

# Development of a Biologically-Based Controlled Growth and Differentiation Model for Developmental Toxicology

Shree Y. Whitaker<sup>1</sup>, Hien T. Tran<sup>1</sup> and Christopher J. Portier<sup>2</sup>

## Abstract

The biologically-based dose-response model for developmental toxicology developed by Leroux *et al.* (1996) is extended. The original model had two basic states; precursor cells and differentiated cells with both states subject to a linear birth-death process. In this paper, a mathematical model, which is biologically and statistically based, is developed with a highly controlled birth and death process for precursor cells. This model limits the number of replications allowed in the development of a tissue or organ and more closely reflects the presence of a true stem cell population. The mathematical formulation of the Leroux *et al.* (1996) model was derived from a partial differential equation for the generating function that limits further expansion into more realistic models of mammalian development. The same formulae for the probability of a defect (a system of ordinary differential equations) can be derived through the Kolmogorov forward equations due to the nature of this Markov process. This modified approach is easily amenable to the expansion of more complicated models of the developmental process such as the one presented here. Comparisons between the Leroux *et al.* (1996) model and the controlled growth and differentiation (CGD) model as developed in this paper are also discussed.

**KEY WORDS:** teratology; multistate process; cellular kinetics

## 1. Introduction

There has been considerable research into the mechanistic basis through which environmental exposures can initiate and promote disease processes. Much of this research has focused on the molecular and biochemical basis describing the interaction of chemical and physical agents with healthy tissue. Most environmental health risk assessments are focused on the rates of morbidity and mortality in human populations following an environmental exposure. The linkage between basic biology and disease incidence in environmental health is best described using a tool which is focused on the incidence of disease and which can fully utilize the emerging science <sup>[1,2]</sup>. Disease incidence is generally described by counting events (e.g. disease prevalence in a population) or by early, functional failure of an entire organ system (e.g. disease incidence per year). Data endpoints such as these require a different mathematical treatment than the mathematical treatment applied to absorption, distribution and metabolism data endpoints <sup>[3]</sup>. While the mechanistic basis for

---

<sup>1</sup> Center for Research in Scientific Computation, Department of Mathematics, Box 8205, North Carolina State University, Raleigh, North Carolina 27695-8205.

<sup>2</sup> Laboratory of Computational Biology for Risk Analysis, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina 27709.

understanding environmentally induced disease has progressed rapidly, biologically-based mechanistic models of morbidity and mortality lag far behind. This gap in development is partially due to the difficulties in the mathematical treatment of these endpoints and partially due to gaps in our understanding of how they occur.

The scientific database for a mechanistic understanding of toxic effects following chemical exposure is probably greatest for the area of carcinogenesis. Several researchers have published work in the area of multistage models for carcinogenesis. The Armitage-Doll model<sup>[4]</sup> is considered the grandfather of the multistage cancer models. This model has been widely used for the analysis of epidemiological data and for cancer risk assessment. From the statistical point of view, this model provides a broad class of hazard functions for the analysis of data. Armitage and Doll<sup>[5]</sup> extended this model to use a deterministic linear birth-death process on the intermediate state in a two-stage model. Several other researchers<sup>[6-23]</sup> have also contributed significant progress to this field. The models developed use biologically based information and basic stochastic processes (generally interconnected birth-death processes with immigration and emigration) to reproduce the behavior of cells as they progress through the stages of cancer. Cancer is viewed as a multi-step process in which cells move from a controlled and systematic state of growth into a state of uncontrollable and chaotic growth<sup>[24, 25]</sup>. The basic assumption – that cancer is a disease of single cells rather than entire organ systems - on which these models were predicated, makes the mathematical modeling of carcinogenesis feasible.

To model a disease process with as much precision as possible, it is imperative to first understand and create a solid foundation for the basic developmental process of a tissue or organ. The linkage is critical in that many of the early markers for carcinogenesis relate back to genes and proteins expressed predominantly during development and cellular replication<sup>[26, 27]</sup>. Several authors have attempted to create models for the analysis of data from developmental toxicity. Much of this effort has focused on the creation of statistical likelihoods for handling the nested variance observed from developmental toxicology studies or for the longitudinal nature of the data<sup>[28-35]</sup>. Others have developed statistical models for the probability of a defect focusing on the role of some other form of toxicity (such as maternal toxicity) in the forms of the equations. Leroux *et al.*<sup>[36]</sup> developed the first biologically-based, stochastic model of developmental toxicity based on an interconnected birth-death process and using the first and second moments for the number of cells in a second stage of the model as an indicator of the probability of toxicity. The general view of development as modeled by Leroux *et al.*<sup>[36]</sup> is a straightforward approach. However, their use of an uncontrolled, linear birth-death process for cellular growth is both biologically unreasonable (for small numbers of cells) and mathematically uncontrollable allowing for no clear numerical constraints on the size of the resulting organ.

The model proposed in this paper, which is biologically and statistically based, is an extension of the Leroux *et al.*<sup>[36]</sup> model with a highly controlled birth and death process for the precursor cells. This formulation limits the number of replications allowed in the development of a tissue or organ and more closely reflects the presence of true stem cell population. The paper is organized as follows. In Section 2.1, we will re-examine the Leroux

*et al.*<sup>[36]</sup> model. We will show that the same linear first-order system of ordinary differential equations describing the expectations, squared expectations and the expected cross product of cells in a two-stage model of developmental toxicology can be derived using, alternatively, the Kolmogorov forward equations. This model is then extended in Section 2.2 to allow greater emphasis on the control of growth and differentiation (CGD). We emphasize that the model proposed by Leroux *et al.*<sup>[36]</sup> is not a sub-model of the CGD model since a linear birth process is not assumed. Finally, in Section 3, a comparison between the Leroux *et al.*<sup>[36]</sup> model and the CGD model is made on the prediction of the effects of methylmercury on brain development in rats and Section 4 contains our concluding remarks.

## **2. Materials and Methods**

Development consists of a series of states during which cells first replicate, then differentiate with some loss of cells and then separate into groups of cells for which a similar function is needed. This cycle is repeated numerous times; some organs and organ systems complete development early while others continue to develop even after birth of the mammal. During the developmental process of a tissue or organ, a cell eventually moves from an uncommitted state to a committed state. This simple characterization of development implies that the rates that drive the formation of a mammal can be classified into one of three categories: birth, death and transformation. This crude, cellular level approach to development can not fully accommodate the subtle biochemical and molecular events, such as expression of key growth factors and hormones, which are really governing overall development, but serves as a phenomenological description of the cellular growth of an organ. These rates (birth, death and transformation) are likely to be functions of these subtler biochemical and hormonal signals.

### **2.1 The Leroux *et al.*<sup>[36]</sup> Model**

In this section, we will review the mathematical model developed by Leroux *et al.*<sup>[36]</sup> for the developmental process for a particular tissue or organ. More specifically, we will show that the same mathematical model, which is a system of ordinary differential equations describing the expectations, squared expectations and expected cross product of cells, can also be derived using the Kolmogorov forward equations. This modified approach is easily amenable to the expansion of more complicated models of the developmental process including the development of a biologically-based controlled growth and differentiation model for developmental toxicology (see Section 2.2).

A schematic diagram of the Leroux *et al.* model is shown in Figure 1. In this model, the basic developmental process is divided into two main subpopulations. Type-*X* cells represent cells before commitment to differentiation whereas type-*Y* cells represent cells that have undergone a transformation and are committed to a phenotype. In addition, the following specific assumptions were made for the mathematical development of the model:

- (i) cells act independently of one another;

- (ii) the probability of more than one event in a single cell in any small time interval  $\Delta t$  is proportional to  $o(\Delta t)$  ;
- (iii) transformation is an irreversible process;
- (iv) a cell in a particular population can only replicate to produce cells of the same population; and
- (v) a malformation results when the number of committed ( $Y(t)$ ) cells is less than a critical number ( $Y_c$ ) at a specified time,  $t_c$  (i.e.,  $Y(t_c) < Y_c$ ).

The type-  $X$  cell has the option to replicate, die or transform while the type- $Y$  cell may either replicate or die. It is noted that the result of a cell replicating is two daughter cells that are a part of the original population (i.e., the daughter cells are members of the same population as the parent cell). The result of a cell dying is either removal of the cell from the population of properly functioning cells or actual death of a cell. A transformation is the product of a type- $X$  cell moving from the population of uncommitted cells to the population of cells committed to differentiation.

