \Delta t$ that one J_u cell at time t would give rise to 1 J_u cell and 1 J_{u+1} cell at time $t + \Delta t$ by genetic changes or epigenetic changes. It follows that the transition of $J_u \rightarrow J_{u+1}$ would not affect the population size of J_u cells but only increase the size of the J_{u+1} population.

3.2. Transition Probabilities and Probability Distributions of Staging Variables. Let $f(x; n, p)$ denote the probability density function of a binomial random variable $X \sim \text{Binomial}(n; p)$, $h(x; \lambda)$ the probability density function of a Poisson random variable $X \sim \text{Poisson}(\lambda)$, and $g(x, y; n, p_1, p_2)$ the probability density function of a bivariate multinomial random vector $(X, Y) \sim \text{Multinomial}(n; p_1, p_2)$. Using the stochastic equations of the staging variables given by (2) and using the probability distributions of the transition variables $\{B_r(t; i), D_r(t; i), M_r(t; i)\}$ in (4), as in Tan et al. [4, 5], we obtain the following transition probabilities of $\{J_r(t + \Delta t; i), r + i, \dots, k - 1\}$ given $\{J_r(t; i), r = i, \dots, k - 1\}$ for $(i = 0, 1, \dots, \text{Min}(k - 1, 2))$:

$$\begin{aligned} P\{J_r(t + \Delta t; i) = v_r, r = i, i + 1, \dots, k - 1 \mid J_r(t; i) \\ = u_r, r = i, i + 1, \dots, k - 1\} \\ = P\{J_i(t + \Delta t; i) = v_i \mid J_i(t; i) = u_i\} \\ \times \prod_{j=i+1}^{k-1} P\{J_j(t + \Delta t; i) = v_j \mid J_j(t; i) \\ = u_j, J_{j-1}(t; i) = u_{j-1}\}, \end{aligned} \quad (12)$$

where

$$\begin{aligned} P\{J_i(t + \Delta t; i) = v_i \mid J_i(t; i) = u_i\} \\ = \sum_{r=0}^{u_i} f(r; u_i, b_i(t) \Delta t) \\ \times f\left(u_i - v_i + r; u_i - r, \frac{d_i(t) \Delta t}{1 - b_i(t) \Delta t}\right), \\ i = 0, 1, \dots, \text{Min}(k - 1, 2), \end{aligned} \quad (13)$$

$$\begin{aligned} P\{J_j(t + \Delta t; i) = v_j \mid J_j(t; i) = u_j, J_{j-1}(t; i) = u_{j-1}\} \\ = \sum_{r_1=0}^{u_j} \sum_{r_2=0}^{u_{j-1}-r_1} g(r_1, r_2; u_j, b_j(t) \Delta t, d_j(t) \Delta t) \\ \times h(v_j - u_j - r_1 + r_2; u_{j-1} \beta_{j-1} \Delta t), \quad j > i. \end{aligned}$$

Define the unobservable transition variables $U_i(t) = \{B_i(t; i), (B_j(t; i), D_j(t; i)), j = i + 1, \dots, k - 1\}'$ ($i = 0, 1, \text{Min}(k - 1, 2)$). Then, we have for the joint probability density function of $\{X_i(t + \Delta t), U_i(t)\}$ given $X_i(t)$

$$\begin{aligned} P\{X_i(t + \Delta t), U_i(t) \mid X_i(t)\} = P\{X_i(t + \Delta t) \mid U_i(t), X_i(t)\} \\ \times P\{U_i(t) \mid X_i(t)\}, \end{aligned} \quad (14)$$

where

$$\begin{aligned} P\{X_i(t + \Delta t) \mid U_i(t), X_i(t)\} \\ = f\left(J_i(t; i) - J_i(t + \Delta t; i) + B_i(t; i); J_i(t; i) \right. \\ \left. - B_i(t; i), \frac{d_i(t) \Delta t}{1 - b_i(t) \Delta t}\right) \\ \times \prod_{j=i+1}^{k-1} h\{J_j(t + \Delta t; i) - J_j(t; i) - B_j(t; i) \\ + D_j(t; i); J_{j-1}(t; i) \beta_{j-1}(t) \Delta t\}, \end{aligned} \quad (15)$$

$$\begin{aligned} P\{U_i(t) \mid X_i(t)\} \\ = f(B_i(t; i); J_i(t; i), b_i(t; i) \Delta t) \\ \times \prod_{j=i+1}^{k-1} g\{B_j(t; i), D_j(t; i); J_j(t; i), b_j(t; i) \Delta t, d_j(t; i) \Delta t\}. \end{aligned} \quad (16)$$

Let $e_i(j)$ be a $(k - i) \times 1$ column vector with 1 in the j th position ($1 \leq j \leq k - i$) and with 0 in other positions. Let $u = (u_i, \dots, u_{k-1})'$ and $v = (v_i, \dots, v_{k-1})'$ be $(k - i) \times 1$ column vectors of nonnegative integers. (i.e., u_j and v_j are nonnegative integers). Then, by using the probability distribution results in (14)–(16) it can readily be shown that

$$\begin{aligned} P\{X_i(t + \Delta t) = v \mid X_i(t) = u\} \\ = [u_j b_j(t) + (1 - \delta_{ji}) u_{j-1} \beta_{j-1}(t)] \Delta t + o(\Delta t), \\ \text{if } v = u + e_i(j), \quad j = i, \dots, k - 1 \\ = u_j d_j(t) \Delta t + o(\Delta t), \\ \text{if } v = u - e_i(j), \quad j = i, \dots, k - 1 \\ = o(\Delta t) \quad \text{if } |1'_{n-k} (u - v)| \geq 2. \end{aligned} \quad (17)$$

The above results imply that $X_i(t)$ is a $(k - i)$ -dimensional birth-death process with birth rates $\{ib_u(t), u = i, \dots, k - 1\}$, death rates $\{id_u(t), u = i, \dots, k - 1\}$, and cross-transition rates $\{\alpha_{u,u+1}(j, t) = j\beta_u(t), u = i, \dots, k - 1, \beta_{u,v}(j, t) = 0 \text{ if } v \neq u + 1\}$. (See Definition 4.1 in Tan ([22], Chapter 4)). Using these results, it can be shown that the Kolmogorov forward equation for the probabilities $P\{J_j(t; i) = u_j, j = i, \dots, k - 1 \mid J_i(0) = m_i, J_j(0) = 0,$

$j > i\} = P(u_j, j = i, \dots, k-1; t) (i = 0, 1, \text{Min}(k-1, 2))$ in the above model is given by

$$\begin{aligned}
& \frac{d}{dt} P(u_j, j = i, \dots, k-1; t) \\
&= P(u_i - 1, u_j, j = i+1, \dots, k-1; t) (u_i - 1) b_i(t) \\
&+ \sum_{j=i+1}^{k-1} P(u_i, u_{i+1}, \dots, u_{j-1}, u_j \\
&\quad -1, u_{j+1}, \dots, u_{k-1}; t) (u_j - 1) b_j(t) \\
&+ \sum_{j=i}^{k-2} P(u_i, u_{i+1}, \dots, u_j, u_{j+1} \\
&\quad -1, u_{j+2}, \dots, u_{k-1}; t) u_j \beta_j(t) \\
&+ \sum_{j=i}^{k-1} P(u_i, u_{i+1}, \dots, u_{j-1}, u_j \\
&\quad +1, u_{j+1}, \dots, u_{k-1}; t) (u_j + 1) d_j(t) \\
&- P(u_i, u_{i+1}, \dots, u_{k-1}; t) \\
&\times \left\{ \sum_{j=i}^{k-1} u_j [b_j(t) + d_j(t)] + \sum_{j=i}^{k-2} u_j \beta_j(t) \right\}, \tag{18}
\end{aligned}$$

for $u_j = 0, 1, \dots, \infty, j = i, \dots, k-1$.

By using the above set of differential equations, one can readily compute the probabilities $P\{J_j(t) = u_j, j = i, \dots, k-1 \mid I_i(0) = m_i\} = P(u_j, j = i, \dots, k-1; t)$ numerically.

3.3. The Probability Distributions of the Number of Detectable Tumors and Times to Tumors. As shown by Yang and Chen [11], malignant cancer tumors arise from primary J_k cells by clonal expansion, where primary J_k cells are J_k cells generated directly by J_{k-1} cells. (J_k cells derived by stochastic birth of other J_k cells are not primary J_k cells). That is, cancer tumors develop from primary J_k cells through stochastic birth-death processes.

To derive the probability distribution for $T(t)$ in J_i people in the population, let $P_T(s, t)$ denote the probability that a primary cancer cell at time s develops into a detectable cancer tumor by time t . (Explicit formula for $P_T(s, t)$ has been given in Tan [22], Chapter 8 and in Tan and Chen [24]). Than, as shown in Tan ([3, 22], chapter 8), the conditional probability distribution of $T(t)$ given $\{J_{k-1}(s; i), s \leq t\}$ in J_i people is Poisson with mean $\omega(t; i)$, where $\omega(t; i) = \int_{t_0}^t J_{k-1}(s; i) \beta_{k-1}(s) P_T(s, t) ds$. That is,

$$T(t) \mid \{J_{k-1}(s; i), s \leq t\} \sim \text{Poisson}(\omega(t; i)). \tag{19}$$

Let $Q_i(j)$ be the probability that cancer tumors develop during $(t_{j-1}, t_j]$ in J_i people in the population. For time homogeneous models with small β_{k-1} , $Q_i(j)$ is then given by

$$\begin{aligned}
Q_i(j) &= E \{e^{-\omega(t_{j-1}; i)} - e^{-\omega(t_j; i)}\} \\
&= e^{-\beta_{k-1} H_i(t_{j-1})} - e^{-\beta_{k-1} H_i(t_j)} + o(\beta_{k-1}), \tag{20}
\end{aligned}$$

where $H_i(t) = \int_{t_0}^t E[J_{k-1}(x; i)] P_T(x, t) dx$.

