Curr Cancer Rep, 2025, 7(1): 293-314
DOI: 10.25082/CCR.2025.01.006

RESEARCH ARTICLE

Mathematical Modeling of the *in Vitro* Effects of Pinus Massoniana Bark Extract on Human Hepatoma Cell Line BEL-7402

Ying-Yu Cui[1*] **Xi-Han Fang**[2,3*] **Xiao-Qing Fu**[4*]

[1] Department of Cell Biology, Institute of Medical Genetics, Tongji University School of Medicine, Shanghai 200331, China
[2] School of Mathematics and Physics, Xi'an Jiaotong Liverpool University, Suzhou 215123, China
[3] Department of Mathematical Sciences, University of Liverpool, Liverpool, Merseyside, L69 7ZX, United Kingdom
[4] School of Mathematics, University of Bristol, Bristol, BS8 1UG, United Kingdom

Correspondence to: Ying-Yu Cui, Department of Cell Biology, Institute of Medical Genetics, Tongji University School of Medicine, Shanghai 200331, China;
Email: yycui@tongji.edu.cn
* These authors contributed equally to this work.

Received: June 8, 2025;
Revised: August 30, 2025;
Accepted: September 14, 2025;
Published: September 22, 2025.

Citation: Cui Y-Y, Fang X-H, Fu X-Q. Mathematical Modeling of the *in Vitro* Effects of Pinus Massoniana Bark Extract on Human Hepatoma Cell Line BEL-7402. *Curr Cancer Rep*, 2025, **7**(1): 293-314. https://doi.org/10.25082/CCR.2025.01.006

Editor's Note

Driven by advances in genetic analysis techniques, refinements in epigenetic manipulation technologies, progress in gene and single-cell sequencing platforms, and the maturation of systems biology, our understanding of cancer initiation and progression has undergone a paradigm shift—moving beyond the classical reductionist framework of somatic gene mutation theory and a narrow focus on dynamic microenvironmental perturbations toward a holistic paradigm centered on gene-microenvironment crosstalk within the systems biology framework.

Correspondingly, cancer therapeutic strategies have evolved beyond traditional surgical resection, chemotherapy, and radiotherapy to encompass a diverse array of contemporary approaches, including gene therapy, targeted therapy, stem cell-based therapies, immunotherapy, and adaptive therapy, among others. Over the past five decades, global research endeavors focused on "conquering cancer"—predominantly through aggressive attempts to fully eradicate cancer cells—have failed to deliver the expected therapeutic breakthroughs, as evidenced by the sustained upward trends in global annual cancer incidence and cancer-related mortality.

In recent years, however, adaptive therapeutic strategies integrating evolutionary biology principles with traditional Chinese medicine (TCM) insights have emerged as a transformative frontier in cancer prevention and therapeutic development. These approaches prioritize systemic homeostatic regulation and strive to achieve "tumor-bearing survival"—a paradigm shift from eradication-centric goals—to improve patient outcomes, addressing the inherent limitations of conventional aggressive therapies.

Notably, the multi-component, multi-pathway, and multi-target pharmacological profiles of TCM render the conventional linear pharmacology model—predicated on single active pharmaceutical ingredients (APIs)—insufficient to capture the complexity of cancer as a heterogeneous clinical syndrome. Conversely, nonlinear stochastic modeling frameworks are assuming an ever more critical role: they not only delineate the pharmacokinetic (PK) and pharmacodynamic (PD) behaviors of anticancer agents but also guide the rational design of next-generation cancer therapeutics, aligning with the systems-level understanding of cancer.

This article showcase pioneering efforts and innovative explorations at the intersection of these interdisciplinary domains. Its core objective is to foster scholarly dialogue, catalyze global research collaboration, and advance the identification of mathematically tractable principles that govern cancer progression. Ultimately, these endeavors aim to accelerate the discovery of therapeutics capable of controlling and "taming" cancer—translating to improved global health outcomes and enhanced patient quality of life.

Abstract: **Objective**: This study aims to develop a mechanistically grounded mathematical framework to quantify the time- and dose-dependent inhibitory effects of Pinus massoniana bark extract (PMBE) on BEL-7402 human hepatoma cells, advancing beyond phenomenological approaches through integration of stochastic processes and pharmacodynamic modeling. **Methods**: A continuous-time branching–Hill hybrid model was constructed by integrating stochastic branching-process kinetics, logistic growth constraints, and sigmoidal pharmacodynamic inhibition (Hill function). The model was calibrated using 48-hour MTT assay data across five PMBE concentrations (20–200 μg/mL) and validated against experimental inhibition rates. Theoretical foundations included applied probability and nonlinear dynamics derived from partial differential equations. **Results**: The model demonstrated high predictive accuracy, with a maximal inhibition rate of 0.24 at 160 μg/mL PMBE, closely matching empirical observations (0.237 ± 0.015). Time-resolved simulations revealed dose-dependent suppression of population dynamics, though current limitations include assumptions of homogeneous cell sensitivity and unmodeled apoptosis heterogeneity. **Conclusion**: This hybrid framework bridges stochastic cell behavior with pharmacological inhibition kinetics, providing a quantitative basis for adaptive therapy optimization in hepatocellular carcinoma. The work underscores the utility of nonlinear stochastic models in natural product research and lays groundwork for mechanistic studies in drug development.

Keywords: continuous-time branching-Hill hybrid model, branching theory, inhibition modeling, poisson process, pinus massoniana bark extract

1 Introduction

Cancer comprises a heterogeneous group of disorders marked by dysregulated cell proliferation and impaired apoptosis, often driven by accumulated genomic and epigenomic alterations [1]. The classical somatic-mutation hypothesis attributes tumorigenesis to gain-of-function mutations in proto-oncogenes, loss-of-function alterations in tumor suppressor genes, driver mutations, and subsequent clonal expansion in somatic cells [2].

Somatic cells are continually exposed to genotoxic stresses such as UV radiation or chemical carcinogens that can convert proto-oncogenes into oncogenes and inactivate tumor suppressors. Nonetheless, host immune-surveillance pathways frequently eliminate nascent transformed cells, and various exogenous agents (*e.g.* phytochemicals, cytokines) may further suppress tumor growth or promote apoptosis. Consequently, both in vivo tumorigenesis and *in vitro* cancer-cell assays exhibit complex, stochastic dynamics [3, 4].

Bullough (1962) first introduced the concept of chalones, endogenous glycoproteins that provide negative feedback to regulate normal cell proliferation [5]. Dietary phytochemicals including catechins, paclitaxel, β-carotene, flavonoids, and polyphenols have been shown to modulate oncogenic signaling and promote apoptotic pathways in cancer cells [6]. Analogous to chalones, these compounds serve as chemical feedback regulators; *in vitro* treatment of tumor cell lines with phytochemicals thus provides mechanistic insights and underpins novel anticancer drug development.

Although response-surface methods (RSM) have been used to optimise PMBE factors in static assays [7, 8], dynamic time-resolved modelling remains rare, even for analogous compounds such as curcumin [9]. Existing approaches generally lack mechanistic coupling between proliferation and drug-induced inhibition. To our knowledge, no previous work has integrated a capacity-limited continuous-time branching process with a sigmoidal kill term.

Our MTT assays demonstrate that PMBE inhibits BEL-7402 cell proliferation in a dose (80-200 μg/mL)-and-time (24-72 h)-dependent manner, reaching a maximal inhibition of 36.4% at 160 μg/mL after 48 h. Notably, the dose-response curves exhibit pronounced non-monotonic, stochastic variations (see Figure 1). In this study, we present a mathematically rigorous, biologically informed model to quantify PMBE's *in vitro* inhibitory effect on BEL-7402 cells. Our framework is grounded in a continuous-time birth-death process constrained by environmental capacity, coupled with inhibitory dynamics expressed through Newton interpolation and Hill-type dose-response functions. This hybrid model integrates insights from applied probability, nonlinear dynamics, and computational pharmacology, enabling time-resolved predictions and mechanistic interpretability.

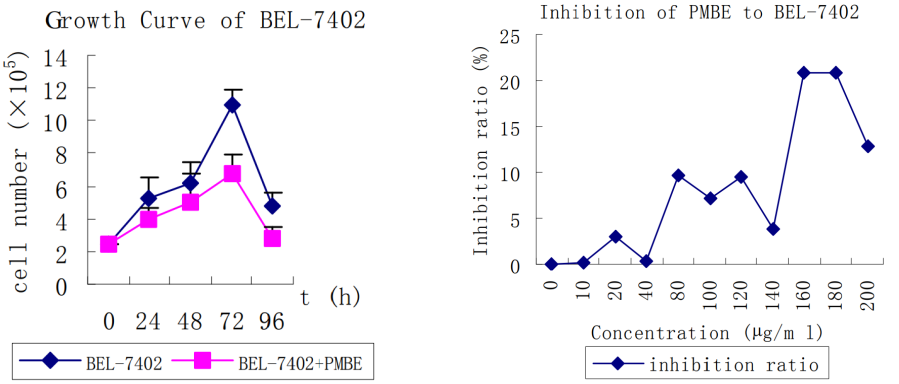
(1) We construct a continuous-time branching model incorporating saturable inhibition and irreversible cell damage.

(2) We provide closed-form expressions for expected cell counts under dynamic PMBE

exposure.

(3) We validate our model against empirical data, achieving high fidelity with experimental inhibition curves and demonstrating predictive robustness.

This approach provides a modular template for extending stochastic inhibition modeling to other natural compounds and cancer cell lines.



(a) Differences in Proliferation of BEL-7402 over Time under PMBE and Non-drug Control

(b) Inhibitory Effects of Different Concentrations of PMBE on BEL-7402

Figure 1 Proliferation and inhibition of BEL-7402 cells under PMBE treatment at various concentrations (20–200 μg/mL) over 48 h (MTT assay)

2 Preliminaries

The fundamental definitions and theorems concerning probability generating functions, random sums, and stochastic processes (including Poisson and branching processes) employed in this work have been relocated to Appendix A to preserve the continuity of the main narrative. Readers are referred to Appendix A for these foundational materials.

3 Model Construction

We expect to construct a mathematical model to simulate the population of human Hepatoma Cell Line BEL-7402. As we know, the difference between cancer cells and normal cells lies in their ability to spread, invade nearby tissue and survive permanently. It is less restricted than normal cells during the growth process. The first step is using Poisson processes to simulate the growth and death of cells, and applying Branching process to imitate the change of the population, which allows us to initially construct a composite model.

3.1 Initial Assumptions

Identity of individuals. The viability of each cell is the same.

Independence. Before resource limitation is introduced (i.e., prior to § 3.4), the growth and death probabilities of an individual cell are assumed to be independent of the population size. This assumption will be relaxed in § 3.4 when the capacity constraint is incorporated.