The developmental rates are denoted by  $\mathbf{b}_1(t)$  (birth rate in  $X$  population),  $\mathbf{d}_1(t)$  (death rate in  $X$  population),  $\mathbf{b}_2(t)$  (birth rate in  $Y$  population),  $\mathbf{d}_2(t)$  (death rate in  $Y$  population) and  $\mathbf{m}(t)$  (transformation rate from  $X$  population to  $Y$  population). The meaning of the rate  $\mathbf{b}_1(t)$ , for example, is that for the time interval  $[t, t + \Delta t)$  where  $\Delta t$  is small, the probability of replication for a type-  $X$  cell is  $\mathbf{b}_1(t)\Delta t + o(\Delta t)$  where

$$\lim_{\Delta t \rightarrow 0} \frac{o(\Delta t)}{\Delta t} = 0. \quad (1.1)$$

**Table 1:** Single Cell Event Probabilities in time  $[t, t + \Delta t)$  for the Leroux et al. <sup>[36]</sup> Model

Event	$t$	$t + \Delta t$	Single cell transition probability
<i>X birth</i>	$x-1, y$	$x, y$	$\mathbf{b}_1(t) \Delta t P_{x_0, y_0, 0}(x-1, y, t)$
<i>X death</i>	$x+1, y$	$x, y$	$\mathbf{d}_1(t) \Delta t P_{x_0, y_0, 0}(x+1, y, t)$
<i>Y birth</i>	$x, y-1$	$x, y$	$\mathbf{b}_2(t) \Delta t P_{x_0, y_0, 0}(x, y-1, t)$
<i>Y death</i>	$x, y+1$	$x, y$	$\mathbf{d}_2(t) \Delta t P_{x_0, y_0, 0}(x, y+1, t)$
<i>transformation</i>	$x+1, y-1$	$x, y$	$\mathbf{m}(t) \Delta t P_{x_0, y_0, 0}(x+1, y-1, t)$
<i>no change</i>	$x, y$	$x, y$	$\{1 - \mathbf{b}_1(t)\Delta t - \mathbf{d}_1(t)\Delta t - \mathbf{m}(t)\Delta t$ $- \mathbf{b}_2(t)\Delta t - \mathbf{d}_2(t)\Delta t\} P_{x_0, y_0, 0}(x, y, t)$

The forward Kolmogorov equation for the transition probability can be expressed generally as

$$P_{\tilde{x}, \tilde{y}, t}(x, y, t + \Delta t) = \Pr[X(t + \Delta t) = x, Y(t + \Delta t) = y \mid X(t) = \tilde{x}, Y(t) = \tilde{y}] \quad (1.2)$$

Since only certain events, which are summarized in Table 1, can occur in a small time interval  $\Delta t$ , the transition probability  $P_{x_0, y_0, 0}$  is given by

$$\begin{aligned}
P_{x_0, y_0, 0}(x, y, t + \Delta t) &= (x-1)\mathbf{b}_1(t)\Delta t P_{x_0, y_0, 0}(x-1, y, t) + (x+1)\mathbf{d}_1(t)\Delta t P_{x_0, y_0, 0}(x+1, y, t) \\
&+ (x+1)\mathbf{m}(t)\Delta t P_{x_0, y_0, 0}(x+1, y-1, t) \\
&+ (y-1)\mathbf{b}_2(t)\Delta t P_{x_0, y_0, 0}(x, y-1, t) + (y+1)\mathbf{d}_2(t)\Delta t P_{x_0, y_0, 0}(x, y+1, t) \\
&+ [1 - x\mathbf{b}_1(t)\Delta t - x\mathbf{d}_1(t)\Delta t - x\mathbf{m}(t)\Delta t - y\mathbf{b}_2(t)\Delta t - y\mathbf{d}_2(t)\Delta t] P_{x_0, y_0, 0}(x, y, t).
\end{aligned} \tag{1.3}$$

We recall that for discrete random variables  $X$  and  $Y$ , the moment generating function for a time interval of length  $\Delta t$  given the initial values  $x_0$  and  $y_0$  is given by

$$E[X^k(t + \Delta t)Y^p(t + \Delta t) | X(0) = x_0, Y(0) = y_0] = \sum_x \sum_y x^k y^p P_{x_0, y_0, 0}(x, y, t + \Delta t). \tag{1.4}$$

Hence, using equations (1.3) and (1.4) we obtain the first moment (mean) for  $X$  as

$$E[X^1(t + \Delta t)Y^0(t + \Delta t) | X(0) = x_0, Y(0) = y_0] = E[X(t + \Delta t) | X(0) = x_0, Y(0) = y_0] \tag{1.5}$$

$$= \sum_x \sum_y x P_{x_0, y_0, 0}(x, y, t + \Delta t) \tag{1.6}$$

$$\begin{aligned}
&= \sum_x \sum_y [x(x-1)\mathbf{b}_1(t)\Delta t P_{x_0, y_0, 0}(x-1, y, t) + x(x+1)\mathbf{d}_1(t)\Delta t P_{x_0, y_0, 0}(x+1, y, t) \\
&+ x(x+1)\mathbf{m}(t)\Delta t P_{x_0, y_0, 0}(x+1, y-1, t) \\
&+ x(y-1)\mathbf{b}_2(t)\Delta t P_{x_0, y_0, 0}(x, y-1, t) + x(y+1)\mathbf{d}_2(t)\Delta t P_{x_0, y_0, 0}(x, y+1, t) \\
&+ x[1 - x\mathbf{b}_1(t)\Delta t - x\mathbf{d}_1(t)\Delta t - x\mathbf{m}(t)\Delta t - y\mathbf{b}_2(t)\Delta t - y\mathbf{d}_2(t)\Delta t] P_{x_0, y_0, 0}(x, y, t).
\end{aligned} \tag{1.7}$$

The first term in equation (1.7) can be simplified by rescaling the summation indices as follows:

$$\sum_x \sum_y x(x-1)\mathbf{b}_1(t)\Delta t P_{x_0, y_0, 0}(x-1, y, t) = \sum_{x'} \sum_y (x'+1)x'\mathbf{b}_1(t)\Delta t P_{x_0, y_0, 0}(x', y, t) \tag{1.8}$$

$$= \sum_{x'} \sum_y x'^2 \mathbf{b}_1(t)\Delta t P_{x_0, y_0, 0}(x', y, t) + \sum_{x'} \sum_y x' \mathbf{b}_1(t)\Delta t P_{x_0, y_0, 0}(x', y, t) \tag{1.9}$$

$$= E[X^2(t)]\mathbf{b}_1(t)\Delta t + E[X(t)]\mathbf{b}_1(t)\Delta t, \tag{1.10}$$

where we used  $E[X^2(t)]$  to denote  $E[X^2(t) | X(0) = x_0, Y(0) = y_0]$ . Proceeding in a similar manner for the remaining terms in equation (1.7) leads to

$$E[X(t + \Delta t)] = \{ \mathbf{b}_1(t)\Delta t - \mathbf{d}_1(t)\Delta t - \mathbf{m}(t)\Delta t + 1 \} E[X(t)]. \tag{1.11}$$

Subtracting  $E[X(t)]$  from both sides of equation (1.11) and then dividing both sides by  $\Delta t$  yields

$$\frac{E[X(t + \Delta t)] - E[X(t)]}{\Delta t} = \{\mathbf{b}_1(t) - \mathbf{d}_1(t) - \mathbf{m}(t)\} E[X(t)], \quad (1.12)$$

which, in the limit as  $\Delta t$  goes to zero, gives

$$\frac{d}{dt} E[X(t)] = \{\mathbf{b}_1(t) - \mathbf{d}_1(t) - \mathbf{m}(t)\} E[X(t)]. \quad (1.13)$$

Analogously, the expected value for the random variable  $Y$  can be derived along with the second moments ( $E[X^2(t) | X(0) = x_0, Y(0) = y_0]$  and  $E[Y^2(t) | X(0) = x_0, Y(0) = y_0]$ ) and the correlation ( $E[X(t)Y(t) | X(0) = x_0, Y(0) = y_0]$ ). The system of ordinary differential equations describing the means, variances and covariances for the number of cells in states  $X$  and  $Y$  at any time  $t$  is given by

$$\frac{d}{dt} E[X(t)] = \{\mathbf{b}_1(t) - \mathbf{d}_1(t) - \mathbf{m}(t)\} E[X(t)] \quad (1.14)$$

$$\frac{d}{dt} E[Y(t)] = \{\mathbf{b}_2(t) - \mathbf{d}_2(t)\} E[Y(t)] + \mathbf{m}(t) E[X(t)] \quad (1.15)$$

$$\begin{aligned} \frac{d}{dt} E[X^2(t)] &= 2\{\mathbf{b}_1(t) - \mathbf{d}_1(t) - \mathbf{m}(t)\} E[X^2(t)] \\ &\quad + \{\mathbf{b}_1(t) + \mathbf{d}_1(t) + \mathbf{m}(t)\} E[X(t)] \end{aligned} \quad (1.16)$$

$$\begin{aligned} \frac{d}{dt} E[Y^2(t)] &= 2\{\mathbf{b}_2(t) - \mathbf{d}_2(t)\} E[Y^2(t)] + \{\mathbf{b}_2(t) + \mathbf{d}_2(t)\} E[Y(t)] \\ &\quad + \mathbf{m}(t) E[X(t)] + 2\mathbf{m}(t) E[X(t)Y(t)] \end{aligned} \quad (1.17)$$

$$\begin{aligned} \frac{d}{dt} E[X(t)Y(t)] &= \{\mathbf{b}_1(t) - \mathbf{d}_1(t) + \mathbf{b}_2(t) - \mathbf{d}_2(t) - \mathbf{m}(t)\} E[X(t)Y(t)] \\ &\quad - \mathbf{m}(t) E[X(t)] + \mathbf{m}(t) E[X^2(t)]. \end{aligned} \quad (1.18)$$

The above mathematical model is identical to the one derived by Leroux *et al.* <sup>[36]</sup> from considering a partial differential equation for the generating function for  $X(t)$  and  $Y(t)$ . In the case of constant rates, the system of differential equations (1.14)-(1.18) is a linear system of ordinary differential equations with constant coefficient of the form

$$\bar{\mathbf{y}}'(t) = A\bar{\mathbf{y}}(t) \quad (1.19)$$

where  $\bar{\mathbf{y}}(t) = (E[X(t)], E[Y(t)], E[X^2(t)], E[Y^2(t)], E[X(t)Y(t)])^T$  and

$$A = \begin{bmatrix} \mathbf{b}_1 - \mathbf{d}_1 - \mathbf{m} & 0 & 0 & 0 & 0 \\ \mathbf{m} & \mathbf{b}_2 - \mathbf{d}_2 & 0 & 0 & 0 \\ \mathbf{b}_1 + \mathbf{d}_1 + \mathbf{m} & 0 & 2(\mathbf{b}_1 - \mathbf{d}_1 - \mathbf{m}) & 0 & 0 \\ \mathbf{m} & \mathbf{b}_2 + \mathbf{d}_2 & 0 & 2(\mathbf{b}_2 - \mathbf{d}_2) & 2\mathbf{m} \\ -\mathbf{m} & 0 & \mathbf{m} & 0 & \mathbf{b}_1 - \mathbf{d}_1 - \mathbf{m} + \mathbf{b}_2 - \mathbf{d}_2 \end{bmatrix}. \quad (1.20)$$

The exact solution can be written in terms of the eigenvalues and eigenvectors of the matrix  $A$  (see e.g., Braun <sup>[37]</sup>).

