arXiv:2603.06751v1 [q-bio.CB] 6 Mar 2026

有限实验数据下细胞周期年龄结构模型的参数可识别性

Ruby E.Nixson[1*]，Helen M. Byrne[1, 2†]，Joe M. Pitt-Francis[3†]，Philip K. Maini[1†]1沃尔夫森数学生物学中心，牛津大学，牛津，英国。2牛津大学路德维希癌症研究所，牛津，英国。3牛津大学计算机科学系，牛津，英国。

*通讯作者。电子邮箱：ruby.nixson@maths.ox.ac.uk；
†这些作者对本研究的贡献相同。

摘要

有丝分裂细胞周期调控DNA复制和细胞分裂。放射治疗和化疗的效果取决于细胞周期位置；在DNA复制和有丝分裂期间，细胞的耐受性会增强。因此，准确的细胞周期数学模型对于理解和预测治疗反应至关重要。然而，数学建模者常常面临一个问题：缺乏公开可用且分辨率足够高的时间序列数据集，无法用于模型参数化。在本研究中，我们探讨如何整合文献中来自不同细胞系和/或实验设置的群体汇总测量，进而影响细胞周期模型参数的可识别性。

最初同步化的细胞群体会在连续多个周期中逐渐失去同步，最终趋于平衡指数增长（BEG）；其特征是种群指数增长，且各时相所占比例稳定、与时间无关。这些比例可通过荧光激活细胞分选（FACS）数据获得。荧光泛素化细胞周期指示剂（FUCCI）的应用日益广泛，可提供更高分辨率的时相动力学信息，例如最短时相持续时间和变异性。

我们提出了一种年龄结构偏微分方程模型，其中细胞周期阶段的进展遵循延迟伽马分布。我们推导出 BEG 阶段比例及其他 FUCCI 可观测量的解析表达式，并用它们评估

数据可用性如何影响参数可识别性。当参数无法被唯一识别时，我们确定可识别的参数组合，从而确定成功拟合细胞周期结构化种群模型所需的最小数据量。

关键词： Cell周期、数学建模、结构化偏微分方程、年龄结构模型、FUCCI、FACS、数据限制

1 引言

有丝分裂细胞周期调控细胞内容物的复制，以及一个细胞分裂为两个子细胞的过程。细胞周期由四个阶段组成，必须按顺序完成后才能发生分裂。首先，细胞进入 G_1 期，在此期间细胞体积增大，并开始为DNA复制做准备。在这一阶段，细胞面临一种“选择”：要么保持活跃的周期循环状态，要么以可逆的方式暂停细胞周期进程，这种状态称为静止期，或 G_0 期 [1]。DNA复制发生在第二阶段，即 S期，随后进入 G_2 期，在此期间细胞为分裂做准备，并修复在 S 期累积的任何损伤。之后，细胞进入有丝分裂期，即 M期，这是一个以细胞分裂告终的短暂阶段。

癌症的一种常见治疗方法是放射治疗，实验结果表明，放射治疗的反应会受到细胞周期时相的影响 [2–4]。M期细胞的放射敏感性最高，而S期细胞的放射抗性最强 [5]。静止期细胞也具有放射抗性 [3]。此外，针对癌细胞的化疗也会与细胞周期相互作用：一些化疗药物会调节细胞周期进程，另一些则会诱发依赖于细胞周期时相的反应 [6–8]。因此，在利用细胞周期模型理解和预测治疗反应时，准确的数学模型至关重要。

传统上，流式细胞术一直是报告细胞群体中细胞周期时相最常用的技术。通过固定细胞并对DNA含量进行染色，流式细胞术可以区分每条染色体各有一份的细胞（表明细胞处于 G_1/G_0），与每条染色体各有两份的细胞（复制后，因此是 G_2 和 M 细胞），还可以区分在 S 期 [9–12]复制过程中DNA含量处于中间水平的细胞。碘化丙啶（PI）染色可显示细胞的DNA含量，但无法准确区分处于晚期G_1 期和早期S 期的细胞，也无法区分处于晚期S 期和早期G_2 期的细胞。因此会使用5-溴-2'-脱氧尿苷（BrdU）染色，它可识别处于 S 期的细胞，因为BrdU会掺入新复制的DNA中。还可结合Ki-67染色来区分增殖细胞和静息细胞，从而报告各时相的细胞数量 [13, 14]。也可以使用无需在实验前固定细胞的染色化合物，以便纵向追踪细胞群体层面细胞周期时相的进程 [11, 15, 16]。然而，这些活细胞染料

会给细胞群体带来更大的应激，而且与固定细胞相比，所得测量结果的分辨率较低 [15]。

肿瘤生长与细胞周期的数学建模是一个活跃的研究领域，涵盖多种方法，包括基于智能体的模型、偏微分方程（PDE）和常微分方程（ODE）。 [17]对细胞周期的数学建模及其对治疗策略的影响进行了综述。其中若干研究 [18–23]，通过ODE模型中不同时相之间的恒定转移速率来刻画细胞周期进程。这隐含地假定，潜在的细胞周期时相长度服从指数分布。实验数据表明，最初在某一细胞周期时相同步的细胞群体，经过若干完整周期后最终会去同步，从而产生平衡指数增长（BEG）。该增长状态的特征是细胞群体数量呈指数增长，但细胞周期阶段比例会趋于与时间无关的稳态。对这些时相比例的流式细胞术测量随后可用于确定细胞周期ODE模型中的恒定转移速率。

过去15年来，基于荧光泛素化的细胞周期指示剂（FUCCI）已成为分析细胞周期进程的常用技术 [24, 25]。与对DNA含量进行染色不同，该技术利用荧光标记来指示细胞周期调控蛋白的积累和降解。与流式细胞术方法不同，FUCCI可以对单个细胞进行实时追踪，从而提供分辨率更高的数据，例如各个时相的持续时间、谱系追踪，以及不同细胞之间时相长度的异质性。FUCCI早期的一项局限是无法区分 $S/G_2/M$ 期细胞，因为这些细胞均呈绿色荧光，而 G_1 期细胞呈红色荧光。不过，现在使用FUCCI4 [26],便可报告全部四个阶段；该技术利用与DNA复制和有丝分裂相关联的细胞周期标记物，分别区分 S 期和 M期细胞。

此前，研究人员已将细胞周期的年龄结构模型拟合到FUCCI成像数据和BrdU脉冲标记数据。Weber等人 [27] 假设阶段时长服从延迟指数分布，并发现，在标准的单脉冲BrdU实验中，可以识别阶段时长的均值，但方差仍无法识别。他们提出重新设计双脉冲实验来解决这一问题，并强调参数可识别性明显取决于实验设计。

FUCCI及其更高分辨率的衍生技术的出现，使得研究更复杂的潜在时相长度分布的参数可识别性成为可能。与Weber等人 [27], 的研究不同，我们不再局限于设法解读流式细胞术方法生成的群体水平数据，而是可以尝试利用FUCCI实验中的单细胞数据。例如，Billy等人 [28] 使用FUCCI时间序列数据推断时相转移分布。随后，这些分布被纳入偏微分方程组，并用于计算不同实验条件下的细胞群体生长速率。此类方法侧重于根据实验数据进行参数估计，而未考察给定实验数据下参数的结构可辨识性或实际可辨识性。在另一种脉冲标记方法中，Sherer等人 [29] 采用了一种基于

提取瞬时亚群的实验设计，以推断细胞周期各时相之间年龄依赖性转变率。

综上，这些早期研究表明，只要有足够丰富的实验数据，就能为这类年龄结构模型设定参数，无论数据来自 BrdU 标记还是 FUCCI 成像。最重要的是，这三项先前研究都使用时间序列数据来拟合模型参数。然而，数学建模者往往无法获取详细的实验数据集，而只能依赖文献中报道的细胞群体汇总测量值；当缺少目标细胞系的完整数据时，这些测量值通常来自多种细胞系。因此，当无法获得这些丰富的数据集时，哪些模型参数能够通过最常见的数据类型识别，以及数据的可用性如何影响模型参数和其他具有生物学意义的量的识别能力，仍不清楚。不同于以往证明更丰富的实验设计能够揭示更多结构的研究，我们关注的是受限汇总数据条件下的参数推断这一互补问题。

在本研究中，我们考虑一种细胞周期年龄结构群体模型，其中每个阶段都对应一个偏微分方程，向下一阶段的转换取决于细胞在当前阶段所经历的时间。我们假设每个阶段时长都服从延迟伽马概率分布，并推导出关键可测量量的解析表达式。这些量包括 BEG比例和阶段时长的变异系数；它们都可以通过流式细胞术或 FUCCI 实验测得，但发表时往往没有附上用于推导这些量的完整时间序列数据。通过假设我们能够获得数量逐渐增加的这些量，我们考察模型在不同情形下的参数可识别性。因此，通过要求模型达到特定的细节程度，我们可以评估需要多少个此类群体汇总值，才能获得足以实现参数可识别性的数据。

2 引入年龄结构模型

我们将细胞周期划分为三个区室，分别对应处于 G_1、 S 和G_2/M的细胞。引入年龄结构变量，用于追踪细胞在各个阶段所经历的时间。

定义 $G_1(a,\ t)$、 $S(b,\ t)$ 和 $G_2(c,\ t)$ 分别表示已处于 $G_1/S/G_2$ 期 a、 b 和 c 小时的细胞密度，我们提出以下偏微分方程组来描述它们的演变：

$$\frac{\partial G_1}{\partial t} + \frac{\partial G_1}{\partial a} = -\mu_1(a)G_1(a,t) + \beta_4 Q(a,t) - \beta_5 G_1(a,t), \tag{1}$$

$$\frac{\partial S}{\partial t} + \frac{\partial S}{\partial b} = -\mu_2(b)S(b,t), \tag{2}$$

$$\frac{\partial G_2}{\partial t} + \frac{\partial G_2}{\partial c} = -\mu_3(c)G_2(c,t), \tag{3}$$

$$\frac{\partial Q}{\partial t} = -\beta_4 Q(a,t) + \beta_5 G_1(a,t), \tag{4}$$

其中，时间 $t > 0$ 时细胞总数 $P(t)$ 为

$$P(t) = \int_0^\infty (G_1(a,t) + Q(a,t))\, da + \int_0^\infty S(b,t)\, db + \int_0^\infty G_2(c,t)\, dc, \tag{5}$$

并且 $Q(a,t)$是以“年龄” $a > 0$退出 G_1 的静息细胞密度。此处，$\mu_1(a)$、 $\mu_2(b)$和 $\mu_3(c)$是年龄为 a、 b 和 c的 $G_1/S/G_2$ 细胞进入细胞周期下一阶段的速率。我们注意到，$Q(a,t)$的偏微分方程中不含 a导数，因为细胞周期进程在静止期暂停，仅当细胞返回 G_1时才恢复。此类细胞以进入静止期时所具有的相同 G_1年龄恢复细胞周期。

我们封闭方程（1）-（4），施加以下边界条件：

$$G_1(0,t) = 2\int_0^\infty \mu_3(c) G_2(c,t)\, dc \qquad \text{(cell division)}, \tag{6}$$

$$S(0,t) = \int_0^\infty \mu_1(a) G_1(a,t)\, da \qquad \text{(progression from } G_1 \text{ to } S), \tag{7}$$

$$G_2(0,t) = \int_0^\infty \mu_2(b) S(b,t)\, db \qquad \text{(progression from } S \text{ to } G_2/M). \tag{8}$$