To derive $Q_i(j)$, denote by

$$\begin{aligned}
\theta_{i(k-1)} &= E[J_i(t_0; i-1)] \beta_i \prod_{u=i+1}^{k-1} \left(\frac{\beta_u}{\gamma_u}\right), \\
i &= 1, \dots, \text{Min}(3, k-1), \\
\lambda_{u(k-1)} &= E[J_u(t_0; u)] \beta_u \prod_{v=u+1}^{k-1} \left(\frac{\beta_v}{\gamma_v}\right), \\
u &= 0, 1, \dots, \text{Min}(2, k-1), \tag{21}
\end{aligned}$$

and define the functions

$$\begin{aligned}
\psi_{i(k-1)}(t) &= \left\{ \prod_{u=i+1}^{k-1} \gamma_u \right\} \sum_{v=i}^{k-1} A_{i(k-1)}(v) \\
&\quad \times \int_{t_0}^t e^{\gamma_u(x-t_0)} P_T(x, t) dx, \tag{22} \\
i &= 0, 1, \dots, \text{Min}(2, k-1).
\end{aligned}$$

Applying results of $E[J_{k-1}(t; i)]$ given in (11), for time homogeneous models with $\gamma_i \neq \gamma_j$ if $i \neq j$ we obtain $Q_i(j)$'s as follows.

- (1) If $k = 2$, then $k-1 = 1$. Hence, $Q_2(0) = 1$, $Q_i(0) = Q_2(j) = 0$ for $(i = 0, 1, j > 0)$ and for $j > 0$,

$$\begin{aligned}
Q_0(j) &= \{e^{-\theta_{11} \psi_{11}(t_{j-1}) - \lambda_{01} \psi_{01}(t_{j-1})} \\
&\quad - e^{-\theta_{11} \psi_{11}(t_j) - \lambda_{01} \psi_{01}(t_j)}\} + o(\beta_1), \tag{23}
\end{aligned}$$

$$\begin{aligned}
Q_1(j) &= (1 - \alpha_1) \{e^{-\lambda_{11} \psi_{11}(t_{j-1})} \\
&\quad - e^{-\lambda_{11} \psi_{11}(t_j)}\} + o(\beta_1). \tag{24}
\end{aligned}$$

(2) If $k \geq 3$, then we have $Q_i(0) = 0$ for $i = 0, 1$ and $Q_2(0) = \delta_{2k}\alpha$ and for $j > 0$,

$$Q_0(j) = \left\{ e^{-\theta_{1(k-1)}\Psi_{1(k-1)}(t_{j-1}) - \lambda_{0(k-1)}\Psi_{0(k-1)}(t_{j-1})} - e^{-\theta_{1(k-1)}\Psi_{1(k-1)}(t_j) - \lambda_{0(k-1)}\Psi_{0(k-1)}(t_j)} \right\} + o(\beta_{k-1}), \quad (25)$$

$$Q_1(j) = \left\{ e^{-\theta_{2(k-1)}\Psi_{2(k-1)}(t_{j-1}) - \lambda_{1(k-1)}\Psi_{1(k-1)}(t_{j-1})} - e^{-\theta_{2(k-1)}\Psi_{2(k-1)}(t_j) - \lambda_{1(k-1)}\Psi_{1(k-1)}(t_j)} \right\} + o(\beta_{k-1}), \quad (26)$$

$$Q_2(j) = \delta_{2k}(1 - \alpha) \left\{ e^{-\lambda_{22}\Psi_{22}(t_{j-1})} - e^{-\lambda_{22}\Psi_{22}(t_j)} \right\} + (1 - \delta_{2k}) \left\{ e^{-\theta_{3(k-1)}\Psi_{3(k-1)}(t_{j-1}) - \lambda_{2(k-1)}\Psi_{2(k-1)}(t_{j-1})} - e^{-\theta_{3(k-1)}\Psi_{3(k-1)}(t_j) - \lambda_{2(k-1)}\Psi_{2(k-1)}(t_j)} \right\} + o(\beta_{k-1}), \quad (27)$$

where $\delta_{2k} = 1$ if $k = 2$ and $= 0$ if $k \neq 2$.

Notice that if $\gamma_0 = 0$, then $\psi_{0(k-1)}(t)$ reduces to

$$\begin{aligned} \psi_{0(k-1)}(t) &= \left(\prod_{u=1}^{k-1} \gamma_u \right) \sum_{v=0}^{k-1} A_{0(k-1)}(v) \\ &\quad \times \int_{t_0}^t e^{\gamma_v(x-t_0)} P_T(x, t) dx \\ &= \left(\prod_{u=1}^{k-1} \gamma_u \right) \sum_{v=1}^{k-1} A_{1(k-1)}(v) \\ &\quad \times \int_{t_0}^t [e^{\gamma_v(x-t_0)} - 1] P_T(x, t) dx. \end{aligned} \quad (28)$$

Notice also that if $P_T(s, t) \approx 1$ for $t > s$ and if $\gamma_0 = 0$, then the above $\psi_{ij}(t)$'s reduce, respectively, to

$$\begin{aligned} \psi_{ii}(t) &= \frac{1}{\gamma_i} \{ e^{\gamma_i(t-t_0)} - 1 \}, \quad i = 1, \dots, k-1, \\ \psi_{0(k-1)}(t) &= \left(\prod_{u=1}^{k-1} \gamma_u \right) \sum_{v=1}^{k-1} A_{1(k-1)}(v) \\ &\quad \times \frac{1}{\gamma_v^2} \{ e^{\gamma_v(t-t_0)} - 1 - (t-t_0)\gamma_v \}, \quad (29) \\ \psi_{i(k-1)}(t) &= \left(\prod_{u=i+1}^{k-1} \gamma_u \right) \sum_{v=i}^{k-1} A_{i(k-1)}(v) \\ &\quad \times \frac{1}{\gamma_v} \{ e^{\gamma_v(t-t_0)} - 1 \}, \quad i = 1, \dots, k-1. \end{aligned}$$

4. Probability Distribution of Observed Cancer Incidence Incorporating Hereditary Cancer Cases

For estimating unknown parameters and to validate the model, one would need real data generated from the model. For studies of carcinogenesis such data are usually given by cancer incidence. For example, in the SEER data of NCI/NIH of USA, the data are given by $\{(y_0, n_0), (y_j, n_j), j = 1, \dots, t_N\}$, where y_0 is the number of cancer cases at birth and n_0 the total number of birth, and where for $j \geq 1$, y_j is the number of cancer cases developed during the j th age group of a one-year period (or 5 years periods) and n_j is the number of noncancer people who are at risk for cancer and from whom y_j of them have developed cancer during the j th age group. Given in Table 1 are the SEER data of uveal melanoma (adult eye cancer) during the period 1973–2007. In Table 1, notice that there are some cancer cases at birth implying some inherited cancer cases. In this section, we will develop a statistical model for these types of data sets from the stochastic multistage model with hereditary cancers as given in Section 2. As in previous sections, let n_{ij} be the number of individuals who have genotype J_i ($i = 0, 1, 2$) at the embryo stage among the n_j people at risk for the cancer in question. Then, as showed above, $(n_{1j}, n_{2j}) \mid n_j \sim \text{Multinomial}\{n_j; p_1, p_2\}$. It follows that $n_{ij} \mid n_j \sim \text{Binomial}\{n_j, p_i\}, i = 0, 1, 2$. In what follows, we let Y_j denote the random variable for y_j unless otherwise stated.

4.1. The Probability Distribution of Y_0 . As shown in Figure 4, J_i ($i = 0, 1, 2$) people would only generate J_i stage cells and J_{i+1} stage cells at birth. Thus, for cancers to develop at or before birth, the number of stages for the stochastic model of carcinogenesis must be 3 or less. It follows that if $\gamma_0 > 0$, the appropriate model of carcinogenesis must be either a 2-stage model or a 3-stage model. Since $n_{20} \mid n_0 \sim \text{Binomial}(n_0; p_2)$ and $n_{10} \mid (n_0, n_{20}) \sim \text{Binomial}(n_0; p_1) + o(p_2)$, the probability distribution of Y_0 is therefore

$$Y_0 \sim \text{Poisson}(\chi_k) \quad k = 2, 3, \quad (30)$$

where

$$\begin{aligned} \chi_k &= n_0(p_2 + p_1\alpha) \quad \text{if } k = 2 \\ &= n_0 p_2 \alpha \quad \text{if } k = 3. \end{aligned} \quad (31)$$

The expected number of Y_0 given n_0 is $E(Y_0 \mid n_0) = n_0(p_2 + p_1\alpha) = \chi_2$ if $k = 2$ and $E(Y_0 \mid n_0) = n_0 p_2 \alpha = \chi_3$ if $k = 3$. Hence, for the 2-stage model (i.e., $k = 2$) or the 3-stage model (i.e., $k = 3$), the maximum likelihood estimate of χ_k is $\hat{\chi}_k = y_0$ and the deviance $D_0(k)$ from the conditional probability distribution of Y_0 given n_0 is

$$\begin{aligned} D_0(k) &= -2 \{ \log \{h(y_0; \chi_k)\} - \log \{h(y_0; \hat{\chi}_k)\} \} \\ &= \{\chi_k - y_0\} - y_0 \log \left\{ \frac{\chi_k}{y_0} \right\}, \quad k = 2, 3. \end{aligned} \quad (32)$$

TABLE 1: The SEER incidence data (1973–2007) of uveal melanoma from NCI/NIH (over all races and genders).