## 2.2 The Controlled Growth and Differentiation (CGD) Model

As already discussed elsewhere in this paper, the mathematical model formulated by Leroux *et al.* <sup>[36]</sup>, which is based on a linear birth-death process, is both biologically and mathematically unreasonable because it allows no constraint on the size of the resulting organ or tissue. We now describe modifications to the Leroux *et al.* model that put greater emphasis on the control of growth and differentiation (CGD). More specifically, we consider a multistate developmental process with separate growth phases in the type-  $X$  cells and a linear birth-death process on the type-  $Y$  cells. A schematic diagram of the CGD model is illustrated in Figure 2. In this system, there are  $k+1$  specific type- $X$  populations denoted as type-  $X_i$ , where  $i = 0, 1, 2, \dots, k$ . Each of the type-  $X$  cell populations represents cells prior to commitment to differentiation. Moreover, as a cell moves from one stage, for example from  $X_i$ , to the next stage,  $X_{i+1}$ , it matures in the developmental process. By setting  $\mathbf{m} = 0$ , for one or more values of  $i$ , we allow for greater control in the type-  $X$  cells prior to differentiation. Eventually, the uncommitted cell will undergo a change and join the population of cells already committed to performing a specific function of a tissue or organ. This process by which an uncommitted cell becomes a cell committed to differentiation is termed *transformation*. Type- $Y$  cells denote the population of cells committed to differentiation.

The assumptions and the parameters in the model have the same basic definition as those in the Leroux *et al.* model (see Section 2.1). More specifically, for each type-  $X_i$  cell, one of three outcomes is possible: division of the cell into two type-  $X_{i+1}$  daughter cells (i.e., a birth in the type-  $X_i$  population), transformation of the cell into the type- $Y$  population or death of the cell. Cells of the type- $Y$  population are allowed only one of two outcomes: division into two type- $Y$  daughter cells or death prior to division. In the type- $Y$  population death is understood to include actual death of a cell or removal from the pool of fully committed cells contributing to the development and function of the tissue or organ in consideration.

The developmental rates are again denoted by  $\mathbf{b}_i(t)$  (birth rate in a  $X_i$  population),  $\mathbf{d}_i(t)$  (death rate in a  $X_i$  population),  $\mathbf{b}_y(t)$  (birth rate in the  $Y$  population),  $\mathbf{d}_y(t)$  (death rate in the  $Y$  population) and  $\mathbf{m}_i(t)$  (transformation rate from the  $X_i$  population to the  $Y$  population) where  $i=0,1,\dots,k$ . Just as before, it is assumed that the probability of more than one event occurring during small time interval  $\Delta t$  is  $o(\Delta t)$ , which by the condition (1.1), is negligible. Table 2 includes all of the possible events for a single cell in the CGD developmental system.

**Table 2:** Single Cell Event Probabilities in time  $[t, t+\Delta t)$  for the CGD Model

Event	$t$	$t+\Delta t$	Single cell transition probability
birth from $X_i$ to $X_{i+1}$ ( $i=0,\dots,k-1$ )	$x_0,\dots,x_i+1,x_{i+1}-2,\dots,y$	$x_0,\dots,x_k,y$	$\mathbf{b}_i(t)\Delta t P(x_0,\dots,x_i+1,x_{i+1}-2,\dots,y,t)$
birth in state $X_k$	$x_0,\dots,x_k-1,y$	$x_0,\dots,x_k,y$	$\mathbf{b}_k(t)\Delta t P(x_0,\dots,x_k-1,y,t)$
death in $X_i$ ( $i=0,\dots,k$ )	$x_0,\dots,x_i+1,\dots,x_k,y$	$x_0,\dots,x_k,y$	$\mathbf{d}_i(t)\Delta t P(x_0,\dots,x_i+1,\dots,x_k,y,t)$
transformation from $X_i$ to $Y$ ( $i=0,\dots,k$ )	$x_0,\dots,x_i+1,\dots,x_k,y-1$	$x_0,\dots,x_k,y$	$\mathbf{m}_i(t)\Delta t P(x_0,\dots,x_i+1,\dots,x_k,y-1,t)$
birth in state $Y$	$x_0,\dots,x_k,y-1$	$x_0,\dots,x_k,y$	$\mathbf{b}_y(t)\Delta t P(x_0,\dots,x_k,y-1,t)$
death in state $Y$	$x_0,\dots,x_k,y+1$	$x_0,\dots,x_k,y$	$\mathbf{d}_y(t)\Delta t P(x_0,\dots,x_k,y+1,t)$
no change	$x_0,\dots,x_k,y$	$x_0,\dots,x_k,y$	$\left\{1 - \left[ \sum_{i=0}^k (\mathbf{b}_i + \mathbf{d}_i + \mathbf{m}_i) + \mathbf{b}_y + \mathbf{d}_y \right] \Delta t \right\}$ $\cdot P(x_0,\dots,x_k,y,t)$

Based on the information presented in Table 2, only certain events may occur in a small time interval,  $\Delta t$ . Thus, the transition probability function for a CGD developmental system is given by

$$\begin{aligned}
& P_{x_0, \dots, x_{k0}, y_0, 0}(x_0, \dots, x_k, y, t + \Delta t) \\
&= \sum_{i=0}^{k-1} (x_i + 1) \mathbf{b}_i \Delta t P_{x_0, \dots, x_{k0}, y_0, 0}(x_0, \dots, x_i + 1, x_{i+1} - 2, \dots, x_k, y, t) \\
&+ (x_k + 1) \mathbf{b}_k \Delta t P_{x_0, \dots, x_{k0}, y_0, 0}(x_0, \dots, x_k + 1, y, t) \\
&+ \sum_{i=0}^k (x_i + 1) \mathbf{d}_i \Delta t P_{x_0, \dots, x_{k0}, y_0, 0}(x_0, \dots, x_i + 1, \dots, x_k, y, t) \\
&+ \sum_{i=0}^k (x_i + 1) \mathbf{m}_i \Delta t P_{x_0, \dots, x_{k0}, y_0, 0}(x_0, \dots, x_i + 1, \dots, x_k, y - 1, t) \\
&+ (y - 1) \mathbf{b}_y \Delta t P_{x_0, \dots, x_{k0}, y_0, 0}(x_0, \dots, x_k, y - 1, t) \\
&+ (y + 1) \mathbf{d}_y \Delta t P_{x_0, \dots, x_{k0}, y_0, 0}(x_0, \dots, x_k, y + 1, t) \\
&+ \left[ 1 - \sum_{i=0}^k x_i (\mathbf{b}_i + \mathbf{d}_i + \mathbf{m}_i) \Delta t - y (\mathbf{b}_y + \mathbf{d}_y) \Delta t \right] P_{x_0, \dots, x_{k0}, y_0, 0}(x_0, \dots, x_k, y, t).
\end{aligned} \tag{2.1}$$

Since the objective is to estimate the average number of cells that accumulate in stage  $Y$  at a particular time of the developmental process, the expected value (first moment) for each random variable will be calculated. This will be done in the same manner as for the re-derivation the Leroux *et al.* <sup>[36]</sup> model in Section 2.1.

For discrete random variables  $X_j$  and  $Y$ , the moments for a time interval of length  $\Delta t$  given the initial values  $x_{00}, x_{10}, \dots, x_{k-1,0}, x_{k0}$  and  $y_0$  is defined to be

$$\begin{aligned}
& E \left[ X_j^k (t + \Delta t) Y^p (t + \Delta t) \mid X_j(0) = x_{j0}, Y(0) = y_0 \right] \\
&= \sum_{x_0} \sum_{x_1} \dots \sum_{x_k} \sum_y x_j^k y^p P_{x_0, \dots, x_{k0}, y_0, 0}(x_0, x_1, \dots, x_k, y, t).
\end{aligned} \tag{2.2}$$

Following the same process presented in Section 2.1, the first moment for the random variable  $X_j, j = 0, 1, \dots, k - 1$ , of the CGD model is given by

$$\frac{d}{dt} E[X_j(t)] = -\{\mathbf{b}_j(t) + \mathbf{d}_j(t) + \mathbf{m}_j(t)\} E[X_j(t)] + 2\mathbf{b}_{j-1}(t) E[X_{j-1}(t)]. \tag{2.3}$$

We refer the reader to Section A.1 in the Appendix for a detailed derivation of equation (2.3).