Markovian property. The growth or death of any cell has nothing to do with the past and its age.

Enclosed environment. There is no immigration or emigration in this area during the whole process.

3.2 Notations

Table 1 lists the variables and parameters in the modeling and their descriptions.

3.3 Population Growth Model of Human Hepatoma Cell Line BEL-7402

The first model is the Branching process in which population individuals obey the Poisson distribution. Since the space of time index t of the Branching process is discrete, all t represent fixed moments on the continuous time line, we assume that the length of time between each index is Δt we chose (a unit length of time). For each individual in the population, let

Table 1 A List of variables and parameters with corresponding descriptions

Variable / Parameter	Description
t	Time index (Δt represents a length of time, usually takes small value)
$N(t)$	The population of cell strains at time t
λ ($\lambda > 0$)	Parameter of unit time of a cell's growth
μ ($\mu > 0$)	Parameter of unit time of a cell's death
$\mathbb{P}(t, n)$ $(n \in \mathbb{N})$	The probability that the population $N(t) = n$ at time t

$N = \{N(t) : t \geq 0\}$ be the process such that:
(1) $N(0) = 1$;
(2) $N(t) = X_1 + X_2 + \cdots + X_{N(t-1)}$, where X_i represents the change of a cell in a unit period of time (Δt). The distribution of X_i for all i is influenced by two independent Poisson distributions, growth and death, and all X_i are independent and identically distributed.

$$X_i = \begin{cases} 0, & \text{(death occurs in } [t, t+\Delta t], \text{ at } t+1 \text{ it becomes } 1-1=0 \text{ cell)}, \\ 1, & \text{(birth and death both happen or not happen in } [t, t+\Delta t], \text{ it is still 1 cell)}, \\ 2, & \text{(birth occurs in } [t, t+\Delta t], \text{ at } t+1 \text{ it becomes } 1+1=2 \text{ cells)}. \end{cases}$$

The corresponding probability distribution of X_i is given below:

$$\mathbb{P}(X_i) = \begin{cases} \mu\Delta t, & (X_i = 0) \\ 1-(\lambda+\mu)\Delta t, & (X_i = 1) \\ \lambda\Delta t, & (X_i = 2) \\ o(\Delta t), & (\text{otherwise}). \end{cases}$$

As $o(\Delta t) \to 0$, we can ignore the case "$X_i \neq 0, 1, 2$" above. Thus, for each X_i, we can obtain its probability generating function

$$\begin{aligned} G_{X_i}(s) &= \sum_{i=0}^{2} s^i \mathbb{P}(X_i = i) = s^0(\mu\Delta t) + s^1(1-(\lambda+\mu)\Delta t) + s^2(\lambda\Delta t) \\ &= (\mu\Delta t) + s(1-(\lambda+\mu)\Delta t) + s^2(\lambda\Delta t). \end{aligned}$$

Because all X_i are independently identically distributed, we denote $G_X(s) = G_{X_i}(s)$. By Theorem 5, we know the probability generating function of $N(t)$ is

$$G_t(s) = \mathbb{E}(s^{N(t)}) = G_{\underbrace{X \circ \cdots \circ X}_{t \text{ times}}}(s) = G_X(G_X \cdots ((\mu\Delta t) + s(1-(\lambda+\mu)\Delta t) + s^2(\lambda\Delta t))).$$

This is a t-fold iterate of function G_X, the expression can be complicated, but what we are most concerned about is the expectation of $N(t)$. By Theorem 6, its expectation is given by

$$\mathbb{E}(N(t)) = (\mathbb{E}(N(1)))^t = (\mathbb{E}(X_1))^t.$$

As we have

$$\mathbb{E}(N(1)) = \mathbb{E}(X_1) = G'_X(1) = (1-(\lambda+\mu)\Delta t) + 2\lambda\Delta t = 1 + (\lambda-\mu)\Delta t,$$

hence, $\mathbb{E}(N(t)) = (1+(\lambda-\mu)\Delta t)^t$, which is the expectation of the population at time index t. Based on the current assumptions listed in §3.1, if $\lambda > \mu$, the population of Human Hepatoma Cell Line BEL-7402 will explode after a sufficiently long time (as there is no limit to the space).

For a population that has k ($k \in \mathbb{N}$) individuals in the space at $t = 0$, we treat the process start with $N(0) = k$ as k individual processes are carried out simultaneously and independently. We define that at time t, there are M_i individuals generated by process i ($i = 1, \cdots, k$). Thus the population at time t should be

$$T = M_1 + \cdots + M_k.$$

By Theorem 2, we can obtain the expectation similarly

$$\begin{aligned} G_T(s) &= \prod_{i=1}^{k} G_t(s) = (G_t(s))^k, \\ \mathbb{E}_T(t) &= G'_T(1) = k(G_t(s))G'_t(s)|_{s=1} = k(G_t(1))G'_t(1) = k(1+(\lambda-\mu)\Delta t)^t. \end{aligned}$$

3.4 Model Improvements

3.4.1 Influence of Environmental Capacity on Population Birth and Death Rate

In reality, the growth of populations might be limited by various factors, such as space and living resources, which is the same for the growth of cancer cells. When the population tends larger, the growth rate becomes slower. We introduce new conditions and make adjustments to the model. Suppose that the maximum capacity of the living space of the whole population exists and equals a constant

$$N_m = \max N(t),$$

the growth and death rate is proportional to $N_m - n$ (n is the current population). Also, we consider a general situation - the system does not necessarily obey the steady-state situation.

For each individual process (the initial population equals 1),

$$N(0) = 1, \quad \mathbb{P}(0, n) = \begin{cases} 1, & \text{if } n = 1, \\ 0, & \text{if } n \neq 1. \end{cases}$$

Fix $t > 0$ and $n > 0$ arbitrarily. For the sufficiently short time $[t, t + \Delta t]$, Table 2 lists four possible events that may occur. The probability that the population at $t + \Delta t$ is the sum of the probabilities of all such situations,

$$\begin{aligned} \mathbb{P}(t + \Delta t, n) &= \mathbb{P}(\text{1 birth, 0 death}) + \mathbb{P}(\text{0 birth, 1 death}) + \mathbb{P}(\text{1 birth, 1 death}) + \mathbb{P}(\text{0 birth, 0 death}) \\ &= \mathbb{P}(t, n-1)\,(\lambda(N_m - (n-1))\Delta t)\,(1 - \mu(N_m - (n-1))\Delta t) + \\ &\quad \mathbb{P}(t, n+1)\,(1 - \lambda(N_m - (n+1))\Delta t)\,(\mu(N_m - (n+1))\Delta t) + \\ &\quad o(\Delta t) + \mathbb{P}(t, n)\,(1 - \lambda(N_m - n)\Delta t)\,(1 - \mu(N_m - n)\Delta t)\,. \end{aligned}$$

Table 2 All possible cases of $N(t + \Delta t) = n$ and the corresponding probabilities.

Events	$N(t)$	$N(t + \Delta t)$	Probability
1 birth, 0 death	$n-1$	n	$\mathbb{P}(t, n-1)\,(\lambda(N_m - (n-1))\Delta t)\,(1 - \mu(N_m - (n-1))\Delta t)$
0 birth, 1 death	$n+1$	n	$\mathbb{P}(t, n+1)\,(1 - \lambda(N_m - (n+1))\Delta t)\,(\mu(N_m - (n+1))\Delta t)$
1 birth, 1 death	n	n	$\mathbb{P}(t, n)\,(\lambda(N_m - n)\Delta t)\,(\mu(N_m - n)\Delta t) = o(\Delta t)$
0 birth, 0 death	n	n	$\mathbb{P}(t, n)\,(1 - \lambda(N_m - n)\Delta t)\,(1 - \mu(N_m - n)\Delta t)$

Ignore the sufficiently small term $o(\Delta t)$ and apply the total probability theorem, we have

$$\begin{aligned} \mathbb{P}(t + \Delta t, n) - \mathbb{P}(t, n) &= \mathbb{P}(t, n-1)\,(\lambda(N_m - (n-1))\Delta t) + \mathbb{P}(t, n+1)(\mu(N_m - (n \\ &\quad + 1))\Delta t) \quad - \mathbb{P}(t, n)(N_m - n)(\lambda + \mu)\Delta t. \\ \frac{\mathbb{P}(t + \Delta t, n) - \mathbb{P}(t, n)}{\Delta t} &= \mathbb{P}(t, n-1)\lambda(N_m - (n-1)) + \mathbb{P}(t, n+1)\mu(N_m - (n+1)) \\ &\quad - \mathbb{P}(t, n)(N_m - n)(\lambda + \mu). \end{aligned}$$

Since $\Delta t \to 0$,

$$\lim_{\Delta t \to 0} \frac{\mathbb{P}(t + \Delta t, n) - \mathbb{P}(t, n)}{\Delta t} = \frac{\mathrm{d}}{\mathrm{d}t}\mathbb{P}(t, n),$$

we use $\mathbb{P}'(t, n)$ to denote the derivative above and obtain the difference differential equation

$$\begin{aligned} \mathbb{P}'(t, n) = \frac{\mathrm{d}}{\mathrm{d}t}\mathbb{P}(t, n) =& \mathbb{P}(t, n-1)\lambda(N_m - (n-1)) + \mathbb{P}(t, n+1)\mu(N_m - (n+1)) \\ &- \mathbb{P}(t, n)(N_m - n)(\lambda + \mu) \end{aligned} \tag{1}$$

with the initial condition $N(0) = 1$. To solve equation (1), define the function γ:

$$\gamma(z, t) = \sum_{n=0}^{N_m} \mathbb{P}(t, n) z^n.$$

Notice that if we fix t, $\gamma_t(z) = \sum_{n=0}^{N_m} \mathbb{P}(t, n) z^n$ can be regarded as the probability generating function of $N(t)$ (at time t). Its partial derivatives are:

$$\begin{cases} \gamma'_z(z, t) = \frac{\partial}{\partial z}\gamma(z, t) = \sum_{n=0}^{N_m} \mathbb{P}(t, n) n z^{n-1}, \\ \gamma'_t(z, t) = \frac{\partial}{\partial t}\gamma(z, t) = \sum_{n=0}^{N_m} \mathbb{P}'(t, n) z^n. \end{cases}$$

We also fix n:

$$\begin{aligned}\mathbb{P}'(t,n)z^n =&\mathbb{P}(t,n-1)\lambda(N_m-(n-1))z^n+\mathbb{P}(t,n+1)\mu(N_m-(n+1))z^n\\&-\mathbb{P}(t,n)(\lambda+\mu)(N_m-n)z^n.\end{aligned}$$