Age groups	Number of people at risk	Observed incidence	Model-F predicated	Model-1 predicated	Two-stage predicated
0	12,495,777	34	36	36	38
1	12,221,582	20	11	11	16
2	12,120,990	14	7	8	17
3	12,112,995	9	5	6	18
4	12,146,174	4	4	5	20
5	12,161,336	3	3	4	22
6	12,111,854	2	2	4	25
7	12,160,452	2	2	4	28
8	11,942,586	2	2	4	30
9	12,381,299	1	2	5	34
10	12,512,703	2	3	6	38
11	12,410,338	5	3	6	41
12	12,449,244	2	3	6	44
13	12,527,781	5	4	7	48
14	12,602,883	3	5	7	51
15	12,719,598	5	5	8	55
16	12,766,107	7	6	9	59
17	12,831,400	9	7	9	62
18	12,382,047	8	7	9	63
19	12,581,638	10	8	10	68
20	12,636,509	7	9	11	71
21	12,682,601	6	10	10	75
22	12,840,510	12	11	11	79
23	13,075,528	17	13	13	84
24	13,358,635	16	15	14	89
25	13,473,849	12	16	16	94
26	13,426,340	17	18	18	97
27	13,525,264	28	20	20	101
28	13,149,674	17	22	21	102
29	13,812,811	23	25	25	110
30	13,886,874	24	28	27	114
31	13,488,332	37	30	29	115
32	13,460,286	32	32	32	118
33	13,256,067	38	35	35	119
34	13,428,827	37	39	38	124
35	13,220,037	40	41	41	126
36	12,870,265	30	44	44	126
37	12,689,592	43	47	47	127
38	12,157,014	42	49	49	125
39	12,494,081	46	55	54	131
40	12,272,125	49	58	58	132
41	11,826,573	56	61	60	130
42	11,663,153	54	65	64	131
43	11,407,082	53	68	68	131
44	11,296,848	70	73	72	133

TABLE 1: Continued.

Age groups	Number of people at risk	Observed incidence	Model-F predicated	Model-1 predicated	Two-stage predicated
45	11,016,369	57	76	76	132
46	10,651,593	71	79	79	130
47	10,475,708	87	84	83	131
48	9,994,684	82	86	85	127
49	10,138,908	78	93	93	131
50	9,836,359	87	97	96	130
51	9,475,641	95	100	99	127
52	9,250,985	113	104	104	126
53	9,027,382	106	108	108	125
54	8,883,737	117	113	113	125
55	8,547,883	129	116	116	123
56	8,279,648	107	119	119	121
57	8,062,368	119	123	123	119
58	7,654,610	132	124	124	115
59	7,563,706	118	130	129	115
60	7,232,719	131	131	130	112
61	6,927,332	116	132	132	109
62	6,708,273	133	134	134	107
63	6,543,931	143	138	137	106
64	6,404,652	130	141	141	105
65	6,168,486	145	142	142	102
66	5,913,479	138	142	142	99
67	5,746,766	151	144	144	98
68	5,480,517	147	142	142	94
69	5,363,912	149	144	144	94
70	5,110,728	136	142	142	90
71	4,925,076	144	141	141	88
72	4,696,825	140	138	138	85
73	4,512,136	146	135	135	82
74	4,345,300	126	133	133	80
75	4,148,801	122	128	129	78
76	3,900,900	124	122	122	74
77	3,681,587	114	116	116	70
78	3,481,918	93	110	110	67
79	3,243,631	102	102	102	63
80	2,961,234	79	92	93	58
81	2,724,984	93	83	84	54
82	2,495,219	82	75	75	50
83	2,271,595	92	66	67	46
84	2,041,351	64	57	58	42
85	10,466,605	304	282	285	216

Note 1: for age 0 and 1–10 years old, the cancer incidence are derived by subtracting incidence of retinoblastoma from the original SEER data (see Tan and Zhou [9]).

Note 2: the observed uveal melanoma incidence rates per 10^6 individuals are derived from the SEER eye cancer incidence by subtracting retinoblastoma incidence as given in Tan and Zhou [9].

4.2. *The Probability Distribution of Y_j ($j \geq 1$).* To derive the probability distribution of Y_j ($j \geq 1$) in the j th age group, let Y_{ij} ($i = 0, 1, 2$) be the number of cancer cases generated by people who have genotype I_i at the embryo stage among these Y_j cancer cases. Then $Y_j = \sum_{i=0}^2 Y_{ij}$ and Y_{0j} is the number of cancer cases generated by the $n_{0j} = n_j - n_{1j} - n_{2j}$ normal people in the population. The conditional probability distribution of Y_{ij} given n_{ij} is

$$Y_{ij} | n_{ij} \sim \text{Poisson} \{n_{ij} Q_i(j)\}, \quad i = 0, 1, 2. \quad (33)$$

Notice that if $k = 2$ (a 2-stage model), then all J_2 individuals would develop tumor at or before birth. Thus, if $k = 2$, then $Y_{2j} = 0$ for all $j > 0$ so that if $j > 0$, cancer cases develop only from normal people ($N = J_0$ people) and J_1 people. On the other hand, if $k > 2$, then with positive probability $Y_{ij} > 0$ for all ($i = 0, 1, 2, j = 0, 1, \dots, n_T$), where n_T is the last time point in the data. Let $\delta_{2k} = 1$ if $k = 2$ and $\delta_{2k} = 0$ if $k \neq 2$. Then, $Y_j | (n_{ij}, i = 0, 1, 2) \sim \text{Poisson}(Q_T(j))$, where $Q_T(j) = \sum_{i=0}^1 n_{ij} Q_i(j) + (1 - \delta_{2k}) n_{2j} Q_2(j)$. Since $(n_{0j}, n_{1j}) | n_j \sim \text{Multinomial}(n_j; p_0, p_1)$, we have for the conditional probability density function $P(y_j | n_j)$ of Y_j given n_j

$$P(y_j | n_j) = \sum_{n_{0j}=0}^{n_j} \sum_{n_{1j}=0}^{n_j - n_{0j}} g(n_{0j}, n_{1j}; n_j, p_0, p_1) h\{y_j; Q_T(j)\}, \quad (34)$$

where $g(n_{0j}, n_{1j}; n_j, p_0, p_1)$ is the probability density function of $(n_{0j}, n_{1j}) | n_j \sim \text{Multinomial}(n_j; p_0, p_1)$ and $h\{y_j; Q_T(j)\}$ the probability density function of $Y_j | (n_{ij}, i = 0, 1, 2) \sim \text{Poisson}\{Q_T(j)\}$.

The probability density function $P(y_j | n_j)$ given by (34) is a mixture of Poisson probability density functions with mixing probability density function given by the multinomial probability distribution of $\{n_{0j}, n_{1j}\}$ given n_j . This mixing probability density function represents individuals with different genotypes at the embryo stage in the population.

Let Θ be the set of all unknown parameters (i.e., the parameters (p_1, p_2, α) and the birth rates, the death rates, and the mutation rates of J_j cells). Based on data $(y_j, j = 0, 1, \dots, n_T)$, the likelihood function of Θ is

$$L\{\Theta | y_j, j = 0, 1, \dots, n_T\} = h(y_0; \chi(k)) \prod_{j=1}^{n_T} P(y_j | n_j). \quad (35)$$

Notice that because the mutation rates are very small, one may practically assume $\beta_i(t) = \beta_i$ for $i = 0, 1, \dots, k-1$. Also, because the stage-limiting genes are basically tumor suppressor genes which act recessively (see Tan [3], Weinberg [6], and Tan et al. [5, 8, 23]), one may practically assume $\{b_i(t) = b_i, d_i(t) = d_i, \gamma_i(t) = b_i - d_i = \gamma_i, i = 0, 1, \dots, k-1\}$ (see Tan et al. [4, 5, 8]).

4.3. *The Joint Probability Distribution of Augmented Variables and Cancer Incidence.* For applying the mixture distribution

of Y_j in (34) to make inference about the unknown parameters, one needs to expand the model to include the unobservable augmented variables $(n_{0j}, n_{1j}, y_{0j}, y_{1j}, j = 1, \dots, t_N)$ and derives the joint probability distribution of these variables. For these purposes, observe that for ($j = 1, \dots, n_T$) and for $k > 2$, the conditional probability distribution $P\{y_{ij}, i = 0, 1 | y_j, n_{ij}, i = 0, 1, n_j\}$ of $(Y_{ij}, i = 0, 1)$ given $(Y_j = y_j, n_{ij}, i = 0, 1, n_j)$ is

$$\begin{aligned} & (y_{ij}, i = 0, 1) | (y_j, n_{ij}, i = 0, 1, n_j) \\ & \sim \text{Multinomial} \left(y_j; \frac{n_{ij} Q_i(j)}{Q_T(j)}, i = 0, 1 \right). \end{aligned} \quad (36)$$

Since the conditional probability distribution of Y_j given $(n_{ij}, i = 0, 1, 2)$ for $k > 2$ is Poisson with mean $Q_T(j)$, we have for the joint conditional probability density function $P\{y_{ij}, i = 0, 1, y_j | n_{ij}, i = 0, 1, n_j\}$ of $(Y_{ij}, i = 0, 1, Y_j)$ given $(n_{ij}, i = 0, 1, n_j)$

$$\begin{aligned} & P\{y_{ij}, i = 0, 1, y_j | n_{ij}, i = 0, 1, n_j\} \\ & = P\{y_{ij}, i = 0, 1 | y_j, n_{ij}, i = 0, 1, 2\} \\ & \quad \times P\{y_j | n_{ij}, j = 0, 1, 2\} = \frac{1}{y_j!} e^{-Q_T(j)} \{Q_T(j)\}^{y_j} \\ & \quad \times g\left\{y_{0j}, y_{1j}; y_j, \frac{n_{ij} Q_i(j)}{Q_T(j)}, i = 0, 1\right\} \\ & = \prod_{i=0}^2 h\{y_{ij}; n_{ij} Q_i(j)\}, \quad (k > 2), \end{aligned} \quad (37)$$

where $y_{2j} = y_j - \sum_{u=0}^1 y_{uj}$ and $n_{2j} = n_j - \sum_{u=0}^1 n_{uj}$.