The remaining first moments and the second moments for the discrete random variables are calculated in a similar manner. The result is the following system of ordinary differential equations with time dependent developmental rates (unless otherwise stated,  $j = 0, 1, 2, \dots, k - 1$ ):

$$\frac{d}{dt}E[X_j(t)] = -\{\mathbf{b}_j(t) + \mathbf{d}_j(t) + \mathbf{m}_j(t)\}E[X_j(t)] + 2\mathbf{b}_{j-1}(t)E[X_{j-1}(t)] \quad (2.4)$$

$$\frac{d}{dt}E[X_k(t)] = \{\mathbf{b}_k(t) - \mathbf{d}_k(t) - \mathbf{m}_k(t)\}E[X_k(t)] + 2\mathbf{b}_{k-1}(t)E[X_{k-1}(t)] \quad (2.5)$$

$$\frac{d}{dt}E[Y(t)] = \{\mathbf{b}_y(t) - \mathbf{d}_y(t)\}E[Y(t)] + \sum_{i=0}^k \mathbf{m}_i(t)E[X_i(t)] \quad (2.6)$$

$$\begin{aligned} \frac{d}{dt}E[X_j^2(t)] &= -2\{\mathbf{b}_j(t) + \mathbf{d}_j(t) + \mathbf{m}_j(t)\}E[X_j^2(t)] \\ &\quad + \{\mathbf{b}_j(t) + \mathbf{d}_j(t) + \mathbf{m}_j(t)\}E[X_j(t)] \\ &\quad + 4\mathbf{b}_{j-1}(t)E[X_{j-1}(t)] + 4\mathbf{b}_{j-1}(t)E[X_{j-1}(t)X_j(t)] \end{aligned} \quad (2.7)$$

$$\begin{aligned} \frac{d}{dt}E[X_k^2(t)] &= 2\{\mathbf{b}_k(t) - \mathbf{d}_k(t) - \mathbf{m}_k(t)\}E[X_k^2(t)] \\ &\quad + \{\mathbf{b}_k(t) + \mathbf{d}_k(t) + \mathbf{m}_k(t)\}E[X_k(t)] \\ &\quad + 4\mathbf{b}_{k-1}(t)E[X_{k-1}(t)] + 4\mathbf{b}_{k-1}(t)E[X_{k-1}(t)X_k(t)] \end{aligned} \quad (2.8)$$

$$\begin{aligned} \frac{d}{dt}E[Y^2(t)] &= 2\{\mathbf{b}_y(t) - \mathbf{d}_y(t)\}E[Y^2(t)] \\ &\quad + \{\mathbf{b}_y(t) + \mathbf{d}_y(t)\}E[Y(t)] \\ &\quad + \sum_{i=0}^k \mathbf{m}_iE[X_i(t)] + 2\sum_{i=0}^k \mathbf{m}_iE[X_i(t)Y(t)] \end{aligned} \quad (2.9)$$

$$\begin{aligned} \frac{d}{dt}E[X_l(t)X_j(t)] &= -\{\mathbf{b}_l(t) + \mathbf{d}_l(t) + \mathbf{m}_l(t) + \mathbf{b}_j(t) + \mathbf{d}_j(t) + \mathbf{m}_j(t)\}E[X_l(t)X_j(t)] \\ &\quad + 2\mathbf{b}_{l-1}(t)E[X_{l-1}(t)X_j(t)] + 2\mathbf{b}_{j-1}(t)E[X_l(t)X_{j-1}(t)] \end{aligned} \quad (2.10)$$

where  $l, j = 0, 1, 2, \dots, k$  and  $l, j$  are not consecutive integers  $l \neq j$

$$\begin{aligned} \frac{d}{dt}E[X_{j-1}(t)X_j(t)] &= -2\mathbf{b}_{j-1}(t)E[X_{j-1}(t)] + 2\mathbf{b}_{j-1}(t)E[X_{j-1}^2(t)] \\ &\quad + 2\mathbf{b}_{j-2}(t)E[X_{j-2}(t)X_j(t)] \\ &\quad - \{\mathbf{b}_{j-1}(t) + \mathbf{d}_{j-1}(t) + \mathbf{m}_{j-1}(t) + \mathbf{b}_j(t) + \mathbf{d}_j(t) + \mathbf{m}_j(t)\}E[X_{j-1}(t)X_j(t)] \end{aligned} \quad (2.11)$$

$$\begin{aligned} \frac{d}{dt}E[X_{k-1}(t)X_k(t)] &= -2\mathbf{b}_{k-1}(t)E[X_{k-1}(t)] + 2\mathbf{b}_{k-1}(t)E[X_{k-1}^2(t)] \\ &\quad + 2\mathbf{b}_{k-2}(t)E[X_{k-2}(t)X_k(t)] \\ &\quad + \{\mathbf{b}_k(t) - \mathbf{d}_k(t) - \mathbf{m}_k(t) - \mathbf{b}_{k-1}(t) - \mathbf{d}_{k-1}(t) - \mathbf{m}_{k-1}(t)\}E[X_{k-1}(t)X_k(t)] \end{aligned} \quad (2.12)$$

$$\begin{aligned}
\frac{d}{dt}E[X_j(t)Y(t)] &= \{b_y(t) - d_y(t) - b_j(t) - d_j(t) - m_j(t)\}E[X_j(t)Y(t)] \\
&+ 2b_{j-1}(t)E[X_{j-1}(t)Y(t)] - m_j(t)E[X_j(t)] \\
&+ \sum_{i=0}^k m_i(t)E[X_j(t)X_i(t)]. \quad \text{for } j = 0, 1, 2, \dots, k
\end{aligned} \tag{2.13}$$

### 3. Comparison of the Leroux *et al.* <sup>[36]</sup> model with the CGD model

In this section, the Leroux *et al.* <sup>[36]</sup> model and the CGD model are examined and compared in an effort to highlight the most useful and biologically realistic model for the developmental process as it relates to animals. Both models are applied to predict the effects of methylmercury on brain development in rats. For a description on the experimental aspects and how the *in vitro* data were used to estimate cell kinetic rates in both models see the Leroux *et al.* paper<sup>[36]</sup>. In addition, all computations were done using computer codes written in the MathWorks' Matlab/Simulink environment.

The two models are not only biologically and mathematically different but the model proposed by Leroux *et al.* is not a submodel of the CGD model since a linear birth process is not assumed. However, setting a specific birth, death or transformation rate to zero in the CGD model creates a model that can give similar predictions as those given by the Leroux *et al.* model. Specifically, consider the very simple developmental CGD model depicted in Figure 3. The system of ordinary differential equations describing the means, variances and covariances for the number of cells in states  $X_j$  and  $Y$  are given in Section A.2 of the Appendix. Using the parameters given in Table 3, the CGD and the Leroux *et al.* <sup>[36]</sup> model predict similar expectations albeit different variances at the time  $t = 5$  (see Figure 4). Figures 5 and 6 further support the ability of the CGD model to closely parallel the Leroux *et al.* model.

**Table 3:** Parameters Used to Predict Similar Results Among the Two Models

	Leroux <i>et al.</i> Model		CGD Model		
	X	Y	X <sub>0</sub>	X <sub>1</sub>	Y
Initial number of cells (time $t=0$ )	$1.68 \times 10^5$	0	$1.68 \times 10^5$	0	0
Birth rate, $\mathbf{b}$	0.4	0.4	0.4	0.4	0.4
Death rate, $\mathbf{d}$	0.01	0.01	0.01	0.01	0.01
Transformation rate, $\mathbf{m}$	0.15	--	0.15	0.15	--

Expected number of cells (time $t=5$ )	$5.578 \times 10^5$	$6.23 \times 10^5$	$1.022 \times 10^4$	$5.476 \times 10^5$	$6.23 \times 10^5$
Total number of cells	$1.181 \times 10^6$		$1.181 \times 10^6$		

One major difference between the two models is the ability of the CGD model to deplete the population of  $X$  cells under constant rates while still maintaining the same distribution of  $Y$  cells. This can be illustrated again using the simple CGD model in Figure 3. As before, both the Leroux *et al.* model and the simple CGD model (Figure 3) were started with the same initial conditions. The  $X$  (Leroux *et al.*) and  $X_0$  (CGD model) type cells consisted of  $1.68 \times 10^8$  cells at time  $t=0$  while the remaining states started with a cell population of zero. The variances and covariances of each random variable were also set equal to zero at time  $t=0$ . Each model ran for a total of five time units. The set of parameter values used in the models is shown in Table 4. The only changes are setting  $\mathbf{m}_0$  to zero and increasing  $\mathbf{d}_1$  and  $\mathbf{m}_1$ .

**Table 4:** Parameters Used and Estimates Obtained Among the Two Models

	Leroux <i>et al.</i> Model		CGD Model		
	X	Y	$X_0$	$X_1$	Y
Initial number of cells (time $t=0$ )	$1.68 \times 10^5$	0	$1.68 \times 10^5$	0	0
Birth rate, $\mathbf{b}$	0.4	0.4	0.4	0.4	0.4
Death rate, $\mathbf{d}$	0.01	0.01	0.01	0.4	0.01
Transformation rate, $\mathbf{m}$	0.15	--	0	0.5174	--
Expected number of cells (time $t=5$ )	$5.58 \times 10^5$	$6.23 \times 10^5$	$2.16 \times 10^4$	$6.69 \times 10^4$	$6.23 \times 10^5$
Total number of cells	$1.181 \times 10^6$		$7.115 \times 10^5$		

Figure 7 shows that the two models have the same basic density shape as before (see Figure 4) with the variance from the CGD model smaller than the variance from the Leroux *et al.* model. The expected value of  $Y$  cells for both models agree at time  $t=5$  as depicted in Figure 8 but differ slightly at other time points. This implies that the CGD model can indeed be tailored to produce the same number of committed (type  $Y$ ) cells as the Leroux *et al.* model. However, Figure 9 reveals that the models produce different expected numbers of uncommitted (type  $X$ ) cells. As seen from Table 4, both models start out with the same

number of uncommitted cells and the birth rates in each type  $X$  population are the same. However, due to the structure of the CGD model, the state  $X_1$  cells are able to have higher death and transformation rates. Biologically, this means that all of the committed (type  $Y$ ) cells have type  $X_1$  precursor cells and that the population of uncommitted cells will eventually die out. This eventually results in a stable population of committed cells at greater times, unlike the Leroux *et al.* model which will continue to grow.