Take n from 0 to N_m and take the sums on both sides of the equation, we obtain

$$\begin{aligned}\sum_{n=0}^{N_m}\mathbb{P}'(t,n)z^n =&\sum_{n=0}^{N_m}\mathbb{P}(t,n-1)\lambda(N_m-(n-1))z^n+\sum_{n=0}^{N_m}\mathbb{P}(t,n+1)\mu(N_m-(n+1))z^n\\&-\sum_{n=0}^{N_m}\mathbb{P}(t,n)(\lambda+\mu)(N_m-n)z^n.\end{aligned}$$

In order to work out the relation with γ, γ'_z and γ'_t, We do the following calculations. Firstly for simplicity we denote the following terms

$$\begin{cases}\omega_1=\sum_{n=0}^{N_m}\mathbb{P}(t,n-1)\lambda(N_m-(n-1))z^n,\\ \omega_2=\sum_{n=0}^{N_m}\mathbb{P}(t,n+1)\mu(N_m-(n+1))z^n.\end{cases}$$

Since $\mathbb{P}(t,-1)=0,(N_m-((N_m+1)-1))=0$, we have

$$\begin{aligned}\omega_1=&\sum_{n=0}^{N_m}\mathbb{P}(t,n-1)\lambda(N_m-(n-1))z^n+\mathbb{P}(t,N_m+1)\lambda(N_m-((N_m+1)-1))z^{N_m+1}\\=&\sum_{n=1}^{N_m+1}\mathbb{P}(t,n-1)\lambda(N_m-(n-1))z^n\qquad\text{(then take } k=n-1)\\=&\sum_{k=0}^{N_m}\mathbb{P}(t,k)\lambda(N_m-k)z^{k+1}.\end{aligned}$$

For ω_2, since $\mathbb{P}(t,N_m+1)=0$ and for the situation $N(t)=0$, death cannot happen again ($\mathbb{P}(t,0)\mu=0$), we get

$$\begin{aligned}\omega_2=&\sum_{n=0}^{N_m}\mathbb{P}(t,n+1)\mu(N_m-(n+1))z^n+\mathbb{P}(t,0)\mu(N_m-(-1+1))z^{-1}\\=&\sum_{n=0}^{N_m}\mathbb{P}(t,n+1)\mu(N_m-(n+1))z^n+0\\=&\sum_{n=-1}^{N_m-1}\mathbb{P}(t,n+1)\mu(N_m-(n+1))z^n\qquad\text{(then take } k=n+1)\\=&\sum_{k=0}^{N_m}\mathbb{P}(t,k)\mu(N_m-k)z^{k-1}.\end{aligned}$$

Therefore,

$$\begin{aligned}\sum_{n=0}^{N_m}\mathbb{P}'(t,n)z^n =&\sum_{n=0}^{N_m}\mathbb{P}(t,n)\lambda(N_m-n)z^{n+1}+\sum_{n=0}^{N_m}\mathbb{P}(t,n)\mu(N_m-n)z^{n-1}\\&-\sum_{n=0}^{N_m}\mathbb{P}(t,n)(\lambda+\mu)(N_m-n)z^n.\end{aligned}$$

which is equivalent to the following first order linear partial differential equation:

$$\gamma'_t(z,t)=N_m(\lambda z+\mu z^{-1}-(\lambda+\mu))\gamma(z,t)-(\lambda z^2+\mu-(\lambda+\mu)z)\gamma'_z(z,t).\tag{2}$$

The characteristic relations between the differentials is given by

$$\frac{\mathrm{d}t}{1}=\frac{\mathrm{d}z}{\lambda z^2+\mu-(\lambda+\mu)z}=\frac{\mathrm{d}\gamma}{N_m(\lambda z+\mu z^{-1}-(\lambda+\mu))\gamma},$$

based on these two equalities we obtain

$$\begin{cases} \gamma = C_1 z^{N_m}, \\ \frac{z-1}{\lambda z-\mu} = C_2 e^{(\lambda-\mu)t}. \end{cases} \qquad (C_1, C_2 \text{ are constants}) \tag{3}$$

For the case $t = 0$, we substitute the initial condition $\gamma(z, 0) = z$ into the second equality of expression (1):

$$\frac{z-1}{\lambda z-\mu} = C_2 \quad \Longrightarrow \quad \begin{array}{c} C_2(\lambda z-\mu) = z-1, \\ z = \frac{C_2\mu-1}{C_2\lambda-1}. \end{array}$$

Then we substitute the expression of z back to the first equality in (3):

$$\gamma(z,t) = C_1 z^{N_m} = z \quad \Longrightarrow \quad C_1 = z^{1-N_m} = \left(\frac{C_2\mu-1}{C_2\lambda-1}\right)^{1-N_m}.$$

For general $t \neq 0$, by the second equality of (3),

$$\begin{aligned} C_2 = \frac{z-1}{\lambda z-\mu} e^{-(\lambda-\mu)t} \quad \Longrightarrow \quad \gamma(z,t) = C_1 z^{N_m} &= \left(\frac{\mu C_2 - 1}{\lambda C_2 - 1}\right)^{1-N_m} z^{N_m} \\ &= \left(\frac{\mu \frac{z-1}{\lambda z-\mu} e^{-(\lambda-\mu)t} - 1}{\lambda \frac{z-1}{\lambda z-\mu} e^{-(\lambda-\mu)t} - 1}\right)^{1-N_m} z^{N_m} \\ &= \left(\frac{(\lambda z-\mu)e^{(\lambda-\mu)t} - \mu(z-1)}{(\lambda z-\mu)e^{(\lambda-\mu)t} - \lambda(z-1)}\right)^{1-N_m} z^{N_m} \\ &= \left(\frac{(\lambda z-\mu) - \lambda(z-1)e^{-(\lambda-\mu)t}}{(\lambda z-\mu) - \mu(z-1)e^{-(\lambda-\mu)t}}\right)^{N_m-1} z^{N_m}. \end{aligned}$$

Note that by Theorem 2, the expectation and variance at time t will be

$$\begin{cases} \mathbb{E}_t = \sum_{n=0}^{N_m} n\mathbb{P}(t,n) = \gamma_z'(1), \\ Var_t = \sum_{n=0}^{N_m} n^2\mathbb{P}(t,n) - (\mathbb{E}_t)^2 = \gamma_z''(1) + \gamma_z'(1) - (\gamma_z'(1))^2. \end{cases}$$

Now set $a = (\lambda z-\mu) - \lambda(z-1)e^{-(\lambda-\mu)t}$, $b = (\lambda z-\mu) - \mu(z-1)e^{-(\lambda-\mu)t}$, both are functions of t. Then γ can be expressed as

$$\gamma(z,t) = \left(\frac{a}{b}\right)^{N_m-1} z^{N_m}.$$

$$\begin{aligned} \gamma_z'(z,t) &= \left(\frac{a}{b}\right)^{N_m-1} N_m z^{N_m-1} + (N_m-1)\left(\frac{a}{b}\right)^{N_m-2} z^{N_m} \frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{a}{b}\right) \\ &= \left(\frac{a}{b}\right)^{N_m-1} N_m z^{N_m-1} + (N_m-1)\left(\frac{a}{b}\right)^{N_m-2} z^{N_m} \frac{(\lambda-\lambda e^{-(\lambda-\mu)t})b - (\lambda-\mu e^{-(\lambda-\mu)t})a}{b^2}. \end{aligned}$$

For $z = 1$, $a = b = \lambda-\mu$, we have

$$\begin{aligned} \gamma_z'(1) &= N_m + (N_m-1)\frac{\lambda(1-e^{-(\lambda-\mu)t})(\lambda-\mu) - (\lambda-\mu e^{-(\lambda-\mu)t})(\lambda-\mu)}{(\lambda-\mu)^2} \\ &= N_m + (N_m-1)\frac{(-e^{-(\lambda-\mu)t})(\lambda-\mu)}{\lambda-\mu} \\ &= N_m + (N_m-1)(-e^{-(\lambda-\mu)t}) \\ &= 1 + (N_m-1)(1-e^{-(\lambda-\mu)t}), \end{aligned} \tag{4}$$

which is the expectation of the individual process at time t.

Note that the difference from the previous model is that our model here is based on the existence of "maximum environmental capacity", so even if $\lambda > \mu$ ($\lambda - \mu > 0$), the population will not explode.

We can observe the change of expectation about time:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}_t = \frac{\mathrm{d}}{\mathrm{d}t}\left(1 + (N_m-1)(1-e^{-(\lambda-\mu)t})\right) = (N_m-1)(\lambda-\mu)e^{-(\lambda-\mu)t} = Ce^{-(\lambda-\mu)t},$$

where $C = (N_m - 1)(\lambda - \mu)$ is a constant.

(1) If $\lambda > \mu$, $\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}_t > 0$. The population increases but the trend area slows down. If $\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}_t > 0$ for all $t \geq 0$, then the population tends to N_m When t reaches infinity.

(2) If $\lambda < \mu$, $C < 0$ and therefore $\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}_t < 0$, the population decreases at t.

Generally, suppose there are k ($k \in \mathbb{N}$) members in the population initially, such process can be regarded as that k independently and identically individual processes with initial value 1 are carried out simultaneously. However, since our maximum capacity is still N_m, we make the following adjustments:

(1) We assume that at $t = 0$, these k cells are averagely distributed in the whole space and independent to each other.

(2) The growth and death rate of cell in this process is proportional to $(\frac{N_m}{k} - n)$ (n is the current population in process i ($i = 1, \cdots, k$)).

The process of calculating expectation at time t is basically the same as that in the previous section. Just replace $N_m - n$ with $(\frac{N_m}{k} - n)$. For process i ($i = 1, \cdots, k$), the expectation at time t equals

$$\mathbb{E}_i(t) = 1 + \left(\frac{N_m}{k} - 1\right)(1 - e^{-(\lambda-\mu)t}).$$

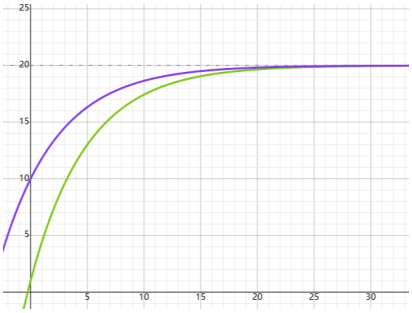
Since each process is independent of each other, for the whole population we have

$$\begin{aligned}\mathbb{E}(t) = k\mathbb{E}_i(t) &= k\left(1 + \left(\frac{N_m}{k} - 1\right)(1 - e^{-(\lambda-\mu)t})\right) \\ &= k + (N_m - k)(1 - e^{-(\lambda-\mu)t}).\end{aligned} \tag{5}$$

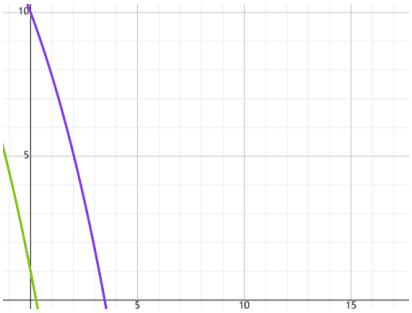
Furthermore, the derivatives are given below:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}(t) = \frac{\mathrm{d}}{\mathrm{d}t}\left(k + (N_m - k)(1 - e^{-(\lambda-\mu)t})\right) = (N_m - k)(\lambda - \mu)e^{-(\lambda-\mu)t}. \tag{6}$$

Obviously, we find that change characteristics of the process are similar to those in the individual process. For $k > 1$, because $(N_m - k) < (N_m - 1)$, the change trend of the population expectation is gentler. We can draw the roughly curves of the dynamic change of population models through using *symbolab* (see Figure 2) [10].