If $k = 2$, then $Y_{2j} = 0$ so that $Y_j = \sum_{i=0}^1 Y_{ij}$. Thus, we have for $k = 2$,

$$Y_j | (n_{ij}, i = 0, 1, n_j) \sim \text{Poisson} \left\{ \sum_{i=0}^1 n_{ij} Q_i(j) \right\}, \quad (38)$$

$$\begin{aligned} & Y_{ij} | (Y_j = y_j, n_{ij}, i = 0, 1, n_j) \\ & \sim \text{Binomial} \left(y_j; \frac{n_{ij} Q_i(j)}{\sum_{u=0}^1 n_{uj} Q_u(j)} \right), \quad i = 0, 1. \end{aligned} \quad (39)$$

It follows that if $k = 2$, then $\sum_{i=0}^1 Y_{ij} = Y_j$ and the joint probability density function of (Y_{0j}, Y_j) given $(n_{uj}, u = 0, 1, 2)$ is

$$\begin{aligned} & P\{y_{0j}, y_j | n_{uj}, u = 0, 1, n_j\} \\ & = P\{y_j | n_{ij}, i = 0, 1, n_j\} \\ & \quad \times P\{y_{0j} | y_j, n_{uj}, u = 0, 1, n_j\} \\ & = \prod_{i=0}^1 h\{y_{ij}; n_{ij} Q_i(j)\}, \quad \text{if } k = 2, \end{aligned} \quad (40)$$

where $y_j = \sum_{i=0}^1 y_{ij}$ and $n_j = \sum_{i=0}^2 n_{ij}$.

Put $\{\mathbf{Y} = (y_{ij}, i = 0, 1, j = 1, \dots, t_N), \mathbf{N} = (n_{ij}, i = 0, 1, j = 1, \dots, t_N), \mathbf{y} = (y_j, j = 0, 1, \dots, t_N)\}$. From (37) and (40), we have for the conditional joint probability density function of $(\mathbf{Y}, \mathbf{y})$ given $(\mathbf{N}, \mathbf{y})$

$$P\{\mathbf{Y}, \mathbf{y} | \mathbf{N}, \mathbf{y}, \Theta\} = h(y_0; \chi(k)) \prod_{j=1}^{t_N} \{h[y_{2j}; n_{2j} Q_2(j)]\}^{1-\delta_{2k}} \\ \times \prod_{i=0}^1 h\{y_{ij}; n_{ij} Q_i(j)\}. \quad (41)$$

It follows that the joint conditional probability density function of $(\mathbf{N}, \mathbf{Y}, \mathbf{y})$ given $(\mathbf{y}, \Theta)$ is

$$P\{\mathbf{N}, \mathbf{Y}, \mathbf{y} | \mathbf{y}, \Theta\} = h(y_0; \chi(k)) \prod_{j=1}^{t_N} g(n_{0j}, n_{1j}; n_j, p_0, p_1) \\ \times \{h[y_{2j}; n_{2j} Q_2(j)]\}^{1-\delta_{2k}} \\ \times \prod_{i=0}^1 h\{y_{ij}; n_{ij} Q_i(j)\}. \quad (42)$$

Notice that the above probability density function is a product of multinomial probability density functions and Poisson probability density functions. For this joint probability density function, the deviance $\text{Dev} = -2\{\log P[\mathbf{Y}, \mathbf{y}, \mathbf{N} | \mathbf{y}, \Theta] - \log P[\mathbf{Y}, \mathbf{y}, \mathbf{N} | \mathbf{y}, \hat{\Theta}]\}$ is

$$\text{Dev} = D_0(k) + \text{Dev}(p_1, p_2) + \sum_{j=1}^{t_N} D_j, \quad (43)$$

where

$$D_0(k) = 2 \left\{ \chi(k) - y_0 - y_0 \log \left\{ \frac{\chi(k)}{y_0} \right\} \right\}, \quad (44)$$

$$\text{Dev}(p_1, p_2) = 2 \sum_{j=1}^{t_N} \left\{ n_{0j} \log \left\{ \frac{\hat{p}_0}{(1-p_1-p_2)} \right\} \right. \\ \left. + \sum_{i=1}^2 n_{ij} \log \left\{ \frac{\hat{p}_i}{p_i} \right\} \right\}, \quad (45)$$

$$D_j = 2 \sum_{i=0}^1 \left\{ n_{ij} Q_i(j) - y_{ij} - y_{ij} \log \left\{ \frac{n_{ij} Q_i(j)}{y_{ij}} \right\} \right\} + 2(1 - \delta_{2k}) \\ \times \left\{ n_{2j} Q_2(j) - y_{2j} - y_{2j} \log \left\{ \frac{n_{2j} Q_2(j)}{y_{2j}} \right\} \right\}, \quad (46)$$

where $\hat{p}_i = ((\sum_{j=0}^{t_N} n_{ij}) / (\sum_{j=0}^{t_N} n_j))$ ($i = 0, 1, 2$).

The joint probability density function $P\{\mathbf{Y}, \mathbf{y}, \mathbf{N} | \mathbf{y}, \Theta\}$ of $(\mathbf{Y}, \mathbf{y}, \mathbf{N})$ given by (42) will be used as the kernel for the Bayesian method to estimate the unknown parameters and to predict the state variables.

4.4. Fitting of the Model to Cancer Incidence Data. To fit the model to real data, as in Tan [3–5], we let $\Delta t \sim 1$ to correspond to a fixed time interval such as 6 months in human cancer studies. (Tan et al. [4] has assumed 3 months as one-time unit while Luebeck and Moolgavkar [26] has assumed one year as one-time unit). Then, because the proliferation rate of the last stage cells is quite large, one may practically assume $P_T(s, t) = 1$ for $t - s \geq 1$. Hence, noting that $\beta_{k-1}(t) = \beta_{k-1}$ is usually very small (see [3–5]), the $Q_i(j)$ is approximated by

$$Q_i(j) \approx E\{e^{-\beta_{k-1} G_i(t_{j-1})} - e^{-\beta_{k-1} G_i(t_j)}\} = e^{-\beta_{k-1} E[G_i(t_{j-1})]} \\ - e^{-\beta_{k-1} E[G_i(t_j)]} + o(\beta_{k-1}), \quad (47)$$

where $G_i(t) = \sum_{s=t_0}^{t-1} J_{k-1}(s; i)$.

Under discrete time approximation, the $E[I_{k-1}(t; i)]$'s have been derived in the appendix. Using these results of expected numbers and using the result $\sum_{i=0}^{t-1} a^i = (a^t - 1)/(a - 1)$ for $a \neq 0$, we obtain

$$\beta_{k-1} E[G_i(t)] = \sum_{r=i}^{i+1} E[J_r(t_0; i)] \left(\prod_{u=r}^{k-1} \beta_u \right) \\ \times \sum_{v=r}^{k-1} A_{r(k-1)}(v) \left\{ \sum_{s=t_0}^{t-1} (1 + \gamma_v)^{s-t_0} \right\} \\ = \theta_{(i+1)(k-1)} \phi_{(i+1)(k-1)}(t) + \lambda_{i(k-1)} \phi_{i(k-1)}(t), \quad (48)$$

where $(\theta_{i(k-1)}, i = 1, 2, 3)$ and $(\theta_{i(k-1)}, i = 0, 1, 2)$ are defined in Section 3.3 and where the $\phi_{i(k-1)}(t)$ ($i = 0, 1, 2, 3$)'s are given by

$$\phi_{i(k-1)}(t) = \left(\prod_{u=i+1}^{k-1} \gamma_u \right) \sum_{v=i}^{k-1} A_{i(k-1)}(v) \\ \times \frac{1}{\gamma_v} \{(1 + \gamma_v)^{t-t_0} - 1\}, \quad i = 0, 1, 2, 3. \quad (49)$$

Notice that if $\gamma_0 = 0$, then $\phi_{0(k-1)}(t)$ reduces to

$$\phi_{0(k-1)}(t) = \left(\prod_{u=1}^{k-1} \gamma_u \right) \sum_{v=1}^{k-1} A_{1(k-1)}(v) \\ \times \frac{1}{\gamma_v^2} \{(1 + \gamma_v)^{t-t_0} - 1 - (t - t_0) \gamma_v\}. \quad (50)$$

Applying these results, for time homogeneous models with $\gamma_i \neq \gamma_j$ if $i \neq j$, the $Q_i(j)$'s under discrete approximation are given as follows.