The expected final size of the tissue/organ is simply computed as the total number of cells present in the system at time  $t = 5$ . The total number of cells, which is defined to be the sum of all expectations of type  $X$  and type  $Y$  cells, present in the Leroux *et al.* model exceeds the number of cells present in the CGD model. This implies that, for sufficiently large time, the size of the tissue/organ might become unrealistically large. It is also important to note that in order to reduce the number of cells in the Leroux *et al.* model, there are a few options available:

- (i) reduce the initial number of type  $X$  cells
- (ii) reduce the birth rate of the type  $X$  and/or type  $Y$  cells
- (iii) increase the death rate of the type  $X$  and/or type  $Y$  cells
- (iv) use more complicated time-varying rates

In each case, the ability to accurately reflect the biological dynamics of the development process is compromised. If options (i) or (iii) are pursued, the experimental data and the mathematical model may not agree initially. If options (ii) or (iv) are incorporated, the cellular kinetics of the system may not be modeled accurately. In any event, the end result is that the mathematical model does not realistically duplicate the biological dynamics of the system. Option (iv) requires significantly more knowledge of the time-varying nature of the developmental rates in the system.

On the other hand, the CGD model can be amended to incorporate all of the biological dynamics of the developmental process with time-constant parameters while also yielding a feasible number of cells that will ultimately dictate the expected final size of the tissue/organ. Even if the birth, death and/or transformation rates are reduced/increased in one state of the CGD model, a different state may be amended to reflect a different aspect of the system dynamics. In essence, the concept of homeostasis can be incorporated into the CGD model while still yielding biologically accurate prediction at each time step.

#### **4. Conclusion**

The two models discussed in this manuscript were derived with the use of the Kolmogorov forward equation. The model originally developed by Leroux *et al.* <sup>[36]</sup> was re-derived without the use of partial differential equations. This approach was then extended to develop a controlled growth and differentiation (CGD) model for an unspecified number of states in the developmental process.

Due to the generalization of the CGD model, the system may be tailored to fit a host of simpler or more complicated developmental processes. Growth kinetics may be assumed to apply to both, none or a combination of the committed and uncommitted cell populations, which allows controlled growth and differentiation. The objective of such a model is to be able to predict the number of cells present in a committed state after a specified time of development. If the number of cells in the committed state is less than a specified number at a critical time, then malformation has occurred. Similarly, it is also unfavorable if there are too many cells in the committed state at a specified time of development. The CGD model has the added benefit of being able to model a biological system with the end result being a tissue/organ of realistic size. The CGD model is also able to model a wide variety of developmental systems whereas the Leroux *et al.* model is limited in this aspect.

Determining the number of states to include in the model and which states should have growth kinetics are not trivial. Once the specifications and details of a model have been adopted, the ultimate test of the model lies in the validation of the model using experimental data. Ultimately, the insight gained from the CGD model may be used to aid in the discussion of toxic agents that pose a threat to the proper development of a tissue or organ.

## Appendix

### A.1 Development of the CGD model

In this section, we provide detailed calculations of the expected value of  $X_j$ , for  $j = 0, 1, \dots, k-1$ . The following process can be repeated to obtain the expectations of type- $Y$  cells, squared expectations and the expected cross products.

$$\begin{aligned}
E[X_j(t + \Delta t)] &= \sum_{x_0} \sum_{x_1} \dots \sum_{x_k} \sum_y \left\{ x_j \sum_{\substack{i=0 \\ i \neq j-1 \\ i \neq j}}^{k-1} (x_i + 1) \mathbf{b}_i \Delta t P(x_0, \dots, x_i + 1, x_{i+1} - 2, \dots, x_k, y, t) \right. \\
&\quad + x_j(x_{j-1} + 1) \mathbf{b}_{j-1} \Delta t P(x_0, \dots, x_{j-1} + 1, x_j - 2, \dots, x_k, y, t) \\
&\quad + x_j(x_j + 1) \mathbf{b}_j \Delta t P(x_0, \dots, x_j + 1, x_{j+1} - 2, \dots, x_k, y, t) \\
&\quad + x_j(x_k + 1) \mathbf{b}_k \Delta t P(x_0, \dots, x_k + 1, y, t) \\
&\quad + x_j \sum_{\substack{i=0 \\ i \neq j-1 \\ i \neq j}}^k (x_i + 1) \mathbf{d}_i \Delta t P(x_0, \dots, x_i + 1, \dots, x_k, y, t) \\
&\quad + x_j(x_{j-1} + 1) \mathbf{d}_{j-1} \Delta t P(x_0, \dots, x_{j-1} + 1, \dots, x_k, y, t) \\
&\quad + x_j(x_j + 1) \mathbf{d}_j \Delta t P(x_0, \dots, x_j + 1, \dots, x_k, y, t) \\
&\quad + x_j \sum_{\substack{i=0 \\ i \neq j-1 \\ i \neq j}}^k (x_i + 1) \mathbf{m}_i \Delta t P(x_0, \dots, x_i + 1, \dots, x_k, y - 1, t) \\
&\quad + x_j(x_{j-1} + 1) \mathbf{m}_{j-1} \Delta t P(x_0, \dots, x_{j-1} + 1, \dots, x_k, y - 1, t) \\
&\quad + x_j(x_j + 1) \mathbf{m}_j \Delta t P(x_0, \dots, x_j + 1, \dots, x_k, y - 1, t) \\
&\quad + x_j(y - 1) \mathbf{b}_y \Delta t P(x_0, \dots, x_k, y - 1, t) \\
&\quad + x_j(y + 1) \mathbf{d}_y \Delta t P(x_0, \dots, x_k, y + 1, t) \\
&\quad \left. + x_j \left[ 1 - \sum_{i=0}^k x_i (\mathbf{b}_i + \mathbf{d}_i + \mathbf{m}_i) \Delta t - y (\mathbf{b}_y + \mathbf{d}_y) \Delta t \right] P(x_0, \dots, x_k, y, t) \right\}. \tag{A.1}
\end{aligned}$$

We now rescale all the probability density functions and denote  $P(x_0, \dots, x_k, y, t)$  simply by

$$P(\bar{x}, y, t). \tag{A.2}$$

Thus, equation (A.1) becomes

$$\begin{aligned}
E[X_j(t + \Delta t)] &= \sum_{\bar{x}} \sum_y \left\{ x_j \sum_{\substack{i=0 \\ i \neq j-1 \\ i \neq j}}^{k-1} x_i \mathbf{b}_i \Delta t P(\bar{x}, y, t) + (x_j + 2)x_{j-1} \mathbf{b}_{j-1} \Delta t P(\bar{x}, y, t) \right. \\
&\quad + (x_j - 1)x_j \mathbf{b}_j \Delta t P(\bar{x}, y, t) + x_j x_k \mathbf{b}_k \Delta t P(\bar{x}, y, t) \\
&\quad + x_j \sum_{\substack{i=0 \\ i \neq j-1 \\ i \neq j}}^k x_i \mathbf{d}_i \Delta t P(\bar{x}, y, t) + (x_j - 1)x_{j-1} \mathbf{d}_{j-1} \Delta t P(\bar{x}, y, t) \\
&\quad + (x_j - 1)x_j \mathbf{d}_j \Delta t P(\bar{x}, y, t) \\
&\quad + x_j \sum_{\substack{i=0 \\ i \neq j-1 \\ i \neq j}}^k x_i \mathbf{m}_i \Delta t P(\bar{x}, y, t) + x_j x_{j-1} \mathbf{m}_{j-1} \Delta t P(\bar{x}, y, t) + (x_j - 1)x_j \mathbf{m}_j \Delta t P(\bar{x}, y, t) \\
&\quad + x_j y \mathbf{b}_y \Delta t P(\bar{x}, y, t) + x_j y \mathbf{d}_y \Delta t P(\bar{x}, y, t) \\
&\quad \left. + x_j P(\bar{x}, y, t) - x_j \sum_{i=0}^k x_i (\mathbf{b}_i + \mathbf{d}_i + \mathbf{m}_i) \Delta t P(\bar{x}, y, t) - x_j y (\mathbf{b}_y + \mathbf{d}_y) \Delta t P(\bar{x}, y, t) \right\}. \tag{A.3}
\end{aligned}$$

Further simplifications reduce to

$$\begin{aligned}
E[X_j(t + \Delta t)] &= \sum_{\bar{x}} \sum_y \{ 2\mathbf{b}_{j-1} \Delta t x_{j-1} P(\bar{x}, y, t) \\
&\quad - (\mathbf{b}_j + \mathbf{d}_j + \mathbf{m}_j) \Delta t x_j P(\bar{x}, y, t) + x_j P(\bar{x}, y, t) \}. \tag{A.4}
\end{aligned}$$

Recall the mathematical definition of the moments and correlations of the discrete random variables  $X_j$  and  $Y$  at time  $t$  is given by

$$\begin{aligned}
E[X_j^n(t + \Delta t) Y^p(t + \Delta t) | X_j(0) = x_{j0}, Y(0) = y_0] \\
= \sum_{x_0} \sum_{x_1} \dots \sum_{x_k} \sum_y x_j^n y^p P_{x_{00}, x_{10}, \dots, x_{k-1,0}, x_{k0}, y_0, 0}(x_0, x_1, \dots, x_k, y, t + \Delta t). \tag{A.5}
\end{aligned}$$