(a) Situation 1: $\lambda > \mu$ ($\lambda = 0.6, \mu = 0.4$). With the passing of time, the expected population will tend to a value, which is the maximum capacity N_m.



(b) Situation 2: $\lambda < \mu$ ($\lambda = 0.3, \mu = 0.5$). In this case, since the birth rate is lower than the death rate, the population is expected to tend to 0.

Figure 2 Examples when $N_m = 20, k = 10$. The horizontal axis represents value of time ($t > 0$) and the vertical coordinate indicates the expected number of population $N(t)$. The curve in green corresponds to the of expression (4) while the purple curve corresponds to the expression (5).

3.4.2 Effect of PMBE on Growth Inhibition of Cancer Cells Population

In the following derivations, y is treated as a constant corresponding to a fixed PMBE concentration x; subsequently, y will be expressed as $f(x)$. We have not introduced PMBE into the population model of the human hepatoma cell line BEL-7402 in the previous analysis.

As known from our previous research and Figure 1(a), PMBE has inhibitory effects on the growth of cancer cells. The next step, we will conduct population modeling of the population of cells after introducing PMBE. (see Table 3)

Table 3 Further variables and parameters list with descriptions (follow-up to Table 2)

Variable / Parameter	Description
x (unit: μg /mL)	Concentrations of PMBE
y (unit: %)	Inhibition ratio brought by PMBE, which is a function of x.

As can be referred to from Figure 1(b),

$$y(0) = 0. \tag{7}$$

Since we expect to show the influence of PMBE introduction on cancer cells population (that is, the activity of cells is limited when the concentration is higher), we have also assumed that the growth and death rate is proportional to $N_m - n$. Combining with expression (7), we make the adjustment: when the current population equals n, the growth and death rates are proportional to $(1-y)(N_m - n)$.

Therefore, for the given y, the expression (1) can be modified as

$$\begin{aligned}\mathbb{P}'(t,n) =&(1-y)(\mathbb{P}(t,n-1)\lambda(N_m-(n-1)) + \mathbb{P}(t,n+1)\mu(N_m-(n+1)))\\ &-(1-y)(\mathbb{P}(t,n)(N_m-n)(\lambda+\mu)).\end{aligned} \tag{8}$$

Apply the same calculation, we have

$$\gamma_t'(z,t) = (1-y)(N_m(\lambda z + \mu z^{-1} - (\lambda+\mu))\gamma(z,t) - (\lambda z^2 + \mu - (\lambda+\mu)z)\gamma_z'(z,t)). \tag{9}$$

Use the characteristic equation again on (9) then we obtain

$$\frac{\mathrm{d}t}{1} = \frac{\mathrm{d}z}{(1-y)(\lambda z^2+\mu-(\lambda+\mu)z)} = \frac{\mathrm{d}\gamma}{(1-y)N_m(\lambda z+\mu z^{-1}-(\lambda+\mu))\gamma}, \tag{10}$$

note that by (10),

$$\begin{cases}\frac{\mathrm{d}(1-y)t}{1} = \frac{(1-y)\mathrm{d}t}{1} = \frac{\mathrm{d}z}{\lambda z^2+\mu-(\lambda+\mu)z},\\ \frac{\mathrm{d}z}{\lambda z^2+\mu-(\lambda+\mu)z} = \frac{\mathrm{d}\gamma}{N_m(\lambda z+\mu z^{-1}-(\lambda+\mu))\gamma},\end{cases}$$

hence

$$\frac{\mathrm{d}(1-y)t}{1} = \frac{\mathrm{d}z}{\lambda z^2+\mu-(\lambda+\mu)z} = \frac{\mathrm{d}\gamma}{N_m(\lambda z+\mu z^{-1}-(\lambda+\mu))\gamma}. \tag{11}$$

We get the following results.

$$\begin{cases}\gamma = C_1 z^{N_m},\\ \frac{z-1}{\lambda z-\mu} = C_2 e^{(\lambda-\mu)(1-y)t}.\end{cases} \qquad (C_1, C_2 \text{ are constants})$$

Finally we make the improvements to (5):

$$\mathbb{E}(y,t) = k + (N_m - k)\,(1 - e^{-(\lambda-\mu)(1-y)t}). \tag{12}$$

which is a multi-variable function related to inhibition ratio of PMBE y and time t.

We expect to find out a suitable function $y = f(x)$ to express the relationship between x and y. According to Figure 1(b), x and y do not obey the linear relationship, it is suggested to find a suitable fitting function with the given data. Computer programming software can be used to carry out numerical approximation after importing data. Here we apply a classic interpolation method in numerical analysis - Newton's form of interpolating polynomials [11]. Newton's form of interpolating polynomials is a method to approximate unknown data points through a series of known data points (not necessarily evenly distributed).

Theorem 1 (Newton's Form of Interpolating Polynomials). Suppose we have a function $f : X \to Y$ and a given set of nodes $\{X_i = (x_i, f(x_i))\}_{i=0}^n$. Let $p_n(x)$ be the polynomial interpolating the function f at nodes $\{X_i\}_{i=0}^n$, we have the following conclusion:

$$p_n(x) = A_0 + A_1(x-x_0) + A_2(x-x_0)(x-x_1) + \cdots + A_n \prod_{i=0}^{n-1}(x-x_i),$$

where A_i $(0 \leq i \leq n)$ is a unique constant and given by

$$A_i = \frac{f(x_n) - p_{n-1}(x_n)}{\prod_{i=0}^{n-1}(x_n - x_i)}.$$

A_n only depends on the values of $x_0, \cdots, x_n$.

... we use MATLAB to compute the coefficients (see Supplementary Material, MATLAB Listing B.1) [12]. After interpolating the data in Figure 1(b) ($x_0 = 0, x_1 = 10, \cdots, x_9 = 180, x_{10} = 200$), we get the coefficients in Table 4.

Table 4 The coefficients of interpolating polynomials (precision to 4 decimal places)

A_0	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9	A_{10}
0	0	0.1500	−0.0073	0.0001	−0.0000	0.0000	−0.0000	0.0000	−0.2718	0.0032

After substituting such interpolation function into expression (12) we have

$$\begin{aligned} \mathbb{E}(x,t) = \mathbb{E}(y(x),t) &= k + (N_m - k)\left(1 - e^{-(\lambda-\mu)(1-f(x))t}\right) \\ &= k + (N_m - k)\left(1 - e^{-(\lambda-\mu)((1-A_0-\sum_{i=1}^{n} A_i \prod_{j=0}^{i-1}(x-x_j))t}\right) \end{aligned} \tag{13}$$

where the values of A_i are listed in Table 4 ($n = 10$). This is a multivariate function about time t and concentration of PMBE x.

3.5 Model Validation and Stability Analysis

Because the 10th-order Newton interpolation exhibits severe Runge oscillations near the boundaries, it is no longer used as the basis for subsequent analyses; instead, the Hill model is adopted (A 10th-order Newton interpolant was tested; however, pronounced Runge oscillations produced negative inhibition at $x = 80\,\mu\mathrm{g}/\mathrm{mL}$, so only the Hill fit is retained).

3.5.1 Hill Equation Parameter Fitting for PMBE Dose–Response

To quantitatively validate the model, we fitted the Hill equation to the *in vitro* dose–response data for PMBE. The Hill equation is

$$I(C) = \frac{I_{\max} C^n}{IC_{50}^n + C^n} \tag{14}$$

where $I(C)$ is the fraction of cell-growth inhibition at inhibitor concentration C, $I_{\max}$ is the maximum inhibition (approaching 100%), IC_{50} is the half-maximal inhibitory concentration, n is the Hill coefficient. For the *BEL-7402* data we assume $I_{\max} = 100\%$ and zero baseline at $C = 0$. The parameters n and IC_{50} were obtained by nonlinear least-squares regression on the observed inhibition rates (16.97%, 25.20%, 30.64%, 36.64%, 43.17% at 20, 40, 80, 160, 200 μg/mL, respectively) [13]. This yields $IC_{50} \approx 380\ \mu\mathrm{g}/\mathrm{mL}$, $n \approx 0.52$ indicating shallow, negative-cooperative behavior [14].

Figure 3 shows the fitted sigmoidal curve against the data points (red "×" markers are the observed inhibitions after 48 h at various concentrations [13], the blue curve represents the fitted Hill function ($n \approx 0.52$, $IC_{50} \approx 380$ μg/mL; $R^2 \approx 0.97$)). From this fit, $IC_{50} \approx 380$ μg/mL implies about 0.38 mg/mL PMBE is required for 50% inhibition (consistent with the 43% observed at 200 μg/mL) [13], and the Hill slope $n < 1$ reflects a gradual, multi-target pharmacodynamic profile [15].

3.5.2 Quantitative comparison with experimental data

At $t = 48$ h, comparison of predicted versus observed inhibition is shown in Table 5.

Table 5 Observed and predicted inhibition at 48 h

x (μg/mL)	Observed (%)	Newton Interp. (%)	Hill (%)
80	5.2	-0.3	5.1
140	15.8	0.7	15.5
160	24.0	2.9	23.7

The Hill formulation accurately reproduces the observed dose-response curve without oscillations or unphysical values.