- (1) If $k = 2$, then $k - 1 = 1$. Hence, $Q_2(0) = 1$, $Q_i(0) = Q_2(j) = 0$ for $(i = 0, 1, j > 0)$ and for $j > 0$,

$$\begin{aligned} Q_0(j) &\approx \left\{ e^{-\theta_{11}\phi_{11}(t_{j-1}) - \lambda_{01}\phi_{01}(t_{j-1})} \right. \\ &\quad \left. - e^{-\theta_{11}\phi_{11}(t_j) - \lambda_{01}\phi_{01}(t_j)} \right\} + o(\beta_1), \\ Q_1(j) &\approx (1 - \alpha_1) \left\{ e^{-\lambda_{11}\phi_{11}(t_{j-1})} \right. \\ &\quad \left. - e^{-\lambda_{11}\phi_{11}(t_j)} \right\} + o(\beta_1). \end{aligned} \quad (51)$$

- (2) If $k \geq 3$, then we have $\{Q_2(0) = \delta_{2(k-1)}\alpha, Q_i(0) = 0, i = 0, 1\}$ and for $j > 0$,

$$\begin{aligned} Q_0(j) &\approx \left\{ e^{-\theta_{1(k-1)}\phi_{1(k-1)}(t_{j-1}) - \lambda_{0(k-1)}\phi_{0(k-1)}(t_{j-1})} \right. \\ &\quad \left. - e^{-\theta_{1(k-1)}\phi_{1(k-1)}(t_j) - \lambda_{0(k-1)}\phi_{0(k-1)}(t_j)} \right\} \\ &\quad + o(\beta_{k-1}), \\ Q_1(j) &\approx \left\{ e^{-\theta_{2(k-1)}\phi_{2(k-1)}(t_{j-1}) - \lambda_{1(k-1)}\phi_{1(k-1)}(t_{j-1})} \right. \\ &\quad \left. - e^{-\theta_{2(k-1)}\phi_{2(k-1)}(t_j) - \lambda_{1(k-1)}\phi_{1(k-1)}(t_j)} \right\} \\ &\quad + o(\beta_{k-1}), \\ Q_2(j) &\approx \delta_{2(k-1)} (1 - \alpha) \left\{ e^{-\lambda_{22}\phi_{22}(t_{j-1})} - e^{-\lambda_{22}\phi_{22}(t_j)} \right\} \\ &\quad + (1 - \delta_{2(k-1)}) \left\{ e^{-\theta_{3(k-1)}\phi_{3(k-1)}(t_{j-1}) - \lambda_{2(k-1)}\phi_{2(k-1)}(t_{j-1})} \right. \\ &\quad \left. - e^{-\theta_{3(k-1)}\phi_{3(k-1)}(t_j) - \lambda_{2(k-1)}\phi_{2(k-1)}(t_j)} \right\} \\ &\quad + o(\beta_{k-1}). \end{aligned} \quad (52)$$

Notice that if one replaces $[1 + \gamma_v]^{t-t_0}$ by $[1 + \gamma_v]^{t-t_0} = e^{(t-t_0)\log(1+\gamma_v)} \approx e^{(t-t_0)\gamma_v}$, the above $Q_i(j)$'s from discrete time model are exactly the same as from the continuous model, respectively, as given in equations (23)–(27) under the assumption that $P_T(s, t) = 1$ for $t - s > 0$. Notice that the assumption $P_T(s, t) = 1$ for $t - s > 0$ is equivalent to assuming that the last stage cancer cells grow instantaneously into cancer tumors as soon as they are generated; see Remark 1.

5. The Fitting of the Model to Cancer Incidence and the Generalized Bayesian Inference Procedure

Given the model in Sections 2 and 3 and cancer incidence, one may use results in Section 4 to fit the model. By using this model and the distribution results in Section 4, one can readily estimate the unknown genetic parameters, predict cancer incidence, and check the validity of the model by using the generalized Bayesian inference and Gibbs sampling procedures; for more detail, see Tan [3, 22] and Tan et al. [4, 5].

The generalized Bayesian inference is based on the posterior distribution $P\{\Theta \mid \mathbf{N}, \mathbf{Y}, \mathbf{y}, \mathbf{n}\}$ of Θ given $\{\mathbf{N}, \mathbf{Y}, \mathbf{y}, \mathbf{n}\}$. This posterior distribution is derived by combining the prior

distribution $P\{\Theta\}$ of Θ with the joint probability distribution $P\{\mathbf{N}, \mathbf{Y}, \mathbf{y} \mid \mathbf{n}, \Theta\}$ given $\{\mathbf{n}, \Theta\}$ given by (42). It follows that this inference procedure would combine information from three sources: (1) previous information and experiences about the parameters in terms of the prior distribution $P\{\Theta\}$ of the parameters, (2) biological information of inherited cancer cases via genetic segregation of cancer genes in the population ($P\{\mathbf{N} \mid \mathbf{n}, p_i, i = 1, 2\}$; see Section 2), and (3) information from the expanded data ($\mathbf{Y}$) and the observed data ($\mathbf{y}$) via the statistical model from the system ($P\{\mathbf{Y}, \mathbf{y} \mid \mathbf{N}, \Theta\}$) given by (37) and (40). Because of additional information from the genetic segregation of the cancer genes, this inference procedure provides an efficient procedure to extract information of effects of genotypes of individuals at the embryo stage.

5.1. The Prior Distribution of the Parameters. For the prior distributions of Θ , because biological information has suggested some lower bounds and upper bounds for the mutation rates and for the proliferation rates, we assume

$$P(\Theta) \propto c \quad (c > 0), \quad (53)$$

where c is a positive constant if these parameters satisfy some biologically specified constraints and are equal to zero for otherwise. These biological constraints are as follows.

- (i) $0 < p_1 < 10^{-2}$, $0 < p_2 < 10^{-6}$, and $-0.01 < \gamma_i < 1$ ($i = 1, 2$).
- (ii) For β_j ($j = 0, 1, \dots, k-1$), we let $1 < \omega_0 = N(t_0)\beta_0 < 1000$ ($N \rightarrow I_1$) and $10^{-8} < \beta_i < 10^{-3}$, $i = 1, \dots, k-1$.
- (iii) For the $\lambda_{j(k-1)}$ ($j = 0, 1, 2$), we let $0 < \lambda_i < 10$, $i = 0, 1$ and $0 < \lambda_2 < 10^3$.
- (iv) For the θ_i ($i = 1, 2, 3$), we let $0 < \theta_1 < 10^{-2}$ and $0 < \theta_2 < 10^{-1}$.

We will refer to the above prior as a partially informative prior which may be considered as an extension of the traditional noninformative prior given in Box and Tiao [28].

5.2. The Posterior Distribution of the Parameters Given $\{Y, N, \mathbf{y}, \mathbf{n}\}$. Denote by $\Theta_1 = \Theta - \{p_1, p_2, \alpha\}$. From the posterior distribution $P\{\Theta \mid \mathbf{N}, \mathbf{Y}, \mathbf{y}, \mathbf{n}\}$, we obtain

$$\begin{aligned} &P\{p_i, i = 1, 2, \alpha \mid \Theta_1, \mathbf{N}, \mathbf{Y}, \mathbf{y}, \mathbf{n}\} \\ &\propto (\chi(k))^{y_0} e^{-\chi(k)} p_1^{\sum_{j=1}^{t_N} n_{1j}} p_2^{\sum_{j=1}^{t_N} n_{2j}} \\ &\quad \times (1 - p_1 - p_2)^{\sum_{j=1}^{t_N} n_{0j}}, \quad 0 < p_1, p_2, \alpha < 1, \end{aligned} \quad (54)$$

$$P\{\Theta_1 \mid (p_1, p_2, \alpha), \mathbf{N}, \mathbf{Y}, \mathbf{y}, \mathbf{n}\}$$

$$\propto \prod_{j=1}^{t_N} \prod_{i=0}^2 e^{-Q_i(j)} \{Q_i(j)\}^{y_{ij}}, \quad \Theta_1 \in \Omega,$$

where Ω is the parameter space of Θ_1 provided by the biological constraints in Section 5.1.

For computational convenience, we notice that the log of $P\{p_1, p_2, \alpha \mid \Theta_1, \mathbf{N}, \mathbf{Y}, \underline{y}, \underline{n}\}$ is proportional to the negative of $D_0(k) + \text{Dev}(p_1, p_2)$ given by (44)-(45); similarly, the log of $P\{\Theta_1 \mid (p_1, p_2, \alpha), \mathbf{N}, \mathbf{Y}, \underline{y}, \underline{n}\}$ is proportional to the negative of $\sum_{j=1}^k D_j$ given by (46).

5.3. The Multilevel Gibbs Sampling Procedure For Estimating Unknown Parameters. Given the posterior probability distributions, we will use the following multilevel Gibbs sampling procedure to derive estimates of the parameters. We notice that numerically, the Gibbs sampling procedure given below is equivalent to the EM-algorithm from the sampling theory viewpoint with Steps 1 and 2 as the *E*-Step and with Steps 3 and 4 as the *M*-Step, respectively [29]. These multilevel Gibbs sampling procedures are given by the following.