(with  $n = 1, p = 0$ ). Thus, equation (A.4) reduces to

$$\begin{aligned}
E[X_j(t + \Delta t)] &= 2\mathbf{b}_{j-1} \Delta t E[X_{j-1}(t)] \\
&\quad - (\mathbf{b}_j + \mathbf{d}_j + \mathbf{m}_j) \Delta t E[X_j(t)] + E[X_j(t)]. \tag{A.6}
\end{aligned}$$

Finally, subtracting  $E[X_j(t)]$  from both sides of equation (A.6), dividing by  $\Delta t$  and letting  $\Delta t$  approaches zero to obtain

$$\lim_{\Delta t \rightarrow 0} \left[ \frac{E[X_j(t + \Delta t)] - E[X_j(t)]}{\Delta t} \right] = 2\mathbf{b}_{j-1} E[X_{j-1}(t)] - (\mathbf{b}_j + \mathbf{d}_j + \mathbf{m}_j) E[X_j(t)], \tag{A.7}$$

or

$$\frac{d}{dt}E[X_j(t)] = 2\mathbf{b}_{j-1}E[X_{j-1}(t)] - (\mathbf{b}_j + \mathbf{d}_j + \mathbf{m}_j)E[X_j(t)]. \quad (\text{A.8})$$

## A.2 Mathematical model for the CGD system shown in Figure 3

$$\frac{d}{dt}E[X_0(t)] = -\{\mathbf{b}_0(t) + \mathbf{d}_0(t) + \mathbf{m}_0(t)\}E[X_0(t)] \quad (\text{A.9})$$

$$\frac{d}{dt}E[X_1(t)] = \{\mathbf{b}_1(t) - \mathbf{d}_1(t) - \mathbf{m}_1(t)\}E[X_1(t)] + 2\mathbf{b}_0(t)E[X_0(t)] \quad (\text{A.10})$$

$$\frac{d}{dt}E[Y(t)] = \{\mathbf{b}_y(t) - \mathbf{d}_y(t)\}E[Y(t)] + \mathbf{m}_0(t)E[X_0(t)] + \mathbf{m}_1(t)E[X_1(t)] \quad (\text{A.11})$$

$$\begin{aligned} \frac{d}{dt}E[X_0^2(t)] &= -2\{\mathbf{b}_0(t) + \mathbf{d}_0(t) + \mathbf{m}_0(t)\}E[X_0^2(t)] \\ &\quad + \{\mathbf{b}_0(t) + \mathbf{d}_0(t) + \mathbf{m}_0(t)\}E[X_0(t)] \end{aligned} \quad (\text{A.12})$$

$$\begin{aligned} \frac{d}{dt}E[X_1^2(t)] &= 2\{\mathbf{b}_1(t) - \mathbf{d}_1(t) - \mathbf{m}_1(t)\}E[X_1^2(t)] \\ &\quad + \{\mathbf{b}_1(t) + \mathbf{d}_1(t) + \mathbf{m}_1(t)\}E[X_1(t)] \\ &\quad + 4\mathbf{b}_0(t)E[X_0(t)] + 4\mathbf{b}_0(t)E[X_0(t)X_1(t)] \end{aligned} \quad (\text{A.13})$$

$$\begin{aligned} \frac{d}{dt}E[Y^2(t)] &= 2\{\mathbf{b}_y(t) - \mathbf{d}_y(t)\}E[Y^2(t)] + \{\mathbf{b}_y(t) + \mathbf{d}_y(t)\}E[Y(t)] \\ &\quad + \mathbf{m}_0E[X_0(t)Y(t)] + \mathbf{m}_1E[X_1(t)Y(t)] \\ &\quad + 2\{\mathbf{m}_0E[X_0(t)Y(t)] + \mathbf{m}_1E[X_1(t)Y(t)]\} \end{aligned} \quad (\text{A.14})$$

$$\begin{aligned} \frac{d}{dt}E[X_0(t)X_1(t)] &= -2\mathbf{b}_0(t)E[X_0(t)] + 2\mathbf{b}_0(t)E[X_0^2(t)] \\ &\quad + \{\mathbf{b}_1(t) - \mathbf{d}_1(t) - \mathbf{m}_1(t) - \mathbf{b}_0(t) - \mathbf{d}_0(t) - \mathbf{m}_0(t)\}E[X_0(t)X_1(t)] \end{aligned} \quad (\text{A.15})$$

$$\begin{aligned} \frac{d}{dt}E[X_0(t)Y(t)] &= \{\mathbf{b}_y(t) - \mathbf{d}_y(t) - \mathbf{b}_0(t) - \mathbf{d}_0(t) - \mathbf{m}_0(t)\}E[X_0(t)Y(t)] \\ &\quad - \mathbf{m}_0(t)E[X_0(t)] + \mathbf{m}_0(t)E[X_0^2(t)] + \mathbf{m}_1(t)E[X_0(t)X_1(t)] \end{aligned} \quad (\text{A.16})$$

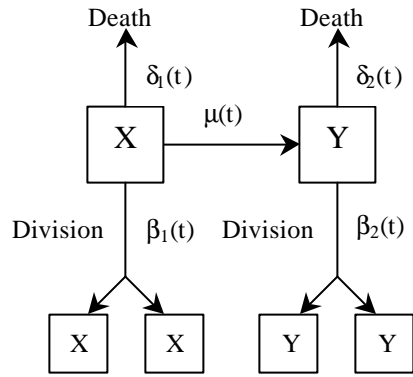
$$\begin{aligned} \frac{d}{dt}E[X_1(t)Y(t)] &= \{\mathbf{b}_y(t) - \mathbf{d}_y(t) + \mathbf{b}_0(t) - \mathbf{d}_1(t) - \mathbf{m}_1(t)\}E[X_1(t)Y(t)] \\ &\quad - \mathbf{m}_1(t)E[X_1(t)] + \mathbf{m}_1(t)E[X_1^2(t)] \\ &\quad + \mathbf{m}_0(t)E[X_0(t)X_1(t)] + 2\mathbf{b}_0(t)E[X_0(t)Y(t)] \end{aligned} \quad (\text{A.17})$$

## References

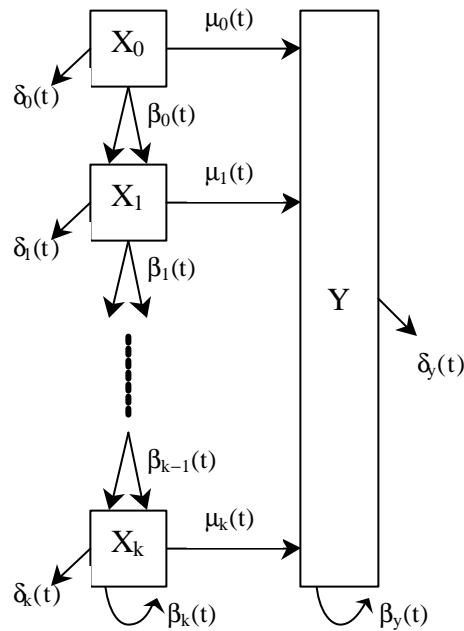
1. Wilson, JD, *Biological bases for cancer dose-response extrapolation procedures*. Environ Health Perspect, **90**: p. 293-6, 1991.
2. Portier, CJ, *Mechanistic modelling and risk assessment*. Pharmacol Toxicol, **72**(Suppl 1): p. 28-32, 1993.
3. Kohn, M and Lucier, G, *Mechanism-based toxicology in cancer risk assessment: biologically-based models*. in **2nd World Congress on Alternatives and Animal use in the Life Sciences**: Elsevier. 1997.
4. Armitage, P and Doll, R, *The age distribution of cancer and a multi-stage theory of carcinogenesis*. British Journal of Cancer, **8**: p. 1-12, 1954.
5. Armitage, P and Doll, R, *A two-stage theory of carcinogenesis in relation to the age distribution of human cancer*. British Journal of Cancer, **11**: p. 161-169, 1957.
6. Moolgavkar, SH, *The multistage theory of carcinogenesis and the age distribution of cancer in man*. J Natl Cancer Inst, **61**(1): p. 49-52, 1978.
7. Moolgavkar, SH, Day, NE and Stevens, RG, *Two-stage model for carcinogenesis: Epidemiology of breast cancer in females*. J Natl Cancer Inst, **65**(3): p. 559-69, 1980.
8. Chu, KC, *A nonmathematical view of mathematical models for cancer*. J Chronic Dis, **40**(Suppl 2): p. 163S-170S, 1987.
9. Moolgavkar, SH, Dewanji, A and Venzon, DJ, *A stochastic two-stage model for cancer risk assessment. I. The hazard function and the probability of tumor*. Risk Anal, **8**(3): p. 383-92, 1988.
10. Dewanji, A, Venzon, DJ and Moolgavkar, SH, *A stochastic two-stage model for cancer risk assessment. II. The number and size of premalignant clones*. Risk Anal, **9**(2): p. 179-87, 1989.
11. Portier, CJ and Kopp-Schneider, A, *A multistage model of carcinogenesis incorporating DNA damage and repair*. Risk Anal, **11**(3): p. 535-43, 1991.
12. Luebeck, EG, Moolgavkar, SH, Buchmann, A and Schwarz, M, *Effects of polychlorinated biphenyls in rat liver: quantitative analysis of enzyme-altered foci*. Toxicol Appl Pharmacol, **111**(3): p. 469-84, 1991.
13. Kopp-Schneider, A, Portier, CJ and Sherman, CD, *The exact formula for tumor incidence in the two-stage model*. Risk Anal, **14**(6): p. 1079-80, 1994.
14. Sherman, CD, Portier, CJ and Kopp-Schneider, A, *Multistage models of carcinogenesis: an approximation for the size and number distribution of late-stage clones*. Risk Anal, **14**(6): p. 1039-48, 1994.
15. Kopp-Schneider, A and Portier, CJ, *A stem cell model for carcinogenesis*. Math Biosci, **120**(2): p. 211-32, 1994.
16. Kopp-Schneider, A and Portier, CJ, *Carcinoma formation in NMRI mouse skin painting studies is a process suggesting greater than two stages*. Carcinogenesis, **16**(1): p. 53-9, 1995.
17. Sherman, CD and Portier, CJ, *Quantitative analysis of multiple phenotype enzyme-altered foci in rat hepatocarcinogenesis experiments: the multipath/multistage model*. Carcinogenesis, **16**(10): p. 2499-506, 1995.