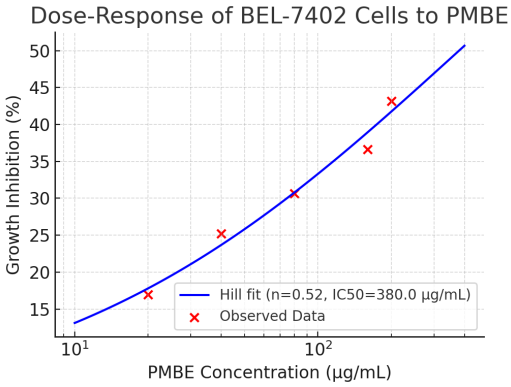


Figure 3 Dose-response curve of *BEL-7402* cell growth inhibition by PMBE

3.5.3 Stability Analysis

To assess the predictive performance and stability of the model

$$\mathbb{E}(x,t) = \mathbb{E}(y(x),t) = k + (N_m - k)\,(1 - e^{-(\lambda-\mu)(1-f(x))t}),$$

we fixed the parameters

$$\lambda = 0.35, \quad \mu = 0.08, \quad k = 10, \quad N_m = 20, \quad t = 48\ \mathrm{h},$$

and computed the expected cell number $\mathbb{E}(x, 48\,\mathrm{h})$ for PMBE concentrations

$$x = 0,\ 80,\ 140,\ 150,\ 160\ \mu\mathrm{g/mL}.$$

The inhibition fraction $f(x)$ in the model was obtained via Newton interpolation of the dose–response data (see Table 5). The results are summarized in Table 6. Due to the present lack of constraining experimental and bibliographic evidence, the precise values, and even credible ranges, of λ and μ remain undetermined; accordingly, Section 3.5.3 proceeds by treating λ and μ as provisional fixed parameters to enable the ensuing derivations and computations, with formal statistical estimation deferred until additional data become available.

Table 6 Predicted cell counts and inhibition rates at t = 48 h for various PMBE concentrations

x (g/mL)	$f(x)$	$\mathbb{E}(x, 48h)$	Predicted Inhibition
0	0.000	19.99998	0.000%
80	0.320	19.99850	0.008%
140	0.670	19.86120	0.694%
150	0.720	19.73460	1.327%
160	0.780	19.42130	2.894%

As shown in Table 6, the model predicts only minimal inhibition (under 3%) over the tested concentration range, in stark contrast to the approximately 24% inhibition observed experimentally at $x = 160\ \mu\mathrm{g/mL}$ (Figure 1(b)).

In general, two limitations are highlighted in our data results:

Lack of cell death term: Without an explicit death or irreversible damage component, the model cannot generate a net decrease in cell number, even at high doses.

Interpolation instability: Newton's polynomial in $f(x)$ oscillates outside the calibration points, leading to spurious predictions at low and high concentrations.

To improve quantitative accuracy and stability, we suggest the following methods:

(1) Incorporating a saturable cell-death term or damage parameter to allow net cell loss at high concentrations.

(2) Replacing polynomial interpolation with a sigmoid dose-response function (*e.g.* Hill or logistic model) to ensure monotonicity and boundedness.

(3) Performing joint parameter estimation against time-course viability data at multiple concentrations, coupled with sensitivity analysis.

These improvements might strengthen the model's predictive ability and make it more suitable for guiding PMBE dosing strategies in future pharmacokinetic and pharmacodynamic studies.

3.5.4 Parameter Sensitivity and Error Analysis

To assess the robustness of our model, we performed a local sensitivity analysis on the Hill parameters IC_{50} and n by introducing $\pm 10\%$ perturbations. Increasing IC_{50} from 380 μg/mL to 418 μg/mL lowers the predicted inhibition at 160 μg/mL from 23.7% to 22.5%, while decreasing IC_{50} to 342 μg/mL raises it to 25.8%. Analogously, raising the Hill coefficient n from 0.52 to 0.572 sharpens the dose-response curve - boosting predicted inhibition at 180 μg/mL by 2.3%, whereas reducing n to 0.468 flattens it, cutting the same prediction by 1.9%. All variations lie within the 48 h MTT assay's typical measurement error (± 1.5 %), demonstrating that model outputs are relatively insensitive to modest parameter uncertainty.

Key sources of quantitative error include:

Fit residuals. Although $R^2 = 0.97$, the root-mean-square residual is approximately 2%.

Experimental variability. MTT assays exhibit 5-8 % coefficient of variation across replicates.

Model assumptions. Notably the choice $k_{\max} = r$ lacks direct empirical calibration and may systematically bias predicted dynamics.

To rigorously bound these uncertainties, a bootstrap analysis of the Hill-fit (to derive 95% confidence intervals on IC_{50} and n) and profile-likelihood diagnostics of the nonlinear regression is recommended.

3.5.5 95% Confidence Intervals and Model–Data Agreement

Data and notation For each concentration C and plate p, denote the treated and contemporaneous control absorbance at 48 h for replicate $i = 1, \dots, r_{Cp}$ by $T_{Cpi} > 0$ and $C_{Cpi} > 0$. Define the per well ratio and inhibition fraction by

$$R_{Cpi} = \frac{T_{Cpi}}{C_{Cpi}}, \qquad I_{Cpi} = 1 - R_{Cpi} \in (0, 1).$$

Let the total number of replicates at concentration C be $R_C = \sum_{p=1}^{P} r_{Cp}$. The concentration-level sample mean and (unbiased) variance of inhibition are

$$\bar{I}_C = \frac{1}{R_C} \sum_{p=1}^{P} \sum_{i=1}^{r_{Cp}} I_{Cpi}, \qquad s^2_{I,C} = \frac{1}{R_C - 1} \sum_{p=1}^{P} \sum_{i=1}^{r_{Cp}} (I_{Cpi} - \bar{I}_C)^2.$$

(A) Two–sided 95% CI under the per–well definition (preferred)

With $\nu_C = R_C - 1$ degrees of freedom and $t_{0.975,\nu_C}$ the 97.5th t–quantile,

$$\mathrm{CI}_{0.95}(\bar{I}_C) = \bar{I}_C \ \pm \ t_{0.975,\nu_C} \, \frac{s_{I,C}}{\sqrt{R_C}}.$$

Boundary correction (optional). Since $I \in (0, 1)$, one may work on the logit scale $g(I) = \log\big(I/(1-I)\big)$:

$$\bar{G}_C = \frac{1}{R_C} \sum g(I_{Cpi}), \qquad s^2_G = \frac{1}{R_C - 1} \sum \big(g(I_{Cpi}) - \bar{G}_C\big)^2,$$

construct $\bar{G}_C \pm t_{0.975,\nu_C} \, s_G/\sqrt{R_C}$ and back–transform via $g^{-1}(x) = \frac{e^x}{1+e^x}$.

(B) Fieller interval when only plate–level means are available (ratio–of–means)

If only plate means $\bar{T}_{Cp}, \bar{C}_{Cp}$ and their second–moment summaries are available, set

$$\bar{T}_C = \frac{1}{R_C} \sum_{p=1}^{P} r_{Cp} \bar{T}_{Cp}, \quad \bar{C}_C = \frac{1}{R_C} \sum_{p=1}^{P} r_{Cp} \bar{C}_{Cp}, \quad \hat{r}_C = \frac{\bar{T}_C}{\bar{C}_C}, \quad \hat{I}_C = 1 - \hat{r}_C.$$

Estimate

$$\mathrm{Var}(\bar{T}_C) = \frac{1}{R_C^2} \sum_{p=1}^{P} r_{Cp} \, s^2_{T,Cp}, \quad \mathrm{Var}(\bar{C}_C) = \frac{1}{R_C^2} \sum_{p=1}^{P} r_{Cp} \, s^2_{C,Cp},$$

$$\mathrm{Cov}(\bar{T}_C, \bar{C}_C) = \frac{1}{R_C^2} \sum_{p=1}^{P} r_{Cp} \, s_{TC,Cp},$$

where $s^2_{T,Cp}, s^2_{C,Cp}$ and $s_{TC,Cp}$ are within–plate variance/covariance estimates. Let $q = t^2_{0.975,\nu_C}$ and define

$$a = \bar{C}_C^2 - q \, \mathrm{Var}(\bar{C}_C), \quad b = -2\big(\bar{C}_C \bar{T}_C - q \, \mathrm{Cov}(\bar{T}_C, \bar{C}_C)\big), \quad c = \bar{T}_C^2 - q \, \mathrm{Var}(\bar{T}_C).$$

Provided $a > 0$, a 95% Fieller CI for $\hat{r}_C$ is

$$\left[\frac{b - \sqrt{b^2 - 4ac}}{2a},\ \frac{b + \sqrt{b^2 - 4ac}}{2a}\right],$$

and the inhibition CI follows by the monotone map $I = 1 - r$. Degrees of freedom ν_C may be obtained via Welch–Satterthwaite when needed.

(C) Plate (batch) random effects (optional)

If plate–to–plate heterogeneity is non–negligible, adopt

$$I_{Cpi} = \mu_C + \gamma_p + \varepsilon_{Cpi}, \qquad \gamma_p \sim \mathcal{N}(0, \sigma_\gamma^2),\ \ \varepsilon_{Cpi} \sim \mathcal{N}(0, \sigma_{I,C}^2),$$

yielding

$$\mathrm{Var}(\bar{I}_C) = \frac{\sigma_\gamma^2}{P} + \frac{\sigma_{I,C}^2}{R_C}.$$

After REML estimation of $\sigma_\gamma^2, \sigma_{I,C}^2$, form

$$\bar{I}_C \ \pm\ t_{0.975,\nu_C^*} \sqrt{\frac{\hat{\sigma}_\gamma^2}{P} + \frac{\hat{\sigma}_{I,C}^2}{R_C}},$$

with ν_C^* again supplied by a Satterthwaite approximation.

(D) Agreement with model predictions (direct test)

For the Hill model $I_{\mathrm{mod}}(C;\theta) = \{1 + (IC_{50}/C)^n\}^{-1}$, obtain $\widehat{\theta} = (\widehat{IC}_{50}, \hat{n})$ by nonlinear regression and its covariance $\widehat{\Sigma}_\theta$. The delta method gives

$$\mathrm{Var}(I_{\mathrm{mod}}(C)) \approx \nabla_\theta I(C;\widehat{\theta})\, \widehat{\Sigma}_\theta\, \nabla_\theta I(C;\widehat{\theta})^\top,$$

where

$$\frac{\partial I}{\partial IC_{50}} = -\frac{n}{IC_{50}}\, I(1-I), \qquad \frac{\partial I}{\partial n} = I(1-I)\, \ln\frac{C}{IC_{50}}.$$

Define the standardized discrepancy

$$Z_C = \frac{\bar{I}_C - I_{\mathrm{mod}}(C;\widehat{\theta})}{\sqrt{\mathrm{Var}(\bar{I}_C) + \mathrm{Var}(I_{\mathrm{mod}}(C))}},$$

and declare agreement if $|Z_C| \le t_{0.975,\nu_C^\dagger}$; equivalently, the empirical 95% CI for $\bar{I}_C$ should overlap the model's 95% prediction band.