许多健康的人类细胞会经历接触抑制；随着细胞密度增加，接触抑制会减缓细胞周期进程和细胞群体的生长速率。然而，本文工作基于癌细胞的实验数据，癌细胞通常会丧失阻止正常细胞群体持续指数增长的接触抑制 [30]。因此，我们避免在偏微分方程的转移速率中施加任何密度依赖性。

为了描述单个细胞有丝分裂间期时长（即细胞周期进程所需的时间）中的细胞异质性，我们假设细胞周期进程速率和细胞分裂速率取决于细胞在相应区室中停留的时间。研究表明，细胞在各细胞周期时相中的停留时间因细胞系而异，可用多种概率分布描述，包括伽马分布、指数分布和对数正态分布 [27, 29, 31–33]。

这里我们采用延迟伽马分布，因为它在解析上易于处理。具体而言，我们假设处于 G_1 期的细胞以服从 $a > 0$ 中平移伽马分布的概率转入 S 期；此外，转变过程 S 到 G_2 以及 G_2 到 G_1 也分别由平移伽马分布描述。我们引入初始时移，以确保细胞在每个时相中至少停留一段固定的最短时间，避免细胞例如从 G_2 直接跳转到 S。

因此，我们假设细胞在 G_1 期必须至少停留 $T_1 > 0$，之后其向 S 推进的概率服从 Gamma(α_1, β_1) 分布。参数集（T_2, α_2, β_2）和（T_3, α_3, β_3）分别定义了 S和G_2的推进速率。然后，我们使用根据概率分布推导年龄结构化转移率的一般公式（由 Stukalin [34], Chaffey [35] 和 Sherer [29]提出），得到以下函数形式：

$$\mu_i(a) = \begin{cases} \frac{\beta_i^{\alpha_i}(a-T_i)^{\alpha_i-1}\exp(-\beta_i(a-T_i))}{\Gamma(\alpha_i,\beta_i(a-T_i))}, & a > T_i \\ 0, & a < T_i \end{cases} \tag{9}$$

其中

$$\Gamma(\alpha, x) = \int_x^\infty s^{\alpha-1}\exp(-s)ds \tag{10}$$

即上不完全伽马函数。采用这些函数形式时，细胞只有在当前阶段停留至少 T_i 时间后，才能进入下一个细胞周期时相；随着其在该时相中停留的时间增加，其进程速率渐近于一个常值。

3 平衡指数增长（BEG）

我们预计，在营养丰富且没有密度限制的环境中，未经处理的癌细胞群体将呈指数增长。实验证据表明，随着细胞群体的增长，处于某一细胞周期时相的细胞计数占总细胞计数的比例最终将稳定为一个恒定的稳态 [36–40]。因此，细胞群体最终将进入平衡指数增长（BEG）状态：总细胞计数呈指数增长，但各细胞周期时相中的细胞比例保持不变。由于可通过流式细胞术测量处于某一细胞周期时相的细胞比例 [9, 11], ，我们将在根据实验数据拟合偏微分方程模型时利用 BEG。

在 BEG 条件下，方程（1）-（8）具有如下形式的可分离解 $X(a,t) = \tilde{X}(a)\exp(\lambda t)$，其中 $X = G_1, S, G_2, Q$；偏微分方程模型因此化简为一组一阶常微分方程。我们求解每个 $X(a)$，并在阶段年龄的半无限域上积分，得到

$$X^{\text{tot}} = \int_0^\infty \tilde{X}(a)\,da, \qquad \text{for} \quad X = G_1, S, G_2, Q. \tag{11}$$

通过写出 $P^{\text{tot}} = G_1^{\text{tot}} + S^{\text{tot}} + G_2^{\text{tot}} + Q^{\text{tot}}$，我们计算出稳态 BEG 比例（$G_1$， S， G_2， Q）为

$$\bar{G}_1 = \frac{G_1^{\text{tot}}}{P^{\text{tot}}} = \frac{2\lambda}{\lambda + f(\lambda)}\left(1 - F(\lambda)\right), \tag{12}$$

$$\bar{S} = 2\left(1 - \left(\frac{\beta_2}{\beta_2+\lambda}\right)^{\alpha_2}\exp(-\lambda T_2)\right)F(\lambda), \tag{13}$$

$$\bar{G}_2 = \left(\frac{\beta_3}{\beta_3+\lambda}\right)^{-\alpha_3}\exp(\lambda T_3) - 1, \tag{14}$$

$$\bar{Q} = \frac{2f(\lambda)}{\lambda + f(\lambda)}\left(1 - F(\lambda)\right), \tag{15}$$

其中

$$f(\lambda) = \frac{\lambda\beta_5}{\lambda + \beta_4} \tag{16}$$

and

$$F(\lambda) = \exp(-(\lambda + f(\lambda))T_1)\left(\frac{1}{1 + \frac{\lambda + f(\lambda)}{\beta_1}}\right)^{\alpha_1}. \tag{17}$$

特征方程

$$1 = 2F(\lambda)\left(\frac{\beta_2}{\beta_2 + \lambda}\right)^{\alpha_2}\left(\frac{\beta_3}{\beta_3 + \lambda}\right)^{\alpha_3}\exp(-(T_2 + T_3)\lambda), \tag{18}$$

定义了唯一的指数增长率 λ，并确保各时相比例之和为 1。

4 拟合实验数据

现在，我们考虑如何将 PI 和 BrdU 数据与上述 BEG分析结合起来，以估算模型参数。我们使用 Celora等人报告的 RKO细胞系（源自一种结肠癌）的流式细胞术数据。[20]这些数据包含九组时相比例测量值，每组均在细胞于常氧条件（21% O_2）下进入 BEG 状态后测得。这些数据不包含任何静止期比例测量值，仅包含 G_1， S 和 G_2/M比例；实验流程也未尝试区分G_1 与静止期。因此，静止期比例是一个隐变量，我们假定的静止期比例必须取自测得的 G_1 比例。在此，我们假定静止期细胞占细胞总数的比例为 0。061，这一数值依据 Corvaisier 等人对 RKO细胞的实验结果 [41] 。我们对九次重复实验取平均值，并在表 1 中列出完整的一组假定 BEG 比例值。

	描述	值
$\bar{g}_1$	BEG处于 G_1 期的细胞比例。	0.246
$\bar{s}$	BEG处于 S 期的细胞比例。	0.539
$\bar{g}_2$	BEG处于 G_2 期的细胞比例。	0.154
$\bar{q}$	BEG处于 Q期的细胞比例。	0.061

表1 Celora等人 [20]中9次实验重复的平均BEG阶段比例。

对于所关注的特定细胞系而言，与脉冲标记实验中获得的时相比例测量时间序列（如 Weber等人 [27]所采用的）相比，通常更容易在文献中找到这些单一的长期时相比例值。由于四个 BEG比例不足以确定全部九个模型参数，因此（α_i、 β_i、 T_i），i =1、2、3我们希望将 BEG比例与通过 FUCCI 获得的细胞群体测量相结合

前提是无法获得完整的 FUCCI成像数据。

通过使用FUCCI技术记录单个细胞的阶段时长，可以获得细胞群体内异质性的数据。在此，我们使用阶段时长的变异系数（CV），以及各时相的最小长度，记为 T_i，见于偏微分方程模型（1）-（4）。对于均值为 μ、标准差为 σ 的概率分布，变异系数定义为

$$\mathrm{CV} = \frac{\sigma}{\mu}, \tag{19}$$

因此，对于位置参数为 T、形状参数为α 、尺度参数为 β的平移伽马分布，我们得到

$$\mathrm{CV}_i = \frac{\sqrt{\alpha_i}}{\alpha_i + T_i\beta_i} \qquad i = 1, 2, 3. \tag{20}$$

由方程（12）和（15），我们得到

$$f(\lambda) = \frac{\bar{Q}}{\bar{G}_1}\lambda. \tag{21}$$

由于我们已有 G_1 和 Q的值，且 β_4、 β_5 仅通过 f（λ）出现在方程（12）-（15）中，因此在本文余下部分中，我们暂不考虑 β_4 和 β_5 。

在 T_i、 $i = 1$、2、3已知，且暂不考虑 β_4、 β_5 的情况下，还剩下六个未知参数值： α_i 和 β_i、 $i = 1$、2、3。方程（12）-（14）以及（20）定义了六个代数方程，我们可用它们确定这六个未知参数。

这些表达式促使我们考察模型的结构可辨识性和实际可辨识性。如果给定模型输出后，存在唯一的输入参数集 [42–44]，则该模型是结构可辨识的。在我们的案例中，如果六个未知参数均可根据 BEG比例、变异系数值和最小阶段长度唯一确定，我们就称该模型具有结构可辨识性。

实际可辨识性是指从含噪输出数据 [42–44] 中唯一估计模型参数的能力。这相当于能够根据含噪的 BEG比例测量值准确识别真实参数值。

数据少于九项时，无法唯一确定全部九个模型参数，因此该模型在结构上不可辨识。不过，即使可用数据有限，我们仍会考察哪些参数分组可以保持可辨识。

4.1 案例1：仅有 BEG 数据

首先，我们假设无法获得由单细胞数据得出的值，因此只能使用流式细胞术测得的 BEG 状态下的平均阶段比例。这对应于最简单、成本最低且最快捷的实验设置，因此这些值通常也最容易在文献中找到。我们还注意到，生长速率λ 与该细胞系的倍增时间 T_D之间的关系为 $\lambda = \log(2)/T_D$。下文中，

我们假设倍增时间为 $T_D = 22$ 小时，这与针对 RKO细胞的实验估计值 21–23 小时一致 [45, 46]。

由于该情形下模型在结构上不可辨识，我们考虑从此类数据中可辨识的参数分组。仅有三项数据可用，因此我们预计最多能得到三个可识别参数分组。在此情形下，我们拟合全部九个模型参数，因此 $\vec{x} = (\alpha_1,\ \alpha_2,\ \alpha_3,\ \beta_1,\ \beta_2,\ \beta_3,\ T_1,\ T_2,\ T_3)$。事实上，从方程（12）–（15）中时相比例的解析表达式即可看出这些可识别参数分组，具体为

$$\eta_1(\alpha_1,\beta_1,\gamma_1) = \left(\frac{\beta_1}{\beta_1 + \lambda(1+\bar{Q}/\bar{G}_1)}\right)^{\alpha_1} \exp(-\lambda(1+\bar{Q}/\bar{G}_1)T_1) = 1 - \frac{\bar{Q}+\bar{G}_1}{2}, \quad (22)$$

$$\eta_2(\alpha_2,\beta_2,\gamma_2) = \left(\frac{\beta_2}{\beta_2 + \lambda}\right)^{\alpha_2} \exp(-\lambda T_2) = \frac{1+\bar{G}_2}{2-\bar{Q}-\bar{G}_1}, \quad (23)$$

$$\eta_3(\alpha_3,\beta_3,\gamma_3) = \left(\frac{\beta_3}{\beta_3 + \lambda}\right)^{\alpha_3} \exp(-\lambda T_3) = \frac{1}{1+\bar{G}_2}. \quad (24)$$