18. Dewanji, A, Luebeck, EG and Moolgavkar, SH, *A biologically based model for the analysis of premalignant foci of arbitrary shape*. Math Biosci, **135**(1): p. 55-68, 1996.
19. Denes, J and Krewski, D, *An exact representation for the generating function for the Moolgavkar- Venzon-Knudson two-stage model of carcinogenesis with stochastic stem cell growth*. Math Biosci, **131**(2): p. 185-204, 1996.
20. Kopp-Schneider, A, Portier, C and Bannasch, P, *A model for hepatocarcinogenesis treating phenotypical changes in focal hepatocellular lesions as epigenetic events*. Math Biosci, **148**(2): p. 181-204, 1998.
21. Sherman, CD and Portier, CJ, *Multistage carcinogenesis Models*, in **Encyclopedia of Biostatistics**, P. Armitage and T. Colton, Editors. Wiley and Sons: Sussex, England. p. 2808-14, 1998.
22. Portier, CJ, Sherman, CD and Kopp-Schneider, A, *Calculating Tumor Incidence from Nonhomogenous Multistage Models of the Cancer Process*. Risk Analysis, (submitted).
23. Portier, CJ, Kopp-Schneider, A and Sherman, CD, *Calculating tumor incidence rates in stochastic models of carcinogenesis*. Math Biosci, **135**(2): p. 129-46, 1996.
24. El-Masri, H and Portier, CJ, *Replication potential of cells via the protein kinase C-MAPK pathway: Application of a mathematical model*. Bulletin of Mathematical Biology, **61**(2): p. 379-398, 1999.
25. Melnick, RL, Kohn, MC and Portier, CJ, *Implications for risk assessment of suggested nongenotoxic mechanisms of chemical carcinogenesis*. Environ Health Perspect, **104 Suppl 1**: p. 123-34, 1996.
26. Kimmel, CA, *Health assessment of exposure to developmental toxicants*. Dev Toxicol Environ Sci, **12**: p. 39-48, 1986.
27. Kavlock, RJ and Setzer, RW, *The road to embryologically based dose-response models*. Environ Health Perspect, **104 Suppl 1**: p. 107-21, 1996.
28. Kupper, LL, Portier, C, Hogan, MD and Yamamoto, E, *The impact of litter effects on dose-response modeling in teratology*. Biometrics, **42**(1): p. 85-98, 1986.
29. Ryan, L, *The use of generalized estimating equations for risk assessment in developmental toxicity*. Risk Anal, **12**(3): p. 439-47, 1992.
30. Ryan, L, *Using historical controls in the analysis of developmental toxicity data*. Biometrics, **49**(4): p. 1126-35, 1993.
31. Lefkopoulou, M and Ryan, L, *Global tests for multiple binary outcomes*. Biometrics, **49**(4): p. 975-88, 1993.
32. Carr, GJ and Portier, CJ, *An evaluation of some methods for fitting dose-response models to quantal-response developmental toxicology data*. Biometrics, **49**(3): p. 779-91, 1993.
33. Bieler, GS and Williams, RL, *Cluster sampling techniques in quantal response teratology and developmental toxicity studies*. Biometrics, **51**(2): p. 764-76, 1995.
34. George, EO and Bowman, D, *A full likelihood procedure for analysing exchangeable binary data*. Biometrics, **51**(2): p. 512-23, 1995.
35. Ten Have, TR, Kunselman, A and Zharichenko, E, *Accommodating negative intracluster correlation with a mixed effects logistic model for bivariate binary data*. J Biopharm Stat, **8**(1): p. 131-49, 1998.

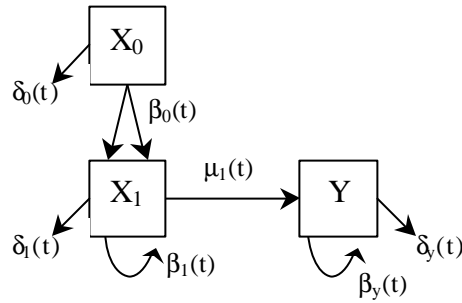
36. Leroux, BG, Leisenring, WM, Moolgavkar, SH and Faustman, EM, *A biologically-based dose-response model for developmental toxicology*. Risk Anal, **16**(4): p. 449-58, 1996.
37. Braun, M., **Differential Equations and Their Applications**, Applied Mathematical Sciences 15, Springer-Verlag, 1986.



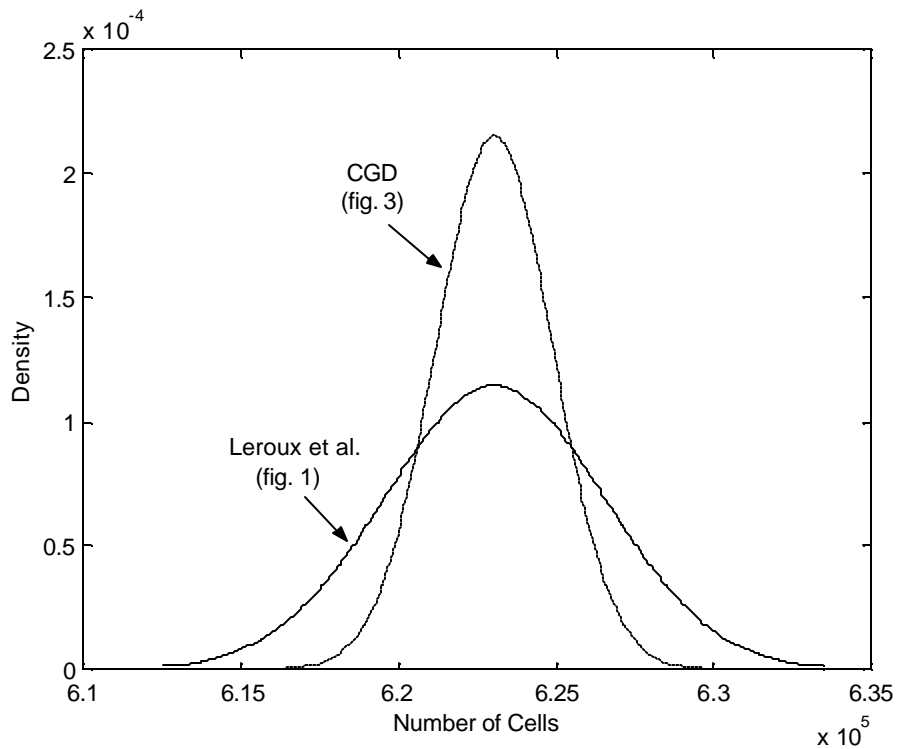
**Figure 1:** Developmental process of a tissue or organ as described by Leroux *et al.* <sup>[36]</sup>. The developmental parameters are birth ( $\mathbf{b}_i(t)$  for  $i=1,2$ ), death ( $\mathbf{d}_i(t)$  for  $i=1,2$ ) and transformation ( $\mathbf{m}(t)$ ).



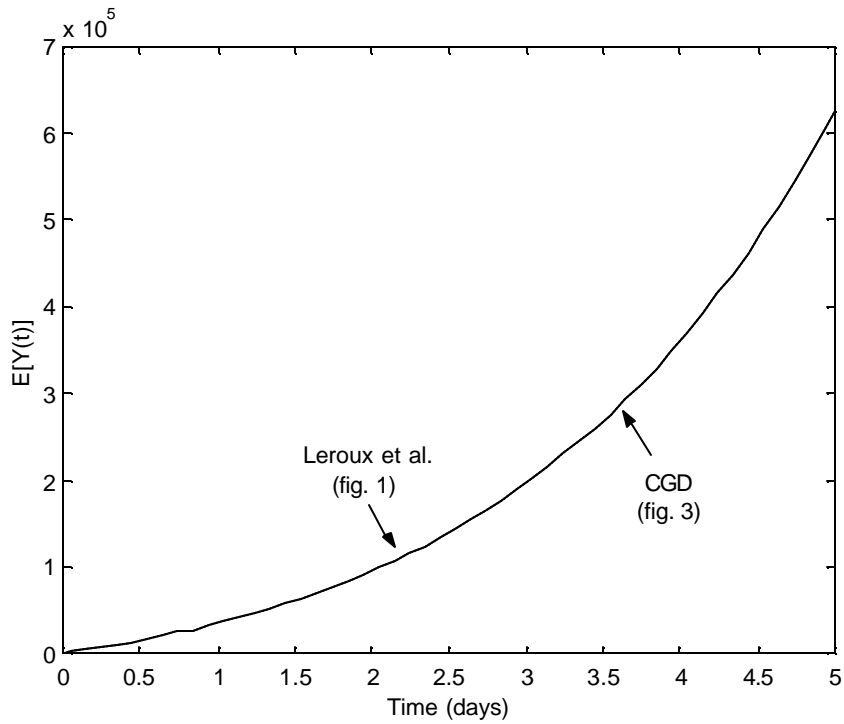
**Figure 2:** A general CGD system with  $k+2$  states. This system allows the cells to go through various levels of maturation before committing to the differentiation process.