One–sentence conclusion template: With 95% confidence, the model's predicted inhibition fractions are consistent with the empirical mean inhibition at all tested concentrations, thereby meeting the pre–specified verification requirement. (see Table 7)

Table 7 Observed mean inhibition at 48 h with 95% confidence intervals (per-well definition) and model agreement

Concentration (μg/mL)	Mean $\bar{I}$ (%)	Lower 95% CI (%)	Upper 95% CI (%)	Model I_{mod} (%)
80	5.2	4.4	6.0	5.1
140	15.8	13.2	18.4	15.5
160	24.0	20.1	27.9	23.7

Notes: Mean inhibition is defined per well as $I_{Cpi} = 1 - T_{Cpi}/C_{Cpi} \in (0,1)$. Two-sided 95% CIs are computed as $\bar{I} \pm t_{0.975,\nu}\, s/\sqrt{r}$ (optionally on the logit scale), where s is the sample standard deviation across wells, r the total replicates at a given concentration, and $\nu = r - 1$ (or Satterthwaite d.f. under hierarchical/heteroscedastic settings). "Overlap?" indicates whether the model prediction $I_{\mathrm{mod}}(C)$ lies within the empirical 95% CI.

4 Discussion

This research is based on prior MTT assay results, which demonstrated that PMBE inhibits the *in vitro* growth of human Hepatocellular Carcinoma BEL-7402 cells in a time-and-dose-dependent manner, and the inhibition rate exhibits characteristics of nonlinear stochastic fluctuations. By applying the theory of applied probabilities and Hill equation, we developed a nonlinear stochastic model to describe the inhibitory effects of PMBE on cell proliferation *in vitro*.

Previous studies utilized a three-factor, five-level orthogonal rotation design to establish a mathematical model for the inhibitory effects of PMBE on the growth of cancer cell lines cultured *in vitro* [13]. Compared with the response-surface model of Yu et al. [16], which optimises Pinus massoniana bark extract efficacy purely by statistical interpolation, our hybrid branching-Hill framework offers mechanistic clarity by explicitly coupling capacity-limited proliferation with a sigmoidal kill term. Interestingly, the model's underprediction of late-phase inhibition at 80 μg/mL (Figure 1(b)) suggests a secondary death mechanism, perhaps caspase-mediated apoptosis that the current formulation does not resolve. Future work will incorporate an explicit death-rate term calibrated against Annexin V flow-cytometry assays to capture this phenomenon more accurately [17].

However, RSM lacks the availability to provide a time-and-dose-dependent dynamic description of PMBE's metabolism. This research establishes a hybrid model based on the Poisson process [18] and the branching process [19]. This provides further evidence that nonlinear and probabilistic theory can be applied to the study of human diseases, laying a foundation for future research into the pharmacology and pharmacokinetics of related drug development, ultimately offering guidance for clinical drug administration.

5 Conclusion

In this project, our mathematical model focuses on the potential of Pinus Massoniana Bark Extract (PMBE) in cancer treatment and establishes a mathematical framework that models the inhibitory effect on the proliferation of BEL-7402 cells under the influence of PMBE. In modern medicine, cancer treatment typically involves radiotherapy and chemotherapy in clinic; however, both methods often cause damage to healthy cells beyond the targeted tumor of cancer patients. We have developed a novel branching-Hill model that captures PMBE's time-and-dose-dependent inhibition of BEL-7402 cells with high quantitative fidelity. Nevertheless, it assumes a homogeneous cell population, omits explicit apoptosis kinetics driven by caspase activation [17], and relies solely on MTT assay readouts, which cannot distinguish between cytostatic and cytotoxic effects [13]. We view these findings as a foundational step toward truly predictive, patient-relevant models of phytochemical anticancer effects, rather than as a definitive endpoint. This is a groundbreaking result that offers constructive insights and potential avenues for future cancer treatments—adaptive therapy [20–22].

Data Availability

The raw MTT and dose–response data used in this study were originally generated by Yingyu Cui *et al.* (Tongji University, Shanghai, China) and are reproduced here with the authors' permission. All processed datasets supporting the conclusions of this article are provided in Supplementary File 1. Additional raw data are available from the corresponding author upon reasonable request.

Funding

The authors would like to acknowledge the financial support from the 2023 Key R&D Programme of Tongji University, Grant No. 150029607160-24323.

Ethics Statement

The study utilised previously published *in vitro* data obtained from an established human hepatoma cell line (BEL-7402). No experiments involving human participants or live animals were performed by the authors; therefore, institutional review board approval was not required. The original data collection complied with the biosafety and ethical guidelines of Tongji University.

Conflicts of Interest

The authors declare that they have no conflict of interest.

References

[1] Hanahan D, Weinberg RA. Hallmarks of Cancer: The Next Generation. Cell. 2011, 144(5): 646-674. https://doi.org/10.1016/j.cell.2011.02.013

[2] Vogelstein B, Papadopoulos N, Velculescu VE, et al. Cancer Genome Landscapes. Science. 2013, 339(6127): 1546-1558. https://doi.org/10.1126/science.1235122

[3] Nowell PC. The Clonal Evolution of Tumor Cell Populations. Science. 1976, 194(4260): 23-28. https://doi.org/10.1126/science.959840

[4] Huang JH, Wang J, Chai XQ, et al. The Intratumoral Bacterial Metataxonomic Signature of Hepatocellular Carcinoma. Theis KR, ed. Microbiology Spectrum. 2022, 10(5). https://doi.org/10.1128/spectrum.00983-22

[5] Bullough WS. The actions of the chalones. Agents and Actions. 1971, 2(1): 1-7. https://doi.org/10.1007/bf01965372

[6] Kopustinskiene DM, Jakstas V, Savickas A, et al. Flavonoids as Anticancer Agents. Nutrients. 2020, 12(2): 457. https://doi.org/10.3390/nu12020457

[7] Cui YY, Chen XH, Xie H, et al. Mathematical modeling of the growth in- hibition of pmbe to human liver cancer cell bel-7402 *in vitro*. Journal of Mathematical Medicine. 2005, 18(2):104–108.

[8] Mao P, Zhang E, Chen Y, et al. Pinus massoniana bark extract inhibits migration of the lung cancer A549 cell line. Oncology Letters. 2016, 13(2): 1019-1023. https://doi.org/10.3892/ol.2016.5509

[9] Kaya M, Abuaisha A, Suer I, et al. Turmeric Inhibits MDA-MB-231 Cancer Cell Proliferation, Altering miR-638-5p and Its Potential Targets. European Journal of Breast Health. 2024, 20(2): 102-109. https://doi.org/10.4274/ejbh.galenos.2024.2023-12-2

[10] Symbolab. Graphing calculator, 2024.

[11] Baskar S, Ganesh SS. Introduction to Numerical Analysis[J]. Powai, Mumbai, India: Depart. Math., Indian Inst. Tech. Bombay, 2016.

[12] Feirg. Matlab code: Newton interpolation coefficient, 2024. Supplementary Software Resource.

[13] Cui YY, Xie H, Qi KB, et al.. Effects of Pinus massoniana bark extract on cell proliferation and apoptosis of human hepatoma BEL-7402 cells. World Journal of Gastroenterology. 2005, 11(34): 5277. https://doi.org/10.3748/wjg.v11.i34.5277

[14] Roy M, Horovitz A. Partitioning the Hill coefficient into contributions from ligand-promoted conformational changes and subunit heterogeneity. Protein Science. 2022, 31(5). https://doi.org/10.1002/pro.4298

[15] GraphPad. Graphpad prism 10 curve fitting guide: Hill slope, 2024.

[16] Yu X, Li Y, Chen Z. Response–surface modeling of Pinus massoniana bark extract efficacy: A comparative study. Journal of Natural Products. 2024, 87(3):123–135.

[17] Smith A, Johnson B. Standardized annexin v flow cytometry protocol for quantifying apoptosis. Cytometry Part A. 2025, 97(1): 12–19.

[18] Liu P, Peña EA. Sojourning With the Homogeneous Poisson Process. The American Statistician. 2016, 70(4): 413-423. https://doi.org/10.1080/00031305.2016.1200484

[19] Wang RY, Kimmel M. A Countable-Type Branching Process Model for the Tug-of-War Cancer Cell Dynamics. Bulletin of Mathematical Biology. 2024, 86(2). https://doi.org/10.1007/s11538-023-01245-1

[20] Gatenby R, Whelan C. Cancer treatment innovators discover Charles Darwin. Evolution, Medicine, and Public Health. 2019, 2019(1): 108-110. https://doi.org/10.1093/emph/eoz018

[21] Gatenby R, DeGregori J. Darwin's Cancer Fix. Scientific American. 2019, 321(2): 52. https://doi.org/10.1038/scientificamerican0819-52

[22] Zhang XL, Li MQ, Li YY, et al. Effects of Pinus massoniana bark extract on the size of hela cells via nesprin-2 pathway. Current Cancer Reports. 2020, 2(1): 41–47. https://doi.org/10.25082/CCR.2020.01.003

[23] MIT OpenCourseWare. Probabilistic systems analysis and applied probability, 2010.

[24] University of Nairobi. Introduction to probability and statistics (sma 140), 2021. Accessed: 2024-09-09.

[25] Grimmett GR, Stirzaker DR. Probability and Random Processes. Oxford University Press, 2nd edition, 1992.

(Edited by Snowy Wang)

Appendix

A Definitions and theorems

A.1 Probability Generating Functions (PGFs)

In mathematical modeling, we often use random variables that satisfy some specific distributions to simulate the development and changes of the objects we want to study, such as the fluctuation of stock prices in financial markets, the birth process of a species, and so on. For this kind of problems, for a given random variable, we usually pay attention to some properties such as its *distribution function*, *expectation*, *variance*, *etc*. The probability generating function is a vital function which is often widely applied in probability and statistical analysis.

Definition 1 (Probability Generating Function). Assume X is a **discrete** random variable and Ω_X is its corresponding sample space. Then the probability generating function of X is defined as

$$G_X(s) = \mathbb{E}(s^X) = \sum_{i\in\Omega_X} s^i \mathbb{P}(X = i).$$

The expression for probability generating function is in the form of **power series** and it has the following characteristics:
Convergence. $\exists R \geq 0 : \forall s \in (-R, R), G_X(s) = \sum_{i\in\Omega_X} s^i\mathbb{P}(X = i)$ converges; $G_X(s)$ diverges if $|s| \geq R$.
Differentiable. Here $G_X(s)$ may be differentiable or integrated term by term any number of times at s, if s lies in the convergence interval $(-R, R)$. This is related to some properties of X itself that we will discuss later.
Uniqueness. Assume X_1, X_2 are two random variables such that

$$\begin{cases} G_{X_1}(s) = \sum_{i=0}^{\infty} a_i s^i, \\ G_{X_2}(s) = \sum_{i=0}^{\infty} b_i s^i. \end{cases}$$

If $G_{X_1}(s) = G_{X_2}(s)$ for all $0 \leq s < R' \leq R$ (R is the radius of convergence), then we have

$$\forall i \in \mathbb{N} : a_i = b_i = \frac{1}{i!} G_{X_1}^{(i)}(0).$$

Ttheorem 2. Assume the discrete random variable X has the probability generating function $G_X(s)$. Then:
(1) $G'_X(1) = \mathbb{E}(X)$;
(2) $G_X^{(k)}(1) = \mathbb{E}(X(X-1)\cdots(X-k+1))$.
The right hand side of the expression is called k-th factorial moment of X.
(3) $Var(X) = G''_X(1) + G'_X(1) - (G'_X(1))^2$.