对于给定的一组时相比例（$\bar{G}_1,\ \bar{S},\ \bar{G}_2,\ \bar{Q}$），这些分组将元组（$\alpha_i,\ \beta_i,\ T_i$）、$i = 1,\ 2,\ 3$ 限定在参数空间中的一个曲面上。自然边界 $\alpha_i > 1$ 和 $\beta_i,\ T_i > 0$ 对进一步约束参数空间几乎没有帮助。我们转而考察参数空间中各时相长度的均值和方差是否可辨识。

回想一下，具有时滞 T 的 Gamma($\alpha,\ \beta$) 分布的均值和方差为

$$\text{Mean} = \frac{\alpha}{\beta} + T, \qquad \text{Var} = \frac{\alpha}{\beta^2}, \quad (25)$$

为了评估这些与生物学相关的量在数据有限的情况下受到多大约束，我们根据上述方程（22）-（24）中的约束，考察这些量的最大值和最小值。完整的数学细节见补充材料；值得注意的是，结果与 Weber等人 [27] 在简化的延迟指数情形 $\alpha = 1$ 下所得结果一致。我们将重点放在仅有 $\bar{G}_1$ 时相的情形，因此使用方程（22），并考察（$\alpha_1,\ \beta_1,\ T_1$）。其他时相也可以用完全相同的方法分析。

图 1 展示了在满足方程（ T_1 ）关系的条件下，移位 Gamma 分布的均值、方差和取值如何随 $\bar{G}_1$ 时相长度在（$\alpha_1,\ \beta_1$）参数空间内变化；其中采用表（22）所示的 BEG 比例。虽然图示参数空间的大部分区域中，均值和方差都处于其总体取值范围的低端，但我们发现，在 $\alpha_1 = 1$ 附近且 β_1 较小时，有一小片区域的均值和方差达到最大值。不过，整个参数空间内的均值彼此相差不超过 0.4 小时，因此相对于最小值，最大扰动约为 10%。

另一方面，方差在参数空间中的取值范围要大得多，较高值集中在均值也接近最大值的区域。这与Weber等人的结果一致：他们发现，仅凭BEG比例，延迟指数分布的均值就比方差受到的约束严格得多。 [27],

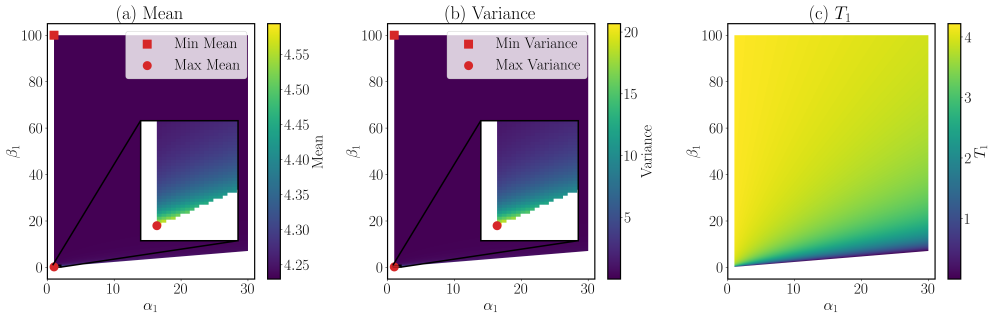

图 1 热图展示了（a）均值、（b）方差和（c）最小时相延迟的数值 T_1与 G_1 时相长度的均值数值方差较小，而方差和 T_1 在（α_1，β_1）参数空间中的差异较大。 T_1 的数值根据方程（22）使用表（1）中的BEG比例计算，随后根据方程（25）计算均值和方差。我们排除不可行区域，其中 $T_1 < 0$。

均值和方差的最大值与最小值在参数空间中出现在同一位置。观察这两个点处的 T_1值，可以看出这一影响。当均值和方差均取最小值时， T_1 远大于均值和方差均取最大值时的情况。此外，均值处于最小值时的 T_1值接近这一最小均值。这意味着，表达式（25）中的均值主要由 T_1（即固定延迟）贡献，因此也解释了为何此时 G_1 时相长度的方差较小。另一方面，在均值最大时（其值也与最小均值相对接近）， T_1 对均值的贡献要小得多。因此，由于此时均值更多地由随机伽马(α_1, β_1)分布而非固定延迟贡献，我们可以理解为何 G_1 时相长度的方差也要大得多。

在有效参数空间内改变（α_1， β_1）对均值 G_1 时相进程影响甚微，但上述结果表明，参数选择会影响底层偏微分方程的动力学。方程（1）-（4）中的偏微分方程组从任意初始分布出发，经过足够长时间后都会进入 BEG状态，但 G_1 时相长度的分布会影响达到这一稳态行为所需的时间。由于 G_1 对均值的贡献较大，时相长度的方差较小，这表明在给定时刻进入 T_1 的所有细胞，其 G_1 离开时刻将更为同步，因此整个细胞周期中的同步性也会更高；相比之下，若 T_1 对平均时相长度的贡献较小，则同步性较低。

如图2所示，我们模拟了完整的偏微分方程组（1）-（4），初始时所有细胞均处于 G_1 ，并考察 G_1阶段时长方差较小和较大的两种极端情形。正如预期， G_1 阶段时长的方差越小，系统到达 BEG 所需的时间就越长。此外，方差较小时，时相比例的中间振荡幅度要大得多。

本研究旨在为一个结构化偏微分方程组设定参数，以便随后利用它模拟细胞对不同抗肿瘤治疗的响应，并考察潜在的细胞周期动力学如何驱动这些响应。由于这些治疗可能会

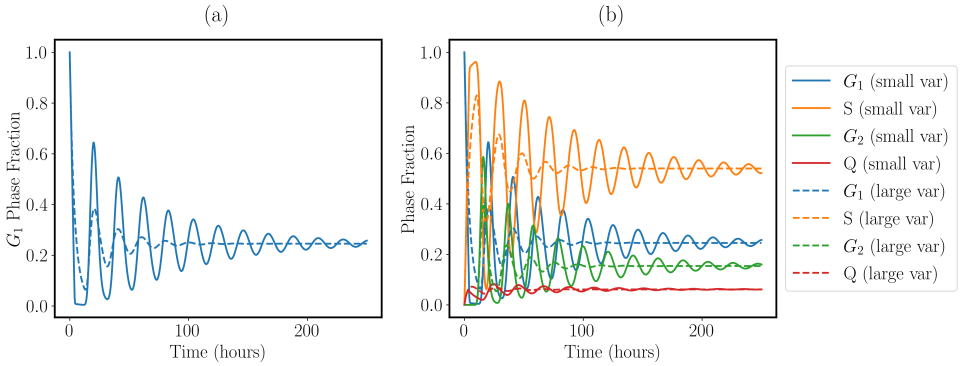

图 2 完整偏微分方程组 (1)-(4) 的模拟结果表明，与方差较大的情况相比，G_1 阶段时长取值的方差较小，会显著延迟恢复至 BEG。在两种情况下，(α_i, β_i, T_i)、 $i = 2$、3 均列于表 (4) 中。(a) 两种不同方差情况下，随时间变化的 G_1 时相比例。(b) 两种不同方差情况下，四个时相比例随时间的变化。方差较小的情况下，$(\alpha_1, \beta_1, T_1) = (1, 5, 4.03)$；方差较大的情况下，$(\alpha_1, \beta_1, T_1) = (1, 0.22, 0.046)$。两次模拟的初始条件相同，所有细胞均从 G_1开始。

通过针对特定阶段的细胞，将细胞周期动力学从 BEG 状态扰动开来，图2 展示了在图（α， β）所示的参数空间中任意选择参数的风险，1尽管这不会影响 G_1 阶段时长的均值。一旦实施治疗并扰动系统，阶段时长参数的选择会影响系统恢复稳态所需的时间。图2 显示，系统恢复至 BEG 状态的时滞可能从约 3 天到 10 天不等；对于间隔几天实施的分次治疗，这可能会影响模型的预测性能和准确性。

如上所述，对于表1中的BEG比例，图1 表明， G_1 期的平均时长可确定到0.4小时以内；当最小均值约为4.2小时时，这一差异相对较小。现在，我们考察这一结果对BEG比例变化的稳健性。对于可行¯ ¯参数空间（$G_1 + Q < 1$）中的每个 G_1 和 Q，我们使用方程（22）和（25）计算（α_1, β_1）空间内的最小和最大均值及方差，方法与上述表（1）中单组BEG比例的处理相同。图3显示了均值范围的百分比差异；该差异通过计算最大和最小均值之差占最小值的百分比，并在可行的阶段比例范围内进行计算得到。¯ ¯我们看到，对于较小的组合（G_1, Q），均值的相对范围更小。因此，如果我们仅希望使用结构化偏微分方程模型来确定平均阶段长度，而不是精确确定单个阶段持续时间分布参数，那么对于 G_1 长度而言，实验测得的 G_1 和 Q值较小的细胞系，其估计结果将更加精确，这表明该模型对某些细胞系比对其他细胞系更适用。

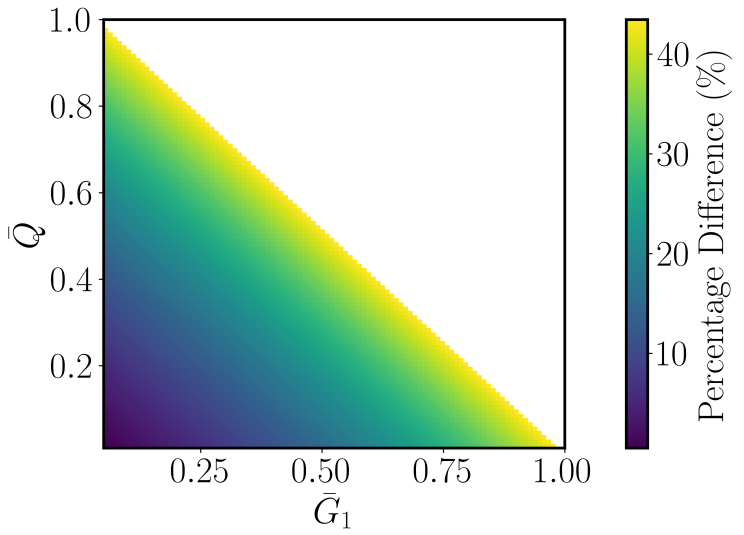

图 3 偏微分方程所捕获的 G_1 阶段时长最大均值与最小均值之间的百分比差异，会随着 BEG $\bar{G}_1$ 和 $\bar{Q}$ 比例的变化而减小。百分比差异的定义见正文。

4.2 案例2：可用数据包括 BEG比例和细胞阶段时长变异系数值

现在，我们考察这样一种情况：可以获得一些源自FUCCI数据的细胞群体水平测量值。鉴于FUCCI成像记录了单个细胞的阶段时长，对于我们的模型（1）-（4）而言，一个相关的细胞群体测量值应当是每个阶段观测到的最小时长。然而，要通过单细胞追踪实验准确估计最小阶段长度，就需要在实验过程中进行频繁成像。根据所需的分辨率，可能需要每10分钟甚至更短时间成像一次，才能捕捉到最小阶段长度。在这种情况下，我们假设成像频率不足以识别最小阶段长度。

相反，我们考虑一个受成像分辨率影响不如最小阶段时长那么大的单一细胞群体指标。一个自然的选择可能是各阶段时长的均值。然而，正如案例1中所见，仅知道BEG比例就足以确定一个范围很小的可能均值。对上文的 G_1 分析进行扩展，并对 S 和 G_2 阶段重复分析后，我们发现，对于表 $[4.23, 4.60]$1 $[12.17, 14.84]$ 中的BEG比例，平均阶段长度分别限定在 $[4.55, 4.89]$ 小时。因此，如果我们要在后续分析中纳入 FUCCI估计值的平均阶段长度，它们必须落在这些区间内，分析才可行。