Step 1 (Generating $\mathbf{N}$ Given $(\mathbf{Y}, \underline{y}, \underline{n}, \Theta)$ (The Data-Augmentation Step 1)). Given Θ and given $\underline{n}$, use the multinomial distribution of $\{n_1, n_2\}$ given $\underline{n}_j$ in Section 3 to generate a large sample of $\mathbf{N}$. Then, by combining this large sample with $P\{\mathbf{Y}, \underline{y} \mid \mathbf{N}, \underline{n}, \Theta\}$ in (37) and (40) to select $\mathbf{N}$ through the weighted bootstrap method due to Smith and Gelfand [30]. This selected $\mathbf{N}$ is then a sample from $P\{\mathbf{N} \mid \mathbf{Y}, \underline{y}, \underline{n}, \Theta\}$ even though the latter is unknown. (For proof, see Tan [22], Chapter 3). Call the generated sample $\widehat{\mathbf{N}}$.

Step 2 (Generating $\mathbf{Y}$ Given $(\mathbf{N}, \underline{y}, \underline{n}, \Theta)$ (The Data-Augmentation Step 2)). Given $\{\underline{y}, \underline{n}, \Theta\}$ and given $\mathbf{N} = \widehat{\mathbf{N}}$ generated from Step 1, generate $\widehat{\mathbf{Y}}$ from the probability distribution $P\{\mathbf{Y} \mid \widehat{\mathbf{N}}, \underline{y}, \underline{n}, \Theta\}$ given by (36) and (38). Call the generated sample $\widehat{\mathbf{Y}}$.

Step 3 (Estimation of $\{p_i, i = 1, 2, \alpha\}$ Given $\{\Theta_1, \mathbf{N}, \mathbf{Y}, \underline{y}, \underline{n}\}$). Given $\{\underline{y}, \underline{n}, \Theta_1\}$ and given $(\mathbf{N}, \mathbf{Y}) = (\widehat{\mathbf{N}}, \widehat{\mathbf{Y}})$ from Steps 1 and 2, derive the posterior mode of $\{p_i, i = 1, 2, \alpha\}$ by maximizing the conditional posterior distribution $P\{p_i, i = 1, 2, \alpha \mid \Theta_1, \widehat{\mathbf{N}}, \widehat{\mathbf{Y}}, \underline{y}, \underline{n}\}$. Under the partially informative prior, this is equivalent to maximize the negative of the deviance $D_0(k) + \text{Dev}(p_1, p_2)$ given by (44)-(45) in Section 4.3 under the constraints given in Section 5.1. Denote this generated mode by $\{\widehat{p}_i, i = 1, 2, \widehat{\alpha}\}$.

Step 4 (Estimation of Θ_1 Given $(p_i, i = 1, 2, \alpha, \mathbf{N}, \mathbf{Y}, \underline{y}, \underline{n})$). Given $(\underline{y}, \underline{n})$ and given $(\mathbf{N}, \mathbf{Y}, p_i, i = 1, 2, \alpha) = (\widehat{\mathbf{N}}, \widehat{\mathbf{Y}}, \widehat{p}_i, i = 1, 2, \widehat{\alpha})$ from Steps 1-3, derive the posterior mode of Θ_1 by maximizing the conditional posterior distribution $P\{\Theta_1 \mid \widehat{p}_i, i = 1, 2, \widehat{\alpha}, \widehat{\mathbf{N}}, \widehat{\mathbf{Y}}, \underline{y}, \underline{n}\}$. Under the partially informative prior, this is equivalent to maximize the negative of the deviance $\sum_{j=1}^k D_j$ in (46) under the constraints. Denote the generated mode as $\widehat{\Theta}_1$.

Step 5 (Recycling Step). With $\{(\mathbf{N}, \mathbf{Y}, p_i, i = 1, 2, \alpha, \Theta_1) = (\widehat{\mathbf{N}}, \widehat{\mathbf{Y}}, \widehat{p}_i, i = 1, 2, \widehat{\alpha}, \widehat{\Theta}_1)\}$ given above, go back to Step 1 and continue until convergence.

The proof of convergence of the above steps can be derived by using procedure given in Tan ([22], Chapter 3). At convergence, the $\widehat{\Theta} = \{\widehat{p}_i, i = 1, 2, \widehat{\alpha}, \widehat{\Theta}_1\}$ are the generated values from the posterior distribution of Θ given $\{\underline{y}, \underline{n}\}$ independently of $(\mathbf{N}, \mathbf{Y})$ (for proof, see Tan [22], Chapter 3). Repeat the above procedures once then generate a random sample of Θ from the posterior distribution of Θ given $\{\underline{y}, \underline{n}\}$; then one uses the sample mean as the estimates of (Θ) and use the sample variances and covariances as estimates of the variances and covariances of these estimates.

6. A New Multistage Stochastic Model for Adult Eye Cancer (Uveal Melanoma)—An Example

The human eye cancers consist of pediatric eye cancers and adult eye cancers. The most common pediatric eye cancer is the retinoblastoma which develops from the retinal pigment epithelium cells underlying the retina that do not form melanoma. The most common adult eye cancers are the uveal melanomas involving the iris, the ciliary body, and the choroid (collectively referred to as the uveal). These cancers develop from melanocytes (pigment cells) which reside within the uveal giving color to the eye. In Tan and Zhou [9] we have developed a modified two-stage model for retinoblastoma. Based on results from molecular biology (see Landreville et al. [14], Mensink et al. [15], and Loercher and Harbour [31]), Landreville et al. [14] have proposed a three stage model for uveal melanoma as given in Figure 2. As an example of applications of this paper, in this section we will apply this model of uveal melanoma to the NCI/NIH eye cancer data from the SEER project. We notice that the same methods can be applied to model other human cancers as well, but this will be our future research.

Given in Table 1 are the numbers of people at risk and the eye cancer cases in the age groups together with the predicted cases from the models. These data give cancer incidence at birth and incidence for 85 age groups ($k = 85$) with each group spanning over a 1-year period except the last age group (≥ 85 years old). For human eye cancer, because the incidence at birth and for age groups from 1 to 10 years old is basically generated by the pediatric eye cancer-retinoblastoma (see [9]), to account for inherited cancer cases of uveal melanomas, the incidence for age 0 (birth) and for age periods from 1 to 10 years old in Table 1 for uveal melanoma is derived by subtracting incidence of retinoblastoma from SEER data (see Tan and Zhou [9]).

To fit the data, we let one-time unit be 6 months after birth and let $t_0 = 1$. To compare different models and to assess different assumptions, we will consider the following 2-3-stage mixture models: (1) the complete 3-stage mixture model (Model-F) in which no assumptions are made on the parameters. (2) the Type-1-3-stage mixture model in which we assume that $\gamma_1 = 0$ and that normal people and J_1 at the embryo stage will remain normal people and J_1 people, respectively, at birth (Model-1). For comparison purposes, we also fit a 2-stage model as defined in Tan and Zhou [9]. We will apply the methods in Section 6 to fit these models to the SEER data given in Table 1.

TABLE 2: The log-likelihood, AIC, and BIC of the fitted models.

Models	Log-likelihood	AIC	BIC
Three-stage			
Model-F	-1312.02	2648.04	2677.35
Model-1	-1295.76	2609.53	2631.51
Two-stage	-4392.421	8796.843	8811.569

FIGURE 5: Curve fitting of SEER data by the Model-F, Model-1, and the two-stage model.

Given in Table 2 are the natural logs of the likelihood functions, the AIC (Akaike Information Criterion) and the BIC (Bayesian Information Criterion) for these models. Given in Table 3 are the estimates of parameters in the 3-stage models. Given in Figure 5 are the plots of predicted cancer cases from the 3-stage mixture models (Model-F and Model-1) and the 2-stage model. For comparison purposes, in Table 1, we also provide numbers of predicted cancer cases from the 3-stage mixture models and the 2-stage model together with the observed cancer cases over time from SEER. From these results, we have made the following observations.

- (a) As shown by results in Table 1 and Figure 5, it appeared that both Model-F and Model-1 fitted the SEER data well, although Model-1 fitted the data slightly better from values of AIC and BIC. The Chi-square test statistics $\chi^2 = \sum_{i=0}^{85} ((y_i - \hat{y}_i)^2 / \hat{y}_i)$ for Model-F and Model-1 are given by 88.43, and 94.48 respectively, giving a P -value of 0.12 ($df = 86 - 12 = 74$) for Model-F and a P value of 0.11 ($df = 86 - 7 = 79$) for Model-1. On the other hand, the 2-stage model fitted the data very poorly; the Chi-statistic value for the 2-stage model is 2747.69 giving a P -value less than 10^{-3} . The AIC (Akaike Information Criteria) and BIC (Bayesian Information Criteria) values of Model-1 are given by (AIC = 2609.53, BIC = 2631.51) which are slightly smaller than those of Model-F, respectively; however, the AIC and BIC values (8796.84, 8811.57) of the two-stage model are considerably greater than those of the 3-stage models, respectively. These results suggest that uveal melanoma may best be described by a 3-stage model with inherited component and that

one may practically assume $\gamma_1 = 0$ and that normal people and J_1 people at the embryo stage will remain normal people and J_1 people, respectively, at birth.