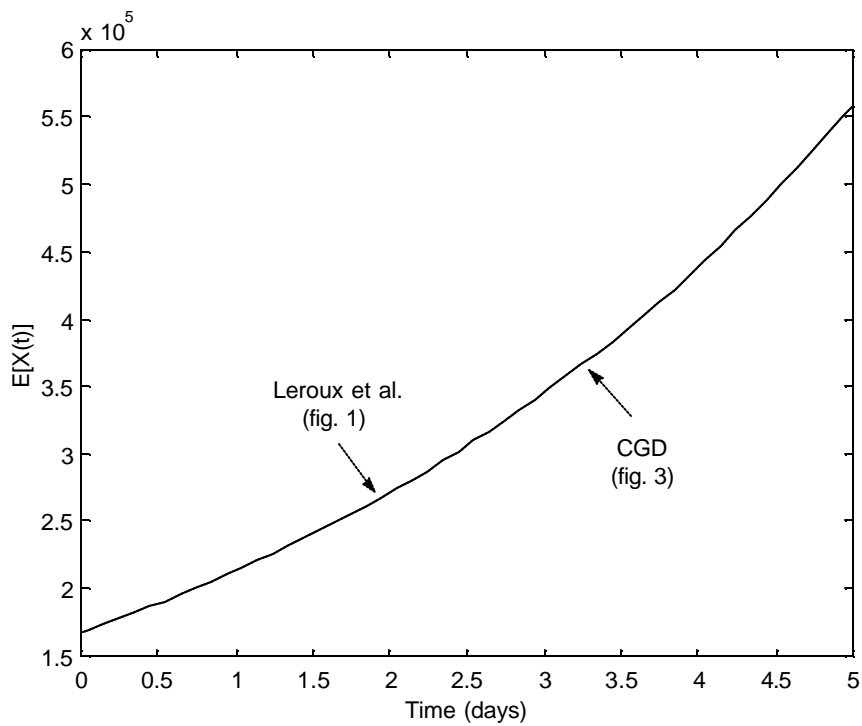
**Figure 3:** The specific case of the CGD model when  $k=1$ . In this example, the transformation rate for the  $X_0$  state is set equal to zero; thus, only cells from the  $X_1$  state transform to the  $Y$  state. Table 3 lists all of the developmental rates associated with this particular CGD model.



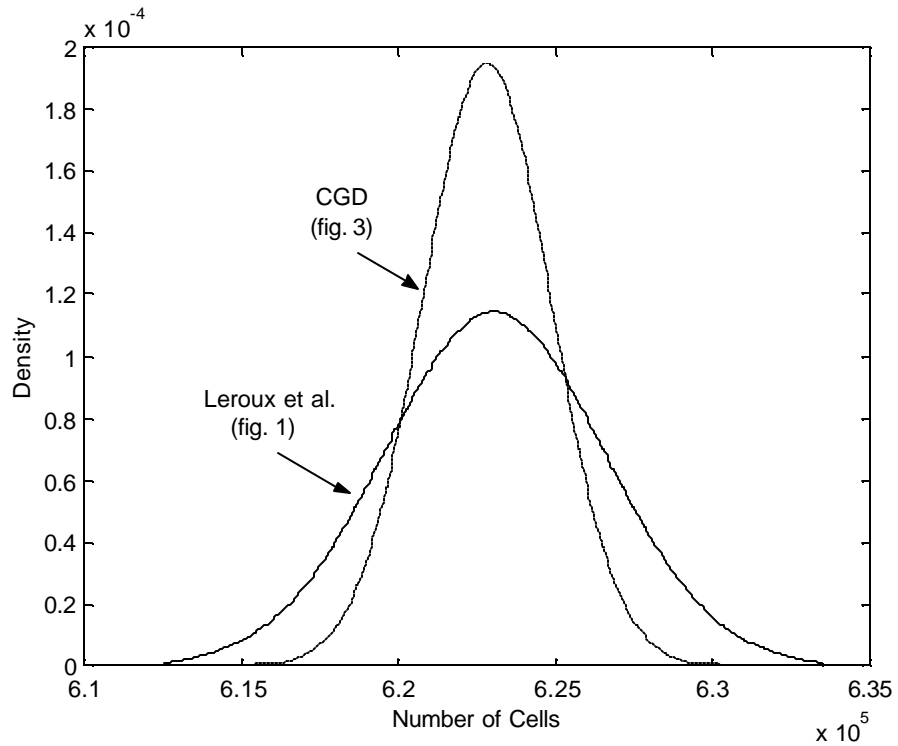
**Figure 4:** Distribution of type  $Y$  cells at  $t=5$  for the Leroux *et al.* model and the CGD model assuming a normal density with mean and variance from the first and second moments using the parameters in Table 3.



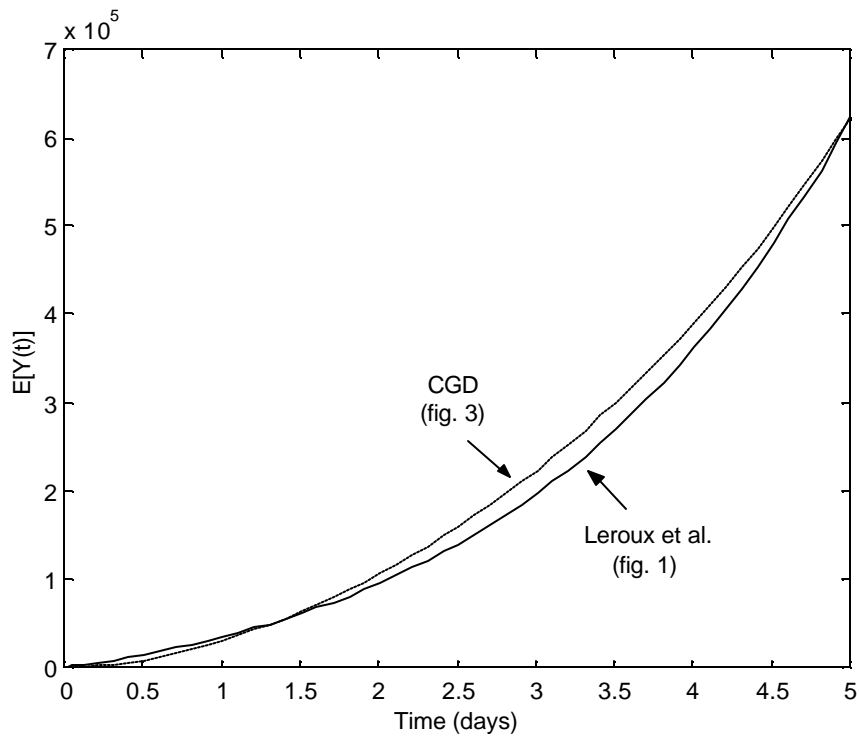
**Figure 5:** Expected number of type  $Y$  cells as predicted by the Leroux *et al.* model and the CGD model using parameters in Table 3.



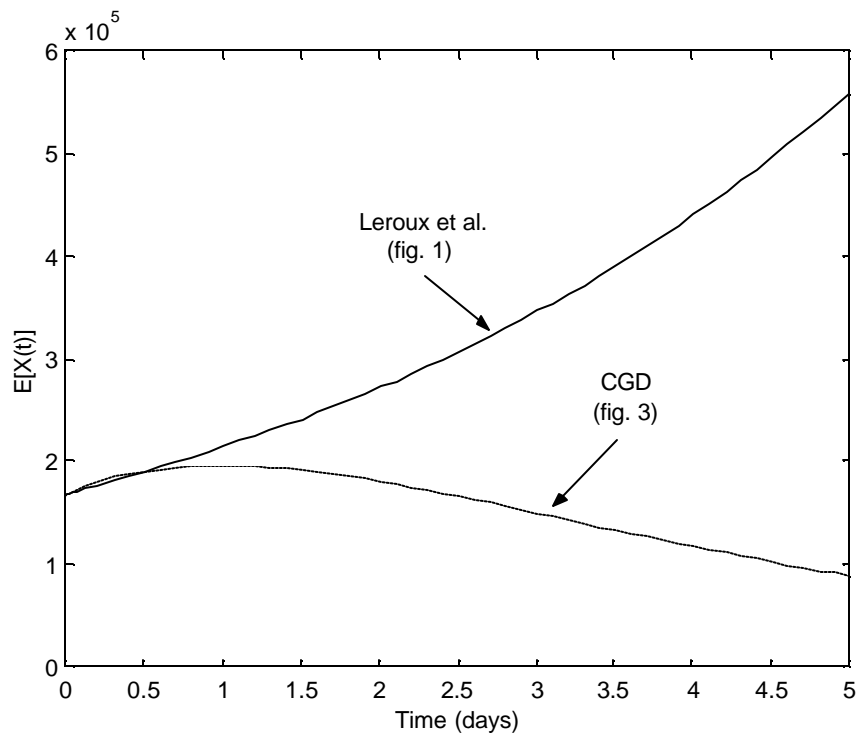
**Figure 6:** Expected number of type  $X$  cells as predicted by the Leroux *et al.* model and the CGD model ( $E[X_1(t) + X_2(t)]$ ) using parameters in Table 3.



**Figure 7:** Distribution of type  $Y$  cells at  $t=5$  for the Leroux *et al.* model and the CGD model assuming a normal density with mean and variance derived from the first and second moments using parameters in Table 4.



**Figure 8:** Expected number of type  $Y$  cells as predicted by the Leroux *et al.* model and the CGD model using parameters in Table 4.



**Figure 9:** Expected number of type X cells as predicted by the Leroux *et al.* model and the CGD model using parameters in Table 4.