由于无法获取RKO细胞的FUCCI细胞周期数据，甚至无法获得细胞群体层面的汇总数据，我们转而寻找具有类似特征的细胞系的单细胞数据，以便与表1中的BEG比例相配套。RKO细胞的特征是具有相对较长的 S期 [20]，而Chao等人的数据 [33] 表明，U2OS细胞系也具有较长的

期（平均 ≈ 10 小时， G_1 和 G_2 期约为4小时）。然而，平均 S 期长度超出了RKO细胞 $[12.17, 14.84]$ 小时的BEG推导区间，因此，U2OS细胞的平均阶段长度不适合作为后续分析的候选值。在一种情况下，如果所选细胞系具有可用的FUCCI推导平均阶段长度值，并且这些值也与BEG推导的平均时相长度区间一致，那么后续分析便可通过拟合平均时相长度值来重复进行。

由于缺乏 RKO细胞的 FUCCI汇总数据，我们采用时相时长的变异系数而非均值；在不同细胞系中，S 时相的时相持续时间变异性最低 [47, 48], ，因此根据方程（19）中的定义， S 时相的变异系数最小。不同细胞系中， G_1 和 G_2 时相长度的变异系数相对大小顺序有所不同，但二者均大于 S时相的变异系数。鉴于本研究旨在作为概念验证，因此我们采用 [33] U2OS细胞的单细胞时相长度数据来估计各时相长度的变异系数，结果见表2。

	描述	值
CV_1	G_1 时相长度的变异系数。	0.42（小时）
CV_2	S 时相长度的变异系数。	0.13（小时）
CV_3	阶段时长的变异系数 G_2 。	0.33（小时）

表 2 我们根据 Chao 等人关于 U2OS 细胞的单细胞数据计算得到的变异系数参数值。 [33]

在此案例中，我们需要拟合六个参数，即 $\vec{x} = （\alpha_1， \alpha_2， \alpha_3， T_1， T_2， T_3）$ ，其中 β_i （$i = 1$，2， 3）为

$$\beta_i = \frac{\sqrt{\alpha_i}(1 - \mathrm{CV}_i\sqrt{\alpha_i})}{\mathrm{CV}_i \times T_i}. \tag{26}$$

与案例1相同，我们也利用 BEG比例数据来确定三个参数分组，如方程（22）-（24）所示，这些方程将 T_i 表示为 α_i 和 β_i的函数。再次聚焦于 G_1 期，对每个（α_1，β_1， T_1），这两个代数关系会生成一条满足自然约束 $\alpha_1 > 1$、 $\beta_1 > 0$、$T_1 > 0$ 的参数曲线（完整的数学分析见补充材料）。通过考察沿此曲线的 $T_1 = T_1（\beta_1）$ 和 $\alpha_1 = \alpha（\beta_1）$ ，我们再次分析 G_1 阶段时长的均值和方差。

图4（a）和（b）显示， α_1 和 T_1 在β_1 的可行取值范围内变化很大；这些范围对应于表1和2中的 BEG 比例和变异系数值。不过，（c）和（d）图显示，均值的精度可达 0.002 小时，方差的精度可达 0.03 以内。为研究细胞周期时相长度而设计的 FUCCI 成像通常以 0.1–0.2 小时的间隔拍摄图像。因此，即使阶段转换时间是瞬时的（即细胞从确定被归类为某一时相，到确定被归类为下一时相所需的时间），这些实验所得的精度仍不如我们的模型估计。

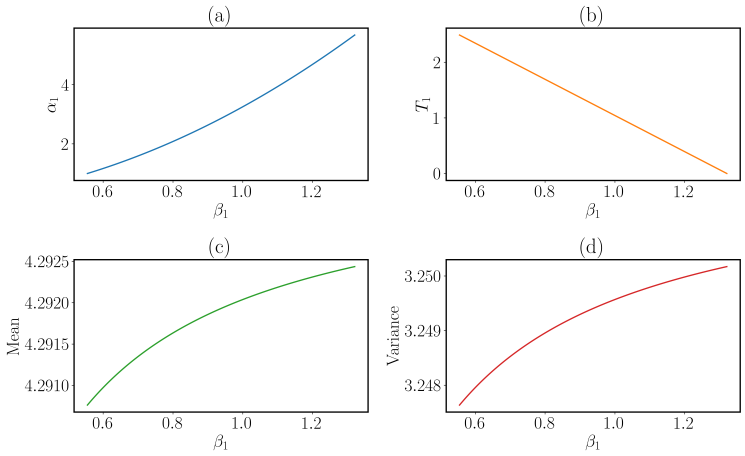

图4 阶段时长的均值和方差 G_1 在可行参数空间内保持一致，尽管 α_1、 β_1 和 T_1 会发生变化，前提是变异系数值和 BEG 值固定。(a) $\alpha_1(\beta_1)$，(b) $T_1(\beta_1)$，(c,d) 通过式（ G_1 25）计算得到的在所有可行的 β_1 值上的均值和方差。

，这些实验所得的精度仍不如我们的模型估计。

总之，时间分辨率较低的 FUCCI 成像无法估算最小阶段时长，但可以估算变异系数值和 BEG比例；即使模型不具备结构可辨识性，这些信息仍使我们能够确定底层伽马分布的前两个矩。因此，如果模型只需提供这两个矩的估计值，那么时间分辨率较低的 FUCCI 成像所得汇总数据便已足够。

4.3 情况3：可用数据包括 BEG比例、细胞阶段时长变异系数值和最小阶段长度

最后，我们考虑这样一种情况：可以获得时间分辨率更高的成像数据（每隔 $0.1-0.2$ 小时拍摄一次图像）所对应的 FUCCI 汇总数据；此类实验成像频率足够高，能够确定最小阶段长度。我们再次使用 Chao等人 [33], 的 U2OS 单细胞数据，相关值列于表 3。

	描述	值
T_1	G_1的最小阶段时长。	1.8（小时）
T_2	S的最小阶段时长。	7.7（小时）
T_3	G_2的最小阶段时长。	1.7（小时）

表3 最小值的参数值
阶段时长，这是我们根据
Chao等人关于 U2OS 细胞的单细胞数据计算得出的。 [33]

4.3.1 拟合的唯一性

根据表中给出的 BEG 阶段比例、最小阶段长度和变异系数值的估计值 1、2 和 3，我们的优化问题是使下式最小化

$$n(\vec{x}) = (\bar{G}_1(\vec{x}) - \bar{g}_1)^2 + (\bar{S}(\vec{x}) - \bar{s})^2 + (\bar{G}_2(\vec{x}) - \bar{g}_2)^2, \tag{27}$$

其中 $\vec{x} = (\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3)$ 是六个未知参数。我们使用 SciPy 的差分进化 函数进行最小化。该方法会搜索参数空间的大范围区域，并且可以并行化以提高计算效率 [49]。

算法只需拟合 α_i，$i = 1$，2、3，因为 β_i 可通过方程（26）计算得出。我们对 $1 < \alpha_i < 1/CV_i^2$ 在 $i = 1$，2、3 时施加约束，使得 β_i 以及方程（ $\mu_i(a)$ 中的系数在方程（26）和（9）中分别为正且有限。

我们运行最小化算法 100 次，以确认满足表4所示最小化问题的参数集是唯一的。该唯一参数集的 $n(\vec{x})$ 为 $< 10^{-30}$，其中 $n(\vec{x})$ 由方程（27）给出。由于 T_i由表3中的值指定，而 β_i 由方程（26）定义，这保证了对于任意给定的 α_i，延迟伽马分布都具有所需的变异系数值，因此， $n(\vec{x})$ 的值衡量了由表达式（12）—（14）计算出的各 BEG 阶段比例与表1所示数据的接近程度。方程（27）中的每个平方项都是一个解析BEG相比例与对应数据值之差，二者都必须介于 0 和 1 之间。因此， $n(\vec{x})$ $< 10^{-30}$ 表明各解析阶段比例与 BEG 数据高度吻合。

	描述	值
α_1	伽马分布的形状参数 G_1 。	1.92
α_2	伽马分布的形状参数 S 。	8.08
α_3	伽马分布的形状参数 G_2 。	3.63
β_1	伽马分布的速率参数 G_1 。	0.78
β_2	S 伽马分布的速率参数。	1.79
β_3	G_2 伽马分布的速率参数。	1.26

表4 将模型拟合至表1中的 BEG阶段比例，以及表2 和3中的变异系数和最小阶段长度后，得到的伽马分布参数唯一组合。

4.3.2 最小阶段时长对拟合优度的影响

如上所示，表 T_i 中的值3 能够使模型参数与 BEG 数据实现极佳拟合。现在，我们考虑在最小化 T_i 时，探查 $n(\vec{x})$ 对最大拟合优度的影响，见方程（27）。

为了直观展示 T_i 对模型拟合质量的影响，我们通过合并 S 和 G_2 阶段来构建一个简化模型，从而无需使用模型中的这些参数

(α_3, β_3, T_3)。这样一来，系统变为

$$\frac{\partial G_1}{\partial t} + \frac{\partial G_1}{\partial a} = -\mu_1(a)G_1(a,t) + \beta_4(P(t))Q(a,t) - \beta_5(P(t))G_1(a,t), \tag{28}$$

$$\frac{\partial S}{\partial t} + \frac{\partial S}{\partial b} = -\mu_2(b)S(b,t), \tag{29}$$

$$\frac{\partial Q}{\partial t} = -\beta_4(P(t))Q(a,t) + \beta_5(P(t))G_1(a,t), \tag{30}$$

with

$$G_1(0,t) = 2\int_0^\infty \mu_2(b)S_2(b,t)\,db, \qquad S(0,t) = \int_0^\infty \mu_1(a)G_1(a,t)\,da. \tag{31}$$

对于这个简化模型，有三个 BEG 比例，我们的目标仅是估计 $i = 12$ 的 (α_i, β_i)。

在图 5中，我们将 G_1 和 $S/G_2/M$ 阶段时长的变异系数分别固定为 0.42 和 0.13，并将 BEG 比例分别设为 0.247 和 0.693。我们展示了当 T_1 和 T_2 变化时，拟合 (i =1，2) 的 (α_i, β_i) 时拟合优度如何变化。对于所有 (T_1, T_2) 参数对，差分进化算法都收敛到全局最小值。然而，检查这些参数集对应的 $n(\vec{x})$ 值后发现，在参数空间的大部分区域内，拟合效果都很差。具体而言，我们将拟合参数集下评估得到的 $n(\vec{x}) > 0.0004$ 定义为“不可接受”的拟合。如果允许 $n(\vec{x}) = 0.0004$，那么 BEG 阶段比例与数据的差异可能高达 0.02。这个定义在很大程度上是任意的，但考虑到静止期 BEG 比例 0.061 与其相差 0.02 就意味着 33% 的误差，因此这一界定是合理的。可以根据所需的 BEG 比例调整此定义。在这里，我们认为任何 $n(\vec{x}) \in [10^{-6}, 0.0004]$ 值都属于可接受的拟合，其中下限对应于数据与拟合 BEG 比例之间的最大差异为 0.001；低于此值则属于最优拟合。