- (b) From Table 3, it is observed that the estimate of γ_1 is close to zero (the estimate is of order 10^{-5}) indicating that the phenotype of J_1 is almost identical to that of $N = J_0$ further confirming that the staging-limiting genes are basically tumor suppressor genes and that there is no haploinsufficiency for these tumor suppressor genes. On the other hand, the estimate of γ_2 is of order 10^{-2} which is about 10^3 times greater than those of cells with genotype J_1 .
- (c) From Table 3, the estimates $\hat{\lambda}_j$ ($j = 0, 1, 2$) of λ_j are of order $\{10^{-1}, 10^{-1}, 10^2 \sim 10^3\}$, respectively. Because $\{\lambda_i = E[J_i(t_0; i)]\beta_i \prod_{u=i+1}^2 (\beta_u/\gamma_u), i = 0, 1, 2\}$, assuming some values of $\{E[J_i(t_0; i)], i = 0, 1, 2\}$ from some biological observations, one can have some rough ideas about the magnitude of β_j ($j = 0, 1, 2$). For example, if we follow Potten et al. [32] to assume $(E[N(t_0)] = E[J_i(t_0; i)] \sim 10^8, i = 0, 1, 2)$, then $\beta_j \approx 10^{-6} \sim 10^{-5}$.
- (d) From Table 3, the estimates of p_1 and p_2 from the SEER data are of orders $10^{-4} \sim 10^{-3}$ and $10^{-7} \sim 10^{-6}$, respectively. This indicates that in the US population, the frequency of the staging limiting cancer gene for uveal melanoma is approximately around 10^{-3} . Table 3 also showed that the estimate of α was 0.8411, indicating that most individuals with genotype J_2 would develop cancer at birth. This may help to explain why there are observed cancer incidences at birth for uveal melanoma in the SEER data even though the estimate of the frequency p_2 is of order $10^{-7} \sim 10^{-6}$.

7. Discussion and Conclusions

To account for inherited cancer cases, in this paper we have developed some general multistage models involving hereditary cancer cases. For human cancer incidence, these models are basically generalized mixture models. In these mixture models, the mixing probability is a multinomial distribution to account for genetic segregation of the staging-limiting tumor suppressor genes. This mixture model allows us to estimate for the first time the frequency of the staging-limiting tumor suppressor gene in human populations. As an example of applications, in this paper we have developed a general 3-stage stochastic multistage model of carcinogenesis for adult human eye cancer. To account for inherited cancer cases in the stochastic model of human eye cancer, we have also developed a generalized mixture model for uveal melanoma in human beings.

For using the proposed models to fit the cancer incidence data, in this paper we have developed a generalized Bayesian inference procedure to estimate the unknown parameters and to predict cancer cases. This inference procedure is advantageous over the classical sampling theory inference (i.e., maximum likelihood method) because the procedure combines

TABLE 3: Estimates of parameters for the 3-stage stochastic models.

Parameters		λ_0	λ_1	λ_2	γ_1	γ_2	p_1	p_2	α	θ_1	θ_2
Model-F	Estimates	2.88E-01	4.81E-01	6.55E+02	2.71E-05	3.37E-02	9.95E-04	9.68E-07	8.41E-01	3.29E-04	3.41E-01
	St.D	1.21E-01	4.65E-02	1.12E+02	8.86E-06	1.92E-04	1.75E-06	6.48E-09	4.19E-02	3.54E-05	1.07E-02
	95% CL-Lower	5.03E-02	3.89E-01	4.34E+02	5.34E-06	3.33E-02	9.91E-04	9.55E-07	7.59E-01	2.60E-04	3.20E-01
	95% CL-Upper	5.25E-01	5.72E-01	8.75E+02	4.01E-05	3.41E-02	9.98E-04	9.81E-07	9.23E-01	3.98E-04	3.62E-01
Model-1	Estimates	6.55E-02	3.72E-01	1.98E+02	NA	3.49E-02	9.98E-04	1.00E-06	8.07E-01	NA	NA
	St.D	2.54E-02	4.63E-02	3.38E+01	NA	3.58E-03	6.27E-05	4.53E-08	7.06E-02	NA	NA
	95% CL-Lower	1.57E-02	2.82E-01	1.32E+02	NA	2.79E-02	8.75E-04	9.11E-07	6.68E-01	NA	NA
	95% CL-Upper	1.15E-01	4.63E-01	2.65E+02	NA	4.19E-02	1.12E0-03	1.09E-06	9.45E-01	NA	NA

Note: NA: assumed nonexistence.

information from three sources: (1) previous information and experiences about the parameters in terms of the prior distribution $P\{\Theta\}$ of the parameters, (2) biological information of inherited cancer cases via the genetic segregation of staging-limiting tumor suppressor genes in the population, and (3) information from the expanded data ($\mathbf{Y}$) and the observed data ($\mathbf{y}$) via the statistical model from the system ($P\{\mathbf{Y}, \mathbf{y} | \mathbf{N}, \Theta\}$) given by (37) and (40).

To illustrate the usefulness and applications of our models and methods, we have applied our models and methods to the eye cancer SEER data of NCI/NIH. Our analysis clearly showed that the proposed 3-stage model with inherited cancer cases fitted the data nicely (see Table 2 and Figure 5); on the other hand, the classical 2-stage model cannot fit the data at all (see Table 2 and Figure 5). These results clearly have confirmed results from molecular biology that the human eye cancer is derived by a 3-stage model with inherited cancer component. Notice, however, our 3-stage multistage model is more general than the classical 3-stage model as described in Little [1], Tan [2], and Zheng [7] in that we postulate that cancer tumors develop from primary I_3 cells by clonal expansion (see Yang and Chen [11]). (Note that the stochastic multistage models in the literature assume that cancer tumors develop from last stage cells immediately as soon as they are generated, ignoring completely cancer progression; see Remark 1). As a matter of fact, we had assumed $\Delta t = 1$ for a period of three months and found that the 3-stage models then did not fit the SEER data, clearly indicating that $P_t(s, t) < 1$ over a period of three months.

Applying our models and methods to the SEER data of human eye cancer, we have derived for the first time some useful pieces of information. Specifically, we mention (1) for the first time that we have estimated the frequency of the staging-limiting tumor suppressor gene in the US population ($\hat{p}_1 \sim .9948 \times 10^{-3}$). (2) With the estimate of α as $\hat{\alpha} = 0.8411$, the predicted number of uveal melanoma at birth is $\hat{y}_0 = n_0 \hat{\alpha}_1 \hat{p}^2 = 36$ by 3-stage models with inherited cancer component (Model-F and Model-1). (The observed number of eye cancer at birth is 34). (3) The estimate of the proliferation rate (γ_1) of J_1 cells using Model-F is $\hat{\gamma}_1 = 2.271 \times 10^{-5} \sim 0$. (The estimate is 8.603×10^{-5} using Model-1). This confirms that the stage-1 limiting gene is a tumor suppressor gene, and unlike the p53 gene in chromosome 17p (see [33]), there is little or no haploid insufficiency for this gene in cells with genotype J_1 .

Using models and methods of this paper, one can easily predict future cancer cases for human eye cancer. Thus, by comparing results from different populations, our models and methods can be used to assess cancer prevention and control procedures. This will be our future research topics; we will not go any further here.

Appendix

The Expected Numbers of State Variables under Discrete Time Approximation

Under discrete time, the stochastic differential equations of state variables reduce to the following stochastic difference

equations of state variables, respectively:

$$\begin{aligned} J_i(t+1; i) &= J_i(t; i) [1 + \gamma_i(t)] \\ &\quad + e_i(t+1; i), \quad i = 0, 1, \dots, \text{Min}(k-1, 2), \\ J_j(t+1; u) &= J_j(t; u) [1 + \gamma_j(t)] + J_{j-1}(t; u) \beta_{j-1}(t) \\ &\quad + e_j(t+1; u), \quad 0 \leq u < j \leq k-1, \end{aligned} \quad (\text{A.1})$$

where $e_i(t+1; i) = [B_i(t; i) - J_i(t; i)b_i(t)] - [D_i(t; i) - J_i(t; i)d_i(t)]$ and $e_j(t+1; i) = [B_j(t; i) - J_j(t; i)b_j(t)] - [D_j(t; i) - J_j(t; i)d_j(t)] + [M_{j-1}(t; i) - J_{j-1}(t; i)\beta_{j-1}(t)]$ for $j > i$.

The initial conditions at birth (t_0) for the above stochastic difference equations are $J_j(t_0; i) > 0$ if ($j = i, i+1$) and $J_j(t_0; i) = 0$ if $j > i+1$. The solution of the above difference equations under these initial conditions is given, respectively, by

$$\begin{aligned} J_i(t; i) &= J_i(t_0; i) \prod_{s=t_0}^{t-1} (1 + \gamma_i(s)) \\ &\quad + \sum_{s=t_0+1}^t e_i(s; i) \prod_{u=s}^{t-1} (1 + \gamma_i(u)), \\ &\quad i = 0, 1, \dots, \text{Min}(k-1, 2), \\ J_j(t; i) &= J_j(t_0; i) \prod_{s=t_0}^{t-1} (1 + \gamma_j(s)) \\ &\quad + \sum_{s=t_0}^{t-1} J_{j-1}(s; i) \beta_{j-1}(s) \prod_{u=s+1}^{t-1} (1 + \gamma_j(u)) \\ &\quad + \sum_{s=t_0+1}^t e_j(s; i) \prod_{u=s}^{t-1} (1 + \gamma_j(u)), \\ &\quad 0 \leq i < j \leq k-1. \end{aligned} \quad (\text{A.2})$$