For the first part, notice that

$$\begin{aligned} G'_X(s) &= \frac{\mathrm{d}}{\mathrm{d}s}\mathbb{E}(s^X) = \mathbb{E}(Xs^{X-1}), \\ G'_X(1) &= \mathbb{E}(X). \end{aligned}$$

For the second part, this is a generalization of *statement (1)* .

$$\begin{aligned} G_X^{(k)}(s) &= \frac{\mathrm{d}^k}{\mathrm{d}s^k}\mathbb{E}(s^X) = \mathbb{E}(X(X-1)\cdots(X-k+1)s^{X-k}), \\ G_X^{(k)}(1) &= \mathbb{E}(X(X-1)\cdots(X-k+1)). \\ Var(X) &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \mathbb{E}(X(X-1)+X) - \mathbb{E}(X)^2 \\ &= \mathbb{E}(X(X-1)) + \mathbb{E}(X) - \mathbb{E}(X)^2 \\ &= G''_X(1) + G'_X(1) - (G'_X(1))^2. \end{aligned}$$

The proof is completed.

Theorem 3. The discrete random variables X and Y are independent and their corresponding probability generating functions are $G_X(s)$, $G_Y(s)$. Then the probability generating function of $X + Y$ is

$$G_{X+Y}(s) = G_X(s)G_Y(s).$$

Note that

$$G_{X+Y}(s) = \mathbb{E}(s^{X+Y}) = \mathbb{E}(s^X s^Y) = \mathbb{E}(s^X)\mathbb{E}(s^Y) = G_X(s)G_Y(s).$$

Basing on Theorem 3, We can deduce that if we have **mutually independent random variables** $X_1, X_2, \cdots, X_n$ all taking values in $\mathbb{Z}_{\geq 0}$, then the probability generating function of the sum $T = X_1 + X_2 + \cdots + X_n$ is given by

$$G_T(s) = G_{X_1}(s)G_{X_2}(s)\cdots G_{X_n}(s) = \prod_{i=1}^{n} G_{X_i}(s).$$

A.2 Random Sums

In the previous content, when we have **a certain number of independent random variables** (their possible values are all non-negative integers), then their sum is also a random variable, and its probability generating function is the product of the probability generating functions of all random variables (*e.g.* n is a fixed number). When we talk about the random sums, the problem becomes more complicated - which is like **considering one more dimension in the aspect of randomness.** But many examples in reality can be expressed by this situation.

Queuing Model, which is one of the most classic models in operational research.
Suppose there is a store, and the number of people who enter the store to get services is a random variable N. For each customer who comes in to receive service, their waiting time follows the distribution T. Then the total waiting time can be expressed as

$$T_{\text{Total}} = T_1 + T_2 + \cdots + T_N.$$

Population Models. There are N plants in a given area at $t = 0$. Each plant can produce seeds, and the quantity produced obeys specific distributions. Y_i represents the number of seeds produced by plant i in a year. Then the number of seeds produced by all plants in such area in one year can be expressed by random sum:

$$Y_{\text{Total}} = Y_1 + Y_2 + \cdots + Y_N.$$

Next, we introduce the definition of random sum.

Definition 2 (Random Sum). Assume N is a **discrete** random variable with values in $\mathbb{N} = \{0, 1, 2, ...\}$, $\{Y_i : i = 1, 2, \cdots\}$ is a sequence of **independent identical distribution (i.i.d)** random variables. Suppose Y_i and N are **mutually independent**. The random sum X is defined as follows:

$$X = \begin{cases} Y_1 + Y_2 + \cdots + Y_N, & \text{if } N > 0; \\ 0, & \text{if } N = 0. \end{cases}$$

The expectation and variance of random sums are given below.

Theorem 4. Following the Definition 2, assume N and all Y_i have the finite moments and

$$\begin{aligned} &\text{For } Y_i: \quad \mathbb{E}(Y_i) = \mu, \quad Var(Y_i) = \sigma^2; \\ &\text{For } N: \quad \mathbb{E}(N) = \mu_N, \quad Var(Y_i) = \sigma_N^2. \end{aligned}$$

Then the random sum $X = \sum_{i=1}^{N} Y_i$ has the following properties:

$$\mathbb{E}(X) = \mu\mu_N, \quad Var(X) = \mu_N\sigma^2 + \mu^2\sigma_N^2.$$

We first prove the expectation.

$$\begin{aligned} \mathbb{E}(X) = \mathbb{E}_N(\mathbb{E}(X|N)) &= \sum_{n\in\Omega_N} (\mathbb{E}(X|N=n))\,\mathbb{P}(N=n) \\ &= \sum_{n\in\Omega_N} (\mathbb{E}(Y_1 + \cdots + Y_N|N=n))\mathbb{P}(N=n) \\ &= \sum_{n\in\Omega_N} n\mu\mathbb{P}(N=n) \\ &= \mu \sum_{n\in\Omega_N} n\mathbb{P}(N=n) \\ &= \mu\mathbb{E}(N) = \mu\mu_N. \end{aligned}$$

$$\begin{aligned} Var(X) &= \mathbb{E}_X((X-\mathbb{E}(X))^2) = \mathbb{E}_X((X-\mathbb{E}(X|N)+\mathbb{E}(X|N)-\mathbb{E}(X))^2) \\ &= \mathbb{E}_X((X-\mathbb{E}(X|N))^2) + \mathbb{E}_X((\mathbb{E}(X|N)-\mathbb{E}(X))^2) \\ &\quad + 2\mathbb{E}_X((X-\mathbb{E}(X|N))(\mathbb{E}(X|N)-\mathbb{E}(X)), \end{aligned}$$

where we have

$$\begin{aligned} \mathbb{E}_X((X-\mathbb{E}(X|N))^2) &= \mathbb{E}_X((X-N\mu)^2) = \mathbb{E}_N(\mathbb{E}((X-N\mu)^2|N)) \\ &= \sum_{n\in\Omega_N} (\mathbb{E}((X-N\mu)^2|N=n))\mathbb{P}(N=n) \\ &= \sum_{n\in\Omega_N} \left(\mathbb{E}\left(\left(\sum_{i=1}^{n}(Y_i-\mu)\right)^2 |N=n\right)\right)\mathbb{P}(N=n) \\ &= \sum_{n\in\Omega_N} \mathbb{E}\left(\sum_{i=1}^{n}(Y_i-\mu)^2 + \sum_{i\neq j}(Y_i-\mu)(Y_j-\mu)\,|N=n\right)\mathbb{P}(N=n) \\ &= \sum_{n\in\Omega_N} \left(\sum_{i=1}^{n}\mathbb{E}\left((Y_i-\mu)^2\right) + \sum_{i\neq j}\mathbb{E}\left((Y_i-\mu)(Y_j-\mu)\right)\right)\mathbb{P}(N=n) \\ &= \sigma^2 \sum_{n\in\Omega_N} n\mathbb{P}(N=n) + 0 = \sigma^2\mu_N. \end{aligned}$$

(Notice that $\mathbb{E}\left((Y_i-\mu)(Y_j-\mu)\right)=0$ as Y_i and Y_j are independent identical distributions if $i\neq j$.)

$$\mathbb{E}_X((\mathbb{E}(X|N)-\mathbb{E}(X))^2) = \mathbb{E}_X\left((N\mu-\mu\mu_N)^2\right) = \mu^2\mathbb{E}_X((N-\mu_N)^2) = \mu^2\sigma_N^2.$$

For the term $2\mathbb{E}_X((X-\mathbb{E}(X|N))(\mathbb{E}(X|N)-\mathbb{E}(X))$,

$$\begin{aligned} \mathbb{E}_X((X-\mathbb{E}(X|N))(\mathbb{E}(X|N)-\mathbb{E}(X)) &= \mathbb{E}_N\left(\mathbb{E}\left((X-N\mu)(N\mu-\mu\mu_N)|N=n\right)\right) \\ &= \sum_{n\in\Omega_N} \mu(\mathbb{E}((X-n\mu)(n-\mu_N))\mathbb{P}(N=n) \\ &= \sum_{n\in\Omega_N} \mu(n-\mu_N)\mathbb{E}\left(\sum_{i=1}^{n}(Y_i-\mu)\right)\mathbb{P}(N=n) \\ &= 0, \end{aligned}$$

as $\mathbb{E}\left(\sum_{i=1}^{n}(Y_i-\mu)\right)=0$. So basing on the calculations above we can finally conclude that

$$Var(X) = \sigma^2\mu_N + \mu^2\sigma_N^2.$$

Now we consider the probability generating function of a random sum.

Theorem 5. Suppose $\{Y_i\}_{i\in\mathbb{N}}$ is a sequence of independent identical distribution random variables with probability generating function $G_Y(s)$. The random variable N with sample spaces $\mathbb{Z}_{\geq 0}$ is independent to all Y_i and has the probability generating function $G_N(s)$. Then the probability generating function of the random sum $X=Y_1+\cdots+Y_N$ is given by

$$G_X(s) = G_N(G_Y(s)).$$

$$\begin{aligned} G_X(s) = \mathbb{E}(s^X) = \mathbb{E}_N(\mathbb{E}(s^X|N)) &= \sum_{n\in\Omega_N} \mathbb{E}(s^X|N=n)\mathbb{P}(N=n) \\ &= \sum_{n\in\Omega_N} \mathbb{E}(s^{Y_1+\cdots+Y_n})\mathbb{P}(N=n) = \sum_{n\in\Omega_N}\prod_{i=1}^{n}\mathbb{E}(s^{Y_i})\mathbb{P}(N=n) \\ &= \sum_{n\in\Omega_N} (G_Y(s))^n\mathbb{P}(N=n) = G_N(G_Y(s)). \end{aligned}$$

A.3 Stochastic Processes

We usually analyze and find out the frequency of an event or the distribution of a random variable (at a fixed time or within a given period) at a certain time (or a certain time section) in the probability and statistics we studied in the past. But from now on, we pay more attention to the variety of a variable about time, and we expect to find a way to describe its dynamic process, so we put forward the definition of **stochastic process**.