事实上，图5 表明，在拟合数据不可接受的最小阶段长度与能够实现最优拟合的最小阶段长度之间，界限很窄（将最优拟合定义为 $n(\vec{x}) < 10^{-6}$）。这些模拟显示，允许实现这一最优拟合的 (T_1, T_2) 空间矩形区域约为 (T_1, T_2) $\in$([0, 4.13], [0, 16.60])。因此，我们看到 (T_1, T_2) 参数空间中存在一个定义明确的区域，可对数据进行最优且唯一的拟合。

此外，与 Weber等人 [27], 一样，我们可以为 T_i在四阶段模型中推导出一个解析上界。例如，对方程（12）关于 β_1 重新排列，并注意到所得表达式随 T_1 递减，且渐近线位于 $T_1 = T_1^* > 0$，我们得到

$$T_1 < T_1^* = -\frac{1}{\lambda + f(\lambda)} \log\left(\frac{2 - \bar{G}_1 - \bar{Q}}{2}\right). \tag{32}$$

同样地，我们可以得到 T_2 和 T_3 的以下界限:

$$T_2 < T_2^* = -\frac{1}{\lambda} \log\left(\frac{2 - \bar{G}_1 - \bar{Q} - \bar{S}}{2 - \bar{G}_1 - \bar{Q}}\right), \quad T_3 < T_3^* = \frac{1}{\lambda} \log(1 + \bar{G}_2). \tag{33}$$

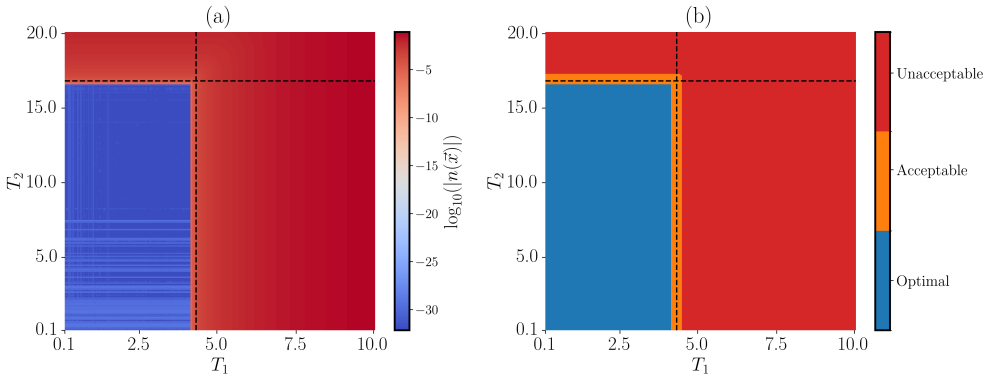

图 5 在简化的两阶段细胞周期模型中改变所施加的最小阶段长度（T_1，T_2），可在参数空间中确定一个矩形区域，在该区域内模型能够对预设的 BEG 阶段比例进行出色且唯一的拟合。**(a)** 拟合参数下 $n(\vec{x})$ 对数变换值的热图；**(b)** 按照正文中的定义，将参数空间中的每个点分类为对 BEG 比例的“不可接受”“可接受”或“最优”拟合。黑色虚线表示由方程（32）和（33）定义的解析界限 T_1^*和 T_2^*。

在图（5）所考虑的三阶段模型中，我们可以设定 $G_2 = 0$ 并得到$T_1 < 4.25$ 以及 $T_2 < 16.72$。我们可以看到，最佳拟合区域包含在这些界限之内。

总之，假设我们已知各阶段时长的变异系数以及通过实验获得的 BEG 比例，那么为了使模型（1）-（4）获得最优参数拟合，我们能够施加的最小细胞周期时相长度 T_i的取值范围就受到限制。根据方程（32）和（33），我们可以根据已知的 BEG 比例推导出 T_i的上界。然而，图 5 的结果表明，要使数据拟合达到最优，T_i 的取值必须受到进一步限制。如果数据中的 T_i 不在这一受限区域内，就需要做出选择：是优先精确拟合 BEG 比例（采用 T_i 值，使 $n(\vec{x})$ 较低），还是允许 BEG 比例偏离测量值，以便使用精确的 T_i数据值。

4.3.3 拟合含噪生物数据

以上结果表明，通过固定变异系数值和最小阶段长度，可以确定唯一的最佳拟合参数集；只要恰当选择最小阶段长度，拟合结果就会非常吻合。我们据此得出结论：在本例中，该模型是结构可辨识的 [42]。现在，我们将考察含噪数据，以检验实际可辨识性 [42]。这里的含噪数据是 Celora等人 [20]提供的 RKO细胞 BEG比例的九次重复测量。

由于 Celora等人 [20] 的实验数据中缺少静止期比例，我们首先考虑以适当的方式模拟这些比例。鉴于数据是通过将每个细胞周期阶段的细胞数除以细胞群体中的细胞总数计算得出的，因此，描述细胞在四个阶段间分布的合理候选模型是二项模型 [50]。当细胞总数固定时，该模型可以捕捉比例可能出现的离散组合。不过，

G_1	S	G_2	Q
0.2556	0.5570	0.1518	0.0356
0.2305	0.5338	0.1685	0.0672
0.2214	0.6021	0.1216	0.0549
0.3103	0.4757	0.1324	0.0817
0.1920	0.5400	0.1900	0.0780
0.2675	0.4985	0.1476	0.0864
0.2465	0.5157	0.1879	0.0499
0.2386	0.5714	0.1391	0.0510
0.2480	0.5607	0.1475	0.0438

表5 基于 Celora 等人数据的细胞周期阶段比例 [20], , 并扩展以纳入静止期区室。

由于数据中缺少细胞总数，我们使用狄利克雷分布 [51]。这使我们能够直接对比例组合进行抽样，因为该分布将比例之和必须为一这一条件内置其中。由此生成的 计算机模拟数据集见表 5，有关如何利用狄利克雷分布模拟静止期比例的完整细节，请参阅补充材料。

我们采用贝叶斯推断方法，评估该模型针对这一扩展计算机模拟数据集的实际可辨识性；完整细节请参见补充材料。

我们使用马尔可夫链蒙特卡罗（MCMC）方法，通过 Python软件包PINTS（噪声时间序列上的概率推断）估计未知参数（α_1、 α_2、 α_3）的后验分布 [52]。图 6 中的边缘后验分布表明，每个参数都围绕单个明确定义的峰值呈对称分布。我们得出结论：根据表 5 [53, 54]中的数据，可以实际辨识这些参数。此外，每个边缘后验分布的均值与通过对表 1中的数据取平均细胞周期阶段比例，并最小化方程（27）而得到的参数值高度一致，即与通过最小化平方误差和得到的 α_i 逐点估计一致。这表明，对于这些数据，取表（5）中 BEG 数据的均值，并通过最小化方程（ α_i （$i = 1$、2、3） n（$\vec{x}$）中的 27）来拟合 α_i ，等价于通过完整的 MCMC 分析获得的

我们进一步通过对每个参数进行剖面似然分析，检验这些参数的实际可辨识性（α_1， α_2， α_3），结果如图 7[55]（完整细节见补充材料）。通过这一分析，我们可以利用似然比统计量 [42, 55]估计参数值的置信区间。在图7中，水平线表示每个参数最大似然估计的 95%置信区间，其定义如下：

$$\mathrm{CL}_{95}(\alpha_i) = \{\alpha_i : |\mathrm{PL}(\alpha_i) - \mathrm{PL}(\hat{\alpha}_i)| < \chi^2_{1,0.95}\}, \tag{34}$$

其中，$\hat{\alpha}_i = \arg\max^{\alpha_i \in [0, 1/CV_i^2]} \mathrm{PL}(\alpha_i)$ 是 α_i 的最大似然估计，$\mathrm{PL}(\alpha_i)$ 是 α_i 的剖面似然函数。参数 α_i 被称为

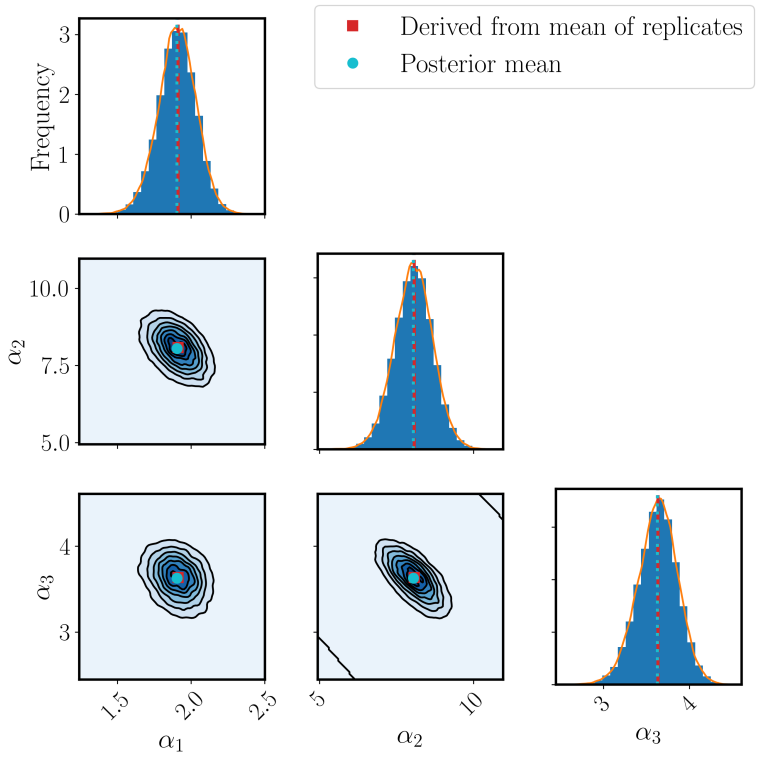

图 6 参数的边缘后验分布 α_1，α_2，α_3 表明每种情况下都具有实际可辨识性。对于对角线上的每个分布，红色虚线对应于通过最小化方程（27）得到的参数值，其中输入比例为表（5）中每列的均值。蓝色点线表示边缘后验分布的均值。非对角线图展示了每一对 α_i 的联合分布等高线。

在实际意义上可辨识，如果区间 $\mathrm{CL}_{95}(\alpha_i)$ 有限。我们发现，对于每个 α_i，$i \in \{1, 2, 3\}$ 情况确实如此，因此每个参数都可以根据表 5 中的数据在实际意义上辨识。

5 讨论

我们的分析表明，在细胞周期数学模型的参数化过程中，群体平均数据的可用性会显著影响模型的可辨识性。然而，我们发现，即使可用数据有限，仍可辨识出关键参数组合，这些组合对应于阶段长度分布的第一阶矩。

由于不同细胞系的细胞周期实验通常缺少可用的时间序列数据，数学建模者往往必须汇集多个来源的汇总数据来参数化模型。由于细胞周期数据主要来自群体层面的实验（DNA染色流式细胞术）和单细胞阶段追踪（FUCCI）实验，