If the model is time homogeneous so that $\{\beta_j(t) = \beta_j, \gamma_j(t) = \gamma_j \ (j = i, \dots, k-1)\}$ and if $\gamma_i \neq \gamma_j$ if $i \neq j$, then the above solutions under the initial conditions ($J_j(t_0; i) = 0, j > i+1$) reduce to

$$\begin{aligned} J_i(t; i) &= J_i(t_0; i) (1 + \gamma_i)^{t-t_0} \\ &\quad + \eta_i^{(0)}(t; i), \quad i = 0, 1, \dots, \text{Min}(k-1, 2), \\ J_j(t; i) &= J_j(t_0; i) (1 + \gamma_j)^{t-t_0} \\ &\quad + \sum_{u=i}^{j-1} J_u(t_0; i) \left\{ \prod_{v=u}^{j-1} \beta_v \right\} \sum_{r=u}^j A_{uj}(r) (1 + \gamma_r)^{t-t_0} \\ &\quad + \eta_j^{(0)}(t; i) + \sum_{u=1}^{j-i} \left\{ \prod_{r=j-u}^{j-1} \beta_r \right\} \eta_j^{(u)}(t; i) \end{aligned}$$

$$\begin{aligned}
&= \sum_{u=i}^{i+1} J_u(t_0; i) \left\{ \prod_{v=u}^{j-1} \beta_v \right\} \sum_{r=u}^j A_{ur}(r) (1 + \gamma_r)^{t-t_0} \\
&\quad + \eta_j^{(0)}(t; i) + \sum_{u=1}^{j-i} \left\{ \prod_{r=j-u}^{j-1} \beta_r \right\} \eta_j^{(u)}(t; i), \\
&\quad 0 \leq i < j \leq k-1,
\end{aligned} \tag{A.3}$$

where $\eta_j^{(0)}(t; i) = \sum_{s=t_0+1}^t e_j(s; i) (1 + \gamma_i)^{t-s}$, $\eta_j^{(u)}(t; i) = \sum_{s=t_0+1}^t e_{j-u}(s; i) \{ \sum_{v=u}^j A_{uv}(v) (1 + \gamma_v)^{t-s} \}$ ($u = 1, \dots, j-i$).

Thus, if the model is time homogeneous and if $\gamma_i \neq \gamma_j$ if $i \neq j$, the $E[J_{k-1}(t; i)]$'s in discrete time models under the initial conditions ($J_j(t_0; i) = 0$, $j > i + 1$) are given, respectively, by

$$\begin{aligned}
E[J_{k-1}(t; i)] &= \sum_{u=i}^{i+1} E[J_u(t_0; i)] \left(\prod_{v=u}^{k-2} \beta_v \right) \\
&\quad \times \sum_{r=u}^{k-1} A_{u(k-1)}(r) (1 + \gamma_r)^{t-t_0}, \\
&\quad i = 0, 1, \dots, \text{Min}(k-1, 2).
\end{aligned} \tag{A.4}$$

References

- [1] M. P. Little, "Cancer models, ionization and genomic instability: a review," in *Handbook of Cancer Models with Applications*, W. Y. Tan and L. Hanin, Eds., chapter 5, World Scientific, River Edge, NJ, USA, 2008.
- [2] W. Y. Tan, *Stochastic Models of Carcinogenesis*, Marcel Dekker, New York, NY, USA, 1991.
- [3] W. Y. Tan, "Stochastic multi-stage models of carcinogenesis as hidden Markov models: a new approach," *International Journal of Systems and Synthetic Biology*, vol. 1, pp. 313–337, 2010.
- [4] W. Y. Tan, L. J. Zhang, and C. W. Chen, "Stochastic modeling of carcinogenesis: state space models and estimation of parameters," *Discrete and Continuous Dynamical Systems B*, vol. 4, no. 1, pp. 297–322, 2004.
- [5] W. Y. Tan, C. W. Chen, and L. J. Zhang, "Cancer biology, cancer models and stochastic mathematical analysis of carcinogenesis," in *Handbook of Cancer Models and Applications*, W. Y. Tan and L. Hanin, Eds., chapter 3, World Scientific, River Edge, NJ, USA, 2008.
- [6] R. A. Weinberg, *The Biology of Human Cancer*, Garland Sciences, Taylor and Frances, New York, NY, USA, 2007.
- [7] Q. Zheng, "Stochastic multistage cancer models: a fresh look at an old approach," in *Handbook of Cancer Models and Applications*, W. Y. Tan and L. Hanin, Eds., chapter 2, World Scientific, River Edge, NJ, USA, 2008.
- [8] W. Y. Tan, L. J. Zhang, W. Chen, and J. M. Zhu, "A stochastic model of human colon cancer involving multiple pathways," in *Handbook of Cancer Models with Applications*, W. Y. Tan and L. Hanin, Eds., chapter 11, World Scientific, River Edge, NJ, USA, 2008.
- [9] W. Y. Tan and H. Zhou, "A new stochastic model of retinoblastoma involving both hereditary and non-hereditary cancer cases," *Journal of Carcinogenesis and Mutagenesis*, vol. 2, no. 2, article 117, 2011.
- [10] X. W. Yan, *Stochastic and State Space Models of carcinogenesis Under Complex Situation [Ph.D. thesis]*, Department of Mathematical Sciences, University of Memphis, Memphis, Tenn, USA.
- [11] G. L. Yang and C. W. Chen, "A stochastic two-stage carcinogenesis model: a new approach to computing the probability of observing tumor in animal bioassays," *Mathematical Biosciences*, vol. 104, no. 2, pp. 247–258, 1991.
- [12] H. Osada and T. Takahashi, "Genetic alterations of multiple tumor suppressors and oncogenes in the carcinogenesis and progression of lung cancer," *Oncogene*, vol. 21, no. 48, pp. 7421–7434, 2002.
- [13] I. I. Wistuba, L. Mao, and A. F. Gazdar, "Smoking molecular damage in bronchial epithelium," *Oncogene*, vol. 21, no. 48, pp. 7298–7306, 2002.
- [14] S. Landreville, O. A. Agapova, and J. W. Hartbour, "Emerging insights into the molecular pathogenesis of uveal melanoma," *Future Oncology*, vol. 4, no. 5, pp. 629–636, 2008.
- [15] H. W. Mensink, D. Paridaens, and A. De Klein, "Genetics of uveal melanoma," *Expert Review of Ophthalmology*, vol. 4, no. 6, pp. 607–616, 2009.
- [16] W. Y. Tan and X. W. Yan, "A new stochastic and state space model of human colon cancer incorporating multiple pathways," *Biology Direct*, vol. 5, article 26, 2010.
- [17] L. B. Klebanov, S. T. Rachev, and A. Y. Yakovlev, "A stochastic model of radiation carcinogenesis: latent time distributions and their properties," *Mathematical Biosciences*, vol. 113, no. 1, pp. 51–75, 1993.
- [18] A. Y. Yakovlev and A. D. Tsodikov, *Stochastic Models of Tumor Latency and Their Biostatistical Applications*, World Scientific, River Edge, NJ, USA, 1996.
- [19] H. Fakir, W. Y. Tan, L. Hlatky, P. Hahnfeldt, and R. K. Sachs, "Stochastic population dynamic effects for lung cancer progression," *Radiation Research*, vol. 172, no. 3, pp. 383–393, 2009.
- [20] A. G. Knudson, "Mutation and cancer: statistical study of retinoblastoma," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 68, no. 4, pp. 820–823, 1971.
- [21] J. F. Crow and M. Kimura, *An Introduction to Population Genetics Theory*, Harper and Row, New York, NY, USA, 1970.
- [22] W. Y. Tan, *Stochastic Models With Applications to Genetics, Cancers, AIDS and Other Biomedical Systems*, World Scientific, River Edge, NJ, USA, 2002.
- [23] W. Y. Tan, C. W. Chen, and L. J. Zhang, "Cancer risk assessment by state space models," in *Handbook of Cancer Models and Applications*, W. Y. Tan and L. Hanin, Eds., Chapter 13, World Scientific, River Edge, NJ, USA, 2008.
- [24] W. Y. Tan and C. W. Chen, "Stochastic modeling of carcinogenesis: some new insights," *Mathematical and Computer Modelling*, vol. 28, no. 11, pp. 49–71, 1998.
- [25] W. Y. Tan and C. W. Chen, "Cancer stochastic models," in *Encyclopedia of Statistical Sciences, Revised edition*, John Wiley and Sons, New York, NY, USA, 2005.
- [26] E. G. Luebeck and S. H. Moolgavkar, "Multistage carcinogenesis and the incidence of colorectal cancer," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 23, pp. 15095–15100, 2002.
- [27] R. Durrett, J. Mayberry, S. Moseley, and D. Schmidt, "Probability models for cancer development and progression," Google Search, Google 2012.
- [28] G. E. P. Box and G. C. Tiao, *Bayesian Inference in Statistical Analysis*, Addison-Wesley, Reading, Mass, USA, 1973.

- [29] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm (with discussion)," *Journal of the Royal Statistical Society B*, vol. 39, pp. 1–38, 1977.
- [30] A. F. M. Smith and A. E. Gelfand, "Bayesian statistics without tears: a samplingresampling perspective," *American Statistician*, vol. 46, pp. 84–88, 1992.
- [31] A. E. Loercher and J. W. Harbour, "Molecular genetics of uveal melanoma," *Current Eye Research*, vol. 27, no. 2, pp. 69–74, 2003.
- [32] C. S. Potten, C. Booth, and D. Hargreaves, "The small intestine as a model for evaluating adult tissue stem cell drug targets," *Cell Proliferation*, vol. 36, no. 3, pp. 115–129, 2003.
- [33] C. J. Lynch and J. Milner, "Loss of one p53 allele results in four-fold reduction of p53 mRNA and protein: a basis for p53 haplo-insufficiency," *Oncogene*, vol. 25, no. 24, pp. 3463–3470, 2006.