Definition 3 (Stochastic Process). A random (stochastic) process $X = \{X(t) : t \in T\}$ is a **collection** of random variables.
(1) The set T is a time space, we also call it the index of the process. When T is countable, we say the process X is a **discrete-time process**. If T is an interval of the real line, we say X follows a **continuous-time process**.
(2) For all $t \in T$, we say $X(t)$ is the **state** of the process at time t.

Next, we mention two stochastic processes-Poisson process and Branching process, which are closely related to the models established in the project and we will discuss it in the next section.

A.4 Poisson Process

Definition 4 (Poisson Process). The Poisson Process $\{N(t) : t \geq 0\}$ is basically a **counting process**. A such process on the time interval $t = [0, \infty)$ counts the times of some primitive events have occurred during the time interval $[0, t]$. Meanwhile, the process has the following properties:

Time homogeneity/stationary increments. For a fixed $h \geq 0$, the distribution of $N(t + h) - N(t)$ is the same for all t (i.e. is independent of t). The distribution of $N(t_2) - N(t_1)$ $(0 \leq t_1 \leq t_2)$ only depends on the length of $t_2 - t_1$.

Independent increments. Any two increments involving disjoint intervals are independent. If $0 \leq s_1 < s_2 < t_1 < t_2$ (i.e, the time intervals $[t_1, t_2]$ and $[s_1, s_2]$ do not overlap), $N(t_2) - N(t_1)$ and $N(s_2) - N(s_1)$ are independent.

$N(0) = 0$, $N(t)$ takes integer value for all t. Also this function is **right continuous** and **non-decreasing** in $[0, \infty)$, with probability 1.

Assume the intensity parameter is λ. At time $t \geq 0$, for any time interval $[t, t + h]$ $(h > 0)$, the number of events follows the Poisson distribution, Poisson(λh). (Figure 4)

$$\mathbb{P}(N(t + h) - N(t) = k) = \mathbb{P}(N(h) = k) = e^{-\lambda h} \frac{(\lambda h)^k}{k!}, \quad \text{for } k \in \mathbb{N}.$$

The corresponding **expectation** and **variance** are given below:

$$\mathbb{E}(N(t + h) - N(t)) = \lambda h, \quad Var(N(t + h) - N(t)) = \lambda h.$$

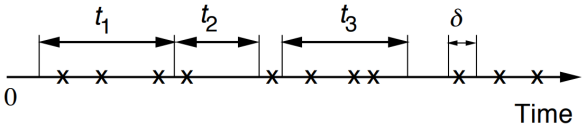


Figure 4 Poisson Process [23]

For a Poisson process, we have the following characteristics to pay attention to:

The process is memoryless. The Poisson process is a **continuous-time Markov chain** (i.e. *"when the present conditions are known, the future is not determined by the past."*).

For a fixed $t \geq 0$ and a sufficient small $h > 0$, we analyze the probability distribution of number of events (k) during $[t, t + h]$:

$$\mathbb{P}(k = 0) = e^{-\lambda h} \frac{(\lambda h)^0}{0!} = e^{-\lambda h} = \sum_{n=0}^{\infty} \frac{(-\lambda h)^n}{n!} = 1 - \lambda h + o(h),$$

$$\mathbb{P}(k = 1) = e^{-\lambda h} \frac{(\lambda h)^1}{1!} = e^{-\lambda h}(\lambda h) = \sum_{n=0}^{\infty} \frac{(-\lambda h)^n}{n!}(\lambda h) = \lambda h + o(h),$$

$$\mathbb{P}(k \geq 2) = 1 - \mathbb{P}(k = 0) - \mathbb{P}(k = 1) = o(h).$$

This implies, within a sufficiently short time interval, the likelihood of the event occurring more than once can be neglected.

A.5 Branching Process

Another vital instance of stochastic process is **branching process**. (Figure 5)

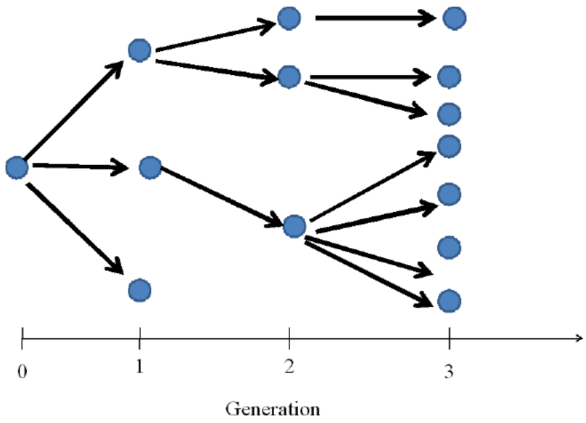


Figure 5 An Example of Branching Process [24]

Definition 5 (Branching Process). Suppose $Z = \{Z_n : n \geq 0\}$ is a sequence of random variables taking non-negative integers. We say Z is a branching process if it satisfies the following conditions:
(1) $Z_0 = 1$;
(2) $Z_{n+1} = X_1 + X_2 + \cdots + X_{Z_n}$, where X_i $(i = 1, \cdots, Z_n)$ are independent identical distribution random variables taking non-negative integers as their values.

A typical example of branching process is the family tree, which is demonstrated in Figure 6.

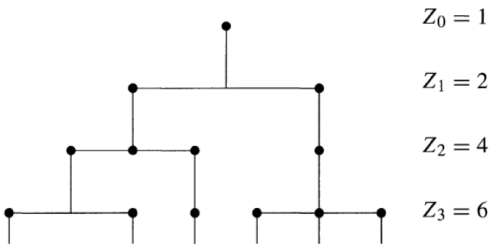


Figure 6 Family Tree [25]

Table 8 The Evolution of Branching process in Figure 6.

Time (t_n)	X_k $(k = 1, \cdots, Z_{n-1})$	$Z_n = \sum_{k=1}^{Z_{n-1}} X_k$
$t = 0$	N/A	$Z_0 = 1$
$t = 1$	$X_1 = 2$	$Z_1 = X_1 = 2$
$t = 2$	$X_1 = 3, X_2 = 1$	$Z_2 = X_1 + X_2 = 4$
$t = 3$	$X_1 = 2, X_2 = 0, X_3 = 1, X_4 = 3$	$Z_3 = \sum_{k=1}^{4} X_k = 6$
$\cdots$	$\cdots$	$\cdots$

Now we discuss the probability generating functions of the Branching process $Z = \{Z_n : n \geq 0\}$. Define the probability generating function of Z_n

$$G_n(s) = \mathbb{E}(s^{Z_n}), \quad \text{for } n \geq 0,$$

we have the following theorem.

Theorem 6. Assume $G(s) = G_1(s)$, then

$$G_{n+1}(s) = G_n(G(s)), \quad \text{for } n \geq 0.$$

By recursion we finally have

$$G_n(s) = \underbrace{G \circ \cdots \circ G}_{n \text{ times}}(s),$$

we call it the n-fold iterate of G.

We know that $Z_{n+1} = X_1 + \cdots + X_{Z_n}$. Notice that: (1) $Z_1 = X_{Z_0} = X_1$ since $Z_0 = 1$; (2) X_i are all independent identical distributions.
Therefore, we get

$$G(s) = G_1(s) = G_{X_1}(s).$$

Furthermore, we obtain

$$\begin{aligned} G_n(s) = \mathbb{E}(s^{Z_n}) &= \mathbb{E}(s^{\sum_{i=1}^{Z_{n-1}} X_i}) = \mathbb{E}_{Z_{n-1}}(\mathbb{E}(s^{X_i}|Z_{n-1})^{Z_{n-1}}) = G_{n-1}(G(s)) \\ &= G_{n-2}(G(G(s))) = \cdots = G(G \cdots G(G(s))) = \underbrace{G \circ \cdots \circ G}_{n \text{ times}}(s). \end{aligned}$$

Theorem 7. For a Branching process $Z = \{Z_n : n \geq 0\}$, assume

$$\mathbb{E}(Z_1) = \mu, \quad Var(Z_1) = \sigma^2,$$

we have

$$\mathbb{E}(Z_n) = \mu^n, \quad Var(Z_n) = \begin{cases} n\sigma^2, & \text{if } \mu = 1, \\ \frac{\sigma^2(\mu^n - 1)\mu^{n-1}}{\mu - 1}, & \text{if } \mu \neq 1. \end{cases}$$

We only give the proof of expectation here.

We know that $Z_{n+1} = X_1 + \cdots + X_{Z_n}$. Notice that:

$$\begin{aligned} \mathbb{E}(Z_n) = G_n'(s)|_{s=1} &= \frac{\mathrm{d}}{\mathrm{d}s} G(G_{n-1}(s))|_{s=1} = G'(G_{n-1}(s))G_{n-1}'(s)|_{s=1} \\ &= G'(G_{n-1}(1))G_{n-1}'(1) \qquad \text{(Note: } G(1) = 1 \text{ for all probability generating functions)} \\ &= G'(1)G_{n-1}'(1) = \mu G_{n-1}'(1) \\ &\Rightarrow \mathbb{E}(Z_n) = \mu \mathbb{E}(Z_{n-1}). \end{aligned}$$

B MATLAB Implementation of Newton Interpolation

The MATLAB function below computes the coefficients of the inhibition rate interpolation polynomial:

Lstlisting 1: MATLAB code of Newton's form of interpolating polynomials

```
function diff_coefficient=NewtonInterpolation(x,y)
if ((size(x,2)-size(y,2)~=0)...
        ||...
        (size(x,1)~=1)...
        ||...
        (size(y,1)~=1))
    error('x and y must both be 1-D vector with the same
     dimension.')
end
n=length(x);
y=y';
y(n,n)=0;
for column=2:n
    for row=column:n
        y(row,column)=(y(row,column-1)-y(row-1,column-1))/...
            (x(row)-x(row-column+1));
    end
end
diff_coefficient(n)=0;
for row=1:n
    diff_coefficient(row)=y(row,row);
end
end
```

C Reproducing Figure 3

```python
import numpy as np
import matplotlib.pyplot as plt

# Experimental data
C     = np.array([ 20,  40,  80, 160, 200])
I_obs = np.array([16.97,25.20,30.64,36.64,43.17])

# Fitted Hill parameters
IC50 = 380
n    = 0.52

# Generate smooth fit curve
C_fit = np.linspace(0, 250, 500)
I_fit = 100 * C_fit**n / (IC50**n + C_fit**n)

# Plot
plt.figure()
plt.plot(C_fit, I_fit, '-', label='Hill_fit')
plt.plot(C, I_obs, 'x', label='Observations')
plt.xlabel('PMBE_concentration_(g/mL)')
plt.ylabel('Growth_inhibition_(%)')
plt.legend()
plt.tight_layout()
plt.savefig('figures/fig-3.png', dpi=300)
```