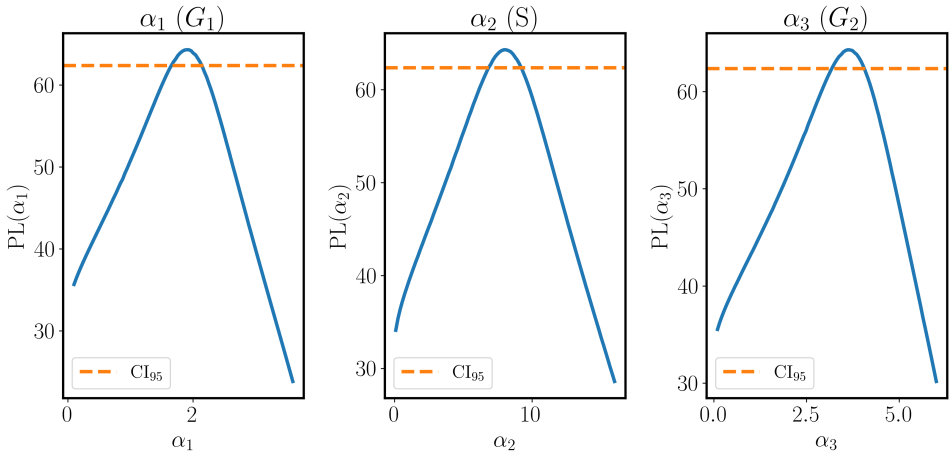

图 7 每个模型参数的剖面似然分析， α_1， α_2 和 α_3，同样表明了实际可辨识性。每个图均以蓝色曲线显示模型参数 α_1、 α_2 和α_3 的剖面似然。橙色虚线与剖面似然的交点表示由方程（34）定义的95%置信区间界限。

合适的数学模型会以阶段持续时间来描述细胞周期动力学。我们提出的年龄结构偏微分方程表述采用具有时滞的伽马分布阶段长度，既便于解析处理，也能提供足够的细节来整合群体层面和单细胞测量。此外，这一表述自然产生平衡指数增长状态，这对于反映细胞群体中的实验观测至关重要 [39, 56], ，同时也提供了一种自然的方式，将模型与更简单、往往更容易获取的流式细胞术数据联系起来。

本研究表明，由于缺乏适当的数据而导致的结构不可辨识性，并不会妨碍模型的实用性。对于数据量不足以支持唯一参数估计的情况，我们仍然能够恢复具有生物学意义的量的估计值，例如每个细胞周期阶段持续时间的均值和方差。因此，我们得出结论：与单个分布参数相比，这些阶段驻留时间分布的矩具有更稳健的可辨识性。

此外，即使模型在结构上似乎可辨识，即存在一个使目标函数最小化的唯一参数集，参数拟合的质量仍取决于不同数据之间的一致性。具体而言，通过流式细胞术获得的 BEG比例足以确定一个相对较窄的可能平均阶段长度范围。因此，这些比例为能够产生良好拟合的单细胞最短阶段长度数据设定了界限。当这些数据彼此不一致时，基于拼接数据的这一框架自然会产生较差的拟合。相较之下，我们的框架指出，使用一种在不同细胞系之间更为一致的度量（例如变异系数），可以解决这一问题。

本研究说明，在寻找实验数据时，考虑模型的预期用途十分重要。如果目标是估计细胞周期时相长度的均值，那么在本文提出的框架下，仅获取 BEG比例就足以估计这些值。然而，在选择个别模型参数以供进一步使用（例如模拟治疗）时，必须谨慎。参数空间中阶段时长的方差存在较大差异，会导致细胞恢复 BEG生长时呈现不同的瞬态动力学，从而影响对分次治疗模拟的预测。

当需要阶段时长方差或单个模型参数等额外信息时，就需要更详细的实验数据。我们使用了时相长度的变异系数和最小时相长度，这两项均根据单细胞FUCCI追踪数据计算得出。这类数据在实验中收集起来更复杂、成本更高且耗时更长，而且文献中往往没有公开。因此，我们认为，在获取所需数据的难易程度与通过这一年龄结构偏微分方程模型了解底层过程所能获得的信息量之间，存在一种权衡。

由于提供数据的癌细胞群体不存在接触抑制，分析假设可以忽略密度依赖性细胞周期进程效应。因此，后续研究可尝试扩展我们的分析，将正常细胞纳入考量，因为密度依赖性在细胞周期调控中起着至关重要的作用。然而，在推导BEG阶段比例表达式时，这很可能会降低解析可处理性，并增加重复分析所需的计算成本。如果将来能够获得合适的数据，我们还希望针对单一细胞系重复第 4.3.3 节所述的实际可辨识性分析，同时获取实验测得的静止期比例和单细胞分析数据。

当然，鉴于 Weber等人、Billy等人和Sherer等人提供了分辨率良好的时间序列数据 [27–29], ，我们会采用这些研究中的方法来参数化年龄依赖性阶段转变率。然而，本文提出的框架是由数据限制所驱动，而非旨在为完整数据集带来方法上的进展。通过逐步增加假设在各个案例中可用的群体汇总数据量，我们的框架表明，汇集不同细胞系和实验设置的数据，能够有效地为细胞周期的年龄结构偏微分方程模型进行参数化。

A 补充信息

A.1 根据可识别参数分组优化均值和方差

A.1.1 案例1：仅有 BEG 数据

如正文所述，检查方程（12）–（17）可知，我们可以找到三组可识别的参数组，即

$$\eta_1(\alpha_1,\beta_1,T_1)=\left(\frac{\beta_1}{\beta_1+L}\right)^{\alpha_1}\exp(-LT_1)=1-\frac{\bar{Q}+\bar{G}_1}{2}, \tag{35}$$

$$\eta_2(\alpha_2,\beta_2,T_2)=\left(\frac{\beta_2}{\beta_2+\lambda}\right)^{\alpha_2}\exp(-\lambda T_2)=\frac{1+\bar{G}_2}{2-\bar{Q}-\bar{G}_1}, \tag{36}$$

$$\eta_3(\alpha_3,\beta_3,T_3)=\left(\frac{\beta_3}{\beta_3+\lambda}\right)^{\alpha_3}\exp(-\lambda T_3)=\frac{1}{1+\bar{G}_2}, \tag{37}$$

其中 $L=\lambda(1+Q/G_1)$。现在我们希望考虑参数值 $(\alpha_i,\ \beta_i,\ T_i)$，$i=1,2,3$，使伽马分布的均值和方差分别达到最小值和最大值，同时满足 $\eta_i(\alpha_i,\ \beta_i,\ T_i)=\eta_i$ 为常数值。

为简洁起见，我们仅介绍 G_1 参数 $(\alpha_1,\ \beta_1,\ T_1)$ 的方法，并指出其他参数集 $(\alpha_i,\ \beta_i,\ T_i)$，$i=2,3$ 也可用相同方法进行分析。

对于已知的固定值 Q 和 G_1，我们可以重新整理方程（35）以得到

$$T_1=\frac{\alpha_1\log\left(\frac{\beta_1}{\beta_1+L}\right)-\gamma_1}{L}, \tag{38}$$

其中 $\gamma_1=\log(1-(Q+G_1)/2)<0$。

利用形状参数为 α、速率参数为 β 且时滞为 T 的延迟伽马分布的均值为 $\alpha/\beta+T$ 这一事实，我们定义

$$m(\alpha_1,\beta_1)=\alpha_1\left(\frac{1}{\beta_1}+\frac{\log(\frac{\beta_1}{\beta_1+L})}{L}\right)-\frac{\gamma_1}{L}. \tag{39}$$

因此，我们的问题简化为在 $\beta_1>0$ 和 $\alpha_1>1$ 上对 $m(\alpha_1,\beta_1)$ 求最小值或最大值。通过施加 $T_1>0$，还可以导出另一个自然边界，由此可得

$$\alpha_1<\frac{\gamma_1}{\log(\frac{\beta_1}{\beta_1+L})}, \tag{40}$$

并将其与 $\alpha_1>1$ 结合，我们还可得到

$$\beta_1>\frac{L}{\exp(-\gamma_1)-1}. \tag{41}$$

计算偏导数可知，$m(\alpha_1,\beta_1)$ 随 α_1 增大而增大，随 β_1 增大而减小。综合以上结果，我们发现

$$\min m(\alpha_1,\beta_1)=-\frac{\gamma_1}{L}\qquad\text{at}\quad\alpha_1=1,\beta_1=\infty, \tag{42}$$

and

$$\max m(\alpha_1, \beta_1) = \frac{\exp(-\gamma_1) - 1}{L} \qquad \text{at} \quad \alpha_1 = 1, \beta_1 = \frac{L}{\exp(-\gamma_1) - 1}, \tag{43}$$

从而可以计算最小值和最大值的范围，如正文图 3 所示。我们注意到，这些最小值和最大值的表达式与 Weber等人 [27], 考虑的延迟指数模型（等同于我们延迟伽马表述中的 $\alpha = 1$ ）一致。

A.1.2 案例2：可用数据包括 BEG比例和变异系数值

当我们还假设已知各时相长度对应的变异系数值时，(α_1, β_1) 的可行参数空间会缩小。根据平移伽马分布变异系数的定义，我们得到

$$T_1 = \frac{\sqrt{\alpha_1}(1 - CV_1\sqrt{\alpha_1})}{CV_1\beta_1}. \tag{44}$$

将其与方程（38）结合，对于给定的 α_1 ，我们得到以下两个可能的 β_1值:

$$\alpha_1^{\pm}(\beta_1) = \left(\frac{\frac{1}{\beta_1 CV_1} \pm \sqrt{\frac{1}{(CV_1\beta_1)^2} + \frac{4\gamma_1}{L}\left(\frac{H(\beta_1)}{L} + \frac{1}{\beta_1}\right)}}{2\left(\frac{H(\beta_1)}{L} + \frac{1}{\beta_1}\right)} \right)^2, \tag{45}$$

其中 $H(\beta_1) = \log(\beta_1/(\beta_1 + L))$。

与案例1一样，我们需要 $\beta_1 > 0$、 $T_1 > 0$ 和 $\alpha_1 > 1$。约束 $\beta_1 > 0$ 意味着 $\alpha_1 < 1/CV_1^2$，经过一些变换后可知，对于所有 $\beta_1 > 0$， $\alpha_1^+(\beta_1)$ 都不可行。

注意到 $\alpha_1^-(\beta_1)$ 随 β_1递增，我们的可行区域受到以下约束: $\beta_1 \in [\hat{}\beta_1, \beta_1^*]$，其中 $\alpha_1^-(\hat{}\beta_1) = 1$ 且 $\alpha_1^-(\beta_1^*) = 1/CV_1^2$。求解方程（45），我们发现 $\hat{}\beta_1$ 是以下方程的唯一正解:

$$1 + \frac{L}{\beta_1} = \exp\left(\frac{L(CV_1 - 1)}{\beta_1 CV_1} - \gamma_1\right), \tag{46}$$

and

$$\beta_1^* = \frac{L}{\exp(-CV_1^2\gamma_1) - 1}. \tag{47}$$

$\hat{}$ 可以证明，对于所有 $\beta_1 \in [\beta_1, \beta_1^*]$， $T_1 > 0$。

A.2 模拟计算机模拟静止比例

如正文所述，RKO细胞周期时相数据不包含静止期细胞比例。为生成这些比例的计算机模拟估计值，我们假设每次

G_1	S	G_2
0.2912	0.5570	0.1518
0.2978	0.5338	0.1685
0.2763	0.6021	0.1216
0.3919	0.4757	0.1324
0.2700	0.5400	0.1900
0.3539	0.4985	0.1476
0.2964	0.5157	0.1879
0.2895	0.5714	0.1391
0.2918	0.5607	0.1475

表6 Celora等人记录的平衡指数生长期比例 [20]。

$^{-}$ $^{-}$ $^{-}$ $^{-}$比例测量值（G_1， S， G_2， Q）均来自狄利克雷分布[51]。

更具体地，对于维数为 M 的狄利克雷分布，我们写作

$$(X_1, X_2, X_3, ..., X_M) \sim \mathrm{Dir}(c_1, c_2, c_3, ..., c_M), \tag{48}$$

其中

$$\mathbb{E}(X_i) = \frac{c_i}{c_0}, \qquad \mathrm{Var}(X_i) = \frac{c_i(c_0 - c_i)}{c_0^2(1 + c_0)}, \qquad \text{for } i = 1, 2, 3, ...M, \tag{49}$$

以及 $c_0 = \sum_{i=1}^{i=M} c_i$。

我们首先对表6所示9次实验重复的基础狄利克雷分布参数进行近似。我们确定 S 和 G_2 阶段比例的标准差分别为 $\sigma_S = 0.038794$ 和 $\sigma_{G_2} = 0.023666$。Corvaisier等人 [41] 的实验结果表明，$^{-}$ $^{-}$ $^{-}$ $^{-}$静止期的标准差约为 $\sigma_Q = 0.018$。我们假设$(G_1, S, G_2, Q) \sim \mathrm{Dir}(c_1, c_2, c_3, c_4)$，其中 $c_1 = \bar{g}_1 c_0$、 $c_2 = \bar{s} c_0$、 $c_3 = \bar{g}_2 c_0$ 和 $c_4 = \bar{q} c_0$（$(\bar{g}_1, \bar{s}, \bar{g}_2, \bar{q})$定义于正文表1中），并且选择 c_0，使观测标准差与计算标准差之间的差异在最小二乘意义下最小。更具体地，利用公式（49），我们选择 c_0 以最小化

$$\mathrm{Error}_{SD}^2(c_0) = \left(\sigma_S - \sqrt{\frac{\bar{s}(1-\bar{s})}{1+c_0}}\right)^2 + \left(\sigma_{G_2} - \sqrt{\frac{\bar{g}_2(1-\bar{g}_2)}{1+c_0}}\right)^2 + \left(\sigma_Q - \sqrt{\frac{\bar{q}(1-\bar{q})}{1+c_0}}\right)^2. \tag{50}$$

我们估计 $c_0 \approx 141.3$。利用狄利克雷分布的条件分布性质 [51]，我们进一步推导出

$$\frac{1}{s + g_2}(\bar{G}_1, \bar{Q}) | (\bar{S} = s, \bar{G}_2 = g_2) \sim \mathrm{Dir}(c_1, c_4), \tag{51}$$

因此，给定 s 和 g_2 各重复实验对应的值（见表6），我们可以从上述分布中抽样，以估计相应的 G_1 和 Q比例。为确保九个样本具有 $\bar{q} = 0.061$ 和 $\sigma_Q = 0.018$，且这些值可复现，我们在抽样前生成 10,000 个随机种子，并选取能使.....最小的随机种子

$$\text{Error}^2_{\text{mean}} = (\bar{q} - 0.061)^2 + (\sigma_Q - 0.018)^2, \tag{52}$$

其中，$\bar{q}$ 和 σ_Q 分别为样本均值和标准差。此过程生成了正文表5中的计算机模拟数据集。

A.3 贝叶斯推断

我们采用贝叶斯推断方法，评估模型对正文表5中计算机模拟数据的实际可辨识性。

按照 Celora 等人 [20], 的记法，我们用 F_j^i（$j \in$ (1，2，3，4)）表示表5中第 i行、第 j列的测量值。由于我们假设变异系数、最小阶段长度和倍增时间均为固定值，因此未知参数集为$\theta = (\alpha_1, \alpha_2, \alpha_3)$。我们将表5中的数据汇总到集合 $\mathcal{E} = \{(F_1^i, F_2^i, F_3^i, F_4^i)\}_{i=1}^{9}$，并用 $\vec{w}(\theta) = (G_1(\theta) \; S(\theta) \; G_2(\theta) \; Q(\theta))$ 表示参数值为 θ时的细胞周期阶段比例。

假设表5中的数据来自狄利克雷分布模型下 $\vec{w}(\theta)$ 的含噪实现，我们还在贝叶斯程序中考虑一个未知的 c_0 参数，其定义见方程（49）。我们假设 $\log(c_0)$ 在 $[\log(2), \log(100000)]$ 上服从均匀分布（请注意，c_0 控制狄利克雷分布的方差），并对 α_i $(i = 1, 2, 3)$ 的 $[0.1, 1/CV_i]$ 使用均匀先验。

因此，该问题的对数似然为

$$\ell(\theta, c_0) = 9\log(\Gamma(c_0)) - 9\sum_{j=1}^{4}\log(\Gamma(c_j(\theta))) + \sum_{i=1}^{9}\sum_{j=1}^{4}(c_j(\theta) - 1)\log(F_j^i), \tag{53}$$

其中 $(c_1(\theta) \; c_2(\theta) \; c_3(\theta) \; c_4(\theta)) = c_0\vec{w}(\theta)$。

我们在 Python 软件包 PINTS（含噪时间序列概率推断）[52] 中实现了该先验和对数似然，并采用马尔可夫链蒙特卡罗（MCMC）方法估计相应的后验分布。为初始化链，我们使用 (θ, c_0) 的最大后验（MAP）估计值；该估计值通过最小化由对数似然（53）以及均匀／对数均匀先验得到的负对数后验来计算。对于 MCMC 算法，我们使用 PINTS 库中的 `HaarioBardenetACMC` 运行三条链，迭代 100,000 次，并将其中 5000 次迭代设为预烧入。

为评估MCMC运行的结果，我们首先计算Gelman-Rubin ^统计量 R，以及 θ 和 c_0中每个参数的有效样本量（有效样本量）。对于^每个参数，R 通过比较后验样本的链内方差与链间方差来表明MCMC链是否混合良好，^ ^随着迭代次数的增加，对于一次收敛的运行，$R \to 1$。我们计算 R

为（1.00022627,1.00008,1.00006592, 1.00015525），分别对应参数（$\alpha_1, \alpha_2, \alpha_3, c_0$）。这些值低于建议的上限阈值1.01，表明这3条链混合良好。有效样本量也被用于评估链混合。由于MCMC算法的连续样本存在自相关，有效样本量是指能够产生相同信息的等效独立样本数。因此，较高的有效样本量代表更好的链混合。我们计算得到的有效样本量值为（20000,20908, 21041,17028），再次高于建议阈值400。

对于后续的剖面似然分析，我们在 α_i (请注意， α_i 的支撑集为 $[0, 1/CV_i^2]$。我们在每个取值处固定α_i ，并通过最大化方程（ c_0 和 α_j， $j \neq i$）中的对数似然 $\ell(\theta, c_0)$，以拟合其余参数。方程（53）。对于 α_i的每个取值，剖面似然定义为

$$\mathrm{PL}(\alpha_i) = \max_{\theta_{-i}, c_0} \ell(\theta, c_0), \tag{54}$$

其中 $\theta_{-i} = \theta \backslash \{\alpha_i\}$。

致谢。RN 获工程与物理科学研究委员会资助（资助编号 EP/W524311/1）。

声明

经费

RN 获工程与物理科学研究委员会资助（资助编号 EP/W524311/1）。为实现开放获取，作者已对因本次投稿而产生的任何作者接受稿件采用知识共享署名许可协议（ CC BY）公共版权许可。

利益冲突

作者声明，与本文内容相关的利益冲突均不存在。

代码可用性

如需获取用于生成本文图表的代码，请联系 ruby.nixson@maths.ox.ac.uk。

作者贡献

所有作者均参与了研究构思与设计。Ruby E. Nixson 负责正式分析与调查。论文初稿由 Ruby E. Nixson 撰写，所有作者均对论文的先前版本提出了意见。所有作者均已阅读并批准最终稿。

参考文献

[1] Matthews HK, Bertoli C, De Bruin RAM. 癌症中的细胞周期调控。《自然综述：分子细胞生物学》。2022年1月;23(1):74–88. https://doi.org/10.1038/s41580-021-00404-3.[2] Pawlik TM, Keyomarsi K. 细胞周期在介导放射治疗敏感性中的作用。《国际放射肿瘤学·生物学·物理学杂志》。2004年7月;59(4):928–942. https://doi.org/10.1016/j.ijrobp.2004.03.005.[3] Lonati L, Barbieri S, Guardamagna I, Ottolenghi A, Baiocco G. 辐射诱导的细胞周期扰动：经流式细胞术数据验证的计算工具。《科学报告》。2021年1月;11(1):925. https://doi.org/10.1038/s41598-020-79934-3.[4] Muz B, De La Puente P, Azab F, Azab AK. 缺氧在癌症进展、血管生成、转移及治疗耐药中的作用。缺氧。2015年12月;第83页。https://doi.org/10.2147/HP.S93413.[5]Yashar CM. 妇科放射治疗基本原理。载于：《临床妇科肿瘤学》。爱思唯尔；2018年。第586–605.e3页。[6] Schwartz GK, Shah MA. 靶向细胞周期：癌症治疗的新方法。《临床肿瘤学杂志》。2005年12月;23(36):9408–9421. https://doi. org/10.1200/JCO.2005.01.5594.[7] Sun Y, Liu Y, Ma X, Hu H. 细胞周期调控对化疗的影响。《国际分子科学杂志》。2021年6月;22(13):6923. https: //doi.org/10.3390/ijms22136923.[8] Otto T, Sicinski P. 细胞周期蛋白：癌症治疗中前景广阔的靶点。《自然综述：癌症》。2017年2月;17(2):93–115. https://doi.org/10.1038/ nrc.2016.138.[9] Pozarowski P, Darzynkiewicz Z. 通过流式细胞术分析细胞周期。载于：Sch¨onthal AH，编。《检查点调控与癌症：第2卷：激活与调控实验方案》。新泽西州托托瓦：Humana出版社；2004年。第301–311页。[10] Cecchini MJ, Amiri M, Dick FA. 哺乳动物细胞中的细胞周期位置分析。《可视化实验杂志：JoVE》。2012年1月;(59):3491. https: //doi.org/10.3791/3491.[11] Kim KH, Sederstrom JM. 使用流式细胞术检测细胞周期状态。《分子生物学现行实验方案》。2015;111(1):28.6.1–28.6.11. https://doi.org/10.1002/0471142727.mb2806s111.[12] Terry NHA, Whiter RA. 第16章：通过分析溴脱氧尿苷掺入估算细胞周期动力学。载于：《细胞生物学方法》。第63卷

《细胞计量学》。学术出版社；2001年，第355–374页。

[13] Eminaga S, Teekakirikul P, Seidman CE, Seidman JG. 通过免疫荧光染色和显微成像检测石蜡包埋组织切片中的细胞增殖标志物. 《分子生物学现行实验方案》. 2016年7月;115:14.25.1–14.25.14. https://doi.org/10.1002/cpmb.13.[14] Miller I, Min M, Yang C, Tian C, Gookin S, Carter D, 等. Ki67是增殖与静止期的分级标志物, 而非二元标志物. 《细胞报告》. 2018年7月;24(5):1105–1112.e5. https://doi.org/10.1016/j.celrep.2018.06.110.[15] Smith PJ, Darzynkiewicz Z, Errington RJ. 细胞核计量学与染色质组织. 《细胞计量学A辑》. 2018年;93(8):771–784. https://doi.org/10.1002/ cyto.a.23521.[16]Bradford JA, Whitney P, Huang T, Pinson P, Cheung CY, Yue S, 等. 新型Vybrant® DyeCycle TM染料可利用紫光、蓝光和绿光激发的流式细胞术对活细胞进行细胞周期分析. 《血液》. 2006年11月; 108(11):4234. https://doi.org/10.1182/blood.V108.11.4234.4234.[17] Ma C, Gurkan-Cavusoglu E. 指导癌症治疗策略的计算细胞周期模型全面综述. 《系统生物学与应用》. 2024年7月;10(1):71. https://doi.org/10.1038/s41540-024-00397-7.[18] Smith JA, Martin L. 细胞会进行周期循环吗? 《美国国家科学院院刊》. 1973年4月; 70(4):1263–1267. https://doi.org/10.1073/pnas.70.4.1263.[19] Falc´o C, Cohen DJ, Carrillo JA, Baker RE. 量化组织拥挤对细胞周期调控的作用. 《生物物理学杂志》. 2025 年 3 月 ;124(6):923–932. https://doi. org/10.1016/j.bpj.2024.05.003.[20] Celora GL, Bader SB, Hammond EM, Maini PK, Pitt-Francis JM, Byrne HM. 周期性缺氧中细胞周期进程的DNA结构化数学模型. 《理论生物学杂志》. 2022年7月; 545:111104. https://doi.org/ 10.1016/j.jtbi.2022.111104.[21] Altinok A, Gonze D, L´evi F, Goldbeter A. 细胞周期的自动机模型. 《界面焦点》. 2011年2月;1(1):36–47. https://doi.org/10.1098/rsfs.2010.0009.[22] Burns FJ, Tannock IF. 论细胞周期中 G_0期的存在. 《细胞增殖》. 1970年;3(4):321–334. https://doi.org/10.1111/j.1365-2184.1970. tb00340.x.[23] Nixson RE, Byrne HM, Pitt-Francis JM, Maini PK. 刻画细胞周期结构化偏微分方程模型相对于相应常微分方程系统的行为. Bulletin of Mathematical Biology. 2025年6月;87(7):93. https://doi. org/10.1007/s11538-025-01472-8.

[24]Zielke N, Edgar BA. FUCCI传感器：细胞增殖分析的强大新工具. Wiley Interdisciplinary Reviews Developmental Biology. 2015年;4(5):469–487. https://doi.org/10.1002/wdev.189.[25] Sakaue-Sawano A, Kurokawa H, Morimura T, Hanyu A, Hama H, Osawa H, 等. 可视化多细胞细胞周期进程的时空动态. Cell. 2008年2月;132(3):487–498. https://doi.org/10.1016/j.cell.2007.12.033.[26] Bajar BT, Lam AJ, Badiee RK, Oh YH, Chu J, Zhou XX, 等. 用于同时报告全部四个细胞周期时相的荧光指示剂. Nature Methods. 2016年12月;13(12):993–996. https://doi.org/10.1038/nmeth.4045.[27] Weber TS, Jaehnert I, Schichor C, Or-Guil M, Carneiro J. 用随机模型和双核苷脉冲标记量化真核细胞周期时相的时长与方差. PLOS Computational Biology. 2014年7月;10(7):e1003616. https://doi.org/10.1371/journal.pcbi.1003616.[28] Billy F, Clairambault J, Delaunay F, Feillet C, Robert N, Billy F, 等. 用于研究生长因子对细胞周期动力学影响的年龄结构细胞群体模型. Mathematical Biosciences and Engineering. 2013年;10(1):1– 17. https://doi.org/10.3934/mbe.2013.10.1.[29] Sherer E, Tocce E, Hannemann RE, Rundell AE, Ramkrishna D. 年龄结构模型的辨识：细胞周期时相转换. Biotechnology and Bioengineering. 2008年3月;99(4):960–974. https://doi.org/10.1002/bit.21633.[30] Pavel M, Renna M, Park SJ, Menzies FM, Ricketts T, F¨ullgrabe J, 等. 接触抑制通过YAP/TAZ-自噬轴调控细胞存活与增殖. Nature Communications. 2018年7月;9(1):2961. https://doi.org/10.1038/s41467-018-05388-x.[31] Yates CA, Ford MJ, Mort RL. 将细胞增殖表示为马尔可夫过程的多阶段模型. Bulletin of Mathematical Biology. 2017年12月;79(12):2905–2928. https://doi.org/10.1007/s11538-017-0356-4.[32] Golubev A. 指数修正伽马分布作为与细胞增殖和基因表达相关的时间变异性模型的应用与意义. 《理论生物学杂志》. 2016年3月;393:203–217. https: //doi.org/10.1016/j.jtbi.2015.12.027.[33] Chao HX, Fakhreddin RI, Shimerov HK, Kedziora KM, Kumar RJ, Perez J, 等. 证据表明人类细胞周期是一系列解耦的无记忆时相. Molecular Systems Biology. 2019年3月;15(3):e8604. https://doi.org/10. 15252/msb.20188604.[34] Stukalin EB, Aifuwa I, Kim JS, Wirtz D, Sun SX. 用于理解连续培养中群体波动的年龄依赖性随机模型

细胞。《英国皇家学会界面杂志》。2013年8月；10(85)：20130325。https: //doi.org/10.1098/rsif.2013.0325.[35] Chaffey GS, Lloyd DJB, Skeldon AC, Kirkby NF。G1期 - S期转换检查点对年龄结构化细胞周期模型的影响。《PLoS ONE》。2014年1月；9(1)：e83477。https://doi.org/10.1371/journal.pone.0083477.[36] Milotti E, Del Fabbro A, Dalla Pellegrina C, Chignola R。细胞去同步化数据分析的统计方法。Physica A：统计力学及其应用。2008年7月；387(16)：4204–4214。https://doi.org/10.1016/j. physa.2008.03.006.[37]Chiorino G, Metz JAJ, Tomasoni D, Ubezio P。细胞群体中的去同步化速率：数学建模与实验数据。《理论生物学杂志》。2001年1月；208(2)：185–199。https://doi.org/10.1006/jtbi.2000. 2213.[38] Olofsson P, McDonald TO。细胞周期去同步化的随机模型。《数学生物科学》。2010年2月；223(2)：97–104。https://doi.org/10.1016/j. mbs.2009.11.003.[39] Vittadello ST, McCue SW, Gunasingh G, Haass NK, Simpson MJ。纳入多阶段细胞周期的数学模型再现了细胞增殖中通常隐藏的内在同步性。《英国皇家学会界面杂志》。2019年8月；16(157)：20190382。https://doi.org/10.1098/rsif.2019.0382.[40] Gavagnin E, Vittadello ST, Gunasingh G, Haass NK, Simpson MJ, Rogers T, 等。人口统计噪声解释了生长中的细胞群体内的同步振荡。《生物物理学杂志》。2021年4月；120(8)：1314–1322。https: //doi.org/10.1016/j.bpj.2021.02.017.[41] Corvaisier M, Bauzone M, Corfiotti F, Renaud F, Amrani ME, Mont´e D, 等。YAP/TAZ和细胞周期蛋白E1对结肠癌细胞静止的调控：对化疗耐药性和癌症复发的影响。《Oncotarget》。2016年8月；7(35)：56699–56712。https://doi.org/10.18632/oncotarget.11057.[42] Wieland FG, Hauber AL, Rosenblatt M, T¨onsing C, Timmer J。论结构可辨识性与实际可辨识性。《系统生物学当前观点》。2021年3月；25：60–69。https://doi.org/10.1016/j.coisb.2021.03.005.[43] Liu Y, Suh K, Maini PK, Cohen DJ, Baker RE。细胞侵袭偏微分方程模型的参数可识别性与模型选择。《英国皇家学会界面杂志》。2024年3月；21(212)：20230607。https://doi.org/10.1098/rsif.2023.0607.[44]Raue A, Kreutz C, Maiwald T, Bachmann J, Schilling M, Klingm¨uller U, 等。部分观测动力学模型的结构与实际可辨识性分析

通过剖面似然加以利用。《生物信息学》。2009年8月；25(15)：1923–1929。https://doi.org/10.1093/bioinformatics/btp358.[45] Haileka V, George S, Steinbrecht S, Jung F, Reinehr R, K¨upper JH。在高渗条件下培养的结肠癌细胞作为体外模型，用于研究脱水效应对癌症药物敏感性的影响。临床血液流变学与微循环。2019年11月；73(1)：169–176。https://doi.org/10.3233/CH-199210.[46]Witzel F, Fritsche-Guenther R, Lehmann N, Sieber A, Bl¨uthgen N。基于阻抗的细胞生长测定分析。《生物信息学》。2015年8月；31(16)：2705–2712。https://doi.org/10.1093/bioinformatics/btv216.[47] Weiner AC, Williams MJ, Shi H, V´azquez-Garc´ıa I, Salehi S, Rusk N, 等。从单细胞 DNA 测序数据推断复制时序与增殖动力学。Nature Communications。2024年10月；15(1)：8512。https://doi. org/10.1038/s41467-024-52544-7.[48] Cameron IL, Greulich RC。关于成年小鼠更新中的上皮组织中 DNA 合成持续时间基本恒定的证据。细胞生物学杂志。1963年7月；18(1)：31–40。https://doi.org/10.1083/jcb.18.1.31.[49] Storn R, Price K。差分进化——一种用于连续空间全局优化的简单高效启发式方法。全球优化杂志。1997年12月；11(4)：341–359。https://doi.org/10.1023/A:1008202821328.[50] McCullagh P, Nelder JA。广义线性模型。统计学与应用概率专著：广义线性模型；37。佛罗里达州博卡拉顿：泰勒与弗朗西斯出版集团旗下劳特里奇出版社；2019年。[51] Albert I, Denis JB。狄利克雷分布与多项分布：性质及其在 JAGS 中的用途。法国国家农业、食品与环境研究院，数学与应用信息学部；2012年。技术报告 2012-5。[52] Clerx M, Robinson M, Lambert B, Lei CL, Ghosh S, Mirams GR, 等。噪声时间序列的概率推断（PINTS）。开放研究软件杂志。2019年7月；7(1)：23。https://doi.org/10.5334/jors.252.[53] Daly AC, Gavaghan D, Cooper J, Tavener S。基于推断的非线性生物模型参数可辨识性评估。《英国皇家学会界面杂志》。2018年7月；15(144)：20180318。https://doi.org/10.1098/rsif.2018. 0318.[54] Browning AP, Warne DJ, Burrage K, Baker RE, Simpson MJ。系统生物学中随机微分方程模型的可辨识性分析。《英国皇家学会界面杂志》。2020年12月；17(173)：20200652。https://doi.org/10.1098/rsif.2020.0652。

[55] Kreutz C、Raue A、Kaschek D、Timmer J。系统生物学中的剖面似然。FEBS期刊。2013年;280(11):2564–2571。 https://doi.org/10.1111/febs.12276。[56] Nowak CM、Quarton T、Bleris L。细胞周期变异性对细胞群体动力学的影响。PLOS Computational Biology。2023年6月;19(6):e1011080。https://doi.org/10.1371/journal.pcbi.1011080。